AIChE Symposium Series No. 326
Volume 98

CHEMICAL PROCESS CONTROL-VI

Assessment and New Directions for Research

Proceedings of the Sixth International Conference on
Chemical Process Control
Tucson, Arizona, January 7-12, 2001

Editors

James B. Rawlings
University of Wisconsin

Babatunde A. Ogunnaike
DuPont Company

John W. Eaton
University of Wisconsin

Production Editor, CACHE Publications
Brice Carnahan
University of Michigan

CACHE
American Institute of Chemical Engineers

2002

> Library of Congress Control Number: 2002102094
>
>

ISBN 0-8169-0869-9

Preface

The sixth International Chemical Process Control meeting (CPC VI) was held January 7–12, 2001 at the Westward Look Resort in Tucson, Arizona. The uniqueness of the CPC meeting series in the process control community is illustrated by the following features: the meetings are held at five year intervals, all talks are invited and are organized in a single track so the entire audience can attend each presentation, and most importantly, a significant fraction of the audience, usually more than half, is comprised of leaders from the chemical process industries.

The overall goal of the CPC conference series is to evaluate the current progress in this field and identify the new intellectual challenges and frontiers that may have fundamental impact on future industrial practice.

At CPC VI, of the 135 registered attendees, about 65 were from the chemical process industries, and the rest were from universities and federal agencies.

The program consisted of the following technical sessions with session organizers in parenthesis.

- Opening Session (J. Brian Froisy)
- Modeling and Identification (Jay H. Lee)
- Life Sciences (Francis J. Doyle)
- Control Theory (Frank Allgower)
- Hybrid Systems (Manfred Morari)
- Controller Performance Monitoring and Maintenance (Derrick J. Kozub)
- Chemical Reactors/Separators (Lorenz T. Biegler and B. Wayne Bequette)
- Modeling and Control of Complex Products (Babatunde A. Ogunnaike)
- Contributed Paper Poster Session (Kenneth R. Muske)
- Closing Session (James B. Rawlings)

The meeting also included a Vendor and Software Display organized by S. Joe Qin.

A detailed meeting summary has been compiled including graphical representation of detailed pre-session and post-session evaluations. This summary is is available at the meeting website:

`http://www.che.wisc.edu/cpc-6`

Finally, we should attempt to draw conclusions from the meeting. These are the editors' opinions, but are informed by the week long CPC VI meeting discussions.

The chemical process control community (indeed the entire chemical engineering profession) is going through a transition period, in which the industries served by the chemical engineering profession are diversifying. During this period, it is essential that process control researchers continue to identify new opportunities in which applications of systems theory concepts provide significant added value. These proceedings document that the younger researchers in the community are already engaged in vigorous efforts to build significant collaborations to support these new application areas. Those activities portend well for the future health of this research community. Several people noted that discussions at previous CPCs were marked by more conflict than those at CPC VI, and one explanation seems to be that we have been extremely successful in bringing the best advanced control concepts and methods into practice during this short time period. We should not become victims of this success, but must embrace both the controversial new ideas, as well as the critical evaluation of these ideas, as we create the concepts that will sparkle at future CPC meetings.

Secondly, it seems timely to return to and embrace the broad meaning of systems theory and the tools it brings to chemical process modeling, dynamics and control. The significant distinction in systems engineering is whether we address the online challenge—using data and models for best real time operations decisions, or whether we address the offline challenge—to synthesize and design the new products and the new processes. Other less vital distinctions and categories that have grown up over the years do not serve us well in the current climate, because they dilute rather than deepen our focus.

Finally, as representatives of the chemical engineering profession with some of the closest connections to applications and technology, this community has a unique opportunity to play a leadership role in identifying the significant new directions in areas such as: computing tools, information processing, and measurement technologies, to name just three areas in which we are all experts. Given the rapid pace of technological change, the chemical engineering profession needs people who can distinguish the fundamental changes that present new opportunities from the blinding array of incremental and ephemeral changes. This community should embrace the challenge of providing that leadership. We have the knowledge, the tools, the experience, the vision and the people to play a key role.

James B. Rawlings, University of Wisconsin
Babatunde A. Ogunnaike, DuPont
John W. Eaton, University of Wisconsin

Acknowledgments

CPC VI was sponsored by the CACHE Corporation and the Computing and Systems Technology (CAST) Division of AIChE.

Generous financial support of the meeting was provided by:

Aspen Technology, Inc.
DuPont
Equilon Enterprises LLC
ExxonMobil Chemical Company
National Science Foundation

Brice Carnahan provided invaluable assistance in preparing the final printed proceedings.

Contents

Business Process Control: The Outer Loop

Lowell B. Koppel
Value Techniques, LLC
16 Hastings Road
Winchester, MA 01890

Abstract

Business process control (BPC) applies the principles of chemical process control (CPC) to enable an entire enterprise to achieve peak performance. Viewed from the chemical process perspective, BPC is the outer loop that manages the CPC targets with a specific objective of meeting and exceeding business targets. Viewed from the business process perspective, CPC is one specific set of the many inner loops requiring strategic direction. Typically, CPC is the set of inner loops that is extensively automated with valves, switches, sensors, and analyzers. CPC has sophisticated, usually computerized, control algorithms. However, its ability to influence the enterprise is limited to what can be accomplished by moving valves and switches.

Characteristics of BPC loops are: enterprise-wide scope; information systems as feedback sensors; and, knowledge workers as control strategists and as actuators. The scope of influence on the enterprise is everything that can be affected by humans, which scope is virtually unlimited. This paper will review the current state of BPC with some example applications. A key theme is to note that the shift in scope from the valves and switches of the automated inner loops, to the enterprise-wide business processes of the outer loops, increases the potential benefits by orders of magnitude. The CPC professional community is the main repository of intellectual capital required to translate and extend the mature CPC technology to the emerging BPC discipline.

Keywords

Business process control, Value metrics, Project planning, Project scheduling

Overview of Business Process Control

The premises of Business Process Control are:

- Most enterprises have untapped potential value
- Potential value and current value can be measured
- There is a gap between potential and current value
- Increasing stakeholder value to its full potential is an enterprise goal
- The goal is reachable when knowledge workers can observe both the current gap, and the reduction in the gap that results from their activities

Business process control (BPC) is the management science that employs measurements of the gap between current and potential value to create new value.

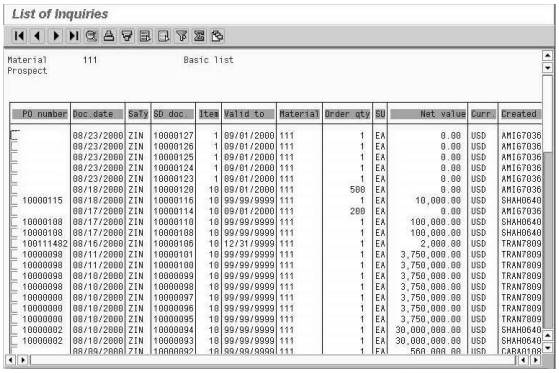
Measurement is key to BPC. Value Metrics provide measurements needed to create new value. We will give several examples of Value Metrics.

Knowledge workers at all levels of the enterprise use metrics to decide and to perform value-creating actions. People manage the enterprise as controllers and actuators in BPC loops.

The basic structure of BPC is identical to the proven structure of a classical feedback loop. Metrics are the feedback sensors. Knowledge workers and decision support tools are the controllers and actuators.

Value Metrics

Enterprises are increasingly deploying key performance indicators (KPIs) as they recognize that measurement is a prerequisite to improvement. A recent book on the subject became a best seller ("Balanced Scorecard").


List of Inquiries

Material 111 Basic list
Prospect

PO number	Doc. date	SaTy	SD doc.	Item	Valid to	Material	Order qty	SU	Net value	Curr.	Created
	08/23/2000	ZIN	10000127	1	09/01/2000	111	1	EA	0.00	USD	AMI67036
	08/23/2000	ZIN	10000126	1	09/01/2000	111	1	EA	0.00	USD	AMI67036
	08/23/2000	ZIN	10000125	1	09/01/2000	111	1	EA	0.00	USD	AMI67036
	08/23/2000	ZIN	10000124	1	09/01/2000	111	1	EA	0.00	USD	AMI67036
	08/23/2000	ZIN	10000123	1	09/01/2000	111	1	EA	0.00	USD	AMI67036
	08/18/2000	ZIN	10000120	10	09/01/2000	111	500	EA	0.00	USD	AMI67036
10000115	08/18/2000	ZIN	10000116	10	99/99/9999	111	1	EA	10,000.00	USD	SHAH0640
	08/17/2000	ZIN	10000114	10	09/01/2000	111	200	EA	0.00	USD	AMI67036
10000108	08/17/2000	ZIN	10000110	10	99/99/9999	111	1	EA	100,000.00	USD	SHAH0640
10000108	08/17/2000	ZIN	10000108	10	99/99/9999	111	1	EA	100,000.00	USD	SHAH0640
100111482	08/16/2000	ZIN	10000106	10	12/31/9999	111	1	EA	2,000.00	USD	TRAN7809
10000099	08/11/2000	ZIN	10000101	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000098	08/11/2000	ZIN	10000100	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000098	08/10/2000	ZIN	10000099	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000098	08/10/2000	ZIN	10000098	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000000	08/10/2000	ZIN	10000097	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000000	08/10/2000	ZIN	10000096	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000000	08/10/2000	ZIN	10000095	10	99/99/9999	111	1	EA	3,750,000.00	USD	TRAN7809
10000002	08/10/2000	ZIN	10000094	10	99/99/9999	111	1	EA	30,000,000.00	USD	SHAH0640
10000002	08/10/2000	ZIN	10000093	10	99/99/9999	111	1	EA	30,000,000.00	USD	SHAH0640
	08/09/2000	ZIN	10000092	10	99/99/9999	111	1	EA	560,000.00	USD	CARA0108

Figure 1: Conventional KPI listing customer inquiries and their resolution.

Value Metrics are a distinctive class of KPI. The distinct characteristics of Value Metrics are

- Direct economic measure
- Root-cause analysis
- Defined business process

We next present examples of metrics. The examples will illustrate these distinct characteristics, and how they significantly increase value creation, compared to the capabilities of conventional KPIs.

Direct Economic Measure

We first consider the importance of direct economic measure. The business process is Customer Service. Figure 1 shows a conventional KPI. It reports Customer Inquiries

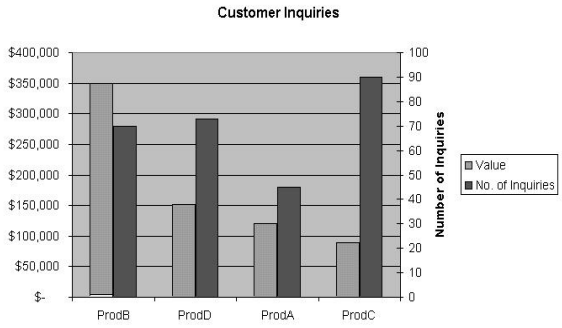


Figure 2: Pareto chart indicating the total lost *value* due to inquiries on each product.

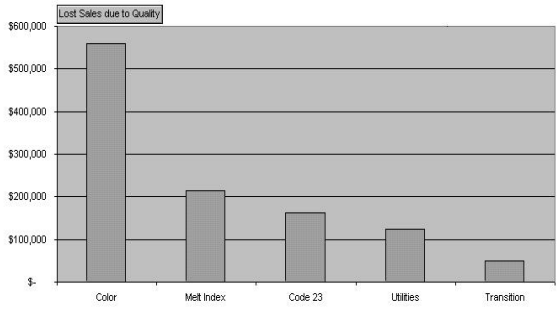


Figure 3: For the current week, Color quality is the leading cause of value loss.

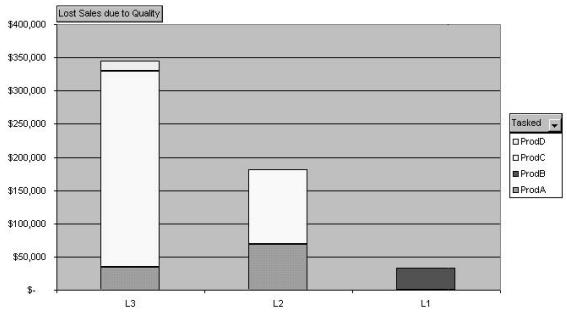


Figure 4: The bulk of the problem occurs when Product C is being made on Line 3.

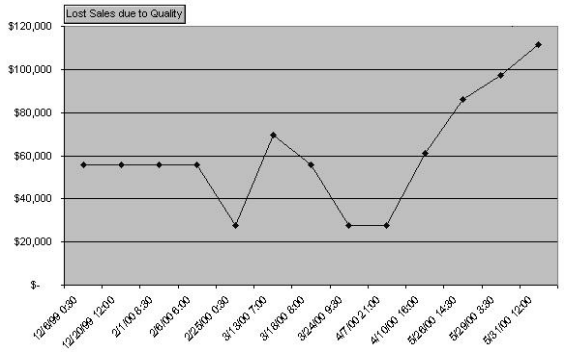


Figure 5: Metric trend plot for Product C.

and their resolution. Each inquiry is valued in terms of product weight, and directly converted to USD. However, this does not necessarily reflect the value of the inquiry.

The Pareto chart in Figure 2 shows a metric that supersedes this KPI. The left ordinate is a metric, indicating the total lost value due to inquiries on each product. The right ordinate is a conventional KPI, indicating only the number of inquiries.

The metric reflects the value of each inquiry and its resolution. It considers not only obvious factors such as weight of product, magnitude of price reduction, and extra shipping and handling costs. It also considers the value of the particular customer's annual revenue stream, the goals and current status of market penetration in the specific product line, and possibly other business-related factors.

The difference between the blue bars representing value, versus the red bars representing frequency, shows one difference between a Value Metric and a conventional KPI. A Customer Service manager using the conventional KPI would be influenced by problems in Product C, the squeaky wheel, rather than by the problems in Product B. But, problems in Product B are contributing three to four times as much to the enterprise value gap than are the more frequent problems in Product C.

Root Cause Analysis

Feedback measurements that lead to root cause detection give knowledge workers important support toward their control of the business process. Value Metrics provide feedback with as much detail as is practical to understand the cause of an observed value loss.

The next example relates to the business process of Quality Management. The metric illustrated in Figure 3 feeds back that, for the current week, Color quality is the leading cause of value loss.

BPC occurs when the knowledge worker determines the cause, and corrects the problem. Choosing Color for the first drill down produces the metric shown in Figure 4. It shows that the bulk of the problem occurs when Product C is being made on Line 3.

To further trace the cause we drill-down by choosing Product C for the metric trend plot shown Figure 5.

This metric identifies the problem as a new one. Each piece of additional information from the metric helps move people to the ultimate goal—BPC to recapture the lost value.

The drill down can go a step further, providing de-

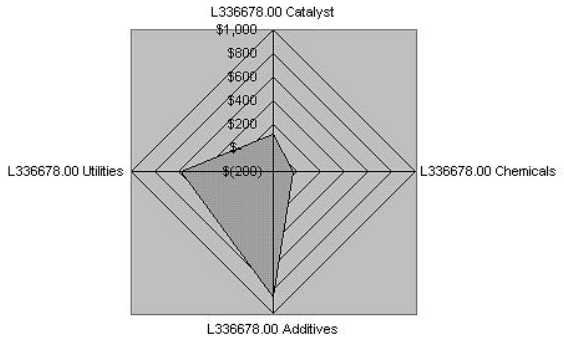


Figure 6: Variable cost components—planned vs. actual.

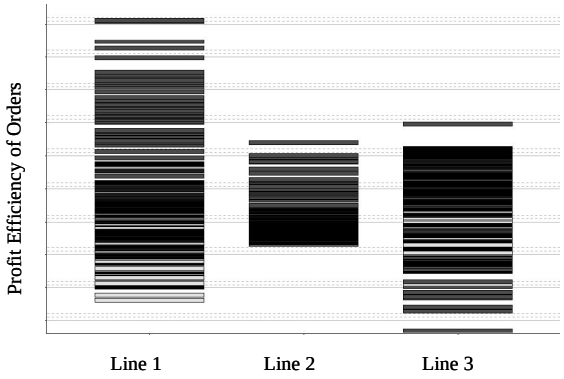


Figure 7: A metric for production management.

tailed feedback on each batch shown on the metric trend. The most recent batch, shown by the arrow, is the likely drill-down candidate.

The spider-web metric shown in Figure 6 displays the results. This metric shows deviation of actual versus planned for each variable cost component.

The root-cause feedback suggests that excess additives may have caused the original color quality losses. Value is being lost both through excess additive cost and possibly through reduced revenues due to the impact on color quality.

Defined Business Processes

The Value Metrics drill-down sequence discussed above is an example of defined business processes. In other words, for the specific example cited above, the defined business process that becomes the best practice for the knowledge worker is:

1. Identify site-wide lost sales due to quality
2. Identify on which production line and product the largest losses were incurred for a specific quality code
3. Identify specific batches within this line/product combination that led to these losses
4. Identify feedstock components that could be associated with the quality issue

Other business processes defined for this metric include specification of the decision and correction/actuation tasks, and their ownership, that will follow each type of feedback signal received from the metric.

Business process definition serves three key purposes:

- Business process definition is essential to forecasting the actual value creation capabilities of a metric. In other words, the defined business processes are the basis for estimating the benefits of implementing a metric.
- Value Metrics with defined business processes are better designed. The requirement to specify in advance exactly how people will use the information leads to better information design.
- Value Metrics with pre-defined business processes have longer useful lifecycles. Execution of the value-creating decision tasks in the initially defined business processes stimulates knowledge workers to devise new business processes for the metric. The discipline of this best practice is rewarded by continual creation of new value, steadily moving the enterprise assets toward their full potential.

Current Applications

We next give three illustrations of current practice in Business Process Control, and the role of metrics. The illustrations progress from substantial automation content, with limited knowledge worker involvement, to zero automation content, with full decision and actuation by knowledge workers.

Production Management

The business process to be controlled in this simple example is execution of an order to produce a specified quantity of specified product for a specified customer on a specified production line at a specified time.

A production order appears at the knowledge worker's station. The automated production management system has filled in most, possibly all, of the information required to fulfill the order. The information is sometimes referred to as the recipe, and includes target values to be downloaded to the process controlling field devices.

The knowledge worker adjusts the information based on his/her experience with value creation for this combination of product, customer, and line. He/she approves the information and triggers its download for execution.

During execution, the knowledge worker monitors and adjusts the recipe and targets based on feedback from the field devices. In most cases, these are not Value Met-

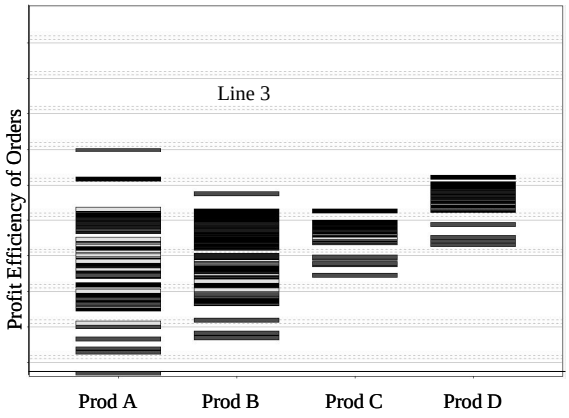


Figure 8: Products A and B are the source of value loss on Line 3.

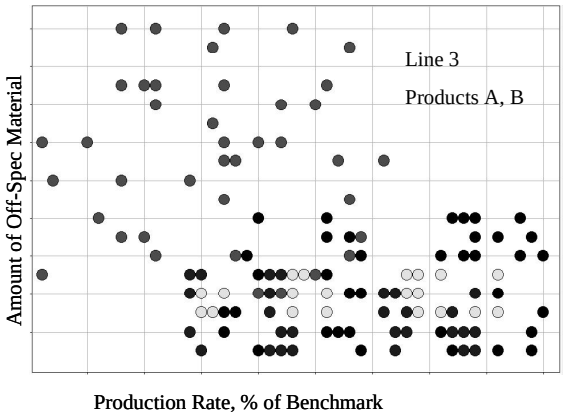


Figure 9: The Red Team is consistently under performing the others on both quality and rate.

rics because they are in physical, rather than economic, units.

A metric for Production Management is illustrated in Figure 7, built on production data collected for each order produced during the previous month.

The direct economic unit on the ordinate is the profit efficiency of each order, defined as the fractional deviation of actual profit from recipe profit. Profit efficiency is sorted against the production lines, and color-coded by the Team responsible for the production order. This metric provides Production Management knowledge workers with the information they need to add value by improving Lines, Teams, and business processes.

A drill down into the relatively poor performance of Line 3 produces the metric shown in Figure 8.

This drill down focuses attention on Products A and B as the source of value loss on Line 3.

The drill-down shown in Figure 9 looks at Products A and B on Line 3. It shows that the Red Team is consistently under performing the others on both quality and rate. This last view is not strictly a Value Metric, because it shows physical rather than economic measures.

However, the previous metrics led the way to this view, which provides Production Management knowledge workers with feedback information for exercising BPC and creating value.

Production Planning

Production planning is a BPC activity. A decision support tool such as a linear program can provide knowledge workers with the feedback information required to set the production plan.

The purpose of this example is to illustrate the advantage of using a true Value Metric as the feedback measurement.

Production planning requires specification of minimum and maximum inventory levels. Often these are set qualitatively, based on experience. The linear (or other) programming tool requires a quantitative objective function to optimize the production plan. The challenge is to choose inventory constraint levels on true economic grounds.

In addition to the obvious requirements that storage facilities be neither flooded nor drained, it is desirable to have a business-driven basis for setting minimum and maximum inventory constraints. These will be targets such as less than an $x\%$ chance that any customer orders will go unfilled, and less than a $y\%$ chance that any supplier receipts must be delayed.

Here is one set of defined business processes and metric to support Production Planning BPC.

- Customer orders and supplier receipts are specified as distributions rather than as fixed numbers. Examples are expected values with standard deviations; or, minimum, likely, and maximum values.
- The optimization is nested. The outer optimization chooses a set of inventory constraints for the inner optimization. The inner optimization optimizes a criterion such as expected profit, or median profit, for the given set of constraints.
- The outer optimization drives to maximum expected profit or similar business-driven criterion. It does this by choosing inventory policy.

Note the absence of any artificial penalty functions on inventory violations. All specifications are expressed in actual business targets.

The metric shown in Figure 10 illustrates the feedback for BPC. It is a plot of the expected profit as a function of the inventory constraint on one particular item, with all other inventory constraints held at their overall optimum levels.

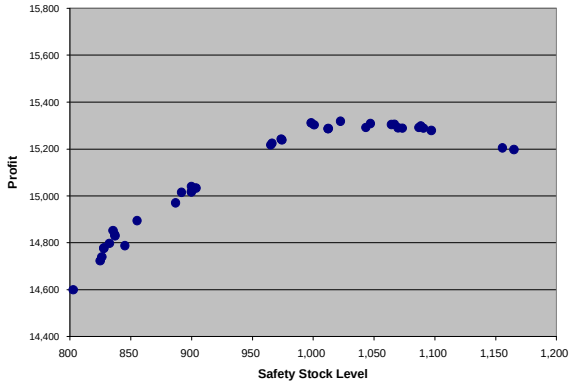


Figure 10: Expected profit as a function of the inventory constraint on one particular item, with all other inventory constraints held at their overall optimum levels.

Project ID	Project Title	Cost, thousands			Benefits, thousands/yr			Duration Quarters			Required Precedents
		Minimum	Forecasted	Maximum	Minimum	Forecasted	Maximum	Minimum	Forecasted	Maximum	
1	Database	970	970	1,940	0	0	0	3	3	5	
2	Production Management System	1,390	1,390	2,780	0	1,710	1,710	6	6	8	1
3	Dock Scheduling System Migration	1,000	1,000	1,400	500	1,000	1,000	3	3	4	
4	Shipments and Receipts System Integration	220	220	440	0	2,780	2,780	3	3	5	2
5	Order Entry System Integration	100	100	200	0	7,320	7,320	3	3	5	2
6	Operations Scheduling System										8
7	Operations Scheduling System Integration										2,7
8	Production Planning System Integration										2
9	Lab System Integration										2
10	Environmental Monitoring and Reporting System										
11	Energy and Utility Management System										
12	Maintenance System Integration				Data not copied to protect client confidentiality						2
13	Materials Procurement System Integration										2
14	Financial Systems Integration										2
15	Health and Safety Management System										14
16	HR Systems Migration to Corporate System										
17	Document Management System										
18	Technology Management System										17

Figure 11: Sample of data on projects that were candidates for capital.

Project Scheduling and Capital Allocation

Allocation of capital to projects and planning their rollout is an ongoing business process. Projects are often prioritized on the basis of their individual cost/benefit attributes because these individual metrics are readily available.

This example presents a metric that evaluates the total slate of candidate projects to support BPC on allocation and scheduling. The metric also has the quality of reflecting uncertainties in the expected costs, benefits, and durations of the individual projects.

Figure 11 is a sample of data on projects that were candidates for capital.

A program of 18 projects over 5 years was proposed to create a site wide information system. Figure 11 shows costs, benefits, durations, and precedents for each project. Some points to note:

- The site preferred conservative uncertainty ranges.
- Actual costs, benefits, and durations were assumed to be equally likely to end up anywhere in between the minimum and maximum levels shown. Other less conservative distributions could have been chosen so that actual outcomes were more likely to be near the forecasted value than near the range extremes.
- Project 1, a database infrastructure project, was assigned no direct benefits at all. However, it is a prerequisite for Project 2, which in turn is a prerequisite for seven other benefit-producing projects. This is a realistic way to handle the benefits of infrastructure projects.
- The numbers shown represent 1991 dollars.

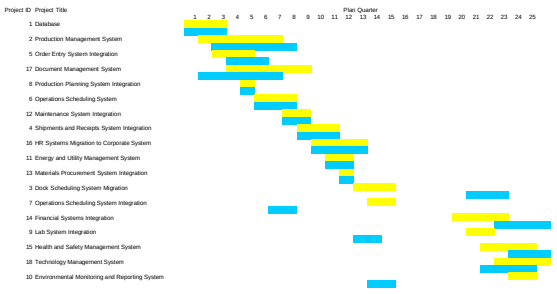


Figure 12: Initial project plan (yellow) and optimized project plan (blue) found by BPC using the feedback from the metric.

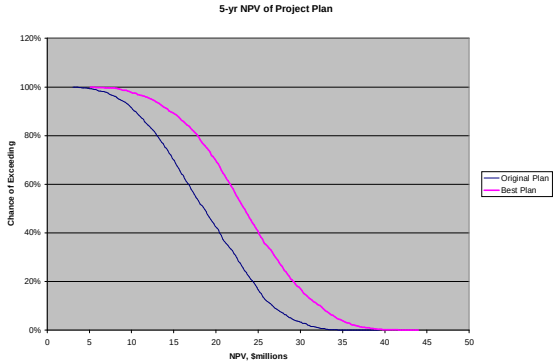


Figure 13: The metric that supported the BPC comparing the distributions of the original plan and the plan that maximized the median value of the NPV.

Figure 12 shows in yellow the initial project plan, and in blue the optimized project plan found by BPC using the feedback from the metric.

BPC reinstated projects 7, 9, and 10, originally omitted from the 5-yr plan. Projects 3 and 14 were removed from the 5-yr plan by BPC.

The metric that supported the BPC is illustrated in Figure 13. The direct economic unit is the 5-year Net Present Value (NPV) of the project plan. The Value Metric compares the distributions of the original plan, and the plan that maximizes the median value of the NPV. The best plan adds about 25% to the NPV.

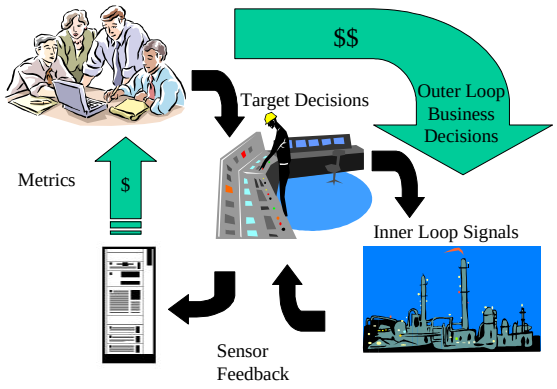


Figure 14: The outer and inner loops that form the familiar cascade configuration.

Cascade Control

The outer and inner loops we have discussed form the familiar cascade configuration shown in Figure 14.

Summary

Table 1 summarizes the diverse business processes considered here, and the Value Metrics that were used.

The unifying theme is BPC. To control each of these business processes requires a feedback metric. Attempting to control without a metric loses value for the enterprise, just as does attempting CPC without valid feedback sensors.

Opportunity

The unifying theme of BPC is Value Metric feedback supporting decision and control/actuation by knowledge workers. The scope covers virtually all enterprise business processes. The potential value creation for this scope far exceeds that of the scope of traditional process control.

An opportunity lies in translating the science of control from chemical processes to business processes. In the near term, it should be possible to develop algorithms that support knowledge workers use of the metrics feedback, by narrowing the possible control actions to those most likely to succeed in value creation. In the long term, once we get the metric feedback right, much more is possible.

Intellectual capital to achieve these goals lies in the community of researchers on CPC. A key goal of this paper is to make a case for the much greater value creation opportunity in BPC than in CPC. The scope of economic influence of knowledge workers far exceeds that of valves and switches.

Business Process	Value Metric
Customer Service	Value of each incident
Quality Management	Value of lost sales
Production Management	Order profit efficiency
Production Planning	Expected plan profit
Capital Project Planning	Expected NPV

Table 1: Summary of the diverse business processes considered and the Value Metrics that were used.

Influence of Computers and Information Technology on Process Operations and Business Processes—A Case Study

J. Patrick Kennedy and Osvaldo Bascur
OSI Software, Inc.
San Leandro, CA 94577

Abstract

Managing a company for longevity is different than managing for the stock price, the former requires flexibility, the latter requires a very short term (quarterly) focus. This paper describes some of the challenges to the management of industrial complexes brought about by the explosion of innovations in the informations system world. It focuses on a case study of a large, integrated mining company in Chile, Codelco, and describes how they have very successfully addressed this problem with largely packaged software and their own initiatives.

Keywords

Information systems, Mining, Optimization, Desktop, Infrastructure

Prices, products, and markets are changing more rapidly every year as information based management replaces the experience based management. Arie P. de Geus, in a short article called Strategy and Learning, observed that the average company survives only 40 years. This study was conducted long before the acceleration caused by the Internet Age so I suspect the number is lower now. He points out that no matter how well a company is managed (e.g. Sears, Penny's), epoch external changes can occur without warning and influence the business in sudden and unpredictable ways. The issue is one of human psychology—"You can not see what your mind has not experienced before, and you will not see that which calls forth unpleasant emotions." This especially applies to new Internet companies that have never experienced the old world values of profits, investment, competition, and growth from within. It also applies to old manufacturing companies that have not experienced new world values of e-Business, direct relations with customers, mind share and globalization.

To survive, companies must create what de Geus called "The Living Company." He also studied companies that had survived over 100 and up to 700 years looking for the secrets of longevity. In his conclusions he noted that "a more open company that involves everybody needed for execution in the planning and decision making process will be more successful in a world which it does not control." It is also clear that the strategy of managing for quarterly profits is quite different than managing for longevity. This changes the basic role of management in an information-based world. In the past, its main role was to allocate the available money generated from profits between the long term good of the company and the short-term demands of the shareholder. With the new demanding shareholder, that role has changed to taking the money left over after giving the shareholders enough to maintain the stock price and using it carefully to keep the company viable. With an Internet company unable to generate profits, the task is more difficult as management also needs to continuously borrow money at usurious rates.

A company is actually its intellectual property, not the mortar and bricks. Its future lies in its ability to create a collaborative environment of its people. It seems so simple—enable and motivate the people that work for the company. They already know how raw materials are acquired, the applicable environmental law, the processes, local conditions, even the limitations of the company and its people. Is it any surprise that they might know more than outside consultants on how to improve the company? If such an environment could be generated, then the 100's of smaller projects would net out large changes and sustained improvement. This is a dramatically different tact than betting the company on a giant software package, whether ERP, CRM, SCM, e-Business, or other "magic bullet." These large, vertically integrated packages do NOT increase the flexibility and enable people, they clamp down on procedures in the hope that in the automation of the business processes lies the efficiency needed to survive. The fatal flaw, as pointed out by de Geus is that it is not efficiency but flexibility that is needed to survive.

An alternative approach could be called IDEA (Infrastructure, Data Collection, Engineering, and Analysis (see `www.osisoft.com` white paper—Just an IDEA). It requires companies to have the methods, vision and faith to put in a system that enables the existing staff. People can then alter the direction of the company with many small moves within and outside their domain. When I wrote the article I had no idea that I would see IDEA in action at a large, state run mining operation in South America. They began this vision in 1997 and were well on their way to success. For all industries, it is worthwhile to analyze their experience.

The company is Codelco and they visualized an Industrial Desktop using Microsoft technology for integration. They combines Microsoft Office and OSI Software PI System to provide a very easy to use, real time development environment suitable for micro projects in the harshest conditions imaginable. Located in Chile,

Codelco has five operating Divisions (Chuquicamata, Radomiro Tomic, El Salvador, Andina, El Teniente), each a full mining/metallurgical complex—e.g. mining, concentrators, smelting, refining, leaching, solvent extraction, electrowinning, utilities, and maintenance. In 1997, the group of representatives from each division went to management and made the case that building this people-enabling environment was a strategic move and should be installed without bids, project justification or other methods associated with capital projects. It would be like a PC or telephone, it was just needed—an amazing move for a conservative manufacturer owned by the government. It has been an outstanding success.

It is almost impossible to describe how unusual and dramatically different this is from business as usual. Mining facilities are very large and diverse; they have different instrumentation, different systems, different equipment, different infrastructure, different standards and are actually in different industries (e.g. mining vs. refining) at each site. These sites tend to be remote—one is in an area called Moon Valley because it looks like the surface of Moon, located in the Atacama Desert. The unique feature of this implementation was that it was not just a common environment, but also a common DEVELOPMENT environment that encouraged mini projects by the users with only minimal help from outside or systems people. This was unique in the world for heavy manufacturers. They bet the future of the company on their employees, not an outside consultant working in London. The only question was would it work.

To say that they have won is an understatement. Recently Microsoft unveiled Microsoft.NET—their new initiative. This is an environment, totally compatible with the desktop environment above, that extends the development environment to the Web. This is a remarkable advancement in its own right. Whereas the first generation of browser users were content to look through information easily (i.e. browse), the next generation user will want to perform actions with the same, universal, and easy to use interface. Microsoft.NET is not only the best, it is the ONLY development environment that will allow people to build this next generation system and incorporate their PC. Imagine that you are able to purchase, set up the shipping, track the arrival, receive the invoice, and pay the invoice for a raw material from one integrated browser screen from any location. The remarkable feature is that even though this vision was not even embryonic in 1997, everything that Codelco has done supports this new vision.

I had the privilege in May 2000 of attending one of our Technology Seminars in Santiago and, even speaking no Spanish I could see the creation of "The Living Company." Codelco was to provide a user paper, but instead they produced on quick order not one but several applications done by their own people. These applications drove to the heart of the problems that exist in mining operations located in the high desert—saving water, saving oxygen consumption, reducing environmental emissions, better coordination of resources, saving energy, increasing tonnage, reducing waste, improving quality. These successes were in no small part a result of the low incremental cost of projects, which were not burdened with the cost of the infrastructure.

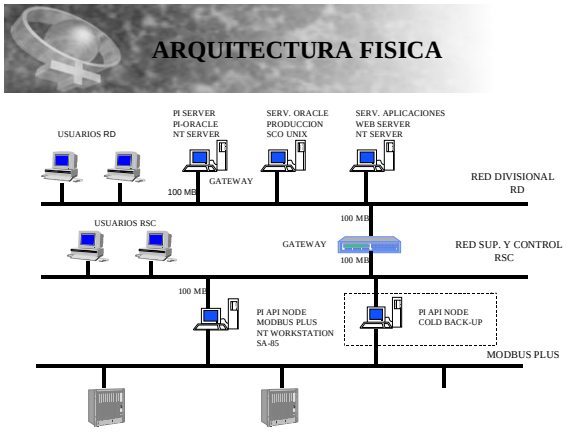


Figure 1: Radomiro Tomic architecture.

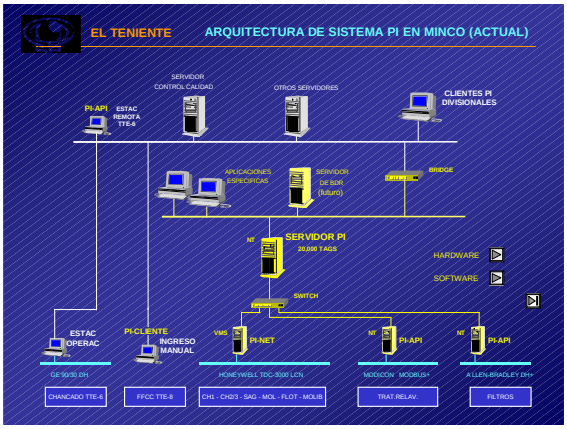


Figure 2: Teniente mine/concentrator plants architecture.

From a software perspective, the key was the openness and extensibility of the Microsoft DNA strategy. Not only did most of the people all ready know how to use Excel, Access, and Word, the PI System client (ProcessBook) is based on Microsoft COM/ActiveX technology and contains VBA from Microsoft. Codelco did not complain that they do not have the manpower to improve the operation. Although, like all other industries, they are minimizing manpower, using these Microsoft tools are as natural as writing a memo or building a spreadsheet.

This presentation was only part of the program—it is clear that there are hundreds more of these mini projects that will be envisioned and implemented by the people of Codelco. The first step was to build an infras-

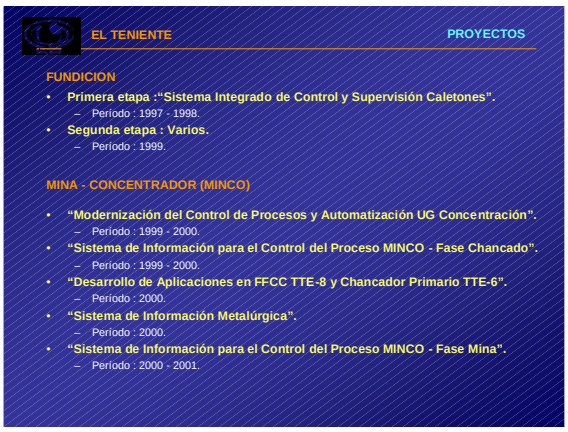


Figure 3: El Teniente projects.

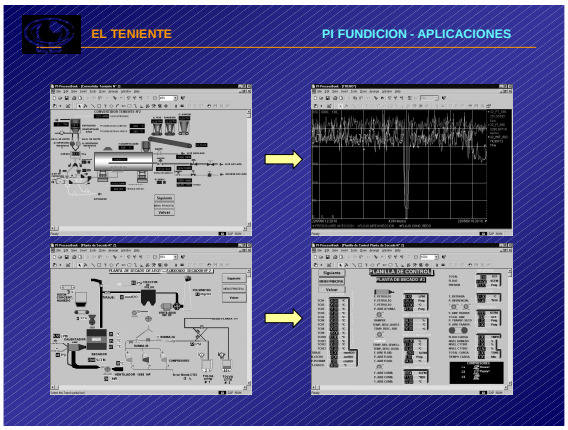


Figure 4: Same displays at El Teniente.

tructure. They have installed 15 PI Systems since 1997 and rolled Microsoft desktops everywhere. They are also cross-pollinating the implementations between Divisions because they have a common environment.

Figure 1 shows what the base infrastructure at Radomiro Tomic Division and Figure 2 shows the oldest systems at Teniente. They presented four separate papers from the different sites and head office with more than 10 separate applications—and a list of future applications.

In Figure 1 you can see the separation between the Control and Divisional Network—the basic data collection is done with a redundant system connected to their Modbus Plus network. There are workstations on the Control Network and the Divisional Network, but the servers that have the real time data (PI), Production data (Oracle) and Web Server are on the Divisional Network. In this way, there can be no interference to the control system or its consoles.

Figure 2 is a similar drawing for the Teniente Mine/Concentrator Plants—the first fully integrated site. The integration means from the crusher at the mine, transfer of ore from mine to crusher and concentrators, tailings management and concentrate filtering. The same basic structure is used except that there is not a separate Control Network. In some cases (e.g. Honeywell TDC3000) this network is incorporated into the control systems. A Modicon Modbus local control network is used at the tailings plant, an Allen Bradley local system is used at the filter plant. Of note here is that this system has continued to expand, e.g. extending the data collection to manual inputs, the railroad car transportation system and the mines, but these changes are incremental and do not require any modifications of the existing systems. For example, if Codelco were to add another DCS or an upgrade of an existing DCS, none of the information infrastructure would have to be changed.

The integration of the railroad car system (FFCC TTE-8) is simplifying the material handling into stockpiles and smoothing the major disturbances in these very large metallurgical complexes.

Figure 3 shows the projects at the El Teniente Division. The initial system was installed in late 1997 at the Smelter for $350K and through the end of 1999 they did various projects that took advantage of the real time data. The integration effort at the Teniente Smelter was large. They have more than 10 PI-API nodes collecting information and events from several areas (Drying, Smelting, Converters, Acid Plants, Anode Plants). The sum of all these projects was only $90K and this number included any additional software (e.g. ProcessBooks or DataLink for additional people), hardware (computers, routers, networks, analyzers), engineering (design of the project) and implementation. Clearly complex problems are being solved at such a low cost that there is almost no limit to the return. In parallel, they are implementing the MINCO System combined with PI for the Mine, Crushing, Concentrators (Sag Grinding, Conventional Grinding, Copper Flotation Plant, Retreatment Plant, Reagent Plant, and Molybdenite Flotation Plant) including the transportation system, tailings and concentrate filter plants.

In the next year, they have some larger projects that will involve adding more of the site plus smaller projects that are targeted at specific problems such as water conversation, air conservation, environmental monitoring, production reporting, integration with training system and others. These projects are all funded and managed locally. Codelco personnel do most of the work, but they also use local integrators (e.g. Contac Ingenieros) as needed.

Figure 4 shows some of the displays built for these projects. PI-ProcessBook from OSI Software was used as the development environment on these displays. Since this package contains VBA, any display can be upgraded to include interactive input and actions such as creating a hierarchy of displays or allow the user to "drill down." A key to the success of these projects is the ease of use of the Microsoft desktop. Very little training was needed

since displays are built just like sketches on PowerPoint and macros can be built just like they are in Excel. It is coincidental that just the week after this presentation the US Government decided that this seamless integration by Microsoft has hurt consumers and must stop—to say that those users that are gaining competitive edge using US software were confused is an understatement.

Figure 4 shows the performance monitoring screens for the Teniente Converter. They have implemented their process models using Datalink and ProcessBook. They also show their preparation and fluidized bed drying of the concentrate to optimize the smelting process. They show all performance, quality control, and events for coordinated decision-making. All personnel have real-time information access at the plant floor, business office, maintenance and metallurgical engineering. Some of the unique features of the software is that all of these display are live—curves are updating, equipment changes color on state, and alarms are seen throughout the site, if desire, within seconds. Any of these displays can be seen on the company Intranet using an ActiveX Viewer called PI-ActiveView.

Figure 5 shows a schematic of the Secondary-Tertiary Crushing plant and all conveyors at Teniente. This overview gives an overall view of the performance of the bins inventories, secondary and tertiary crushers, feeders, conveyors and screens at any time. Special queries can position the performance indicators at any shift, day, and monthly aggregation of the information on request. Spreadsheets can be used to compare the performance indicators by ore type or shifts.

There is a large amount of information in this real time graphic. However, by using the Microsoft VBA in ProcessBook they have each object connected to the tags by areas and units for the trend to get activated. They can monitor the inventories of these large bins to stabilize the level, thus minimizing segregation of the ore, which can cause large disturbances in the crushers, screening and conveying systems. The conveying systems are very large and require special monitoring. Several simple applications have been developed like counting how many times a crusher has hit a high oil pressure and motor amps, similar for the conveyors and screen. These slow moving variables are the cause of long-term degradation of the equipment. They use the historical and event driven routines for real time asset management and implementing condition based monitoring.

These systems are much different than existing instrumentation such as SCADA, DCS and PLC's. These plant systems can "see" and support projects that span multiple areas and help make the enterprise much more flexible. The PI System is the bridge between the plants and the enterprise business systems.

Figure 6 shows the real time production reporting using ProcessBook while Figure 7 shows reporting using Excel and DataLink (the PI real time data Add-In). Figure 6 shows the transfer of ore between the mine and the concentrator. These reports have macro capability (e.g. click on a number and see a trend or execute an analysis program or export the data), but Excel has the same rich set of features, familiar to all. In addition, most third party software can use Excel as data input to run new and legacy analysis programs. This shows the flexibility of the integrated environment from Microsoft and all of its partners like OSI Software. The full integration of all tiers, common menu operations (e.g. to Send a display as an object or to Save As is done the same everywhere) overcomes the training issue—always a problem for remote, industrial sites.

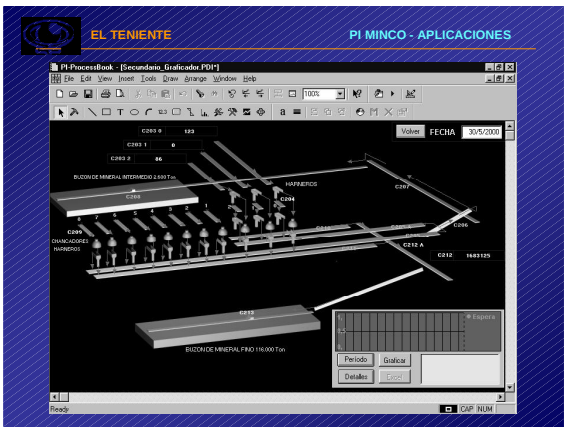


Figure 5: Secondary and tertiary crushing plant.

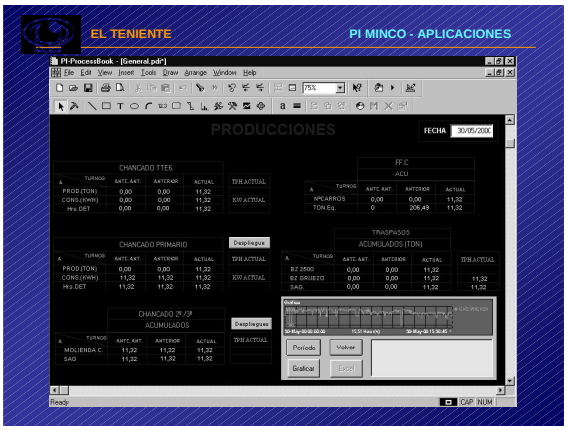


Figure 6: Real time production reporting.

Figure 8 is a ProcessBook display from the Plant Systems Operations computer at Radomiro Tomic. One advantage of having a common infrastructure is that everyone is familiar with the systems both across Divisions and sites. This had lead to a significant sharing of applications and knowledge that was never possible before due to the difference in the systems. Codelco presented a paper at the Latin American Mining Perpectives describing the strategic plan for automation and process management and work being done at Chuquicamata (Rojas

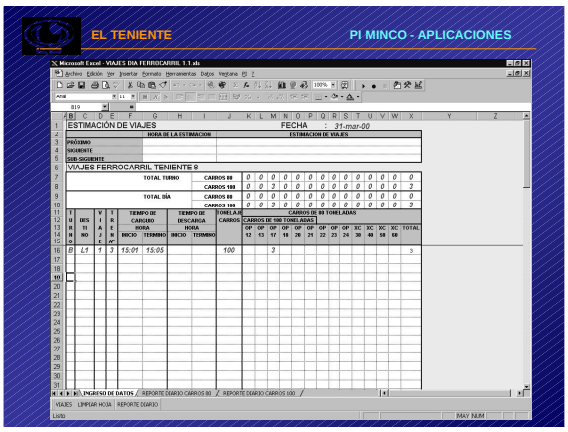


Figure 7: Excel and DataLink manual data entry tools.

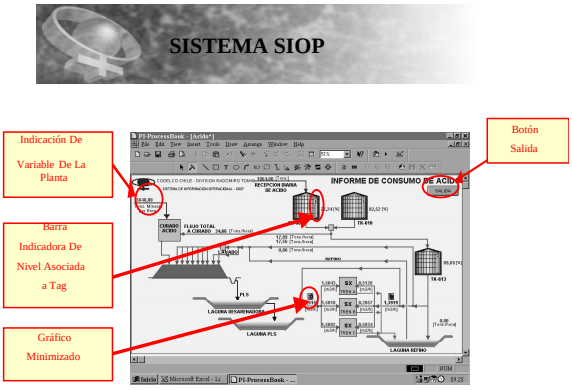


Figure 8: Sample display from plant operations system at Radomiro Tomic.

and Valenzuela, 1998). Montoya (1998) described the Radomiro Tomic hydrometallurgical complex. The integration of operational data and easy access are the key for improved leaching, improved equipment availability, minimized organic losses in the solvent extraction plants, and improved copper harvest cycle times at the electrowinning plant. Codelco personnel have developed many creative ways of transforming raw process data into highly valuable actionable information for fast operational decision-making. The real time operational information infrastructure is empowering operators, technologists, and managers to work in collaborative environment of continuous improvement.

It is collaboration of the employees which is making a change in paradigm on how we operate in these remote plants.

They have not stopped yet, in the next few years they will use their flexible infrastructure and the RLINK SAP gateway to support the real time applications in SAP's R/3—the corporate standard.

The decision in 1997 to use PI/Microsoft development environment to enable the employees was a masterstroke for Codelco.

It is with no little pride that I saw our highly complex software running flawlessly in this environment and delivering benefits unheard of in the mining industry. My only regret is that I do not see this more in other industrial sites. This is a very good example of how the integration of the desktop software from Microsoft has helped the manufacturing industry. By opening up their technology (e.g. ActiveX, COM, VBA) they have made it possible for smaller vendors to take advantage of the ease of use and built in training one gets with the Microsoft desktop. The technology used at Codelco is far more sophisticated than the most of the telco's and Internet companies that have been held back by their slow adoption of the Microsoft technology.

Conclusions

Many other companies are transforming themselves using Microsoft technology and all of those will be prepared for the next surge referred to earlier, Microsoft.NET. The missing element has been the willingness of companies to believe in their own people instead of the consultants. Web Enabled E-Business is a hoax of the same order as the Y2K hoax of last decade and Codelco is the proof.

Additional discussions about Arie de Geus methods and strategies on how to build an intelligent organization can be found in the book, Fifth Discipline Field Book by Peter Senge and collaborators (Senge et al., 1994).

References

Bascur, O. A. and J. P. Kennedy, Overall Process Effectiveness in Industrial Complexes, In Bascur, O., editor, *Latin American Perspectives*, pages 293–306, Littleton, CO. SME (1998).

Bascur, O. A. and J. P. Kennedy, Mineral Processing Process Optimization and Collaboration at the Industrial Desktop, In *XXI International Mineral Processing Congress*, Rome (2000).

Bascur, O. A. and J. P. Kennedy, Web Enabled Industrial Desktop Enterprise Process Collaboration, In *IFAC Conference*, Helsinki (2000).

Bascur, O. A., Bridging the Gap between Plant Management and Process Control, In Scheiner, B. J., editor, *Emerging Computer Techniques for the Mineral Industry*, pages 73–81, Littleton, CO. SME (1993).

De Geus, A., "Planning as Learning," *Harvard Business Review* (1988).

Goly-Probst, A. and U. Sturner, Innovative Coarse Ore Conveyor System in Chilean Copper Mine, In *Metals, Mining and More*. Siemens (2000).

Kennedy, J. P., Just an Idea (2000). `www.osisoft.com`.

Montoya, R., Division Radomiro Tomic: Its Hydrometallurgical Process, In Bascur, O. A., editor, *Latin American Perspectives*, pages 217–227, Littleton, CO. SME (1998).

Rojas, J. H. and H. M. Valenzuela, Strategic Plan for Automation and Process Management at the Chuquicamata Mine, In Bascur, O. A., editor, *Latin American Perspectives*, pages 281–292, Littleton, CO. SME (1998).

Senge, P., A. Kleiner, C. Roberts, R. Ross, and S. B., *The Fifth Discipline Fieldbook*. Doubleday, NY (1994).

Nonlinear Model Reduction for Optimization Based Control of Transient Chemical Processes

Wolfgang Marquardt*
Lehrstuhl für Prozesstechnik
RWTH Aachen
D-52064 Aachen, Germany

Abstract

Optimization based control aims at maximizing the economical performance of a plant in a transient environment employing nonlinear models. Model quality is crucial for achieving good economical performance. The models have to represent plant behavior with sufficient detail, but the computational complexity must be limited to facilitate real-time optimization on a time horizon oriented at the dominating time constant of the process. This contribution reviews nonlinear model reduction techniques from an optimization based control perspective. The use of different variants of process models in a structured control system is particularly emphasized. Challenging research directions are identified.

Keywords

Fundamental modeling, Nonlinear model reduction, Nonlinear model predictive control, Real-time optimization, Transient operation, Process dynamics

Introduction

Optimization based control refers to the class of control techniques which predict the process behavior by means of a dynamic plant model and *optimize the economical performance* of a process while *satisfying operational constraints*. For continuous processes, predominantly operated at steady-state, this concept is implemented in industrial practice by a real-time optimization system (Marlin and Hrymak, 1997), which by means of a steady-state model, computes the setpoints for the model predictive controller (Henson, 1998; Morari and Lee, 1999; Allgöwer et al., 1999; Rawlings, 2000; Mayne et al., 2000) which itself provides setpoints for the base control systems. In industrial applications, model predictive controllers (MPC) are almost exclusively based on a linear plant model determined from experimental identification using plant test data (Qin and Badgwell, 1996). Nonlinear model predictive control has not yet gained significant industrial interest (Qin and Badgwell, 2000) despite the inherently nonlinear behavior of most process plants. This is partly due to the state of technology which yet neither provides mature tools to assist model development nor sufficiently robust algorithms to reliably solve the optimization problem on-line. However, even if these shortcomings could be overcome, nonlinear control technology will only be applied if the significantly increased effort in designing, implementing and maintaining such controllers leads to a significant improvement of plant economics as compared to state of the art linear control technology.

If only stationary operational phases are considered, the need for nonlinear optimizing control technology can hardly be justified, since only few practical situations such as non-minimum phase behavior or steady-state multiplicity may call for a nonlinear controller. However, chemical process systems are often operated in transient phases, where all process variables are *intentionally time-varying*. Transient operation is not only limited to processes which are of an inherent dynamic nature such as batch and periodically forced processes, but it finds increasing attention also in continuous processes to implement feedstock switching or product grade transitions (Helbig et al., 2000a), to realize cross-functional coordination between units in a plant or a site (Lu, 2000) or to exploit low frequency disturbances in the dynamically changing environment of a plant in supply chain conscious plant operation (Backx et al., 1998).

In *transient plant operation*, the operational envelope of the plant naturally covers a large region of the state space. The dynamics can therefore not adequately be represented by a linear model. Hence, *nonlinear models and nonlinear control techniques* are indispensible to achieve satisfactory performance. However, it is not only the nonlinearity which distinguishes the control problem in transient operation from its stationary counterpart. In transient operation, the control task must be considered from a wider perspective. Instead of maintaining a setpoint or tracking a trajectory given by a superior decision layer and rejecting disturbances, the control system has to achieve the *best possible economical performance* within a constrained operational region despite the dynamically changing environment of the plant (Backx et al., 1998). Hence, the targets in stationary phases (Rawlings, 2000) are replaced by an economical objective in transient phases (Helbig et al., 2000b) which consequently results in an integration of optimizing control and on-line economical (dynamic) optimization by means of receding horizon techniques (Backx et al., 2000). We introduce the term *operation support system* to emphasize the wider scope which goes beyond mere setpoint control, trajectory tracking, and disturbance rejection.

These trends in process operations and control necessitate models of sufficient rigor, which are suitable for

*marquardt@lfpt.rwth-aachen.de

implementation in an optimization based operation support system. Since these models must cover the whole operational envelope of a plant, purely *empirical process models* seem to be unfavorable due to the immense effort required for plant testing (Pearson and Ogunnaike, 1997). Consequently, models for optimization and control should capture the major physico-chemical effects in a mechanistic manner at least to the extent accessible. The more a-priori-knowledge can be built into a *fundamental process model* the less experimental effort will be required to fit the model to plant data in order to obtain good extrapolation capabilities in a large region of the state space.

Modeling is considered one of the major bottlenecks of nonlinear model predictive control (Henson, 1998; Lee, 2000) or, more generally, of optimization based process operations and control in the sense of Backx et al. (1998) and Helbig et al. (2000a). Lee (2000) discusses the requirements, the current status and the needs of nonlinear modeling and identification for control and operations with an emphasis on *experimental identification*. This paper aims at complementing and partly detailing Lee's assessment by focusing on *nonlinear model reduction* for the implementation of optimization based operations support systems.

A comprehensive and sensible review of this subject is a formidable task which can hardly be achieved given the limited space available. Hence, the selection of the material presented reflects at least to some extent the interest and the ignorance of the author. The references given should be taken as exemplary rather than as comprehensive. They have been chosen to point the interested reader to relevant approaches and results and to provide a first guide to a more detailed literature study.

Modeling always has to be oriented towards a projected use in an application (Foss et al., 1998). Hence, the next section introduces first a mathematical formulation of the *optimization based operation support problem*, suggests some decomposition strategies and derives general requirements the models have to fulfill. The major phases of a *systematic work process for the development of (fundamental) process models* are given in the following section. The resulting models show a natural structure which should be exploited in model application. This structure can be related to *hybrid modeling* which is discussed in the following section.

Fundamental models are typically of a high computational complexity which is difficult to handle by online optimization algorithms (Henson, 1998). Therefore, nonlinear model reduction techniques are of significant relevance if large-scale applications have to be tackled. Consequently, model reduction techniques are discussed in great detail next. We consider both, *model order reduction* to reduce the number of equations, and *model simplification* to reduce the complexity of the dynamic process model. With the final section, we return to the *optimization based operation support system* and discuss which type of models are good candidates for implementing the different modules in a potentially decomposed system. We conclude with a summary of important *open research problems*.

Optimization Based Operation Support

We introduce a general problem formulation for optimization based operation support and discuss potential decomposition strategies for implementing such a control system in an industrial environment (see Backx et al. (1998), Helbig et al. (2000b) and Backx et al. (2000) for a more detailed discussion). Resulting requirements on the models are summarized to guide fundamental modeling and model order reduction and simplification.

Mathematical Problem Formulation

The goal of optimal process control and operations is the minimization of some economic cost function Φ_i over a certain time horizon $\Delta_i = [t_{0,i}, t_{f,i}]$, both set by a decision maker on a higher level in the automation hierarchy (e.g. a planner or a scheduler), in face of unknown parametric or unstructured model uncertainty and time-varying exogenous inputs represented by the disturbance vector $\boldsymbol{d}_i(t)$. The minimization is subject to the model equations $\boldsymbol{f}(\cdot)$ and to production and process constraints $\boldsymbol{m}_i(\cdot)$ and $\boldsymbol{g}_i(\cdot)$ such as product quality and capacity restrictions or equipment limitations. The constraints $\boldsymbol{m}_i(\cdot), \boldsymbol{g}_i(\cdot)$ can be either path, point or periodicity constraints. The final time $t_{f,i}$ of the operational phase Δ_i is determined by the operational objectives. It could, for example, be the final batch time or the time after which a grade change in a continuous process has been completed. Feedback of available measurements $\boldsymbol{\eta}_i(t)$ of the system outputs $\boldsymbol{y}_i(t)$, inputs $\boldsymbol{u}_i(t)$ and disturbances $\boldsymbol{d}_i(t)$ is introduced to achieve satisfactory performance. Typically, an estimation of the states $\boldsymbol{x}_i(t)$ and the disturbances $\boldsymbol{d}_i(t)$ is required to implement output feedback with high performance. We drop the index i denoting a specific operational phase subsequently to ease notation in the sequel.

The overall output feedback problem can be separated into a dynamic data and model reconciliation problem and into an optimal control problem. At any time instance $t_c \in \Delta$, the reconciliation problem

$$\min_{x_{r,0}, d_r} \Phi_r(\boldsymbol{y}_r, \boldsymbol{\eta}, \boldsymbol{x}_{r,0}, \boldsymbol{d}_r, t_r, t_c) \tag{1}$$

subject to

$$\begin{aligned} 0 &= \boldsymbol{f}_r(\dot{\boldsymbol{x}}_r, \boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{d}_r) \ , \\ \boldsymbol{x}_r(t_r) &= \boldsymbol{x}_{r,0} \ , \\ \boldsymbol{y}_r &= \boldsymbol{h}_r(\boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{d}_r) \ , \\ \boldsymbol{u}_r &= \mathcal{U}[\boldsymbol{u}_c(\cdot)] \ , \\ 0 &\geq \boldsymbol{g}_r(\boldsymbol{x}_r, \boldsymbol{u}_r, \boldsymbol{d}_r) \ , \end{aligned}$$

$$t \in [t_r, t_c]$$

and the control problem

$$\min_{u_c} \Phi_c(\boldsymbol{x}_c, \boldsymbol{u}_c, t_c, t_f) \tag{2}$$

subject to

$$\begin{aligned} 0 &= \boldsymbol{f}_c(\dot{\boldsymbol{x}}_c, \boldsymbol{x}_c, \boldsymbol{u}_c, \boldsymbol{d}_c) \,, \\ \boldsymbol{x}_c(t_c) &= \boldsymbol{x}_r(t_c) \,, \\ \boldsymbol{y}_c &= \boldsymbol{h}_c(\boldsymbol{x}_c, \boldsymbol{u}_c, \boldsymbol{d}_c) \,, \\ \boldsymbol{d}_c &= \mathcal{D}[\boldsymbol{d}_r(\cdot)] \,, \\ 0 &\geq \boldsymbol{g}_c(\boldsymbol{x}_c, \boldsymbol{u}_c, \boldsymbol{d}_c) \,, \\ 0 &\geq \boldsymbol{m}_c(\boldsymbol{x}_c, \boldsymbol{u}_c, \boldsymbol{d}_c) \,, \end{aligned}$$

$$t \in [t_c, t_f]$$

have to be solved in real-time on horizons $[t_r, t_c]$ and $[t_c, t_f]$, respectively. The indices r and c refer to quantities in the reconciliation and the control problem, respectively. Purposely, we have assumed that neither the models nor the production and process constraints are the same in the reconciliation and control problems. For the sake of simplicity, we have not explicitly introduced the discrete nature of measurements $\boldsymbol{\eta}$ and controls $\boldsymbol{u}_c, \boldsymbol{u}_r$. These vectors could, however, be thought of being concatenations of the respective vectors at discrete times t_k in either $[t_r, t_c]$ or $[t_c, t_f]$. Further, though we do not want to exclude hybrid continuous-discrete systems (e.g. Barton et al., 1998; Bemporad and Morari, 1999), we have not accounted for any discrete variables in the problem formulation explicitly for the sake of a simpler notation.

The problems (1) and (2) are coupled and cannot be solved independently. The states $\boldsymbol{x}_r(t_c)$ and the disturbances $\boldsymbol{d}_r(t)$ are estimated in the reconciliation problem (1) and are passed to the control problem (2) to facilitate state and disturbance prediction via the control model and the predictor $\mathcal{D}$. On the other hand, the control variables $\boldsymbol{u}_c(t)$ are passed from the control problem (2) to the reconciliation problem (1) and are processed by $\mathcal{U}$ to update the controls needed for state and disturbance estimation.

From a control perspective, this problem is an output feedback optimal control problem with general (instead of least-squares) objectives reflecting process economics. Since there are no operational targets in the sense of setpoints or reference trajectories (characteristic for a model predictive control problem, see Rawlings (2000)) the problem may also be interpreted from an operational perspective. Hence, the problem can be considered a generalization of state of the art (steady-state) real-time optimization (Marlin and Hrymak, 1997) which aims at establishing economically optimal transient plant operation (Backx et al., 1998, 2000). In any case, the solution of this *operation support problem* would achieve an integration of advanced (predictive constrained) process control and economical optimization in a transient environment.

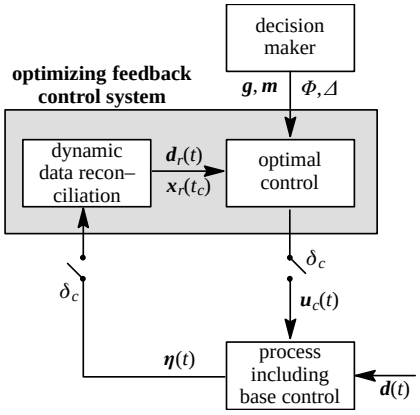


Figure 1: Direct (centralized) optimization approach. δ_c refers to the feedback control sampling time.

Decomposition Strategies

In principle, the problem (1), (2) can be solved simultaneously on the controller sampling frequency δ_c (see Figure 1). This *centralized* or *direct approach* is only computationally tractable if small-scale processes (such as single process units) are considered and/or strongly simplified models are applicable. Obviously, highly efficient solution algorithms and sophisticated model reduction techniques are extremely important to push the frontier of computational complexity (Biegler and Sentoni, 2000). However, even if the problem could be solved easily for large-scale processes, it is questionable whether such an unstructured approach would be accepted by both, the industrial plant operators and the control system vendors. Problems could be associated with a lack of redundancy, reliability and transparency as well as with a high engineering complexity and maintenance effort.

Decomposition of the overall problem seems to be unavoidable in particular if large-scale plant or even site wide control problems with cross-functional integration (Lu, 2000) are considered (Backx et al., 1998). The development of such decomposition strategies can be built on the theory of multi-level hierarchical systems (Mesarovic et al., 1970; Singh, 1977; Findeisen et al., 1980; Morari et al., 1980; Jose and Ungar, 2000) which has been worked out before the mid-eighties. Many of the concepts have not widely been implemented at that time due to a lack of computational power. Though this bottleneck has been largely overcome today, the theory has not yet found adequate attention in the process control community.

Two fundamentally different decomposition strategies can be distinguished. *Horizontal decomposition* refers to

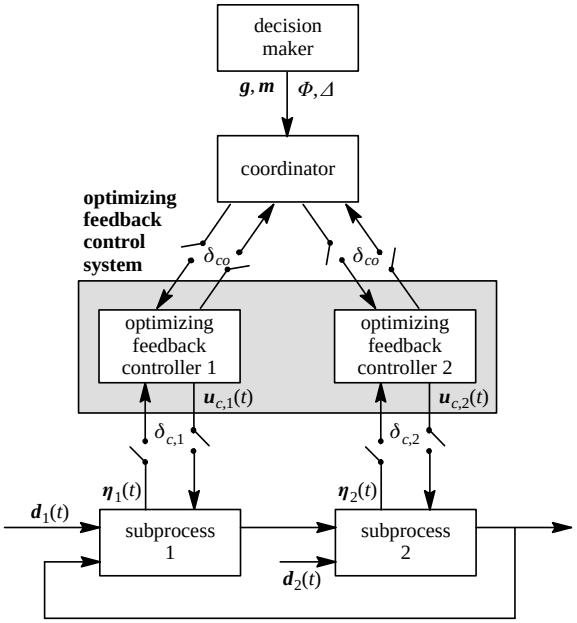


Figure 2: Horizontal decomposition, decentralized optimization approach. δ_c and δ_{co} refer to the feedback control and to the coordination sampling times.

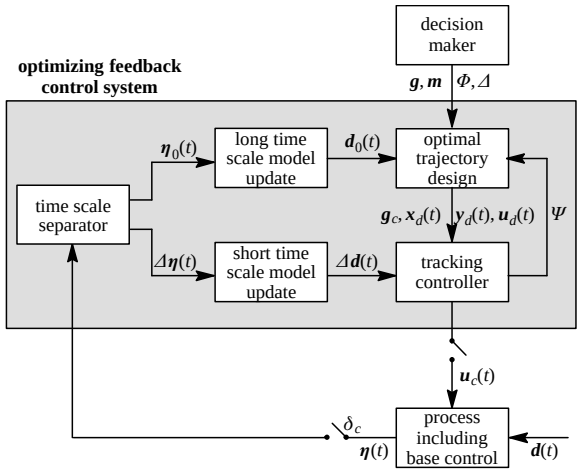


Figure 3: Vertical, two time-scale decomposition of optimization based operations support for transient processes. δ_c refers to the sampling time of the tracking controller, Ψ refers to a process performance indicator.

a decentralization of the control problem typically (but not necessarily, e.g. Lee et al. (2000)) oriented at the functional constituents of a plant (e.g. the process units). Coordination is required to guarantee that the optimal value of the objective reached by a centralized optimizing control system (see Figure 1) can also be achieved by decentralized dynamic optimization. Various coordination strategies for dynamic systems have been described, for example, by Findeisen et al. (1980). Figure 2 shows one possible structure, where the coordinator adjusts the objective functions of the decentralized optimizing feedback controllers to achieve the "true" optimum of the centralized approach.

Vertical decomposition refers to a multi-level separation of the problem (1), (2) with respect to different time-scales. Typically, base control, predictive reference trajectory tracking control, and dynamic economic optimization could be applied with widely differing sampling rates in the range of seconds, minutes, and hours (see Findeisen et al. (1980) for example). According to Helbig et al. (2000b), the feasibility of a multiple time-scale decomposition does not only depend on the dynamic properties of the autonomous system but also on the nature of the exogenous inputs and disturbances. If, for example in a stationary situation, the disturbance can be decomposed into at least two contributions,

$$\boldsymbol{d}(t) = \boldsymbol{d}_0(t) + \Delta\boldsymbol{d}(t)\ , \tag{3}$$

a *slow trend* $\boldsymbol{d}_0(t)$ fully determined by slow frequency contributions and an additional *zero mean contribution* $\Delta\boldsymbol{d}(t)$ containing high frequencies, some sort of decomposition should be feasible. Figure 3 shows a possible structure of the optimization based operations support system in this case. The upper level is responsible for the design of a desired optimal trajectory $\boldsymbol{x}_d(t), \boldsymbol{u}_d(t), \boldsymbol{y}_d(t)$ whereas the lower level is tracking the trajectory set by the upper level. Due to the time varying nature of the disturbances $\boldsymbol{d}(t)$, feedback is not only necessary to adjust the action of the tracking controller but also to adjust the optimal trajectory design to compensate for variations in $\boldsymbol{d}_0(t)$ and $\Delta\boldsymbol{d}(t)$, respectively. The control action $\boldsymbol{u}_c(t)$ is the sum of the desired control trajectory $\boldsymbol{u}_d(t)$ and the tracking controller output $\Delta\boldsymbol{u}(t)$. Reconciliation is based on the slow and fast contributions $\boldsymbol{\eta}_0(t)$ and $\Delta\boldsymbol{\eta}(t)$ separated by a time-scale separation module. Control and trajectory design are typically executed on two distinct sampling intervals δ_c and $k\delta_c$ with integer $k > 1$. The performance of the controller, coded in some indicator Ψ, needs to be monitored and communicated to the trajectory design level to trigger an update of the optimal trajectory in case the controller is not able to achieve acceptable performance. Though this decomposition scheme is largely related to so-called composite control in the singular perturbation literature Kokotovic et al. (1986), the achievable performance will be determined by the way the time-scale separator is implemented.

Model Requirements

Optimization based control requires appropriate models to implement solutions to the reconciliation and control problems.

The model in (2) must predict the cost function, out-

puts and states over the time horizon $[t_c, t_f]$ with sufficient accuracy. Any notable inaccuracy will inevitably result in an economical loss because of a violation of the constraints or a deviation from the true economic optimum. The requirements on the prediction quality of disturbances on $[t_c, t_f]$ are high, since they influence the cost function in (2). This is in contrast to model predictive control, where even a crude disturbance prediction is sufficient to eliminate offset in trajectory tracking. The same (high) accuracy requirements hold in the whole operational envelope covered during nominal plant operation.

The prediction accuracy on the control horizon $[t_c, t_f]$ crucially depends on the quality of the state and disturbance estimates on the reconciliation horizon $[t_r, t_c]$ as determined from a solution of the reconciliation problem (1) employing the accessible measurements. Due to unavoidable model uncertainty, the model needs to be reconciled simultaneously. In the most simple case, a number of carefully chosen model parameters has to be updated periodically. Often, unknown time-varying exogenous functions or plant upsets and operator interaction at discrete time instances complicate the reconciliation problem.

Since all the models are used as part of an optimization algorithm, the gradients with respect to $\boldsymbol{x}_{0,r}$ and the parameterization of $\boldsymbol{d}_r$ in (2) and the parameterization of $\boldsymbol{u}_r$ in (1) must be of high accuracy, too, to avoid unnecessary iterations or even convergence to the wrong optimum (see Biegler et al. (1985) and Ganesh and Biegler (1987) for a discussion in steady-state optimization).

Accuracy requirements are much higher here as compared to setpoint or trajectory tracking feedback control. Since any model uncertainty directly influences plant economics, we cannot rely on feedback only to cope with model uncertainty. From a plant economics point of view, a quantification of the model error would be desirable (Tatrai et al., 1994) though hardly achievable in practice. Model validation against plant and cost data is extremely important and needs to be an integral part of the model development activity. We should keep in mind that the predictive quality of the model has to be assessed in a closed-loop rather than an open-loop mode. For one, the gain and frequency characteristics of the model usually differ in open- and closed-loop. Further, plant measurements can be eventually used to update the model as part of the reconciliation problem (1) to compensate for deficiencies in the predictive capabilities of the model.

Due to the high complexity of any large-scale industrial plant, model reduction has always to be considered. Inevitably, any reduction of the model complexity introduces inaccuracies, which—if not significantly smaller in magnitude than the mismatch between plant and original model—will lead to a loss of economical performance and may even give rise to instability. Despite this fact, a compromise between model complexity and predictive quality must always be achieved (Tatrai et al., 1994) since low sampling frequency control action as a consequence of the high complexity of a very accurate model would also reduce the performance of the operation support system.

Though controllability and observability are properties of the plant rather than the model, these structural properties may get lost in case of simple plant models if not properly accounted for. The same holds for the identifiability of model parameters which is not only a matter of the available measurements but also of the detail built into the model. Of course, these requirements are obvious but difficult to assess in the nonlinear case. There is some evidence that fast time-scales should be eliminated from the model to yield better robustness and more favorable stability properties (Christofides and Daoutidis, 1996; Kumar et al., 1998).

To facilitate numerical treatment, the models should be of differential index one (Brenan et al., 1996; Martinson and Barton, 2000), proper boundary conditions have to provided in case of distributed parameter models (Martinson and Barton, 2000), and singular arcs (Bryson and Ho, 1975; Kadam, 2000) should be avoided. Discontinuities are still difficult (and will be at least expensive) to handle numerically (Barton et al., 1998) and should therefore be avoided if possible.

In summary, in order to implement the optimization based control system (1), (2), sufficiently accurate models are required for the nominal intrinsic process dynamics, for the economical objectives (including product quality), for the exogenous disturbances, and for sensor and actuator characteristics if they are relevant for process economics. Major sources of structured or unstructured uncertainty should at least be identified as a prerequisite for appropriate model updating. Different tailored models are necessary for the control and reconciliation subproblems even in a direct (centralized) approach (cf. Figure 1). If some decomposition of the operation support system is employed (cf. Figures 2 and 3), model decomposition is an additional major concern.

Validation of the model quality has to be accomplished under closed-loop conditions. A brute force approach could rely on the formulation and solution of an optimization problem which assesses feasibility, flexibility, and controllability according to the classification given by Abel et al. (1998). The problem formulation leads to a large-scale bilevel dynamic optimization problem the reliable solution of which is out of reach at this point for most industrially relevant processes.

Systematic Model Development

The discussion of the last section clearly reveals the complexity of the model requirements. It is therefore not

surprising that modeling and the proper validation of the resulting models is (and will be for a long time) the major bottleneck in introducing model-based operation support systems into industrial application.

Recently, Foss et al. (1998) undertook an industrial field study to identify current industrial practice in process modeling. They identified the same major steps organized nearly the same way in a work process in all the companies included in the study, if only a coarse task granularity is considered. There seem to be no generally practiced patterns on the subtask level. A first analysis of the modeling work process on a detailed level has been attempted recently (Marquardt, 1995; Lohmann and Marquardt, 1996; Lohmann, 1998) in the context of work process centered computer-aided modeling support systems (Jarke and Marquardt, 1996; Bogusch et al., 2001). The approach pursued in these studies is promising but does not yet address the requirements of optimization based control sufficiently. More emphasis has to be put on model transformations (including model order reduction and simplification), model structure discrimination and parameter identification as well as closed-loop model validation in the future.

We are far from a recommended work process which would lead us to a reasonable set of models for a decomposed optimization based operations support system at minimal cost. This section presents the major modeling steps on a coarse granular task level to guide the development of more elaborate modeling work processes (see Foss et al., 1998, for details) and to put model reduction and model application as discussed in the remainder of the paper into perspective.

(a) *Requirements analysis:* A precise problem formulation is necessary but often omitted in process modeling since most of the requirements are still vague. As in any design activity these requirements have to evolve with the model during the modeling process. Major issues are the purpose of modeling and the intended application of the model, the quantities to be computed from the model, their dependency on time and non-time coordinates, the accuracy to be attained, the available resources for model construction and the available computational resources for model interpretation.

(b) *Abstraction of the process:* The boundaries of the process under consideration are specified by stating all external connections to the process' environment first. Subsequently, the process is decomposed hierarchically into more and more refined interconnected model objects until a desired level of resolution is reached. The properties of the model objects are described in detail. The information collected comprises an *informal descriptive representation* of the model. The extensive quantities to be balanced, the assumptions on the physico-chemical phenomena and the level of detail to be considered are for example part of this description. Canonical model objects and a recommended procedure have been defined to guide this abstraction process (e.g. Marquardt, 1995).

(c) *Formulation of model equations:* For every model object, the descriptive model of step (b) is cast into a set of model equations to precisely define the model object's dynamics. The *informal descriptive model* is converted into a *formal mathematical model*. First, the balance equations are determined accounting for the desired spatial resolution. The process quantities occurring in the balance equations are classified as states, parameters or state functions. Parameters are fixed together with an uncertainty interval. State functions are refined by additional constitutive equations. Appropriate initial and boundary conditions are specified. Simultaneously, a consistency check of physical dimensions and units, an analysis of the remaining degrees of freedom or of the differential index can be carried out (Bogusch et al., 2001).
The resulting process model comprises partial differential equations to cope with spatially distributed model objects, integro-differential-equations to represent particulate model objects by population balances, as well as differential-algebraic equations to describe spatially averaged (well-mixed) model objects. Often, the models are in addition of a hybrid discrete-continuous nature (Barton and Pantelides, 1994) to represent physical state events (such as a phase change) or discrete control action (such as a switching controller).

(d) *Aggregation of model equations:* The equations of the whole process model are deduced by an aggregation of those of every model object. This aggregation process follows the hierarchical structure introduced during the abstraction in step (b). Additional constraints due to the aggregation may be introduced. Again, the model is checked for its index as well as for a proper specification of degrees of freedom and initial conditions. The resulting model may be very large-scale and may comprise of some hundred thousand equations if plant-wide or even site-wide optimization is envisaged. The model structure may be exploited later to accomplish horizontal decomposition (cf. Figure 2).

(e) *Model transformation:* Usually, the model equations are not solved as derived during the modeling process. Instead, the model equations are reformulated with different objectives. For example, reformulation or even model reduction are performed to reduce the computational complexity or the index in case of high index models.

(f) *Implementation and verification:* The model is implemented by means of a modeling and simulation tool. Instead of a formal verification to check whether the model satisfies the intent of the modeler (expressed in the requirements formulated in step (a)) the model is run, the simulation results are checked for plausibility and the computational resources are determined.

(g) *Structure discrimination, parameter estimation, and model validation:* An appropriate model structure has to be chosen from a set of candidates and unknown model parameters have to be determined from experimental data. Optimization based experimental design may be applied to reduce the number of experiments required. Typically, the model fitting process works bottom-up starting with the model objects on the lowest level of the aggregation hierarchy. Another data set is used subsequently to validate the nominal model. Note, that only open-loop experiments are possible at this point.

(h) *Documentation:* A complete documentation of the modeling process, the resulting process model, its implementation, and its validation is provided to facilitate the use of the model or of its parts in a later application.

(i) *Model application:* The model is employed for the intended application, i.e. for the implementation of one of the functional modules in an optimization based control system. An objective function, a disturbance model and a set of constraints have to be defined in addition. The application is validated as a whole which, most importantly boils down to a validation of the model in closed-loop. Not only the modeling, but also the quality of the numerical solution and the measurement data have to be accounted for at this stage.

There is a very close link and a large degree of interdependency between steps (a) to (i). Consequently, a large number of iterations cannot be avoided to meet the complicated and widely varying requirements set out at the beginning. Though, it would be extremely useful to better manage the modeling process in order to come up with a satisfactory solution the first time right, a formalization of the modeling process as a prerequisite for proper work process management seems to be completely out of reach today. This is largely due to a lack of understanding of the modeling process as whole. Still, nonlinear modeling for control is rather an art than an engineering science (cf. also Aris, 1991).

The remainder of the paper will largely deal with steps (c), (e) and (i) with an emphasis on a reduction of model complexity as part of the model transformations in step (e).

Hybrid Modeling

The formulation of model equations in step (c) of the work process above is largely depending on the level of process knowledge available. This knowledge can—at least in part—be organized along the natural structure displayed in model equation sets as derived during fundamental modeling (Marquardt, 1995). On the uppermost level, there are the balances (e.g. a component mass balance) which are composed of generalized fluxes (e.g. a reaction rate), which may be computed from constitutive equations (e.g. the reaction rate expression). Recursively, these constitutive equations contain variables (e.g. a rate constant) which may result from other constitutive equations (e.g. an Arrhenius law). The equation system can be organized as bipartite graph with equations and variables representing the two types of nodes (Bogusch and Marquardt, 1997).

For every process quantity occurring in an equation we can decide whether it is treated as a constant or even time-varying parameter (to be eventually estimated online) or whether it will be refined by another equation. In the latter case, these "constitutive equations" may have a mechanistic basis, or alternatively, they may be of a physically motivated semi-empirical or even of a completely empirical nature. Obviously, the process quantities occurring in these equations can be treated the same way on a next refinement level. In most cases, we are not able or (for complexity reasons) not interested in incorporating truly mechanistic knowledge (on the molecular level) to determine the constitutive equations. Instead, we correlate unknown process quantities by means of a (semi-)empirical equation.

Hence, all process models are by definition *hybrid models*, since they comprise fundamental as well as empirical model constituents. The fundamental model constituents typically represent the balances of mass, energy and momentum and at least part of the constitutive equations required to fix fluxes and thermodynamic state functions as functions of state variables. Empirical model constituents, on the other hand, are incorporated in the overall process model to compensate for a lack of understanding of the underlying physico-chemical mechanisms. These empirical model constituents are typically formed by some regression model such as a linear multivariate or an artificial neural network model. The parameters in the regression model are identified from plant data. Therefore, hybrid modeling is also often referred to as combining a fundamental and a data-driven (or experimental) approach.

Empirical Regression Models

Many different ways of combining empirical (or data-driven) and fundamental modeling have been proposed in recent years (see van Can et al., 1996; Agarwal, 1997), for various alternatives). The most important structures

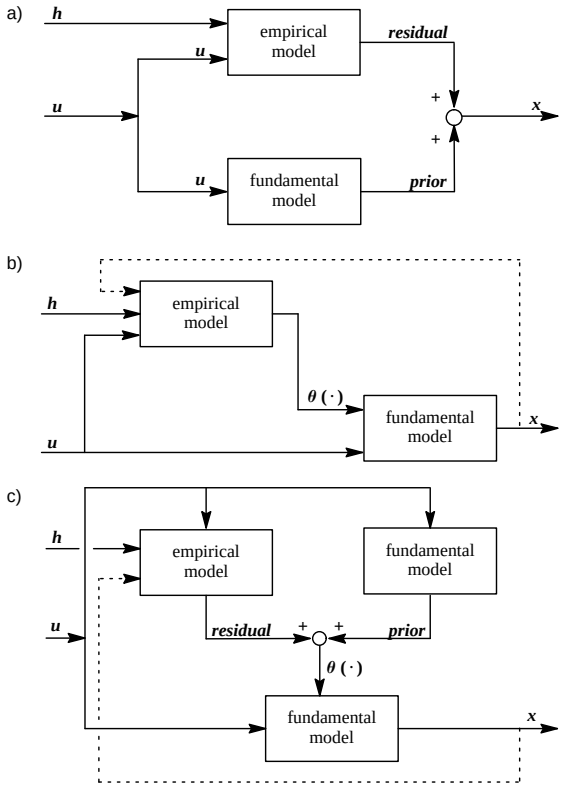


Figure 4: Different structures of hybrid models. (a) parallel structure, (b) serial structure, and (c) an example of a mixed structure.

of combining a fundamental and an empirical model are depicted in Figure 4.

In the *parallel structure*, independently introduced by Su et al. (1992), Kramer and Thompson (1992), Thompson and Kramer (1994), and Johansen and Foss (1992), the model output is a weighted sum of an empirical and a fundamental constituent (cf. Figure 4 (a)). Usually, both of these models are dynamic. The empirical model is often implemented as some type of neural network. It acts as an error model and compensates for any unstructured uncertainty in the fundamental model.

In contrast to this ad-hoc approach to hybrid modeling, the *serial structure* (cf. Figure 4 (b)) is fully consistent with a fundamental model structure. For lumped parameter systems, we find in general on some level of refinement the equation structure

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\theta}(\cdot), \boldsymbol{p}_1, \boldsymbol{u}) \tag{4}$$

with parameters $\boldsymbol{p}_1$ and the unknown function $\boldsymbol{\theta}(\cdot)$. The quantities $\boldsymbol{\theta}$ represent physical quantities which are difficult to model mechanistically. Examples are a flux, a kinetic constant, or a product quality indicator. Typically, $\boldsymbol{f}(\cdot)$ includes the balances of mass and energy, which can always be formulated easily. Obviously, instead of postulating some model structure

$$\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{x}, \boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{p}_2) \tag{5}$$

which is based on some physical hypotheses as in fundamental modeling, any other purely mathematically motivated model structure can be chosen to implement the constitutive equation. This function and the probably unknown parameters $\boldsymbol{p}_1$ in the fundamental model have to be estimated from the measured outputs $\boldsymbol{\eta}$ and the known inputs $\boldsymbol{u}$ after appropriate parameterization by $\boldsymbol{\theta}(\cdot)$ and $\boldsymbol{p}_2$. This approach to hybrid modeling has been introduced first by Psichogios and Ungar (1992) who suggested the use of a feedforward neural network as the regression model. Identifiability is a serious concern, if $\boldsymbol{\theta}(\cdot)$ not only depends on known inputs $\boldsymbol{u}$ and measured outputs $\boldsymbol{\eta}$ as in most reported studies but also on states $\boldsymbol{x}$.

Any combination of the *parallel and serial structures* in Figure 4 is conceivable. An example, reported by Simutis et al. (1997), is shown in Figure 4 (c).

Hybrid models with serial structure have got a lot of attention. They have found numerous applications in different areas such as in catalysis and multiphase reaction engineering (e.g. Molga and Westerterp, 1997; Zander et al., 1999; Molga and Cherbański, 1999), biotechnology (e.g. Saxén and Saxén, 1996; van Can et al., 1998; Shene et al., 1999; Thibault et al., 2000), polymerization (e.g. Tsen et al., 1996), minerals processing (e.g. Reuter et al., 1993; Gupta et al., 1999), drying (e.g. Zbiciński et al., 1996) or in environmental processes (e.g. de Veaux et al., 1999). In most cases reported so far, the hybrid model has been determined off-line. Satisfactory prediction quality can be obtained if sufficient data are available for training. Typically, the interpolation capabilities are comparable to fully empirical models but the extrapolation capabilities are far superior (e.g. van Can et al., 1998; de Veaux et al., 1999). A tutorial introduction to hybrid modeling with emphasis on the serial structure has been given by te Braake et al. (1998).

On-line updating of the neural network model in the context of model-based control has also been suggested in cases where a high predictive quality cannot be obtained by off-line training due to a lack of sufficient data or to a time-varying nature of the process. An example has been recently reported by Costa et al. (1999) who apply optimal control to a fed-batch fermentation. These authors employ a functional link (neural) network to model the reaction rates in a hybrid model with serial structure and update the parameters on-line to improve control performance.

Empirical Trend Models

Employing nonlinear regression with neural networks for the determination of $\boldsymbol{\theta}(\cdot)$ is not the only one approach to hybrid modeling of uncertain systems. For example, the structure of the model (4), (5) has been explored before

the introduction of hybrid neural network models in the area of reaction calorimetry (see Schuler and Schmidt (1992) for a review). In this context, the unknown function $\boldsymbol{\theta}(\cdot)$ refers to the heat of reaction which is inferred from temperature measurements by some state estimation technique. Instead of the static model (5) a dynamic model

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\vartheta}_1(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{p}_3) \,, \tag{6}$$

$$\dot{\boldsymbol{\pi}} = \boldsymbol{\vartheta}_2(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{p}_4) \tag{7}$$

is chosen to complement a fundamental model in the serial structure shown in Figure 4 (b). Here, the quantities $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$ are interpreted as part of the (extended) state vector rather than as a nonlinear state function. Often, due to a lack of mechanistic knowledge, the dynamic models for $\boldsymbol{\theta}$ and $\boldsymbol{\pi}$ are chosen in a simple manner. In many cases, constant or linear trends are sufficient to obtain an estimate, which is completely satisfactory for monitoring and control though the predictive capabilities of the model are very limited in theses cases.

This approach is not restricted to reaction calorimetry. For example, Helbig et al. (1998) report on an extension of this idea to real-time optimization of a two-phase polymerisation reactor operated in semi-batch mode to maximize productivity despite lacking knowledge on the detailed physico-chemical phenomena occurring. Empirical models for the overall reaction rate and the interfacial mass transfer rate have been included in a fundamental model comprising mass and energy balances as well as physical property models of the two-phase system. Simple models (6), (7), have been chosen which guarantee observability and controllability even when control variables are constrained.

Wiener/Hammerstein Type Hybrid Models

Hybrid models of a completely different type may be built as follows. In many cases, a fundamental steady-state process model is available either from process design of from steady-state real-time optimization. Often, the effort of converting the existing steady-state model implemented in some process modeling environment to a dynamic model to be employed for the support of transient process operations is quite high. Dynamic real-time optimization may be accomplished at least approximately by means of a hybrid model which combines the nonlinear fundamental steady-state model with an empirical linear dynamic model in a serial manner. The resulting structure corresponds—in the most simple case of a SISO system—to a Hammerstein or a Wiener model depending on the sequence order of the linear dynamic and the nonlinear static submodels.

Three hybrid models of this type—a Wiener, a Hammerstein and a feedback structure, have been treated recently under the assumption of a known nonlinear static map and unknown linear dynamics by Pearson and Pottmann (2000). These authors report tailored identification algorithms which exploit the knowledge of the nonlinear static map given by the fundamental model. In their binary distillation case study, the nonlinear map between reflux and top composition has been approximated by simple spline functions.

In a similar study, reported previously by Norquay et al. (1999), a hybrid model of Wiener type has been developed and employed in a model predictive control strategy for dual composition control of an industrial C_2-splitter in an ethylene plant. First-order plus dead-time linear transfer functions are used to implement an empirical 2×2 linear model followed by two static maps which result from cubic spline approximation of a fundamental steady-state process model. The Wiener model based controller has been successfully implemented on an industrial plant in Australia.

Discussion

Hybrid trend as well as regression models (in particular with serial structure) are not only compensating for lacking mechanistic knowledge. In most cases these models also result in a much lower complexity as compared to representative detailed mechanistic models. Hybrid modeling is therefore closely related to model reduction which is discussed in more detail below. In many cases, the validity of the model can only be guaranteed if online dynamic model reconciliation by either parameter estimation in case of regression models (e.g. Costa et al. (1999)) or by combined state and disturbance estimation in case of trend models (e.g. Helbig et al. (1998)) is employed.

Though there is some evidence that the serial structure has more favorable extrapolation properties than the parallel structure (van Can et al., 1996), methods of systematically designing hybrid model structures—either employing a static regression or a dynamic trend model—and their ranking with respect to prediction and extrapolation quality under open- and closed-loop conditions are still lacking.

The Hammerstein/Wiener type nonlinear hybrid models are very promising in those cases where a fundamental steady-state real-time optimization model is already available. The work of Norquay et al. (1999) and Pearson and Pottmann (2000) are good starting points for further investigation which should aim at an extension of the concept to real-time dynamic optimization and control. The engineering effort can be reduced tremendously if the optimization model implementation can be directly integrated with a dynamic linear model in a modular fashion by a heterogeneous simulation and optimization platform (e.g. von Wedel and Marquardt (2000)).

Model Order Reduction

We have seen that model complexity can be reduced significantly by introducing empirical components in a model. Here, we will review order reduction techniques for a given nonlinear lumped parameter models which may be of whatever origin. We restrict the discussion to models of type

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \ , \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \ , \quad \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) \ , \tag{8}$$

though more general differential-algebraic models should be considered. However, there seems to be only the paper of Löffler and Marquardt (1991), which treats this relevant class. We attempt to unify—at least to the extent possible—the great many variants of reported nonlinear order reduction techniques. Such an exercise will hopefully uncover hidden relationships and foster new developments in the future. Order reduction is always possible for large-scale models because the most significant contribution to the dynamics originates in a subspace of the complete state space. The key questions are what we rate as a significant contribution and how we are going to reveal it by systematic means ultimately under closed loop conditions.

Projection Methods

Projection methods have been suggested in a great variety in the recent literature. A generic procedure can be formulated as follows:

1. Transform the original state space into a state space better revealing the important contributions to process dynamics, i.e.

$$\boldsymbol{x} - \boldsymbol{x}^* = \boldsymbol{T}(\boldsymbol{z}) \ , \tag{9}$$

   with a general diffeomorphism $\boldsymbol{T}$ and the transformed state vector $\boldsymbol{z} \in R^n$. The reference state $\boldsymbol{x}^*$ is often a non-zero nominal operating point. Note that $\boldsymbol{T}$ is a non-singular square matrix in the linear case.

2. To achieve order reduction we decompose the transformed space into two complementary subspaces with state vectors $\boldsymbol{z}_1 \in R^m$ and $\boldsymbol{z}_2 \in R^{n-m}$, respectively. Hence,

$$\boldsymbol{T}(\boldsymbol{z}) = \boldsymbol{T}(\boldsymbol{z}_1, \boldsymbol{z}_2) \ , \tag{10}$$

   or

$$\boldsymbol{T}\boldsymbol{z} = \boldsymbol{T}_1\boldsymbol{z}_1 + \boldsymbol{T}_2\boldsymbol{z}_2 \ , \quad \boldsymbol{T} = [\boldsymbol{T}_1, \boldsymbol{T}_2] \ , \tag{11}$$

   in case of a linear transformation. We call $\boldsymbol{z}_1$ the dominant states and refer to $\boldsymbol{z}_2$ as the non-dominant states. Note, that the new states are linear combinations of the original states in the linear case.

3. Finally, we have to deduce a nonlinear dynamic model for the dominant states

$$\dot{\boldsymbol{z}}_1 = \boldsymbol{f}_1(\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{u}) \tag{12}$$

   and a hopefully simple algebraic model for the non-dominant states

$$\boldsymbol{0} = \boldsymbol{f}_2(\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{u}) \ . \tag{13}$$

4. Approximate states $\tilde{\boldsymbol{x}}$ and outputs $\tilde{\boldsymbol{y}}$ of the original system can be easily computed from $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ using Equation 9 with Equations 10 or 11. Equation 9 can be viewed as an output equation of the model (12), (13).

The variants of projection methods differ mainly in steps 1–3. The major techniques are introduced and put into perspective next.

Model transformation

Scherpen (1993) suggested nonlinear balancing as a tool for nonlinear model reduction. Her method generalizes balancing of linear systems as introduced by Moore (1981). The idea of balancing is the transformation of a system into an equivalent form which allows the assessment of the importance of the state variables with respect to the energy in its input and output signals. According to Scherpen (1993), we define the controllability and observability functions

$$\begin{aligned} L_c(\boldsymbol{x}_0) = \min_{u, x(0)=x_0} \frac{1}{2} \int_{-\infty}^{0} \boldsymbol{u}(t)^T \boldsymbol{u}(t) dt \ , \\ \boldsymbol{x}(-\infty) = \boldsymbol{0} \ , \quad \boldsymbol{u} \in L_2(-\infty, 0) \ , \end{aligned} \tag{14}$$

$$\begin{aligned} L_o(\boldsymbol{x}_0) = \frac{1}{2} \int_{0}^{\infty} \boldsymbol{y}(t)^T \boldsymbol{y}(t) dt \ , \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \ , \ \boldsymbol{u}(t) \equiv \boldsymbol{0} \ , \ 0 \le t < \infty \ , \end{aligned} \tag{15}$$

for linear and nonlinear systems. These functions denote the amount of input energy required to reach the state $\boldsymbol{x}_0$ and the amount of output energy generated by the state $\boldsymbol{x}_0$, respectively. For a linear system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$, these functions are quadratic forms

$$L_c = \frac{1}{2} \boldsymbol{x}_0^T \boldsymbol{M}_c^{-1} \boldsymbol{x}_0 \ , \quad L_o = \frac{1}{2} \boldsymbol{x}_0^T \boldsymbol{M}_o^{-1} \boldsymbol{x}_0 \tag{16}$$

of the controllability and observability Gramians

$$\boldsymbol{M}_c = \int_0^\infty e^{\boldsymbol{A}t} \boldsymbol{B}\boldsymbol{B}^T e^{\boldsymbol{A}^T t} \, dt \ , \tag{17}$$

$$\boldsymbol{M}_o = \int_0^\infty e^{\boldsymbol{A}^T t} \boldsymbol{C}^T \boldsymbol{C} e^{\boldsymbol{A}t} \, dt \ , \tag{18}$$

which, for stable systems, can be computed from the Lyapunov equations

$$\boldsymbol{A}\boldsymbol{M}_c + \boldsymbol{M}_c\boldsymbol{A}^T = -\boldsymbol{B}\boldsymbol{B}^T \ , \tag{19}$$

$$\boldsymbol{A}^T\boldsymbol{M}_o + \boldsymbol{M}_o\boldsymbol{A} = -\boldsymbol{C}^T\boldsymbol{C} \ . \tag{20}$$

A system is called internally balanced (Moore, 1981) if

$$\boldsymbol{M}_o = \boldsymbol{M}_c = \boldsymbol{\Sigma} = \text{diag}\,\{\sigma_i\} \tag{21}$$

where $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_n$ are the Hankel singular values. There always exists an orthogonal matrix $\boldsymbol{T}$ (with $\boldsymbol{T}^{-1} = \boldsymbol{T}^T$) which transforms a (stable) linear system $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$[1] into its balanced equivalent $(\boldsymbol{T}^T\boldsymbol{A}\boldsymbol{T}, \boldsymbol{T}^T\boldsymbol{B}, \boldsymbol{C}\boldsymbol{T})$ with $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{z}, \boldsymbol{z} = \boldsymbol{T}^T\boldsymbol{x}$ (Moore, 1981). More precisely, there are two transformation applied subsequently which employ the eigenvectors and eigenvalues of both Gramians to arrive at diagonalized controllability and observability Gramians of the transformed system.

The generalization to stable nonlinear systems (8) requires the determination of the nonlinear controllability and observability functions $L_c(\boldsymbol{x}_0)$ and $L_o(\boldsymbol{x}_0)$, respectively. Scherpen (1993) derives two nonlinear partial differential equations, a Lyapunov and a Hamilton-Jacobi type of equation, to determine these functions. Two nonlinear transformations can be derived from these solutions. If they are applied to $L_c(\boldsymbol{x}_0)$ and $L_o(\boldsymbol{x}_0)$ again, n singular value functions, the nonlinear analogs to the singular values in the linear case, can be identified. Hence, these transformations can be used to balance a nonlinear system. A similar technique based on a different definition of the energy of the input and output signals related to the nonlinear H_∞ control problem as well as extensions to nonlinear systems have been reported more recently (Scherpen and van der Schaft, 1994; Scherpen, 1996). The drawback of these approaches is obvious: the required analytical computations to determine $L_c(\boldsymbol{x}_0)$ and $L_o(\boldsymbol{x}_0)$ are rarely feasible in practice and a fully constructive method for the determination of the nonlinear transformation is not yet available. Hence, an approximation of the analytical balancing method of Scherpen seems to be the only way forward.

Consequently, Newman and Krishnaprasad (1998) explored the possibility of determining $L_c(\boldsymbol{x}_0)$ and $L_o(\boldsymbol{x}_0)$ by a Monte-Carlo technique for a two-dimensional problem.

An alternative approach to the computation of approximate Gramians can be deduced from a remark given by Moore (1981, p.21). He suggested to sample the impulse response matrix at a finite number of times to empirically construct an approximate $\boldsymbol{M}_c$ and $\boldsymbol{M}_o$ if solutions to the Lyapunov equations (19), (20) do not exist or are hard to compute.

Pallaske (1987) seems to be the first who made use of this idea in the context of model order reduction for nonlinear systems (without referring to Moore's original work). He introduced the *covariance matrix*

$$\boldsymbol{M} = \int_{\mathcal{G}} \int_0^\infty (\boldsymbol{x}(t) - \boldsymbol{x}^*)(\boldsymbol{x}(t) - \boldsymbol{x}^*)^T \, dt \, d\mathcal{G} \tag{22}$$

[1] See Skogestad and Postlethwaite (1996), pp. 464, for extensions to unstable systems.

as the basis for the derivation of a linear transformation employing an orthogonal matrix $\boldsymbol{T}$. The symbol $\mathcal{G}$ denotes a set of representative trajectories resulting from a variation of initial conditions and input signals. Hence, the covariance matrix averages the behavior of the nonlinear system over a set of representative trajectories.

In order to compute the covariance matrix, the set $\mathcal{G}$ must be parameterized. Pallaske (1987) did not use impulse responses but suggested to keep the controls at constant values matching to the nominal operating point $\boldsymbol{x}^*$ and to parameterize the initial conditions by

$$\boldsymbol{x}_0 = \boldsymbol{x}^* + \boldsymbol{F}\boldsymbol{r} \tag{23}$$

with $\boldsymbol{F} \in R^{n \times s}$ and $\boldsymbol{r} \in R^s$ with $s \ll n$. Pallaske (1987) did not give any recommendations on suitable choices of $\boldsymbol{F}$. Löffler and Marquardt (1991) suggested to select $\boldsymbol{F}$ as the averaged steady-state gain matrix of the system in some region including $\boldsymbol{x}^*$. This suggestion results from a choice of $\mathcal{G}$ as a set of step responses at $\boldsymbol{x}^*, \boldsymbol{u}^*$. Other choices of trajectories $\mathcal{G}$ (or rather matrices $\boldsymbol{F}$) resulting in different covariance matrices $\boldsymbol{M}$ are obviously possible, if they are representative for the system dynamics in some region of the state space and for the intended use of the reduced model.

For nonlinear systems, the covariance matrix can be determined, in principle, by numerical quadrature of the multi-dimensional integral in Equation 22 after a suitable choice of $\boldsymbol{F}$. Highly accurate adaptive parallel algorithms are becoming available to solve this computationally demanding task even on heterogeneous workstation clusters (Čiegis et al., 1997). However, the number of simulations required scales with 2^n. These algorithms can therefore only be used for small to moderate n. Instead, Monte-Carlo techniques may be used to get a coarser approximation for large-scale problems (Jadach, 2000).

In many cases $\boldsymbol{M}$ has to be computed from a linear approximation to get acceptable results (Löffler and Marquardt, 1991). The covariance matrix can be easily determined from an algebraic Lyapunov equation for any parameterization (23). To show this property, we linearize (8) at the reference state $\boldsymbol{x}^*, \boldsymbol{u}^*$ and evaluate the integral in (22) for the set of trajectories $\mathcal{G}$ resulting from the linearized system after a variation of the initial conditions (23) with $\|\boldsymbol{r}\| < \rho$. The resulting covariance matrix $\boldsymbol{M}_l$ is then given by

$$\boldsymbol{M}_l = \int_0^\infty e^{\boldsymbol{A}t} \, \boldsymbol{S} \, e^{\boldsymbol{A}^T t} \, dt \tag{24}$$

with the Jacobian $\boldsymbol{A} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}|_{\boldsymbol{x}^*, \boldsymbol{u}^*}$ and

$$\boldsymbol{S} = k\boldsymbol{F}\boldsymbol{F}^T \,, \quad k = \frac{2\pi^{\frac{s}{2}} \rho^{\frac{s}{2}}}{s(s+2)\Gamma(\frac{s}{2})} \,. \tag{25}$$

The integration can be replaced by the solution of a Lyapunov equation. Löffler and Marquardt (1991) suggested

to use $\boldsymbol{F} = -\boldsymbol{A}^{-1}\boldsymbol{B}$, the static gain at the reference point $\boldsymbol{x}^*, \boldsymbol{u}^*$ to emphasize the input-state relation of the dynamics. An averaged gain in some neighborhood of the reference point can be used instead, i.e. $\boldsymbol{F} = -\overline{\boldsymbol{A}^{-1}\boldsymbol{B}}$, to reflect some of the nonlinearity in the calculation of $\boldsymbol{M}$.

Motivated by the stiffness occurring in many large-scale systems, Pallaske (1987) suggests to choose the transformation such that the dynamics of the states can be approximately captured in a lower dimensional subspace. This translates to a minimization of the variance of the state in the directions of the coordinate axes of the reduced space. The transformation $\boldsymbol{T} = [\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_n]$, with $\boldsymbol{d}_i$ being the normalized eigenvectors of the covariance matrix $\boldsymbol{M}$, results in such a choice (see below).

The close relation of Pallaske's method to model reduction by balancing can be identified as follows. A choice of $\boldsymbol{F} = \boldsymbol{F}_c = \boldsymbol{B} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}|_{\boldsymbol{x}^*,\boldsymbol{u}^*}$ and $\boldsymbol{F} = \boldsymbol{F}_o = \boldsymbol{C}^T = (\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}|_{\boldsymbol{x}^*,\boldsymbol{u}^*})^T$ and ρ determined from (25) with $k = 1$ results in the local controllability or observability Gramian (17), (18) of the system (8) at $\boldsymbol{x}^*$ emphasizing the input-state or the state-output relation of the dynamics. These matrices are used in linear balancing (Moore, 1981) to construct the transformation $\boldsymbol{T}_{c,o}$. This transformation aims at a removal of the weakly controllable and observable subspaces. It is in contrast to Pallaske's objective which is a removal of the fast non-dominant states. Obviously, a transformation determined from a linearization of the nonlinear model does not exactly balance the nonlinear system in the sense of Scherpen (1993) but may qualify as a useful empirical approximation which at least is consistent with the linear theory. In fact, Wisnewski and Doyle III (1996a) successfully demonstrate the applicability of a related approach. They compute the transformation $\boldsymbol{T}$ from the left and right eigenvectors of the Hankel matrix, i.e. the product of the observability and controllability Gramians, of a linearization of the nonlinear model at a stationary reference point.

Empirical balancing of nonlinear systems has been recently introduced by Lall et al. (1999). They suggest empirical controllability and observability Gramians, $\boldsymbol{M}_c$ and $\boldsymbol{M}_o$, which are closely related to Equation 22. Impulse responses of varying magnitude are chosen in the set $\mathcal{G}$ to compute $\boldsymbol{M}_c$, whereas responses to different initial conditions are used in the set $\mathcal{G}$ to compute $\boldsymbol{M}_o$. In this case, Lall et al. (1999) use the covariances of the outputs $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x})$ in Equation 22 instead of the states $\boldsymbol{x}$. These Gramians are used to empirically balance the nonlinear system as in the linear case. This approach has been adopted recently by Hahn and Edgar (Hahn and Edgar, 1999, 2000).

A completely different approach of determining the transformation matrix is reported in structural dynamics (Slaats et al., 1995). As in modal reduction techniques for linear systems (Litz, 1979; Bonvin and Mellichamp, 1982), the transformation matrix $\boldsymbol{T}$ is formed by the dominating modes of the second order model (the eigenvectors associated with complex conjugate eigenvalues). However, these modes are taken as functions of the displacement of a node in a mechanical structure to account for the nonlinearities. Analytical expressions are derived to determine the modes for the model linearized at the initial condition and for first and second order sensitivities of the modes with respect to the nodal positions. The transformation matrix is then computed from these quantities. The method carries over to first order process systems models, but seems only appropriate for less severe nonlinearities in the vicinity of an operating point.

State space decomposition

According to step 2 of the general projection method, the transformed space has to be decomposed next into two subspaces capturing the dominant and the non-dominant states, respectively.

Scherpen (1993) suggests to use the magnitude of the singular value functions occuring in the transformed observability and controllability functions in some domain of the state space as an indication for weakly controllable and observable subspaces. Those states with large values of the singular value functions are grouped into the vector of dominant states $\boldsymbol{z}_1$, and the remaining state variables form $\boldsymbol{z}_2$. The same strategy is employed by Lall et al. (1999) and Hahn and Edgar (1999, 2000). They analyse the singular values of the empirical Gramians and delete those states with small singular values indicating weak observability and controllablility as in the linear case (Moore, 1981).

Pallaske (1987) poses an optimization problem to bound the normalized L_2-error between the approximate and the original state vectors $\tilde{\boldsymbol{x}}$ and $\boldsymbol{x}$ to a user-defined tolerance ε_0 by varying m, the dimension of $\boldsymbol{z}_1$. The solution to this problem is

$$m = \min k \tag{26}$$

subject to

$$\sum_{i=1}^{k} \mu_i \geq (1 - \varepsilon_0^2)\, \mathrm{trace}\,\{\boldsymbol{M}\}$$

with $\mu_1 > \mu_2 \cdots \geq \mu_n$ being the eigenvalues of $\boldsymbol{M}$. Hence, the first m normalized eigenvectors $\boldsymbol{d}_i\,, i = 1, \ldots n$ of $\boldsymbol{M}$ span the subspace for the dominant transformed states. Consequently, $\boldsymbol{T}_1 = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_m]$ and $\boldsymbol{T}_2 = [\boldsymbol{d}_{m+1}, \ldots, \boldsymbol{d}_n]$ in Equation 11. The same approach has been adopted by Löffler and Marquardt (1991).

In some cases, the choice of the dominant states may be based solely on physical insight. This selection is typically done without transformation in the original coordinates.

Formulation of the reduced model

There are different strategies of determining the reduced model equations after the dominant states have been identified. In complete analogy to the linear case, Scherpen (1993, 1996) suggests to simplify the balanced nonlinear model by *truncation* of the balanced state. The non-dominant states $\boldsymbol{z}_2$ are equated to zero, accounting for their negligible influence on the input-output behavior of the system. Since the fully nonlinear balancing method is not applicable in practice, we present this approach in more detail for a linear transformation (11) resulting for example from empirical nonlinear balancing or from Pallaske's method.

After transformation of (8) and subsequent decomposition into two subsystems, we obtain

$$\dot{\boldsymbol{z}}_1 = \boldsymbol{T}_1^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{T}_1\boldsymbol{z}_1 + \boldsymbol{T}_2\boldsymbol{z}_2, \boldsymbol{u}) \,, \tag{27}$$

$$\dot{\boldsymbol{z}}_2 = \boldsymbol{T}_2^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{T}_1\boldsymbol{z}_1 + \boldsymbol{T}_2\boldsymbol{z}_2, \boldsymbol{u}) \,, \tag{28}$$

$$\boldsymbol{z}_1(0) = \boldsymbol{T}_1^T(\boldsymbol{x}_0 - \boldsymbol{x}^*) \,, \tag{29}$$

$$\boldsymbol{z}_2(0) = \boldsymbol{T}_2^T(\boldsymbol{x}_0 - \boldsymbol{x}^*) \tag{30}$$

Truncation of the transformed state yields

$$\dot{\tilde{\boldsymbol{z}}}_1 = \boldsymbol{T}_1^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{T}_1\tilde{\boldsymbol{z}}_1, \boldsymbol{u}) \,, \tag{31}$$

$$\tilde{\boldsymbol{z}}_2 = \boldsymbol{0} \,, \tag{32}$$

a reduced model of order $m < n$. Alternatively, we may assume

$$\dot{\tilde{\boldsymbol{z}}}_2 = \boldsymbol{0} = \boldsymbol{T}_2^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{T}_1\tilde{\boldsymbol{z}}_1 + \boldsymbol{T}_2\tilde{\boldsymbol{z}}_2, \boldsymbol{u}) \tag{33}$$

to form the reduced model (27), (33). This concept, often referred to as *residualization*, is closely related to singular perturbation discussed in more detail below. It results in a differential-algebraic system of the same order as the original model which is still difficult to solve in general. However, this way the (small) contribution of $\boldsymbol{z}_2$ to $\tilde{\boldsymbol{x}}$ is captured at least to some extent. Combinations of truncation and residualization have been suggested in the linear case (Liu and Anderson, 1989) and obviously carry over to the nonlinear case. Truncation as well as residualization strategies have been suggested and successfully applied to nonlinear systems by Pallaske (1987), Löffler and Marquardt (1991) as well as by Hahn and Edgar (1999, 2000).

Approximation of original states

An approximation of the original state $\tilde{\boldsymbol{x}}$ is obtained from

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}^* + \boldsymbol{T}_1\tilde{\boldsymbol{z}}_1 + \boldsymbol{T}_2\tilde{\boldsymbol{z}}_2 \tag{34}$$

according to Equations 9 and 11. The second term vanishes identically in case of truncation. Steady-state accuracy cannot be guaranteed in the general nonlinear case.

Proper Orthogonal Decomposition

Proper orthogonal decomposition (POD), often also called Karhunen-Loeve expansion or method of empirical eigenfunctions (Fukunaga, 1990; Holmes et al., 1996), is somehow related to the projection methods discussed above. This method has gained much attention in fluid dynamics in the context of discovering coherent structures in turbulent flow patterns. Later, the method has been worked out for the construction of low-order models for dynamic (usually distributed parameter) systems with emphasis on fluid dynamical problems (Sirovich, 1987; Aubry et al., 1988; Holmes et al., 1996). POD has found many applications in simulation and optimization of reactive and fluid dynamical systems (e.g. Graham and Kevrekidis, 1996; Kunisch and Volkwein, 1999; Afanasiev and Hinze, 1999) or chemical vapor deposition processes (e.g. Banerjee and Arkun, 1998; Baker and Christofides, 1999). The method comes in a number of variants. We will summarize one of them following the presentation of Ravindran (1999) next.

Assume we have a representative trajectory of (8) for a certain initial condition $\boldsymbol{x}_0$ and control $\boldsymbol{u}(t)$ defined on a finite time interval $[t_0, t_1]$. The trajectory is uniformly sampled for simplicity to form the ensemble $\mathcal{S} = \{\boldsymbol{x}(t_k) - \boldsymbol{x}^*\}_{k=1}^p = \{\Delta\boldsymbol{x}(t_k)\}_{k=1}^p$ containing p data sets of length n which are often called *snapshots*. As before, $\boldsymbol{x}^*$ is the reference which can be either a steady-state or the ensemble average of the snapshots. We are interested in a unit vector $\boldsymbol{d}$ which is in some sense close to the snapshots in $\mathcal{S}$. We may request that $\boldsymbol{d}$ is as parallel as possible to all the snapshots. This requirement leads to the optimization problem

$$\max \frac{1}{p}\sum_{k=1}^{p} \frac{(\Delta\boldsymbol{x}(t_k)^T\boldsymbol{d})^2}{\boldsymbol{d}^T\boldsymbol{d}} . \tag{35}$$

We assume $\boldsymbol{d}$ to be a linear combination of the data, i.e.

$$\boldsymbol{d} = \sum_{k=1}^{p} w_k \Delta\boldsymbol{x}(t_k) \,, \tag{36}$$

and determine the weights w_k to solve the optimization problem (35). Solving this optimization problem is the same as finding the eigenvectors of the *correlation matrix* $\boldsymbol{N}$ with elements

$$N_{i,j} = \Delta\boldsymbol{x}(t_i)^T \Delta\boldsymbol{x}(t_j) \,. \tag{37}$$

Since this matrix is nonnegative Hermitian, it has a complete set of orthogonal eigenvectors $\{\boldsymbol{w}_1, \ldots \boldsymbol{w}_p\}$ along with a set of eigenvalues $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_p$. We can now construct an orthogonal basis $\mathrm{span}\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p\}$ by means of (36) with

$$\boldsymbol{d}_i = \frac{1}{\sqrt{\lambda_i}} \sum_{k=1}^{p} w_{i,k}\, \Delta\boldsymbol{x}(t_k); \,, \quad i = 1, \ldots p \tag{38}$$

where $w_{i,k}$ denote the elements of the eigenvector $\boldsymbol{w}_i$. It can be shown that any approximation of $\boldsymbol{x}(t_k)$ in a subspace spanned by the first $p_1 < p$ basis vectors $\boldsymbol{d}_i$ maximizes the captured energy $\boldsymbol{x}(t_k)^T\boldsymbol{x}(t_k)$ of the data set. Due to this property, we may just use a reduced basis span$\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_{p_1}\}$ with $p_1 \ll p$ to obtain sufficient approximation quality. The value of p_1 is determined after some experimentation. The ratio

$$\kappa = \frac{\sum_{k=1}^{p_1} \lambda_k}{\sum_{k=1}^{p} \lambda_k} \tag{39}$$

indicates the percentage of energy contained in the first p_1 basis vectors. Obviously this ratio should be close to unity. Note, that we therefore do not have to match the number of snapshots (or basis vectors) p to the dimension n of the dynamic system. Often, we want to use $p \ll n$ for convenience, if very large-scale systems are considered, which, for example, may arise after discretization of a distributed parameter system.

There are at least two common ways of determining the basis vectors. Banerjee and Arkun (1998) employ a singular value decomposition and construct the basis from the left singular vectors of $\boldsymbol{N}$. An alternative approach does not rely on the correlation matrix $\boldsymbol{N}$ but on the $n \times p$ *snapshot matrix*

$$\boldsymbol{X} = [\Delta\boldsymbol{x}(t_1), \Delta\boldsymbol{x}(t_2), \ldots \Delta\boldsymbol{x}(t_p)] \tag{40}$$

the columns of which are the snapshots $\Delta\boldsymbol{x}(t_k)$ at t_k (Aling et al., 1996; Shvartsman and Kevrekidis, 1998). Again, the basis is formed by those $p_1 \ll p$ left singular vectors of $\boldsymbol{X}$ which are associated with the largest singular values and hence capture most of the energy in the data set.

The basis constructed from the correlation or snapshot matrices gives rise to a representation of an approximate solution $\tilde{\boldsymbol{x}}$ to (8) by a linear combination of the basis vectors. Hence,

$$\tilde{\boldsymbol{x}} - \boldsymbol{x}^* = \sum_{k=1}^{p_1} a_k \, \boldsymbol{d}_k \; . \tag{41}$$

If the expansion coefficients a_k and the basis vectors $\boldsymbol{d}_k$ are collected in a vector $\boldsymbol{a}_1 = [a_1, a_2, \ldots a_{p_1}] \in R^{p_1}$ and a matrix $\boldsymbol{U}_1 = [\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots \boldsymbol{d}_{p_1}] \in R^{n \times p_1}$ we can rewrite this equation as

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}^* + \boldsymbol{U}_1\boldsymbol{a}_1 \tag{42}$$

which has the same structure as Equation 34 in case of truncating the non-dominant states. The reduced model is

$$\dot{\boldsymbol{a}}_1 = \boldsymbol{U}_1^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{U}_1\boldsymbol{a}_1, \boldsymbol{u}) \; , \tag{43}$$

$$\boldsymbol{a}_1(0) = \boldsymbol{U}_1^T(\boldsymbol{x}_0 - \boldsymbol{x}^*) \; , \tag{44}$$

which has exactly the same appearance as the truncated model (31) resulting from model reduction by projection.

As in the projection methods, truncation is not the only possibility of developing a reduced model. Rather, the full basis spanned by $p < n$ vectors can be employed by summing to p instead to p_1 in the approximation (41). This approach would result in a model structure completely analogous to the system (27), (33). Residualization (e.g. setting $\dot{\boldsymbol{a}}_2$ to zero) or more sophisticated *slaving methods* can be used to reduce the computational effort (e.g. Aling et al., 1997; Shvartsman and Kevrekidis, 1998; Baker and Christofides, 2000). A more detailed discussion will be provided in the next section.

The major difference between POD and projection is the lack of a convergence proof which would guarantee the reduced model to match the original model in case the number of basis functions p_1 approaches the state dimension n. Further, there is no diffeomorphism between the original state space and the space spanned by the empirical eigenvectors $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_p$ in POD which is available in the projection methods discussed above. Otherwise, both types of methods are very similar as they heavily rely on data taken from a series of simulations of the original model (8). However, the data is organized and used differently depending on the intention of model reduction. Here, the basis is constructed from the correlation or snapshot matrix $\boldsymbol{N}$ or $\boldsymbol{X}$ (cf. (37) and (40)) whereas the covariance matrix $\boldsymbol{M}$ (cf. (22)) and its specializations are used in the projection methods.

Slaving

Residualization in projection results in a set of differential-algebraic equations (cf. Equations 27, 33). A similar system is obtained in POD, if all basis vectors are employed but only the first p_1 are used to build the dynamic subsystem. If the algebraic subsystem (cf. (33) in projection method) cannot be solved explicitly, the computational effort cannot be reduced and the model reduction largely fails. Truncation could be employed instead, but a significant loss in approximation accuracy would result inevitably.

A reduction of the computational effort is possible, if we could find an explicit relation between the algebraic and the dynamic variables ($\boldsymbol{z}_2$ and $\boldsymbol{z}_1$ in projection or $\boldsymbol{a}_2$ and $\boldsymbol{a}_1$ in POD methods). Hence, we are looking for a function

$$\boldsymbol{z}_2 = \boldsymbol{\sigma}(\boldsymbol{z}_1) \tag{45}$$

in case of residualization in projection methods, or for equivalent functions in case of residualization in POD. This approach has also been called *slaving* in the recent literature. Its roots are in nonlinear dynamics, where the computation of *approximate inertial manifolds* has some tradition (e.g. Foias and Témam, 1988; Foias et al., 1988). Aling et al. (1997) as well as Shvartsman and Kevrekidis (1998) use this idea in their case studies

on nonlinear model reduction based on POD. Here, we present the concept in the context of projection methods, where it does not seem to have been applied yet.

The algebraic equation (33) forms the starting point of defining the family of maps

$$\begin{aligned}\tilde{\boldsymbol{z}}_2^{(k+1)} &= \boldsymbol{\sigma}_k(\boldsymbol{z}_1, \tilde{\boldsymbol{z}}_2^{(k)}) \\ &=: \tilde{\boldsymbol{z}}_2^{(k)} - \boldsymbol{T}_2^T \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{T}_1\boldsymbol{z}_1 + \boldsymbol{T}_2\tilde{\boldsymbol{z}}_2^{(k)}, \boldsymbol{u}) \; .\end{aligned} \tag{46}$$

If the map is a contraction, Equation 33 can be solved iteratively from the initial guess $\tilde{\boldsymbol{z}}_2^{(0)} = \boldsymbol{0}$ which would be used in truncation. The computational effort can be reduced significantly, if only few iterations are carried out to improve on the initial value. In case of two iterations we find for example

$$\tilde{\boldsymbol{z}}_2 = \boldsymbol{\sigma}_2(\boldsymbol{z}_1, \boldsymbol{\sigma}_1) \; , \quad \boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_1(\boldsymbol{z}_1, \boldsymbol{0}) \; . \tag{47}$$

This approximation can now be used to eliminate $\boldsymbol{z}_2 \approx \tilde{\boldsymbol{z}}_2$ in Equation 27 and to compute $\tilde{\boldsymbol{x}}$ from Equation 34. An accuracy similar to residualization can be achieved by this method but only a model of the same order as in truncation has to be solved.

Equation Residual Minimization Methods

Equation residual minimization methods have been extensively studied in the linear case (e.g. Eitelberg, 1982). A nice generalization to the nonlinear case has been given recently by Lohmann (1994, 1995). He defines a nonlinear reduced model[2] of (8) as

$$\dot{\boldsymbol{z}}_1 = \boldsymbol{V}\boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{W}\boldsymbol{z}_1, \boldsymbol{u}) \; , \tag{48}$$

to compute the approximate states

$$\tilde{\boldsymbol{x}} = \boldsymbol{x}^* + \boldsymbol{W}\boldsymbol{z}_1 \; . \tag{49}$$

This reduced model has the same structure as that resulting from a truncated projection method (cf. Equation 31) or a truncated POD method (cf. Equation 43). The matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ are, however, determined differently. Their elements are the decision variables in a parameter optimization problem which minimizes the residuals of Equations 48, 49. First, $\boldsymbol{W}$ is determined from minimizing the sum of weighted errors

$$\sum_{k=1}^{p} q_{1,k} \left\| \boldsymbol{x}(t_k) - \boldsymbol{x}^* - \boldsymbol{W}\boldsymbol{z}_1(t_k) \right\|^2 \tag{50}$$

using a number of representative trajectories sampled at discrete times t_k. A set of carefully chosen step responses is chosen for this purpose. Next, given $\boldsymbol{W}$, the sum of weighted equation residuals

$$\sum_{k=1}^{p} q_{2,k} \left\| \dot{\tilde{\boldsymbol{z}}}_1(t_i) - \boldsymbol{V}\boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{W}\tilde{\boldsymbol{z}}_1(t_i), \boldsymbol{u}(t_i)) \right\|^2 \tag{51}$$

is minimized to fix the elements of $\boldsymbol{V}$. In contrast to other nonlinear model reduction techniques, steady-state accuracy can be guaranteed by incorporating a steady-state condition as an equality constraint in this residual minimization problem.

The choice of the dominant states $\boldsymbol{z}_1$ can be done on physical insight in the most simple case. However, Lohmann (1994, 1995) also suggests a systematic alternative to the selection of dominant states which usually gives better approximation results. He introduces a transformation of the state space which reveals the dominant modes which is similar to both, the balancing transformation of Moore (1981) and to the transformation of Pallaske (1987). The transformation is applied first, the dominant states are identified, and the residual minimization technique is applied to the transformed system.

A modified version of the Lohmann method has been reported recently by Kordt (1999). It applies to systems where only the dominant states of the original model are entering the nonlinear terms of the system equations whereas the non-dominant states are confined to the linear part. An even more special case has been treated earlier by Hasenjäger (1991). He assumes that nonlinearities occur only in the equations of the dominant states and that these nonlinearities only depend on the dominant states. In this case, it is suggested to first neglect the nonlinear terms, reduce the linear part of the model by any of the linear reduction techniques (e.g. Litz, 1979; Moore, 1981; Eitelberg, 1982; Bonvin and Mellichamp, 1982; Glover, 1984; Samar et al., 1995; Muscato, 2000), and then add the nonlinear terms to the reduced linear model equations.

Perturbation Methods

In many cases, chemical processes are characterized by phenomena on separate time scales. Multiple time scales occur due to orders of magnitude differences in the densities of contacting vapor (or gas) and liquid phases in multi-phase processes, in the thermal capacitances of reactors, in the time constant of chemical reactions, or in the transfer rates of material or energy across phase boundaries. The models (8) describing such multiple time scale systems usually incorporate process parameters varying in a wide range. In the simplest case, where there are only two time scales present, we may identify a small parameter $\epsilon \ll 1$ in (8) to result in

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \epsilon) \; . \tag{52}$$

Let us assume that this equation can be reformulated as

$$\dot{\boldsymbol{x}}_s = \boldsymbol{g}_s(\boldsymbol{x}_s, \boldsymbol{x}_f, \boldsymbol{u}, \epsilon) \; , \tag{53}$$

$$\epsilon\dot{\boldsymbol{x}}_f = \boldsymbol{g}_f(\boldsymbol{x}_s, \boldsymbol{x}_f, \boldsymbol{u}, \epsilon) \; . \tag{54}$$

Here, $\boldsymbol{x}_s \in R^{n_s}, \boldsymbol{x}_f \in R^{n-n_s}$ are denoting the so-called slow and fast state variables, respectively. Obviously,

[2] Lohmann introduces a refined parameterization of the reduced model by splitting the right hand sides into linear and nonlinear terms.

this reformulation is identical to the partitioning of the state vector $\boldsymbol{x}$ into $\boldsymbol{x}_s$ and $\boldsymbol{x}_f$ by some strategy. This system representation is called the *standard form of a singularly perturbed system* (Kokotovic et al., 1986). This model comprises two time scales t and $\tau = \frac{t-t_0}{\epsilon}$ of different magnitude and is therefore called a two-time-scale model.

Basic singular perturbation approach

Since ϵ is a small parameter, the solution of (53), (54) can be determined by means of perturbation methods in the limit of $\epsilon \to 0$. It is given as

$$\boldsymbol{x}_s = \tilde{\boldsymbol{x}}_s + O(\epsilon) \;, \quad \boldsymbol{x}_f = \tilde{\boldsymbol{x}}_f + \boldsymbol{\mu} + O(\epsilon) \tag{55}$$

with $\tilde{\boldsymbol{x}}_s, \tilde{\boldsymbol{x}}_f$ computed from (53), (54) for $\epsilon = 0$. The notation $O(\epsilon)$ is used in the usual sense and refers to small terms of order ϵ. The so-called boundary layer correction $\boldsymbol{\mu}$ can be computed from

$$\frac{d\boldsymbol{\mu}}{d\tau} = \boldsymbol{g}_f(\tilde{\boldsymbol{x}}_s, \tilde{\boldsymbol{x}}_f + \boldsymbol{\mu}, \boldsymbol{u}, 0) \;. \tag{56}$$

It quickly dies out in a short time interval at the beginning of a transient. For an asymptotically stable system (56), the boundary layer correction may be neglected together with the $O(\epsilon)$ contributions to the solution. These considerations result in the simplified differential-algebraic model

$$\dot{\tilde{\boldsymbol{x}}}_s = \boldsymbol{g}_s(\tilde{\boldsymbol{x}}_s, \tilde{\boldsymbol{x}}_f, \boldsymbol{u}, 0) \;, \tag{57}$$

$$\boldsymbol{0} = \boldsymbol{g}_f(\tilde{\boldsymbol{x}}_s, \tilde{\boldsymbol{x}}_f, \boldsymbol{u}, 0) \tag{58}$$

to determine approximations $\tilde{\boldsymbol{x}} = [\tilde{\boldsymbol{x}}_s^T, \tilde{\boldsymbol{x}}_f^T]^T$ of the state $\boldsymbol{x}$. This approximation is often called the quasi-steady-state approximation (QSSA) of (8). Further simplification is possible, if the nonlinear algebraic equations can be solved analytically for $\tilde{\boldsymbol{x}}_f$, i.e.

$$\tilde{\boldsymbol{x}}_f = \boldsymbol{g}_f^{-1}(\tilde{\boldsymbol{x}}_s, \boldsymbol{u}, 0) \;. \tag{59}$$

This case can be interpreted as an exact slaving approach (cf. Equation 45).

If this reduction method is applied to a process model in practice, two issues have to be addressed. First, we have to derive the standard singularly perturbed system (53),(54). This boils down to the determination of n_s and to a proper association of the states $x_i, i = 1 \ldots n$, to the vectors $\boldsymbol{x}_s$ and $\boldsymbol{x}_f$. Second, since the QSSA does often not lead to models of sufficient accuracy, corrections to the QSSA solutions are of particular interest. A number of suggestions found in the literature are discussed in the following paragraphs.

Standard singularly perturbed systems and QSSA

Often, the QSSA is derived heuristically based on a thorough understanding of the underlying physics. It is well known, however, that such an approach does not always work out well (e.g. Tatrai et al., 1994). More rigorous techniques are therefore required. The number of slow states n_s is usually determined from an analysis of the eigenvalues of the linearized system (8) along some representative trajectories. Robertson and Cameron (1997b) relate the separation ratio denoting the distance of two separated clusters of eigenvalues to the model reduction error. Their analysis gives quantitative recommendations for determining n_s and the eigenvalues in the cluster to be retained in a reduced model. The same authors present a homotopy-continuation technique to identify the eigenvalue-state association to determine the set of states concatenated into $\boldsymbol{x}_s$ and $\boldsymbol{x}_f$, respectively. Both techniques are demonstrated on various non-trivial examples, where the fast or slow states could not be identified merely on the basis of physical insight (Robertson and Cameron, 1997a,b; Tatrai et al., 1994).

Duchene and Rouchon (1996) demonstrate on a simple example from reaction kinetics that the QSSA is not coordinate-free, because the method may not lead to useful results in the original state space. These and other authors suggest a simple linear transformation for special reaction systems (e.g. van Breusegem and Bastin, 1991) and for binary distillation columns (Levine and Rouchon, 1991) which lead to more favorable coordinates for a QSSA. The linear transformation suggested by Pallaske (1987) and introduced above could qualify for the general case. By construction, the dominant states $\boldsymbol{z}_1$ represent the dominant (and often slow) states in the transformed state space, whereas $\boldsymbol{z}_2$ denote the non-dominant (often fast) states $\boldsymbol{z}_2$. Residualization in projection methods or POD discussed above is identical with a QSSA in transformed coordinates.

In general, however, the change of coordinates is nonlinear. Existence of a nonlinear map transforming any nonlinear system (8) into standard singularly perturbed form (53),(54) has been investigated by Marino and Kokotovic (1988). These authors give conditions, which assure the two-time-scale property of (52). They also provide general criteria, which guarantee the existence of an ϵ-independent diffeormorphism to transform the two-time-scale system (52) into the standard form (53),(54). A procedure for constructing such a diffeomorphism is given. It reveals a set of integrability conditions the transformation has to suffice. A more refined analysis has been developed more recently by Krishnan and McClamroch (1994) which has been adopted and extended by Kumar et al. (1998). These authors study systems (8) which are affine in the control variables $\boldsymbol{u}$ and in integer powers of a large parameter $\frac{1}{\epsilon}$. Krishnan and McClamroch (1994) give properties sufficient for the system to reveal two-time-scale characteristics and characterize the slow and fast dynamics. They also generalize their results to systems with more than one large parameter. The analysis of Kumar et al. (1998) reveals two distinct cases depending on the properties of the model nonlin-

earities. Both, the transformation as well as the region of the state-space in which the system shows two-time-scale behavior, may either depend on ϵ or not. Again, the nonlinear transformation has to satisfy the set of integrability conditions already identified by Marino and Kokotovic (1988). The analytical determination of the transformation may be restricted to special nonlinear systems of low to moderate order. It is often difficult to obtain.

The methods of Marino and Kokotovic (1988), Krishnan and McClamroch (1994) as well as of Kumar et al. (1998) require the identification of a large parameter. The quality of the resulting reduced model will crucially depend on this choice. While this parameter can often be found based on physical insight, it would be advantageous to have a transformation available which does not require such a maybe arbitrary choice. Nonlinear balancing as introduced by Scherpen (1993, 1996) may qualify as such a transformation at least in those cases where the fast states coincide with the weakly observable and controllable states. The computation of the transformation is, however, even more involved than that suggested by Kumar et al. (1998). Empirical nonlinear balancing (Lall et al., 1999; Hahn and Edgar, 2000) could be used instead, sacrificing however the nonlinear transformation in favor of a linear transformation similar to the approach of Pallaske (1987). The relation between singular perturbation and nonlinear balancing can only be conjectured at this point. The generalization of the linear result of Liu and Anderson (1989) to the nonlinear case is yet an open problem.

Coordinate-free perturbation methods

The quasi steady-state approximation is widely employed in original coordinates (e.g. Robertson and Cameron, 1997a,b) or in heuristically introduced transformed coordinates (e.g. Levine and Rouchon, 1991; van Breusegem and Bastin, 1991; Kumar et al., 1998). In both cases the approximation quality may be limited due to coordinates which are not appropriately revealing the time-scale separation of the fast and slow variables. Therefore, coordinate-free perturbation methods are attractive which do not rely on coordinate transformation but still come up systematically with satisfactory reduced models of type (57),(58).

There are various coordinate-free model reduction methods which not only result in reasonable approximate models but which go beyond the accuracy of the QSSA.

A first coordinate free method is for example reported by Genyuan and Rabitz (1996). They improve on the QSSA by expanding the fast variables in the regular perturbation series $\tilde{\boldsymbol{x}}_f = \tilde{\boldsymbol{x}}_f^{(0)} + \epsilon\tilde{\boldsymbol{x}}_f^{(1)} + \epsilon^2\tilde{\boldsymbol{x}}_f^{(2)} + \ldots$ where the functions $\tilde{\boldsymbol{x}}_f^{(k)}$ only depend on $\tilde{\boldsymbol{x}}_s$. The equations determining these functions follow from a regular perturbation method. The method is computationally attractive in those cases, where the fast equations are linear in the fast states.

A series of methods has been based on a geometrical interpretation of the global nonlinear system dynamics. In fact, they approximate the trajectories of the original system by trajectories on an *attractive invariant manifold* $\mathcal{M}$. A manifold $\mathcal{M}$ is invariant with respect to the vector field $\boldsymbol{f}$ in Equation 52 if $\boldsymbol{f}$ is tangent to $\mathcal{M}$. It is (locally) attractive, if any trajectory (starting close to $\mathcal{M}$) tends to $\mathcal{M}$ as $t \to \infty$. Model reduction means restriction of the system dynamics to this manifold. The key problem is to (at least approximately) obtain a set of equations defining $\mathcal{M}$.

Duchene and Rouchon (1996) present a solution to this problem. Their reduced model can be written as

$$\dot{\tilde{\boldsymbol{x}}}_s = \boldsymbol{C}(\tilde{\boldsymbol{x}}_s, \boldsymbol{\nu})\, \boldsymbol{g}_s(\tilde{\boldsymbol{x}}_s, \boldsymbol{\nu}) \;, \tag{60}$$

$$\boldsymbol{0} = \boldsymbol{g}_s(\tilde{\boldsymbol{x}}_s, \boldsymbol{\nu}) \;, \tag{61}$$

employing the auxiliary variables $\boldsymbol{\nu}$ provided the decomposition of the space into a fast and a slow subspace can be determined beforehand (for example from physical considerations or with the method of Robertson and Cameron (1997a,b). Duchene and Rouchon provide an explicit formula for the symbolic computation of the matrix $\boldsymbol{C}$ which comprises derivatives of the vector fields $\boldsymbol{g}_s$ and $\boldsymbol{g}_f$ with respect to the slow and fast variables. An explicit expression is also provided to determine approximations to the fast states as a function of $\tilde{\boldsymbol{x}}_s$ and $\boldsymbol{\nu}$ if they are of interest. As with all the perturbation techniques, model reduction is most effective if the auxiliary variables can be eliminated symbolically in the differential equations. Otherwise, slaving technique may be used additionally.

Duchene and Rouchon (1996) state that their method is completely equivalent to a method reported earlier by Maas and Pope (1992) if the system truly admits two time-scales. The method of Maas and Pope is also based on the idea of computing the manifold $\mathcal{M}$. However, instead of deriving a system of equations for approximation of the dynamics on $\mathcal{M}$ as Duchene and Rouchon (1996), their algorithm only determines a series of points to approximate $\mathcal{M}$ itself. These data points have to satisfy a set of n_f constraints. Rhodes et al. (1999) have suggested just recently to use this data and employ black-box identification to relate the fast states $\boldsymbol{x}_f$ to the slow states $\boldsymbol{x}_s$ by some explicit nonlinear function

$$\boldsymbol{x}_f = \boldsymbol{\sigma}(\boldsymbol{x}_s) \;. \tag{62}$$

This way a reduced order model of dimension n_s can be obtained without the need of deriving a singularly perturbed system in standard form and regardless whether the fast subsystem can be solved explicitly. Note that this technique is completely equivalent to the idea of slaving as employed in the context of POD (e.g. Aling et al., 1997; Shvartsman and Kevrekidis, 1998). Instead of the

Maas and Pope algorithm, the computational method reported by Davis and Skodje (1999) could be used to generate data points of an approximation to $\mathcal{M}$ which is then used to build the correlation (62) as suggested by Rhodes et al. (1999).

Obviously, this technique can be applied to any other singular perturbation or projection method. The algebraic equation (58) of the QSSA or (61) can be sampled for given values of $\boldsymbol{x}_s$. This data set can then be used to determine an expression (62) which can be used to eliminate the fast states $\boldsymbol{x}_f$ in (58) or the auxiliary variables $\boldsymbol{\nu}$ in (60).

It should be noted that all the methods discussed in this section require the partitioning of the original state vector $\boldsymbol{x}$ into fast and slow variables $\boldsymbol{x}_f$ and $\boldsymbol{x}_s$. The method reported by Robertson and Cameron (1997b)—though cumbersome—seems to be most suitable for this purpose.

Remarks on Distributed Parameter Systems

So far, distributed parameter systems have been represented by a model of type (8) employing either averaging over a spatial domain as part of the modeling procedure or by discretizing the spatial coordinates of a partial differential equation (PDE) model. Hence, the infinite-dimensional model has been first reduced to some potentially high order model (8) which then may be subject to model order reduction as reviewed above. Alternatively, order reduction could directly be applied to the infinite-dimensional PDE model to avoid often heuristic finite-dimensional approximate modeling. There is a lot of literature dealing with this problem which would justify a review in its own. Only a few references are given here as a starting point for the interested reader.

The rigorous model reduction approaches for nonlinear PDE models include Galerkin projection involving empirical (e.g. Holmes et al., 1996) or modal eigenfunctions (e.g. Armaou and Christofides, 2000) as well as weighted residuals methods of various kinds (e.g. Villadsen and Michelsen, 1978; Cho and Joseph, 1983; Stewart et al., 1985; Tali-Maamar et al., 1994). Better prediction quality can usually be obtained if the truncated contributions in the series expansion are captured by the approximate inertial manifold (e.g. Christofides and Daoutidis, 1997; Shvartsman and Kevrekidis, 1998; Armaou and Christofides, 2000). These techniques are closely related to projection and proper orthogonal decomposition as discussed above. Singular perturbation techniques have also been applied directly to PDE models. For example, Dochain and Bouaziz (1994) propose a low order model for the exit concentrations of a flow bioreactor.

Extremely compact low order models can be derived for those distributed parameter systems which show wave propagation characteristics such as separation and reaction processes (Marquardt, 1990). The major state variables comprise the spatial position of the wave front and some properties of the shape of the wave. This concept has been applied most notably to fixed-bed reactors (e.g. Gilles and Epple, 1981; Epple, 1986; Doyle III et al., 1996) as well as to binary or multi-component as well as reactive distillation columns (e.g. Gilles et al., 1980; Marquardt and Gilles, 1990; Hwang, 1991; Han and Park, 1993; Balasubramhanya and Doyle III, 2000; Kienle, 2000).

The results available indicate that reduction techniques for PDE models should be seriously considered at least in the sense of a first model reduction step in particular in a plant-wide model reduction problem to derive an approximate lumped parameter model of type (8) for some of the process units.

Discussion

The review in this section shows a large variety of nonlinear model reduction techniques stemming from different scientific areas. This is largely due to the lack of a unifying theory which could be used to guide the model reduction process. Truly nonlinear approaches with a sound theoretical basis are those singular perturbation techniques, which rely on some approximation of the attractive invariant manifold of the dynamical system (e.g. Duchene and Rouchon, 1996; Rhodes et al., 1999; Davis and Skodje, 1999) and nonlinear balancing techniques (Scherpen, 1993, 1996). As always in nonlinear theory, the computations are tedious or even infeasible—in particular if large-scale problems have to be tackled. An interesting alternative are those projection methods which incorporate the nonlinearity of the system in the reduction procedure at least to some extent (e.g. Pallaske, 1987; Hahn and Edgar, 1999, 2000). Some theoretical justification is available from their close relation to a more general nonlinear theory. POD has gained significant interest in recent years in particular for very large-scale processes which often occur as a result of discretizing distributed parameter systems despite their lack of theoretical foundation.

At this point, there is neither evidence whether any of the nonlinear model reduction techniques could qualify as a generic method which gives good results for any process system, nor are there guidelines available which of them to prefer for a particular class of process systems. A selection of results for type (8) models are presented in Table 1. The reductions presented are those suggested by the authors to give satisfactory results. A quantitative comparison is almost impossible and should not be attempted. Obviously, significant order reduction has only been achieved for those processes which have a distributed nature (i.e. the distillation column, fixed bed reactor, pulp digester and rapid thermal processing cases in Table 1). In these cases, model reduction based on nonlinear wave propagation can lead to even higher levels of reduction. However, in all reported studies, the

authors	system	original	reduced
Löffler and Marquardt (1991)	fixed-bed reactor	80 DAE	6 DAE
Tatrai et al. (1994)	FCC unit	20 ODE	15 ODE
Lohmann (1994)	vehicle suspension	10 ODE	7 ODE
Wisnewski and Doyle III (1996a)	continuous pulp digester	210 ODE	17 ODE
Robertson and Cameron (1997b)	evaporator	15 ODE, 30 AE	8 ODE, 34 AE
Robertson and Cameron (1997a)	compressor	51 ODE, 70 AE	20 ODE, 91 AE
Kumar et al. (1998)	CSTR	5 ODE	3 ODE
Aling et al. (1997)	rapid thermal processing	5060 ODE	10 ODE
Hahn and Edgar (1999)	distillation column	32 ODE	3 ODE
Kordt (1999)	aircraft	29 ODE	10 ODE
Hahn and Edgar (2000)	CSTR	6 ODE	4 ODE

Table 1: Selected results of nonlinear model reduction.

complexity of the reduced order model equations is significantly higher than that of the original model. For projection and POD methods, this is due to the linear combination of all the right hand sides of the original model in any equation of the reduced order model (cf. Equations 27, 43). Despite this increase in complexity, significant reductions in computational time have been observed in dynamic simulation in most cases summarized in Table 1. Plant wide models have not yet been considered. Also, the computational complexity as well as the model quality under closed loop conditions and in particular in optimization based controllers has not yet been studied in the context nonlinear model reduction.

More theoretical analysis with an emphasis on closed-loop properties and a reduction of the computational load in dynamic optimization as well as comparative studies on realistic large-scale problems are required to build up more experience which could guide the model order reduction process in a concrete context. Order reduction must be considered to be complemented by model simplification which is introduced in the following section.

Model Simplification

Model simplification is a special type of model reduction where the order of the model is preserved but the complexity of the functional expressions in the model equations is reduced. Since the computational effort is to a large extent determined by the function evaluation of the model equations, such methods are at least as important as model order reduction techniques. We will briefly present promising approaches which are applicable in general and two exemplary areas specific to chemical process systems models.

Linearization

The classical approach to the simplification of nonlinear models for control is a linearization at some nominal operating point. In particular, in a model predictive control framework, the computational complexity can be reduced drastically and the reliability and robustness of the optimization algorithms can be improved significantly, since a (convex) quadratic program has to be solved online instead of a (nonconvex) nonlinear program.

However, since the control system is required to operate in a large operational envelope with satisfactory performance, a linear model resulting from mere Jacobian linearization of the fundamental model at a nominal operating point will not suffice to adequately predict process dynamics. Instead, feedback linearization (Isidori, 1989) of the nonlinear fundamental model can be applied to produce a linear system with a set of state dependent constraints (Nevistić and Morari, 1995; Kurtz and Henson, 1997, 1998). An algorithm close to linear model predictive control can then be applied to handle the constraints. Though conceptually attractive, these techniques are limited to feedback linearizable (small-scale) processes (Morari and Lee, 1999). Therefore, they are not expected to get significant attention for optimization based control of industrial processes.

Instead of feedback linearization, Jacobian linearization at different reference points along a transient trajectory can be envisioned. Many variants of this modeling approach have been reported in the recent literature (e.g. García, 1984; Gattu and Zafiriou, 1992; Lee and Ricker, 1994) to limit the complexity of the optimization in nonlinear model predictive control. Most of the reported studies have been limited to low order models. If large-scale systems are considered, linear model reduction (e.g. Litz, 1979; Moore, 1981; Eitelberg, 1982; Bonvin and Mellichamp, 1982; Glover, 1984; Samar et al., 1995; Muscato, 2000) can be applied to reduce the computational load in a predictive control strategy. Wisnewski and Doyle III (1996b) and Doyle III and Wisnewski (2000) use such a strategy. They keep the reduced nonlinear model constant to avoid the computational burden of on-line linear model reduction along the trajectory.

An alternative to successive linearization along the tra-

jectory are interpolated piecewise linear models which are valid only locally in a certain region of the operational envelope. Different realizations of this idea have been reported for example by Banerjee et al. (1997), Johansen and Foss (1997), Chikkula et al. (1998), Banerjee and Arkun (1998), Lakshmanan and Arkun (1999), Foss et al. (2000), or Dharaskar and Gupta (2000). Though these authors largely aim at experimental identification to develop the local linear models, they could also be constructed by linearization of the fundamental model at a number of reference points followed by subsequent linear model reduction. The local models are then glued together by some interpolation strategy to provide an aggregated model valid in the whole operational envelope. At a first glance, these approaches seem to be more favorable than successive linearization and subsequent on-line model reduction, since model building can be done off-line. This is, however, not completely true, since the parameters of the fundamental model have to be adjusted on-line as part of the optimizing control system to reduce plant-model mismatch. Therefore, a tailored approach is required to adapt the piecewise linear model on-line to plant data or to the updated fundamental model.

Nonlinear Approximation of Functional Expressions

Often, the model equations contain quite complicated nonlinear expressions which result from detailed fundamental modeling and/or from subsequent nonlinear model reduction. In many cases, the right hand sides of the differential-algebraic models are formed by some nonlinear function, which comprises a number of additive, mostly nonlinear terms according to

$$\phi(\boldsymbol{x},\boldsymbol{u}) = \sum_{j=1}^{n_t} \alpha_j \phi_j(\boldsymbol{x},\boldsymbol{u}) \; . \tag{63}$$

Here, ϕ and ϕ_j are scalar functions, α_j are constant weights and $\boldsymbol{x}$ and $\boldsymbol{u}$ are vectors of given states and inputs which vary with time. Functions of this type arise, for example, in reaction kinetic models, or inevitably in reduced order models if derived by projection (cf. Equations 31, 43). In these cases, the right hand sides of the model are always linear combinations of nonlinear functions, which either comprise the right hand sides of the original model in case of a reduced order model (cf. Equation 31) or the reaction rates of elementary reactions in a reaction kinetic model.

We are interested in systematic methods which replace the probably complex functional expression of ϕ by a simpler functional expression $\tilde{\phi}$ which approximates ϕ up to a user specified tolerance for a set of trajectories denoted by $\mathcal{G}$. Obviously, the problem can be generalized by replacing the linear combination in (63) by a general nonlinear expression, which would lead to the problem of approximating a general function $\phi(\boldsymbol{x},\boldsymbol{u},\boldsymbol{\alpha})$ by some simpler function $\tilde{\phi}(\boldsymbol{x},\boldsymbol{u},\tilde{\boldsymbol{\alpha}})$.

At a first glance, this seems to be a classical problem of multi-variate nonlinear approximation, for which many solution techniques should be readily available. However, the problem is quite complicated due to the probably large number of independent variables occurring as arguments of ϕ_j and due to the fact that the approximation should cover a set of trajectories $\mathcal{G}$. Further, we have a combinatorial component in the problem because there is no preferred candidate functional structure for $\tilde{\phi}$ a priori.

Desrochers and Al-Jaar (1985) have studied a closely related problem in a discrete-time setting. Their problem formulation can be met, if the trajectories in $\mathcal{G}$ are combined to one composite trajectory (by putting them in a sequence in time). The composite trajectory is then sampled on some time grid to result in a sequence $\{\boldsymbol{x}_k,\boldsymbol{u}_k,\phi_k\},\ k = 1\ldots K$. Their approximation problem can be reformulated as the mixed-integer nonlinear programming problem (MINLP)

$$\min_{y,\tilde{\alpha}} \sum_{k=1}^{K} e_k^2 + \boldsymbol{w}^T \boldsymbol{y} \tag{64}$$

subject to

$$\begin{aligned} e_k &= \phi_k - \tilde{\phi}_k \\ &:= \phi_k - \sum_{j=1}^{n_t} y_j \tilde{\alpha}_j \phi_j(\boldsymbol{x}_k,\boldsymbol{u}_k) \; , \\ 0 < \sum_{j=1}^{n_t} y_j &< n_t \; , \qquad \boldsymbol{y} \in \{0,1\}^{n_t} \; . \end{aligned}$$

This problem can be solved (after an appropriate reformulation to replace the disjunctions by a more favorable constraint set) by any MINLP method (at high computational expense). However, an elegant tailored solution technique has been reported by Desrochers and Al-Jaar (1985). Their method completely decouples the combinatorial part of the problem from the parameter identification problem. They first identify the most promising combination of functions ϕ_j in the simplified model on the basis of the residual error and then solve a single parameter estimation problem for the most favorable model structure. The penalty term in the objective can be chosen to account for those terms (and variables) which are preferably eliminated to directly influence the sparsity pattern of a model equation. It should be noted that—at least in some cases—some variables (and hence equations) may be eliminated simultaneously, if they only occur in the discarded functions ϕ_j.

A related technique to model simplification, specifically tailored to rapid thermal processing, a microelectronics manufacturing process, has been reported by Aling et al. (1997). Their objective is to further simplify

a nonlinear model stemming from proper orthogonal decomposition.

Obviously, there is no need to rely on expressions ϕ_j which are already present in ϕ to form the approximation $\tilde{\phi}$. Rather, any functional structure could be postulated for $\tilde{\phi}$. For example, Duchene and Rouchon (1996) suggest to consider multivariate interpolation and approximation techniques. However, such an approach seems to be impractical if the number of arguments of ϕ is large.

A more promising approach could be built on methods developed for nonlinear empirical modeling. For example McKay et al. (1997) and Marenbach et al. (1997) present a method for the identification of the structure and the parameters of nonlinear models for steady-state and dynamic processes, respectively. In their approach, process models are postulated to consist of a given set of elementary functional building blocks $\phi_i(\boldsymbol{x}, \boldsymbol{u}, \tilde{\boldsymbol{\alpha}})$. In contrast to Equation 63, these functions depend nonlinearly on unknown parameters $\tilde{\boldsymbol{\alpha}}$. They are combined in a nonlinear fashion to form the approximation $\tilde{\phi}$. Genetic programming is applied to select the best combination and to determine appropriate parameters to get the best fit of measurements. Obviously, this formulation generalizes problem (64) at the expense of a significantly higher computational complexity.

The principle advantage of this kind of methods lies in the ability to use knowledge on favorable functional forms available from fundamental modeling in defining a set of candidate building blocks. Alternatively, one could employ truly black-box nonlinear identification methods such as neural networks. For example, Shvartsman et al. (2000) report on simplification of a reduced order model of a distributed reaction system derived by proper orthogonal decomposition.

Simplification of Chemical Kinetics Models

Large-scale chemical kinetics models arise in many applications. Model complexity stems from the large number of reactions and components. For the simplification of reaction mechanisms we assume a reaction network with n_r reactions and n_s species. The complexity of the reaction kinetics model can be reduced by eliminating both, reactions and species from the reaction network. Elimination of reactions corresponds to model simplification, whereas elimination of species is a special case of order reduction. Here, we focus therefore on the first problem.

Sensitivity analysis is the most classical approach to assess the importance of individual reactions on the evolution of the concentration of all species (e.g. Seigneur et al., 1982; Brown et al., 1997, and the references cited therein). These methods determine the effect of a perturbation in a kinetic rate constant on the concentrations at some point in time or on average during the course of the reaction. Those reactions with rate constants resulting in a large sensitivity of the concentrations are considered important and should be retained in the reaction kinetics model, whereas those reactions leading to small sensitivity can be eliminated without sacrificing prediction accuracy. Sensitivity methods have been successfully applied to a variety of large-scale reaction mechanisms. However, sensitivity analysis may lead to wrong results as illustrated by means of a simple example by Petzold and Zhu (1999). Therefore, optimization based techniques have been suggested more recently by Edwards et al. (1998), Petzold and Zhu (1999), Edwards and Edgar (2000), Edwards et al. (2000), and Androulakis (2000) for a reduction of the number of reactions in a network. The problem formulations presented by these authors are variants of the MINLP in the previous section and aim at facilitating the numerical solution for large-scale problems. However, the parameters α_i in (64) are the stoichiometric coefficients of the reaction model and are (usually) not considered as degrees of freedom in model simplification.

Simplification of Physical Property Models

A classical example of reducing the computational complexity of a process model is related to the simplification of physical property models. The development of local thermodynamic models dates back into the seventies. This research has been initiated by the observation of the large fraction of computational time spent with physical property calculations in steady-state flowsheeting (Grens, 1983). The calculation of K-values in phase equilibrium models

$$y_i = K_i(\boldsymbol{x}, \boldsymbol{y}, p, T)\, x_i \;, \quad i = 1 \ldots n_c \;, \tag{65}$$

for ideal as well as strongly nonideal mixture has got particular attention due the high complexity of the models. Here, $\boldsymbol{x}, \boldsymbol{y}, p$ and T are the liquid and vapor concentrations as well as pressure and temperature under equilibrium conditions. A local model is intended to approximate the K-values as well as their derivatives with a functional expression of strongly reduced complexity. The structure of the local models is derived on physical arguments. For example, Leesley and Heyen (1977) neglect concentration dependencies and suggest a modification of Raoult's law, whereas Chimowitz and coworkers (Chimowitz et al., 1983; Chimowitz and Lee, 1985), Hager (1992) and Ledent and Heyen (1994) consider concentration dependencies by modified Porter or Margules models. Hager's equation, for example, is

$$\ln(K_i p) = A_{i,1} + \frac{A_{i,2}}{T} + (B_{i,1} + \frac{B_{i,2}}{T})(1 - x_i)^2 + B_{i,3}(1 - x_i)^2(1 + 2x_i) \;. \tag{66}$$

The local models are only valid in a limited region of the operating envelope. Hence, at least some of the model parameters ($B_{i,1}, B_{i,2}, B_{i,3}$ in the example given) must be updated along a trajectory in order to retain sufficient approximation accuracy. The parameter update can be

triggered by the simulation or optimization algorithm or by an estimate of the error between the approximate local and the original models. Model parameters are obtained from some least-squares fit of data obtained from the original model. Various variants of updating schemes have been reported by Leesley and Heyen (1977), Macchietto et al. (1986), Hillestad et al. (1989) and by Storen and Hertzberg (1997). Obviously, parameter updates result in model discontinuities. If not properly handled, these discontinuities will make simulation and optimization algorithms fail or converge to wrong solutions (Barton et al., 1998). Hence, explicit discontinuity handling or discontinuity smoothing is a necessity with these models. For the latter approach, interpolation strategies employed in linear multiple models (e.g. Foss et al., 2000; Johansen and Foss, 1997)) could be adopted here.

Significant savings in computational time have been reported for steady-state simulation and optimization (Chimowitz et al., 1984; Perregaard, 1993), dynamic simulation (Macchietto et al., 1986; Hager, 1992; Perregaard, 1993; Ledent and Heyen, 1994) and dynamic optimization (Storen and Hertzberg, 1997) if local thermodynamic models are applied.

Discussion

Nonlinear model simplification has not yet got significant attention in the systems and control literature. It is particularly suited to simplify reduced order models arising from projection methods with the objective to regain at least to some extent sparsity in the reduced model Jacobian. There has been significant activity in the context of chemical kinetics and physical property models. The variety of techniques tailored to these special problems deserve careful analysis in order to assess the potential of applying the specific concept after generalization to other model simplification problems.

For example, sensitivity analysis as worked out in chemical kinetics, is applicable in principle to the simplification of any parametric model (cf. the derivation by Seigneur et al., 1982) but—to the author's knowledge—it has not been explored for general model simplification problems. This is also true for optimization based methods given the close correspondence between reaction model and general model simplification.

On the other hand, the success of local physical property models suggests to consider similar strategies in a more general setting. The simplified model should be based on a fundamental principle rather than on some arbitrary empirical ansatz. In many cases, the simple models are only of sufficient accuracy in a limited region of the operating envelope. Then, adaptive updating of the parameters of a simple model structure using data from a rigorous model can be considered as an interesting alternative to globally valid simplified models. Obviously, a compromise needs to be established between model complexity and range of model validity. For example, a globally valid but complex neural network model (e.g. Kan and Lee (1996) for a liquid-liquid equilibrium model or Molga and Cherbański (1999) for a liquid-liquid reaction model) can be used instead of a simpler (local) model with a limited region of validity which requires parameter updating along the trajectory.

There is an obvious relation between model simplification and hybrid models as discussed above. Hybrid models are motivated by a lack of knowledge on the mechanistic details of some physico-chemical phenomena. A nonlinear regression model such as a neural network is used instead of a fundamental model to predict some process quantity (such as a reaction rate, a mass transfer rate or phase equilibrium concentrations). Model simplification on the other hand aims at reducing the complexity of a given fundamental model. Hence, hybrid modeling in the sense of Psichogios and Ungar (1992) can be readily applied to model simplification. The (typically algebraic) mechanistic model, which—for example—determines a flux (e.g. a reaction rate, see Molga and Cherbański, 1999), or a separation product flow rate, (see Safavi et al., 1999), a kinetic coefficient (e.g. a flotation rate constant, see Gupta et al., 1999), some state function (e.g. a holdup in a two-phase system, see Gupta et al., 1999) is replaced by some nonlinear regression model (such as a neural network). This regression model is typically explicit in the quantity of interest and hence can be evaluated extremely efficiently. Note, that these models can be designed for a large or a small region of validity. In the latter case, parameter updating (using the rigorous model to produce the data required) is required along the trajectory.

Model Application

We assume that a detailed dynamic model is available for example from the process design activities. This model can be simplified or reduced by physical insight, by one of the techniques discussed above, or by a combination thereof to meet the requirement of the various model-based tasks in integrated dynamic optimization and control system following a direct or some decomposition approach. This section summarizes some thoughts about the type of reduced and/or simplified models which might be used most appropriately in a certain context.

Direct Approach

We can make use of any model and apply an optimizing predictive control and a suitable reconciliation scheme to realize dynamic real-time optimization by the direct approach. Instead of following a reference trajectory set by some upper decision layer in the control hierarchy, an economical objective is maximized on-line on the receding control horizon to compute the control moves. The computational complexity of the reconciliation and con-

trol problems must, however, be quite low, since both tasks have to be executed roughly with the sampling frequency of the available measurements. Hence, any of the model order reduction and simplification techniques or combinations thereof should be employed to come up with a model of manageable computational complexity under real-time conditions.

Since there exists a huge number of possibilities for reducing a given detailed model of a realistic industrial process, constructive guidelines for designing such models for both, the reconciliation and the control task, would be extremely helpful. Such guidelines do not seem to be available yet. However, hybrid regression or trend models as well as local linearization along a trajectory or in different areas of the operating envelope seem to be attractive candidates provided the prediction horizon is chosen to reflect the prediction quality of the model. The benefit of model order reduction cannot yet be assessed due to a lack of practical experience with challenging plant-wide optimization based control problems. For large-scale systems, a structured approach to order reduction exploiting the natural spatial decomposition of a plant together with an appropriate model simplification procedure as outlined above seems to be crucial for a successful application. Such a structured approach could also be combined with the horizontal decomposition approach introduced before.

Vertical Decomposition Approach

In contrast to the direct approach, two different models of the same process are required to implement both, the dynamic optimizer (DO) and the model predictive controller (MPC) together with their respective estimators on both levels (cf. Figure 3). Ideally, these models should be derived from a detailed master model to guarantee consistency. The requirements on the models are different in both cases:

(a) *Computational constraints:* DO is executed with a much lower frequency (say in the order of once every one or two hours) whereas MPC is executed with a higher frequency (say in the order of once every couple of minutes). Hence, a higher computational complexity can be tolerated for DO as compared to MPC. Obviously, higher frequencies would facilitate better performance in case a sufficiently valid model would be available. Therefore, the complexity of the model should be minimal in both cases, though larger for DO than for MPC, provided the requirements on prediction accuracy are still satisfied.

(b) *Prediction accuracy:* DO should be able to predict economical performance as well as state and output trajectories with sufficient accuracy over the full operating region. Hence, a fairly detailed model incorporating the major process nonlinearities is required. In contrast, the model implemented in the MPC must predict the setpoint deviation, the outputs and possibly the potentially constrained states in the vicinity of the reference trajectory only. Therefore, a simpler model with a much smaller region of validity can be chosen in this case. Even a linear model, updated along the trajectory, may qualify in this case.

(c) *Frequency range:* Due to the different execution frequencies and the different tasks of both levels, the models have to cover the low and high frequency behaviors of the plant for DO and MPC, respectively. Hence, a model with an appropriate prediction quality on a fast time-scale is required for MPC whereas a model a slow time-scale is needed for DO.

Model order reduction by projection, by equation residual minimization or by proper orthogonal decomposition in conjunction with model simplification can be employed for implementation of DO as well as of MPC. Different degrees of reduction should be employed for DO and MPC, however, to account for the specific requirements on prediction errors and computational complexity. A mildly reduced model can be used for DO, whereas a strongly reduced model must be used for MPC to meet the computational complexity constraints. While requirements (a) and (b) could be met by this choice, requirement (c) is definitely in conflict. Though not explicitly incorporated in the model reduction techniques, a mildly reduced model will cover faster time-scales while a strongly reduced model will cover slow time-scales only. This conclusion is based on an interpretation of the strategy employed during model reduction. For example, the integral average of the projection error in Pallaske's method is reduced by a quantifiable amount, if the dimension of the reduced model is increased (Pallaske, 1987; Löffler and Marquardt, 1991). The larger the dimension of the reduced model, the shorter are the time-scales incorporated.

Provided the model used for DO is updated regularly and thus provides updated reference trajectories to the MPC which reflect process economics and comply with constraints the requirements on the model used in the MPC are quite relaxed. In particular, a linearization of some kind (along the reference trajectory for example) together with linear model reduction could be fully sufficient to achieve adequate overall performance. A relatively simple approach for the implementation of DO are the Wiener/Hammerstein hybrid models which build on an available steady-state fundamental (optimization) model. A key issue is in all cases the integration of the model update during reconciliation on the DO and MPC levels.

Some time-scale separation can be achieved by a proper choice of the models used on the DO and MPC levels. However, there is no theoretical basis for keeping

the models on both levels consistent to each other. This problem is well-known even in state of the art (steady-state) real-time optimization and control where serious performance deterioration has been observed in some cases. If time-scale separation is envisioned, singular perturbation methods might be more favorable. They explicitly address this issue by construction to yield two (or even multiple) dynamic models valid on certain time-scales. Though it may be doubted whether exact nonlinear techniques (Marino and Kokotovic, 1988; Krishnan and McClamroch, 1994; Kumar et al., 1998) are widely applicable to construct a singularly perturbed system in standard form, the approximate projection methods of Pallaske (1987) or Lall et al. (1999) employing linear transformations should be applicable to separate even large-scale model into a fast (non-dominant) and a slow (dominant) submodel. If the fast and slow subsystems are envisioned to be used for implementation of DO and MPC on a slow and a fast time-scale, the coupling between both subsystems may lead to serious interaction which could deteriorate control system performance or even stability.

There are cases (e.g. Stiharu-Alexe and O'Shea, 1995; Kumar and Daoutidis, 2000) where not only the state but also the control and output variables are partitioned in the slow and fast subsystems by singular perturbation. In those cases, a completely partitioned DO and MPC level can be implemented (Stiharu-Alexe and O'Shea, 1995), where the control variables of the fast and the slow subsystems are manipulated by MPC and DO completely independently employing the fast and slow output with high and low sampling rates respectively. A consistent time-scale separation can be achieved in this case.

It is still a largely open question, when and how singular perturbation techniques can be applied to partition a model into a fast and a slow submodel to be used in a consistent manner on the DO and MPC levels in time-scale decomposition.

Alternatively, one may employ multi-resolution methods (Binder et al., 1998) to develop models on different time-scales which could be used in DO and MPC, respectively. The basic idea is briefly discussed next using a scalar system (8) for the sake of a simpler notation. The continuous model is projected onto a sparse multi-scale subspace by a Wavelet-Galerkin method. The state and control vectors x and u are expanded in a series according to

$$x = \boldsymbol{d}^T\ \boldsymbol{\psi}(t) = \boldsymbol{d}_s^T \boldsymbol{\psi}_s(t) + \boldsymbol{d}_f^T \boldsymbol{\psi}_f(t)\ , \tag{67}$$

$$u = \boldsymbol{e}^T\ \boldsymbol{\psi}(t) = \boldsymbol{e}_s^T \boldsymbol{\psi}_s(t) + \boldsymbol{e}_f^T \boldsymbol{\psi}_f(t)\ , \tag{68}$$

where $\boldsymbol{\psi}$ denotes the vector of multi-scale basis functions and the vectors $\boldsymbol{d}$ and $\boldsymbol{e}$ contain the expansion coefficients for the state and control variable, respectively. The expansion can be divided into a leading sum referring to the low frequency content (coefficients $\boldsymbol{d}_s, \boldsymbol{e}_s$ and basis functions $\boldsymbol{\psi}_s(t)$) and into a residual sum which covers the high frequency content (coefficients $\boldsymbol{d}_f, \boldsymbol{e}_f$ and basis functions $\boldsymbol{\psi}_f(t)$).

A discretized model of the slow system is obtained as

$$\boldsymbol{\gamma}_s(\boldsymbol{d}_s^s, \boldsymbol{e}_s) = \boldsymbol{0} \tag{69}$$

after Galerkin projection. The vector function $\boldsymbol{\gamma}_s$ results from the projection of the scalar model (8) with the basis functions $\boldsymbol{\psi}_s(t)$. It fixes the expansion coefficients of the states $\boldsymbol{d}_s^s$ as a function of those of the control variable $\boldsymbol{e}_s$ to approximate the low frequency content of the state $x_s = (\boldsymbol{d}_s^s)^T\ \boldsymbol{\psi}_s(t)$. A discretized model of the fast system is obtained from a Galerkin projection with the basis functions $\boldsymbol{\psi}_s(t)$ and $\boldsymbol{\psi}_f(t)$ as

$$\boldsymbol{\gamma}_f(\boldsymbol{d}_s^f, \boldsymbol{d}_f^f, \boldsymbol{e}_s, \boldsymbol{e}_f) = \boldsymbol{0} \tag{70}$$

which fixes the expansion coefficients $\boldsymbol{d}_s^f, \boldsymbol{d}_f^f$ of the state variable as a function of those of the control variable, $\boldsymbol{e}_s, \boldsymbol{e}_f$, to approximate the low and high frequency content of the state

$$\tilde{x} = \tilde{x}_s + \tilde{x}_f = (\boldsymbol{d}_s^f)^T \boldsymbol{\psi}_s(t) + (\boldsymbol{d}_f^f)^T \boldsymbol{\psi}_f(t)\ . \tag{71}$$

We would get a decoupling of the slow and the fast subsystem, if $\boldsymbol{d}_s^s$ and $\boldsymbol{d}_s^f$, the expansion coefficients for the slow contributions to the state in the slow and the fast models, respectively, would be identical. However, due to the wavelet properties, we only find

$$\boldsymbol{d}_s^f = \boldsymbol{d}_s^s + \boldsymbol{\delta}_f \tag{72}$$

with typically small corrections $\boldsymbol{\delta}_f \neq \boldsymbol{0}$, i.e. $\|\boldsymbol{\delta}_f\| = \varepsilon \|\boldsymbol{d}_s^s\|$ with $\varepsilon < 1$.

This approach to time-scale separation may satisfy requirements (b) and (c) above but it is definitely in conflict with requirement (a) since the size of the fast discretized model is much larger than that of the slow. It is still an open issue, whether and how this problem can be solved. In this case, this approach could be an interesting alternative to implement DO and MPC and its associated estimators on both levels.

Closed-loop Model Validation

Obviously, the validity of the various models has to be assessed in the context of their application in the various modules of the operations support system.

A comparison of the *open loop behavior* is possible by a variety of means. Examples are nonlinearity measures (Helbig et al., 2000b, and references cited therein), to compare the loss of nonlinearity between two candidate models, step responses, or nonlinear describing function analysis (Amrhein et al., 1993) to reveal the frequency content of a nonlinear model. One might argue, that measures of uncertainty could be derived at least in principle by comparison of the reduced and the

original model to be used later during estimation and control design. In fact, Andersson et al. (1999) provide a very interesting result on the comparison and also the simplification of two uncertain models. These authors define a simplification error in terms of the L_2-induced gain. They further show that this error can be computed by convex optimization for *linear uncertain systems* and for a certain (broad) class of *nonlinear uncertain systems* with isolated static nonlinearities. The method can be applied to open-loop as well as closed-loop systems. In the linear case, their result generalizes truncation and singular perturbation. Their result is a good basis for comparing and simplifying large-scale nonlinear process models.

Open-loop tests are not sufficient to test the validity of candidate (reduced) models under closed-loop control conditions. A key question is to relate any simplification of the model to the unavoidable loss of economical performance of the feedback control system and to guarantee closed loop stability despite the simplifications made. Stability loss and performance degradation are well-known phenomena if a controller is designed by means of a reduced order model and applied to the plant in a linear setting if no special design technique is applied (see Zhou et al. (1995), Bendotti and Beck (1999), Wortelboer et al. (1999) for recent examples). All these problems will carry over to the nonlinear case at least in principle. However, there is very little knowledge yet about these issues for integrated dynamic optimization and control as investigated in this work. Obviously, the problem could be addressed from a robust control perspective. If we assume, at least for the moment, the detailed (or nominal) model to perfectly match the plant, any model reduction introduces quantifiable uncertainty. This is in contrast to mainstream research in robust (model predictive) control, where the model error cannot be quantified precisely. The knowledge on the uncertainty introduced by model reduction could be employed to robustly accommodate the mismatch between the reduced models (of the estimator and controller) and the real plant (perfectly matched by the nominal model).

There are some starting points for future research in the recent literature on nonlinear control. For example, Scherpen (1993, 1996) proves the stability of reduced models derived from truncation after nonlinear balancing. It is worth noting, that similar stability results are not available for the more empirical methods suggested by Pallaske (1987), Lohmann (1994), Lall et al. (1999), or Kordt (1999) though they are based on linear transformations only and should therefore be simpler to address. The stability problem of a closed loop system with a controller designed by means of a reduced model obtained from nonlinear H_∞ balancing (Scherpen, 1996) is addressed by Pavel and Fairman (1997). They generalize results on closed loop H_∞ balanced truncation by Mustafa and Glover (1991) from the linear to the nonlinear case. The authors provide first a criterion for maintaining the closed loop stability if the controller is designed by solving the nonlinear normalized $H_\infty(L_2)$ control problem for the reduced model and applied to the plant. Further, the degradation of performance is analyzed in closed loop.

For reduced models obtained from some singular perturbation analysis, there are not only strong results available for a number of nonlinear control system design techniques but also for open-loop optimal control. For example, Artstein and Gaitsgory (2000) proved just recently convergence of the value function of the perturbed system to that of the slow system for $\epsilon \to 0$ under mild assumptions (such as controllability of the fast subsystem) for general systems in standard singularly perturbed form. Such results could be a starting point for the analysis under closed-loop conditions in future research.

Conclusions

Optimization-based control of transient processes requires nonlinear models of sufficient predictive quality which can be employed for the various tasks in the feedback control system in a real-time environment. On the basis of a suitable formulation of the control problem and some thoughts on its implementation, we focussed on fundamental modeling and in particular on nonlinear model reduction which comprises both, model order reduction and model simplification. A large variety of methods with differing theoretical justification has been reviewed and put into perspective. Though, there has been significant progress in the last 10 years, a thorough understanding which technique could and should advantageously be used in optimization-based control and how it should be tailored to a specific problem is largely lacking. The situation is even worse, if a (vertical) decomposition of the optimizing control system is envisioned in order to extend the state of the art in (steady-state) real-time optimization where a multi-level architecture is typically implemented. Some of the major open research problems are:

(a) Lumped process systems models are usually of differential-algebraic type. With the exception of the work of Löffler and Marquardt, there are no general model order reduction techniques for this class of systems available.

(b) All of the nonlinear reduction techniques rely on a representative set of trajectories. The selection of this set is crucial for the success of the reduction. To the author's knowledge there are no systematic techniques yet to guide this selection. Obviously, the set should be as close as possible to the trajectories occurring in closed-loop.

(c) In contrast to linear model order reduction, only

truncation reduces the model order significantly. If residualization or some sort of steady-state assumption is introduced, differential-algebraic equations result which are often as difficult to solve as the original problem. Systematic approaches based on slaving as introduced in POD or even nonlinear regression of the set of nonlinear equations with a simple nonlinear map are required.

(d) Most of the model reduction techniques lead to a reduced number or equations which are however of a significantly higher functional complexity. The structure and the sparsity of the original model is lost. In particular, in dynamic optimization, exploitation of model structure is a key to a high performant numerical solution. Any means of preserving structure and sparsity at least to some extent during model order reduction is highly beneficial. In addition, sparsity can be reintroduced by systematic model simplification for example by eliminating most of the nonlinear terms in the right hand sides of the reduced model.

(e) There has been almost no work on the system theoretical properties of resulting nonlinear reduced models. Most of the reduction techniques cannot be expected to preserve stability, observability, or controllability properties. In principle, equation residual minimization (Lohmann, 1994, 1995) could be extended by additional constraints not only preserving steady-state accuracy but also stability (using the approach of Mönnigmann and Marquardt (2000), for formulating appropriate constraints).

(f) All process models are hybrid by nature since they comprise fundamental and empirical parts. The appropriate combination of fundamental and empirical knowledge is still an open issue even in case of open-loop (simulation) applications. In closed-loop, an appropriate parameterization of the uncertainty in both model constituents as a basis for an efficient reconciliation is an even more challenging problem.

(g) Excitation frequencies or the magnitude of the controls and disturbances driving the process in closed-loop are largely unknown in most cases. Ideally, on-line adaptation of the structure and not only the parameters of the reduced model on the basis of actual or historical process data would be most preferable. Computational singular perturbation (Massias et al., 1999) or adaptive Galerkin methods (von Watzdorf and Marquardt, 1997; Briesen and Marquardt, 2000) as developed for the treatment of multi-component separation and reaction processes could be a first starting point for the development of a more general technique.

(h) The validity of the model under closed-loop conditions is critical for the success of integrated dynamic optimization and control. Since there are most likely many models interacting in the various functional blocks of the control and optimization system systematic means of constructing consistent models are required to reach high performance. It is a largely open issue how to efficiently assess the validity of the individual models with respect to the prediction of states and outputs as well as gradient and sensitivity information, the consistency between these models, as well as the stability and performance of the integrated system a priori (i.e. during the early phases of the design phase of the control system).

Hopefully this review and this list of major research challenges rises interest in the systems and control community to work in this fruitful and rewarding area. Obviously, most of the questions are not hard core process control problems, but they are rather at the interfaces to related fields which renders them more interesting and more challenging at the same time.

Acknowledgments

This research is partly supported by the European Commission under grant G1RD–CT–1999–00146 in the INCOOP project. INCOOP strives for the integration of control and optimization of transiently operated chemical processes. Fruitful discussions with all the INCOOP team members, in particular with Ton Backx, Okko Bosgra, Jogchem v.d. Berg and Rob Tousain, are gratefully acknowledged. The support of Elke Krumrück, Jitendra Kadam, Adel Mhamdi, Martin Schlegel, and Thomas Binder during the preparation of the manuscript is also acknowledged.

References

Abel, O., A. Helbig, and W. Marquardt, Optimization approaches to control-integrated design of industrial batch reactors, In Berber, R. and C. Kravaris, editors, *Nonlinear Model Based Process Control*, NATO-ASI Series, pages 513–551. Kluwer, Dordrecht (1998).

Afanasiev, K. and M. Hinze, Adaptive control of a wake flow using proper orthogonal decomposition, Technical Report 648, Fachbereich Mathematik, TU Berlin (1999).

Agarwal, M., "Combining neural and conventional paradigms for modelling, prediction and control," *Int. J. Sys. Sci.*, **28**(1), 65–81 (1997).

Aling, H., R. L. Kosut, A. Emami-Naeini, and J. L. Ebert, Nonlinear model reduction with application to rapid thermal processing, In *Proceedings of the 35th Conf. of Decision and Control*, pages 4305–4310, Kobe (1996).

Aling, H., S. Banerjee, A. K. Bangia, V. Cole, J. Ebert, A. Emami-Naeini, K. F. Jensen, I. G. Kevrekidis, and S. Shvartsman, Nonlinear Model Reduction for Simulation and Control of Rapid Thermal Processing, In *Proceedings of ACC 1997*, volume 4, pages 2233–2238, Albuquerque (1997).

Allgöwer, F., T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright, Nonlinear Predictive Control and Moving Horizon Estimation, In Frank, P. M., editor, *Advances in Control*, pages 391–449. Springer Verlag (1999).

Amrhein, M., F. Allgöwer, and W. Marquardt, Validation and analysis of linear distillation models for controller design, In *Proceedings of the Proc. Europ. Control Conf.*, pages 655–660 (1993).

Andersson, L., A. Rantzner, and C. Beck, "Model comparison and simplification," *Int. J. Robust and Nonlinear Control*, **9**, 157–181 (1999).

Androulakis, I. P., "Kinetic mechanism reduction based on an integer programming approach," *AIChE J.*, **46**(2), 361–371 (2000).

Aris, R., "Manners makyth modellers," *Trans. Inst. Chem. Eng.*, **69**(Part A), 165–174 (1991).

Armaou, A. and P. D. Christofides, "Wave suppression by nonlinear finite-dimensional control," *Chem. Eng. Sci.*, **55**, 2627–2640 (2000).

Artstein, Z. and V. Gaitsgory, "The value function of singularly perturbed control systems," *Applied Mathematics and Optimization*, **41**, 425–445 (2000).

Aubry, N., P. Holmes, J. L. Lumley, and E. Stone, "The dynamics of coherent structures in the wall region of a turbulent boundary layer," *J. Fluid Mech.*, **192**, 115 (1988).

Backx, T., O. Bosgra, and W. Marquardt, Towards Intentional Dynamics in Supply Chain Conscious Process Operations, FOCAPO 98, Snowmass, www.lfpt.rwth-aachen.de/Publication/Techreport/ (1998). Revised version submitted to *J. Proc. Control*, July 2000.

Backx, T., O. Bosgra, and W. Marquardt, Integration of model predictive control and optimization of processes, In *Proceedings ADCHEM 2000*, volume 1, pages 249–260 (2000).

Baker, J. and P. D. Christofides, "Output feedback control of parabolic PDE Systems with nonlinear spatial differential operators," *Ind. Eng. Chem. Res.*, **38**, 4372–4380 (1999).

Baker, J. and P. D. Christofides, "Finite-dimensional approximation and control of non-linear parabolic PDE systems," *Int. J. Control*, **73**(5), 439–456 (2000).

Balasubramhanya, L. S. and F. J. Doyle III, "Nonlinear model-based control of a batch reactive distillation column," *J. Proc. Cont.*, **10**, 209–218 (2000).

Banerjee, A. and Y. Arkun, "Model predictive control of plant transitions using a new identification technique for interpolating nonlinear models," *J. Proc. Cont.*, **8**(5/6), 441–457 (1998).

Banerjee, A., Y. Arkun, B. Ogunnaike, and R. Pearson, "Estimation of nonlinear systems using linear multiple models," *AIChE J.*, **43**(5), 1204–1226 (1997).

Barton, P. I. and C. C. Pantelides, "Modeling of combined discrete/continuous processes," *AIChE J.*, **40**(6), 966–979 (1994).

Barton, P. I., R. J. Allgor, W. F. Feehery, and S. Galán, "Dynamic optimization in a discontinuous world," *Ind. Eng. Chem. Res.*, **37**(3), 966–981 (1998).

Bemporad, A. and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, **35**(3), 407–427 (1999).

Bendotti, P. and C. L. Beck, "On the role of LFT model reduction methods in robust controller synthesis for a pressurized water reactor," *IEEE Trans. Cont. Sys. Tech.*, **7**, 248–257 (1999).

Biegler, L. T. and G. B. Sentoni, "Efficient Formulation and Solution of Nonlinear Model Predictive Control Problems," *Latin American Applied Research*, **30**(4), 315–324 (2000).

Biegler, L. T., I. E. Grossmann, and A. W. Westerberg, "A note on approximation techniques used for process optimization," *Comput. Chem. Eng.*, **9**(2), 201–206 (1985).

Binder, T., L. Blank, W. Dahmen, and W. Marquardt, Towards multiscale dynamic data reconciliation, In Berber, R. and C. Kravaris, editors, *Nonlinear Model Based Process Control*, NATO-ASI Series, pages 623–665, Dordrecht. Kluwer (1998).

Bogusch, R. and W. Marquardt, "A Formal Representation of Process Model Equations," *Comput. Chem. Eng.*, **21**(10), 1105–1115 (1997).

Bogusch, R., B. Lohmann, and W. Marquardt, "Computer-aided process modeling with MODKIT," *Comput. Chem. Eng.* (2001). In press.

Bonvin, D. and D. A. Mellichamp, "A unified derivation and critical review of modal approaches to model reduction," *Int. J. Control*, **35**(5), 829–848 (1982).

Brenan, K. E., S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. Society for Industrial and Applied Mathematics (1996).

Briesen, H. and W. Marquardt, "Adaptive model reduction and simulation of the thermal cracking of multicomponent hydrocarbon mixtures," *Comput. Chem. Eng.*, **24**, 1287–1292 (2000).

Brown, N. J., G. Li, and M. L. Koszykowski, "Mechanism reduction via principal component analysis," *Int. J. Chem. Kinet.*, **29**(6), 393–414 (1997).

Bryson, A. E. and Y.-C. Ho, *Applied Optimal Control*. Taylor and Francis, Bristol (1975).

Chikkula, Y., J. H. Lee, and B. A. Ogunnaike, "Dynamically scheduled MPC of nonlinear processes using hinging hyperplane models," *AIChE J.*, **44**(12), 2658–2674 (1998).

Chimowitz, E. H. and C. S. Lee, "Local thermodynamic models for high pressure process calculations," *Comput. Chem. Eng.*, **9**(2), 195–200 (1985).

Chimowitz, E. H., T. F. Anderson, S. Macchietto, and L. F. Stutzman, "Local models for representing phase equilibria in multicomponent, nonideal vapor-liquid and liquid-liquid systems. 1. Thermodynamic approximation functions," *Ind. Eng. Chem. Proc. Des. Dev.*, **22**, 217–225 (1983).

Chimowitz, E. H., S. Macchietto, T. F. Anderson, and L. F. Stutzman, "Local models for representing phase equilibria in multicomponent, nonideal vapor-liquid and liquid-liquid systems. 2. Application to process design," *Ind. Eng. Chem. Proc. Des. Dev.*, **23**, 609–618 (1984).

Cho, Y. S. and B. Joseph, "Reduced-order steady-state and dynamic models for separation processes. I. Development of the model reduction procedure. II. Application to nonlinear multicomponent systems.," *AIChE J.*, **29**, 261–276 (1983).

Christofides, P. D. and P. Daoutidis, "Feedback control of two-time-scale nonlinear systems," *Int. J. Control*, **63**, 965–994 (1996).

Christofides, P. D. and P. Daoutidis, "Finite-dimensional control of parabolic PDE systems using approximate intertial manifolds," *J. Math. Anal. Appl.*, **216**(2), 398–420 (1997).

Čiegis, R., R. Šablinskas, and J. Waśniewski, Numerical integration on distributed-memory parallel systems, In *Recent Advances in Parallel Virtual Machine and Message Passing Interface. Lecture Notes in Computer Science*, pages 329–336. Springer (1997).

Costa, A. C., A. S. W. Henriques, T. L. M. Alves, R. Maciel Filho, and E. L. Lima, "A hybrid neural model for the optimization of fed-batch fermentation," *Braz. J. Chem. Eng.*, **16**(1) (1999).

Davis, M. J. and R. T. Skodje, "Geometric investigation of low-dimensional in systems approaching equilibrium," *J. Chem. Phys.*, **111**(3), 859–874 (1999).

de Veaux, R. D., R. Bain, and L. H. Ungar, "Hybrid neural network models for environmental process control," *Environmetrics*, **10**, 225–236 (1999).

Desrochers, A. A. and R. Y. Al-Jaar, "A method for high order linear system reduction and nonlinear system simplification," *Automatica*, **21**(1), 93–100 (1985).

Dharaskar, K. P. and Y. P. Gupta, "Predictive control of nonlinear processes using interpolated models," *Chem. Eng. Res. Des.*, **78**(A4), 573–580 (2000).

Dochain, D. and B. Bouaziz, "Approximation of the model of fixed bed reactors via a singular perturbation approach," *Math. Comp. Simul.*, **37**, 165–172 (1994).

Doyle III, F. J. and P. A. Wisnewski, Nonlinear multi-rate MPC with large scale fundamental models: application to a continuous Kamyr Digester, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 419–432. Birkhäuser, Basel (2000).

Doyle III, F. J., H. M. Budman, and M. Morari, ""Linearizing" controller design for a packed-bed reactor using a low-order wave propagation model," *Ind. Eng. Chem. Res.*, **35**, 3567–3580 (1996).

Duchene, P. and P. Rouchon, "Kinetic scheme reduction via geometric singular perturbartion techniques," *Chem. Eng. Sci.*, **51**(20), 4661–4672 (1996).

Edwards, K. and T. F. Edgar, "Reaction set simplification using variable selection techniques," *Chem. Eng. Sci.*, **55**, 551–572 (2000).

Edwards, K., T. F. Edgar, and V. I. Manousiouthakis, "Kinetic model reduction using genetic algorithms," *Comput. Chem. Eng.*, **22**(1-2), 239–246 (1998).

Edwards, K., T. F. Edgar, and V. I. Manousiouthakis, "Reaction mechanism simplification using mixed-integer nonlinear programming," *Comput. Chem. Eng.*, **24**(1), 67–79 (2000).

Eitelberg, E., "Model reduction and perturbation structures," *Int. J. Control*, **35**, 1029–1050 (1982).

Epple, U., Model reduction for nonlinear systems with distributed parameters, In *Proc. IFAC Symposium DYCOPS'86 Dynamics and Control of Chemical Reactors and Distillation Columns, Bournemouth, UK*, pages 279–284 (1986).

Findeisen, W., M. Brdys, K. Malinowski, P. Tatjewski, and A. Wozniak, *Control and Coordination in Hierarchical Systems*. Wiley (1980).

Foias, C. and R. Témam, "Algebraic approximation of attractors: The finite dimensional case," *Physica D*, **32**, 163 (1988).

Foias, C., G. R. Sell, and R. Témam, "Inertial manifolds for nonlinear evolutionary equations," *J. Differential Equations*, **73**, 309 (1988).

Foss, B. A., B. Lohmann, and W. Marquardt, "A field study of the industrial modeling process," *J. Proc. Cont.*, **8**, 325–337 (1998).

Foss, B. A., T. A. Johansen, and A. V. Sørensen, "Nonlinear predictive control using local models: Applied to a batch fermentation process," *Control Eng. Practice*, **3**(3), 389–396 (2000).

Fukunaga, K., *Introduction to Statistical Pattern Recognition*. Academic Press (1990).

Ganesh, N. and L. T. Biegler, "A robust technique for process flowsheet optimization using simplified model approximations," *Comput. Chem. Eng.*, **11**(6), 553–565 (1987).

García, C. E., "Quadratic dynamic matrix control of non-linear processes. An application to a batch reactor process," *AIChe Annual Meeting, San Francisco* (1984).

Gattu, G. and E. Zafiriou, "Nonlinear Quadratic Dynamic Matrix Control with State Estimation," *Ind. Eng. Chem. Res.*, **31**(4), 1096–1104 (1992).

Genyuan, L. and H. Rabitz, "Combined symbolic and numerical approach to constrained nonlinear lumping: with application to an H2/O2 oxidation model," *Chem. Eng. Sci.*, **51**(21), 4801–4816 (1996).

Gilles, E. D. and U. Epple, Nonlinear wave propagation and model reduction of the fixed-bed reactor, In *Modelling of Chemical Reaction Systems*, pages 312–336. Springer-Verlag, Berlin (1981).

Gilles, E. D., B. Retzbach, and F. Silberberger, Mopdeling, simulation and control of an extractive distillation column, In *ACS Symposium Series*, volume 124, pages 481–492 (1980).

Glover, K., "All optimum Hankel-norm approximations of linear multivariable systems and their L_2-error bounds," *Int. J. Control*, **39**(6), 1115–1193 (1984).

Graham, M. D. and I. G. Kevrekidis, "Alternative approaches to the Karhunen-Loeve decomposition for model reduction and data analysis," *Comput. Chem. Eng.*, **20**(5), 495–506 (1996).

Grens, E. A., "Efficient use of thermodynamic models in process calculations," *Proc. Int. Conf. Found. Comp. Aid. Proc. Des., Snowmass*, pages 249–302 (1983).

Gupta, S., P.-H. Liu, S. A. Svoronos, R. Sharma, and N. A. Abdel-Khalek, "Hybrid first-principles/neural networks model for column flotation," *AIChE J.*, **45**(3), 557–566 (1999).

Hager, J. M., *Lokale Modelle zur Berechnung von Gas-Flüssigkeits-Phasengleichgewichten*, PhD thesis, Universität Stuttgart (1992).

Hahn, J. and T. F. Edgar, "An improved method for nonlinear model reduction using balancing of empirical gramians," *Comp. Chem. Eng.* (1999). submitted.

Hahn, J. and T. F. Edgar, Reduction of nonlinear models using balancing of balancing of empirical gramians Galerkin projections, In *Proceedings Amer. Control. Conf. 2000*, pages 2864–2868, Chicago (2000).

Han, H. and S. Park, "Control of high-purity distillation columns using nonlinear wave theory," *AIChE J.*, **39**, 787–796 (1993).

Hasenjäger, E., *Digitale Zustandsregelung von Parabolantennen unter Berücksichtigung von Nichtlinearitäten*. VDI Verlag (1991).

Helbig, A., O. Abel, and W. Marquardt, Model predictive control for on-line optimization of semi-batch reactors, In *Proc. Amer. Control Conf.*, volume 9 (1998).

Helbig, A., W. Marquardt, and F. Allgöwer, "Nonlinearity measurers: definition, computation and applications," *J. Proc. Cont.*, **10**, 113–123 (2000b).

Helbig, A., O. Abel, and W. Marquardt, Structural Concepts for Optimization Based Control of Transient Processes, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 295–311. Birkhäuser, Basel (2000a).

Henson, M. A., "Nonlinear model predictive control: Current status and future directions," *Comput. Chem. Eng.*, **23**, 187–202 (1998).

Hillestad, M., C. Sorlie, T. F. Anderson, I. Olsen, and T. Hertzberg, "On estimating the error of local thermodynamic models—a general approach," *Comput. Chem. Eng.*, **13**(7), 789–796 (1989).

Holmes, P., J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge Univ. Press (1996).

Hwang, Y.-L., "Nonlinear wave theory for dynamics of binary distillation columns," *AIChE J.*, **37**, 705–723 (1991).

Isidori, A., *Nonlinear Control Systems*. Springer (1989).

Jadach, S., "Multi-dimensional general purpose Monte Carlo generator with self-adapting simplical grid," *Computer Physics Communications*, **130**(3), 244–259 (2000).

Jarke, M. and W. Marquardt, "Design and evaluation of computer-aided process modeling tool," *AIChE Symp. Ser.*, **92**(312), 97–109 (1996).

Johansen, T. A. and B. A. Foss, "Representing and learning unmodeled dynamics with neural network memories," *Proc. Amer. Control Conf.*, **3**, 3037 (1992).

Johansen, T. A. and B. A. Foss, "Operating regime based process modeling and identification," *Comput. Chem. Eng.*, **21**(2), 159–176 (1997).

Jose, R. A. and L. H. Ungar, "Pricing interprocess streams using slack auctions," *AIChE J.*, **46**(3), 575–587 (2000).

Kadam, J. V., Index analysis and adaptive refinement in multiscale dynamic optimization, Master's thesis, Dept. of Chem. Engng., Indian Institute of Technology Bombay and Lehrstuhl für Prozesstechnik, RWTH Aachen (2000).

Kan, P. and C. J. Lee, "A neural network model for prediction of phase equilibria in aqueous two-phase extraction," *Ind. Eng. Chem. Res.*, **35**(6), 2015–2023 (1996).

Kienle, A., "Low-order dynamic models for ideal multicomponent distillation processes using nonlinear wave theory," *Chem. Eng. Sci.*, **55**(10), 1817–1828 (2000).

Kokotovic, P. V., H. K. Khalil, and J. O'Reilly, *Singular Perturbations in Control: Analysis and Design*. Academic Press (1986).

Kordt, M., "Nichtlineare Ordnungsreduktion für ein Transportflugzeug," *Automatisierungstechnik*, **47**(11), 532–539 (1999).

Kramer, M. A. and M. L. Thompson, "Embedding theoretical models in neural networks," *Proc. Amer. Control Conf.*, **1**, 475 (1992).

Krishnan, H. and N. H. McClamroch, "On the connection between nonlinear differential- algebraic equations and singularly perturbed control systems in nonstandard form," *IEEE Trans. Auto. Cont.*, **39**(5), 1079–1084 (1994).

Kumar, A. and P. Daoutidis, "Nonlinear dynamics and control of process systems with recycle," *ADCHEM Preprints*, **1**, 13–22 (2000).

Kumar, A., P. D. Christofides, and P. Daoutidis, "Singular perturbation modeling of nonlinear processes with nonexplicit time-scale multiplicity," *Chem. Eng. Sci.*, **53**(8), 1491–1504 (1998).

Kunisch, K. and S. Volkwein, "Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition," *J. Optim. Theo. Appl.*, **102**(2), 345–371 (1999).

Kurtz, M. J. and M. A. Henson, "Input-output linearizing control of constrained nonlinear processes," *J. Proc. Cont.*, **7**(1), 3–17 (1997).

Kurtz, M. J. and M. A. Henson, "Feedback linearizing control of discrete-time nonlinear systems with input constraints," *Int. J. Control*, **70**(4), 603–616 (1998).

Lakshmanan, N. M. and Y. Arkun, "Estimation and model predictive control of non-linear batch processes using linear parameter varying models," *Int. J. Control*, **72**(7/8), 659–675 (1999).

Lall, S., J. E. Marsden, and S. Glavaski, "Empirical model reduction of controlled nonlinear systems," *14th IFAC World Congress, Beijing* (1999).

Ledent, T. and G. Heyen, "Dynamic approximation of thermodynamic properties by means of local models," *Comput. Chem. Eng.*, **18**(Suppl.), S87–S91 (1994).

Lee, J. H. and N. L. Ricker, "Extended Kalman filter based nonlinear model predictive control," *Ind. Eng. Chem. Res.*, **33**(6), 1530–1541 (1994).

Lee, P. L., L. Huaizhong, and I. T. Cameron, "Decentralized control design for nonlinear multi-plants: A gap metric approach," *Chem. Eng. Sci.*, **55**, 3743–3758 (2000).

Lee, J. H., Modeling and identification for nonlinear model predictive control: Requirements, current status and future research needs, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 269–293. Birkhäuser, Basel (2000).

Leesley, M. E. and G. Heyen, "The dynamic approximation method of handling vapor-liquid equilibrium data in computer calculations for chemical processes," *Comput. Chem. Eng.*, **1**, 109–112 (1977).

Levine, J. and P. Rouchon, "Quality control of binary distillation columns via nonlinear aggregated models," *Automatica*, **27**, 463–480 (1991).

Litz, L., "Ordnungsreduktion linearer Zustandsraummodelle durch Beibehaltung der dominanten Eigenbewegungen," *Regelungstechnik*, **28**(3), 80–86 (1979).

Liu, Y. and B. Anderson, "Singular perturbation approximation of balanced systems," *Int. J. Control*, **50**, 1379–1405 (1989).

Löffler, H. P. and W. Marquardt, "Order reduction for nonlinear differential-algebraic process models," *J. Proc. Cont.*, **1**, 32–40 (1991).

Lohmann, B. and W. Marquardt, "On the Systematization of the process of model development," *Comput. Chem. Eng.*, **21**(Suppl.), S213–S218 (1996).

Lohmann, B., "Ordnungsreduktion und Dominanzanalyse nichtlinearer Systeme," *Automatisierungstechnik*, **42**(10), 466–474 (1994).

Lohmann, B., "Application of model order reduction to a hydropneumatic vehicle suspension," *IEEE Trans. Cont. Sys. Tech.*, **3**(1), 102–109 (1995).

Lohmann, B., *Ansätze zur Unterstützung des Arbeitsablaufes bei der rechnerbasierten Modellierung verfahrenstechnischer Prozesse*. VDI Verlag (1998).

Lu, J. Z., Multi-zone control under enterprise optimization: Needs, challenges and requirements, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 393–402. Birkhäuser, Basel (2000).

Maas, U. and S. Pope, "Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space," *Comb. & Flame*, **88**, 252–264 (1992).

Macchietto, S., E. H. Chimowitz, T. F. Anderson, and L. F. Stutzman, "Local models for representing phase equilibria in multicomponent nonideal vapor-liquid and liquid-liquid systems. 3. Parameter estimation and update," *Ind. Eng. Chem. Proc. Des. Dev.*, **25**, 674–682 (1986).

Marenbach, P., K. D. Bettenhausen, S. Freyer, U. Nieken, and H. Rettenmaier, Data-driven structured modelling of a biotechnological fed-batch fermentation by means of genetic programming, In *Proceedings of the Institution of Mechanical Engineers. Part I—J. Sys. Control Engng.*, volume 211, pages 325–332 (1997).

Marino, R. and P. V. Kokotovic, "A geometric approach to nonlinear singularly perturbed control systems," *Automatica*, **24**(1), 31–41 (1988).

Marlin, T. E. and A. N. Hrymak, "Real-time operations optimization of continuous processes," *AIChE Symp. Ser.*, **93**(316), 156–164 (1997).

Marquardt, W. and E. D. Gilles, Nonlinear wave phenoma as fundamentals for model-based control system design in distillation, AIChE Annual Meeting, Chicago, IL (1990).

Marquardt, W., "Traveling waves in chemical processes," *Int. Chem. Engng.*, **30**, 585–606 (1990).

Marquardt, W., Numerical methods for the simulation of differential-algebraic process models, In Berber, R., editor, *Methods of model based control*, volume 293 of *NATO-ASI Ser. E, Applied Sciences*, pages 42–79. Kluwer, Dordrecht (1995).

Martinson, W. S. and A. I. Barton, "A differentiation index for partial differential-algebraic equations.," *SIAM J. Sci. Comp.*, **21**(6), 2295–2315 (2000).

Massias, A., D. Diamantis, E. Mastrorakos, and D. A. Goussis, "An algorithm for the construction of global reduced mechanisms with CSP data," *Combustion and Flame*, **117**, 685–708 (1999).

Mayne, D. Q., J. B. Rawlings, C. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, **36**, 789–814 (2000).

McKay, B., M. Willis, and G. Barton, "Steady-state modelling of chemical process systems using genetic programming," *Comput. Chem. Eng.*, **21**(9), 981–996 (1997).

Mesarovic, M. D., D. Macko, and Y. Takahara, "Two coordination principles and their application in large scale systems control," *Automatica*, **6**, 261–270 (1970).

Molga, E. and R. Cherbański, "Hybrid first-principle—neural-network approach to modelling of the liquid-liquid reacting system," *Chem. Eng. Sci.*, **54**, 2467–2473 (1999).

Molga, E. and K. Westerterp, "Neural network based model of the kinetics of catalytic hydrogenation reactions," *Stud. Surf. Sci. Catal.*, **109**, 379–388 (1997).

Mönnigmann, M. and W. Marquardt, "Normal vectors on manifolds of critical points for parametric robustness of equilibrium solutions of ODE systems," *J. Nonlinear Sci.* (2000). submitted.

Moore, B. C., "Principal component analysis in linear systems: Controllabity, observability, and model reduction," *IEEE Trans. Auto. Cont.*, **26**(1), 17–32 (1981).

Morari, M. and J. H. Lee, "Model predictive control: past, present and future," *Comput. Chem. Eng.*, **23**, 667–682 (1999).

Morari, M., Y. Arkun, and G. Stephanopoulos, "Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the control problem. Process decomposition and the classification of the control tasks. Analysis of the optimizing control structures," *AIChE J.*, **26**(2), 220–232 (1980).

Muscato, G., "Parametric generalized singular perturbation approximation for model order reduction," *IEEE Trans. Auto. Cont.*, **45**(2), 339–343 (2000).

Mustafa, D. and K. Glover, "Controller reduction by H_∞-balanced truncation," *IEEE Trans. Auto. Cont.*, **36**, 668–682 (1991).

Nevistić, V. and M. Morari, Constrained control of feedback-linearizable systems, In *Proceedings of Europ. Control Conf.*, pages 1726–1731, Rome (1995).

Newman, A. and P. S. Krishnaprasad, "Nonlinear model reduction for RTCVD," *IEEE Proc. 32nd Conf. Information Sciences and Systems, Princeton* (1998).

Norquay, S. J., A. Palazoglu, and J. A. Romagnoli, "Application of Wiener model predictive control (WMPC) to an industrial C2-splitter," *J. Proc. Cont.*, **9**, 461–473 (1999).

Pallaske, U., "Ein Verfahren zur Ordnungsreduktion mathematischer Prozeßmodelle," *Chem.-Ing.-Tech.*, **59**(7), 604–605 (1987).

Pavel, L. and F. W. Fairman, "Controller reduction for nonlinear plants: An L_2 approach," *Int. J. Rob. Nonl. Contr.*, **7**, 475–505 (1997).

Pearson, R. K. and B. A. Ogunnaike, Nonlinear process identification, In *Nonlinear process control*, pages 11–110. Prentice-Hall (1997).

Pearson, R. K. and M. Pottmann, "Gray-box identification of block-oriented nonlinear models," *J. Proc. Cont.*, **10**, 301–314 (2000).

Perregaard, J., "Model simplification and reduction for simulation and optimization of chemical processes," *Comput. Chem. Eng.*, **17**(5/6), 465–483 (1993).

Petzold, L. and W. Zhu, "Model reduction for chemical kinetics: an optimization approach," *AIChE J.*, **45**(4), 869–886 (1999).

Psichogios, D. C. and L. H. Ungar, "A hybrid neural-network first principles approach to process modeling," *AIChE J.*, **38**, 1499 (1992).

Qin, S. J. and T. J. Badgwell, An overview of industrial model predictive control technology, In *Chemical Process Control V*, volume 316,93 of *AIChE Symposium Series*, pages 232–256 (1996).

Qin, S. J. and T. J. Badgwell, An overview of nonlinear model predictive control applications, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*. Birkhäuser, Basel, Switzerland (2000).

Ravindran, S. S., Proper orthogonal decomposition in optimal control of fluids, Technical Report NASA/TM-1999-209113, National Aeronautics and Space Administration NASA, Hampton, Virginia (USA) (1999).

Rawlings, J. B., "Tutorial overview of model predictive control," *IEEE Cont. Sys. Mag.*, **20**(3), 38–52 (2000).

Reuter, M., J. van Deventer, and P. van der Walt, "A generalized neural-net rate equation," *Chem. Eng. Sci.*, **48**, 1499 (1993).

Rhodes, C., M. Morari, and S. Wiggins, "Identification of low order manifolds: validating the algorithm of Maas and Pope," *CHAOS*, **9**(1), 108–123 (1999).

Robertson, G. and I. Cameron, "Analysis of dynamic models for structural insight and model reduction: Part 2. A multi-stage compressor shutdown case study," *Comput. Chem. Eng.*, **21**(5), 475–488 (1997a).

Robertson, G. and I. Cameron, "Analysis of dynamic process models for structural insight and model reduction: Part 1. Structural identification measures," *Comput. Chem. Eng.*, **21**(5), 455–473 (1997b).

Safavi, A. A., A. Nooraii, and J. A. Romangnoli, "A hybrid model formulation for a distillation column and the on-line optimisation study," *J. Proc. Cont.*, **9**, 125–134 (1999).

Samar, R., I. Postlethwaite, and D.-W. Gu, "Model reduction with balanced realizations," *Int. J. Control*, **62**(1), 33–64 (1995).

Saxén, B. and H. Saxén, "A neural-network based model of bioreaction kinetics," *Can. J. Chem. Eng.*, **74**(1), 124–131 (1996).

Scherpen, J. M. A. and A. J. van der Schaft, "Normalized coprime factorizations and balancing for unstable nonlinear systems," *Int. J. Control*, **60**(6), 1193–1222 (1994).

Scherpen, J. M. A., "Balancing for nonlinear systems," *Sys. Cont. Let.*, **21**, 143–153 (1993).

Scherpen, J. M. A., "H_∞-balancing for nonlinear systems," *Int. J. Rob. Nonl. Contr.*, **6**, 645–668 (1996).

Schuler, H. and C.-U. Schmidt, "Calorimetric state estimators for chemical reactor diagnosis and control: review of methods and applications," *Chem. Eng. Sci.*, **47**, 899–915 (1992).

Seigneur, C., G. Stephanopoulos, and R. W. Carr Jr., "Dynamic sensitivity analysis of chemical reaction systems: a variational method," *Chem. Eng. Sci.*, **17**, 845–853 (1982).

Shene, C., C. Diez, and S. Bravo, "Neural networks for the prediction of the state of Zymomonas mobilis CP4 batch fermentations," *Comput. Chem. Eng.*, **23**, 1097–1108 (1999).

Shvartsman, S. Y. and I. G. Kevrekidis, "Nonlinear model reduction for control of distributed systems: A computer-assisted study," *AIChE J.*, **44**(7), 1579–1595 (1998).

Shvartsman, S. Y., C. Theorodoropuolos, R. Rico-Martinez, I. G. Kevrekidis, and E. S. Titi, "Order reduction for nonlinear dynamic models of distributed reacting systems," *J. Proc. Cont.*, **10**, 177–184 (2000).

Simutis, R., R. Oliviera, M. Manikowski, S. Feyo de Azevedo, and A. Lübbert, "How to increase the performance of models for process optimization and control," *J. Biotech.*, **59**, 73–89 (1997).

Singh, M. D., *Dynamical Hierarchical Control*. North-Holland (1977).

Sirovich, L., "Turbulence and the dynamics of coherent structures: I, II and III," *Quart. Appl. Math.*, **XLV**(3), 561 (1987).

Skogestad, S. and I. Postlethwaite, *Multivariable Feedback Control*. John Wiley (1996).

Slaats, P. M. A., J. Jongh de, and A. A. H. J. Sauren, "Model reduction tools for nonlinear structural dynamics," *Computers & Structures*, **54**(6), 1155–1171 (1995).

Stewart, W. E., K. L. Levien, and M. Morari, "Simulation of fractionation by orthogonal collocation," *Chem. Eng. Sci.*, **40**, 409–421 (1985).

Stiharu-Alexe, I. and J. O'Shea, "Four-dimensional guidance of atmospheric vehicles," *J. Guidance, Control, and Dynamics*, **19**(1), 113–122 (1995).

Storen, S. and T. Hertzberg, "Local thermodynamic models used in sensitivity estimation of dynamic systems," *Comput. Chem. Eng.*, **21**(Suppl.), S709–S714 (1997).

Su, H. T., P. A. Bhat, P. A. Minderman, and T. J. McAvoy, "Integrating neural networks with first principles models for dynamic modeling," *IFAC Symp. Dynamics and Control of Chem. Reactors, Distillation Columns, and Batch Processes* (1992).

Tali-Maamar, N., M. T. Nihtila, and J. P. Babary, "On control of nonlinear distributed parameter bioreactors," *Math. Comp. Simul.*, **37**, 173–181 (1994).

Tatrai, F., P. Lant, P. Lee, I. Cameron, and R. Newell, "Model reduction for regulatory control: an FCCU case study," *Trans. Inst. Chem. Eng.*, **72**(5), 402–407 (1994).

te Braake, H. A. B., H. J. L. van Can, and H. B. Verbruggen, "Semi-mechanistic modeling of chemical processes with neural networks," *Artificial Intelligence*, **11**, 507–515 (1998).

Thibault, J., G. Acuna, R. Pérez-Correa, H. Jorquera, and P. Molin, "A hybrid representation approach for modelling complex dynamic bioprocesses," *Bioprocess Engineering*, **22**, 547–556 (2000).

Thompson, M. L. and M. A. Kramer, "Modeling chemical processes using prior knowledge and neural networks," *AIChE J.*, **40**(8), 1328–1340 (1994).

Tsen, A. Y.-D., S. S. Jang, D. S. H. Wong, and J. Babu, "Predictive control of quality in batch polymerization using hybrid ANN models," *AIChE J.*, **42**(2), 455–465 (1996).

van Breusegem, V. and G. Bastin, "Reduced order dynamical modelling of reaction systems: a singular perturbation approach," *Proc. 30th Conf. Decision and Control, Brighton*, pages 1049–1054 (1991).

van Can, H. J. L., C. Hellinga, K. C. A. M. Luyben, J. J. Heijnen, and H. A. B. te Braake, "Strategy for dynamic process modeling based on neural networks in macroscopic balances," *AIChE J.*, **42**(12), 3403–3418 (1996).

van Can, H. J. L., H. A. B. te Braake, S. Dubbelmann, C. Hellinga, and K. C. A. M. Luyben, "Understanding and applying the extrapolation properties of serial gray-box models," *AIChE J.*, **44**(5), 1071–1089 (1998).

Villadsen, J. V. and M. L. Michelsen, *Solution of Differential Equation Models by Polynomial Approximation*. Prentice-Hall (1978).

von Watzdorf, R. and W. Marquardt, "Fully adaptive model size reduction for multicomponent separation problems," *Comput. Chem. Eng.*, **21**(Suppl.), S811–S816 (1997).

von Wedel, L. and W. Marquardt, "Cheops: A case study in component-based process simulation," *AIChE Symp. Ser.*, **96**(323), 494–497 (2000).

Wisnewski, P. A. and F. J. Doyle III, Nonlinear model reduction using Hankel-norm approximation, Technical report, University of Delaware (1996a).

Wisnewski, P. A. and F. J. Doyle III, "A reduced model approach to estimation and control of a Kamyr digester," *Comput. Chem. Eng.*, **20**(Suppl.), S1053–S1058 (1996b).

Wortelboer, P. M. R., M. Steinbuch, and O. H. Bosgra, "Iterative model and controller reduction using closed-loop balancing, with application to a compact disc mechanism," *Int. J. Robust and Nonlinear Control*, **9**(3), 123–142 (1999).

Zander, H.-J., R. Dittmeyer, and J. Wagenhuber, "Dynamic modeling of chemical reaction systems with neural networks and hybrid models," *Chem. Eng. Technol.*, **21**(7), 571–574 (1999).

Zbiciński, I., P. Strumiłło, and W. Kamiński, "Hybrid neural model of thermal drying in a fluidized bed," *Comput. Chem. Eng.*, **20**(Suppl.), S695–S700 (1996).

Zhou, K. M., C. D'Souza, and C. J. R., "Structurally balanced controller order reduction with guaranteed closed-loop performance," *Sys. Cont. Let.*, **24**(4), 235–242 (1995).

Model Requirements for Next Generation Integrated MPC and Dynamic Optimization

Ton C. Backx*
IPCOS Technology
Bosscheweg 145a, 5282 WV Boxtel, The Netherlands
Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

This paper outlines the requirements imposed upon models and modeling techniques applied for high performance model based control systems and model based optimizers for support of the operation of (chemical) processes. An overview is given of on-going developments in the area of integrated high performance model predictive control and model based dynamic optimization (INCOOP and IMPACT projects). To enable tight control of processes at given Cpk values and within imposed 6-sigma intervals on a variety of process variables, the models applied for control and for optimization have to be sufficiently accurate both as a function of frequency as well as a function of the time varying operating conditions. Requirements on the models are summarized. Modeling of relevant process behavior for control and optimization is a very significant cost factor in overall application development. Reduction of these costs by extensive use of a-priori knowledge and by integration of various modeling techniques is discussed.

Keywords

Model-based control, Dynamic optimization, Process modeling

Introduction

Industrial processes are subject to continuous improvement of performance with respect to yield, quality, flexibility and innovation. Market developments, legislation, social and environmental requirements have started to create a need for continuously better predictability of process performance and more flexible operation of the processes over the past two decades (García and Prett, 1986; Prett and Garcìa, 1988; Morari, 1988; Backx et al., 1998, 2000). Chemical Processing Industries are currently facing an enormous challenge: Within the next years they have to realize a turnaround in their financial performance to remain attractive for capital investors.

Gradual decline of the productivity of invested capital over the past three decades has become a major point of concern of the chemical processing industries. The financial performance of many companies belonging to the Chemical Processing Industries is lagging economic developments of the market, which makes it hard to compete with industries that do better than average like for example the Information and Communication Technology oriented industries.

The relatively poor performance of the Chemical Processing Industries may be explained from the hesitation of industry to adapt to and anticipate fast changes in the market. Fast developments of new markets, stimulated by rapidly adopted microelectronic and information technology developments, have created a complete turnaround of the market. The market has turned from a regional, mainly supply driven market into a worldwide, demand driven market over the past 15 years. Many of the processing industries and the Chemical Processing Industries in particular did not yet follow this turnaround and are still mainly organized to predominantly produce in a supply driven way. Latest developments of computing, modeling and control technologies of the past decade offer a great opportunity however to quickly realize the changes from the technical side. Organizational adaptations have to be made accordingly though.

Wide application of model based control and optimization technology in chemical process industries is hampered until now by the following limitations of current state-of-the-art technologies:

- Costs of application development are (too) high in relation to direct return on the investments that have to be made for the development of these applications
- Most of the currently applied MPC technologies in industry are based on the use of process models for prediction and control that approximate process dynamics in a linear, time invariant way. Chemical plants are often producing a mix of products with different specifications, which involves transitions between different operating points with corresponding different process dynamics in each of these operating points related to the non-linearities in process behavior
- Performance of state-of-the art MPC technologies in terms of their capabilities to reduce variances of critical process variables and product parameters is restricted by limitations in the models applied for prediction and for control to cover the whole frequency range at which the process is operated. This limitation stems from the techniques and procedures applied for process testing, process identification and model validation
- Extensive (re-)use of available information and knowledge on dynamic process behavior is one of the ways to significantly reduce costs of application

*Ton.Backx@IPCOS.nl

development related to process testing and the engineering effort involved. State-of-the-art technologies hardly support this.
- Currently applied techniques for (closed loop) model based optimization of plant performance use first principles based steady state models. Consistency of these models with the models obtained from process identification and applied inside the model predictive control systems is not guaranteed. As a consequence the performance of the systems is restricted to satisfy robustness requirements over the applicable operating envelope.
- Currently applied techniques for (closed loop) model based optimization of plant performance are limited to steady state only, which implies that they don't support transitions between various operating points and adequate response to plant upsets.

This paper discusses developments that overcome the obstacles summarized above. These developments are done in the context of two development projects: IMPACT and INCOOP. The IMPACT project (Eureka project with label 2063) has the objective of developing dedicated model based control and transition optimization techniques for various types of PE/PP polymer manufacturing processes. The INCOOP project (EC funded 5th framework) focuses on the integration of non-linear high performance model based control and dynamic plant optimization.

Section 2 of the paper first outlines requirements imposed upon process operation in accordance with market requirements. In section 3 these requirements on process operation are translated to requirements on the models that are applied for control and optimization. Section 4 focuses on the dynamic operation of plants and the additional impact on models applied for the support of such operation. Section 5 discusses model requirements and modeling approaches for high performance model based process control. Section 6 shows some preliminary results on a PE polymer manufacturing process. Final remarks and conclusions are given in section 7

Requirements on Process Operations

The main requirement on process operation is: To produce products that meet specifications in the requested volume at the right time and at minimum cost in accordance with an imposed schedule respecting operational and legal permit constraints. The optimization of process performance involves fulfilling these requirements the best possible way. Realization of the requirements implies that the optimization has both to exploit the technical capabilities of the process and to use freedom in process operation to achieve economic optimization without taking too much risk on unplanned process shutdowns. As a consequence the optimization has to blend the functionality of current generation economic optimizers, which generally do not consider process dynamics with state-of-the-art technology from the field of dynamic optimization, which, however, usually does not concentrate on economic problems, but rather on control-oriented problems. A dynamic 'Money Conservation Law' is developed to integrate both worlds. The objective function V is written as a sum of profit made along a trajectory and a capital inventory term (Van der Schot, 1998):

$$\max V = \max \int_{t_0}^{t_f} EUR_{revenues}(t) - EUR_{expenses}(t)\,dt + EUR_{inventory}\Big|_{t_0}^{t_f} \quad (1)$$

This trajectory gives the recipe for an optimal transition between operating conditions or for the optimum path to recover from encountered disturbances in a certain operating point. It comes down to maximizing the added value of the process over a fixed time horizon $[t_0, t_f]$. In this, we define the Euro flows (€/hr) as the product of physical flows (kg/hr) and product prices (€/kg). These prices will depend on product quality and on market conditions. The product quality is governed by process operation and can be simulated by a rigorous dynamic model of the (chemical) process. This model can be included in the overall optimization as a set of equality constraints.

The functions $\Phi(x(t), u(t))$ and $\Psi(x(t_f), u(t_f))$, that link process manipulations $u(t)$ and process states $x(t)$ to performance measures in the applied criterion function (cf. Equation 2) related to this optimization problem, essentially consist of two separate components:
- a smooth non-linear part related to the process transfer characteristics obtained from the DAE model described by the equality constraints in Equation 2. This model connects manipulated variables and disturbances to process outputs,
- a highly non-linear part that connects product properties to market values of these products.

$$\max_{u(t)} \quad V = \int_0^{t_f} \Phi(x(t), u(t))\,dt + \Psi(x(t_f), u(t_f)) \quad (2)$$

subject to

$$0 = f(\dot{x}(t), x(t), u(t))$$
$$0 \leq c(x(t), u(t))$$

The first component—the one that describes the process transfer characteristics—has a high complexity due to the process mechanisms involved. The second part has limited complexity but involves discontinuous functions, which are very nasty from an optimization viewpoint. This is due to the fact that products within high quality specifications have a high, market determined fixed

value. Products that meet looser specifications generally have a lower, also market determined, fixed value. Off-spec products have an even lower value that even may be negative (cf. Van der Schot et al., 1999). The market values of products relate to specified ranges for a selected group of product properties. Products with properties in the specified intervals all have the same value. The product value changes as soon as one or more of the product properties exceed the specified tolerances for these parameters.

Operation of the process is subject to operating constraints on process inputs, process outputs and/or rates of change of these variables. Both performance and robustness of the control systems directly relate to the quality and accuracy of the models used as prediction models in the control systems. The problem to be solved in order to enable process operation as outlined, is to get models that satisfy the following conditions (cf. De Moor and Berckmans, 1996; Jacobsen et al., 1991; Johansen and Foss, 1995; Johansen, 1996; Kemna and Mellichamp, 1995; Lindskog and Ljung, 1995; Ljung et al., 1991; Rivera and Gaikwad, 1995; Skogestad et al., 1991; Tsen and Wong, 1995; Tulleken, 1993; Wei and Lee, 1996):

- describe all process dynamics that are relevant for model predictive control and dynamic optimization of the process in accordance with given specifications
- cover the full operating envelope of the process consisting of specified operating points and transition trajectories between these operating points covered by process optimization

Model Requirements

In order to get sets of models that accurately reflect all control relevant dynamics of the process in all selected operating points of a real plant, extensive plant tests are required, if traditional model predictive control (MPC) system design methodologies are applied. Traditional MPC design techniques based upon the application of system identification techniques require testing of the process in each operating point. The tests involve persistent excitation of all relevant process dynamics with sufficient energy and during a sufficiently long time to ensure good identification results (Ljung, 1987; Ljung and Söderström, 1983). The duration of a plant test is governed by the time to steady state (T_{ss}) of the unit that is tested, by the number of process inputs that have to be tested (N_{inp}) and by the type of test signals applied. A system identification related plant test typically takes a time T_{test} in the range given by (3) for each operating point that needs to be tested, if model quality has to be ensured.

$$3 \cdot Int(\frac{N_{inp}}{5} + 1) \cdot T_{ss} \leq T_{test} \leq 5 \cdot N_{inp} \cdot T_{ss} \qquad (3)$$

The type of test signals applied governs the actual required time for plant testing. The effort involved in plant testing makes application of this approach to processes operated in a broad range of operating points, or to processes operated in specific operating points for a relatively short time compared to the dynamic response times of the process, economically or even technically unfeasible. Many of the slow dynamics, which to a large extent are related to physical phenomena like e.g. material transport, residence times, warming up and cooling down of huge heat capacities with restricted energy flows, etc. can be modeled quite accurately using first principles based modeling techniques. Often these dynamics also don't vary much even under operating condition changes. Relative fast dynamics related to local physical, chemical, biological phenomena frequently are hard to be modeled accurately using first principles modeling techniques. Examples of such hard to model phenomena are for example flows through complex piping systems, turbulent flows around a valve, inhomogenity in mixtures in reactor tanks, specific reaction complexes (e.g. polymerization, cracking, ...), kinetics in complex chemical reaction systems, metabolisms of biomass, etc. The use of validated, first principle model based dynamic process simulators that accurately reflect the slow process dynamics as a reference for the design of the model predictive control system may overcome the problem of the required long plant tests for process identification in each operating point. Relatively short dedicated plant tests at well-selected operating conditions may be applied to accurately model the relevant fast dynamics using traditional process identification techniques. A combination of both modeling approaches in a heuristic way to model all relevant dynamics for process control may be a feasible way to solve the problem. The specific assumption made here is that a high fidelity, first principles based, dynamic process model covers the main process mechanisms that govern the low frequency transfer characteristics of the process over a sufficiently large operating range covering all relevant operating conditions (Backx, 1999).

Model based or model predictive process control in combination with trajectory optimization to find optimum dynamic transition paths can contribute to meeting the new needs on high performance plant operation. These technologies make extensive use of available knowledge on the dynamic behavior of processes for continuously driving the process to desired operating and performance conditions. Model predictive control enables revision of the control strategy on a sample-by-sample-basis using latest information on the status of the plant and its environment. The knowledge of process behavior is represented in the form of a mathematical model that describes the process dynamics, which are relevant for control. This model is explicitly used in the controller for predicting future process responses to past

input manipulations and measured disturbances and to calculate best future input manipulations that satisfy the control objectives. The model is assumed to reflect all significant dynamic properties of process input/output behavior. It furthermore has to enable simulation of the future process outputs on the basis of known past process inputs within a pre-specified operating envelope of the process with limited inaccuracy and uncertainty (Cutler and Ramaker, 1980; Muske and Rawlings, 1993; Froisy, 1994; Qin and Badgwell, 1997; Richalet, 1997).

Dynamic Operation of Processes

Process operation has to become fully dynamic in stead of quasi-steady-state to meet the market requirements described in section 1. Processes in this respect have to be viewed as dynamically operated elements of a supply chain. This supply chain is composed of severa,l mutually interacting processes (Backx et al., 1998). The total supply chain has to meet the requirements of the market. As a consequence specific processes have to be operated in a way that production and products can easily be adapted to the changing and highly fluctuating market demand. Ideally a plant has to be controlled in such a way that production can follow market demand to the extent that 'Just-in-time' production at specifications might be feasible thus minimizing stocks of finished products and intermediates and enabling maximum capital productivity. The consequence of operating plants this way is that processes need to be operated under conditions that are fully synchronized with demand despite a wide range of dynamic effects that govern production behavior. Several sources of dynamics may be discriminated in this respect that all contribute to the overall dynamic behavior of a production plant:

- Marketplace dynamics
- Ecosphere dynamics
- Macro scale plant dynamics
- Meso scale unit process dynamics
- Micro scale reaction dynamics

The *marketplace* dynamics are characterized by long cycles. The cyclic behavior of the market place typically ranges from months to many years. The actual dynamic behavior is affected by various types of discrete events, which cause relatively fast (days to weeks time-frame) fluctuations in the market conditions. Examples of such events are feedstock (or utilities) availability, product demand as well as the prices for both feedstock and product. The market behavior is largely unpredictable due the large number of influencing factors, which are mostly unknown. It is a major disturbance that has to be coped with.

The dynamics of the ecosphere also shows cycles of different time scales. Typical examples of these *ecosphere* dynamics are the sometimes very rapidly changing weather conditions (e.g. a rain shower, a thunderstorm, a clouded sunny day, ...), the fast day-night patterns or the several orders of magnitude slower seasonal cycles with different temperature, humidity, waste water requirements and cooling water conditions. Like market behavior also ecosphere behavior is largely unpredictable. It generally has a large impact on actual process performance. Due to this large influence on actual process behavior it has to be compensated by control systems and optimizers.

The mix of unit processes that together form the plant determines *plant dynamics*. The dynamic behavior of the plant is governed both by the dynamics of each of the unit processes and by the dynamics related to recycles (e.g. recovery and re-use of materials, waste water, ...) and integration (e.g. heat integration, cooling water, ...). These dynamics may span several decades ranging from minutes to several days. As an example a Cracker plant like an Ethylene plant may be mentioned. This plant consists of reactors with dynamics in the minute range as well as distillation columns with dynamics that can range up to a day. Heat integration may cause such a plant to show dynamic behavior spanning days up to weeks.

Unit process dynamics normally span a few decades on a time scale. The actual unit process dynamics highly depend upon the type of unit process and may range from sub-seconds to days. Forming processes (e.g. Steel rolling, Extrusion of Polymers, glass forming, ...) are examples of processes showing sub-second to minute dynamics. Fluidized bed or slurry loop polyethylene and polypropylene reactors, high purity ethylene and propylene distillation columns and glass melting tanks are examples of unit processes with dynamics that may range from several hours up to one or even more days.

Reaction dynamics are usually very fast. They can span several decades of a time scale as well, but often they are in the micro second to second range.

All these overlapping dynamics together form the dynamics that have to be handled by the systems that are applied for operation of the plant. Adequate control of this wide range of dynamics is crucial to meet both flexibility requirements and quality requirements of products in accordance with continuously changing market demand.

The effect of adequate control of a wide range of dynamics in relation to product quality at the unit process level is shown in Figure 1. The bandwidth of the closed loop system determines the reduction in variance due to disturbances of the controlled variables. The reduction in variance that will be achieved by the control system can be estimated by calculating the frequency dependant reduction of the power spectral density (PSD) of the con-

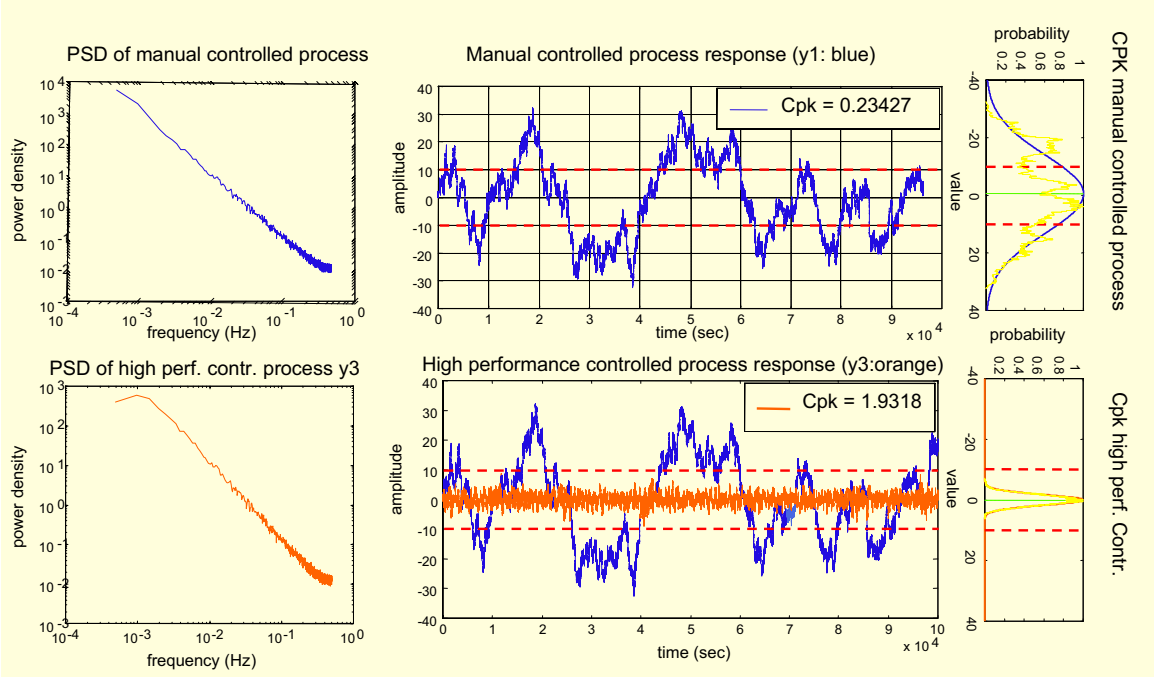


Figure 1: Bandwidth of a 'quasi-steady-state' (q.s.s.) control system and a 'High Performance' control system in relation to performance.

trolled variables Y(f) by closed loop control. Application of the theorem of Parceval (e.g. Papoulis, 1984) gives a direct estimate of the variance reduction achieved by the control system:

$$\begin{aligned}\sigma_{cl}^2 &= \frac{\Phi_{yy}|_{cl}}{\Phi_{yy}|_{ol}} \cdot \sigma_{ol}^2 \\ &= \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty} |Y_{cl}(f)|^2\, df}{\frac{1}{2\pi}\int_{-\infty}^{\infty} |Y_{ol}(f)|^2\, df} \cdot \sigma_{ol}^2 \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |Y_{cl}(f)|^2\, df\end{aligned} \tag{4}$$

$\Phi_{yy}|_c l$ and $\Phi_{yy}|_o l$ respectively denote the closed loop and open loop power spectral densities of the controlled signal. The closed loop control system acts as a high-pass filter with a cut-off frequency determined by the bandwidth of the control system to reject output disturbances.

Process Modeling for High Performance Model Based Process Control

High performance process operation requires fully reproducible and predictable control of process units both during steady operation in a selected operating point as well as during transition between different operating points. This implies the need for models that accurately describe process dynamics over a frequency range that exceeds the intended bandwidth of the control system. As is shown in Figure 1 this bandwidth is governing the performance that can be achieved. Also for close tracking of transition trajectories a large bandwidth control system is necessary to enable high performance disturbance rejection during transition and to make fast transitions possible. The accuracy requirements on the models applied for control are fully dictated by:

- the bandwidth of the control system,
- the disturbance characteristics (amplitude ranges or probability density functions, frequency contents) of the variables that need to be controlled
- the process transfer characteristics (bandwidth of process transfers in relation to the bandwidth of the disturbances) over the operating envelope that needs to be handled by the control system

The bandwidth of the control system and the bandwidth of the model are linked by the role of the model in model predictive and model based control systems:

> *The model is assumed to accurately predict process output behavior thus minimizing the closed loop gain and providing an accurate estimate of the actual output disturbance (cf. Figure 2).*

From the frequency point on where the model looses accuracy in describing the process transfer behavior, the model predictive control system turns from a primarily feedforward driving control system into a classic (multivariable) feedback controller. This feedback controller

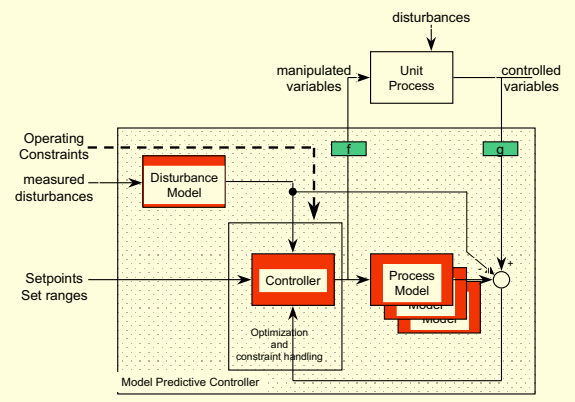


Figure 2: Structure of a model predictive control system.

has all the well known difficulties with stability and performance and is hard to tune (Rosenbrock, 1970; MacFarlane, 1979). Therefore the control system is tuned to not give any control performance any more in this frequency range in most industrial applications. As a consequence it only operates over a restricted (low) frequency range and does not reject higher frequency disturbances. Recovery from process upsets usually also takes much longer than required because of the limited bandwidth of the closed loop controlled process.

High performance control requires models with a bandwidth that cover all relevant process transfer dynamics over a frequency range given by the necessary disturbance rejection. The models have to enable accurate description of the process transfer dynamics in such a way that the model predicted outputs and the actual process outputs coincide both in amplitude and in phase over the relevant frequency range. Especially the requirement that the predicted and actual signal may not have a significant phase error makes that the bandwidth of the model has to exceed the closed loop bandwidth of the controlled process. Sensitivity for modeling errors over the full relevant frequency range is determined by the complexity of the process dynamics. Frequency ranges with large changes in phase shift between process inputs and outputs in general have to be treated with great care due to the sensitivity for poor performance or even instability in these frequency ranges. Processes often show large changes in phase shift between inputs and outputs especially around the cut-off frequencies, which makes that this frequency range almost always needs to be modeled accurately for high performance control.

First principles based modeling techniques in general are not very well suited to accurately model all process mechanisms that govern the higher frequency amplitude and especially phase characteristics. This is caused by the many interacting mechanisms that dominate the process behavior in this frequency range (e.g. inhomogenities in materials, temperatures, concentrations of components, turbulent flow patterns, non-homogeneous mixture of components, ...). Process identification techniques on the contrary are very well capable in capturing this behavior in models, if it is stationary. The following aspects are critical for process identification techniques to model process behavior accurately in the critical frequency ranges for high performance model based process control:

- The process needs to show stationary and reproducible behavior within the operating envelope
- The applied modelsets for system identification have to enable accurate modeling of the amplitude/phase characteristics of all process transfers in the critical ranges; i.e. complexity of the applied modelsets needs to be sufficiently high (Willems, 1986a,b, 1987)
- The process data used for estimation of the model parameters have to contain sufficient information on these process characteristics in ratio to encountered disturbances during testing; the process has to be persistently excited with adequate signal-to-noise ratio's at each of the process outputs (controlled variables) (Ljung, 1987)
- The test signals applied to the process have to excite the process in a balanced way to enable equally accurate modeling of low and high gain directions; this is of particular importance if the control system has to enable high performance in low gain directions of the process (e.g. dual quality control in distillation cf. Figure 3)
- The criterion function applied for estimation of the model parameters has to be consistent with the later use of the model in the model based control system, which usually implies prediction of future process outputs over some future time horizon on the basis of known past process inputs and known past process outputs.

High performance process control starts with the selection of appropriate process input (Manipulated Variables) and process output variables (Controlled Variables). The selected set of manipulated variables together with the information obtained from the measured disturbance variables have to enable compensation of the disturbances that affect the controlled variables. Furthermore the selected set of manipulated variables has to enable fast, predictable and reproducible transition between selected operating points of the process. Multivariate statistical analysis tools support selection of appropriate process inputs and process outputs for control (Yoon and MacGregor, 2000; Clarke-Pringle and MacGregor, 2000).

High performance model based process control in general requires process models that reflect process behavior

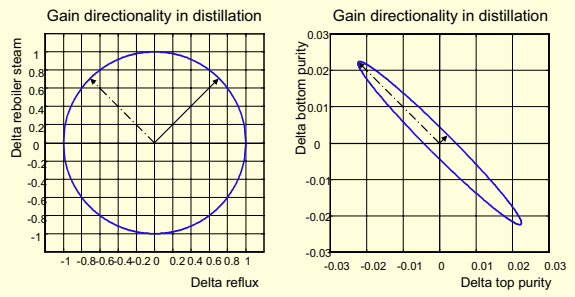


Figure 3: Directionality in process transfer characteristics of a local operating point in distillation.

over a large frequency range. A distillation column may be used as a simple example to illustrate this. An industrial column with 15 trays will have a time to steady state of approximately 1.5 hours. A change in top product quality in response to a top pressure change will be noticed well within a minute however. This means that even for this simple column the range of dynamics to be covered already is ¿100. This number is the ratio of fastest relevant frequency (response to a pressure change in this example) over slowest relevant frequency (dynamics that govern settling towards steady state) for high performance model based control. A finite step response model needs at least 200 relevant samples to cover such a range of dynamics. The further the slowest and fastest relevant frequencies get apart the more parameters will be needed to describe this process behavior with finite step responses. The required sample rate for the model is governed by the highest relevant frequency. The time span that needs to be covered by the model is governed by the low frequencies that determine settling to steady state of the process. The time length of the model divided by the applied sampling time gives the complexity of the model for non-parametric type models like finite step responses or finite impulse responses. The complexity is a measure for the number of parameters of the model. This implies that these non-parametric types of models are requiring an increase in model complexity to cover all relevant dynamics, if the fastest and slowest relevant dynamic modes get further apart. In most of todays model predictive control applications the higher frequency characteristics of the process are not included in the models to prevent the models from becoming too complex.

The critical issue for high performance model based control is accuracy of the applied model in certain frequency ranges. Model sets that support coverage of wide frequency ranges without a necessarily large increase of model complexity are parametric models like e.g. state space models. State space models can handle the wide range of dynamics encountered in most industrial processes with a complexity dictated by the number of dynamic modes in the observed dynamic process behavior. The number of parameters required to accurately describe process behavior does not grow with the ratio of fastest and slowest relevant process dynamics, if parametric models are applied. It depends upon the complexity of the process dynamics i.e. the order of the difference equations that approximately describe the relevant process behavior with sufficient accuracy.

The data used for process identification has to contain information that enables reconstruction of the process transfer characteristics at the critical frequency ranges with the required accuracy. This implies that test signals applied to the process have to span the input space later used by the control system in such a way that all relevant frequencies are well excited and that the variables that will be controlled are showing balanced responses. This also has to hold for the directions that show the largest and the smallest gains over the relevant frequency range. Model accuracy obtained from process identification is governed by the signal-to-noise ratios encountered over the full frequency range over all output directions (Zhu and Backx, 1993; Zhu et al., 1994).

Subspace and orthonormal basis function based techniques are latest developments in multivariable process identification techniques that enable modeling of a wide range of process dynamics, without requiring detailed a-priori knowledge on the order and structure of the multivariable system (Van Overschee and De Moor, 1993, 1994; Verhaegen and Dewilde, 1993a; Viberg, 1995; Heuberger et al., 1995; Van den Hof and Bokor, 1995; Van Donkelaar et al., 1998). Essential in both approaches is that observed input-output behavior of the process as represented by the process data collected and pre-treated for identification are projected to a subspace spanned by a set of (orthonormal) basis functions that can represent all relevant process dynamics. The difference between both methods stems from the type of basis functions applied and by the way the selection of these basis functions is established.

Subspace identification techniques determine the basis functions directly from Hankel U and Y matrices constructed from the process input and process output data applied for process identification. The basis functions are obtained by data compression via an RQ factoriza-

tion (Verhaegen and Dewilde, 1993a,b).

$$\begin{aligned}\begin{bmatrix} U \\ Y \end{bmatrix} &= \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \cdot \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \\ &= \begin{matrix} (n-1)p \rightarrow \\ p \rightarrow \\ (n-1)q \rightarrow \\ q \rightarrow \end{matrix} \begin{bmatrix} \overbrace{}^{np} & \overbrace{}^{nq} \\ R_{11}^1 & 0 \\ * & 0 \\ R_{21}^1 & R_{22}^1 \\ * & * \end{bmatrix} \cdot \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \\ &= \begin{matrix} p \rightarrow \\ (n-1)p \rightarrow \\ q \rightarrow \\ (n-1)q \rightarrow \end{matrix} \begin{bmatrix} \overbrace{}^{np} & \overbrace{}^{nq} \\ r_u & 0 \\ R_{11}^2 & 0 \\ r_{y1} & r_{y2} \\ R_{21}^2 & R_{22}^2 \end{bmatrix} \cdot \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \end{aligned} \tag{5}$$

In these equations n denotes the order of the state space system representation, p denotes the number of inputs and q denotes the number of outputs. Calculation of the singular value decomposition of the matrix $[(R_{21} - H_n \cdot R_{11})\ R_{22}]$ that links future process output behavior to past and current input manipulations:

$$\begin{aligned}\Lambda_n \cdot X_l &= Y - H_n \cdot U \\ &= [(R_{21} - H_n \cdot R_{11})\ R_{22}] \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \\ &= U_s \cdot S_s \cdot V_s^T \cdot \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}\end{aligned} \tag{6}$$

provides the matrix U_s^1 by selection of the first $(n-1) \cdot q$ rows of U_s. H_n is the Hankel matrix constructed from n Markov parameters estimated from the available process input output data.

The state space system matrices are subsequently calculated by least squares solution of the following set of equations:

$$\begin{bmatrix} U_s^{1^*} \cdot (R_{21}^2 - H_{n-1} \cdot R_{11}^2) & U_s^{1^*} \cdot R_{22}^2 \\ r_{y1} & r_{y2} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} U_s^{1^*} \cdot (R_{21}^2 - H_{n-1} \cdot R_{11}^1) & U_s^{1^*} \cdot R_{22}^1 \\ r_u & 0 \end{bmatrix} \tag{7}$$

Orthonormal basis function based algorithms use an initial basis function generation step in which a rough low order approximation of the process dynamics is used in combination with an inner transfer function to generate the applied set of basis functions (Van Donkelaar et al., 1998; Heuberger et al., 1995; Van den Hof and Bokor, 1995).

$$G(z, \vartheta) = \sum_{i=1}^{N} W_i(\vartheta) \cdot F_i(z) \tag{8}$$

represents the process transfer as a function of time shifts z and of the model parameters . Wi() indicates the weights on each of the basis functions and Fi(z) refers to the basis functions applied.

A well known set of basis functions commonly applied in industrial applications is the pulse basis:

$$f_i(z) = z^{-i} \tag{9}$$

Substitution of this basis in (8) gives the Finite Impulse Response representation of the process. This basis requires many parameters Wi to describe all relevant process dynamics, if the process dynamics cover a wide range as discussed above. A significant reduction in the number of parameters required for describing all relevant process dynamics may be obtained by using Laguerre or Kautz sets of basis functions:

Laguerre:

$$f_i(x) = \sqrt{1 - a^2} \cdot \frac{(1 - az)^i}{(z - a)^{i+1}} \tag{10}$$

Kautz:

$$\begin{aligned} f_{2i}(z) &= \\ &z \frac{\sqrt{1 - c^2} \cdot (z - b)}{z^2 + b(c-1)z - c} \cdot \left[\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right]^i \\ f_{2i+1}(z) &= \\ &z \frac{\sqrt{(1 - c^2) \cdot (1 - b^2)}}{z^2 + b(c-1)z - c} \cdot \left[\frac{-cz^2 + b(c-1)z + 1}{z^2 + b(c-1)z - c} \right]^i \end{aligned} \tag{11}$$

A significant reduction in the number of parameters N is achieved by selecting the coefficients of these basis functions—parameter a in the Laguerre basis functions and parameters b, c in the Kautz basis functions—in such a way that the first elements of the set of basis functions closely represents the relevant process dynamics. The Laguerre basis works well for systems that show smoothly damped behavior. The Kautz basis works fine for system with badly damped transfer characteristics. To achieve a close approximation with a few basis functions of actual processes that show more complicated transfer dynamics, a better performance is obtained by making use of a generalized orthonormal basis as described in (Heuberger et al., 1995). This allows the selection of a set of orthonormal basis functions of which the first functions closest approximate the actual observed process dynamics.

Both methods—the subspace method and the orthonormal basis function method—don't directly give an optimum model with minimum output error. In general the models are very good starting points for a final nonlinear output error optimization as described in (Falkus, 1994). This output error optimization allows fine tuning of the state space model to accurately describe process dynamics in the critical frequency ranges. Application of the identification techniques in closed loop process operation allows the identification techniques to automatically concentrate on realizing the highest model accuracy in the critical frequency ranges for closed loop control (Hjalmarsson et al., 1996).

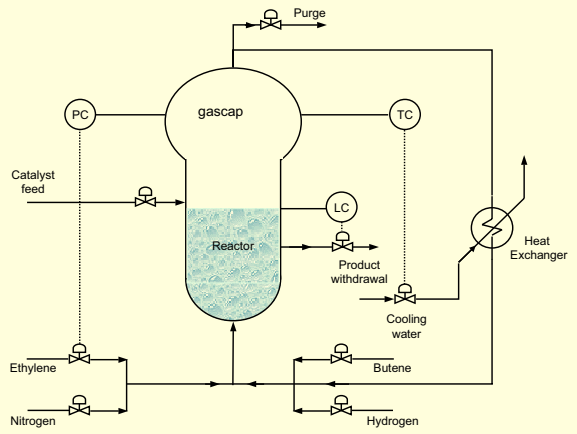


Figure 4: Fluidized bed gas phase HDPE reactor.

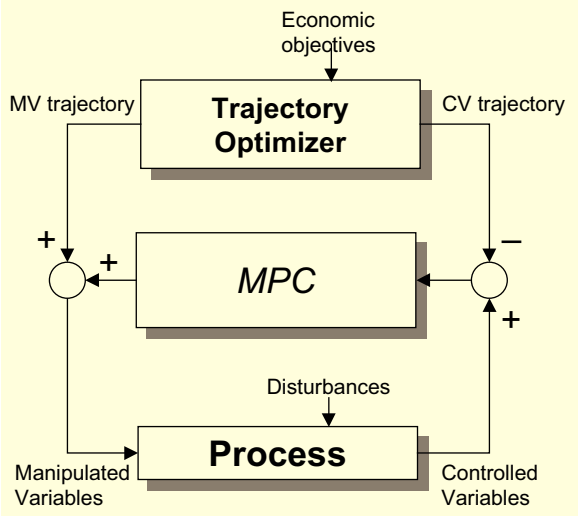


Figure 5: Integration of the delta mode MPC system and the rigorous model based dynamic trajectory optimizer.

Some Results

Control of a gas phase fluidized bed HDPE polymer reactor is used as an example to show some initial results of performance improvement that can be achieved by integrated high performance model based control and dynamic optimization (Figure 4). As discussed in the previous sections performance of the model predictive control system and the dynamic optimizer is governed by accuracy of the applied models.

In this example a set of approximate linear models is applied to realize a high performance model predictive transition control system that optimizes polymer production performance both during normal operation as well as during grade transitions. The dynamic grade transition optimization is done on the basis of a first principles based dynamic model of the process as discussed in section 2 and implemented in gPROMS.

The model predictive control system developed for control of the polymerization reactor has been designed to cover a broad operating range of the process. The control system simultaneously manipulates Monomer/Co-monomer ratio, Hydrogen/Monomer ratio, Catalyst flow, Gascap Pressure and bed Temperature. The model predictive control system controls Density, Melt Index and Production Rate. Direct or inferential measurements of the controlled variables needs to be available for this purpose. The control system operates in delta mode and includes linearizing functions to cover the large, non-linear operating range of the process with sufficient accuracy (Figure 5).

In case no on-line measurement of the controlled variables are available, the controller will calculate the required control actions on the basis of model predictions of these variables between the updates of the real measured process values. This functionality enables robust operation of the control system at various sample rates of the product quality (Melt Index and Density) measurement. It will make the control rather insensitive for changes in the sample rate as long as the model predictions don't show severe errors.

The control system is designed for high performance, robust control of the polymer properties over a broad operating range of the process. Table 1 gives an overview of the total set of grades covered by the integrated control and dynamic optimization system. Performance of the control system is stable over this operating range. Figure 6 shows the results of the transitions from O7 to P3 (left two columns) and R0 to P3 (right two columns) as an example of two grade changes subject to a variety of external disturbances acting on the process. The picture clearly shows the improvements obtained over traditional grade transition control. The transitions are fast and the polymer properties remain well within specifications at both grades.

The transitions shown are shortcuts in the normal grade slate. The transitions involve a large change in the density of the polymer, which implies a large change in the co-monomer/monomer ratio. In order to enable a fast transition the production rate is heavily decreased to get a minimum amount of wide spec product during the transition. The large change in production rate implies that a severe change is process behavior is encountered due to the wide operating range spanned. The models applied for prediction in the control system have to accurately describe these changing process characteristics to enable close tracking of the optimum transition trajectory calculated by the optimizer. The performance improvement achieved by the integrated MPC and dynamic trajectory optimization in this example represents an economic benefit of € 117330 for the transition of O7 to P3. The benefit related to the improvement of R0

Grade Names			DENS 964 966	DENS 954 956	DENS 944 946	DENS 934 936	DENS 924 926	DENS 914 916	DENS 904 906	DENS 894 896
LNMI	-2.4	-2.6	GO	G1	G2	G3	G4	G5	G6	G7
LNMI	-1.4	-1.6	H0	H1	H2	H3	H4	H5	H6	H7
LNMI	-0.4	-0.6	I0	I1	I2	I3	I4	I5	I6	I7
LNMI	0.6	0.4	J0	J1	J2	J3	J4	J5	J6	J7
LNMI	1.6	1.4	K0	K1	K2	K3	K4	K5	K6	K7
LNMI	2.6	2.4	L0	L1	L2	L3	L4	L5	L6	L7
LNMI	3.6	3.4	M0	M1	M2	M3	M4	M5	M6	M7
LNMI	4.6	4.4	N0	N1	N2	N3	N4	N5	N6	N7
LNMI	5.6	5.4	O0	O1	O2	O3	O4	O5	O6	O7
LNMI	6.6	6.4	P0	P1	P2	P3	P4	P5	P6	P7
LNMI	7.6	7.4	Q0	Q1	Q2	Q3	Q4	Q5	Q6	Q7
LNMI	8.6	8.4	R0	R1	R2	R3	R4	R5	R6	R7
LNMI	9.6	9.4	S0	S1	S2	S3	S4	S5	S6	S7

Table 1: Grade definition table.

to P3 represents a value of € 49081 at the given market values. The actual benefit and the corresponding optimum transition strategy strongly depend upon market conditions for first grade and wide spec products.

Conclusions

Changing market conditions enforce chemical processing industries to better utilize process capabilities. Process operation needs to be closer tied with market demand to improve capital productivity. Currently applied state-of-the-art model predictive control and model based optimization techniques don't support close tracking of optimum operating conditions. A main reason for this is the limitation in accuracy of the models applied for control and optimization. High performance control requires models that accurately describe all relevant process dynamics over a wide operating range. Especially the frequency ranges where the process is showing huge changes in its transfer phase characteristics are critical and need to be modeled accurately. High performance in general requires models that cover the full process transfer frequency range accurately. The frequency range covered accurately by the model dictates the ultimate bandwidth of the closed loop model predictive control system. This bandwidth in its turn governs the disturbance rejection capabilities of the control system and therefore the resulting capability to achieve product and process quality requirements. The bandwidth of the control system furthermore governs the capabilities to closely track transition trajectories that enable cost effective changeovers between various operating points.

Accurate modeling of all relevant process dynamics for the entire process operating envelope envisaged requires integration of rigorous modeling techniques with process identification techniques to be economically feasible. Process testing needs to be minimized due to the high cost involved with plant tests. Extensive (re-)use of a-priori knowledge on process dynamics is enabling a significant reduction in test time required. Plant tests have to focus primarily on critical frequency ranges for control. These dedicated tests can be relatively short and less expensive in general.

Optimum transition control is not supported by steady state optimization techniques. New concepts based on the use of dynamic plant models have been discussed that enable exploitation of plant dynamics for optimization of economic performance of plants. Consistency between model-based optimization and model-based control is crucial for high performance in dynamic plant operation. Intential dynamic operation of a plant opens opportunities for very significant improvement of plant economics and capital productivity. Market driven operation of plants becomes feasible, if plant and process designs support it.

References

Backx, T., O. Bosgra, and W. Marquardt, Towards Intentional Dynamics in Supply Chain Conscious Process Operations, FOCAPO 98, Snowmass (1998). Revised version submitted to *J. Proc. Control*, July 2000.

Backx, T., O. Bosgra, and W. Marquardt, Integration of model predictive control and optimization of processes, In *Proceedings ADCHEM 2000*, volume 1, pages 249–260 (2000).

Backx, T., Combining first principles dynamic modeling and plant identification for design of an industrial MPC application to a polymer reactor, In *Proc. IFAC 14th World Congress*, Beijing, China (1999).

Clarke-Pringle, T. and J. F. MacGregor, Reduced dimension control: An overview of approaches for product quality control, In *Proc. ADCHEM 2000*, pages 189–194, Pisa, Itlay (2000).

Cutler, C. R. and B. L. Ramaker, Dynamic Matrix Control—

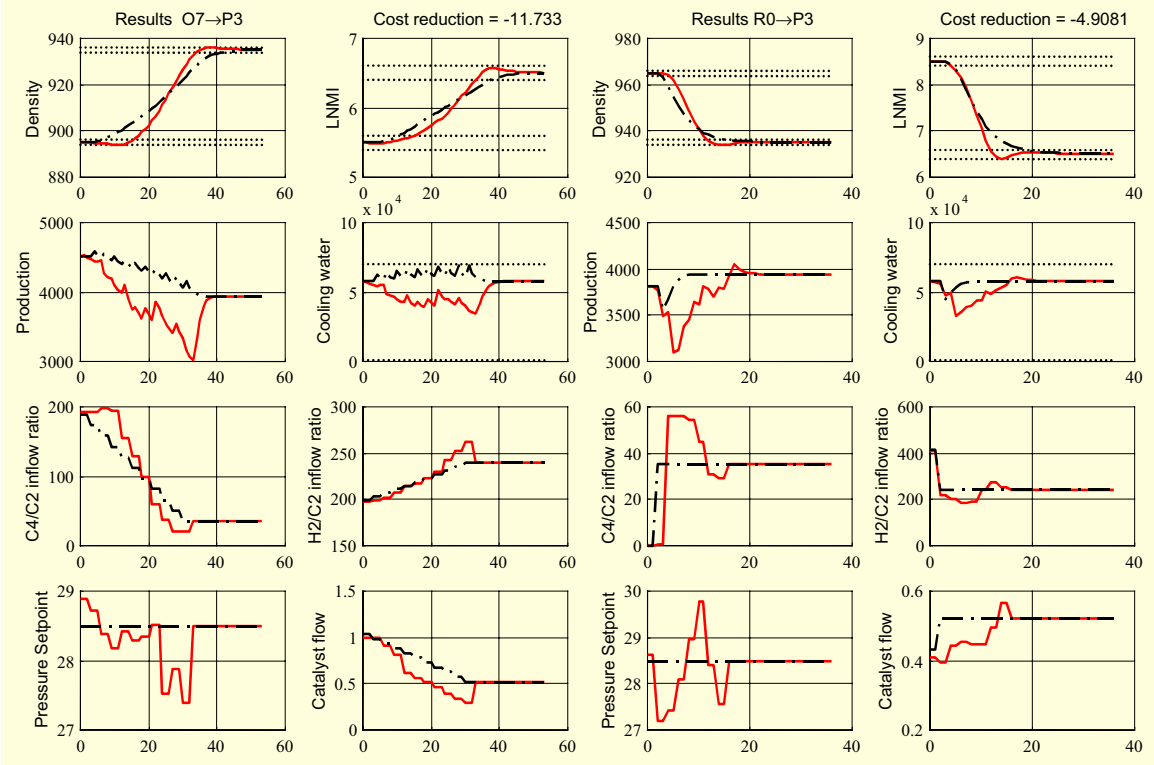


Figure 6: Process values and manipulated variables for two optimized grade changes. The dashed lines represent the initial trajectory, while the solid lines correspond to the optimized controlled trajectory.

A Computer Control Algorithm, In *Proc. of Joint Automatic Control Conference*, volume Paper WP5-B, San Francisco, CA (1980).

De Moor, B. and D. Berckmans, "Building a grey box model to model the energy and mass transfer in an imperfectly mixed fluid by using experimental data," *Math. Comp. Simul.*, **42**(2-3), 233–244 (1996).

Falkus, H. M., *Parametric uncertainty in system identification*, PhD thesis, Eindhoven University of Technology (1994).

Froisy, J. B., "Model predictive control: Past, present and future," *ISA Transactions*, **33**, 235–243 (1994).

García, C. E. and D. M. Prett, Design methodology based on the fundamental control problem formulation, In *The Shell Process Control Workshop*, pages 3–25. Butterworths, Stoneham, MA (1986).

Heuberger, P. S. C., P. M. J. van den Hof, and O. H. Bosgra, "A generalized orthonormal basis for linear dynamical systems," *IEEE Trans. Auto. Cont.*, **40**, 451–465 (1995).

Hjalmarsson, H., M. Gevers, and F. de Bruyne, "For model-based control design, closed-loop identification gives better performance," *Automatica*, **32**(12), 1659–1673 (1996).

Jacobsen, E. W., P. Lundstrom, and S. Skogestad, Modelling and Identification for robust control of ill-conditioned plants -A distillation case study, In *Proc. ACC*, pages 242–248 (1991).

Johansen, T. A. and B. A. Foss, "Semi-empirical modeling of non-linear dynamic systems through identification of operating regimes and local models," *Modelling, Identification and Control*, **16**(4), 213–232 (1995).

Johansen, T. A., "Identification of non-linear systems using empirical data and prior knowledge-An optimization approach," *Automatica*, **32**(3), 337–356 (1996).

Kemna, A. H. and D. A. Mellichamp, "Identification of combined physical and empirical models using non-linear a-priori knowledge," *Control Eng. Practice*, **3**(3), 375–382 (1995).

Lindskog, P. and L. Ljung, "Tools for semi-physical modeling," *Int. J. of Adaptive Control and Signal Processing*, **9**(6), 509–523 (1995).

Ljung, L. and T. Söderström, *Theory and practice of recursive identification*, Series in Signal Processing, Optimization and Control. MIT Press (1983).

Ljung, L., B. Wahlberg, and H. Hjalmarsson, Model quality: The role of prior knowledge and data information, In *Proc. 30th IEEE Conf. Decision and Control*, pages 273–278, Brighton, UK (1991).

Ljung, L., *System Identification: Theory for the User*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1987).

MacFarlane, A. G. J., *Frequency Response Methods in Control Sysems*. IEEE Press (1979).

Morari, M., Process Control Theory: Reflections on the past and goals for the next decade, In *The Second Shell Process Control Workshop*. Butterworths, Stoneham, MA (1988).

Muske, K. R. and J. B. Rawlings, "Model Predictive Control with Linear Models," *AIChE J.*, **39**(2), 262–287 (1993).

Papoulis, A., *Signal Analysis*, Electrical & Electronic Eng. Series. Mc-Graw Hill (1984).

Prett, D. M. and C. E. Garcìa, *Fundamental Process Control*. Butterworths, Boston (1988).

Qin, S. J. and T. A. Badgwell, An overview of industrial model predictive control technology, In Kantor, J. C., C. E. García, and B. Carnahan, editors, *Proceedings of Chemical Process Control—V*, pages 232–256. CACHE, AIChE (1997).

Richalet, J., 3rd Generation Predictive Control of Chemical Reactors, In Kantor, J. C., C. E. García, and B. Carnahan, editors,

Proceedings of Chemical Process Control—V. CACHE, AIChE (1997).

Rivera, D. E. and S. V. Gaikwad, "Systematic techniques for determining modeling requirements for SISO and MIMO feedback control," *J. Proc. Cont.*, **5**(4), 213–224 (1995).

Rosenbrock, H. H., *State-Space and Multivariable Theory*. Thomas Nelson and Sons, London (1970).

Skogestad, S., E. W. Jacobsen, and P. Lundstrom, Modeling requirements for robust control of distillation columns, In *Proc. 11th Triennial World Congress of IFAC*, volume 6, pages 191–197, Tallin (1991).

Tsen, A. Y.-D., S.-S. J. and D. S.-H. Wong, Real-time non-linear model based control of a complex chemical plant via parallel processing rigorous distillation model-based control, In *Proc. of the National Science Council, Republic of China, Part A*, volume 19, pages 199–209 (1995).

Tulleken, H. J. A. F., "Grey-box modeling and identification using physical knowledge and Bayesian techniques," *Automatica*, **29**(2), 285–308 (1993).

Van den Hof, P. M. J., P. S. C. H. and J. Bokor, "System identification with orthonormal basis functions," *Automatica*, **31**, 1821–1834 (1995).

Van der Schot, J. J., A. C. P. M. R. L. Tousain, Backx, and O. H. Bosgra, SSQP for the solution of large-scale dynamic-economic optimization problems, In *ESCAPE 1999*, Budapest, Hungary (1999).

Van der Schot, J. J., A practical approach to dynamic-economic optimization and control of industrial processes, Master's thesis, Delft University of Technology (1998).

Van Donkelaar, E. T., P. S. C. Heuberger, and P. M. J. van den Hof, Identification of a fluidized catalytic cracking unit: An orthonormal basis function approach, In *Proc. American Control Conference*, pages 1914–1917 (1998).

Van Overschee, P. and B. De Moor, "Subspace algorithms for the stochastic identification problem," *Automatica*, **29**, 649–660 (1993).

Van Overschee, P. and B. De Moor, "N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, **30**, 75–93 (1994).

Verhaegen, M. and P. Dewilde, "Subspace model identification. Part I: The output error state space model identification class of algorithms," *Int. J. Control*, **56**, 1187–1210 (1993a).

Verhaegen, M. and P. Dewilde, "Subspace model identification. Part II: Analysis of the elementary output-error state space model identification algorithm," *Int. J. Control*, **56**, 1211–1241 (1993b).

Viberg, M., "On Subspace-Based Methods for the Identification of Linear Time-Invariant Systems," *Automatica*, **31**, 1835–1852 (1995).

Wei, L. and J. H. Lee, "Control relevant identification of ill-conditioned systems: Estimation of gain directionality," *Comput. Chem. Eng.*, **20**(8), 1023–1042 (1996).

Willems, J. C., "From time series to linear system-Part I: Finite dimensional linear time-invariant systems," *Automatica*, **22**, 561–580 (1986a).

Willems, J. C., "From time series to linear system-Part II: Exact modeling," *Automatica*, **22**, 675–694 (1986b).

Willems, J. C., "From time series to linear system-Part III: Approximate modeling," *Automatica*, **23**, 87–115 (1987).

Yoon, S. and J. F. MacGregor, Relationship between statistical and causal model based approaches to fault detection and isolation, In *Proc. ADCHEM 2000*, volume I, pages 81–86, Pisa, Italy (2000).

Zhu, Y. and A. C. P. M. Backx, *Identification of Multivariable Industrial Processes*. Springer Verlag, Berlin (1993).

Zhu, Y., P. Van Overschee, and L. Ljung, Comparison of three classes of identification methods, In *Proc. 10th IFAC Symp. on System Identification*, pages 175–180, Copenhagen (1994).

Recent Advances and Challenges in Process Identification

Sten Bay Jørgensen*
CAPEC
Department of Chemical Engineering
Technical University of Denmark
Lyngby, Denmark

Jay H. Lee†
Department of Chemical Engineering
Georgia Institute of Technology
Atlanta, Georgia, USA

Abstract

Process identification is undergoing tremendous developments as computational capabilities improve. Theories are rapidly catching up with the needs of practical applications but practical process identification experiences still reveal significant gaps between theory and practice. This review attempts to highlight the present gaps and challenges. With this objective, the review treats recent progress in process identification with data gathered in closed loop, and in the tailoring of an entire identification process to a given control objective.

Keywords

Process identification, Model uncertainty, Robust control, Parameter estimation, Closed-loop identification, Identification for control, Multivariable systems, Nonlinear systems, Time-varying systems

Introduction

Process identification is concerned with using observed process data to form a model of the process that can be used in various constructive ways for process improvement. It is unquestionably one of the most important steps of control system design—accounting for as much as 80-90% of the cost and time involved. It encompasses a diverse set of tasks that include plant testing, selection of a model structure, parameter estimation, and model validation. Prior to actual controller implementation, it is the only step that requires direct interaction with the process. Consequently, any erroneous decision here can jeopardize success of an entire control project. This fact demands that all decisions involved in identifying a process model be made carefully and systematically based on sound scientific principles and methods—a fact that perhaps explains the good synergy between researchers and practitioners of the field.

The paper of Åström and Eykhoff (1971) was one of the first to review the research in system identification. During the 70s and 80s, the theoretical foundations for system identification were laid by the pioneering work of Ljung (1987), Söderström and Stoica (1989) and their coworkers. The work of Ljung centered around a particular paradigm called *Prediction Error Minimization (PEM)*, which today is the norm of industrial practice. Process identification at the beginning of the 90s was reviewed by Andersen et al. (1991) in CPC-IV.

The decade has seen several major developments. One of them is the *subspace approach*, which was motivated to overcome some drawbacks of PEM for multivariable system identification (Van Overschee and De Moor, 1996; Verhaegen and Dewilde, 1992; Larimore, 1990). Significant advances have also come along in closed-loop identification and "control-oriented" (or "control-relevant") identification. In these areas the goal is to tailor the whole identification procedure to a given control objective. The past decade has also seen an unprecedented range and volume of applications of process identification in industries, mainly to provide models for predictive control. After a decade of such explosive developments, it is indeed apt to reflect upon the progress and the state of the field at this CPC meeting.

The gap between research and practice, though narrower than in most fields, is nevertheless significant and therefore is worth elaborating a bit:

- *Plant Test:* Industrial plant tests use simple signals like steps or PRBSs. In addition, it is almost always limited to perturbing one input at a time, mostly out of the concern for unpredictable effects on process behavior (Qin and Badgewell, 1997). Literature is replete with optimal test signal design methods including those that attempt to incorporate specific control requirements and process characteristics into design in an iterative manner (Rivera et al., 1993; Asprey and Macchietto, 2000; Pearson, 1998; Cooley and Lee, 2001). However, such tailored and iterative designs have rarely been attempted in practice, if ever. The single-input testing will inevitably emphasize accuracy of individual SISO dynamics but several studies have shown that accurate identification of SISO dynamics may be inadequate for multivariable control of certain types of plants (*e.g.*, ill-conditioned plants) (Andersen et al., 1989; Koung and MacGregor, 1994; Li and Lee, 1996).
- *Model Structure:* Popular structures are Finite Impulse Response (FIR) models and ARX models, both of which lead to linear regression problems. Other structures, like ARMAX models, OE models, and Box-Jenkins models, which require non-convex optimization, are less common but are used in some occasions. In almost all cases, Multiple-Input-Single-Output (MISO) structures are used, in which a separately parameterized model is fitted for each output (Andersen et al., 1991). This practice

*sbj@kt.dtu.dk
†jay.lee@che.gatech.edu

is clearly inefficient, both with respect to model order and accuracy, in view of the fact that most industrial process outputs exhibit significant levels of cross-correlation. The preference for (or insistence on) MISO structure is clearly linked to ease of parameter estimation as explained below.

- *Model Estimation:* PEM is by far the most dominant method for estimating model parameters, perhaps owing to its flexibility and sound theoretical basis as well as the ready availability of software tools. However, with multivariable structures, PEM requires special parameterizations and nonconvex optimization (Ljung, 1987; Van Overschee and De Moor, 1994), a fact that perhaps explains the industry's proclivity toward use of MISO structures. The subspace approach is designed to obviate these problems but requires relatively large data sets. In fact, the two approaches are best combined into one: The subspace approach can be used to provide a good initialization for PEM, which should alleviate the aforementioned problems. However, extensive use of such MIMO identification methods in the industry is not at all evident.
  Statistical methods like the Maximum Likelihood Estimation or Bayesian estimation have not found much use. This is probably because of the lack of probability information or the fact that these complex methods often reduce to the same least squares calculation as PEM under commonly made probability assumptions. Likewise, use of nonparametric methods such as the frequency-domain Empirical Transfer Function Estimation seems rare, probably due to the lack of sufficiently large data sets.
- *Use of Estimated Model:* The PEM approach with many structures (e.g., ARX or ARMAX) as well as the subspace approach yield a combined deterministic-stochastic system model. However, the noise part of the model is seldom used in control system design. This may be due to the belief that disturbances experienced during an identification experiment are not representative of those encountered during real operation. However, this practice bears some danger as the two model parts are identified to work together as a *single predictor*. Also, many process monitoring and soft-sensing schemes require precisely information on how variables are correlated to one another in time due to unknown inputs and noises. Hence, the noise part of a multivariable model, when fitted with appropriate data, can serve a very useful function. In addition, some control applications, such as those involving inferential estimation, should clearly benefit from availability of an accurate noise model (Amirthalingam and Lee, 1998; Amirthalingam et al., 2000).
- *Model Validation:* Standard tests like residual analysis and cross validation are ubiquitous in the industrial practice. However, model uncertainties are seldom captured in a form that can be used for robust controller design. There is an extensive literature on how information on process noises (either statistics or bounds) can be translated into model bounds (Rivera et al., 1993; Ninness and Goodwin, 1995; Böling and Mäkilä., 1995). However, these methods are seldom used as process noise information itself is unavailable or highly inaccurate. In addition, there are many other contributing uncertainties, *e.g.*, process nonlinearities, actuator errors, etc. Good theories with wrong assumptions are just as ineffective as bad theories.
- *Closed-Loop Testing and Iterative Improvement:* Many industrial processes already have several working loops that may not be removed. Hence, we suspect that closed-loop testing has already been practiced to some extent. However, it is not clear whether practitioners are always aware of potential problems that can arise with usage of feedback correlated data. It is also less common to use closed-loop testing as a way to generate data that are highly informative for feedback controller design (Gevers and Ljung, 1986; Gevers, 2000; Forssell and Ljung, 1999a). In addition, many published papers deal with iterative improvement of model and controller through closed-loop testing (Van den Hof and Scharma, 1995; Hjalmarsson and Birkel, 1998). The idea of continually improving the closed-loop performance by using data collected from a working loop is very attractive from a practical viewpoint. It connects well with the industry's growing concern over maintaining performance of advanced controllers. However, it is not clear that the industry at large has seriously considered this possibility.
- *Nonlinear Process Identification:* While systematic tools for formulating first engineering principle models have begun to appear, few systematic methods for *identifying* first engineering principle models are available (Asprey and Macchietto, 2000; Lee, 2000). On the other hand, some implementations of nonlinear model predictive control have been reported (Qin and Badgewell, 1998; Young et al., 2001). The most common industrial approach to deal with process nonlinearities is by use of multiple models. Switching rules among different models are *ad hoc* and seldom systematically designed. Some applications of artificial neural networks are reported but their effectiveness as *causal* models, *i.e.*, as optimization and control would use them, is questionable at best. Despite the vigorous research in this area during the past decade, the field still lacks a basic framework and a unifying theoretical foundation (Johansen and Foss, 1995).

These gaps in Table 1 serve to motivate our selection of topics covered in this review paper. We have

Issue	Practice	Theory
Model Structure	SISO/MISO ARX or FIR	MIMO State Space
Parameter estimation	PEM (Least Squares)	Subspace and PEM
Noise info usage	Rarely used	Disturbance estimation, inferential control
Model validation	Residual analysis	Model error bounds
Plant Test	One input at a time	Simultaneous testing of multiple inputs
Closed Loop test	Little understood	Vigorously researched
Nonlinear Process Identification	Seldom used	Systematic tools lacking

Table 1: Summary of current gaps between practice and theory.

chosen topics, for which theories, in our view, have advanced to a point that some of the aforementioned gaps can be closed to bring significant benefits to the practice. For that, we concentrate mainly on two topics, control-oriented identification and closed-loop identification, which in our view have seen the most significant research and progress related to process control during the past decade. We attempt to highlight the progress as well as the current limitations, and point out obstacles an engineer may face in adopting the new approaches in practice.

We have chosen to leave out the subspace approach in our review, other than developments relevant to closed-loop identification, as the topic has already been well publicized. Current literature on this topic includes a book (Van Overschee and De Moor, 1996) and several review papers, e.g., Viberg (1995) and Shi and MacGregor (2000). We have also chosen to leave out a review on nonlinear process identification, for which only limited progress has occurred but further developments are very much needed. Marquardt (2001)in this conference gives a comprehensive coverage of this topic.

The rest of the paper is organized as follows. In Section 2, we review the current state of knowledge in *closed-loop identification*. The asymptotic behavior of various identification approaches are discussed but our focus remains on practical implications of the theories. In Section 3, we discuss the problem of tailoring the entire identification process to a given specific control objective. We discuss why this consideration naturally leads to iterative identification and review both open-loop and closed-loop strategies. In Section 4, we conclude.

Our main objective for writing this paper is to fuel an honest and substantive debate among researchers and practitioners on the current state of identification theories as related to potential closure of the existing gaps. We hope that such a debate will clarify the strengths and limitations of the existing theories and methodologies for practitioners. We also hope that the factors previously passed over by the researchers but must be accounted for theories to be practicable will also be brought out in the open.

Closed-Loop Identification

Introduction

During the past decade, interest in closed-loop identification by the community has risen noticeably. Closed-loop identification is motivated by the fact that many industrial processes have already in place one or more loops that cannot be removed for safety and/or economic reasons. Beyond this lies the attractive idea of being able to improve the closed-loop performance continually by making use of data being collected from a working loop. It has also been claimed that data collected from a closed-loop operation better reflect the actual situation in which the developed model will be used, and therefore could yield better overall results (Gevers and Ljung, 1986; Hjalmarsson et al., 1996; Gevers, 2000).

On the other hand, a closed-loop condition presents some additional complications for system identification. The fundamental problem is the correlation between the output error and the input through the feedback controller. Because of the correlation, many identification methods that are proven to work with open-loop data can fail. This is true for the prediction error approach as well as the subspace approach and nonparametric approaches like empirical transfer function estimation.

Awareness of the potential failings has engendered significant research efforts, which in turn have led to better understanding of the properties of the existing methods when used with closed-loop data as well as some remedies and special measures needed to circumvent the potential problems. We will try to give a concise and pragmatic summary of the recent developments, concentrating on the practical implications of the theoretical results. See Gustavsson et al. (1987); Van den Hof and Scharma (1995); Forssell and Ljung (1999a) for more comphensive surveys and formal disquisitions on the topic. We will focus on the prediction error approach (Ljung, 1987), which is the standard at the moment, but we will also point to some potential problems and remedies for the subspace approach at the end.

The closed-loop identification methods can be classified into three broad categories. In the *direct* approach, the feedback is largely ignored and the open-loop system is identified directly using measurements of the input and

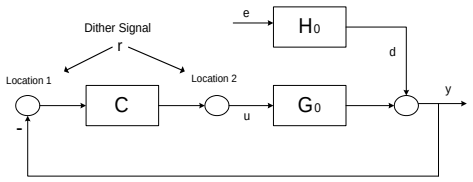


Figure 1: Typical setup for closed-loop identification.

output. In the *indirect* approach, the closed-loop transfer function between an external perturbation signal and the output (or input) is first identified and an open-loop system model is determined from it with knowledge of the controller. Finally, in the *joint input-output* approach, measurements of both the input and output are used to identify a joint system, which has the plant input and output as its outputs and the external perturbation signal as its input. From the joint system, an open-loop system model is extracted. We will examine each approach one at a time, concentrating on asymptotic properties and their implications.

The block diagram in Figure 1 displays a typical setup of closed-loop identification experiment. External perturbations may be added to the setpoint of the controller (at Location 1 in Figure 1) or directly to the controller output (at Location 2 in Figure 1). For the simplicity of exposition, we will assume from now on that external perturbations enter through Location 2. Note that, in the case that the controller is linear, this assumption can be made without loss of generality since perturbations introduced at Location 1 can always be rewritten as equivalent perturbations at Location 2.

Disturbances in the plant output are described collectively as a white noise signal (denoted by e) passed through some linear filter H_0. Without loss of generality, we will assume that H_0 is stably invertible and monic ($H_0(\infty) = I$) and the white noise signal e has the mean of zero and the covariance of P_e.

Assuming the plant G_0 and the controller C are both linear, we can write down the following closed-loop relationships:

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} \underbrace{(I + G_0C)^{-1}}_{S_0} G_0 & (I + G_0C)^{-1}H_0 \\ \underbrace{(1 + CG_0)^{-1}}_{S_0^r} & -C(I + G_0C)^{-1}H_0 \end{bmatrix} \begin{bmatrix} r(t) \\ e(t) \end{bmatrix} \tag{1}$$

In the above, S_0 and S_0^r represent the sensitivity function and reverse sensitivity function respectively.

Of course, both the assumption of linear plant and the particular way of describing the disturbance are great simplifications but are typical of developing and analyzing linear identification methods. In addition, the assumption of linearity of the controller, when made, may be untenable in many industrial situations where the controllers are equipped with various anti-windup and override features.

Direct Approach

In the direct approach, measurements of the plant input and output are used to build a model for the open-loop system directly, as in open-loop identification. The main advantage of the direct method is that the controller is not restricted to linear ones and its identity needs not be known. However, the correlation introduced by the feedback can cause problems and successful application demands some additional knowledge as we shall see.

The Method. The generic model structure used in linear identification is as follows:

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \tag{2}$$

θ is the vector of unknown parameters of the model and e is a white noise sequence. We may restrict the search to $\theta \in \Theta_M$. In this case, $\mathcal{G}_M \triangleq \{G(q, \theta) \mid \theta \in \Theta_M\}$ and $\mathcal{H}_M \triangleq \{H(q, \theta) \mid \theta \in \Theta_M\}$ represent the allowable set for plant model G and noise model H, respectively. Let the available data be denoted by

$$D^N = [y(1), u(1), \cdots, \cdots, y(N), u(N)] \tag{3}$$

The prediction error minimization (PEM) chooses $\hat{\theta}_N$, the estimate of θ based on D^N, according to

$$\hat{\theta}_N = \arg \min_{\theta \in \Theta_M} \left[V_N(\theta, D^N)\right] \tag{4}$$

where

$$V_N(\theta, D^N) = \frac{1}{N} \sum_{t=1}^{N} \|\varepsilon(\theta, t)\|_W \tag{5}$$

$$\varepsilon(\theta, t) = H^{-1}(q, \theta) \left[y(t) - G(q, \theta)u(t)\right] \tag{6}$$

and $\|x\|_W \triangleq x^T W x$ denotes the weighted quadratic norm.

Notation: We will often use symbols $\hat{G}_N$ and $\hat{H}_N$ to denote $G(q, \hat{\theta}_N)$ and $H(q, \hat{\theta}_N)$. The same symbols may also be used to compactly represent $G(e^{j\omega}, \hat{\theta}_N)$ and $H(e^{j\omega}, \hat{\theta}_N)$ when the context makes their meanings clear. Similarly, we will use G_θ and H_θ to denote $G(q, \theta)$ and $H(q, \theta)$, or sometimes $G(e^{j\omega}, \theta)$ and $H(e^{j\omega}, \theta)$.

Convergence Behavior. If all the signals are quasi-stationary, $V_N(\theta, D^N) \to \bar{V}$ and $\hat{\theta}_N \to \bar{\theta}$ w.p. 1 as $N \to \infty$, where

$$\bar{V}(\theta) \triangleq \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} E\{\|\varepsilon(\theta, t)\|_W\} \tag{7}$$

and

$$\bar{\theta} \triangleq \arg\min_{\theta\in\Theta_M} \bar{V}(\theta) \tag{8}$$

From Parseval's relation,

$$\bar{\theta} = \arg\min_{\theta\in\Theta_M} \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{tr}\left[W\Phi_\varepsilon\right] d\omega \tag{9}$$

where Φ_ε is the spectrum of the prediction error ε.

Since from (6),

$$\varepsilon(t) = H_\theta^{-1}\left[G_0u(t) + H_0e(t) - G_\theta u(t) - H_\theta e(t)\right] + e(t) \tag{10}$$

and $e(t)$ is independent of $(G_0 - G_\theta)u(t)$ and $(H_0 - H_\theta)e(t)$ because $(G_0 - G_\theta)$ and $(H_0 - H_\theta)$ both contain at least one delay,

$$\Phi_\varepsilon = H_\theta^{-1}\Delta_{GH}H_\theta^{-1*} + P_e \tag{11}$$

where

$$\Delta_{GH} = \begin{bmatrix}(G_0 - G_\theta) & (H_0 - H_\theta)\end{bmatrix}\begin{bmatrix}\Phi_u & \Phi_{ue}\\ \Phi_{eu} & P_e\end{bmatrix}\begin{bmatrix}(G_0^* - G_\theta^*)\\ (H_0^* - H_\theta^*)\end{bmatrix} \tag{12}$$

Hence,

$$\bar{\theta} = \arg\min_{\theta\in\Theta_M}\int_{-\pi}^{\pi} \mathrm{tr}\left\{\Delta_{GH}H_\theta^{-1*}WH_\theta^{-1}\right\} d\omega \tag{13}$$

The expression in (13) provides some good insights into the convergence behavior under the direct approach.

- Suppose $G_0 \in \mathcal{G}_M$ and $H_0 \in \mathcal{H}_M$ (i.e., the true plant and noise process both lie inside the parameterized model sets). Then, it is clear from the objective function that the minimum corresponds to $G_{\bar{\theta}} = G_0$ and $H_{\bar{\theta}} = H_0$ (implying consistent estimation) if

$$\Phi_x = \begin{bmatrix}\Phi_u & \Phi_{ue}\\ \Phi_{eu} & P_e\end{bmatrix} > 0 \qquad \forall\omega.$$

- In the case of a linear controller,

$$\begin{aligned}&\begin{bmatrix}\Phi_u & \Phi_{ue}\\ \Phi_{eu} & P_e\end{bmatrix}\\ &= \begin{bmatrix}\underbrace{S_0^r\Phi_r S_0^{r*}}_{\Phi_u^r} + \underbrace{CS_0H_0P_eH_0^*S_0^*C^*}_{\Phi_u^e} & \underbrace{-CS_0H_0P_e}_{\Phi_{ue}}\\ \underbrace{-P_eH_0^*S_0^*C^*}_{\Phi_{eu}} & P_e\end{bmatrix}\\ &= \begin{bmatrix}\Phi_u^r + \Phi_{ue}P_e^{-1}\Phi_{eu} & \Phi_{ue}\\ \Phi_{eu} & P_e\end{bmatrix}\\ &= \begin{bmatrix}I & \Phi_{ue}P_e^{-1}\\ 0 & I\end{bmatrix}\begin{bmatrix}\Phi_u^r & 0\\ 0 & P_e\end{bmatrix}\begin{bmatrix}I & 0\\ P_e^{-1}\Phi_{eu} & I\end{bmatrix}\end{aligned} \tag{14}$$

  Note that Φ_u is expressed as sum of two components, Φ_u^r representing the contribution from external dithering and Φ_u^e the contribution from noise feedback. From (14), it is clear that $\Phi_x > 0$ iff $\Phi_u^r > 0$.

- If the controller is nonlinear or time-varying, Φ_x may be positive definite even without $\Phi_r > 0$.
- Suppose $G_0 \in \mathcal{G}_M$ but $H_0 \notin \mathcal{H}_M$ (implying the noise part is undermodelled). Then, from (13), we see that $\hat{G}_N$ does not converge to G_0 as $G_\theta = G_0$ does not achieve the minimum of (13) in this case.
- When the noise model is completely fixed a priori (as in an *Output Error* structure), $\hat{G}_N$ does not converge to G_0 in general. From expression (13), we see that $\hat{G}_N \to G_0$ only when the assumed noise model is perfect or $\Phi_{ue} = 0$, which corresponds to the open-loop case.

Error Analysis. The error behavior in the limit can be formalized as follows.

Bias

Bias refers to the expected error $E\{\theta_0 - \hat{\theta}_N\}$. In terms of our notations, it is $\theta_0 - \bar{\theta}$ (or $G_0 - G_{\bar{\theta}}$ in terms of frequency-domain transfer function). An expression for bias in the limit can be obtained by further manipulating equation (13) into the following expression (Forssell and Ljung, 1999b).

$$\begin{aligned}\bar{\theta} = \arg\min_{\theta\in\Theta_M}\int_{-\pi}^{\pi} \mathrm{tr}\{&[(G_0 + E_\theta - G_\theta)\Phi_u(G_0 + E_\theta - G_\theta)^*\\ &+ (H_0 - H_\theta)(P_e - \Phi_{eu}\Phi_u^{-1}\Phi_{ue})\\ &(H_0 - H_\theta)^*]\,H_\theta^{-1*}WH_\theta^{-1}\}\,d\omega\end{aligned} \tag{15}$$

where

$$E_\theta = (H_0 - H_\theta)\Phi_{eu}\Phi_u^{-1} \tag{16}$$

From (15), we can conclude the followings:

1. If the parameterization of the plant/noise model is flexible enough that $G_\theta \in \mathcal{G}_M$ and $H_\theta \in \mathcal{H}_M$,

   $$G_{\bar{\theta}} = G_0 \text{ and } H_{\bar{\theta}} = H_0$$

   provided that the minimum is unique, which is guaranteed by $\Phi_u^r > 0$. In this case, $\hat{G}_N$ is an unbiased estimate of G_0.
2. Suppose noise model is fixed a priori as H_M, which does not have any dependence on θ. (Alternatively, assume that separate sets of parameters are used for G and H as in Box-Jenkins model). Also suppose that $G_\theta \in \mathcal{G}_M$. Then, we can conclude from the above that

   $$(G_0 - G_{\bar{\theta}}) = (H - H_M)\Phi_{eu}\Phi_u^{-1} \tag{17}$$

   Hence, $(H - H_M)\Phi_{eu}\Phi_u^{-1}$ is the bias. The bias will be zero if one or both of the following conditions is satisfied.

   - The assumed noise model is perfect, i.e., $H_M = H_0$.

- $\Phi_{eu} = 0$, which implies open-loop testing.

We also note that the size of bias depends on the following two things:

- $(H - H_M)$, which is the error in the assumed noise model.
- $\Phi_{eu}\Phi_u^{-1}$, which is affected by, among many things, the power of the external perturbation signal relative to the noise signal. Note that

$$\Phi_{eu}\Phi_u^{-1} = (P_e\Phi_u^{-1}) \times (\Phi_u^e\Phi_u^{-1}) \tag{18}$$

$P_e\Phi_u^{-1}$ is the noise-to-signal ratio in an open-loop sense. What multiplies this is $\Phi_u^e\Phi_u^{-1}$, which can be interpreted as the relative contribution of the noise feedback to the total input power. The larger the noise feedback's contribution, the bigger the bias. Note that increasing the controller gain will decrease $P_e\Phi_u^{-1}$ but will also increase $\Phi_u^e\Phi_u^{-1}$, thus making its effect on the bias inconclusive.

3. In the case of undermodeling such that $G_\theta \notin \mathcal{G}_M$ and/or $H_\theta \notin \mathcal{H}_M$, θ will be chosen to make both $G_0 - G_\theta$ and $H_0 - H_\theta$ small. Bias in the frequency domain will be distributed according to the weightings given in (15).

Variance

The other part of error is variance, which refers to the error due to an insufficient number of data points relative to the number of parameters. This error is mathematically formalized as

$$\mathrm{Cov}(\hat{\theta}_N) \triangleq E\left\{\left(\hat{\theta}_N - \bar{\theta}\right)\left(\hat{\theta}_N - \bar{\theta}\right)^T\right\} \tag{19}$$

in the parameter domain and

$$\mathrm{Cov}\left(\mathrm{vec}\hat{G}_N\right) \triangleq E\left\{\left(\mathrm{vec}\hat{G}_N - \mathrm{vec}G_{\bar{\theta}}\right)\left(\mathrm{vec}\hat{G}_N - \mathrm{vec}G_{\bar{\theta}}\right)^T\right\} \tag{20}$$

in the frequency domain, where the notation $\mathrm{vec}(\cdot)$ refers to a vectorized form of a matrix obtained by stacking the columns of the matrix into a single column.

In Zhu. (1989), it is shown for open-loop identification that, as $n \to \infty$ and $N \to \infty$,

$$\mathrm{Cov}(\mathrm{vec}\hat{G}_N) \sim \frac{n}{N}(\Phi_u)^{-T} \otimes \Phi_d \tag{21}$$

where

$$\Phi_d = H_0 P_e H_0^* \tag{22}$$

and $\otimes$ denotes the Kronecker product. The above expression shows that the asymptotic variance distribution in the frequency domain is shaped by the signal-to-noise ratio.

For a closed-loop system, it follows directly (Forssell and Ljung, 1999b) that

$$\mathrm{Cov}(\mathrm{vec}\hat{G}_N) \sim \frac{n}{N}(\Phi_u^r)^{-T} \otimes \Phi_d \tag{23}$$

Recall that $\Phi_u = \Phi_u^r + \Phi_u^e$. Thus the above says that the excitation contributed by noise feedback does not contribute to variance reduction, at least in the asymptotic case of $n \to \infty$. One can make some sense of this result by considering the extreme case that input excitation is entirely due to noise feedback. In this case, measurements of output and input contain just the information about closed-loop transfer functions $(I + G_0C)^{-1}H_0$ and $C(I + G_0C)^{-1}H_0$. If the model order is allowed to approach infinity, even perfect knowledge of the closed-loop functions yields no information about G_0 and H_0 independently. This is because, for an arbitrary choice of H_0, one can always choose G_0 to match any given closed-loop functions and vice versa (barring an invertibility problem). In such a case, variance would be infinite at all frequencies, which is consistent with expression given in (23).

An exceptional case is when noise model is perfectly known a priori. In this case, one can show that asymptotic variance distribution follows the open-loop case. Hence, noise feedback contributes just as much as external dithering to variance reduction. This is consistent with the foregoing argument as the perfect knowledge of H_0 would enable direct translation of information about the closed-loop functions into that about the open-loop function.

In situations where model order can be constrained to a finite number, the noise feedback would make some contribution to the information content for the estimation of open-loop functions G and H.

Notation: The Kronecker product between $A \in \Re^{n\times m}$ and $B \in \Re^{p\times r}$ is defined as

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \cdots & \cdots & a_{1,m}B \\ a_{1,2}B & \ddots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1}B & \cdots & \cdots & a_{n,m}B \end{bmatrix} \tag{24}$$

Practical Implications. The following are the important points to take away from the foregoing analysis.

- The main advantage of the direct approach is that the controller is completely taken out of the picture in the estimation step. Not only is it unnecessary to know the controller, but it is also not required for the controller to be linear and time-invariant. It is an important advantage considering that most industrial controllers are not adequately represented by a linear, time-invariant operator.

- The most serious problem for the direct approach is that noise model (structure) needs to be perfect in order for it to yield consistent estimates. Fixing noise model completely a priori, as is often done in practice, results in a biased-esimate of G_0 in general. For example, the popular least squares identification of finite impulse response parameters would give a biased model. This is the most important difference from the open-loop identification case.

- The theoretical result indicates that consistent estimates can be obtained by leaving noise model sufficiently flexible (so that $H_\theta \in \mathcal{H}_M$). Though this would indeed help reduce bias, it may not get rid of the bias problem completely in practice as most disturbances in industrial processes are non-stationary and are not accurately captured by a white noise through a linear filter. In addition, increasing the order of model would mean increased variance, thus demanding heavier external dithering.

- In case that noise part is undermodelled, the size of the resulting bias in G depends on the signal-to-noise ratio and the relative contribution of external dithering to the total input power (compared to noise feedback). Hence, in principle, *bias can be made negligible by overwhelming noise feedback with heavy external dithering.* On the other hand, boosting the signal-to-noise ratio by increasing the controller gain has no ameliorating effect on bias as it increases the relative contribution of noise feedback to the input power as well the signal-to-noise ratio .

- External dithering with a persistently exciting signal is necessary for consistent estimation. It is generally not sufficient to have an input that is persistently exciting. The requirement for external dithering may be removed by using a nonlinear controller or a time-varying controller.

- Asymptotic variance of a frequency transfer function estimate is shaped by input excitation *achieved through external dithering.* The excitation through noise feedback contributes little to variance reduction when model order is high. This means relying solely on noise feedback for input excitation can lead a very bad result (an estimate of infinite variance). Exception is the case where the noise model is accurately known a priori. However, such a case would be rare in practice, and any error in the a priori fixed noise model could translate into a bias in the estimate for G_0.

- External dithering is motivated from both the viewpoint of bias and of variance. If the noise structure is left very flexible, variance would be a significant problem without external dithering, no matter how much input excitation is there through noise feedback. On the other hand, if the noise structure is very restricted (for example, fixed completely), external dithering would be needed to reduce bias.

- Choice of location for dithering (between Location 1 and Location 2) is not important as a perturbation made at one location can always be translated into an equivalent perturbation at the other location. On the other hand, the perturbation signal (*e.g.*, its spectrum) should be designed accordingly. To see this, let us compare the expressions for the resulting input spectrum (Φ_u^r), which appears in the asymptotic variance expressions of (23), under the two dithering strategies. For dithering at Location 1, we obtain

$$\Phi_u^r = CS_0\Phi_r S_0^* C^* \tag{25}$$

  Dithering at Location 2 gives

$$\Phi_u^r = S_0^r\Phi_r S_0^{r*} \tag{26}$$

  Loops with integral controllers yield S_0 (or S_0^r) that starts from 0 at $\omega = 0$ and increases to 1 at high frequencies. Hence, a white noise perturbation signal at Location 2 would give Φ_u^r that is zero at $\omega = 0$ and very low in the frequency region well below the controller's bandwidth. In view of (23), this would make the low frequency part of the model very poor. Estimation of the low frequency dynamics is exacerbated by the typical shape of noise spectrum Φ_v, which is high in the low frequency region for most process control problems. On the other hand, white noise dithering at Location 1 does not suffer from this problem to a same degree since $CS_0 \approx G_0^{-1}$ in the low frequency region and $\approx I$ in the high frequency region. However, white noise dithering signal is generally not optimal, regardless of the location, and its spectrum could be designed systematically based on estimated noise spectrum and desired variance distribution.

- In many practical situations, disturbances process sees during an identification experiment are non-stationary and better represented by a model that contains integrators such as the one shown below:

$$y = G(q)u + H(q)\frac{1}{1-q^{-1}}e$$
$$\Rightarrow \quad \Delta y = G(q)\Delta u + H(q)e \tag{27}$$

  From the right-hand-side of the arrow in the above, we see that differencing the input and output data prior to applying the PEM makes this case equivalent to the standard case. All the foregoing discussions regarding the signal spectra hold with respect to the differenced signals. For example, to distribute the variance fairly evenly across the frequency, one should use a dithering strategy where

an *integrated white noise* signal (or an integrated PRBS) is added to the controller's setpoint. This would make Δr a white noise signal in our analysis. On the other hand, differencing of noisy data can amplify the noise effect and make the identification more difficult.

Indirect Approach

In the indirect approach, measurements of plant output (or input) is used along with record of an external dither signal to build a model for the closed-loop system first. Then, based on knowledge of the controller, an estimate for open-loop function is extracted from the estimate of the closed-loop function. Note that

$$y(t) = \underbrace{(I + G_0C)^{-1}G_0}_{T_0^{yr}} r(t) + \underbrace{(I + G_0C)^{-1}H_0}_{T_0^{ye}} e(t) \quad (28)$$

Hence, in the first step, T_0^{yr} is estimated using data record for r and y, and in the second step, G_0 is extracted from the estimate.

The main advantage of the indirect approach is that the first step is in essence the same problem as open-loop identification because one does not have to be concerned with any feedback-induced correlation between the system input (r) and the noise (e). This removes the requirement of perfect noise model (structure) for consistent estimation. On the other hand, the second step requires a mathematical representation of the controller, which generally has to be assumed linear in order to extract the open-loop functions from the closed-loop function. In addition, the resulting model can be of very high order, depending on the parameterization used.

The Method. The generic model structure used here is

$$y(t) = T^{yr}(q, \theta)r(t) + T^{ye}(q, \theta)e(t) \quad (29)$$

The same PEM can be applied to the above model structure with data record of

$$D^N = [y(1), r(1), \cdots, \cdots, y(N), r(N)]$$

to obtain estimates $\hat{T}_N^{yr}$ and $\hat{T}_N^{ye}$. Since $T_0^{yr} = (I + G_0C)^{-1}G_0$, we can obtain an estimate for open-loop function G_0 as follows:

$$\hat{G}_N = \hat{T}_N^{yr}(I - \hat{T}_N^{yr}C)^{-1} \quad (30)$$

The above calculation can result in $\hat{G}_N$ of very high order. One may obviate this problem by using the following parameterization of the closed-loop function:

$$T^{yr}(q, \theta) = (I + G(q, \theta)C)^{-1}G(q, \theta) \quad (31)$$

However, adoption of such a parameterization would make parameter estimation (via PEM) more complex.

Another concern may be that resulting $\hat{G}_N$ may not even give a stable closed loop under the controller C. To ensure that the model is stabilized by the controller, one can first parameterize the set of all linear plants that are stabilized by C using the dual Youla parameterization and then perform the search over the set (de Callafon and Van den Hof, 1996).

By applying the expression of (13) to the indirect identification, convergence behavior can be easily analyzed. It is not discussed here since the problem is essentially the same as the open-loop case. Error analysis can also be carried out in a similar manner as before. Asymptotic variance of $\hat{G}_N$ under the indirect method is the same as in the direct method and follows (23).

Key Points.

- To obtain the closed-loop function, one can use any proven open-loop identification method without modification.
- Consistent estimation is possible even without a perfect noise model (structure), which is the main attractive point for the indrect approach.
- The most serious problem for the indirect approach is the assumption of linear controller, which may be untenable in many industrial situations. Most industrial controllers are equipped with special anti-windup/override features and their behavior may not be represented accurately by a single linear function. One needs to be careful that these special features do not become active during data collection. An error in the controller representation will translate into an error in extracting the open-loop function from the closed-loop function.
- Another potential disadvantage is high model order that often results from the two step calculation. This problem may be obviated by employing a particular parameterization involving a parameterized form of the open-loop function (such as the one in (31)). On the other hand, with such a specialized structure, the prediction error minimization becomes more demanding computationally.
- The asymptotic variance distribution in the frequency domain follows the same expression as in the direct identification case. Hence, all the previous remarks regarding shaping of variance distribution apply here as well.

Joint Input-Output Approach

The idea behind the joint input-output approach is to use measurements of both the output and input to remove the requirement of a known controller. Since

$$y(t) = \underbrace{G_0(I + CG_0)^{-1}}_{T_0^{yr}} r(t) + \underbrace{(I + G_0C)^{-1}H_0}_{T_0^{ye}} e(t) \quad (32)$$

and

$$u(t) = \underbrace{(I + CG_0)^{-1}}_{T_0^{ur}} r(t) + \underbrace{-C(I + G_0C)^{-1}H_0}_{T_0^{ue}} e(t) \quad (33)$$

we have

$$G_0 = T_0^{yr} \left(T_0^{ur}\right)^{-1} \quad (34)$$

The above relation can be used to calculate the open-loop function from the two closed-loop functions without knowledge of the controller.

In effect, in the joint input-output approach, the controller is "identified." Even though the controller is not required to be known, it is implicitly assumed to be linear and time-invariant, an assumption that may not always be justified in practical situations.

The Method. In the first step, the following model form is used:

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = \begin{bmatrix} T^{yr}(q,\theta) \\ T^{ur}(q,\theta) \end{bmatrix} r(t) + \begin{bmatrix} T_M^{ye}(q) \\ T_M^{ue}(q) \end{bmatrix} e(t) \quad (35)$$

In the above, we assumed that the noise part of the model is fixed a priori but it too can be parameterized for estimation. Note that consistent estimates can be obtained with a fixed noise model in this case since r and e are uncorrelated.

To the above, PEM can be applied with the data record of

$$D^N = [y(1), u(1), r(1), \cdots\cdots\cdots, y(N), u(N), r(N)]$$

to obtain estimates $\hat{T}_N^{yr}$ and $\hat{T}_N^{ur}$. Then open-loop plant estimate $\hat{G}^N$ is obtained by

$$\hat{G}_N = \hat{T}_N^{yr} \left(\hat{T}_N^{ur}\right)^{-1} \quad (36)$$

One drawback is that $\hat{G}_N$ so obtained may be of very high order. This can be circumvented by using a model form that recognizes the underlying structure, such as the one below:

$$T^{yr}(q,\theta) = G(q,\theta_1)T^{ur}(q,\theta_2) \quad (37)$$

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

By adopting the above structure, $G(q,\theta_1)$ is identified directly, rather than through the inversion in (36).

Two-Stage Method and Projection Method. Among the variations of the above method are the so called *two-stage method* by Van den Hof and Schrama. (1993) and *projection method* by Forssell and Ljung (2000a). Note that

$$u(t) = \underbrace{(I + CG_0)^{-1}r(t)}_{u^r(t)} + \underbrace{-C(I + G_0C)^{-1}H_0e(t)}_{u^e(t)} \quad (38)$$

Hence, u_r represents the portion of the input generated by signal r and u^e is the portion due to noise feedback. Now,

$$y(t) = G_0u^r(t) + G_0u^e(t) + H_0e(t) \quad (39)$$

Since u^r is uncorrelated with the rest of the right-hand-side, a consistent estimate of G_0 can be obtained with data for y and u^r. This consideration leads to the following method:

1. Start with the parameterized structure

$$u(t) = T^{ur}(q,\theta)r(t) + T_M^{ue}(q)e(t) \quad (40)$$

   Apply PEM to the above model with data record D^N and obtain $\hat{T}_N^{ur}$.

2. Obtain data for u^r through

$$\hat{u}_N^r(t) = \hat{T}_N^{ur}r(t), t = 1, \cdots, N \quad (41)$$

3. Apply PEM to the model structure

$$y(t) = G(q,\theta)u^r(t) + T_M^{ye}(q)e(t) \quad (42)$$

   using the data record of

$$D^N = [\hat{u}_N^r(1), y(1), \cdots, \cdots, \hat{u}_N^r(N), y(N)]$$

   to obtain $\hat{G}_N$.

In the projection method, T^{ur} is allowed to be an *acausal* operator by parameterizing it as a two-sided FIR filter:

$$T^{ur}(q,\theta) = \sum_{i=-n_1}^{n_2} \theta_i q^{-i} \quad (43)$$

This is to ensure that the resulting $\hat{u}_N^r$ is uncorrelated ("orthogonal" in the least squares language) asymptotically with $u - \hat{u}_N^r$, which becomes a part of the residual in the second PEM. Recall that, in the least squares estimation, the residual has to be orthogonal to the regressor in order to obtain a consistent estimate (Ljung, 1987). If the controller is linear, all the preceding argument holds and the orthogonality of the residual can be assured (asymptotically) with a sufficiently rich yet causal T_θ^{ur}. However, with a nonlinear or time-varying controller, $\hat{u}_N^r$ obtained with any casual $\hat{T}_N^{ur}$ and $u - \hat{u}_N^r$ may be correlated, thereby destroying the consistency of estimation in the final step. By employing an acausual FIR filter with sufficiently large n_1 and n_2 for $T^{ur}(q,\theta)$, one gets $\hat{u}_N^r$ that is uncorrelated (orthogonal) with $u - \hat{u}_N^r$, even when the controller is nonlinear or time-varying.

The convergence behavior is essentially the same as for the indirect approach and does not require an elaborate discussion here. Error behavior too is similar to that for the indirect approach. The variance also follows the same expression of (23). The same comments apply to the two-step method and the projection method.

Closed-Loop method	Perfect Noise Model	External Dithering	Linear Controller	Known Controller
Direct	Yes	Yes*	No	No
Indirect	No	Yes	Yes	Yes
Joint I/O	No	Yes	Yes	No
Two-Step	No	Yes	Yes	No
Projection	No	Yes	No	No

*Unless the noise model is perfectly known *a priori*

Table 2: Summary of the requirements of different closed-loop identification methods.

Key Points. Some key points for the joint I/O method and its off-springs are:

- As with the indirect approach, the main advantage of the joint input-output approach over the direct approach is that consistent estimates can be obtained even with an imperfect noise model.
- The main advantage of the joint input-output approach over the indirect approach is that explicit knowledge of the controller is not required. However, it does implicitly assume that the controller is linear.
- The two step method is essentially same as the standard joint input-output method and does not appear to offer any new advantage.
- The projection method, on the other hand, further improves it by removing the requirement of linear controller for consistent estimation. It also retains the aforementioned advantage over the direct approach. The relaxation of the requirement for linear controller behavior is practically significant in view of the fact that most industrial controllers show some degree of nonlinear behavior due to various fixes and add-ons.
- Measurement errors for the input do not destroy the consistency if they are uncorrelated with the dither signal.

The practical requirements of the different closed-loop identification approaches are summarized in table 2.

Subspace Identification with Closed-Loop Data

So far, our discussion has centered around PEM. An alternative to PEM is subspace identification, which has drawn much attention in recent years. The main attraction for the subspace approach is that it yields a *multivariable system model* without the need for a special parameterization, which requires significant prior knowledge and nonconvex optimization. Both represent significant hurdles for using PEM for multivariable system identification, which explains why it is hardly used for this purpose. Most subspace methods in the literature can fail, however, when used with closed-loop data. By "fail", we mean that the guarantee of an asymptotically unbiased estimate, a nice proven property for most subspace methods, is lost. A practical implication is that, with closed-loop data, the method may yield a poor model regardless of number of data points used. Here, we will briefly examine what fundamental problem closed-loop data pose for the subspace method and examine some available modifications to the standard approach in order to circumvent the problem.

Main Idea of Subspace Identification. We first review the subspace method. Though many versions exist in the literature, the essential ideas are same and the practical outcomes from applying different algorithms should not differ much (Van Overschee and De Moor, 1995). We will discuss one of the most popular methods, called N4SID, which was introduced by Van Overschee and De Moor (1995).

The underlying plant is assumed to be

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t) + \nu(t) \end{aligned} \tag{44}$$

where $w(t)$ and $\nu(t)$ are white noise processes. The state-space description is very general and subsumes most of the input-output structures studied in the system identification literature.

The following is an alternative representation of the above when (44) is viewed as an input-output system description:

$$\begin{aligned} x_\infty(t+1) &= Ax_\infty(t) + Bu(t) + K_\infty\varepsilon_\infty(t) \\ y(t) &= Cx_\infty(t) + Du(t) + \varepsilon_\infty(t) \end{aligned} \tag{45}$$

(45) can be interpreted as the steady-state Kalman filter for (44) and hence ε_∞ is the innovation sequence, which is white. The two are equivalent in an input-output sense and (45) is referred to as the innovation form for (44).

The N4SID method attempts to identify the following *non-steady-state Kalman filter* equation (within some state coordinate transformation) to obtain the parameters for the system equation in (45):

$$\begin{aligned} x_{n+1}(t+1) &= Ax_n(t) + Bu(t) + K_n\varepsilon_n(t) \\ y(t) &= Cx_n(t) + Du(t) + \varepsilon_n(t) \end{aligned} \tag{46}$$

From the notation, you may deduce that the above represents a non-steady-state Kalman filter started at $t-n$ with zero initial estimate and initial covariance set equal to the system's open-loop covariance. Successful identification of the above equation yields system matrices (A, B, C, D) for (45). In addition, with a large n, it gives a good approximation for stochastic part of the system in (45) since $K_n \to K_\infty$ and $\mathrm{Cov}\{\varepsilon_n\} \to \mathrm{Cov}\{\varepsilon_\infty\}$ as $n \to \infty$. Here n is assumed to be higher than the intrinsic system order.

For the simplicity of discussion, we will assume from hereafter that $u(t)$ is an independent (temporally uncorrelated) sequence.

The key to subspace identification is the following multi-step prediction equation:

$$\begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+n-1) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x_n(t) + \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ CA^{n-2}B & \cdots & CB & D \end{bmatrix} \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+n-1) \end{bmatrix} + \begin{bmatrix} \varepsilon_n(t|t-1) \\ \varepsilon_n(t+1|t-1) \\ \vdots \\ \varepsilon_n(t+n-1|t-1) \end{bmatrix} \quad (47)$$

$\varepsilon_n(t+j|t-1)$ represents the $j+1$-step-ahead prediction error (based on the Kalman estimate $x_n(t)$). The above will be denoted by

$$\mathcal{Y}_n^{0+}(t) = \Gamma_n^o x_n(t) + H_n^f U_n^{0+}(t) + \mathcal{E}_n^{0+}(t) \quad (48)$$

A key point is that $\mathcal{E}_n^{0+}(t)$ is orthogonal to both $x_n(t)$ and $U_n^{0+}(t)$. Now, since the Kalman filter is linear with respect to the measurement and the input, we can express

$$x_n(t) = M_1 \mathcal{Y}_n^-(t) + M_2 \mathcal{U}_n^-(t) \quad (49)$$

where

$$\mathcal{Y}_n^-(t) = \begin{bmatrix} y(t-n) \\ y(t-n+1) \\ \vdots \\ y(t-1) \end{bmatrix}; \qquad \mathcal{U}_n^-(t) = \begin{bmatrix} u(t-n) \\ u(t-n+1) \\ \vdots \\ u(t-1) \end{bmatrix} \quad (50)$$

Substituting (49) into (48) gives

$$\mathcal{Y}_n^{0+}(t) = \underbrace{\Gamma_n^o \begin{bmatrix} M_1 & M_2 \end{bmatrix}}_{H_n^p} \begin{bmatrix} \mathcal{Y}_n^-(t) \\ \mathcal{U}_n^-(t) \end{bmatrix} + H_n^f U_n^{0+}(t) + \mathcal{E}_n^{0+}(t) \quad (51)$$

By arranging the data for y and u appropriately based on the above equation, one can obtain estimates for H_n^p and H_n^f through linear least squares. The least squares estimation yields consistent estimates of H_n^p and H_n^f as residual $\mathcal{E}_n^{0+}(t)$ is orthogonal to regressors $\mathcal{Y}_n^-(t), \mathcal{U}_n^-(t)$ and $U_n^{0+}(t)$.

Using estimate $\hat{H}_n^p$, one can determine the system order (by examining its rank) and obtain estimates for Γ_n^o,M_1 and M_2 (within a coordinate transformation). Finally, data for the Kalman state x_n can be constructed from the estimates as

$$\hat{x}_n(t) = \begin{bmatrix} \hat{M}_1 & \hat{M}_2 \end{bmatrix} \begin{bmatrix} \mathcal{Y}_n^-(t) \\ \mathcal{U}_n^-(t) \end{bmatrix}, \qquad t = n, \cdots, N-n \quad (52)$$

The whole calculation can be done by performing some matrix projections with appropriately arranged data matrices, which can be implemented in a computationally efficient and robust way (Van Overschee and De Moor, 1996).

By following a similar procedure, one can also obtain data for $x_{n+1}(t+1), t = 1, \cdots, N-n$.

After the data for $x_n(t)$ and $x_{n+1}(t+1)$ are obtained, the system matrices are estimated based on the following equation:

$$\begin{bmatrix} x_{n+1}(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} x_n(t) + \begin{bmatrix} B \\ D \end{bmatrix} u(t) + \begin{bmatrix} K_n \\ I \end{bmatrix} \varepsilon_n(t) \quad (53)$$

Again, since $\varepsilon_n(t)$ is orthgonal to $x_n(t)$ and $u(t)$, linear least squares gives consistent estimates of A, B, C, and D. One can also obtain consistent estimates of K_n and $P_n(\triangleq \mathrm{Cov}\{\varepsilon_n(t)\})$ using the residuals from the least squares.

In overall, one obtains consistent estimates of the Kalman filter matrices in (46) since the first least squares yields a consistent estimate of H_n^p and therefore of $x_n(t)$ (within a coordinate transformation), and the second least squares yields a consistent estimate of (A, B, C, D, K_n, P_n) provided consistent estimates of $x_n(t)$ and $x_{n+1}(t+1)$ are used.

With the obtained estimates denoted hereafter by $\left(\hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{K}, \hat{P}\right)$, one can form the state-space model

$$\begin{aligned} x(t+1) &= \hat{A}x(t) + \hat{B}u(t) + \hat{K}e(t) \\ y(t) &= \hat{C}x(t) + \hat{D}u(t) + e(t) \end{aligned} \quad (54)$$

where e is assumed to be a white noise process of covariance $\hat{P}$. The deterministic part of the model is unbiased and the stochastic part is slightly biased due to mismatch between K_n and K_∞ as well as that between P_n and P_∞. *Note:* If the input u is not white, the assumed initialization of the non-steady-state Kalman filter should be altered slightly in order to make $\mathcal{E}_n^{0+}$ uncorrelated with $x_n(t)$ and $U^{0+}(t)$. In this case, the proper initial estimate to assume for the Kalman Filter is not zero, but a function of $\mathcal{U}_n^-(t)$ and $\mathcal{Y}_n^-(t)$. This creates a slight inconsistency between the assumed initialization of the underlying Kalman filter for $x_n(t)$ and that for $x_{n+1}(t+1)$,

and complicates the procedure somewhat. The details of the modifications needed to save the nice asymptotic property are worked out in (Van Overschee and De Moor, 1994, 1996). This is mostly a theoretical concern, however. Since the effect of initialization on the Kalman estimate becomes negligible as $n \to \infty$, the basic algorithm works well with a large n, even when the input is not white.

Problem with Closed-Loop Data. With closed-loop data, consistency for the first least squares estimation breaks down. This is because $\mathcal{E}_n^{0+}(k)$ is no longer uncorrelated with $\mathcal{U}^{0+}$. For example, $u(t)$ is a function of $y(t)$ and therefore of $\varepsilon_n(t|t-1)$, which is correlated with $\varepsilon_n(t+1|t-1)$. Note that $u(t)$ appears in the regressor for the two-step-ahead prediction model of $y(t+1)$. Hence, the regressor becomes correlated with its residual $\varepsilon_n(t+1|t-1)$. Since $\hat{H}_n^p$ is biased, the whole argument for convergence breaks down.

The Modifications. Since in the indirect approach, system input is an external perturbation signal which has no correlation with system noise, the standard subspace method can be used to obtain a closed-loop system model. Given a state-space representation of the controller, a state-space model for the open-loop system can easily be extracted from it.

Van Overschee and De Moor (1997) discuss a way to modify the projection algorithm they use in their N4SID algorithm to account for closed-loop nature of data. Though this method does not belong to the indirect approach, the method requires knowledge of the controller as is the case for the indirect approach.

The joint input-output approach can also be integrated seamlessly with the subspace method. The subspace method can be used to identify a joint system and then an open-loop system model can be conveniently extracted from it, as shown in (Verhagen, 1993).

In addition, the two-step method and the projection method discussed earlier in the context of PEM should be applicable here. Since the steps for applying these methods would be essentially the same as before, just with PEM replaced by a subspace method, they are not discussed here.

Finally, Ljung and McKelvey (1996) proposed an alternative way to construct state data. Instead of picking the state basis from the projected data matrix, which poses a consistency problem in the case of closed-loop data, they suggested to identify a high-order ARX model and use it to calculate the n-step predictions. Note that identifying a high-order ARX model amounts to identifying just the first of the n-step predictor in (51) with a very large n. The reason for identifying only the first of the n-step predictor is that consistent estimates of the one-step-ahead predictor can be obtained even with closed-loop data, since the regressor is uncorrelated with the residual. The identified ARX model is used successively to construct the n-step predictions of y, which are obtained with all the future inputs set equal to zero (*i.e.*, $u(t) = \cdots, u(t+n-1) = 0$). Denote the resulting n-step predictions by

$$\hat{\mathcal{Y}}_n^{0+}(t) = \begin{bmatrix} \hat{y}^T(t|t-1) & \hat{y}^T(t+1|t-1) \\ & \cdots \quad \hat{y}^T(t+n-1|t-1) \end{bmatrix}^T. \quad (55)$$

From here on, $\hat{\mathcal{Y}}_n^{0+}(t)$ is treated an object equivalent to $\hat{H}_n^p \begin{bmatrix} \mathcal{Y}_n^-(t) \\ \mathcal{U}_n^-(t) \end{bmatrix}$ in the standard method. Hence, the state is created as $\hat{x}_n(t) = L\hat{\mathcal{Y}}_n^{0+}(t)$, where L contains the basis picked by examining the data matrix for $\hat{\mathcal{Y}}_n^{0+}(t)$.

Control-Oriented Process Identification

The term *control-oriented process identification* refers to a tailoring of an entire identification process to requirements of intended control. It points directly to the much needed integration between model development and controller design. The term perhaps was born in the midst of a debate on the apparent disparity between system identification methodologies and robust controller design methodologies. It was pointed out that, since most robust controller designs require both a nominal model and error bounds, more than just a nominal model should be passed from the identification step to the controller design step.

Over the course of time, the term has come to mean much more than simply providing a model description suited to robust controller design. Since achievable performance of a model-based controller, regardless of design strategy one employs, depends on the model quality, it is beneficial to shape the whole identification process based on the ultimate goal of achieving the best possible closed-loop performance. For example, the data generation step has a direct bearing on the size and distribution of model error, which in turn influence the closed-loop performance achievable with a particular controller design strategy. Hence, rather than using some universal testing procedure that ignores the characteristics of underlying plant dynamics and control objective, one should *tailor-design* the procedure to befit the plant characteristics.

Integration of identification and control design has been presented by many authors going back to Ziegler and Nichols (1942) and Åström and Hägglund (1984) who aimed at model free tuning of simple controllers. The desire for development of identification for robust control was indicated by Andersen et al. (1991)at CPC-IV. Theoretical interest in identification for control design has grown significantly during the past decade and the intensive research has yielded some useful insights and methodologies.

One fundamental understanding that emanated from

the research is the necessity of iteration. For example, control-oriented design of a data-gathering experiment demands that one relates the design parameters to model error and then ultimately to the closed-loop performance. Under well-defined assumptions on the prior model set and data set (or prior probability distributions on the model parameters and measurement errors), it is possible to relate design parameters (*e.g.*, test signals) to model error. The more complicated half of the puzzle is how model error degrades closed-loop performance. This degradation depends on the nominal model as well as the controller design method employed (*i.e.*, how the nominal model and possibly model error information gets translated into a controller). Even though the controller design method may be known a priori, the nominal model is not. This leads to the necessity of approximating it through iteration (Cooley and Lee, 2001).

Even though the complex relationship between model error and closed-loop performance is the challenging aspect of the problem from a theoretical viewpoint, the issue of how design parameters influence model quality is by no means simple and in fact may be a more serious barrier to practical use of most available approaches. This is because the question inherently requires prior information on plant and measurement error (as deterministic bounds or probability distributions), which are not readily available or easily expressed in the required mathematical terms.

The problem of control-oriented identification can also be posed in the context of closed-loop identification, and in fact most research in this area has followed this route. The main result is that, to minimize the closed-loop error that will ultimately result from a model-based controller, the identification of the model should use closed-loop data produced with the very controller (Gevers, 2000). However, the controller is not known prior to an actual experiment and therefore iteration (between controller design and closed-loop identification) is performed with the hope that the controller converges through the iteration.

The following section presents key issues and results regarding identification of models with the aim of achieving the best possible closed-loop performance. First, we present some preliminaries related to tradeoffs between two types of identification errors. Understanding of how model error arises from different sources plays an important role during the development of identification for control. Thereafter control designs and model approximations are explained before development of identification for control is described and a methodology for joint optimization of modelling and control is presented.

Preliminaries

Identification. In identification for control, it is preferred to identify low order models, which subsequently may be used for robust controller design. Given that such low order models cannot represent the true plant over the entire relevant frequency range they will have a systematic error, *a bias error*, in addition to the inevitable noise-induced *variance error*. The relative contributions of the two types of error may be expressed in the frequency domain, by defining a model error $\tilde{T}$, where the model $T(q) = [P(q) \quad H(q)]$ contains a process and a noise transfer function $y(t+1) = P(q)u(t) + H(q)e(t)$. Thus a performance objective $\bar{J}$ may be formulated as

$$\tilde{T}(e^{i\omega}) = \hat{T}(e^{i\omega}) - T_0(e^{i\omega}) \tag{56}$$

$$\bar{J}(\tilde{T}) = \int_{-\pi}^{\pi} E\left\{\tilde{T}(e^{i\omega})C(\omega)\tilde{T}^T(e^{-i\omega})\right\} d\omega \tag{57}$$

where the Hermitian 2x2 block matrix weighting function $C(\omega)$ describes the relative importance of a good fit over the frequency range of interest as well as the relative importance of the fit of P and H respectively. Note that the expectation is due to the randomness of $\tilde{T}$. Using a first order approximation of the model error gives:

$$\begin{aligned}\tilde{T}_N(e^{i\omega},\Theta) &= \\ &= \hat{T}_N(e^{i\omega},\Theta) - T_0(e^{i\omega}) \\ &\simeq T(e^{i\omega},\Theta^*) - T_0(e^{i\omega}) + (\hat{\Theta}_N - \Theta^*)^T T'(e^{i\omega},\Theta^*) \\ &= B(e^{i\omega},\Theta) + (\hat{\Theta}_N - \Theta^*)^T T'(e^{i\omega},\Theta^*)\end{aligned} \tag{58}$$

where

$$T'(e^{i\omega},\Theta) = \frac{dT(i\omega,\Theta)}{d\Theta} \tag{59}$$

$$B(e^{i\omega}) = T(e^{i\omega}) - T_0(e^{i\omega}) \tag{60}$$

and Θ^* is the parameter estimate as $N \to \infty$.

Substituting this into (57) gives the following approximate expression Ljung (1999):

$$\bar{J} \simeq J_B + J_P \tag{61}$$

with

$$J_B = \int_{-\pi}^{\pi} B(e^{i\omega})C(\omega)B^T(e^{-i\omega})d\omega \tag{62}$$

$$J_P = \frac{1}{N}\int_{-\pi}^{\pi} tr[P(\omega)C(\omega)]d\omega \tag{63}$$

where

$$P(\omega) = T'^T(e^{-i\omega})[N\mathrm{Cov}(\hat{\Theta_N})]T'(e^{i\omega}) \tag{64}$$

The bias contribution J_B is mainly affected by the model set as well as by signal power spectrum. The variance contribution J_P on the other hand decreases with increasing amount of data and signal power, whereas it increases with the number of parameters estimated. Minimizing the objective function in (61) clearly involves

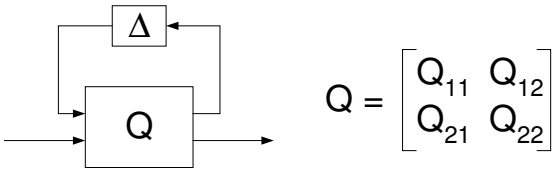


Figure 2: Representation of model perturbation by upper Linear Fractional Transformation (LFT) $F_u(Q, \Delta)$.

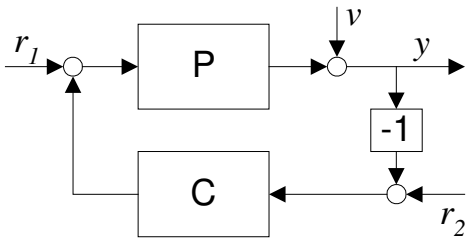


Figure 3: Feedback connection $\mathcal{T}(P_0, C)$ around a plant P_0.

making a tradeoff between bias and variance, i.e. between the model set and the number of parameters used. A proper decision on the tradeoff is key to success of any process identification scheme. The tradeoff is a factor through the design of an entire process identification process, including selection of model set and its parameterization, experimental design (e.g., test signals' power spectra), etc. The development of identification methods for control design has also exploited this tradeoff in various ways as described in the sequel.

Representation of Model Approximations. Knowledge of the nominal plant to be controlled is generally incomplete for robust controller design. Model error can be represented by a set of models. Such a set can be built up from a nominal model $\bar{P}$ along with a model perturbation Δ (Doyle et al., 1992) as illustrated in Figure 2. This Linear Fractional Transformation (LFT) model framework can represent additive as well as multiplicative uncertainty, each of which requires a different Q. In addition Q will contain the necessary information to characterize a set of models, $\mathcal{P}$. In Hansen and Fanklin (1988), a fractional model representation is described by a quotient of two stable factors parameterized via the dual Youla-Kucera parameterization. This approach is able to deal with estimation of a model set that includes both stable and unstable plants, even when operating under feedback controlled conditions. This approach is further explored by introducing a separate coefficient matrix Q (Zhou et al., 1996), where the entries depend on the nominal model $\bar{P}$ and the way in which the allowable perturbation Δ affects the nominal model. The set of models $\mathcal{P}$ may be characterized as

$$\mathcal{P} = \{P | P = F_u(Q, \Delta)\}, \text{ with } ||\Delta||_\infty < 1 \tag{65}$$

where the upper LFT $F_u(Q, \Delta) \equiv Q_{22} + Q_{21}\Delta(I - Q_{11}\Delta)^{-1}Q_{12}$ assuming the inverse exists. With this representation the nominal model is $\bar{P} = Q_{22}$.

Control Design. The performance of a feedback connection $\mathcal{T}(P_0, C)$, where

$$\begin{pmatrix} y \\ u \end{pmatrix} = \mathcal{T}(P_0, C) \begin{pmatrix} r_2 \\ r_1 \end{pmatrix} \tag{66}$$

$$\mathcal{T}(P_0, C) = \begin{pmatrix} P_0 \\ C \end{pmatrix} (I + CP_0)^{-1} \begin{pmatrix} C & I \end{pmatrix} \tag{67}$$

which is constructed from a plant P_0 and a controller C as shown in Figure 3, can be characterized by a norm value of $J(P_0, C)$, which is a closed-loop-relevant operator. Minimization of this norm may be done through controller design directly, provided that measurements from the plant can provide information on the control objective function $J(P_0, C)$. If this norm can be assessed directly on the plant, then the controller may be tuned optimally iteratively. This approach may be termed *model-free controller tuning*. This idea has been used by several researchers. Hjalmarsson. et al. (1994) developed an iterative procedure where two to three batch experiments are performed at each iteration. This procedure has been further developed to the multivariable case by Hjalmarsson and Birkel (1998), and reviewed by Gevers (1998). The direct iterative tuning is restricted to those control objective functions that can be assessed directly from observations. However, this methodology originated from studying the interplay between identification and control. This interplay is further discussed in the following sections, where the key aspect is to relate the purpose of identification, i.e., achieving control performance, to the identification procedure.

Identification of Model for Control

The nominal model $\hat{P}$ within a model set plays an important role. Since the nominal model is the sole basis for many controller designs, its quality is very important. In this section, we review the work on identifying a nominal model that leads to a good control performance when a controller design based on nominal performance is used. The importance of the nominal model has been recognized by many researchers, (e.g., Rivera and Gaikwad, 1992; Scharma and Bosgra, 1993; Zang et al., 1995; Lee et al., 1995).

Problem Formulation. A key idea in *identification for control* is to tune the bias and variance error for control design. Such a tuning may be achieved

by bounding the actual control performance $||J(P_0, C)||$ through utilization of the nominal control performance $||J(\hat{P}, C)||$ and the performance degradation due to approximate modelling $||J(P_0, C) - J(\hat{P}, C)||$. The bounding is achieved through exploitation of the triangular inequality:

$$||J(\hat{P}, C)|| - ||J(P_0, C) - J(\hat{P}, C)|| \leq ||J(P_0, C)|| \leq ||J(\hat{P}, C)|| + ||J(P_0, C) - J(\hat{P}, C)|| \quad (68)$$

Thus a tight upper bound may be achieved by minimizing the performance degradation $||J(P_0, C) - J(\hat{P}, C)||$. For a given controller C, this minimization constitutes a "control relevant identification problem" (Gevers, 1993; Van den Hof and Scharma, 1995), where a nominal model $\hat{P}$ is found by minimizing the difference between the performance of the feedback connections $\mathcal{T}(P_0, C)$ and $\mathcal{T}(\hat{P}, C)$. From the triangular inequality it is clear that both the nominal model $\hat{P}$ and the controller C may be used to minimize the performance cost $||J(P_0, C)||$.

Alternating between minimizing the performance degradation $||J(P_0, C) - J(\hat{P}, C)||$ as an identification problem and minimizing the nominal performance $||J(\hat{P}, C)||$ as a control design problem provides an iterative scheme for subsequent identification and control design. By such a scheme it is hoped that $||J(P_0, C)||$ decreases. The development of these schemes is briefly reviewed below as they have led to several interesting identification schemes.

Exact Modeling. Some of the first attempts towards investigating the interaction between identification and control in case of exact modelling, i.e. without bias error, were presented by Åström and Wittenmark (1971) and Gevers and Ljung (1986). The latter authors mentioned that, in the case of exact modelling, to minimize norm-based performance degradation, a prediction error based estimation method can be devised that uses closed-loop experiments and appropriate data filters. However the data filters contain knowledge about the controller to be designed, and therefore an iterative procedure of identification and control design can be used to gain knowledge of the data filters. This was confirmed by Hjalmarsson et al. (1996) and Forssell and Ljung (2000b). The latter authors investigated particular experimental design issues involved in minimizing the performance degradation due to variance errors in the identified model with constraints on a linear combination of input and output variance. For the case where penalty is only upon the misfit in P, they show that the optimal controller is given by the solution to a standard LQ problem. Also the optimal reference signal is determined.

The idea of minimizing the parameter covariance matrix is also pursued by Cooley and Lee (2001) in an open-loop identification's context. The control oriented design here takes the form of a weighted trace optimal (L-optimal) design, where the weighting matrices depend on desired loop shapes as well as the estimate from the last iteration. The optimal design was originally formulated as a nonconvex optimization for the sampled data values of test inputs directly, but was later reformulated by Samyudia and Lee (2000) as a Linear Matrix Inequality problem cast in terms of the covariance of the inputs. The procedure was shown to perform much superior to the conventional PRBS tests on several ill-conditioned multivariable processes, which served as the main motivation for their development.

Although most of the above results are useful in pointing out the desire for iterative experiments, they suffer from the requirement of exact modelling of the plant P_0. This requirement inevitably leads to the requirement of estimating a high order nominal model (e.g., FIR model used in Cooley and Lee (2001)).

Approximate Modeling. Analysis of the more realistic situation, where $\hat{P}$ is considered to be an approximation of P_0 was presented by Wahlberg and Ljung (1986). These authors showed that a norm based expression can be used to characterize the bias of a model $\hat{P}$, and that this bias could be tuned provided a suitable model structure was used. Liu and Skelton (1990) proposed closed loop experiments to provide the proper weighting filters in an explicitly tunable bias expression from Wahlberg and Ljung (1986). However, *the controller to be used for closed loop experiments is still unknown.* The possibility of using closed loop experiments was proposed by Liu and Skelton (1990) and developed by Zang et al. (1995), Hakvoort et al. (1994) and Tay et al. (1997).

Gevers (2000) in his recent review discusses an approach based on the idea that minimization of the performance degradation caused by model error can be formulated as a prediction error minimization (weighted by the sensitivity function) with closed-loop data. However, for the equivalence to hold, the closed-loop data must be generated by the very controller, which is to be determined later and therefore unknown before the experiment. This naturally brings up an iterative scheme where the controller design with nominal model and the prediction error minimization with closed-loop data are alternated. Even though this approach has been applied with promising results on several examples, the convergence has been shown to fail.

The idea of providing proper weighting filters in a tunable bias expression during identification was also used by Rivera et al. (1993) and Rivera and Gaikwad (1992), where it is assumed that prefiltering of *open-loop* data from the plant can replace the benefits from closed-loop experiments. The variance error was not considered.

Although the model bias during the approximate identification is tuned towards the intended model applica-

tion in the above works, the model plant mismatch is not taken into account during the control design. In order to be able to account for both bias and variance errors in control design and also to judge the quality of the model against the performance requirement, estimation of a model set rather than just a nominal model is needed.

Identification and Robust Control Design

Estimating Model Sets. A set of models can be used to represent the incomplete knowledge of the plant P_0. The incompleteness is due to the limited availability of possible disturbed observations of the plant behaviour. Such a set of models can be considered to consist of all models that are validated by the data (Ljung, 1999). However, Smith et al. (1997) pointed out that it is never possible to validate models solely on the basis of finite number of experiments. Thus a set of models will consist of models that cannot be invalidated by the data from the plant P_0. The available data along with the prior assumptions give rise to a set of feasible models $\mathcal{F}$ (Hakvoort et al., 1994). If the prior assumptions are correct, then $P_0 \in \mathcal{F}$. However, the set $\mathcal{F}$ can be unstructured, and therefore estimation of a structured set $\mathcal{P}$ should be done such that $\mathcal{P}$ outer-bounds $\mathcal{F}$. Performance and robustness are conflicting requirements (Doyle et al., 1992), and in case of conservative control design, this conflict causes the performance of a designed controller to deteriorate. Thus $\mathcal{P}$ should be estimated in such a way that the performance degradation of a controller designed based upon $\mathcal{P}$ is as small as possible.

Hence, to enable incorporation of robustness into the design of a model based controller, a set of models must be estimated. Mainly two different approaches for estimating error bounds have appeared, which differ by the nature of underlying assumptions on the error bounds. When the prior assumptions are stochastic, so called 'soft' error bounds result. Examples here of are Goodwin et al. (1992), Bayard (1992), Rivera et al. (1993), Ninness and Goodwin (1995), and Cooley and Lee (1998). When deterministic assumptions on the data or the plant are used, non-probabilistic error bounds result. Consequently, these are called 'hard' bounds. Examples here of are Wahlberg and Ljung (1992) and Böling and Mäkilä. (1995). Combinations of both types of model error bounding may give the combined advantages (de Vries and Van den Hof, 1995; Hakvoort and Van den Hof, 1994). Some of the error bounding approaches take the intended control application into account by estimating control relevant, possibly low order, nominal models with an additive or multiplicative bound on the modelling error, e.g. Bayard (1992).

Although the approaches referenced in the two previous paragraphs treat bias and variance aspects separately, the estimation of a set of models $\mathcal{P}$ should include both bias and variance aspects. Thus the identification

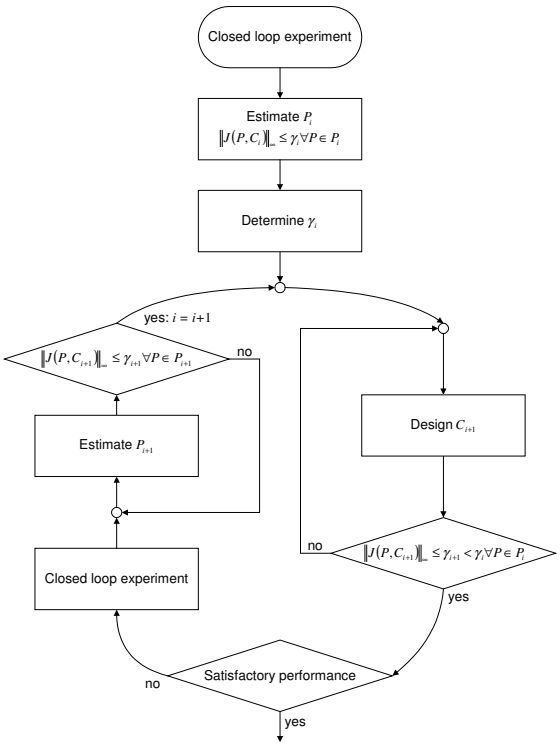


Figure 4: Iterative identification cycle for robust control.

and construction of the Q coefficient matrix in Figure 2, is crucial for the design of a well-performing robust controller (de Callafon, 1998). Furthermore the set must be suitable for robust control design, in order to enable performance improvement of the controlled plant. Thus the structure and the estimation of the set $\mathcal{P}$ should take the performance cost $||J(P_0, C)||$ into account.

Iterative Identification and Robust Control Design Procedure. The ultimate quest for joint optimization of modelling and control is a very complicated task, which still needs considerable research effort. A suboptimal procedure, depicted in Figure 4, has been proposed by de Callafon and Van den Hof (1997) and de Callafon (1998). One step in the procedure is that given a controller C_i at the i^{th} iteration, which together with the plant P_0 forms a stable feedback connection that satisfies the performance specification $||J(P_0, C_i)||_\infty \le \gamma_i$, then design a controller C_{i+1} that satisfies:

$$||J(P_0, C_{i+1})||_\infty \le \gamma_{i+1} \le \gamma_i \tag{69}$$

Another key step of the model-based procedure where the knowledge of the plant P_0 is represented by a set of models $\mathcal{P}_i$ is that evaluation of $||J(P_0, C_i)||_\infty$ can be achieved by evaluating $||J(P, C_i)||_\infty$ for all $P \in \mathcal{P}_i$. This enables specification of a control objective function that

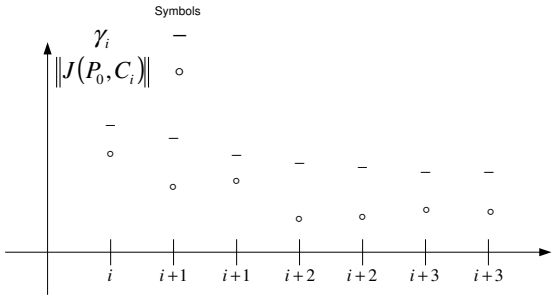


Figure 5: Progress during iterative identification cycle for robust control.

cannot be measured directly from data. Thereby three procedural steps can be outlined:

1. **Initial identification:** Use experimental data from $\mathcal{T}(P_0, C_i)$ and prior information on data or plant, to estimate a set of models $\mathcal{P}_i$ such that γ_i is minimized while $P_0 \in \mathcal{P}_i$, where
$$||J(P, C_i)||_\infty \leq \gamma_i \qquad \forall P \in \mathcal{P}_i \tag{70}$$
2. **Control Design:** Design C_{i+1} such that
$$||J(P, C_{i+1})||_\infty \leq \gamma_{i+1} \leq \gamma_i \qquad \forall P \in \mathcal{P}_i \tag{71}$$
3. **Re-Identification:** Use new experimental data from $\mathcal{T}(P_0, C_{i+1})$ and prior information to estimate a set of models $\mathcal{P}_{i+1}$ such that $P_0 \in \mathcal{P}_{i+1}$, subject to the condition
$$||J(P, C_{i+1})||_\infty \leq \gamma_{i+1} \qquad \forall P \in \mathcal{P}_{i+1} \tag{72}$$

Note that in the first step, a performance assessment is carried out to evaluate $||J(P_0, C_i)||_\infty$ for initialization purposes. Subsequently, the second step contributes a controller and the third step is a modelling validation step to enforce Equation 69. Subsequent to the initialization in step 1, repeated execution of steps 2 and 3 will provide a design procedure where the upper bound γ_i on a predetermined performance cost $||J(P_0, C_i)||_\infty$ can be progressively reduced as illustrated in Figure 5.

Implementation of this model based iterative procedure requires a procedure for estimating a set of models P tuned towards robust control design application. Such estimation of a modelling set may be carried out in a two step procedure described in de Callafon (1998). The set of models depend upon the nominal coprime factorization $(\hat{N}, \hat{D})$ and the weighting functions $(\hat{V}, \hat{W})$ that bound the model uncertainty. For the estimation, the closed loop performance based optimization
$$\min_{N,D,V,W} \sup_{P \in \mathcal{P}} ||J(P, C)||_\infty \tag{73}$$
is considered to ensure a model uncertainty set which is suitable for robust control design. The two estimation steps involve, (1)estimation of a nominal factorization such that the closed loop criterion is being minimized, subjected to internal stability of $\mathcal{T}(\hat{P}, C)$, and (2)estimation of model uncertainty, which consists of characterization of a frequency dependent upper bound on the uncertainty such that the closed-loop criterion is minimized using (V, W) subject to $P_0 \in \mathcal{P}$ (Hakvoort and Van den Hof, 1997).

Discussion. A number of important identification-related issues emerge from the above procedure:

- *Identification can be performed by open-loop experiments but more naturally by closed-loop experiments:* In order to account for the link between identification and control, identification should be performed with data that are representative of the eventual closed-loop condition. Although such data can be generated by carefully designing open-loop test signal, they are more naturally obtained by closing a loop with a sequence of controllers that gets progressively closer to the eventual controller.
- *Need for an iterative scheme:* Accounting for the intended application requires an iterative scheme of identification and feedback controller design. As the iteration continues, the feedback controller's robust performance continues to improve and this in turn generates information on the dynamics of plant that is more and more relevant for the ultimately intended robust control.
- *Model error information:* While performing iterations, information is developed on the complexity of nominal model, shape of allowable model uncertainty, and attainable robust performance. The characterization of modelling errors enables us to avoid performance degradation during the iterations. This also opens the possibility to formulate invalidation criteria for refusing models and controllers during iterations.
- *Unstable plants:* Both stable and unstable plants can be handled using the algebraic framework of stable factorizations, where a possible unstable dynamics is split into two stable factors. This split opens the possibility of an open-loop equivalent identification of the factors. The algebraic framework allows possible model errors to be described in a dual-Youla parameterization. Thus the effect of model perturbations can be studied under feedback controlled conditions. By considering a nominal stable factorization perturbed by an unknown but bounded stable operator, the set of models describe all models that are stabilized by a given feedback controller.

The above iterative procedure builds upon a large body of literature. The proposed procedure may be too complex for most applications, where the desire simply is to retune the feedback controller to improve the plant

performance. Simplifications of the procedure should be considered. Such simplifications could in fact become the standard practice of tomorrow for improving closed loop performance when requested by the operator on a routine basis as the need arises.

Summary

Closed-loop identification has progressed significantly over the past decade. Sound albeit limited theories have emerged to point out the limitations and tradeoffs for various closed-loop identification approaches. Final analysis reveals that there is no single methodology to be preferred on a universal basis and understanding of the advantages and disadvantages for the various available options is indispensable for making a right choice for a given situation. Potential benefits of closed-loop identification are huge and multi-faceted, ranging from the increased safety of the process and reduced harmful disruption to ongoing operation to the engineer's ability to collect more informative data in shorter time.

The area of identification for control holds significant promise to enable efficient, "plant-friendly" identification of multi-variable plants, including ill-conditioned plants, for which the traditional open-loop SISO identification is known to fail. Shaping of an identification process according to a specific control objective is naturally iterative in that various choices involved, e.g., shape of test signals, complexity of nominal model, shape of allowable model uncertainty and attainable performance, are highly dependent on the underlying plant. The shaping can be done in an open-loop experimental context where the data collected from previous experiments are used to improve the design of a new experiment. However, the shaping is more naturally done in a closed-loop experimental context, where data are made more and more representative of the eventual closed loop application by iteratively improving the controller. Despite the promise, more methodological developments are needed to address the practical issues, e.g., the requirement for ensuring the integrity of on-going operation, before routine application can be foreseen. Potential benefits of such methodological developments are significant in that process identification may be substantially facilitated when plants can be operated in their standard closed-loop configuration.

Acknowledgments

JHL gratefully acknowledges the financial support from the National Science Foundation (CTS #0096326), Aspen Technology, Weyerhaeuser, and Owens Corning.

SBJ gratefully acknowledges support from the European Union (Batch-Pro), The Department of Energy, Denmark, and The CAPEC industrial consortium (see `www.capec.kt.dtu.dk`).

References

Amirthalingam, R. and J. H. Lee, "Inferential Control of Continuous Pulp Digester," *J. Proc. Cont.*, **9**, 397–406 (1998).

Amirthalingam, R., S. Sung, and J. H. Lee, "A Two Step Procedure for Data-Based Modeling for Inferential Predictive Control System Design," *AIChE J.*, **46**, 1974–1988 (2000).

Andersen, H. W., M. Kümmel, and S. B. Jørgensen, "Dynamics and Identification of a Binary Distillation Column," *Chem. Eng. Sci.*, **44**, 2571–2581 (1989).

Andersen, H. W., K. Rasmussen, and S. B. Jørgensen, Advances in Process Identification, In Arkun, Y. and W. Ray, editors, *Proc. 4'th Int'l Conf. On Chemical Process Control—CPC IV*, South Padre Isl and, TX (1991).

Asprey, S. P. and S. Macchietto, "Statistical Tools for Optimal Dynamic Model Building," *Comput. Chem. Eng.*, **24**, 1261–1267 (2000).

Åström, K. J. and P. Eykhoff, "System Identification—A Survey," *Automatica*, **7**, 123–162 (1971).

Åström, K. J. and T. Hägglund, "Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins," *Automatica*, **34**, 90–113 (1984).

Åström, K. J. and B. Wittenmark, "Problems of Identification and Control," *J. Math. Anal. Appl.*, **34**, 90–113 (1971).

Bayard, D. S., Statistical Plant Set Estimation Using Schoeder-Phased Multisinosoidal Input Design, In *Proc. American Control Conference*, pages 2988–2995. Chicago,USA (1992).

Böling, J. M. and P. M. Mäkilä., On Control Relevant Criteria in H_∞ Identification, In *Proc. American Control Conference*, pages 3076–3080, Seatle, USA. Proc. American Control Conference (1995).

Cooley, B. L. and J. H. Lee, "Integrated Identification and Robust Control," *J. Proc. Cont.*, **8**, 431–440 (1998).

Cooley, B. L. and J. H. Lee, "Control-Relevant Experiment Design for Multivariable Systems Described by Expansions in Orthonormal Bases," *Automatica*, **37**, 273–281 (2001).

de Callafon, R. A. and P. M. J. Van den Hof, Multivariable Closed-Loop Identification: From Indirect Identification to Dual-Youla Parameterization., In *Proc. Of the 32nd Conference on Decision and Control*, pages 3384–3389 (1996).

de Callafon, R. A. and P. M. J. Van den Hof, "Suboptimal Feeback Control by a Scheme of Iterative Identification and Cotrol Design," *Mathematical Modelling of Systems*, **3**, 77–101 (1997).

de Callafon, R. A., *Feedback Oriented Identification for Enhanced and Robust Control*, PhD thesis, Technical University of Delft (1998).

de Vries, D. K. and P. M. J. Van den Hof, "Quantification of Uncertainty in Transfer Function Estimation: A Mixed Probabilistic—Worst-Case Approach," *Automatica*, **31**, 543–558 (1995).

Doyle, J. C., B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. MacMillan, New York (1992).

Forssell, U. and L. Ljung, "Closed-Loop Identification Revisited," *Automatica*, **35**, 1215–1241 (1999a).

Forssell, U. and L. Ljung, "A Projection Method for Closed-Loop Identification," *IEEE Trans. Auto. Cont.* (1999b). In press.

Forssell, U. and L. Ljung, "A Projection Method for Closed-Loop Identification," *IEEE Trans. Auto. Cont.*, **45**, 2101–2106 (2000a).

Forssell, U. and L. Ljung, "Some Results on Optimal Experiment Design," *Automatica*, **36**, 749–756 (2000b).

Gevers, M. and L. Ljung, "Optimal Experiment Design with Respect to the Intended Model," *Automatica*, **22**, 543–554 (1986).

Gevers, M., Towards a Joint Design of Identification and Control?, In Trentelman, H. and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 111–151. Birkhäuser, Boston (1993).

Gevers, M., "Iterative Feedback Tuning: Theory and Applications," *IEEE control Sections*, pages 26–41 (1998).

Gevers, M., A decade of progress in iterative process control design: from theory to practice, In *ADCHEM 2000*, pages 7–12, Pisa, Italy. IFAC (2000).

Goodwin, G. C., M. Gevers, and B. Ninness, "Quantifying the Error in Estimated Transfer Functions with Application to Model Order Selection," *IEEE Trans. Auto. Cont.*, **AC-37**, 913–928 (1992).

Gustavsson, I., L. Ljung, and T. Söderström, "Identification of Processes in Closed Loop—Identifiability and Accuracy Aspects," *Automatica*, **13**, 59–77 (1987).

Hakvoort, R. G. and P. M. J. Van den Hof, An Instrumental Variable Procedure for the Identification of Probabilistic Frequency Response Uncertainty Regions, In *Proc. 33rd IEEE Conference on Decision and Control*, pages 3596–3601, Lake Buena Vista, USA (1994).

Hakvoort, R. G. and P. M. J. Van den Hof, "Identification of Probabilistic System Uncertainty Regions by Explicit Evaluation of Bias and Variance Errors," *IEEE Trans. Auto. Cont.*, **AC-42**, 15166–1528 (1997).

Hakvoort, R. G., R. J. P. Schrama, and P. M. J. Van den Hof, "Approximate Identification with Closed-Loop Performance Criterion and Application to LQG Feedback Design," *Automatica*, **30**, 679–690 (1994).

Hansen, F. R. and G. F. Fanklin, On a Fractional Representation Approach to Closed-Loop Experiment Design, In *Proc. American Control Conference, Atlanta, USA*, pages 1319–1320 (1988).

Hjalmarsson, H. and T. Birkel, Iterative Feedback Tuning of Linear Time-Invariant MIMO Systems, In *37th IEEE Conf. On Decision and Control.* (1998).

Hjalmarsson., H., S. Gunnarsson, and M. Gevers, A Convergent Iterative Restricted Complexity Control Design Scheme, In *A Convergent Iterative Restricted Complexity Control Design Scheme*, pages p. 1735–1740, Lake Buena Vista, USA. pp. 1735-1740. 33rd IEEE Conference on Decision and Control (1994).

Hjalmarsson, H., M. Gevers, and F. Debruyne, "For Model-Based Control Design, Closed Loop Identification Gives Better Performance," *Automatica*, **32**, 1659–1673 (1996).

Johansen, T. A. and B. A. Foss, "Identifcation of Non-Linear System Structure and Parameters Using Regime Decopmosition," *Automatica*, **31**, 321–326 (1995).

Koung, C. W. and J. F. MacGregor, "Identification for Robust Multivariable Control: The Design of Experiments," *Automatica*, **30**, 1541–1554 (1994).

Larimore, W. E., Canonical Variate Analysis in Identification, Filtering and Adaptive Control, In *Proceedings. 29'th IEEE Conference on Decision and Control*, pages 596–604, Honolulu, Hawaii (1990).

Lee, W. S., B. D. O. Anderson, I. M. Y. Mareels, and R. L. Kosut, "On some Key Issues in the Windsurfer Approach to Adaptive Robust Control," *Automatica*, **31**, pp. 1619–1636 (1995).

Lee, J. H., *Modeling and Identification for Nonlinear Model Predictive Control: Requirements, Current Status and Future Research Needs*, pages 269–293. Birkhauser Verlag, Basel/Switzerland (2000).

Li, W. and J. H. Lee, "Control Relevant Identification of Ill-Conditioned Processes," *Comput. Chem. Eng.*, **20**, 1023–1042 (1996).

Liu, K. and R. E. Skelton, Closed Loop Identification and Iterative Control Design, In *Proc. 29th IEEE Conference on Decision and Control*, pages 482–487 (1990).

Ljung, L. and T. McKelvey, "Subspace Identification from Closed-Loop Data," *Signal Processing*, **52**, 209–215 (1996).

Ljung, L., *System Identification—Theory for the User*. Prentice Hall (1987).

Ljung, L., *System Identification—Theory for the User*. 2nd ed. Prentice Hall, New Jersey, second edition edition (1999).

Marquardt, W., Fundamental Modeling and Model Reduction for Optimization Based Control of Transient Processes, In *Chemical Process Control VI*. CACHE and AIChE (2001).

Ninness, B. and G. C. Goodwin, "Estimation of Model Quality," *Automatica*, **31**, 1771–1795 (1995).

Pearson, R. K., "Input sequences for nonlinear modeling," *Nonlinear Model Based Process Control. Proceedings of the NATO Advanced Study Institute*, pages 599–621 (1998).

Qin, S. J. and T. A. Badgewell, An Overview of Industrial Model Predictive Control Technology, In Kantor, J. C. and C. E. Garcia, editors, *Fifth International Conference on Chemical Process Control*, pages 232–256. Fifth International Conference on Chemical Process Control, AIChE and CACHE (1997).

Qin, S. J. and T. A. Badgewell, An Overview of Nonlinear Model Predictive Control Applications, In *International Symposium on Nonlinear Model Predictive Control: Assessment and Future Directions, Ascona* (1998).

Rivera, D. E. and S. V. Gaikwad, Modeling for Control Design in Combined Feedback/Feedforward Control, In *Proc. American Control Conference*, pages pp. 1445–1446 (1992).

Rivera, D. E., X. Chen, and D. S. Bayard, Experimental Design for Robust Process Control Using Schroeder-Phased Input Signals, In *Proc. American Control Conference,*, pages 895–899 (1993).

Samyudia, Y. and J. H. Lee, A Two-Step Approach to Control Relevant Design of Test Input Signals for Iterative System Identification, In *Proceedings of SYSID 2000* (2000).

Scharma, R. J. P. and O. H. Bosgra, "Adaptive Performance Enhancement by Iterative Identification and Control Design," *Int. Journal of Adaptive Control and Signal Processsing*, **7**, pp. 475–487 (1993).

Shi, R. and J. F. MacGregor, "Modeling of Dynamic Systems Using Latent Variable and Subspace Methods," *J. Chemometrics*, **14**, 423–439 (2000).

Smith, R. G., G. Dullerud, S. Rangan, and K. Poola, "Model Validation for Dynamical Uncertain Systems," *Mathematical Modelling of Systems*, **3**, 43–58 (1997).

Söderström, T. and P. Stoica, *System Identification*. Prentice Hall (1989).

Tay, T.-T., I. Mareels, and J. B. Moore, *High Performance Control*. Birkhauser (1997).

Van den Hof, P. M. J. and R. J. P. Scharma, "Identification and Control—Closed Loop Issues," *Automatica*, **31**, 1751–1770 (1995).

Van den Hof, P. M. J. and R. J. P. Schrama., "An Indirect Method for Transfer Function Estimation from Closed Loop Data," *Automatica*, **29**, 1523–1527 (1993).

Van Overschee, P. and B. De Moor, "N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems," *Automatica*, **30**, 75–93 (1994).

Van Overschee, P. and B. De Moor, "A Unifying Theorem for Three Subspace System Identification Algorithms. ," *Automatica*, **31**, 1877–1883 (1995).

Van Overschee, P. and B. De Moor, *Subspace Identification of Linear Systems. Theory, Implementation, Applications*. Kluwer Academic Publishers (1996).

Van Overschee, P. and B. De Moor, Closed-Loop Subspace Identification Algorithm, In *Proc. Of the 36th IEEE CDC*, pages 1848–1853 (1997).

Verhaegen, M. and P. Dewilde, "Subspace Model Identification. Part I: The Output-Error State Space Model Identification Class of Algorithms. Part II: Analysis of the Elementary Output-Error State Space Model Identification Class of Algorithms.," *Int. J. Control*, **56**, I: 1187–1210; II: 1211–1241 (1992).

Verhagen, M., "Application of a Subspace Identifcation Algorithm to Identify LTI Systems Operating in Closed-Loop," *Automatica*, **30**, 61–74 (1993).

Viberg, M., "On Subspace-Based Methods for the Identification of Linear Time-Invariant Systems," *Automatica*, **31**, 1835–1852 (1995).

Wahlberg, B. and L. Ljung, "Design Variables for Bias Distribution in Transfer Function Estimation," *IEEE Trans. Auto. Cont.*, **AC-37**, 900–912 (1986).

Wahlberg, B. and L. Ljung, "Hard Frequency-Domain Error Bounds from Least Squares Like Identification Techniques," *IEEE Trans. Auto. Cont.*, **AC-37**, 900–912 (1992).

Young, R. E., R. D. Bartusiak, and R. W. Fontaine, Evolution of an Industrial Nonlinear Model Predictive Controller, In *Preprints Chemical Process Control* -, pages 399–410 (2001).

Zang, Z., R. R. Bitmead, and M. Gevers, "Iterative Weighted Least-Squares Identification and Weighted LQG Control Design," *Automatica*, **31**, pp. 1577–1594 (1995).

Zhou, K., J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, New Jersey, USA (1996).

Zhu., Y. C., "Black-Box Identification of MIMO Transfer Functions: Asymptotic Properties of Prediction Error Models.," *J. Adapt. Cont. and Sig. Process.*, **3**, 357–373 (1989).

Ziegler, J. G. and N. B. Nichols, "Optimum Settings for Automatic Controllers," *Trans. ASME*, **42**, 759–768 (1942).

Controlled Biological Processes and Computational Genomics

James S. Schwaber
Daniel Baugh Institute for Functional Genomics/Computational Biology
Thomas Jefferson University College of Medicine
Philadelphia, PA 19107

Francis J. Doyle III and Daniel E. Zak
Department of Chemical Engineering
University of Delaware
Newark, DE 19716

Keywords

Biological processes, Computational genomics, Computational neuroscience

Introduction

The purpose of the present paper is to point out opportunities for systems and process engineering approaches that appear to be arising as a consequence of the explosion of genomic data. In particular we consider our own and other's work in approaching biological systems with modeling and simulation studies as a guide to the contemporary landscape of such opportunities.

It has been impossible to miss the hoopla as the Human Genome Project has moved rapidly forward over the past year. The genome—i.e. the full set of human genes—is the code that guides the development and operation of the human organism. The excitement over the Human Genome Project is certainly justified, but gene sequences are only a first step towards understanding gene function (Schena, 1996). Genes function in highly interconnected, hierarchical, and nonlinear networks. Organismic states and characteristics are often not the result of the expression of single genes but rather the result of interactions of multiple genes, as in the case of some human cancers (Szallasi and Liang, 1998), as well as the past and present intracellular and extracellular environments.

A surprising result of the genome projects has been the similarity of genomes across species, even between man and yeast. This has led many to conclude that the genes alone can not account for species complexity and differences. Regulation of specific gene activities is crucial for creating the complexity of higher organisms. Consider DNA as a huge, flexible macromolecule containing genes and a large number of binding sites. The DNA molecule organizes simultaneous and interacting chemical binding with very large numbers of other molecules at these sites. These sites bind transcription factors that control gene activity with resulting control of gene and protein expression, regulating cell phenotype. The control is the result of inputs from the environment through receptors which trigger signals that guide gene activity. Cell states are dynamic, being constantly influenced by environmental and intercellular signals.

The control engineering approach to biosystems modeling and analysis offers an integrative perspective and brings unique insights and tools. The motivation for such an approach is provided by the richly interconnected feedback layers that underlie much of biological regulation. Integrative analyses across spatial and temporal orders of magnitude are essential for understanding and interpreting the underlying behavior.

Tools from control engineering are also relevant in light of emerging high throughput quantitative techniques, such as DNA microarrays and proteomic methods. These allow the measurement of thousands of intracellular factors, such as messenger RNA transcripts (that are indicative of gene activity) and proteins, in parallel and over time. The data from these methods present significant challenges and opportunities. The challenges are due to the enormity of the systems and their complexity. They present opportunities because never before has it been possible to so sensitively measure the conditions within cells.

Unfortunately it is still a common view that biological systems are beyond the scope of engineering approaches. It is true that there are huge gaps in knowledge and the systems are immensely complex. However, it is instructive to look at the results to date from modeling and simulation studies of biological systems because there are substantial positive results. Two of the most influential, Nobel Prize-winning pieces of biological work, however, are purely theoretical: the Hodgkin-Huxley formalism (Hodgkin and Huxley, 1952) and the Watson-Crick double helix model for DNA (Watson and Crick, 1953). To further illustrate the usefulness of computational approaches for complex biological processes, our successes in computational neuroscience are discussed next. This is followed by a detailed discussion of the opportunities and challenges in high throughput methods, particularly DNA microarrays, that are relevant to control engineering.

Computational Neuroscience

The greatest success for theoretical and modeling approaches in biology has been in Computational Neuroscience. Our own early work in this domain was with a focus on biological control systems. At that time the

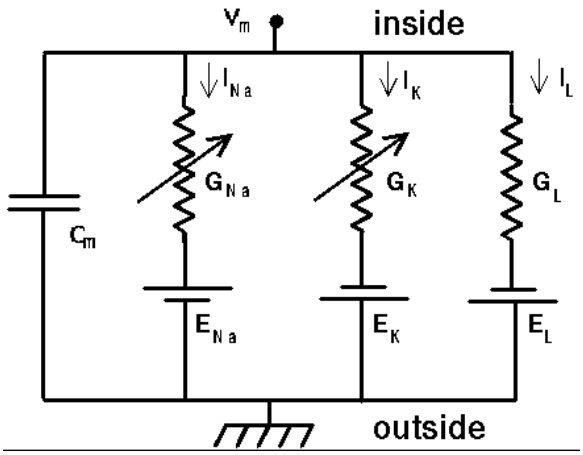


Figure 1: Hodgkin Huxley formalism.

Hodgkin-Huxley (HH) formalism was used in modeling studies of channel kinetics, one channel type at a time. Neuron models were sum and squash, non-spiking variety. We conceived of creating neuron models built up of multiple species of channel types using HH. The problem was computationally intractable, however, and the interaction within an extremely complex set of nonlinear, dynamical systems, including feedback, was certain to be extremely complex. We had excellent success, however, in creating complex, spiking neurons that were robust, occupying large, continuous and high-dimensional parameter spaces (Schwaber et al., 1993; Foster et al., 1993). We took these results to indicate that this class of biological system has evolved to be extremely robust, since its performance was not sensitive to parameter values. Parameter value variations are expected to result from varying environmental conditions in a real system.

We also were encouraged by the success at crossing levels of analysis, in this case from channel kinetics to whole cell behavior. We decided to extrapolate this work to additional levels, using several types of our HH neuron models to create neuronal networks with synaptic, chemical communication. We extended these network models to closed system models controlling peripheral organs. We were able to study the impact of manipulating cellular properties on system behavior in these models (Rybak et al., 1997a,b,c).

We have also been successful in connecting organism-level function (in this case, blood pressure regulatory behavior) with intracellular processing (second messenger pathways: Cheng et al., 1999, 1997; Hardwick et al., 1995; Parsons et al., 1987; Schwaber et al., 1993). Figure 2 shows the proposed reflex circuitry for short-term regulation of cardiovascular system. The local reflex architecture in Figure 2 is consistent with the experimental results in Cheng et al. (1997).

We have constructed a computer simulation model for the local cardiac reflex based on the anatomical experimental results and physiological data in the literature. Simulation results indicate that the local cardiac reflex could be effecting attenuation of the nonlinearity of the relationship between cardiac vagal drive and the arterial blood pressure. We have explored the hypothesis that the functional role of dynamic neuromodulation by SIF cells is an input-output "linearizing" effect on the actuator (heart) dynamics (Vadigepalli et al., 2001). We employed coherence analysis to characterize the nonlinearity between the frequency of vagus nerve pulsatile input and the mean arterial pressure. We also investigated the role of modulatory synaptic transmitters in this function.

We also have driven the levels of analysis question to more molecular and chemical processes, including receptors and signaling cascades through which environmental signals affect system behavior (Brown et al., 1999; Kholodenko et al., 1999; Kholodenko, 2000; Vadigepalli et al., 2001). This latter work has raised the question of connecting from the biochemical level to the gene level.

The cell and system regulation we are studying must, in life, interact with the regulation occurring at the gene level. The question that arises next is, "What would be needed to create an equally successful Computational Genomics?"

Computational Genomics

New analytical techniques have been developed that allow the quantification of many intracellular factors (Gombert and Nielsen, 2000). These include DNA microarray technology (Schena et al., 1995), which allows the relative transcription levels of thousands of genes to be measured in parallel, and gel electrophoresis and mass spectrometry, which allow levels of hundreds of proteins to be quantified (Gygi et al., 2000). Many groups have applied these methods as well as information from the genome sequencing projects to explore gene function on a genomic scale (Fodor et al., 1993; Schena et al., 1995; Ross-Macdonald et al., 1999; Winzeler et al., 1999; Uetz et al., 2000). DNA chips which allow for genome-wide measurements of mRNA levels (Schena, 1996; De Saizieu et al., 1998; Eisen et al., 1998; Marshall and Hodgson, 1998; McKenzie et al., 1998; Ramsay, 1998; Spellman et al., 1998; Brown and Botstein, 1999; Jia et al., 2000).

It is known that the data obtained with these methods have limitations. For example, it has been observed that mRNA and protein levels in do not correlate well enough for relative mRNA transcription levels to be predictive of protein expression levels (Gygi et al., 1999). The value of microarray data, however, should not be understated. It is of great value for generating hypotheses and it is widely accepted that, "The mRNA levels sensitively reflect the state of the cell, perhaps uniquely defining cell types, stages, and responses. To decipher the logic of gene regulation, we should aim to be able to monitor the expression level of all genes simultaneously"

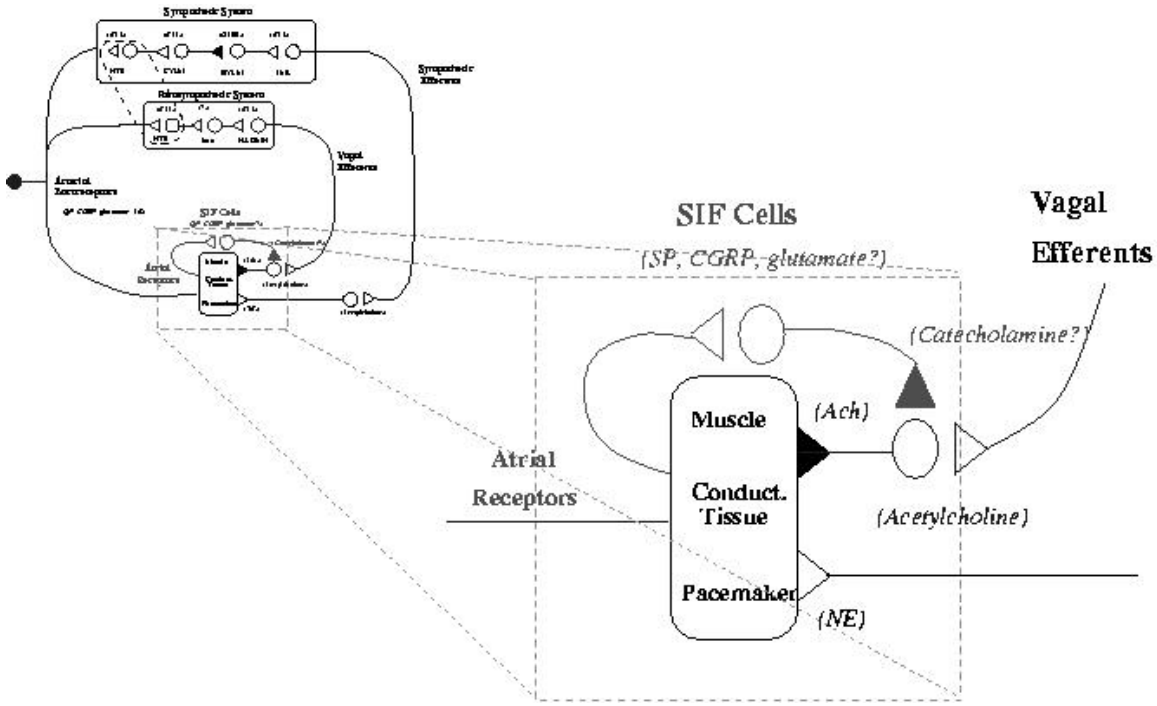


Figure 2: Schematic of local cardiac reflex control. The local cardiac reflex consists mainly of three components. (i) SIF cells receive sensory inputs from the atrial receptors and project to the principal neurons (PNs); (ii) PNs also receive input from vagal efferents from dorsal motor nucleus of the vagus; and (iii) PN activity has a phase-dependent effect on the SA node pacemaker and hence on the heart-beat cycle.

(Lander, 1996).

One motivation behind our group's interest in DNA microarrays and other high throughput quantitative techniques is their ability to provide insight into the genetic networks that dictate cellular responses to intercellular and enviromental stimuli. A simple example of a genetic network is the genetic switch modeled and synthesized by Gardner et al. (2000), shown below in Figure 3. Here, the product of one gene represses the transcription of the other gene, causing the system to have two stable steady states. A review of some of the complex behavior that can be observed for similar systems is given by Smolen et al. (2000). Gene networks in even the simplest organisms are expected to be complex. The bacteria E. Coli, for example, has approximately 2000 genes, with connectivities between the genes averaging between two and three (Thieffry et al., 1998).

Several factors make the determination of gene networks from DNA microarray data a challenge. These are related to the design of the experiments, interpreting the data and incorporating biological information, and determining model structures for gene networks. These are discussed below.

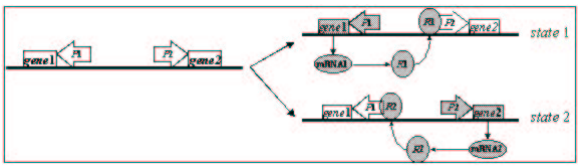


Figure 3: Schematic of the bistable genetic switch (Gardner et al., 2000).

Experimental Approach

The microarray experiments we perform in our lab involve the collection of transcript profiles over time while the organism is undergoing a systemic perturbation (chronic ethanol exposure, sleep deprivation, or chronic hypertension). By collecting array data over time, we expect to gain insights into the genetic basis for the organism's response to the perturbation. Our approach in designing these experiments is to apply the methods of system identification from systems engineering that follow a formal procedure of: (i) variable definition (i.e., which variables can be perturbed for maximum information), (ii) input sequence design (i.e., size and frequency of changes in the input variables), (iii) execution of input sequence, (iv) data cleaning (outlier detection, noise fil-

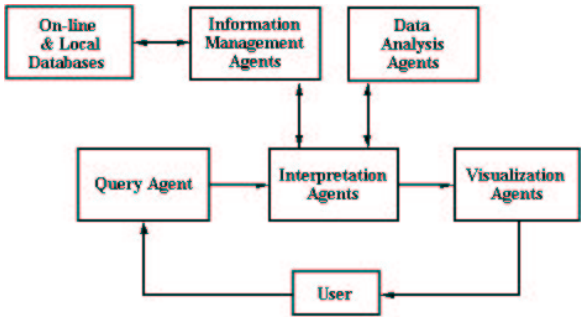


Figure 4: Architecture for analysis of genomic data using intelligent agents.

tering), (v) model building (using, e.g., correlation analysis or other statistical methods), and finally, and most importantly, (vi) validation of the model. Critical elements in this protocol for the biological problem at hand are the definition of suitable perturbation sources, the character of the perturbation signal, and the processing of the noisy data. A "rich" perturbation sequence will be designed to maximize the information content in the resulting signals. Richness in this context will be assessed by both the dynamic character of the forcing signal, as well as the nonlinear character of the forcing signal. The latter is especially important for the identification of nonlinear models, using such techniques as Laguerre kernel expansions (e.g., Marmarelis, 1993). A simple binary sequence (on-off) generates little or no information content at the output of a nonlinear system.

Data Issues

Interpretation of biological data consists of integrating data from several levels in the organismic hierarchy, including genomic sequences, microarray information, proteomic and metabolic information, and physiological information. This data will be analyzed through the use of various algorithms and statistical techniques. However, many of these techniques ignore domain knowledge of the system, which could lead to better and more efficient analysis of the data. Data analysis techniques like clustering are independent of whether they are applied on shopping data, weather data or biological data. We will incorporate domain knowledge from biology for improved and intelligent analysis of data. We have started some preliminary work in this area for the analysis of microarray data.

One of the analysis techniques that we have tested using simulator data is clustering. We will also compare clustering methods with other techniques and also with knowledge based clustering. In knowledge-based clustering, the implementation will use intelligent software agents to obtain knowledge automatically from databases. The intelligent agents will gather the knowledge from different sources on the web such as pathway databases, etc., and will use this knowledge during data analysis. We propose to build a system for microarray data analysis that can be incorporated to related ongoing work in intelligent multi-agent systems for data analysis. Specific agents will be used to gather the knowledge relevant for data analysis. This knowledge will then be integrated with the analysis technique such as clustering. A proposed architecture for this is given in Figure 4 above.

Gene Network Model Structures

Determining network architecture from microarray data is nontrivial. This data can consist of relative transcription levels for thousands of genes over hundreds of time points. Mathematical models, combined with biological knowledge, are necessary to determine the relationships contained within this data. A discussion of some of the key considerations in building such models, notably data requirement, is given in Fuhrman et al. (1999).

Several approaches for building models of transcriptional regulatory gene networks from temporal microarray data have been described in the literature. Top-down modeling approaches have been primarily designed for elucidating network connections from temporal microarray expression data. As mentioned above, there is no clear correlation between protein levels and relative transcription levels and it is therefore not possible to determine network connections from microarray transcription data alone (Gygi et al., 1999). Microarray data can be used, however, to generate hypotheses that can direct future experiments. Also, the methods are generally applicable to any system of large numbers of interacting components, and are thus relevant for interpreting temporal protein level data from gel electrophoresis and mass spectrometry. There are few examples in the literature where these methods have actually been applied (Reinitz and Sharp, 1995; D'Haeseleer et al., 1999), and therefore the utility of each of these methods still requires verification. Four examples of top-down modeling approaches are logical, linear, "linear plus squashing", and differential.

In the logical approach, genes are either "on" or "off" and have a limited number of inputs from other genes (Kauffman, 1993). This approach is appealing because it may give basic structural information and has the smallest data requirement, of order log2(N) time points if the transcription of N genes are only influenced by two genes each (Akutsu et al., 1999). Its main limitations are that the number of regulatory inputs must be limited a priori and genes that can have intermediate expression levels or influence the transcription of other genes to varying degrees are neglected (Weaver et al., 1999). At the next level of complexity is the linear approach (D'Haeseleer et al., 1999), where the transcription levels of the genes

at one time point are linear combinations of the expression levels of all of the genes at the previous time point. The drawbacks of this approach are that it requires at least as many time points as genes since it has N2 parameters, that it poses no economy on interconnections, and that the process it describes is not linear. In spite of this, D'Haeseleer et al. (1999), has had some success with this approach when applying it to expression data for rat nervous system development. An improvement to the linear approach is the "linear plus squashing" approach (Reinitz and Sharp, 1995; Weaver et al., 1999). The input to a gene is still a linear combination of the expression levels of all of the other genes, but now the input and the gene expression level are related by a sigmoidal "squashing function." This is a more realistic model of gene expression. A fourth approach is the differential model proposed by Chen et al. (1999). The time rate of change of mRNA concentration is expressed as a linear combination of the protein concentrations minus a degradation term. The time rate of change of protein concentration is a linear combination of the transcript concentrations minus a degradation term. Since this model includes additional states for the protein concentrations, it is an improvement over the linear model. A drawback is that the number of empirical parameters for this model is nearly twice that of the linear model, giving it a significantly larger data requirement. The system is also not completely determined unless initial protein concentrations are known.

Conclusions

Clearly a new era for biology is emerging that can bring tremendous developments in medicine and understanding. Before these can be realized, however, computational approaches must be developed that can make full use of the data coming from the high throughput technologies. Given the highly regulated and interconnected nature of biological systems, methods from control engineering should be able to contribute significantly towards this goal.

References

Akutsu, T., S. Miyano, and S. Kuhara, "Identification of Genetic Networks from a Small Number of Gene Expression Patterns Under the Boolean Network Model," *Pac. Symp. Biocomput.*, **4**, 7–28 (1999).

Brown, P. O. and D. Botstein, "Exploring the new world of the genome with DNA microarrays," *Nat. Genet.*, **21**, 33–37 (1999).

Brown, L. J., G. E. Gonye, and J. S. Schwaber, "Non-linear PI Control Inspired by Biological Control Systems," *Advances in Neural Information Processing Systems*, **11**, 975–981 (1999).

Chen, T., H. L. He, and G. M. Church, "Modeling Gene Expression with Differential Equations," *Pac. Symp. Biocomput.*, **4**, 29–40 (1999).

Cheng, Z., T. L. Powley, J. S. Schwaber, and F. J. Doyle III, "Vagal afferent innervation of the rat heart atria reconstructed with laser confocal microscopy: A hypothesis of a local reflex circuit," *J. Comp. Neurol.*, **381**, 1–17 (1997).

Cheng, Z., T. L. Powley, J. S. Schwaber, and F. J. Doyle III, "Projections of the dorsal motor nucleus of the vagus to cardiac ganglia of rat atria: an anterograde tracing study," *J. Comp. Neurol.*, **410**, 320–41 (1999).

De Saizieu, A., U. Certa, J. Warrington, C. Gray, W. Keck, and J. Mous, "Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays," *Nat. Biotechnol.*, **16**, 45–48 (1998).

D'Haeseleer, P., X. Wen, S. Fuhrman, and R. Somogyi, "Linear Modeling of mRNA Expression Levels During CNS Development and Injury," *Pac. Symp. Biocomput.*, **4**, 41–52 (1999).

Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster Analysis and Display of Genome-Wide Expression Patterns," *P. Natl. Acad. Sci. USA*, **95**, 14863–14868 (1998).

Fodor, S. P. A., R. P. Rava, X. C. Huang, A. C. Pease, C. P. Holmes, and C. L. Adams, "Multiplexed biochemical assays with biological chips," *Nature*, **364**, 555–556 (1993).

Foster, W. R., L. H. Ungar, and J. S. Schwaber, "Significance of conductances in Hodgkin-Huxley models," *J. Neurophysiol.*, **70**, 2502–18 (1993).

Fuhrman, S., P. D'haeseleer, and R. Somogyi, Tracing genetic information flow from gene expression to pathways and molecular networks, in DNA microarrays: the new frontier in gene discovery and gene expression analysis, Short course, Society for Neuroscience (1999).

Gardner, T. S., C. R. Cantor, and J. J. Collins, "Construction of a Genetic Toggle Switch in Escherichia coli," *Nature*, **403**, 339–342 (2000).

Gombert, A. K. and J. Nielsen, "Mathematical Modeling of Metabolism," *Curr. Opin. Biotech.*, **11**, 180–186 (2000).

Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold, "Correlation Between Protein and mRNA Abundance in Yeast," *Mol. Cell. Biol.*, **19**, 1720–1730 (1999).

Gygi, S. P., G. L. Corthals, Y. Zhang, Y. Rochon, and R. Aebersold, "Evaluation of Two-Dimensional Gel Electrophoresis-Based Proteome Analysis Technology," *P. Natl. Acad. Sci. USA*, **97**, 9390–9395 (2000).

Hardwick, J. C., G. M. Mawe, and R. L. Parsons, "Evidence for afferent fiber innervation of parasympathetic neurons of the guinea-pig cardiac ganglion," *J. Autonom. Nerv. Syst.*, **53**, 166–74 (1995).

Hodgkin, A. L. and A. F. Huxley, "A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve," *J. Physiol.*, **117**, 500–544 (1952).

Jia, M. H., R. A. LaRossa, J. M. Lee, A. Rafalski, E. J. De Rose, G. E. Gonye, and Z. Xue, "Global Expression Profiling of Yeast Treated with Sulfometuron Methyl, an Inhibitor of Amino Acid Biosynthesis," *Physiol. Genomics*, **3**, 83–92 (2000).

Kauffman, S. A., *The Origins of Order*. Oxford University Press, New York (1993).

Kholodenko, B. N., O. V. Demin, G. Moehren, and J. B. Hoek, "Quantification of short-term signaling by the epidermal growth factor receptor," *J. Biol. Chem.*, **274**, 30169–30181 (1999).

Kholodenko, B. N., "Negative feedback and ultrasensitivity can bring about oscillations in the miogen- activated protein kinase cascades," *Eur. J. Biochem.*, **267**, 1583–1588 (2000).

Lander, E. S., "The new genomics: global views of biology," *Science*, **274**, 536–539 (1996).

Marmarelis, V. Z., "Identification of nonlinear biological systems using Laguerre expansions of kernels," *Ann. Biomed. Eng.*, **21**, 573–89 (1993).

Marshall, A. and J. Hodgson, "DNA chips: an array of possibilities," *Nat. Biotechnol.*, **16**, 27–31 (1998).

McKenzie, S. E., E. Mansfield, E. Rappaport, S. Surrey, and P. Fortina, "Parallel molecular genetic analysis," *Eur. J. Hum. Genet.*, **6**, 417–429 (1998).

Parsons, R. L., D. S. Neel, T. W. McKeon, and R. E. Carraway, "Organization of a vertebrate cardiac ganglion: a correlated biochemical and histochemical study," *J. Neurosci.*, **7**, 837–46 (1987).

Ramsay, G., "DNA chips: State-of-the art," *Nat. Biotechnol.*, **16**, 40–44 (1998).

Reinitz, J. and D. H. Sharp, "Mechanism of eve Stripe Formation," *Mech. Develop.*, **49**, 133–158 (1995).

Ross-Macdonald, P., P. S. Coelho, T. Roemer, S. Agarwal, A. Kumar, R. Jansen, K. H. Cheung, A. Sheehan, D. D. Symoniatis, L. Umansky, M. Heidtman, F. K. Nelson, H. Iwasaki, K. Hager, M. Gerstein, P. Miller, G. S. Roeder, and M. Snyder, "Large-scale analysis of the yeast genome by transposon tagging and gene disruption," *Nature*, **402**, 413–8 (1999).

Rybak, I. A., J. F. Paton, and J. S. Schwaber, "Modeling neural mechanisms for genesis of respiratory rhythm and pattern I. Models of respiratory neurons," *J. Neurophysiol.*, **77**, 1994–2006 (1997a).

Rybak, I. A., J. F. Paton, and J. S. Schwaber, "Modeling neural mechanisms for genesis of respiratory rhythm and pattern II. Network models of the central respiratory pattern generator," *J. Neurophysiol.*, **77**, 2007–26 (1997b).

Rybak, I. A., J. F. Paton, and J. S. Schwaber, "Modeling neural mechanisms for genesis of respiratory rhythm and pattern III. Comparison of model performances during afferent nerve stimulation," *J. Neurophysiol.*, **77**, 2027–39 (1997c).

Schena, M., D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray," *Science*, **270**, 467–470 (1995).

Schena, M., "Genome Analysis with Gene Expression Microarrays," *BioEssays*, **18**, 427–431 (1996).

Schwaber, J. S., J. S. Graves, and J. F. R. Paton, "Computational Modeling of Neuronal Dynamics for Systems Analysis: Applications to Neurons of Cardiorespiratory NTS in the Rat," *Brain Res.*, **604**, 126–141 (1993).

Smolen, P., D. A. Baxter, and J. H. Byrne, "Mathematical Modeling of Gene Networks," *Neuron*, **26**, 567–580 (2000).

Spellman, P. T., G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher, "Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization," *Mol. Biol. Cell.*, **9**, 3273–3297 (1998).

Szallasi, Z. and S. Liang, "Modeling the Normal and Neoplastic Cell Cycle with Realistic Boolean Genetic Networks," *Pac. Symp. Biocomput.*, **3**, 66–76 (1998).

Thieffry, D., A. Huerta, E. Perez-Rueda, and J. Collado-Vides, "From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in, E," *Coli. Bioessays*, **20**, 433–440 (1998).

Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. M. Yang, M. Johnston, S. Fields, and J. M. Rothberg, "A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae," *Nature*, **403**, 623–627 (2000).

Vadigepalli, R., F. J. Doyle III, and J. S. Schwaber, Analysis and Neuronal Modeling of the Nonlinear Characteristics of a Local Cardiac Reflex in the Rat, Submitted to Neural Comput. (2001).

Watson, G. D. and F. H. C. Crick, "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid," *Nature*, **171**, 737 (1953).

Weaver, D. C., C. T. Workman, and G. D. Stormo, "Modeling Regulatory Networks with Weight Matrices," *Pac. Symp. Biocomput.*, **4**, 102–111 (1999).

Winzeler, E. A., D. D. Shoemaker, H. Astromoff, K. Liang, B. Anderson, B. Andre, R. Bangham, R. Benito, J. D. Boeke, H. Bussey, A. M. Chu, C. Connelly, K. Davis, F. Dietrich, S. W. Dow, M. El Bakkoury, F. Foury, S. H. Friend, E. Gentalen, G. Giaever, J. H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D. J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, J. L. Rebischung, J. L. Revuelta, L. Riles, C. J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R. K. Storms, V. R. S. M. Voet, G. Volckaert, T. R. Ward, R. Wysocki, G. S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston, and R. W. Davis, "Functional Characterization of the S. cerevisiae Genome by Gene Deletion and Parallel Analysis," *Science*, **285**, 901–906 (1999).

Stochastic and Deterministic Control in Two Bacterial Cellular Networks

Adam Arkin*
University of California
Lawrence Berkeley National Laboratory

Manuscript not submitted.

*aparkin@lbl.gov

Computer-Aided Design of Metabolic Networks

Klaus Mauch, Stefan Buziol, Joachim Schmid and Matthias Reuss *
Institute of Biochemical Engineering
University of Stuttgart, Germany

Abstract

This contribution presents results obtained from a model based design of metabolic networks. In the first part of the paper, topological analysis is used for exploring the metabolic architecture. These investigations—also called pathway analysis or flux space analysis—are aimed at detecting the metabolic routes that lead from anyone starting point to some products. The technique is applied for the computation of maximal yields for amino acids and, for the first time, also for the analysis of metabolic networks in context with the formation of biomass. The latter study leads to an array of mutants with different biomass yields, for which the name "Phenome" has been coined.

In the second part of the contribution, a strategy for the optimization of product formation rates is presented by means of the ethanol formation rate in *Saccharomyces cerevisiae*. A dynamic model based on experimental observations at defined anaerobic conditions serves as a starting point. Non-linear optimization of the distribution of enzyme activities results in a substantial increase of ethanol formation rate. The optimum is mainly constrained by homeostasis and can be characterized by higher activities of strongly rate limiting steps. However, some enzymes exerting almost no control on ethanol flux (e.g. triose phosphate isomerase) are found at higher activities as well. This finding can be explained by the enzyme's ability of counteracting an increase of pool concentrations effectively.

Keywords

Metabolic networks, Flux analysis, Flux optimization

Introduction

One of the fascinating challenges in mastering biosystems is to interpret living processes in a quantitative manner. As such, mathematical modeling is of central importance. At a time of the ballooning amount of data generated by the high throughput technology in genome, transcriptome, proteome and metabolome research, one of the important issues for modeling is to bridge the gap between data and an integrated understanding of the complex functionality of biosystems. Moreover, there is an urgent need for new approaches to strengthen the model based design of biosystems. This activity is of increasing relevance with respect to the optimization of yields, selectivities and productivities in industrial bioprocesses. Furthermore, mathematical modeling will lead to significant insight through an integrative analysis of diseases and support target identification for drug discovery.

While the potential and promise of biological systems modeling is substantial, also several obstacles are encountered. Before benefits can be gathered from biological systems analysis, issues like, for example, the appropriate balance between significance, complexity and availability of quantitative experimental observations need to be addressed. Furthermore, in many cases insufficient emphasis has been placed on fundamental questions of purpose, intended application (Bailey, 1998), predictive power and relevance concerning significant contributions to the solution of the aforementioned problems.

This contribution aims at the design of biosystems within industrial applications. Therefore, modeling is part of a well known engineering cycle depicted in Figure 1.

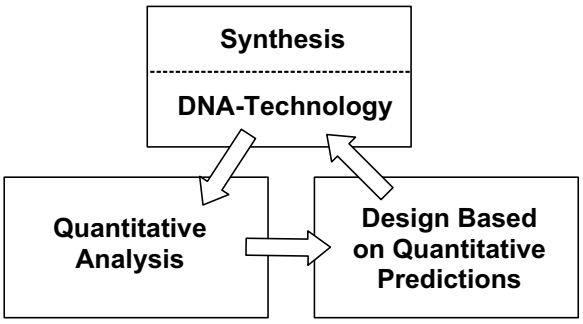


Figure 1: Optimization of bioprocesses through metabolic engineering.

Central to the general modeling framework is its predictive strength. A sound prediction, in turn, must rest upon reliable experimental data. In the majority of cases, implementation of the model-based suggestions for genetic reprogramming is performed with the aid of recombinant DNA-technology. Unfortunately, only few models qualify for the design of metabolic networks. When extrapolation beyond the horizon of experimental observations is required, the missing link mostly is predictive strength. A critical assessment of the state of the art and meaningful discussion of the present limits would require a comprehensive evaluation of several fundamental issues of modeling biosystems. Only two of these issues will be addressed in this contribution: (1) Flux analysis based on topological properties and (2) Flux optimization based both on toplogical and kinetic

*reuss@ibvt.uni-stuttgart.de

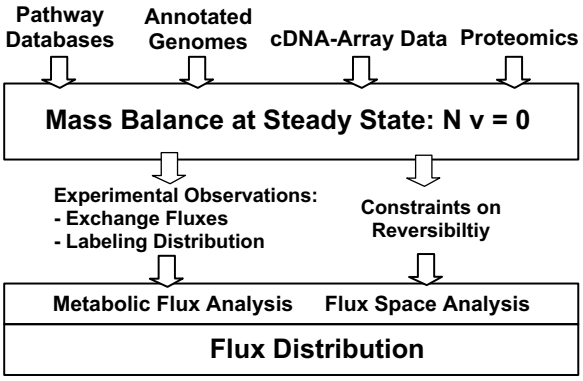


Figure 2: Quantification of flux distributions.

properties of metabolic networks.

Flux Analysis

According to the source of information, balance equations for the metabolic system of interest can be created (see Figure 2).

When balance equations are derived from annotated sequence data, the resulting stoichiometric matrix reflects the metabolic capabilities of the genotype. While depending on physiological conditions, a more specific phenotype is considered when data from cDNA chips or proteome analysis are used. Generally, balance equations of structured metabolic models can be written as

$$\frac{d}{dt}\mathbf{n} = \mathbf{N}\,\mathbf{v}, \tag{1}$$

where matrix $\mathbf{N}\,(m \times n)$ contains the stoichiometric coefficients $\nu_{i,j}$ of the n biochemical reactions. m denotes the number of metabolites, $\mathbf{v}$ the vector of reaction rates, whereas the vector of metabolites is denoted by $\mathbf{n}$. Note that in Equation 1 the transport rates across the various membranes of the cell as well as the dilution of metabolites due to growth are included in the state vector of the reaction rates $\mathbf{v}$. For steady state conditions Equation 1 reads:

$$\mathbf{N}\,\mathbf{v} = \mathbf{0}. \tag{2}$$

Metabolic Flux Analysis

Depending on the data available and the area of application, Equation 2 can be used for the computation of flux distributions in metabolic networks. The first route delivers solutions for flux distributions through experimentally determined exchange fluxes (Stephanopoulos et al., 1998; Mauch et al., 2000).

For the estimation of unknown metabolic fluxes from experimentally determined fluxes, we rewrite Equation 2 as

$$[\,\mathbf{N}_m|\,\mathbf{N}_c]\begin{bmatrix}\mathbf{v}_m\\ \hline \mathbf{v}_c\end{bmatrix} = \mathbf{0}, \tag{3}$$

where vector $\mathbf{v}_m$ consists of q experimentally determined fluxes. The remaining $(q-n)$ unknown fluxes are gathered in vector $\mathbf{v}_c$. Partitioning of $\mathbf{v}$ into the known fluxes $\mathbf{v}_m$ and unknown fluxes $\mathbf{v}_c$ then yields

$$\mathbf{N}_m\mathbf{v}_m + \mathbf{N}_c\mathbf{v}_c = \mathbf{0}, \tag{4}$$

with matrices $\mathbf{N}_m\,(m \times q)$ and $\mathbf{N}_c\,(m \times n-q)$ corresponding to $\mathbf{v}_m$ and $\mathbf{v}_c$, respectively. From Equation 4 we obtain

$$\mathbf{N}_c\mathbf{v}_c = -\mathbf{N}_m\mathbf{v}_m, \tag{5}$$

and upon multiplying Equation 5 by the transposed of matrix $\mathbf{N}_c$, the solution for unknown fluxes $\mathbf{N}_c$ is obtained according to

$$\mathbf{v}_c = -\left(\mathbf{N}_c^T\mathbf{N}_c\right)^{-1}\mathbf{N}_c^T\mathbf{N}_m\mathbf{v}_m, \tag{6}$$

where the superscript -1 indicates matrix inversion. By defining the pseudoinverse $\mathbf{N}_c^{\#}$

$$\mathbf{N}_c^{\#} = \left(\mathbf{N}_c^T\mathbf{N}_c\right)^{-1}\mathbf{N}_c^T, \tag{7}$$

Equation 6 may be rewritten as

$$\mathbf{v}_c = -\mathbf{N}_c^{\#}\mathbf{N}_m\mathbf{v}_m. \tag{8}$$

Mathematically, a determined metabolic system is defined by

$$dim\,(\mathbf{v}_m) = n - rank\,(\mathbf{N})\,. \tag{9}$$

That is, the amount of experimentally determined fluxes q in $\mathbf{v}_m$ equals the degree of freedom of the metabolic network. If no conservation relations are present in a determined system, the pseudoinverse $\mathbf{N}_c^{\#}$ coincides with the inverse of $\mathbf{N}_c$, thus

$$\left(\mathbf{N}_c^T\mathbf{N}_c\right)^{-1}\mathbf{N}_c^T = (\mathbf{N}_c)^{-1}\,. \tag{10}$$

Hence, the fluxes $\mathbf{v}_c$ of a determined system without conservation relations may be obtained by

$$\mathbf{v}_c = -\mathbf{N}_c^{-1}\mathbf{N}_m\mathbf{v}_m. \tag{11}$$

Importantly, a solution for $\mathbf{v}_c$ only exists for a nonsingular matrix $\mathbf{N}_c$, that is

$$det\,(\mathbf{N}_c) \neq 0, \tag{12}$$

and for a determined metabolic system with $0 \leq z < m$ conservation relations:

$$det\left(\mathbf{N}_c^T\mathbf{N}_c\right) \neq 0. \tag{13}$$

Provided the existence of a unique solution for $\mathbf{v}_c$, the metabolic system is called observable. For an underdetermined metabolic network we have

$$dim\,(\mathbf{v}_m) < n - rank\,(\mathbf{N})\,, \tag{14}$$

and since a unique solution for $\mathbf{v}_c$ cannot be obtained, we always find

$$det\left(\mathbf{N}_c^T\mathbf{N}_c\right) = 0. \tag{15}$$

Hence, the dimension l of the solution space for an underdetermined system is given by

$$l = n - rank\,(\mathbf{N}_c) - dim\,(\mathbf{v}_m)\,, \tag{16}$$

whereas an overdetermined (i.e. redundant) metabolic system is characterized by

$$dim\,(\mathbf{v}_m) > n - rank\,(\mathbf{N})\,. \tag{17}$$

Similarly to a determined system, in overdetermined systems a unique solution for the unknown fluxes $\mathbf{v}_c$ can only be derived if Equation 13 holds. Note that an overdetermined system is not necessarily observable. Frequently we meet the situation where experimental information on some fluxes is redundant while part of the metabolic network still cannot be observed. The problem of non-observable fluxes can often be bypassed when experimental data derived from tracer experiments are available. Those measurements of labeled substrates are either performed with NMR (Marx et al., 1996; Szyperski et al., 1999) or gas chromatography/mass-spectroscopy (GC-MS) (Christensen and Nielsen, 2000; Dauner and Sauer, 2000). When metabolite balance equations are extended with balance equations for the metabolite labeling distributions, a system of non-linear equations has to be solved to estimate the associated steady state flux distributions (Wiechert and de Graaf, 1997; Wiechert et al., 1997; Schmidt et al., 1997, 1999).

Flux Space Analysis

Flux space analysis (see Figure 2) summarizes strategies for topological analysis leading to meaningful information on the flux space in metabolic networks. This is an area of increasing importance in integrative functional genomics aimed at a better understanding of the complex relation between genotype and phenotype (Schuster et al., 1999; Schilling et al., 1999, 2000; Schuster et al., 2000a; Edwards and Palsson, 1999; Palsson, 2000). This area—in terms of the "omic" revolution sometimes named "phenomics"—becomes urgent for developing engineering methods to deal with the massive amounts of genetic and expression data in an integrative and holistic way.

Flux space analysis (sometimes also called pathway analysis) aims at detecting the metabolic routes that lead from anyone starting point to the products of interest steady state condtitions. The two strategies applied for this analysis are the null space and the convex cone. While linear algebra is used to determine the null space of the homogeneous system of linear equations, methods of convex analysis (Vanderbei, 1998) are applied to compute the convex cone.

The null space of matrix $\mathbf{N}$ is the subspace spanned by $k = rank\,(\mathbf{N}) - dim\,(\mathbf{v})$ linearly independent vectors $\mathbf{v}$ satisfying Equation 2. Accordingly, every linear independent base vector forms a column of the null space matrix $\mathbf{K}\,(n \times k)$, thus

$$\mathbf{N}\,\mathbf{K} = \mathbf{0}. \tag{18}$$

Within the null space lie all of the flux distributions under which the system can operate at steady state. Thus, the null space allows to describe any flux distribution of a genotype (by superposing it's base vectors). The vectors spanning the nullspace are, however, non-unique solutions of Equation 2. Moreover, the base vectors of the nullspace do not necessarily fulfill reversibility criteria of individual reactions.

A different approach rests upon convex analysis and leads to unique sets of vectors spanning the space of admissible fluxes, for which names like, for example, elementary flux modes (Heinrich and Schuster, 1996) or extreme pathways (Schilling et al., 2000) have been coined for. In contrast to the base vectors of the nullspace, flux vectors obtained by convex analysis always obey sign restrictions of practically irreversible reaction steps. The investigations presented in the following are based on the concepts of the elementary flux modes.

Elementary Flux Modes

Elementary flux modes are non-decomposable flux distributions admissible in steady state, including reaction cycles (Heinrich and Schuster, 1996). Normally, elementary flux modes comprise a set of non-zero fluxes and a set of zero fluxes, with the latter pointing to enzymes which are not used to implement a specific function. An example for an elementary mode which frequently occurs in cellular systems is the complete oxidation of substrate in the respiratory chain. To sustain respiration, enzymes catalyzing anabolic reactions obviously become dispensable. Another example might be an elementary flux mode leading to the formation of an amino acid where, again, larger parts of the respiratory chain are nonessential. Thus, by computing elementary flux modes, the metabolic capacity of a given metabolic network is unitized. In other words, the phenotype of a certain genotype may be characterized by the complete set of elementary flux modes.

The motivation for the study of elementary flux modes arises from various potential applications. In biotechnology, an important objective is to increase the yield of biosynthetic processes where a desired product can often be synthesized by various different routes. It is then of interest to detect and subsequently implement the route on which the product/substrate ratio is maximum. Generally, a flux pattern that uses only the optimal route cannot be obtained in practice. Nevertheless, it is helpful to compute the upper limits for the molar yield from a given network topology.

It has turned out that the problem of maximizing the yield of a biotransformation can be solved by detecting all elementary modes in the system and choosing the mode giving the best yield (Schuster et al., 1999). Al-

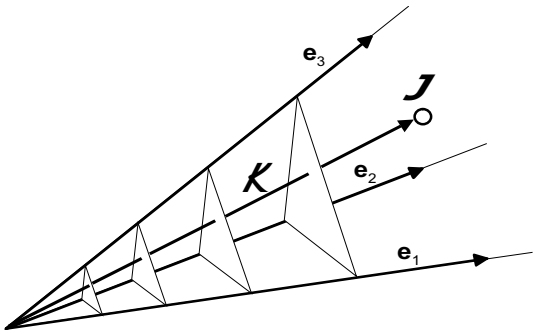


Figure 3: Convex polyhedral cone $\mathcal{K}$ spanned by three generating vectors $\mathbf{e_1}$ to $\mathbf{e_3}$. Flux distribution **J** results from a linear combination of the generating vectors $\mathbf{e_i}$.

ternatively, such optimization problems were tackled by linear programming (Fell and Small, 1986; Savinell and Palsson, 1992; van Gulik and Heijnen, 1995).

In many cases, the net direction of a reaction is known. Therefore, we decompose the flux vector into two subvectors, $\mathbf{v}^{irr}$ and $\mathbf{v}^{rev}$, corresponding to what will be called the irreversible and reversible reactions, respectively. So we have

$$\mathbf{v}^{irr} \geq 0. \tag{19}$$

Equation 2 in conjunction with inequality(19) determines what is called a convex polyhedral cone. The edges of the convex cone are established by the elementary flux modes, and all the points on the interior of the cone can be represented as positive combinations of these fundamental pathways (see Figure 3). From there, the convex cone enfolds all potential stationary flux distributions of a metabolic system.

Pursuing the goal to find basic pathways in biochemical reaction networks, Schuster and Schuster (1993) have earlier developed a method for detecting the simplest flux vectors **v** fulfilling relations (2) and (19) with all reactions assumed to be irreversible. Generalizing the approach in that both reversible and irreversible reactions are allowed for, this has led to the concept of elementary flux modes (Schuster and Hilgetag, 1994).

Elementary flux modes are defined as follows (Heinrich and Schuster, 1996, cf.): An elementary flux mode, $\mathcal{M}$, is defined as the set

$$\mathcal{M} = \{\mathbf{v} \in R^n | \, \mathbf{v} = \lambda \, \mathbf{v}^*, \lambda > 0\}, \tag{20}$$

where $\mathbf{v}^*$ is a n- dimensional vector (unequal to the null vector) fulfilling the following two conditions.

(a) Steady-state condition. $\mathbf{N}\,\mathbf{v}^* = \mathbf{0}$.

(b) Sign restriction.

If the system involves irreversible reactions, then the corresponding subvector $\mathbf{v}^{irr}$ of $\mathbf{v}^*$ fulfils inequality (19).

For any couple of vectors and (unequal to the null vector) with the following properties:

- $\mathbf{v}^{'}$ and $\mathbf{v}^{''}$ obey restrictions (a) and (b),
- both $\mathbf{v}^{'}$ and $\mathbf{v}^{''}$ contain zero elements wherever $\mathbf{v}^*$ does, and they include at least one additional zero component each, $\mathbf{v}^*$ is not a nonnegative linear combination of $\mathbf{v}^{'}$ and $\mathbf{v}^{''}$,$\mathbf{v}^* \neq \lambda_1 \mathbf{v}^{'} + \lambda_2 \mathbf{v}^{''}$, $\lambda_1, \lambda_2 > 0$.

The last condition formalizes the concept of genetic independence introduced by Seressiotis and Bailey (1988). The condition says that a decomposition into two other modes should not involve additional enzymes.

Elementary modes have been determined in a number of biochemical networks, such as the synthesis of precursors of aromatic amino acids (Liao et al., 1996), the tricarboxylic acid cycle and adjacent pathways (Schuster et al., 1999) and glycolysis and alternative pathways in bacteria (Dandekar et al., 1999).

While the promise is substantial, the value of the topological analysis will not be fully utilized until algorithms are developed capable of tackling large metabolic, signaling or gene networks efficiently. What are the limits of the known algorithms in the calculation of elementary flux modes when applied to large systems?

Most of the algorithms applied to biological systems so far are bottom up approaches (Schuster et al., 2000b; Schilling et al., 2000). Initially, the stoichiometric matrix is augmented with the identity matrix. Next, a consecutive computation of matrices through combination of rows is performed until the stoichiometric matrix only contains zero elements. Obviously, this is a consecutive approach where all elementary modes are created at the final step. In addition, a large number of interim solutions are computed which finally disappear. By consequence, since even modest network sizes showing a larger degree of freedom might feature thousands of elementary flux modes, the problem easily becomes computationally intractable. Nevertheless, (Schilling and Palsson, 2000) have recently tackled the problem of prediction of the so called extreme pathways for *Haemophilus influenza*, represented by a network of 461 reactions. The strategy applied rests upon a decomposition of the network into subsystems to get a picture of the structural information for the entire system.

Alternative approaches make use of top down strategies considering the entire system in a more direct way (Happel and Sellers, 1989). As a result, successive solutions containing the complete network information are created. Mauch (to be published) developed an algorithm performing combinations of the base vectors of the null space in pairs. Figure 4 shows an example of the application of this algorithm for a system with 141 reactions in which the polymerization reactions for formation of biomass—and therefore growth—has been taken into account for the prediction of the elementary flux modes.

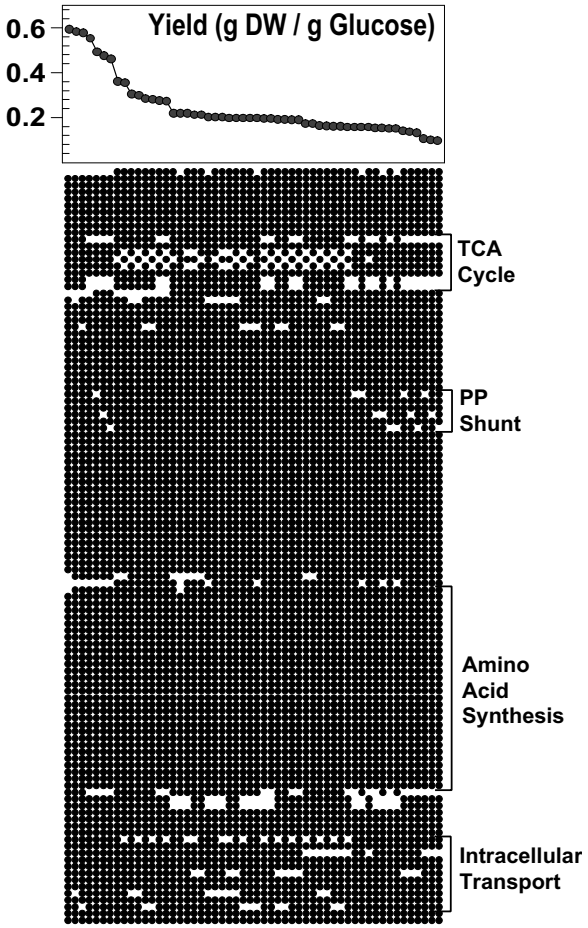


Figure 4: Elementary flux modes in *S. cerevisiae*. Explanation see text.

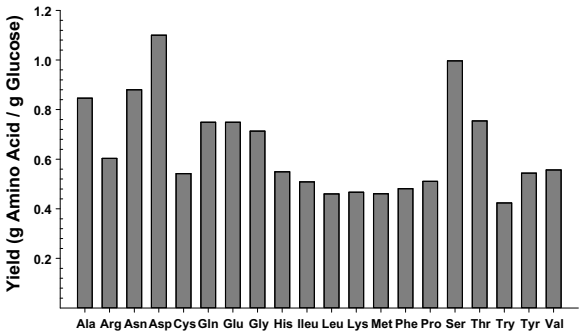


Figure 5: Maximal yields of amino acids on glucose in *S. cerevisiae*.

Each column shown in Figure 4 represents an elementary flux mode utilizing glucose as the sole carbon and energy source while producing biomass. Open symbols represent deleted genes for various reactions in the network. Obviously, these different mutants—or different phenotypes—are still able to grow albeit with varying yield coefficients. Elementary modes given in Figure 4 are ordered in descending sequence with respect to the yield of biomass on glucose. The maximal value of the biomass yield is in agreement to what has been measured for *S. cerevisae* growing with glucose under aerobic conditions. For mutants with defects in the transport systems via the mitochondrial membrane, for example, the biomass yield decreases to a value observed under anaerobic conditions. Thus, structural properties of the network clearly define the upper limit of product yields. Interestingly, these findings are independent of the kinetic properties of the network. Absolute values of fluxes as well as a dynamic response to system perturbations, however, can only be described when kinetic rate equations have been assigned.

By a closer inspection of the solution shown in Figure 4, it is possible to separate more flexible regions of the network from rigid parts. For instance, we cannot delete any reaction in the pathways leading to amino acids. This is a reasonable finding for growth in the presence of synthetic media. In contrast, greater flexibility is observed in the citric acid cycle (TCA), pentose phosphate (PP) shunt and intracellular transport. It is interesting to recognize that these variations also have been identified as key modulations during the process of evolution.

The design aspect of these predictions may be better elaborated by examples of biotechnological relevant product formation. A desired product can often be synthesized by various different routes or pathways within the network. It is then of interest to detect the route on which the product/substrate ratio is maximum. Figure 5 illustrates an example showing the maximal yield of the individual 20 amino acids on substrate glucose in *S. cerevisiae*.

The results shown in Figure 5 have been obtained by detecting all elementary modes in the system and then choosing the mode giving the best yield. Sometimes, however, implementation of an elementary mode leading to a slightly non-optimal yield might be easier from a practical point of view. In contrast to linear optimization, the complete set of elementary flux modes immediately provides the complete spectrum of alternative implementations.

Flux Optimization

The driving force for selection of an optimal pathway is the maximization of the yield of the product. However, economic considerations also require optimization of the product formation rate (productivity). This problem leads to the question of an optimal modulation of enzyme activities in metabolic networks.

At first glance, the example chosen for this discussion—ethanol production with the yeast *S.*

cerevisiae—appears to be rather boring because it has been tackled so many times by yeast geneticists. The primary aim of these empirical attempts to modulate (mostly amplify) the key enzymes within glycolysis is the maximization of ethanol production rate which correlates with carbon dioxide evolution and, in turn, with the baking power of yeast. While addressing this problem it should be emphasized that the interest in answering the question of an optimal redistribution of enzyme activities is much broader. Generally, knowledge of the rate controlling steps in the central metabolism (glycolysis and PP shunt) is of central importance for cell cultures used for producing proteins as well as for the analysis of potential targets in cancer cells. Another field of interest is the identification of potential targets for antitrypanosomal drugs, important for treatment of the african sleeping sickness (Bakker et al., 1999). The last named authors concluded that "Despite the great interest, it is not yet known completely for any organism how the control of the glycolytic flux is distributed".

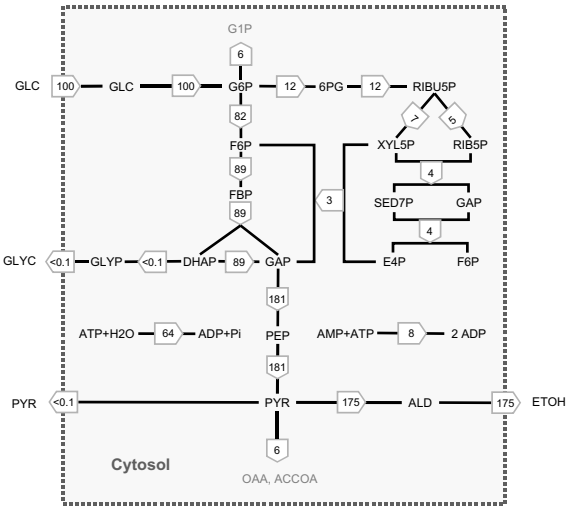


Figure 7: Flux distribution within central metabolic pathways of *S. cerevisiae* under anaerobic conditions at a specific growth rate of $\mu = 0.1\,h^{-1}$. All fluxes are related to the influx of glucose.

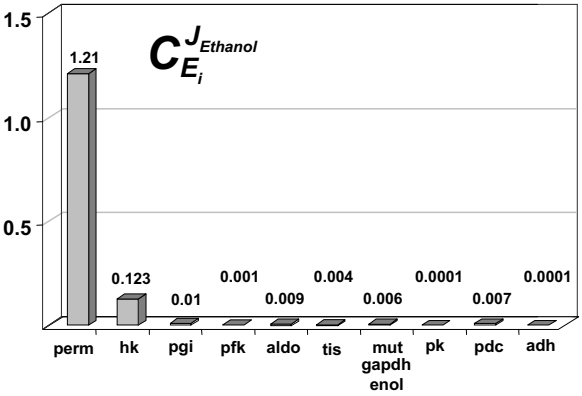


Figure 8: Distribution of flux control coefficients on ethanol formation route. Enzymes involved: Hexose transporter (perm), hexokinase(hk), phosphoglucose isomerase (pgi), phosphofructokinase (pfk), aldolase (aldo), triose phosphate isomerase (tis), gap dehydrogenase (gapdh), phospoglyceromutase (mut), enolase (enol), pyruvate kinase (pk), pyruvatedecarboxylase (pdc) and alcohol dehydrogenase (adh).

Dynamic Model

Similar to the strategy of Bakker et al. (1999), the problem can be approached from the basis of experimentally determined kinetic properties of the key enzymes which are then aggregated to a dynamic model. Individual rate expressions including their kinetic parameter have been identified *in vivo* by a stimulus response methodology (Theobald et al., 1997; Rizzi et al., 1997; Vaseghi et al., 1999): A pulse of glucose or alternative stimuli are introduced into a continuous culture operating at steady state and the transient response of several intracellular metabolite and cometabolite pools is experimentally determined in time spans of seconds or, recently, also milliseconds (Buziol et al., to be published). Within these relatively short time spans, enzyme concentrations are considered to be in a "frozen" state. Figure 6 summarizes examples for some of the experimental observations of metabolites and cometabolites from the yeast *S. cerevisiae* growing under anaerobic conditions.

The metabolome's response due to dynamic system excitation has been used to identify the dynamic system behavior by a stepwise internalization of metabolites similar to the method proposed by Rizzi et al. (1997). To describe the dynamic system behavior, deterministic kinetic rate equations for the pathways for the reactions have been formulated.

The general form of this rate equations can be written as

$$r_i = r_{max,i}\, f\left(\mathbf{c}, \mathbf{p}\right), \tag{21}$$

where the maximal rate (capacity) $r_{max,i}$ is obtained from the vector of model parameters $\mathbf{p}$, the vector comprising metabolite, cometabolite and effector concentrations $\mathbf{c}$ and the flux distribution $\mathbf{J}$ at the systems's steady state (e.g. $\mu = 0.1\,h^{-1}$); accordingly

$$r_{max,i} = 1/r_i^{steady\ state}\, f\left(\mathbf{c}^{steady\ state}, \mathbf{p}\right). \tag{22}$$

Intracellular flux distribution has been estimated by experimentally determined uptake and excretion rates of glucose, carbon dioxide, ethanol, glycerol and biomass. The results are documented in Figure 7.

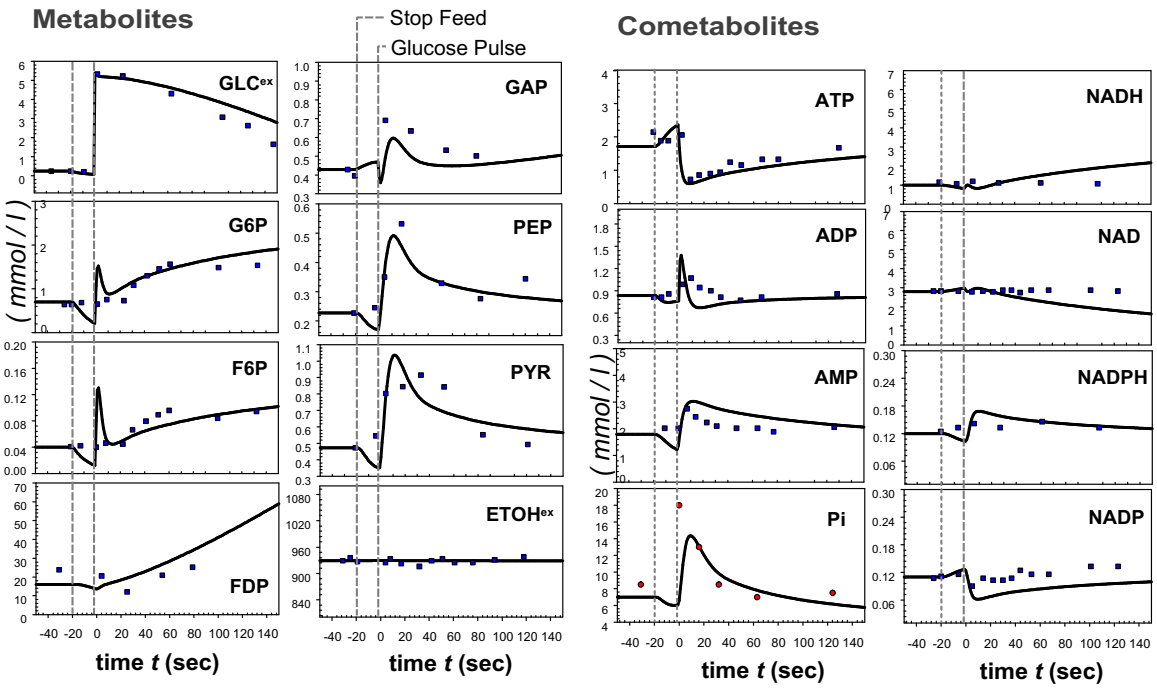


Figure 6: Comparison between model simulation and measured concentrations of glycolytic metabolites and cometabolites after dynamic system excitation. Data shown left from the broken lines represent steady state values at a growth rate of $\mu = 0.1\,h^{-1}$.

Sensitivity Analysis

The next step towards solving the envisaged design problem is to calculate the so called flux control coefficients, or—in the terminology of engineering—sensitivity coefficients. The flux control coefficient $C^J_{E_i}$ has been defined as the fractional change of the network flux J caused by a fractional change in the level of enzyme activity E_i. Thus,

$$Flux\ Control\ Coefficient = \frac{dJ\,(E_i)}{dE_i}. \qquad (23)$$

or normalized

$$C^J_{E_i} = \frac{dJ/J}{dE_i/E_i} = \frac{d\,lnJ\,(E_i)}{d\,lnE_i}. \qquad (24)$$

Figure 8 depicts the results of these calculations for those enzymes involved on the path from glucose to ethanol.

Since only a subset of the flux control coefficients with respect to ethanol formation are shown in Figure 8 (excluding, for example, the coefficients of enzymes involved in the PPP shunt), flux coefficients of this subset do not necessarily sum up to one. From the hierarchy of sensitivities it can be concluded that the enzyme responsible for the transport of glucose via the cell membrane (permease) shows the overwhelming control strength upon the ethanol production rate. Consequently, amplification of this enzyme should lead to an increased flux from glucose to ethanol.

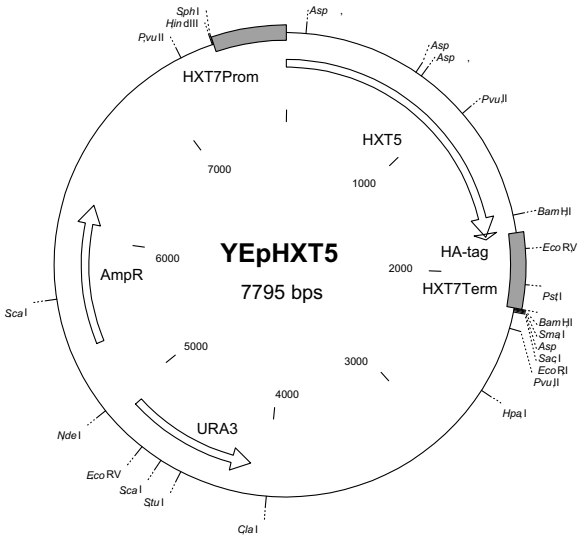


Figure 9: HXT5 multi copy plasmid.

Synthesis—Amplification of Hexose Transporters

Experimental verification of the above-named design proposal has been performed in a collaboration project with Institute of Microbiology from the University of

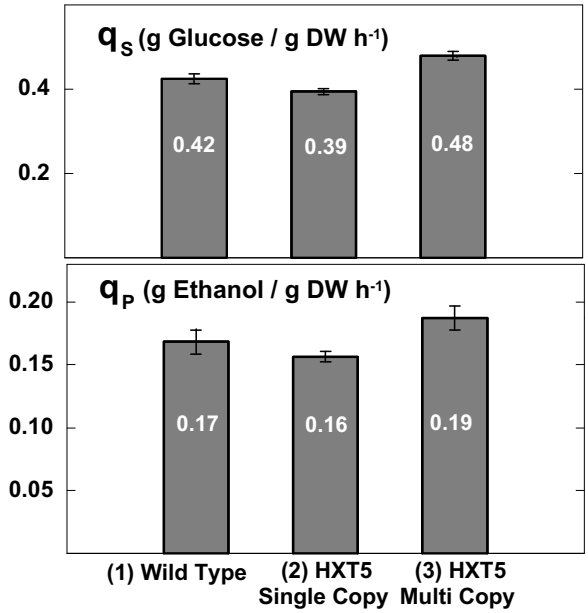


Figure 10: Specific substrate uptake rate q_s and specific product formation rate q_p during anaerobic chemostat cultivation. (1) Wild strain. (2) HXT5 with a single copy of the gene integrated the chromosome and constitutively expressed. (3) Multi copy plasmid with HXT5.

Düsseldorf (Dr. Boles). Within this project, the kinetics of the three most important hexose transporters out of the 17 transporters identified from the yeast genome project are investigated (Boles et al., 1997; Buziol et al., to be published). One of these s—HXT5—has been also expressed with a multi copy plasmid, shown in Figure 9 (Boles, to be published).

The transporter gene in HXT5 is flanked by an enhanced HXT7 promoter and HXT7 terminator, respectively. HXT7 is a high affinity transporter expressed at low glucose concentrations. The construct shown in Figure 9 ensures expression of HXT5 at target growth conditions. The same transporter has been integrated as single copy in the chromosome of a yeast strain in which all the genes of the other transporters has been knocked out (Wieczorke et al., 1999). As a result of these genetic constructions, it is possible to compare the flux of glucose through the glycolysis between two strains differing only in the amount of the hexose transporter. According to the hierarchy of flux control coefficients discussed in context with Figure 8, one would expect a noticeable increase of ethanol flux.

Experiments with three different strains were performed in continuous cultures at a dilution rate of $D = 0.07\,h^{-1}$ (Buziol et al., to be published). The results of these experiments are summarized in Figure 10.

Compared to the wild type, ethanol excretion rate and substrate uptake rate of the strain with higher transporter activity have found to be 10% and 25% higher, respectively. The discussion of the relevance of these results from industrial point of view is beyond the horizon of this paper. However, the rather modest effect prompts the question: Are there any other alternative design possibilities resulting in a substantial increase of the ethanol production rate?

Objective Function

The apparent failure to produce significant increase of the glycolytic flux points to the fundamental question if the underlying assumptions leading to the design suggestion are adequate. Keeping in mind that the pathway of interest is part of a whole—the living cell—, the idea that one need to amplify a single or multiple enzymes according to the hierarchy of flux control coefficients may not correspond to physiological reality. There are two aspects that should attract attention. First, an increased expression of enzymes is linked with energetically expensive protein synthesis. Glycolytic enzymes in yeast are known to contribute in the order of 30% to the total amount of cellular proteins. Thus, it seems to be likely that overexpression of these enzymes result in a stress situation with an unforeseeable impact on cell physiology. Instead of a single or simultaneous elevation of enzyme activities, a more robust strategy should try to keep the total concentration of proteins at a constant value and redistribute the activities according to the required objective. The optimization problem may then be stated as

$$\text{Maximize } J\left(\mathbf{r}_{max}^{Path}\right) \tag{25}$$

subject to

$$\frac{1}{w}\sum_{i=1}^{w}\frac{r_i^{max}}{r_{i,reference}^{max}} \leq \Omega. \tag{26}$$

In writing Equation 26 we assume that the maximal rate r_i^{max} is proportional to the enzyme amount. Therefore, Equation 26 specifies a fixed level for the total enzyme activity Ω.

Another relevant issue concerns the pool concentration of the metabolites within the cell. Attempts to increase metabolic fluxes by changes in individual enzyme concentrations may lead to substantial changes in metabolite concentrations. Again, a substantial change in metabolite concentrations either proves to be cytotoxic or at least leads to an undesired flux diversion (Kell and Mendes, 2000). Therefore, preservation of the metabolite concentrations close to the steady-state values of the wild strain is at any rate desirable to meet the basic property of well established metabolic systems for which the metaphor *homeostasis* has been coined (Reich and Selkov, 1981). It is also known that part of an optimal performance of cells with respect to the use of, for example, energy and carbon sources is a process of adaptation leading to a change of structure to control homeostasis. Interestingly, such structural changes may

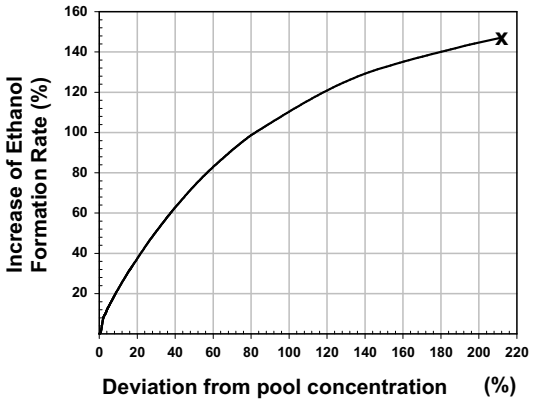


Figure 11: Increase of the Ethanol formation rate as a function of the maximal deviation from initial pool concentrations.

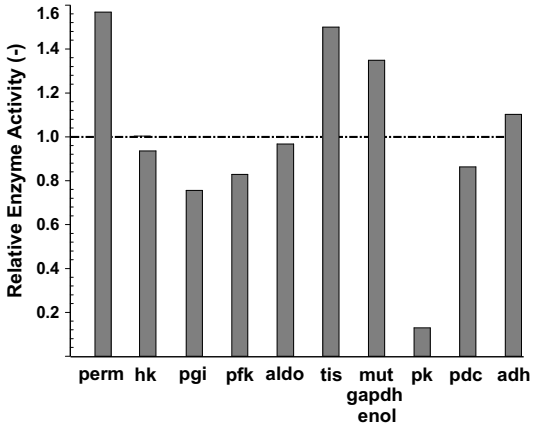


Figure 12: Optimal distribution of enzyme activities. All activities are related to their initial values.

include a change of enzyme concentrations via regulation of enzyme synthesis. Hence, homeostasis puts a further constraint on the metabolic redesign. Mathematically, this could take the form

$$\frac{1}{m}\sum_{i=1}^{m}\frac{\left|c_{i,optimum}^{steady\ state}-c_{i,reference}^{steady\ state}\right|}{c_{i,reference}^{steady\ state}}\leq\Theta. \qquad (27)$$

Finally, the optimum must be constraint to stable steady states, leading to condition (28)

$$\frac{d}{dt}\mathbf{c}=\mathbf{0}\ \text{ and }\ Re\left(\lambda_{i,optimum}\right)\leq 0, \qquad (28)$$

with $Re\left(\lambda_i\right)$ denoting the real part of the system's eigenvalues.

Optimal Solutions

Retaining the total activity at the system's initial state at $\Omega = 1$, non-linear optimization of enzyme activities yielded significantly higher ethanol formation rates (see Figure 11.

As shown in Figure 11, amplification of ethanol formation largely depends on the allowed deviation from the initial pool concentrations Θ. Since the ratios of enzyme levels on the elementary flux mode glucose—ethanol are subject to modifications, no amplification can be expected for $\Theta = 0\%$ while the maximal possible amplification of ethanol formation is found to be as much as 144% at $\Theta = 210\%$. No stable steady states have been detected above this value.

Figure 12 shows the optimal distribution of enzyme activities for the selected example, 40% increase of pool concentrations resulting in an amplification of ethanol formation of 63%.

The optimized modulation of the enzyme activities results in an unexpected and interesting redistribution of enzyme activities. Due to its large share in the control of ethanol formation rate, the activity of the hexose transporter (perm) is found to be significantly higher in the optimized metabolic system. In contrast, even though the control coefficients of tis, gapdh, mut, enol and adh are relatively small, activities of those enzymes are also found to be larger compared to the non-optimized system. This outcome can be explained by the enzymes ability of counteracting an increase of pool concentrations provoked by a risen glucose influx effectively. Obviously, this is the result of the superposition of the three objectives: maximization of flux at more or less homeostatic conditions and unchanged total amount of enzymes.

Concluding Remarks

This paper has presented typical examples of computer aided design problems in Metabolic Engineering. The examples shown refer to the two different characteristics of biological systems: (1) Topological properties and (2) Kinetic properties of the individual reactions.

Topological analysis of metabolic networks turns out to be of immense value for relating genotypes and phenotypes. The further application of the illustrated concept of elementary flux modes critically depends on the development of new and more effective algorithms to treat larger networks. In case of biotechnological production processes, the most important application concerns the prediction of optimal topological properties for maximizing the product yield.

The second example illustrates model based design of an enhanced product formation rate. In addition to information on the network topology, the suggested solution of this important task in Metabolic Engineering requires detailed knowledge on kinetic rate expressions. Armed with a dynamic model for the most important part of the system, it is then possible to apply non-linear

optimization methods. The advantages of this approach are twofold: (1) Effects on large changes in enzyme concentrations can be easily studied and (2) Constraints such as limits on the total amount of enzymes and deviations from steady state metabolite pool concentrations can be taken into account in parallel.

References

Bailey, J. E., "Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities," *Biotechnol. Prog.*, **14**, 8–28 (1998).

Bakker, B. M., F. R. Michels, F. R. Opperdoes, and H. V. Westerhoff, "What controls glycolysis in bloodstream from *Trypanosoma brucei*?," *J. Biol. Chem.*, **274**, 14551–14559 (1999).

Christensen, B. and J. Nielsen, "Metabolic network analysis of *Penicillium chrysogenum* using 13C-labeled glucose," *Biotech. Bioeng.*, **68**, 652–659 (2000).

Dandekar, T., S. Schuster, B. Snel, M. Huynen, and P. Bork, "Pathway alignment: application to the comparative analysis of glycolytic enzymes," *Biochem. J.*, **343**, 115–124 (1999).

Dauner, M. and U. Sauer, "GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing," *Biotechnol. Prog.*, **16**, 642–649 (2000).

Edwards, J. S. and B. O. Palsson, "Systems properties of the *Haemophilus influenzae* Rd metabolic genotype," *J. Biol. Chem.*, **274**(25), 17410–17416 (1999).

Fell, D. A. and J. R. Small, "Fat synthesis in adipose tissue," *Biochem. J.*, **238**, 781–786 (1986).

Happel, J. and P. H. Sellers, "The characterization of complex systems of chemical reactions," *Chem. Eng. Commun.*, **83**, 221–240 (1989).

Heinrich, R. and S. Schuster, *The regulation of cellular systems*. Chapman & Hall, New York (1996).

Kell, D. B. and P. Mendes, Snaphots of systems: Metabolic Control Analysis and biotechnology in the post-genomic era, In *Technological and medical implications of Metabolic Control Analysis*, pages 3–25. Kluwer, Dordrecht (2000).

Liao, J. C., S. Y. Hou, and Y. P. Chao, "Pathway analysis, engineering and physiological considerations for redirecting central metabolism," *Biotech. Bioeng.*, **52**, 129–140 (1996).

Marx, A., A. A. de Graaf, W. Wiechert, L. Eggeling, and H. Sahm, "Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by Nuclear Magnetic Resonance Spectroscopy combined with metabolite balancing," *Biotech. Bioeng.*, **49**, 111–129 (1996).

Mauch, K., S. Vaseghi, and M. Reuss, Quantitative analysis of metabolic and signalling pathways in *Saccharomyces cerevisiae.*, In Schügerl, K. and B. Bellgardt, editors, *Bioreaction Engineering*, pages 435–477. Springer, Berlin (2000).

Palsson, B., "The challenges of in silico biology. Nat. Biotechnol," *Nat. Biotechnol.*, **18**, 1147–1150 (2000).

Reich, J. G. and E. E. Selkov, *Energy metabolism of the cel. A theoretical treatise*. Academic Press, New York (1981).

Rizzi, M., M. Baltes, U. Theobald, and M. Reuss, "*In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II Mathematical model," *Biotech. Bioeng.*, **55**, 592–608 (1997).

Savinell, J. M. and B. O. Palsson, "Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism," *J. Theor. Biol.*, **154**, 421–454 (1992).

Schilling, C. H. and B. O. Palsson, "Assesment of the metabolic capablitites of *Haemophilus influenza* Rd through genome-scale pathway analysis.," *J. Theor. Biol.*, **203**, 249–283 (2000).

Schilling, C. H., S. Schuster, B. O. Palsson, and R. Heinrich, "Metabolic pathway analysis: basic concepts and scientific applications in the post genome era," *Biotechnol. Prog.*, **15**, 296–303 (1999).

Schilling, C. H., D. Letschert, and B. O. Palsson, "Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from pathway-oriented perspective," *J. Theor. Biol.*, **203**, 229–248 (2000).

Schmidt, K., J. Carlsen, J. Nielsen, and J. Villadsen, "Modeling isotopomer distribution in biochemical networks using isotopomer mapping matrices," *Biotech. Bioeng.*, **55**, 831–840 (1997).

Schmidt, K., J. Nielsen, and J. Villadsen, "Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models," *Biotech. Bioeng.*, **71**, 175–190 (1999).

Schuster, S. and C. Hilgetag, "On elementary flux modes in biochemical reaction systems at steady state," *J. Biol. Sys.*, **2**, 165–182 (1994).

Schuster, R. and S. Schuster, "Refined algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed," *CABIOS*, **9**, 79–85 (1993).

Schuster, S., T. Dandekar, and D. A. Fell, "Detection of elemntary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering," *Trends Biotechn.*, **17**, 53–60 (1999).

Schuster, S., D. A. Fell, and T. Dandekar, "A general definition of metabolic patheways useful for systematic organization and analysis of complex metabolic networks," *Nat. Biotechnol.*, **18**, 326–332 (2000b).

Schuster, S., T. Dandekar, K. Mauch, M. Reuss, and D. A. Fell, Recent developments in metabolic pathway analysis and their potential implications for biotechnology and medicine, In *Technological and medical implications of Metabolic Control Analysis*, pages 57–66. Kluwer, Dordrecht (2000a).

Seressiotis, A. and J. E. Bailey, "MPS: an artificially intelligent software system for the analysis and synthesis of metabolic pathways," *Biotech. Bioeng.*, **31**, 587–602 (1988).

Stephanopoulos, G. N., A. A. Aristidou, and J. Nielsen, *Metabolic engineering*. Academic Press, New York (1998).

Szyperski, T., R. W. Glaser, M. Hochuli, and J. Fiaux, "Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional 13C labeling and two-dimensional NMR spectroscopy," *Metabolic Engineering*, pages 189–197 (1999).

Theobald, U., W. Mailinger, M. Baltes, M. Rizzi, and M. Reuss, "*In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations," *Biotech. Bioeng.*, **55**, 305–316 (1997).

van Gulik, W. M. and J. J. Heijnen, "A metabolic network stoichiometry analysis of microbial growth and product formation," *Biotech. Bioeng.*, **48**, 681–698 (1995).

Vanderbei, R. J., *Linear programming: foundations and extensions*. Kluwer, Boston (1998).

Vaseghi, S., A. Baumeister, M. Rizzi, and M. Reuss, "*In vivo* dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*," *Metabolic Engineering*, **1**, 128–140 (1999).

Wiechert, W. and A. A. de Graaf, "Bidirectional reaction steps in metablic networks: I. Modeling and simulation of carbon isotope labeling experiments," *Biotech. Bioeng.*, **55**, 101–117 (1997).

Wiechert, W., C. Siefke, A. A. de Graaf, and A. Marx, "Bidirectional reaction steps in metablic networks: II. Flux estimation and statistical analysis," *Biotech. Bioeng.*, **55**, 118–134 (1997).

Wieczorke, R., S. Krampe, W. T., K. Freidel, C. P. Hollenberg, and E. Boles, "Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexose in *Saccharomyces cerevisiae*," *FEBS Letters*, **464**, 123–128 (1999).

Neuro-Dynamic Programming: An Overview

Dimitri P. Bertsekas*
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139, USA

Abstract

There has been a great deal of research recently on dynamic programming methods that replace the optimal cost-to-go function with a suitable approximation. These methods are collectively known as neuro-dynamic programming or reinforcement learning, and are described in a number of sources, including the books by Bertsekas and Tsitsiklis (1996) and Sutton and Barto (1988). In this paper, we provide an overview of the major conceptual issues, and we survey a number of recent developments, including rollout algorithms which are related to recent advances in model predictive control for chemical processes.

Keywords

Dynamic programming, Neuro-dynamic programming, Reinforcement learning, Optimal control, Suboptimal control

Neuro-dynamic programming (NDP for short) is a relatively new class of dynamic programming methods for control and sequential decision making under uncertainty. These methods have the potential of dealing with problems that for a long time were thought to be intractable due to either a large state space or the lack of an accurate model. They combine ideas from the fields of neural networks, artificial intelligence, cognitive science, simulation, and approximation theory. We will delineate the major conceptual issues, survey a number of recent developments, describe some computational experience, and address a number of open questions.

We consider systems where decisions are made in stages. The outcome of each decision is not fully predictable but can be anticipated to some extent before the next decision is made. Each decision results in some immediate cost but also affects the context in which future decisions are to be made and therefore affects the cost incurred in future stages. Dynamic programming (DP for short) provides a mathematical formalization of the tradeoff between immediate and future costs.

Generally, in DP formulations there is a discrete-time dynamic system whose state evolves according to given transition probabilities that depend on a decision/control u. In particular, if we are in state i and we choose decision u, we move to state j with given probability $p_{ij}(u)$. Simultaneously with this transition, we incur a cost $g(i,u,j)$. In comparing, however, the available decisions u, it is not enough to look at the magnitude of the cost $g(i,u,j)$; we must also take into account how desirable the next state j is. We thus need a way to rank or rate states j. This is done by using the optimal cost (over all remaining stages) starting from state j, which is denoted by $J^*(j)$. These costs can be shown to satisfy some form of Bellman's equation

$$J^*(i) = \min_u E\{g(i,u,j) + J^*(j) \mid i,u\}, \qquad \text{for all } i,$$

where j is the state subsequent to i, and $E\{\cdot \mid i,u\}$ denoted expected value with respect to j, given i and u. Generally, at each state i, it is optimal to use a control u that attains the minimum above. Thus, decisions are ranked based on the sum of the expected cost of the present period, and the optimal expected cost of all subsequent periods.

The objective of DP is to calculate numerically the optimal cost function J^*. This computation can be done off-line, i.e., before the real system starts operating. An optimal policy, that is, an optimal choice of u for each i, is computed either simultaneously with J^*, or in real time by minimizing in the right-hand side of Bellman's equation. It is well known, however, that for many important problems the computational requirements of DP are overwhelming, mainly because of a very large number of states and controls (Bellman's "curse of dimensionality"). In such situations a suboptimal solution is required.

Cost Approximations in Dynamic Programming

NDP methods are suboptimal methods that center around the approximate evaluation of the optimal cost function J^*, possibly through the use of neural networks and/or simulation. In particular, we replace the optimal cost $J^*(j)$ with a suitable approximation $\tilde{J}(j,r)$, where r is a vector of parameters, and we use at state i the (suboptimal) control $\tilde{\mu}(i)$ that attains the minimum in the (approximate) right-hand side of Bellman's equation

$$\tilde{\mu}(i) = \arg\min_u E\{g(i,u,j) + \tilde{J}(j,r) \mid i,u\}.$$

The function $\tilde{J}$ will be called the *scoring function*, and the value $\tilde{J}(j,r)$ will be called the *score* of state j. The general form of $\tilde{J}$ is known and is such that once the parameter vector r is determined, the evaluation of $\tilde{J}(j,r)$ of any state j is fairly simple.

We note that in some problems the minimization over

*bertsekas@lids.mit.edu

u of the expression

$$E\{g(i,u,j) + \tilde{J}(j,r) \mid i,u\}$$

may be too complicated or too time-consuming for making decisions in real-time, even if the scores $\tilde{J}(j,r)$ are simply calculated. In such problems we may use a related technique, whereby we approximate the expression minimized in Bellman's equation,

$$Q(i,u) = E\{g(i,u,j) + J^*(j) \mid i,u\},$$

which is known as the *Q-factor corresponding to* (i,u). In particular, we replace $Q(i,u)$ with a suitable approximation $\tilde{Q}(i,u,r)$, where r is a vector of parameters. We then use at state i the (suboptimal) control that minimizes the approximate Q-factor corresponding to i:

$$\tilde{\mu}(i) = \arg\min_u \tilde{Q}(i,u,r).$$

Much of what will be said about approximation of the optimal cost function also applies to approximation of Q-factors. In fact, we will see later that the Q-factors can also be viewed as optimal costs of a related problem. We thus focus primarily on approximation of the optimal cost function J^*.

We are interested in problems with a large number of states and in scoring functions $\tilde{J}$ that can be described with relatively few numbers (a vector r of small dimension). Scoring functions involving few parameters are called *compact representations*, while the tabular description of J^* are called the *lookup table representation*. Thus, in a lookup table representation, the values $J^*(j)$ are stored in a table for all states j. In a typical compact representation, only the vector r and the general structure of the scoring function $\tilde{J}(\cdot,r)$ are stored; the scores $\tilde{J}(j,r)$ are generated only when needed. For example, $\tilde{J}(j,r)$ may be the output of some neural network in response to the input j, and r is the associated vector of weights or parameters of the neural network; or $\tilde{J}(j,r)$ may involve a lower dimensional description of the state j in terms of its "significant features", and r is the associated vector of relative weights of the features. Thus determining the scoring function $\tilde{J}(j,r)$ involves two complementary issues: (1) deciding on the general structure of the function $\tilde{J}(j,r)$, and (2) calculating the parameter vector r so as to minimize in some sense the error between the functions $J^*(\cdot)$ and $\tilde{J}(\cdot,r)$.

Approximations of the optimal cost function have been used in the past in a variety of DP contexts. Chess playing programs represent a successful example. A key idea in these programs is to use a *position evaluator* to rank different chess positions and to select at each turn a move that results in the position with the best rank. The position evaluator assigns a numerical value to each position, according to a heuristic formula that includes weights for the various features of the position (material balance, piece mobility, king safety, and other factors). Thus, the position evaluator corresponds to the scoring function $\tilde{J}(j,r)$ above, while the weights of the features correspond to the parameter vector r. Usually, some general structure of position evaluator is selected (this is largely an art that has evolved over many years, based on experimentation and human knowledge about chess), and the numerical weights are chosen by trial and error or (as in the case of the champion program Deep Thought) by "training" using a large number of sample grandmaster games.

As the chess program paradigm suggests, intuition about the problem, heuristics, and trial and error are all important ingredients for constructing cost approximations in DP. However, it is important to supplement heuristics and intuition with more systematic techniques that are broadly applicable and retain as much as possible the nonheuristic aspects of DP.

NDP aims to develop a methodological foundation for combining dynamic programming, compact representations, and simulation to provide the basis for a rational approach to complex stochastic decision problems.

Approximation Architectures

An important issue in function approximation is the *selection of architecture*, that is, the choice of a parametric class of functions $\tilde{J}(\cdot,r)$ or $\tilde{Q}(\cdot,\cdot,r)$ that suits the problem at hand. One possibility is to use a neural network architecture of some type. We should emphasize here that in this presentation we use the term "neural network" in a very broad sense, essentially as a synonym to "approximating architecture." In particular, we do not restrict ourselves to the classical multilayer perceptron structure with sigmoidal nonlinearities. Any type of universal approximator of nonlinear mappings could be used in our context. The nature of the approximating structure is left open in our discussion, and it could involve, for example, radial basis functions, wavelets, polynomials, splines, etc.

Cost approximation can often be significantly enhanced through the use of *feature extraction*, a process that maps the state i into some vector $f(i)$, called the *feature vector* associated with the state i. Feature vectors summarize, in a heuristic sense, what are considered to be important characteristics of the state, and they are very useful in incorporating the designer's prior knowledge or intuition about the problem and about the structure of the optimal controller. For example in a queueing system involving several queues, a feature vector may involve for each queue a three-value indicator, that specifies whether the queue is "nearly empty", "moderately busy", or "nearly full". In many cases, analysis can complement intuition to suggest the right features for the problem at hand.

Feature vectors are particularly useful when they can

capture the "dominant nonlinearities" in the optimal cost function J^*. By this we mean that $J^*(i)$ can be approximated well by a "relatively smooth" function $\hat{J}(f(i))$; this happens for example, if through a change of variables from states to features, the function J^* becomes a (nearly) linear or low-order polynomial function of the features. When a feature vector can be chosen to have this property, one may consider approximation architectures where both features and (relatively simple) neural networks are used together. In particular, the state is mapped to a feature vector, which is then used as input to a neural network that produces the score of the state. More generally, it is possible that both the state and the feature vector are provided as inputs to the neural network.

A simple method to obtain more sophisticated approximations, is to partition the state space into several subsets and construct a separate cost function approximation in each subset. For example, by using a linear or quadratic polynomial approximation in each subset of the partition, one can construct piecewise linear or piecewise quadratic approximations over the entire state space. An important issue here is the choice of the method for partitioning the state space. Regular partitions (e.g., grid partitions) may be used, but they often lead to a large number of subsets and very time-consuming computations. Generally speaking, each subset of the partition should contain "similar" states so that the variation of the optimal cost over the states of the subset is relatively smooth and can be approximated with smooth functions. An interesting possibility is to use features as the basis for partition. In particular, one may use a more or less regular discretization of the space of features, which induces a possibly irregular partition of the original state space. In this way, each subset of the irregular partition contains states with "similar features."

Simulation and Training

Some of the most successful applications of neural networks are in the areas of pattern recognition, nonlinear regression, and nonlinear system identification. In these applications the neural network is used as a universal approximator: the input-output mapping of the neural network is matched to an unknown nonlinear mapping F of interest using a least-squares optimization. This optimization is known as *training the network*. To perform training, one must have some training data, that is, a set of pairs $(i, F(i))$, which is representative of the mapping F that is approximated.

It is important to note that in contrast with these neural network applications, in the DP context there is no readily available training set of input-output pairs $(i, J^*(i))$, which can be used to approximate J^* with a least squares fit. The only possibility is to evaluate (exactly or approximately) by simulation the cost functions of given (suboptimal) policies, and to try to iteratively improve these policies based on the simulation outcomes. This creates analytical and computational difficulties that do not arise in classical neural network training contexts. Indeed the use of simulation to evaluate approximately the optimal cost function is a key new idea, that distinguishes the methodology of this presentation from earlier approximation methods in DP.

Using simulation offers another major advantage: it allows the methods of this presentation to be used for systems that are hard to model but easy to simulate; that is, in problems where an explicit model is not available, and the system can only be observed, either as it operates in real time or through a software simulator. For such problems, the traditional DP techniques are inapplicable, and estimation of the transition probabilities to construct a detailed mathematical model is often cumbersome or impossible.

There is a third potential advantage of simulation: it can implicitly identify the "most important" or "most representative" states of the system. It appears plausible that if these states are the ones most often visited during the simulation, the scoring function will tend to approximate better the optimal cost for these states, and the suboptimal policy obtained will perform better.

Neuro-Dynamic Programming

The name *neuro-dynamic programming* expresses the reliance of the methods of this article on both DP and neural network concepts. In the artificial intelligence community, where the methods originated, the name *reinforcement learning* is also used. In common artificial intelligence terms, the methods allow systems to "learn how to make good decisions by observing their own behavior, and use built-in mechanisms for improving their actions through a reinforcement mechanism." In less anthropomorphic DP terms, "observing their own behavior" relates to simulation, and "improving their actions through a reinforcement mechanism" relates to iterative schemes for improving the quality of approximation of the optimal cost function, or the Q-factors, or the optimal policy. There has been a gradual realization that reinforcement learning techniques can be fruitfully motivated and interpreted in terms of classical DP concepts such as value and policy iteration; see the nice survey by Barto et al. (1995), and the book by Sutton and Barto (1988), which point out the connections between the artificial intelligence/reinforcement learning viewpoint and the control theory/DP viewpoint, and give many references.

The currently most popular methodology in NDP iteratively adjusts the parameter vector r of the scoring function $\tilde{J}(j, r)$ as it produces sample state trajectories $(i_0, i_1, \ldots, i_k, i_{k+1}, \ldots,)$ by using simulation. These tra-

jectories correspond to either a fixed stationary policy, or to a "greedy" policy that applies, at state i, the control u that minimizes the expression

$$E\{g(i,u,j) + \tilde{J}(j,r) \mid i,u\},$$

where r is the current parameter vector. A central notion here is the notion of a *temporal difference*, defined by

$$d_k = g(i_k,u_k,i_{k+1}) + \tilde{J}(i_{k+1},r) - \tilde{J}(i_k,r),$$

and expressing the difference between our expected cost estimate $\tilde{J}(i_k,r)$ at state i_k and the predicted cost estimate $g(i_k,u_k,i_{k+1}) + \tilde{J}(i_{k+1},r)$ based on the outcome of the simulation. If the cost approximations were exact, the average temporal difference would be zero by Bellman's equation. Thus, roughly speaking, the values of the temporal differences can be used to make incremental adjustments to r so as to bring about an approximate equality (on the average) between expected and predicted cost estimates along the simulated trajectories. This viewpoint, formalized by Sutton in 1988, can be implemented through the use of gradient descent/stochastic approximation methodology. Sutton proposed a family of methods of this type, called TD(λ), and parameterized by a scalar $\lambda \in [0,1]$. One extreme, TD(1), is closely related to Monte-Carlo simulation and least-squares parameter estimation, while the other extreme, TD(0), is closely related to stochastic approximation. A related method is Q-learning, introduced by Watkins (1989), which is a stochastic approximation-like method that iterates on the Q-factors. While there is convergence analysis of TD(λ) and Q-learning for the case of lookup table representations (see Tsitsiklis, 1994; Jaakkola et al., 1994), the situation is much less clear in the case of compact representations. In our presentation, we will describe results that we have derived for approximate policy and value iteration methods, which are obtained from the traditional DP methods after compact representations of the various cost functions involved are introduced.

A simpler type of methodology for NDP, called *rollout*, is to approximate the optimal cost-to-go by the cost of some reasonably good suboptimal policy, called the *base policy*. Depending on the context, the cost of the base policy may be calculated either analytically, or more commonly by simulation. In a variant of the method, the cost of the base policy is approximated by using some approximation architecture. It is possible to view this approach as a single step of a policy iteration method. The rollout approach is particularly simple to implement, and is also well-suited for on-line replanning, in situations where the problem parameters change over time. The rollout approach may also be combined with rolling horizon approaximations, and in some variations is related to *model predictive control*, and *receding horizon control*; see Keerthi and Gilbert (1988), the surveys by Morari and Lee (1999), and Mayne et al. (2000), and the references quoted there. Despite being less ambitious than the TD and approximate policy iteration methods mentioned earlier, the rollout approach has performed surprisingly well in a variety of studies and applications, often achieving a spectacular improvement over the base policy.

While the theoretical support for the NDP methodology is only now emerging, there have been quite a few reports of successes with problems too large and complex to be treated in any other way. A particularly impressive success is the development of a backgammon playing program as reported by Tesauro (1992). Here a neural network provided a compact representation of the optimal cost function of the game of backgammon by using simulation and TD(λ). The training was performed by letting the program play against itself. After training for several months, the program nearly defeated the human world champion. Variations of the method used by Tesauro have been used with success by us and several other researchers in a variety of applications. In our presentation we will provide some analysis that explains the success of this method, and we will also point to some unanswered questions.

The recent experience of several researchers, involving several engineering applications, has confirmed that NDP methods can be impressively effective in problems where traditional DP methods would be hardly applicable and other heuristic methods would have a limited chance of success. We note, however, that the practical application of NDP is computationally very intensive, and often requires a considerable amount of trial and error. Fortunately, all the computation and experimentation with different approaches can be done off-line. Once the approximation is obtained off-line, it can be used to generate decisions fast enough for use in real time. In this context, we mention that in the machine learning literature, reinforcement learning is often viewed as an "on-line" method, whereby the cost approximation is improved as the system operates in real time. This is reminiscent of the methods of traditional adaptive control. We will not discuss this viewpoint in our presentation, as we prefer to focus on applications involving a large and complex system. A lot of training data is required for such a system. These data typically cannot be obtained in sufficient volume as the system is operating; even if they can, the corresponding processing requirements are typically too large for effective use in real time.

Extensive references for the material of this article are the research monographs by Bertsekas and Tsitsiklis (1996), and by Sutton and Barto (1988). A more limited textbook discussion is given in the DP textbook by Bertsekas (1995a). The 2nd edition of the first volume of this DP text (Bertsekas, 1995b) contains a detailed discussion of rollout algorithms. The extensive survey by Barto et al. (1995), and the overviews by Werbös (1992a,b),

and other papers in the edited volume by White and Sofge (1992) point out the connections between the artificial intelligence/reinforcement learning viewpoint and the control theory/DP viewpoint, and give many references.

References

Barto, A. G., S. J. Bradtke, and S. P. Singh, "Real-Time Learning and Control Using Asynchronous Dynamic Programming," *Artificial Intelligence*, **72**, 81–138 (1995).

Bertsekas, D. P. and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA (1996).

Bertsekas, D. P., *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA (1995a).

Bertsekas, D. P., *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA, 2 edition (1995b).

Jaakkola, T., M. I. Jordan, and S. P. Singh, "On the Convergence of Stochastic Iterative Dynamic Programming Algorithms," *Neural Computation*, **6**, 1185–1201 (1994).

Keerthi, S. S. and E. G. Gilbert, "Optimal Infinite-horizon Feedback Laws for a General Class of Constrained Discrete-time Systems: Stability and Moving-horizon Approximations," *J. Opt. Theory and Appl.*, **57**(2), 265–293 (1988).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and Optimality," *Automatica*, **36**(6), 789–814 (2000).

Morari, M. and J. H. Lee, "Model predictive control: past, present and future," *Comput. Chem. Eng.*, **23**, 667–682 (1999).

Sutton, R. S. and A. G. Barto, *Reinforcement Learning*. MIT Press, Cambridge, MA (1988).

Sutton, R. S., "Learning to Predict by the Methods of Temporal Differences," *Machine Learning*, **3**, 9–44 (1988).

Tesauro, G., "Practical Issues in Temporal Difference Learning," *Machine Learning*, **8**, 257–277 (1992).

Tsitsiklis, J. N., "Asynchronous Stochastic Approximation and Q-Learning," *Machine Learning*, **16**, 185–202 (1994).

Watkins, C. J. C. H., *Learning from Delayed Rewards*, PhD thesis, Cambridge University, England (1989).

Werbös, P. J., Approximate Dynamic Programming for Real-Time Control and Neural Modeling, In White, D. A. and D. A. Sofge, editors, *Handbook of Intelligent Control*. Van Nostrand, NY (1992a).

Werbös, P. J., Neurocontrol and Supervised Learning: an Overview and Valuation, In White, D. A. and D. A. Sofge, editors, *Handbook of Intelligent Control*. Van Nostrand, NY (1992b).

White, D. A. and D. A. Sofge, editors, *Handbook of Intelligent Control*. Van Nostrand, NY (1992).

The Behavioral Approach to Modeling and Control of Dynamical Systems

Jan C. Willems*
Dept. of Electrical Engineering
University of Leuven
Belgium

and

Mathematics Institute
University of Groningen
The Netherlands

Abstract

The behavioral approach provides a mathematical framework for modeling, analysis, and synthesis of dynamical systems. The main difference from the classical view is that it does not the input/output partition as its starting point. In this setting, control is viewed as interconnection.

Keywords

Behaviors, Tearing and zooming, Controllability, Control as interconnection

Introduction

The purpose of this paper is to outline the basics of a mathematical language for the modeling, analysis, and the synthesis of systems. The framework that we present considers the *behavior* of a system as the main object of study. This paradigm differs in an essential way from the input/output paradigm which has dominated the development of the field of systems and control in the 20-th century. This paradigm-shift calls for a reconsideration of many of the basic concepts, of the model classes, of the problem formulations, and of the algorithms in the field.

It is not the purpose to develop mathematical ideas for their own sake. To the contrary, we will downplay mathematical issues of a technical nature. The main aim is to convince the reader that the behavioral framework is a cogent systems-theoretic setting that properly deals with physical systems and that uses modeling as the essential motivation for choosing appropriate mathematical concepts.

It is impossible to do justice to all these aspects in the span of one article. We will therefore concentrate of a few main themes. The behavioral approach is discussed, including the mathematical technicalities, in the recent textbook (Polderman and Willems, 1998), where additional references may be found. We also mention the article (Pillai and Shankar, 1999) where some of these results are generalized to partial differential equations.

The Behavior

The framework that we use for discussing mathematical models views a model as follows. Assume that we have a phenomenon that we wish to model. Nature (that is, the reality that governs this phenomenon) can produce certain events (we will also call them outcomes). The totality of these possible events (*before* we have modeled the phenomenon) forms a set $\mathbb{W}$, called the *universum*. A *mathematical model* of this phenomenon restricts the outcomes that are declared possible to a subset $\mathfrak{B}$ of $\mathbb{W}$; $\mathfrak{B}$ is called the *behavior* of the model. We refer to $(\mathbb{W}, \mathfrak{B})$ (or to $\mathfrak{B}$ by itself, since $\mathbb{W}$ is usually obvious from the context) as a mathematical model.

Examples

1. *The port behavior of an electrical resistor* The outcomes are: pairs (V, I) with V the voltage (say, in volts) across the resistor and I the current (say, in amps) through the resistor. The universum is $\mathbb{R}^2$. After the resistor is modeled, by Ohm's law, the behavior is $\mathfrak{B} = \{(V, I) \in \mathbb{R}^2 | V = RI\}$ with R the value of the resistance (say, in ohms).

2. The *ideal gas law* poses $PV = kNT$ as the relation between the pressure P, the volume V, the number N of moles, and the temperature T of an ideal gas, with k a physical constant. The universum $\mathbb{W}$ is $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{N} \times \mathbb{R}_+$, and the behavior $\mathfrak{B} = \{(P, V, N, T) \in \mathbb{W} | PV = kNT\}$.

In the study of (dynamical) systems we are, more specifically, interested in situations where the events are signals, trajectories, i.e., maps from a set of *independent variables* (time, or space, or time and space) to a set of *dependent variables* (the values taken on by the signals). In this case the universum is the collection of all maps from the set of independent variables to the set of dependent variables. It is convenient to distinguish these sets explicitly in the notation for a mathematical model: $\mathbb{T}$ for the set of independent variables, and $\mathbb{W}$ for the set of dependent variables. $\mathbb{T}$ suggests 'time', but in distributed parameter systems $\mathbb{T}$ is often time and space—we have incorporated distributed systems because of their importance in chemical engineering

*Jan.Willems@esat.kuleuven.ac.be or willems@math.rug.nl

models. Whence we define a *system* as a triple

$$\boxed{\Sigma = (\mathbb{W}, \mathbb{T}, \mathfrak{B})}$$

with $\mathfrak{B}$, *the behavior,* a subset of $\mathbb{W}^{\mathbb{T}}$ ($\mathbb{W}^{\mathbb{T}}$ is the standard mathematical notation for the set of all maps from $\mathbb{T}$ to $\mathbb{W}$). The behavior is the central object in this definition. It formalizes which signals $w : \mathbb{T} \to \mathbb{W}$ are possible, according to the model: those in $\mathfrak{B}$, and which are not: those not in $\mathfrak{B}$.

Examples

1. *Newton's second law* imposes a restriction that relates the position $\vec{q}$ of a point mass with mass m to the force $\vec{F}$ acting on it: $\vec{F} = m\frac{d^2}{dt^2}\vec{q}$. This is a dynamical system with $\mathbb{T} = \mathbb{R}, \mathbb{W} = \mathbb{R}^3 \times \mathbb{R}^3$ (typical elements of $\mathfrak{B}$ are maps $(\vec{q}, \vec{F}) : \mathbb{R} \to \mathbb{R}^3 \times \mathbb{R}^3$), and behavior $\mathfrak{B}$ consisting of all maps $t \in \mathbb{R} \mapsto (\vec{q}, \vec{F})(t) \in \mathbb{R}^3 \times \mathbb{R}^3$ that satisfy $\vec{F} = m\frac{d^2}{dt^2}\vec{q}$. We do not specify the precise sense of what it means that a function satisfies a differential equation (we will pay almost no attention to such secondary issues).
2. One-dimensional *diffusion* describes the evolution of the temperature $T(x,t)$ (with $x \in \mathbb{R}$ position, and $t \in \mathbb{R}$ time) along a uniform bar and the heat $q(x,T)$ supplied to it. Their relation is given by the partial differential equation

$$\frac{\partial}{\partial t}T = \frac{\partial^2}{\partial x^2}T + q$$

   where the constants are assumed to have been chosen appropriately. This defines a system with $\mathbb{T} = \mathbb{R}^2, \mathbb{W} = \mathbb{R}^2$, and $\mathfrak{B}$ consisting of all maps $(T,q) : \mathbb{R}^2 \to \mathbb{R}^2$ that satisfy this partial differential equation.
3. *Maxwell's equations* provide the example of a distributed system with many independent variables. They describe the possible realizations of the electric field $\vec{E} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, the magnetic field $\vec{B} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, the current density $\vec{j} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$, and the charge density $\rho : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$. Maxwell's equations are

$$\begin{aligned}
\nabla \cdot \vec{E} &= \frac{1}{\varepsilon_0}\rho, \\
\nabla \times \vec{E} &= -\frac{\partial}{\partial t}\vec{B}, \\
\nabla \cdot \vec{B} &= 0, \\
c^2 \nabla \times \vec{B} &= \frac{1}{\varepsilon_0}\vec{j} + \frac{\partial}{\partial t}\vec{E},
\end{aligned}$$

   with ε_0 the dielectric constant of the medium and c^2 the speed of light in the medium. This defines the system $(\mathbb{R} \times \mathbb{R}^3, \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}, \mathfrak{B})$, with $\mathfrak{B}$ the set of all fields $(\vec{E}, \vec{B}, \vec{j}, \rho) : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$ that satisfy Maxwell's (partial differential) equations.

These examples fit perfectly our notion of a dynamical system as a triple $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ with $\mathfrak{B} \subseteq \mathbb{W}^{\mathbb{T}}$. In example 1 , the set of independent variables $\mathbb{T}$ is time only, while in the second example, diffusion, and in the third, Maxwell's equations, $\mathbb{T}$ involves time and space. Note that in each of these examples, we are dealing with 'open' systems, that is, systems that interact with their environment (mathematically, systems in which appropriate initial conditions are insufficient to determine the solution uniquely). It has been customary to deal with such systems by viewing them as input/output systems, and by assuming that the input is imposed by the environment. Of course, our first two examples could be thought of as input/output systems. In the case of diffusion, the heat supplied may be thought of as caused by an external heating mechanism that imposes q. But q may also be the consequence of radiation due to the temperature of the bar, making the assumption that it is q that causes the evolution of T untenable, since it is more like T that causes the radiation of heat. It is inappropriate to force Maxwell's equations (where there are clearly free variables in the system: the number of equations, 8, being strictly less than the number of variables, 10) into an input/output setting.

The input/output setting imposes an unnecessary—and unphysical—signal flow structure on our view of systems in interaction with our environment. The input/output point of view has many virtues as a vehicle of studying physical systems, but as a starting point, it is simply inappropriate. First principles laws in physics always state that some events can happen (those satisfying the model equations) while others cannot happen (those violating the model equations). This is a far distance from specifying a system as being driven from the outside by free inputs which together with an initial state specifies the other variables, the outputs. The behavioral framework treats a model for what it is: an exclusion law.

Latent Variables

In the basic equations describing systems, very often other variables appear in addition to those whose behavior the model aims at describing. The origin of these auxiliary variables varies from case to case. They may be state variables (as in automata and input/state/output systems); they may be potentials (as in the well-known expressions for the solutions of Maxwell's equations); most frequently, they are interconnection variables (we will discuss this later). It is important to incorporate these variables in our basic modeling language *ab initio*, and to distinguish clearly between the variables whose behavior the model aims at, and the auxiliary variables introduced in the modeling process. We call the former *manifest* variables, and the latter *latent* variables.

A *mathematical model with latent variables* is defined as a triple $(\mathbb{W}, \mathbb{L}, \mathfrak{B}_{\text{full}})$ with $\mathbb{W}$ the universum of manifest variables, $\mathbb{L}$ the universum of latent variables, and $\mathfrak{B}_{\text{full}} \subseteq \mathbb{W} \times \mathbb{L}$ the *full behavior*. It induces the *manifest model* $(\mathbb{W}, \mathfrak{B})$, with $\mathfrak{B} = \{w \in \mathbb{W} |$ there exists $\ell \in \mathbb{L}$ such that $(w, \ell) \in \mathfrak{B}_{\text{full}}\}$. A *system with latent variables* is defined completely analogously as

$$\boxed{(\mathbb{T}, \mathbb{W}, \mathbb{L}, \mathfrak{B}_{\text{full}})}$$

with $\mathfrak{B}_{\text{full}} \subseteq (\mathbb{W} \times \mathbb{L})^{\mathbb{T}}$. The notion of a system with latent variables is the natural end-point of a modeling process and hence a very natural starting point for the analysis and synthesis of systems. We shall see that latent variables enter also very forcefully in representation questions.

Examples

1. In modeling the port behavior of an *electrical circuit*, the manifest variables are the voltage V across and the current I into the circuit through the port. However, it is usually not possible to come up directly with a model (say in the form of a differential equation) that involves only (V, I). In order to model the port behavior, we usually need to look at the internal structure of the circuit, and introduce the currents through and the voltages across the internal branches as latent variables. Using Kirchhoff's laws and the constitutive equations of the elements in the branches, this readily leads to a latent variable model. Similar situations occur in other systems, for example mechanical systems, and more generally any type of interconnected system.
2. Assume that we want to model the relation between the temperatures and the heat flows radiated at the ends of a uniform bar of length 1. The bar is assumed to be isolated, except at the ends. We wish to model the relation between q_0, T_0, q_1, T_1, the heat flows and temperatures at both ends. In order to obtain a model, it is convenient to introduce the temperature distribution $T(x,t), 0 \leq x \leq 1$, in the bar as latent variables. The full behavior is then described by the partial differential equation

$$\frac{\partial}{\partial t} T = \frac{\partial^2}{\partial x^2} T$$

with the boundary conditions

$$T(0,t) = T_0(t), \qquad T(1,t) = T_1(t),$$
$$\frac{\partial}{\partial x} T(0,t) = -q_0(t), \qquad \frac{\partial}{\partial x} T(1,t) = q_1(t).$$

This defines a latent variable system with $\mathbb{T} = \mathbb{R}, \mathbb{W} = \mathbb{R}^4, \mathbb{L} = \mathfrak{C}^\infty([0,1], \mathbb{R})$, and $\mathfrak{B}_{\text{full}}$ consisting of all maps $((T_0, T_1, q_0, q_1), T) : \mathbb{R} \to \mathbb{R}^4 \times \mathfrak{C}^\infty([0,1], \mathbb{R}))$ such that the above equations are satisfied.

3. In a state model for a dynamical system, the input/output behavior is specified through a system of differential equations as

$$\frac{d}{dt} x = f(x,u), \quad y = h(x,u).$$

This defines a latent variable system $(\mathbb{R}, \mathbb{U} \times \mathbb{Y}, \mathbb{X}, \mathfrak{B}_{\text{full}})$ with $\mathfrak{B}_{\text{full}}$ all trajectories $((u,y),x) : \mathbb{R} \to (\mathbb{U} \times \mathbb{Y}) \times \mathbb{X}$ that satisfy these equations. The manifest behavior is the input/output behavior, that is all trajectories $(u,y) : \mathbb{R} \to \mathbb{U} \times \mathbb{Y}$ that are 'supported' (in the sense made apparent by the full behavior) by a trajectory $x : \mathbb{R} \to \mathbb{X}$.

Situations in which models use latent variables either for mathematical reasons or in order to express the behavioral constraints abound: internal voltages and currents in electrical circuits, momenta in mechanics, chemical potentials, entropy and internal energy in thermodynamics, prices in economics, state variables, the wave function in quantum mechanics in order to explain observables, the basic probability space Ω in probability, etc.

Differential Systems

The 'ideology' that underlies the behavioral approach is the belief that in a model of a dynamical (physical) phenomenon, it is the behavior $\mathfrak{B}$, i.e., a set of possible trajectories $w : \mathbb{T} \to \mathbb{W}$, that is the central object of study. However, as we have seen, in first principles modeling, also latent variables enter *ab initio*. But, the sets $\mathfrak{B}$ or $\mathfrak{B}_{\text{full}}$ of trajectories must be specified somehow, and it is here that differential equations (and difference equations in discrete-time systems) enter the scene. Of course, there are important examples where the behavior is specified in other ways (for example, hybrid systems), but we do not consider these very relevant refinements in the present paper.

For systems described by ODE's (*1-D* systems), with $\mathbb{T} = \mathbb{R}$, and in the case without latent variables, $\mathfrak{B}$ consists of the solutions of a system of differential equations as

$$f_1(w, \frac{d}{dt} w, \ldots, \frac{d^N}{dt^N} w) = f_2(w, \frac{d}{dt} w, \ldots, \frac{d^N}{dt^N} w).$$

We call these *1-D differential systems*. In the case of systems with latent variables these differential equations involve both manifest and latent variables, yielding

$$f_1(w, \frac{d}{dt} w, \ldots, \frac{d^N}{dt^N} w, \ell, \frac{d}{dt} \ell, \ldots, \frac{d^N}{dt^N} \ell) = f_2(w, \frac{d}{dt} w, \ldots, \frac{d^N}{dt^N} w, \ell, \frac{d}{dt} \ell, \ldots, \frac{d^N}{dt^N} \ell),$$

as the equation relating the (vector of) manifest variables w to the (vector of) latent variables ℓ.

Of particular interest (in control, signal processing, circuit theory, econometrics, etc.) are systems with a signal space that is a finite-dimensional vector space and behavior described by linear constant coefficient differential (or difference) equations. Such systems occur not only when the dynamics are linear, but also after linearization around an equilibrium point, when studying the 'small signal behavior'. A *1-D linear time-invariant differential system* is a dynamical system $\Sigma = (\mathbb{R}, \mathbb{W}, \mathfrak{B})$, with $\mathbb{W} = \mathbb{R}^{\mathtt{w}}$ a finite-dimensional (real) vector space, whose behavior consists of the solutions of a system of differential equations of the form

$$R_0 w + R_1 \frac{d}{dt} w + \cdots + R^n \frac{d^n}{dt^n} w = 0,$$

with $R_0, R_1, \ldots, R_n$ matrices of appropriate size that specify the system parameters, and $w = (w_1, w_2, \ldots, w_{\mathtt{w}})$ the vector of (real-valued) system variables. These systems call for polynomial matrix notation. It is convenient to denote the above system of differential equations as

$$\boxed{R(\tfrac{d}{dt})w = 0,}$$

with $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$ a real polynomial matrix with $\mathtt{w}$ columns. The behavior of this system is defined as

$$\mathfrak{B} = \{w : \mathbb{R} \to \mathbb{R}^{\mathtt{w}} \mid R(\frac{d}{dt})w = 0\}.$$

The precise definition of what we consider a solution of $R(\frac{d}{dt})w = 0$ is an issue that we will slide over, but for the results that follow, it is convenient to consider solutions in $\mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{w}})$. Since $\mathfrak{B}$ is the kernel of the differential operator $R(\frac{d}{dt})$, we often write $\mathfrak{B} = \ker(R(\frac{d}{dt}))$, and call $R(\frac{d}{dt})w = 0$ a *kernel representation* of the associated linear time-invariant differential system. We denote the set of differential systems or their behaviors by $\mathfrak{L}^\bullet$, or by $\mathfrak{L}^{\mathtt{w}}$ when the number of variables is $\mathtt{w}$.

Of course, the number of columns of the polynomial matrix R equals the dimension of $\mathbb{W}$. The number of rows of R, which represents the number of scalar differential equations, is arbitrary. In fact, when the row dimension of R is less than its column dimension, as is usually the case, $R(\frac{d}{dt})w = 0$ is an under-determined system of differential equations which is typical for models in which the influence of the environment is taken into account.

In the linear time-invariant case with latent variables, this becomes

$$\boxed{R(\tfrac{d}{dt})w = M(\tfrac{d}{dt})\ell,}$$

with R and M polynomial matrices of appropriate sizes. Define the *full* behavior of this system as

$$\{(w, \ell) : \mathbb{R} \to \mathbb{R}^{\mathtt{w}+\ell} \mid R(\frac{d}{dt})w = M(\frac{d}{dt})\ell\}.$$

Hence the *manifest* behavior of this system is

$$\{w : \mathbb{R} \to \mathbb{R}^{\mathtt{w}} \mid \text{ there exists } \ell : \mathbb{R} \to \mathbb{R}^{\mathtt{l}} \text{ such that } R(\frac{d}{dt})w = M(\frac{d}{dt})\ell\}.$$

We call the $R(\frac{d}{dt})w = M(\frac{d}{dt})\ell$ a *latent variable* representation of the manifest behavior $\mathfrak{B}$.

There is a very extensive theory about these linear differential systems. It is a natural starting point for a theory of dynamical systems. Besides being the outcome of modeling (perhaps after linearization), it incorporates high order differential equations, the ubiquitous first order state systems and transfer function models, implicit (descriptor) systems, etc., as special cases. The study of these systems is intimately connected with the study of polynomial matrices, and may seem somewhat abstract, but this is only because of unfamiliarity. See (Polderman and Willems, 1998) for details.

An important issue that occurs is *elimination:* the question whether the manifest behavior $\mathfrak{B}$ of a latent variable representation belongs to $\mathfrak{L}^{\mathtt{w}}$, i.e., whether it can also be described by a linear constant coefficient differential equation. The following *elimination theorem* holds: For any real polynomial matrices (R, M) with $\text{rowdim}(R) = \text{rowdim}(M)$, there exists a real polynomial matrix R' such that the manifest behavior of $R(\frac{d}{dt})w = M(\frac{d}{dt})\ell$ has the kernel representation $R'(\frac{d}{dt})w = 0$.

The relevance of the elimination problem in object-oriented modeling is as follows. As we will see, a model obtained by tearing and zooming usually involves very many auxiliary (latent) variables and very many equations, among them many algebraic ones originating from the interconnection constraints. The elimination theorem tells us that (for *1-D* linear time-invariant differential systems) the latent variables may be completely eliminated and that the number of equations can be reduced to no more than the number of manifest variables. Of course, the order of the differential equation will go up in the elimination process. We should also mention that there exist very effective, computer algebra based, algorithms for going from a latent variable representation to a kernel representation. The generalization of the elimination theorem and of elimination algorithms to other classes of systems (for example, time-varying or certain classes of nonlinear systems) is a matter of ongoing research. Particularly interesting is the generalization of some of the above concepts and results to systems described by constant coefficient linear *PDE*'s. Define a $\mathtt{n}$-*D* distributed linear shift-invariant differential system as a system $\Sigma = (\mathbb{R}^{\mathtt{n}}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B})$, whose behavior $\mathfrak{B}$ consists of the $(\mathfrak{C}^\infty(\mathbb{R}^{\mathtt{n}}, \mathbb{R}^{\mathtt{d}}))$ solutions of a system of linear constant-coefficient partial differential equations

$$\boxed{R(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_{\mathtt{n}}})w = 0}$$

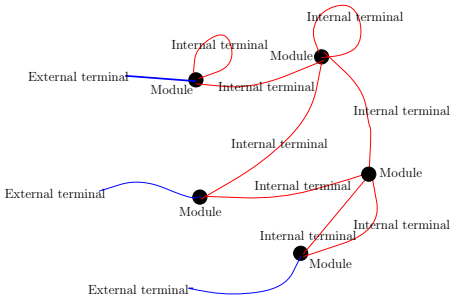


Figure 1: Interconnected system.

viewed as an equation in the w's, in the functions

$$(x_1, \ldots, x_{\mathtt{n}}) = x \in \mathbb{R}^{\mathtt{n}} \mapsto (w_1(x), \ldots, w_{\mathtt{w}}(x)) = w(x) \in \mathbb{R}^{\mathtt{w}}.$$

Here, $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi_1, \ldots, \xi_{\mathtt{n}}]$ is a matrix of polynomials in $\mathbb{R}[\xi_1, \ldots, \xi_{\mathtt{n}}]$, polynomials with real coefficients in $\mathtt{n}$ indeterminates.

For distributed differential systems with latent variables, this leads to equations of the form

$$\boxed{R(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_{\mathtt{n}}})w = M(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_{\mathtt{n}}})\ell,}$$

with R and M matrices of polynomials in $\mathbb{R}[\xi_1, \ldots, \xi_{\mathtt{n}}]$.

It is easy to prove that a *1-D* linear differential system admits an *input/output representation*. This means that for every $\mathfrak{B} \in \mathfrak{L}^{\mathtt{w}}$, there exists a permutation matrix Π and a partition $\Pi w = \mathrm{col}(u, y)$ such that for any $u^* \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{u}})$, there exist a $y \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^{\mathtt{y}})$ such that $(u^*, y) \in \Pi\mathfrak{B}$. Moreover, the y's that such $(u^*, y) \in \Pi\mathfrak{B}$ form a linear finite dimensional variety, implying that such a y is uniquely determined by its derivatives at $t = 0$.

Thus in linear differential systems, the variables can always be partitioned into two groups. The first group act a free inputs, the second group a bound outputs: they are completely determined by the inputs and their initial conditions.

Tearing and Zooming

Systems, especially engineering systems, usually consist of interconnections of subsystems. This feature is crucial in both modeling and design. The aim of this section is to formalize interconnections and to analyze the model structures that emerge from it. The procedure of modeling by *tearing and zooming* is an excellent illustration of the appropriateness of the behavioral approach as the supporting mathematical language. We assume throughout finiteness, i.e., that we interconnect a finite number of modules (subsystems), each with a finite number of terminals, etc. See figure 1.

We view an interconnected system as a collection of *modules* with *terminals*, interconnected through an *interconnection architecture*. The building blocks, called *modules*, of an interconnected system are systems with *terminals*. Each of these terminals carries variables from a universum, and the (dynamical) laws that govern the module are expressed by a behavior that relates the variables at the various terminals. Finally, the terminals of the modules are assumed to be interconnected, expressed by an *interconnection architecture*. The interconnection architecture imposes relations between the variables on these terminals.

After interconnection, the architecture leaves some terminals available for interaction with the environment of the overall system. The behavior of the interconnected system consists of the signals that satisfy both the module behavior laws and the interconnection constraints. In specifying the behavior of an interconnected system, we consider the variables on the interconnected terminals as latent variables, and those on the terminals that are left for interaction with the environment as manifest variables. We will occasionally call the interconnected variables *internal* variables, and the exposed variables *external* variables. It is important to note immediately the hierarchical nature of this procedure. The modules thus become subsystems. The paradigmatic example to keep in mind is an electrical circuit. The modules are resistors, capacitors, inductors, transformers, etc. The terminals are the wires attached to the modules and are electrical terminals, each carrying a voltage (the potential) and a current. The interconnection architecture states how the wires are connected. We now formalize all this, assuming that we are treating continuous time dynamical systems (hence, with time set $\mathbb{T} = \mathbb{R}$). Of course, for process engineering, generalization to distributed systems and to 'distributed' terminals, as in interconnection along surfaces, is mandatory.

A *terminal* is specified by its *type*. Giving the type of a terminal identifies the kind of a physical terminal that we are dealing with. The type of terminal implies a universum of *terminal variables*. These variables are physical quantities that characterize the possible 'signal states' on the terminal, it specifies how the module interacts with the environment through this terminal.

A *module* is specified by its *type*, and its *behavior*. Giving the type of a module identifies the kind of a physical system that we are dealing with. Giving a *behavior specification* of a module implies giving a *representation* and the values of the associated *parameters* a representation. Combined these specify the behavior of the variables on the terminals of the module. The type of a module implies an ordered set of terminals. Since each of the terminals comes equipped with a universum of terminal variables, we thus obtain an ordered set of variables associated with that module. The module behavior then specifies what time trajectories are possible for these variables. Thus a module defines a dynamical system $(\mathbb{R}, \mathbb{W}, \mathfrak{B})$ with $\mathbb{W}$ the Cartesian product over the terminals of the universa of the terminal variables. How-

ever, there are very many ways to specify a behavior (for example, as the solution set of a differential equation, as the image of a differential operator, through a latent variable model, through a transfer function, etc.). The behavioral representation picks out one of these. These representations will then contain unspecified parameters (for example, the coefficients of the differential equation, or the polynomials in a transfer function). Giving the parameter values specifies their numerical values, and completes the specification of the behavior of the signals that are possible on the terminals of a module.

Formally, a system Σ of a given type with T terminals yields $\mathbb{W} = \mathbb{W}_1 \times \mathbb{W}_2 \times \cdots \times \mathbb{W}_T$, with $\mathbb{W}_k$ the universum associated with the k-th terminal. The behavioral specification yields the behavior $\mathfrak{B} \subseteq \mathbb{W}^{\mathbb{R}}$. If $(w_1, w_2, \ldots, w_T) \in \mathfrak{B}$, then we think of $w_k \in (\mathbb{W}_k)^{\mathbb{R}}$ as a signal that can be realized on the k-th terminal.

An interconnected system is composed of modules, its building blocks. They serve as subsystems of the overall system. Each module specifies an ordered set of terminals. By listing the modules, and the associated terminals, we obtain the Cartesian product of all the terminals in the interconnected system. The manner in which these terminals, and hence the associated modules, are interconnected is specified by the *interconnection architecture*. This consists of a set of disjunct pairs of terminals, and it is assumed that each such pair consists of terminals of adapted type. Typical 'adapted' type means that the are the same physical nature (both electrical, or both 1-D mechanical, both thermal, etc. But, when the terminal serves for information processing (inputs to actuators, output of sensors) it could also mean that one variable must be an input to the module to which it is connected (say, the input of an actuator), and the other must be an output to the module to which it is connected (say the output of a sensor). Note that the interconnection architecture involves only the terminals of the modules and their type, but not the behavior. Also, the union of the terminals over the pairs that are part of the interconnection architecture will in general be a strict subset of the union of the terminals of all the modules. We call the terminals that are not involved in the interconnection architecture the *external (or exposed) terminals*. It is along these terminals that the interconnected system can interact with its environment. The terminals that enter in the interconnection architecture are called *internal terminal*. It is along these terminals that the modules are interconnected.

Pairing of terminals by the interconnection architecture implies an *interconnection law*. Some examples of interconnection laws are $V_1 = V_2, I_1 + I_2 = 0$ for electrical terminals, $Q_1 + Q_2 = 0, T_1 = T_2$ for thermal terminals, $p_1 = p_2, f_1 + f_2 = 0$ for fluidic terminals, etc.

The physical examples of interconnection laws all involve equating of 'across' variables and putting the sum of 'through' variables to zero. This is in contrast to the input-output identification for information processing terminals. The latter is actually the only interconnection that is used in flow diagram based modeling, as implemented, for example, in SIMULINK. It is indeed very much based on the input/output thinking that has permeated systems theory and control throughout the past century. Unfortunately, this is of limited interest when it comes to modeling interconnected physical systems. As such the ideas developed in the *bond-graph* literature and the modeling packages that use this philosophy are bound to be much more useful in the long run. Interconnection of physical systems involves across and through variables, efforts and flows, extensive and intensive quantities, and not in first instance flow diagrams. These considerations are the main motivation for the development of the behavioral approach.

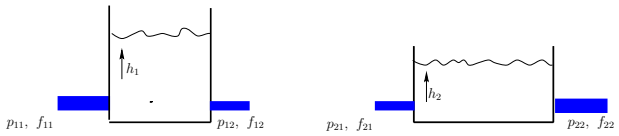


Figure 2: Tanks.

We now formalize the interconnected system. The most effective way to proceed is to specify it as a latent variable system, with as manifest variables the variables associated with the external terminals, and as latent variables the internal variables associated with the terminals that are paired by the interconnection architecture. This latent variable system is specified as follows. Its universum of manifest variables equals $\mathbb{W} = \mathbb{W}_{e_1} \times \cdots \times \mathbb{W}_{e_{|E|}}$, where $E = \{e_1, \ldots, e_{|E|}\}$ is the set of external terminal. Its universum of latent variables equals $\mathbb{L} = \mathbb{W}_{i_1} \times \cdots \times \mathbb{W}_{i_{|I|}}$, where $I = \{i_1, \ldots, i_{|I|}\}$ is the set of internal terminals. Its full behavior behavior consists of the behavior as specified by each of the modules, combined by the interconnection laws obtained by the interconnection architecture. The behavior of each of the modules involves a combination of internal and external variables that are associated with the module. The interconnection law of a pair in the interconnection architecture involves the internal variables associated with these terminals.

Modeling interconnected via the above method of *tearing and zooming* provides the prime example of the usefulness of behaviors and the inadequacy of input/output thinking. Even if our system, after interconnection, allows for a natural input/output representation, it is unlikely that this will be the case of the subsystem and of the interconnection architecture. We illustrate this by means of an example.

Example: Consider two tanks filled with a fluid, both equipped with two tubes through which the fluid can flow in or out (see figure 2). Assume that the pressures (p_{11}, p_{12}) and the flows (f_{11}, f_{12}) at the end of these tubes

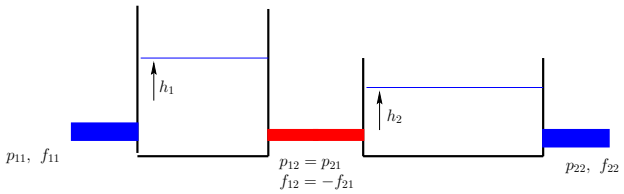


Figure 3: Connected tanks.

of the first tank tank are governed by a differential equation of the form

$$\frac{d}{dt}h_1 = F_1(h_1, p_{11}, p_{12}),$$
$$f_{11} = H_{11}(h_1, p_{11}), \qquad f_{12} = H_{12}(h_1, p_{12}),$$

where h_1 denotes the height of the fluid in the first tank. Similarly for the second tank:

$$\frac{d}{dt}h_2 = F_2(h_2, p_{21}, p_{22}),$$
$$f_{21} = H_{21}(h_2, p_{21}), \qquad f_{22} = H_{22}(h_2, p_{22})$$

It is quite reasonable, by all accounts, to consider in the first system the pressures p_{11}, p_{12} as inputs and the flows f_{11}, f_{12} as outputs, and for for the second system the pressures p_{21}, p_{22} as inputs and the flows f_{21}, f_{22} as outputs. Now assume that we interconnect the tube 12 to 21. This yields the interconnection laws of a fluidic terminal:

$$p_{12} = p_{21}, f_{12} + f_{21} = 0.$$

Note that this comes down to equating two inputs, and equating two outputs. Precisely the opposite that what is supposed to happen in the output-to-input identification that signal flow modeling wants us to do! A similar situation holds in mechanics: interconnection equates two positions (often both outputs), and puts the sum of two forces (often both inputs) equal to zero.

If the field of systems and control wants to take modeling seriously, is should retrace the *faux pas* of taking input/output thinking as the basic framework, and cast models in the language of behaviors. It is only when considering the more detailed signal flow graph structure of a system that input/output thinking becomes useful. Signal flow graphs are useful building blocks for interpreting information processing systems, but physical systems need a more flexible framework.

Controllability and Observability

An important property in the analysis and synthesis of systems is controllability. Controllability refers to be ability of transferring a system from one mode of operation to another. By viewing the first mode of operation as undesired and the second one as desirable, the relevance to control and other areas of applications becomes clear. The concept of controllability has originally

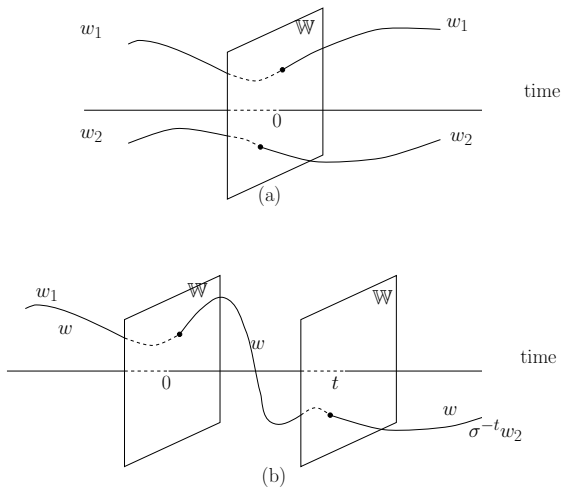


Figure 4: Controllability.

been introduced in the context of state space systems. The classical definition runs as follows. The system described by the controlled vector-field $\frac{d}{dt}x = f(x,u)$ is said to be controllable if for all states a, b, there exists an input$u$ and a time $T \geq 0$ such that the solution to $\frac{d}{dt}x = f(x,u)$ and $x(0) = a$ yields $x(T) = b$. One of the elementary results of system theory states that the finite-dimensional linear system $\frac{d}{dt}x = Ax + Bu$ is controllable if and only if the matrix $[B\ AB\ A^2B \cdots A^{\dim(x)-1}B]$ has full row rank. Various generalizations of this result to time-varying, to nonlinear (involving Lie brackets), and to infinite-dimensional systems exist.

A disadvantage of the notion of controllability as formulated above is that it refers to a particular representation of a system, notably to a state space representation. Thus a system may be uncontrollable either for the intrinsic reason that the control has insufficient influence on the system variables, or because the state has been chosen in an inefficient way. It is clearly not desirable to confuse these reasons. In the context of behavioral systems, a definition of controllability has been put forward that involves the manifest system variables directly.

Let $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ be a dynamical system with $\mathbb{T} = \mathbb{R}$ or $\mathbb{Z}$, and assume that is time-invariant, that is $\sigma^t\mathfrak{B} = \mathfrak{B}$ for all $t \in \mathbb{T}$, where σ^t denotes the t-shift (defined by $(\sigma^t f)(t') := f(t'+t)$); Σ is said to be *controllable* if for all $w_1, w_2 \in \mathfrak{B}$ there exists $T \in \mathbb{T}$, $T \geq 0$ and $w \in \mathfrak{B}$ such that $w(t) = w_1(t)$ for $t < 0$ and $w(t) = w_2(t-T)$ for $t \geq T$. Thus controllability refers to the ability to switch from any one trajectory in the behavior to any other one, allowing some time-delay. This is illustrated in figure 4.

Two questions that occur are the following: What conditions on the parameters of a system representation imply controllability? Do controllable systems admit a particular representation in which controllability becomes apparent? For linear time-invariant differential systems,

these questions are answered in the following theorem. Let $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathbb{w}}, \mathfrak{B})$ be a linear time-invariant differential system. The following are equivalent:

1. $\mathfrak{B} \in \mathfrak{L}^{\mathbb{w}}$ is controllable.
2. The polynomial matrix R in a kernel representation $R(\frac{d}{dt})w = 0$ of $\mathfrak{B} \in \mathfrak{L}^{\mathbb{w}}$ satisfies $\text{rank}(R(\lambda)) = \text{rank}(R)$ for all $\lambda \in \mathbb{C}$.
3. The behavior $\mathfrak{B} \in \mathfrak{L}^{\mathbb{w}}$ is the image of a linear constant-coefficient differential operator, that is, there exists a polynomial matrix $M \in \mathbb{R}^{\mathbb{w} \times \bullet}[\xi]$ such that $\mathfrak{B} = \{w \mid w = M(\frac{d}{dt})\ell \text{ for some } \ell\}$.

There exist various algorithms for verifying controllability of a system starting from the coefficients of the polynomial matrix R in a kernel (or a latent variable) representation of Σ.

A point of the above theorem that is worth emphasizing is that, as stated in the above theorem, controllable systems admit a representation as the manifest behavior of the latent variable system of the special form

$$\boxed{w = M(\tfrac{d}{dt})\ell.}$$

We call this an *image* representation of the system with behavior

$$\mathfrak{B} = \{w \mid \text{ there exists } \ell \text{ such that } w = M(\frac{d}{dt})\ell\}.$$

It follows from the elimination theorem that every system in image representation can be brought in kernel representation. But not every system in kernel representation can be brought in image representation: it are precisely the controllable ones for which this is possible.

The controllability question has been pursued for many other classes of systems. In particular (more difficult to prove) generalizations have been derived for differential-delay (Rocha and Willems, 1997; Glüsing-Lüerssen, 1997), for nonlinear systems, and, as we will discuss soon, for *PDE's*. Systems in an image representation have received much attention recently for nonlinear differential-algebraic systems, where they are referred to as *flat* systems (Fliess and Glad, 1993). Flatness implies controllability, but the exact relation remains to be studied.

The notion of observability is always introduced hand in hand with controllability. In the context of the input/state/output system $\frac{d}{dt}x = f(x,u), y = h(x,u)$, it refers to the possibility of deducing, using the laws of the system, the state from observation of the input and the output. The definition that is used in the behavioral context is more general in that the variables that are observed and the variables that need to be deduced are kept general.

In observability, we ask the question: Can the trajectory w_1 be deduced from the trajectory w_2? (See figure 5). Let $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ be a dynamical system, and assume that $\mathbb{W}$ is a product space: $\mathbb{W} = \mathbb{W}_1 \times \mathbb{W}_2$. . Then w_1 is said to be *observable* from w_2 in Σ if $(w_1, w_2') \in \mathfrak{B}$ and $(w_1, w_2'') \in \mathfrak{B}$ imply $w_2' = w_2''$. Observability thus refers to the possibility of deducing the w_1 from observation of w_2 and from the laws of the system ($\mathfrak{B}$ is assumed to be known).

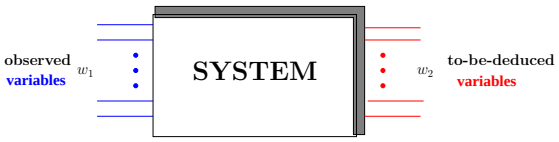


Figure 5: Observability.

The theory of observability runs parallel to that of controllability. We mention only the result that for linear time-invariant differential systems, w_1 is observable from w_2 if and only if there exists a set of differential equations satisfied by the behavior of the system (i.e., a set of consequences) of the following form, that puts observability into evidence: $w_1 = R_2'(\frac{d}{dt})w_2$. This condition is again readily turned into a standard problem in computer algebra.

Many of the results for controllability and observability have recently been generalized to distributed systems (Pillai and Shankar, 1999). We mention them briefly. The system $\mathfrak{B} \in \mathfrak{L}_{\mathbb{n}}^{\mathbb{w}}$ is said to be *controllable* if for all $w_1, w_2 \in \mathfrak{B}$ and for all open non-overlapping subsets $O_1, O_2 \subseteq \mathbb{R}^{\mathbb{n}}$, there exists $w \in \mathfrak{B}$ such that $w|_{O_1} = w_1|_{O_1}$ and $w|_{O_2} = w_2|_{O_2}$, i.e., if its solutions are 'patch-able'.

Note that it follows from the elimination theorem for $\mathfrak{L}_{\mathbb{n}}^{\bullet}$ that the manifest behavior of a system in image representation, i.e., a latent variable system of the special form

$$\boxed{w = M(\tfrac{\partial}{\partial x_1}, \ldots, \tfrac{\partial}{\partial x_{\mathbb{n}}})\ell}$$

can be described as the solution set of a system of constant coefficient *PDE's*. Whence, every image of a constant coefficient linear partial differential operator is the kernel of a constant coefficient linear partial differential operator. However, not every kernel of a constant coefficient linear partial differential operator is the image of a constant coefficient linear partial differential operator. It turns out that it are precisely the controllable systems that admit an image representation (Pillai and Shankar, 1999).

Note that an image representation corresponds to what in mathematical physics is called a *potential function*, with ℓ the potential and $M(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_{\mathbb{n}}})$ the partial differential operator that generates elements of the behavior from the potential. An interesting aspect of the above theorem therefore is the fact that it identifies the existence of a potential function with the system theoretic property of controllability (patch-ability of trajectories in the behavior).

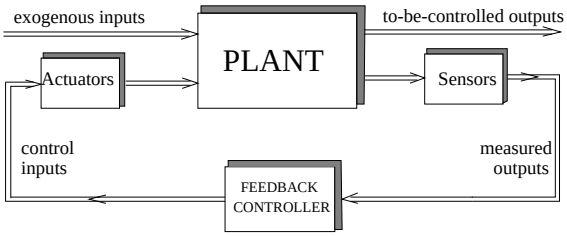


Figure 6: Intelligent control.

It can be shown that Maxwell's equations define a controllable distributed differential system. Indeed, in the case of Maxwell's equations, there exists a well-known image representation using the scalar and vector potential.

Control

The behavioral point of view has received broad acceptance as an approach for modeling dynamical systems. It is generally agreed upon that when modeling a dynamic component it makes no sense to prejudice oneself (as one would be forced to do in a transfer function setting) as to which variables should be viewed as inputs and which variables should be viewed as outputs. We have argued this point extensively throughout the previous sections of this paper. This is not to say, by any means, that there are no situations where the input/output structure is natural. Quite to the contrary. In fact, whenever logic devices are involved, in information processing, the input/output structure is often a must.

The behavioral approach has, until now, met with much less acceptance in the context of control. We can offer a number of explanations for this fact. Firstly, there is something very natural in viewing control variables as inputs and measured variables as outputs. Control then becomes decision making on the basis of observations. When subsequently a controller is regarded as a feedback processor, one ends up with the feeling that the input/output structure is in fact an essential feature of control. Secondly, since, as mentioned in a previous section, it is possible to prove that every linear time-invariant system admits a component-wise input/output representation, one gets the impression that the input/output framework can be adopted without second thoughts, that nothing is lost by taking it as a universal starting point.

Present-day control theory centers around the signal flow graph shown in figure 6. The plant has four terminals (supporting variables which will typically be vector-valued). There are two input terminals, one for the control, one for the other exogenous variables (disturbances, set-points, reference signals, tracking inputs, etc.) and there are two output terminals, one for the measurements, and one for the to-be-controlled variables. By using feed-through terms in the plant equations this configuration accommodates, by incorporating these variables in the measurements, for the possible availability to the controller of set-point settings, reference signals, or disturbance measurements for feed-forward control, and, by incorporating the control input in the to-be-controlled outputs, for penalizing excessive control action. The control inputs are generated by means of actuators and the measurements are made available through sensors. Usually, the dynamics of the actuators and of the sensors are considered to be part of the plant. We call this structure *intelligent control*. In intelligent control, the controller is thought of as a micro-processor type device which is driven by the sensor outputs and which produces the actuator inputs through an algorithm involving a combination of feedback, identification, and adaptation. Also, often loops expressing model uncertainty are incorporated in the above as well. Of course, many variations, refinements, and special cases of this structure are of interest, but the basic idea is that of supervisor reacting in an intelligent way to observed events and measured signals.

The paradigm embodied in figure 6 has been universally prevalent ever since the beginning of the subject, from our interpretation of the Watt regulator, Black's feedback amplifier, and Wiener's cybernetics, to the ideas underlying modern adaptive and robust control. It is indeed a deep and very appealing paradigm, which will undoubtedly gain in relevance and impact as logic devices become ever more prevalent, reliable, and inexpensive. This paradigm has a number of features which are important for considerations which will follow. Some of these are:

1. There is an asymmetry between the plant and the controller: it remains apparent what part of the system is the plant and what part is the controller. This asymmetry disappears to some extent in the closed loop.

2. The intelligent control paradigm tells us to be wary of errors in the measurements. Thus it is considered as being ill-advised to differentiate measurements, presumably, because this will lead to noise amplification.

3. The plant and the controller are dynamical systems which can be interconnected at any moment in time. If for one reason or another the feedback controller temporarily fails to receive an input signal, then the control input can be set to a default value, and later on the controller can resume its action.

We will now analyze an example of a down-to-earth controller, a very wide-spread automatic control mechanism, namely the traditional device which ensures the automatic closing of doors. There is nothing peculiar

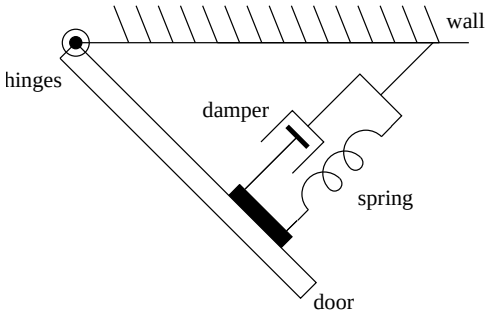


Figure 7: Door-closing mechanism.

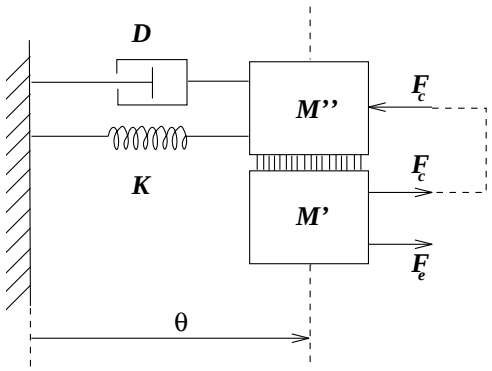


Figure 8: Spring/damper representation.

about this example. Devices based on similar principles are used for the suspension of automobiles and the points which we make through this example could also be made just as well through many temperature or pressure control devices. A typical automatic door-closing mechanism is schematically shown in figure 7.

A door-closing mechanism usually consists of a spring to ensure the closing of the door and a damper in order to make sure that it closes gently. In addition, these mechanisms often have considerable weight so that their mass cannot be neglected as compared to the mass of the door itself. The automatic door-closing mechanism can be modeled as a mass/spring/damper combination. In good approximation, the situation can be analyzed effectively as the mechanical system shown in figure 8. We model the door as a mass M', on which, neglecting friction in the hinges, two forces act. The first force, F_c, is the force exerted by the door-closing device, while the second force, F_e, is the exogenous force exerted for example by a person pushing the door in order to open it. The equation of motion for the door becomes

$$M' \frac{d^2}{dt^2}\theta = F_c + F_e,$$

where θ denotes the opening angle of the door. The automatic door-closing mechanism, modeled as a mass/spring/damper combination, yields

$$M'' \frac{d^2}{dt^2}\theta + D\frac{d}{dt}\theta + K\theta = -F_c.$$

Here, M'' denotes the mass of the door-closing mechanism, D its damping coefficient, and K its spring constant. Combining these equations yields

$$(M' + M'') \frac{d^2}{dt^2}\theta + D\frac{d}{dt}\theta + K\theta = F_e.$$

In order to ensure proper functioning of the door-closing device, the designer can to some extent choose M'', D and K (all of which must, for physical reasons, be positive). The desired response requirements are: small overshoot (to avoid banging of the door), fast settling time, and a reasonably high steady state gain (to avoid having to exert excessive force when opening the door). This is an example of an elementary control design exercise. A good design will be achieved by choosing a light mechanism (M'' small), with a reasonably strong spring (K large), but not too strong so as to avoid having to use excessive force in order to open the door, and with the value of D chosen so as to achieve slightly less than critical damping (ensuring gentle closing of the door).

It is completely natural to view in this example the door as the plant and the door-closing mechanism as the controller. Then, if we insist on interpreting this plant/controller combination in terms of control system configurations as figure 6, we obtain.

$$\text{Plant:} \quad M' \frac{d^2}{dt^2}\theta = u + v;\ y = \theta;\ z = \theta$$

with u the control input ($u = F_c$), v the exogenous input ($v = F_e$), y the measured output, and z the to-be-controlled output.

$$\text{Controller:} \quad u = -M'' \frac{d^2}{dt^2} y - D\frac{d}{dt} y - Ky.$$

This yields the controlled system, described by:

$$\text{Controlled system:} \quad (M' + M'') \frac{d^2}{dt^2} z + D\frac{d}{dt} z + Kz = v.$$

Observe that in the control law, the measurement y should be considered as the input, and the control u should be considered as the output. This suggests that we are using what would be called a *PDD-controller* (a proportional and twice differentiating controller), a singular controller which would be thought of as causing high noise amplification. Of course, no such noise amplification occurs in reality. Further, the plant is second order, the controller is second order, and the closed loop system is also second order (unequal to the sum of the order of the plant and the controller). Hence, in order to connect the controller to the plant, we will have to 'prepare' the initial states of the controller and the plant. Indeed, in attaching the door-closing mechanism to the

door, we will make sure that at the moment of attachment the initial values of θ and $\frac{d}{dt}\theta$ are zero *both* for the door and the door-closing mechanism.

We now come to our most important point concerning this example. Let us analyze the signal flow graph. In the plant, it is natural to view the forces F_c and F_e as inputs and θ as output. This input/output choice is logical both from the physical and from the cybernetic, control theoretic point of view. In the controller, on the other hand, the physical and the cybernetic points of view clash. From the cybernetic, control theoretic point of view, it is logical to regard the opening angle θ as input and the control force F_c as output. From the physical point of view, however, it is logical to regard (just as in the plant) the force F_c as input and θ as output. It is evident that as an interconnection of two mechanical systems, the door and the door-closing mechanism play completely symmetric roles. However, the cybernetic, control theoretic point of view obliges us to treat the situation as asymmetric, making the force the cause in one mechanical subsystem, and the effect in another.

In our opinion, this simple but realistic example permits us to draw the following conclusions. Notwithstanding all its merits, the intelligent control paradigm of figure 6 gives an unnecessarily restrictive view of control theory. In many practical control problems, the signal-flow-graph interpretation of figure 6 is untenable. The solution which we propose to this dilemma is the following. We will keep the distinction between plant and controller with the understanding that this distinction is justified only from an *evolutionary* point of view, in the sense that it becomes evident only *after* we comprehend the genesis of the controlled system, after we understand the way in which the closed loop system has come into existence as a purposeful system. However, we will abandon the intelligent control signal flow graph as a paradigm for control. We will abandon the distinction between control inputs and measured outputs. Instead, we will view *interconnection of a controller to a plant* as the central paradigm in control theory. However, we by no means claim that the intelligent control paradigm is without merits. To the contrary, it is extremely useful and important. Claiming that the input/output framework is *not always* the suitable framework to approach a problem does not mean that one claims that it is *never* the suitable framework. However, a good universal framework for control should be able to deal both with interconnection, with designing subsystems, and with intelligent control. The behavioral framework does, the intelligent control framework does not.

In order to illustrate the nature of control that we would like to transmit in this presentation, consider the system configuration depicted in figure 9. In the top part of the figure, there are two systems, shown as black-boxes with terminals. It is through their terminals that

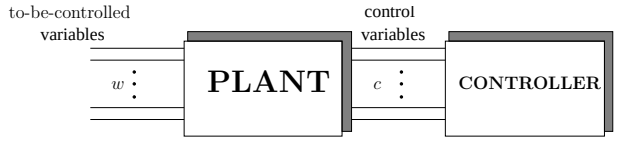


Figure 9: Control as interconnection.

systems interact with their environment. The black-box imposes relations on the variables that 'live' on its terminals. These relations are formalized by the behavior of the system in the black-box. The system to the left in figure 9 is called the *plant*, the one to the right the *controller*. The terminals of the plant consist of *to-be-controlled variables* w, and *control variables* c. The controller has only terminals with the control variables c. In the bottom part of the figure, the control terminals of the plant and of the controller are connected. Before interconnection, the variables w and c of the plant have to satisfy the laws imposed by the plant behavior. But, after interconnection, the variables c also have to satisfy the laws imposed by the controller. Thus, after interconnection, the restrictions imposed on the variables c by the controller will be transmitted to the variables w. Choosing the black-box to the right so that the variables w have a desirable behavior in the interconnected black-box is, in our view, the basic problem of control. This point of view is discussed with examples in (Polderman and Willems, 1998).

This leads to the following mathematical formulation. The *plant* and the *controller* are both dynamical systems $\Sigma_{plant} = (\mathbb{T}, \mathbb{W} \times \mathbb{V}, \mathfrak{B}_{plant})$ and $\Sigma_{controller} = (\mathbb{T}, \mathbb{V}, \mathfrak{B}_{controller})$ where $\mathbb{W}$ denotes the signal space of the to-be-controlled variables, $\mathbb{V}$ denotes the signal space of control variables, and both systems are assumed to have the same time axis $\mathbb{T}$. The interconnection of Σ_{plant} and $\Sigma_{controller}$ leads to the system $\Sigma_{full} = (\mathbb{T}, \mathbb{W}, \mathbb{V}, \mathfrak{B}_{\text{full}})$ which is a system with latent variables ($\mathbb{V}$) and full behavior defined by

$$\mathfrak{B}_{\text{full}} = \{(w,c) : \mathbb{T} \to \mathbb{W} \times \mathbb{V} \mid (w,c) \in \mathfrak{B}_{plant} \text{ and } c \in \mathfrak{B}_{controller}\}$$

The manifest system obtained by Σ_{full} is the *controlled system* and is hence defined as $\Sigma_{controlled} = (\mathbb{T}, \mathbb{W}, \mathfrak{B}_{controlled})$ with

$$\mathfrak{B}_{controlled} = \{w : \mathbb{T} \to \mathbb{W} \mid \text{ there exists } c : \mathbb{T} \to \mathbb{V} \text{ such that} \in \mathfrak{B}_{plant} \text{ and } c \in \mathfrak{B}_{controller}\}$$

The notion of a controller put forward by the above view considers interconnection as the basic idea of control. The special controllers that consist of sensor-outputs to actuator-inputs signal processors emerge as (very important) special cases. We think of these as controllers as *feedback*, or *intelligent*, controllers. However, our view of control as the design of suitable subsys-

tems greatly enhances the applicability of control, since it views control as integrated subsystem design.

A question that arises in this context is the following. Assume that Σ_{plant} is given. What systems $\Sigma_{controlled}$ can be obtained by suitably choosing $\Sigma_{controller}$? This question can be answered very explicitly, at least for linear time-invariant differential systems. Assume that the plant is given by $\Sigma_{plant} = (\mathbb{R}, \mathbb{R}^{\mathtt{w}} \times \mathbb{R}^{\mathtt{c}}, \mathfrak{B}_{plant}) \in \mathfrak{L}^{\mathtt{w}+\mathtt{c}}$ with $\mathfrak{B}_{plant}$ described by a system of linear constant differential equations $R(\frac{d}{dt})w = R(\frac{d}{dt})c$. Let $\Sigma_{controller} = (\mathbb{R}, \mathbb{R}_c, \mathfrak{B}_{controller}) \in \mathfrak{L}^{\mathtt{c}}$, and assume that $\mathfrak{B}_{controller}$ is similarly described by a system of linear constant coefficient differential equations $C(\frac{d}{dt})c = 0$. Then, by elimination theorem $\Sigma_{controlled} = (\mathbb{R}, \mathbb{R}^{\mathtt{w}}, \mathfrak{B}_{controlled})$ has also a behavior that is described by a system of linear constant coefficient differential equations. It turns out that the behaviors $\mathfrak{B}_{controlled}$ that can be obtained this way can be characterized very nicely.

Define therefore two behaviors, $\mathcal{P}$ and $\mathcal{N}$; $\mathcal{P}$ is called the *realizable (plant) behavior* and $\mathcal{N}$ the *hidden behavior*. They are defined as follows: $\mathcal{P}$ is the manifest behavior of the system, i.e.,

$$\boxed{\mathcal{P} = \{w : \mathbb{R} \to \mathbb{R}^{\mathtt{w}} \mid \text{ there exists } c : \mathbb{R} \to \mathbb{R}^{\mathtt{c}} \text{ such that } (w, c) \in \mathfrak{B}_{\text{full}}\},}$$

and $\mathcal{N}$ is defined as

$$\boxed{\mathcal{N} = \{w : \mathbb{R} \to \mathbb{R}^{\mathtt{w}} \mid (w, 0) \in \mathfrak{B}_{plant}\}.}$$

Hence $\mathcal{N}$ is the behavior of the plant variables that are compatible with the control variables equal to zero. We say that $\mathfrak{B}_{controller}$ *implements* $\mathfrak{B}_{controlled}$ if there exists a controller such that the controlled behavior after interconnecting the controller with behavior $\mathfrak{B}_{controller}$ to the plant, yields $\mathfrak{B}_{controlled}$ as the controlled behavior.

The controller implementability problem asks what behaviors $\mathfrak{B}_{controlled}$ can be obtained this way. For linear time-invariant systems it is possible to prove that $\mathfrak{B}_{controlled}$ is implementable if and only if

$$\boxed{\mathcal{N} \subseteq \mathfrak{B}_{controlled} \subseteq \mathcal{P}.}$$

This result shows that the behaviors that are implementable by means of a controller are precisely those that are wedged in between the hidden behavior $\mathcal{N}$ and the realizable plant $\mathcal{P}$. The problem of control can therefore be reduced to finding such a behavior. Of course, the issue of how to construct $\Sigma_{controller}$ (for example, as a signal processor from the sensor outputs to the actuator inputs) must be addressed as well, but this can be done. In (Willems and Trentelman, 2002) this approach is used for the design of $\mathcal{H}_\infty$-controllers, and we discuss several results on the implementability by feedback controllers as well.

We believe that the point of view of control that emerges from this, as designing a subsystem (with feedback control as a special case) greatly enhances the scope and applicability of control as a discipline. In this setting, control comes down to sub-system design.

Conclusions

In this paper, we have covered some highlights of the behavioral approach to systems and control. We view a mathematical model as a subset of an universum. However, in engineering applications, models are invariably obtained by interconnecting subsystems. This leads to the presence in mathematical models of manifest variables (the variables whose behavior the model aims at) and latent variables (the auxiliary variables introduced in the modeling process). The central object in behavioral systems theory is a system with latent variables.

The concept of controllability becomes an intrinsic systems property related to patch-ability of system trajectories. In the context of latent variable systems, observability refers to the possibility of deducing the latent variables in a system from observation of the manifest variables. In this way, these important concepts are extended far beyond their classical state space setting.

We view control as the design of a subsystem in an interconnected system, a subsystem that interacts with the plant through certain pre-specified variables, the control variables. For a linear time-invariant differential plant, it is possible to prove that a behavior is implementable by a linear time-invariant controller if and only if its behavior is wedged in between the hidden behavior and the realizable plant behavior.

The pre-occupation of systems and control with input/output systems does not do proper justice to the nature of physical systems: most physical systems are simply not a signal processors. Notwithstanding the importance of signal processors, the universal view of a system as an input/output device is simply a *faux pas*. And an unneccesary one at that: the behavioral approach offers a viable alternative.

References

Fliess, M. and S. T. Glad, An algebraic approach to linear and nonlinear control, In Trentelman, H. L. and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and Its Applications*, pages 223–267. Birkhaäuser (1993).

Glüsing-Lüerssen, H., A behavioral approach to delay-differential systems, *SIAM Journal on Control and Optimization*, Volume 35, 480-499 (1997).

Pillai, H. K. and S. Shankar, A behavioral approach to control of distributed systems, *SIAM Journal on Control and Optimization*, Volume 37, 388-408 (1999).

Polderman, J. W. and J. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer-Verlag (1998).

Rocha, P. and J. C. Willems, Behavioral controllability of delay-differential Systems, *SIAM Journal on Control and Optimization*, Volume 35, 254-264 (1997).

Willems, J. C. and H. L. Trentelman, "Synthesis of dissipative systems using quadratic differential forms," *IEEE Trans. Auto. Cont.*, **47**(3) (2002).

Input to State Stability and Related Notions

Eduardo D. Sontag*
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903

Abstract

The analysis of input/output stability is one of the fundamental issues in control theory. External inputs might represent disturbances, estimation errors, or tracking signals, and outputs may correspond to the entire state, or to a more general quantity such as a tracking or regulation error, or the distance to a target set of states such as a desired periodic orbit.

For linear systems, one characterizes i/o stability through finiteness of gains (operator norms). A nonlinear generalization is provided by *input to state stability (ISS)*. This paper summarizes some of the main theoretical results concerning ISS and related notions such as integral ISS (energy bounds), output to input stability (a notion of "minimum-phase" system), and input/output to state stability (a notion of detectability). It also describes, as an illustrative application, an observer design methodology for certain kinetic networks which is based on ISS ideas.

Keywords

Input-to-state stability, ISS, Detectability, Observers, Minimum-phase, Lyapunov functions, Dissipation inequalities, H_∞ control

Introduction

Analyzing how external signals influence system behavior is one of the fundamental issues in control theory. In particular, a central concern is that of input/output stability, that is, stability from inputs u to outputs y in a system.

$$u(\cdot) \;\rightarrow\; \boxed{\;x(\cdot)\;} \;\rightarrow\; y(\cdot)$$

Inputs u might represent disturbances, estimation errors, or tracking signals, while outputs y may correspond to the entire state, or a more general quantity such as a tracking or regulation error, or the distance to a target set of states such as a desired periodic orbit.

The classical approach to i/o stability questions, for linear systems, relies upon transfer functions, which are closely related to more "modern" formulations in terms of operator norms (H_∞ control and the like). However, these approaches have a limited utility when used in a nonlinear context. A new paradigm which emerged within the last 10 or so years, for understanding input/output stability for general nonlinear systems, is that of *input to state stability (ISS)*.

This paper summarizes some of the main theoretical results concerning ISS and related notions such as input to output stability (IOS), integral ISS (iISS, which deals with "energy," as opposed to uniform bounds), mixed notions of integral and uniform stability, output to input stability (which is a notion of "minimum-phase" system), and input/output to state stability (a notion of detectability).

Also described are some illustrative applications, including an observer design methodology for kinetic networks based on ISS ideas.

The paper is written in a very informal style, and readers should consult the references for precise statements and proofs of results.

*Research supported in part by US Air Force Grant F49620-98-1-0242. URL: `http://www.math.rutgers.edu/\protect\nobreakspace{}sontag`, email: `sontag@control.rutgers.edu`

Input to State Stability

There are two desirable, and complementary, features of stability from inputs u to outputs y, one asymptotic and the other one on transients:

- *asymptotic stability*, which can be summarized by the implication "u small $\Rightarrow$ y small," and
- *small overshoot*, which imposes a boundedness constraint on the behavior of internal states x.

Of course, "small" and "boundedness" must be precisely defined, and to that goal we turn next. In order to explain the main ideas as simply as possibly, we begin with the case when the output y is the full state x (which will lead us to "input to state stability"); later we explain the general case ("input to output stability"). In addition, and also in the interest of exposition, we deal with notions relative to equilibria, instead of stability with respect to more arbitrary attractors.

We start by recalling the basic concept of internal stability for linear systems

$$\dot{x} = Ax + Bu\,, \;\; y = Cx\,. \tag{1}$$

Internal stability means that A is a Hurwitz matrix, i.e., $x(t) \to 0$ as $t \to +\infty$ for all solutions of $\dot{x} = Ax$, or equivalently, that $x(t) \to 0$ whenever $u(t) \to 0$. For internally stable systems, one has the explicit estimate

$$|x(t)| \;\leq\; \beta(t)|x^0| + \gamma\,\|u\|_\infty$$

where

$$\beta(t) = \left\|e^{tA}\right\| \to 0 \;\text{ and }\; \gamma = \|B\| \int_0^\infty \left\|e^{sA}\right\| ds$$

and $\|u\|_\infty$ = (essential) sup norm of u restricted to $[0,t]$. For t large, $x(t)$ is bounded by $\gamma\|u\|_\infty$, independently of initial conditions; for small t, the effect of initial states may dominate. Note the superposition of transient and asymptotic effects. Internal stability will now be generalized to "ISS," with the linear functions of $|x^0|$ and $\|u\|_\infty$ replaced by nonlinear ones.

We consider systems of the form

$$\dot{x} \;=\; f(x,u)\,,\;\; y \;=\; h(x)$$

evolving in finite-dimensional spaces $\mathbb{R}^n$, and we suppose that inputs u take values in $\mathbb{R}^m$ and outputs y are $\mathbb{R}^p$-valued. An *input* is a Lebesgue-measurable locally essentially bounded $u(\cdot):[0,\infty)\to\mathbb{R}^m$. We employ the notation $|x|$ for Euclidean norms, and use $\|u\|$, or $\|u\|_\infty$ for emphasis, to indicate the essential supremum of a function $u(\cdot)$, usually (depending on context) restricted to an interval of the form $[0,t]$. The map $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ is locally Lipschitz and satisfies $f(0,0)=0$. The map $h:\mathbb{R}^n\to\mathbb{R}^p$ is locally Lipschitz and satisfies $h(0)=0$.

The internal stability property for linear systems amounts to the "$L^\infty\to L^\infty$ finite-gain condition" that

$$|x(t)| \;\le\; c|x^0|e^{-\lambda t} \;+\; c\sup_{s\in[0,t]}|u(s)| \qquad (2)$$

holds for all solutions (assumed defined for all $t>0$), where c and $\lambda>0$ and appropriate constants. What is a reasonable nonlinear version of this?

Two central characteristic of the ISS philosophy are (a) the use of nonlinear gains rather than linear estimates, and (b) the fact that one does not ask what are the exact values of gains, but instead asks *qualitative* questions of *existence*. This represents a "topological" vs. a "metric" point of view (the linear analogy would be to ask only "is the gain $<\infty$?" or "is an operator bounded?").

Our general guiding principle may be formulated thus:

notions of stability should be invariant under (nonlinear) changes of variables.

By a *change of variables* in $\mathbb{R}^\ell$, let us mean here any transformation $z=T(x)$ with $T(0)=0$, where $T:\mathbb{R}^\ell\to\mathbb{R}^\ell$ is a homeomorphism whose restriction $T|_{\mathbb{R}^\ell\setminus\{0\}}$ is a diffeomorphism. (We allow less differentiability at the origin in order to state elegantly a certain converse result later.)

Let us see where this principle leads us, starting from the $L^\infty\to L^\infty$ finite-gain condition (2) and taking both state and input coordinate changes $x=T(z)$, $u=S(v)$. For any input u and initial state x^0, and corresponding trajectory $x(t)=x(t,x^0,u)$, we let $x(t)=T(z(t))$, $u(t)=S(v(t))$, $z^0=z(0)=T^{-1}(x^0)$.

For suitable functions $\underline{\alpha},\overline{\alpha},\overline{\gamma}\in\mathcal{K}_\infty$, we have:

$$\underline{\alpha}(|z|) \;\le\; |T(z)| \;\le\; \overline{\alpha}(|z|)\;\;\forall\, z\in\mathbb{R}^n$$

$$|S(v)| \;\le\; \overline{\gamma}(|v|)\;\;\forall\, v\in\mathbb{R}^m\,.$$

The condition $|x(t)|\le c|x^0|e^{-\lambda t}+c\sup_{s\in[0,t]}|u(s)|$ becomes, in terms of z,v:

$$\underline{\alpha}(|z(t)|) \;\le\; c\,e^{-\lambda t}\overline{\alpha}(|z^0|) \;+\; c\sup_{s\in[0,t]}\overline{\gamma}(|v(s)|)\;\;\forall\, t\ge 0\,.$$

Using again "x" and "u" and letting $\beta(s,t):=ce^{-\lambda t}\overline{\alpha}(s)$ and $\gamma(s):=c\overline{\gamma}(s)$, we arrive to this estimate, with $\beta\in\mathcal{KL}$, $\gamma\in\mathcal{K}_\infty$:

$$\underline{\alpha}\,(|x(t)|) \;\le\; \beta(|x^0|,t)\;+\;\gamma\,(\|u\|_\infty)\;.$$

(It is shown in (Sontag, 1998a) that, for any $\mathcal{KL}$ function β, there exist $\alpha_1,\alpha_2\in\mathcal{K}_\infty$ with

$$\beta(r,t) \;\le\; \alpha_2\left(\alpha_1(r)e^{-t}\right)\;\;\forall\, s,t\,,$$

so the special form of β adds no extra information.) Equivalently, one may write (for different β,γ)

$$|x(t)| \;\le\; \beta(|x^0|,t)\;+\;\gamma\,(\|u\|_\infty)$$

or one may use "max" instead of "+" in the bound.

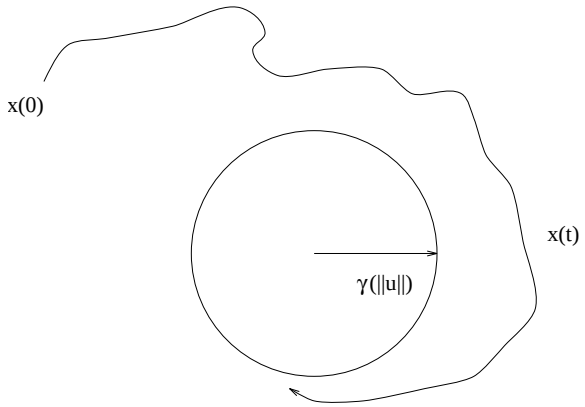
A system is *input to state stable (ISS)* if such an estimate holds, for some $\beta\in\mathcal{KL}$, $\gamma\in\mathcal{K}_\infty$. More precisely, for each x^0, u, the solution $x(t)=x(t,x^0,u)$ is defined for all $t\ge 0$, and the estimate holds.

Asymptotic Gain Characterization

For $u\equiv 0$, the estimate reduces to $|x(t)|\le\beta(|x^0|,t)$, so ISS implies that the unforced system $\dot{x}=f(x,0)$ is globally asymptotically stable (with respect to $x=0$), or as one usually says, "GAS", and in particular stable.

In addition, an ISS system has a well-defined *asymptotic gain*: there is some $\gamma\in\mathcal{K}_\infty$ so that, for all x^0 and u:

$$\limsup_{t\to+\infty}|x(t,x^0,u)| \;\le\; \gamma\,(\|u\|_\infty)\;.$$



A far less obvious converse holds:

Theorem. *("Superposition principle for ISS") A system is ISS if and only if it admits an asymptotic gain and the unforced system is stable.*

This result is nontrivial, and constitutes the main contribution of the paper (Sontag and Wang, 1996), which

establishes as well many other fundamental characterizations of the ISS property. The proof hinges upon a relaxation theorem for differential inclusions, shown in that paper, which relates global asymptotic stability of an inclusion $\dot{x} \in F(x)$ to global asymptotic stability of its convexification.

Dissipation Characterization of ISS

A smooth, proper, and positive definite $V : \mathbb{R}^n \to \mathbb{R}$ is an *ISS-Lyapunov function* for $\dot{x} = f(x,u)$ if, for some $\gamma, \alpha \in \mathcal{K}_\infty$,

$$\dot{V}(x,u) = \nabla V(x)\, f(x,u) \;\leq\; -\alpha(|x|) + \gamma(|u|) \quad \forall\, x, u$$

i.e., one has the dissipation inequality

$$V(x(t_2)) - V(x(t_1)) \;\leq\; \int_{t_1}^{t_2} w(u(s), x(s))\, ds$$

along all trajectories of the system, with "supply" function $w(u,x) = \gamma(|u|) - \alpha(|x|)$.

The following is a fundamental result in ISS theory:

Theorem. (Sontag and Wang, 1995a) *A system is ISS if and only if it admits an ISS-Lyapunov function.*

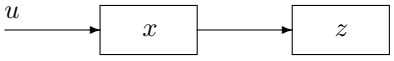
Sufficiency is easy: a differential inequality for V provides an estimate on $V(x(t))$, and hence on $|x(t)|$. Necessity follows by applying the converse Lyapunov theorem from (Lin et al., 1996) for GAS uniform over all $\|d\|_\infty \leq 1$, to a system of the form $\dot{x} = g(x,d) = f(x, d\rho(|x|))$, for an appropriate "robustness margin" $\rho \in \mathcal{K}_\infty$. This is in effect a smooth converse Lyapunov theorem for locally Lipschitz differential inclusions.

ISS is Natural for Series Connections

As a further motivation for the concept of ISS, and as an illustration of the characterizations given, we remark that any cascade (series connection) of ISS systems is again ISS. Consider a cascade connection of ISS systems

$$\begin{aligned} \dot{z} &= f(z,x) \\ \dot{x} &= g(x,u) \end{aligned}$$

where the z-subsystem is ISS with x as input and the x-subsystem is ISS.



The fact that cascades of ISS systems are ISS is one of the reasons that the concept is so useful in recursive design. (In the particular case in which the system $\dot{x} = g(x)$ has no inputs, we conclude that cascading an ISS with a GAS system we obtain a system which is GAS with respect to the state $(x,z) = (0,0)$.)

This fact can be established in several manners, but a particularly illuminating approach is as follows. We start by picking *matching* (cf. (Teel and Sontag, 1995)) ISS-Lyapunov functions for each subsystem:

$$\begin{aligned} \dot{V}_1(z,x) &\leq \theta(|x|) - \alpha(|z|) \\ \dot{V}_2(x,u) &\leq \tilde{\theta}(|u|) - 2\theta(|x|)\,. \end{aligned}$$

Then, $W(x,z) := V_1(z) + V_2(x)$ is an ISS-Lyapunov function:

$$\dot{W}(x,z) \;\leq\; \tilde{\theta}(|u|) - \theta(|x|) - \alpha(|z|)$$

and so a cascade of ISS systems is indeed ISS.

Generalization to Small Gains

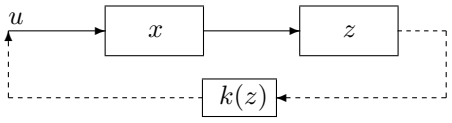
In particular, when $u = 0$, one obtains that a cascade of a GAS and an ISS system is again GAS. More generally, one may allow inputs u fed-back with "small gain": if $u = k(z)$ is so that $|k(z)| \leq \tilde{\theta}^{-1}((1-\varepsilon)\alpha(|z|))$, i.e.

$$\tilde{\theta}(|u|) \leq (1-\varepsilon)\alpha(|z|)$$

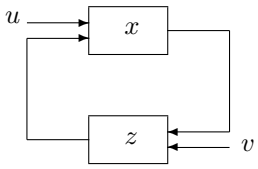
then

$$\dot{W}(x,z) \leq -\theta(|x|) - \varepsilon\alpha(|z|)$$

and the closed-loop system is still GAS.



Even more generally, under suitable conditions on gains (Small-Gain Theorem (Jiang et al., 1994) of Jiang, Praly, and Teel) the closed loop system obtained from an interconnection of two ISS systems $\dot{x} = f(x,z,u)$ and $\dot{z} = g(z,x,v)$, is itself ISS with respect to (u,v).



Series Connections: An Example

As a simple illustration of the cascade technique, consider the angular momentum stabilization of a rigid body controlled by two torques acting along principal axes (for instance, a satellite controlled by two opposing jet pairs). If $\omega = (\omega_1, \omega_2, \omega_3)$ is the angular velocity of a body-attached frame with respect to inertial coordinates, and $I = \mathrm{diag}(I_1, I_2, I_3)$ are the principal moments of inertia, we obtain the equations:

$$I\dot{\omega} \;=\; \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} I\omega + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} v\,.$$

We assume $I_2 \neq I_3$; then, introducing new state and input coordinates via $(I_2 - I_3)x_1 = I_1\omega_1$, $x_2 = \omega_2$, $x_3 =$

ω_3, $I_2u_1 = (I_3-I_1)\omega_1\omega_3+v_1$, and $I_3u_2 = (I_1-I_2)\omega_1\omega_2+v_2$, we obtain a system on $\mathbb{R}^3$, with controls in $\mathbb{R}^2$:

$$\begin{aligned}\dot{x}_1 &= x_2x_3\\ \dot{x}_2 &= u_1\\ \dot{x}_3 &= u_2\,.\end{aligned}$$

Then the following feedback law globally stabilizes the system:

$$\begin{aligned}u_1 &= -x_1 - x_2 - x_2x_3 + v_1\\ u_2 &= -x_3 + x_1^2 + 2x_1x_2x_3 + v_2\end{aligned}$$

when $v_1 = v_2 \equiv 0$. The feedback was obtained arguing in this way: with $z_2 := x_1+x_2$, $z_3 := x_3-x_1^2$, the system becomes:

$$\begin{aligned}\dot{x}_1 &= -x_1^3 + \alpha(x_1, z_2, z_3)\\ \dot{z}_2 &= -z_2 + v_1\\ \dot{z}_3 &= -z_3 + v_2\,.\end{aligned}$$

The x_1-subsystem is easily seen to be ISS, because $\deg_{x_1}\alpha \leq 2$ and hence the cubic term dominates, for large x_1. Thus the cascade is also ISS; in particular, it is GAS if $v_1 = v_2 \equiv 0$. (We also proved a stronger result: ISS implies a global robustness result with respect to actuator noise.)

Generalizations of Other Gains

ISS generalizes finite $L^\infty \to L^\infty$ gains ("L^1 stability") but other classical norms often considered are induced $L^2 \to L^2$ ("H_∞") or $L^2 \to L^\infty$ ("H_2").

Nonlinear transformations starting from "H_∞" estimates:

$$\int_0^t |x(s)|^2\,ds \;\leq\; c|x^0|^2 \,+\, c\int_0^t |u(s)|^2\,ds \quad \forall\, t \geq 0$$

lead to (for appropriate comparison functions):

$$\int_0^t \underline{\alpha}(|x(s)|)\,ds \;\leq\; \kappa(|x^0|) + \int_0^t \gamma(|u(s)|)\,ds \quad \forall\, t \geq 0\,.$$

Theorem. *There is such an "integral to integral" estimate if and only if the system is ISS.*

The proof of this unexpected result is based upon the nontrivial characterizations of the ISS property obtained in (Sontag and Wang, 1996); see (Sontag, 1998a).

On the other hand, "$L^2 \to L^\infty$" stability:

$$|x(t)| \;\leq\; c|x^0|e^{-\lambda t} \,+\, c\int_0^t |u(s)|^2\,ds \quad \text{for all } t \geq 0$$

leads to (for appropriate comparison functions):

$$\underline{\alpha}\,(|x(t)|) \;\leq\; \beta(|x^0|, t) \,+\, \int_0^t \gamma(|u(s)|)\,ds \quad \text{for all } t \geq 0\,.$$

This is the *iISS* (*integral* ISS) property to which we'll return later.

An Application: Observers for Kinetic Networks

We now describe some recent work which we recently carried out with our graduate student Madalena Chaves, see (Chaves and Sontag, Chaves and Sontag), dealing with the design of observers (deterministic Kalman filters) for chemical reaction networks of the "Feinberg-Horn-Jackson" type (cf. (Feinberg, 1987, 1995) as well as an exposition in (Sontag, Sontag)), when seen as systems $\dot{x} = f(x)$ with outputs $y = h(x)$.

For such a system, the dynamics $\dot{x} = f(x)$ are n-dimensional (n is the number of species), and assumed to be given by ideal mass action kinetics, the reaction graph is weakly reversible, and the "deficiency" is zero: $m - \ell - d = 0$, where m is the number of complexes in the network, ℓ is the number linkage classes (connected components in the reaction graph), and d is the dimension of the stoichiometric subspace. We assume also that there are no boundary equilibria on positive classes. It is possible to write such systems as follows (we assume for simplicity here that $\ell = 1$, i.e. the graph is connected):

$$\dot{x} \;=\; f(x) \;=\; \sum_{i=1}^m \sum_{j=1}^m a_{ij}\, x_1^{b_{1j}} x_2^{b_{2j}} \dots x_n^{b_{nj}}\,(b_i - b_j),$$

where the constants a_{ij} are all nonnegative, and the matrix $A = (a_{ij})$ irreducible, each b_j is a column vector in $\mathbb{R}^n$ with entries $b_{1j}, b_{2j}, \dots, b_{nj}$, which are nonnegative integers, and the matrix $B := [b_1, b_2, \dots, b_m]$ has rank $m \leq n$, and no row of B vanishes. We are interested in trajectories which evolve in the positive orthant.

We consider output functions $h : \mathbb{R}^n \to \mathbb{R}^p$ given by vectors of monomials; this includes situations in which concentrations (x_1, x_2, etc.) or reaction rates (proportional to $x_1x_2^3$, etc) are measured. That is, $h : \mathbb{R}^n \to \mathbb{R}^p$ (typically, $p < n$) is of this form:

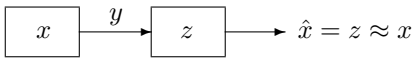
$$h(x) = \begin{pmatrix} x_1^{c_{11}} x_2^{c_{12}} \cdots x_n^{c_{1n}} \\ \vdots \\ x_1^{c_{p1}} x_2^{c_{p2}} \dots x_n^{c_{pn}} \end{pmatrix},$$

where

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \ddots & \\ c_{p1} & c_{p2} & \cdots & c_{pn} \end{bmatrix}$$

has nonnegative integer entries.

A (full-state) *observer* for $\dot{x} = f(x)$, $y = h(x)$ is a system $\dot{z} = g(z, y)$, with state-space $\mathbb{R}^n_+$, such that, for all $x(0)$ and $z(0)$ in $\mathbb{R}^n_+$, the composite system obtained by feeding $y = h(x)$ has solutions defined for $t > 0$, and $|z(t) - x(t)| \to 0$ as $t \to +\infty$.



Generally speaking, the problem of constructing nonlinear observers is extremely difficult. An obvious necessary condition for the existence of observers is detectability: A system $\dot{x} = f(x)$, $y = h(x)$ is *detectable* (on $\mathbb{R}^n_+$) if for all pairs of solutions $x_1(\cdot)$ and $x_2(\cdot)$ in $\mathbb{R}^n_+$:

$$h(x_1(t)) \equiv h(x_2(t)) \;\Rightarrow\; |x_1(t) - x_2(t)| \to 0 \text{ as } t \to \infty\,.$$

Let us introduce the stoichiometric subspace, i.e. the linear span of the "reaction vectors":

$$\mathcal{D} \;:=\; \text{span}\{b_i - b_j \mid i,j = 1,\ldots,m\}\,.$$

Theorem. *The system $\dot{x} = f(x)$, $y = h(x)$ is detectable if and only if*

$$\mathcal{D}^\perp \bigcap \ker C \;=\; \{0\}\,.$$

This condition is simple to check, involving only linear algebra computations. Our main result in (Chaves and Sontag, Chaves and Sontag) shows that this is in fact sufficient.

Theorem. *There exixts an observer if and only if the system is detectable.*

Moreover, in that case, we showed that the following system is an observer:

$$\dot{z} \;=\; f(z) + C'(y - h(z))$$

Note the formal analogy to Luenberger (deterministic Kalman filters) observers for linear systems (in which case the C' matrix is a replaced by a gain L which stabilizes $A + LC$).

In the context of the present paper, the most interesting feature of this observer's construction is the proof that it indeed provides unbiased estimates. The proof is based, roughly, upon the following idea. We first consider the system with inputs

$$\dot{z} \;=\; f(z) + C'(u - h(z))$$

and prove that this system is ISS, relative not to "$z = 0$ and $u = 0$" but each steady-state $z = x_0$ of $\dot{x} = f(x)$ and the associated input $u = h(x_0)$. This is established using an ISS-Lyapunov function (based on relative entropy). Next, we invoke the fact, discussed above, that a cascade of ISS systems is again ISS, plus the fact that $x(t) \to x_0$ for some equilibrium, to conclude the observer property. As a "bonus" from the construction, one gets an automatic property of robustness to small observation noise.

The paper (Chaves and Sontag, Chaves and Sontag) illustrated the observer with the example (studied in (Sontag, Sontag)) of the class of systems corresponding to the kinetic proofreading model for T-cell receptor signal transduction due to McKeithan in (McKeithan, 1995). Let us show here the simplest case of that class of models, with $n = 3$. The kinetics correspond to

$$X_1 + X_2 \underset{k_-}{\overset{k_+}{\rightleftarrows}} X_3\,.$$

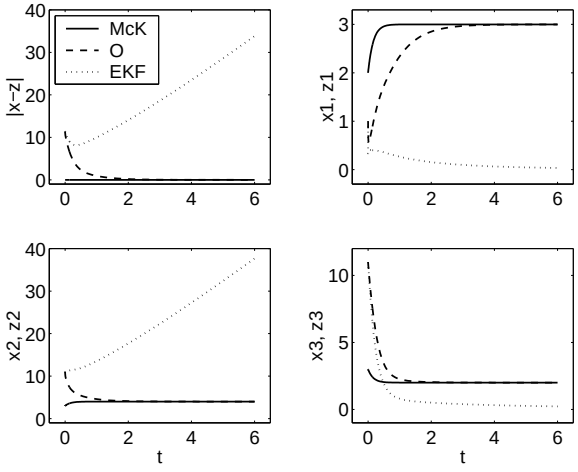
The equations for the system are as follows:

$$\begin{aligned}
\dot{x}_1 &= f_1(x) = -k_+ x_1 x_2 + k_- x_3\\
\dot{x}_2 &= f_2(x) = -k_+ x_1 x_2 + k_- x_3\\
\dot{x}_3 &= f_3(x) = \;\; k_+ x_1 x_2 - k_- x_3
\end{aligned}$$

and we pick for example the following measurement function: $y = h(x) = \begin{pmatrix} x_1 x_2^2 \\ x_1 x_3 \end{pmatrix}$. Then, our theory results in the following observer:

$$\begin{aligned}
\dot{z}_1 &= f_1(z) + (x_1 x_2^2 - z_1 z_2^2) + (x_1 x_3 - z_1 z_3)\\
\dot{z}_2 &= f_2(z) + 2(x_1 x_2^2 - z_1 z_2^2)\\
\dot{z}_3 &= f_3(z) + (x_1 x_3 - z_1 z_3)
\end{aligned}$$

This observer turns out to work surprisingly well (at least in sumulations), even when measurements are very noisy or when there are unobserved step or periodic disturbances on the states of the system. As an example, we show here a simulation which displays the convergence of the observer to the true solution, while an Extended Kalman Filter diverges for the same example. (Here, k_+=0.5, k_-=3, and the initial conditions are $x(0)$=$(2,3,3)'$ and $z(0)$=$(1,11,11)'$.)



Remark: Reversing Coordinate Changes

The "integral to integral" version of ISS arose, in the above discussion, from coordinate changes when starting from L^2-induced operator norms. Interestingly, this result from (Grune et al., 1999) shows that the reasoning can be reversed:

Theorem. *Assume $n \neq 4, 5$. If $\dot{x} = f(x,u)$ is ISS, then, under a coordinate change, for all solutions one has:*

$$\int_0^t |x(s)|^2\, ds \;\leq\; |x^0|^2 \;+\; \int_0^t |u(s)|^2\, ds\,.$$

The cases $n = 4, 5$ are still open. The proof is based on tools from "h-cobordism theory," developed by Smale and Milnor in the 1960s in order to prove the validity of the generalized Poincaré conjecture.

Integral-Input to State Stability

Recall that the "$L^2 \to L^\infty$" operator gain property led us, under coordinate changes, to the iISS property expressed by the estimate:

$$\underline{\alpha}\,(|x(t)|) \;\leq\; \beta(|x^0|,t) \;+\; \int_0^t \gamma(|u(s)|)\,ds\,.$$

There is a dissipation characterization here as well.

A smooth, proper, and positive definite $V : \mathbb{R}^n \to \mathbb{R}$ is an *iISS-Lyapunov function* for $\dot{x} = f(x,u)$ if for some *positive definite* continuous α and $\gamma \in \mathcal{K}_\infty$

$$\nabla V(x)\, f(x,u) \;\leq\; -\alpha(|x|) + \gamma(|u|) \quad \forall\, x \in \mathbb{R}^n,\, u \in \mathbb{R}^m$$

—observe that we are not requiring now $\alpha \in \mathcal{K}_\infty$. (Intuitively: even for constant u one may have $\dot{V} > 0$, but $\gamma(|u|) \in \mathcal{L}^1$ means that $\dot{V}$ is "often" negative.)

A recent result from (Angeli et al., 2000a) is this:

Theorem. *A system is iISS if and only if it admits an iISS-Lyapunov function.*

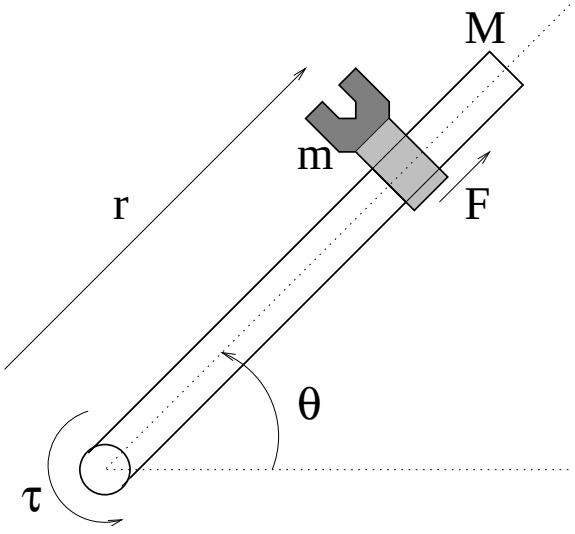
Since any $\mathcal{K}_\infty$ function is positive definite, every ISS system is also iISS, but the converse is false. For example, a bilinear system

$$\dot{x} = (A + \sum_{i=1}^m u_i A_i)x + Bu$$

is iISS if and only if A is a Hurwitz matrix, but in general it is not ISS—e.g., if $B = 0$ and $A + \sum_{i=1}^m u_i^0 A_i$ is not Hurwitz for some u^0. As another example, take $\dot{x} = -\tan^{-1} x + u$. This is not ISS, since bounded inputs may produce unbounded trajectories; but it is iISS, since $V(x) = x\,\tan^{-1} x$ is an iISS-Lyapunov function.

An Application of iISS Theory

Let us illustrate the iISS results through an application which, as a matter of fact, was the one that originally motivated much of the work in (Angeli et al., 2000a). Consider a rigid manipulator with two controls:



The arm is modeled as a segment with mass M and length L, and the hand as a point with mass m. Denoting by r the position and by θ the angle of the arm, the resulting equations are:

$$(mr^2 + ML^2/3)\;\ddot{\theta} + 2mr\dot{r}\dot{\theta} = \tau\,, \quad m\ddot{r} - mr\dot{\theta}^2 = F$$

where F and τ are the external force and torque. In a typical passivity-based tracking design, one takes

$$\tau := -k_{d_1}\dot{\theta} - k_{p_1}(\theta - \theta_d)$$
$$F := -k_{d_2}\dot{r} - k_{p_2}(r - r_d)$$

where r_d and θ_d are the desired signals and the gains $(k_{d_1}, \ldots)$ are > 0. For constant reference θ_d, r_d, there is tracking: $\theta \to \theta_d$, $\dot{\theta} \to 0$, and analogously for r.

But, what about time-varying θ_d, r_d? Can these destabilize the system? The answer is yes: there are bounded inputs which produce "nonlinear resonance," so the system cannot be ISS (not even bounded-input bounded-state). Such examples are presented in (Angeli et al., 2000a). On the other hand, one reason that standard tracking design is useful is that many inputs are *not* destabilizing, and one would like to find a way to formulate qualitatively that aspect. One way is by showing that the system is iISS.

The closed-loop system is 4-dimensional, with states $(q, \dot{q})$, $q = (\theta, r)$ and $u = (k_{p_1}\theta_d, k_{p_2}r_d)$:

$$(mr^2 + ML^2/3)\ddot{\theta} + 2mr\dot{r}\dot{\theta} = u_1 - k_{d_1}\dot{\theta} - k_{p_1}\theta$$
$$m\ddot{r} - mr\dot{\theta}^2 = u_- k_{d_2}\dot{r} - k_{p_2}r\,.$$

To prove that it is iISS, we consider the mechanical energy V, and note the following passivity-type estimate:

$$\frac{d}{dt}V(q(t),\dot{q}(t)) \;\leq\; -c_1|\dot{q}(t)|^2 + c_2|u(t)|^2$$

for sufficiently small $c_1 > 0$ and large $c_2 > 0$.

In general, we say that a system is *h-dissipative* with respect to an output function $y = h(x)$ (continuous and with $h(0) = 0$) if, for some $\mathcal{C}^\infty$ positive definite, proper $V : \mathbb{R}^n \to \mathbb{R}$, and for some γ, α as above,

$$\nabla V(x)\, f(x,u) \;\leq\; -\alpha(h(x)) + \gamma(|u|) \quad \forall\, x \in \mathbb{R}^n,\, u \in \mathbb{R}^m$$

and that it is *weakly h-detectable* if, for all trajectories, $y(t) = h(x(t)) \equiv 0$ implies that $x(t) \to 0$ as $t \to \infty$.

This is proved in (Angeli et al., 2000a):

Theorem. *A system is iISS if and only if it is weakly h-detectable and h-dissipative for some output h.*

With output $\dot{q}$, our example is weakly zero-detectable and dissipative, since $u \equiv 0$ and $\dot{q} \equiv 0$ imply $q \equiv 0$. Thus the system is indeed iISS, as claimed.

Mixed Notions

Changes of variables transformed "finite L^2 gain" to an "integral to integral" property, which turns out to be

equivalent to ISS. Finite gain as operators between L^p and L^q spaces, with $p \neq q$ both finite, lead instead to this type of "weak integral to integral" estimate:

$$\int_0^t \underline{\alpha}(|x(s)|)\,ds \;\leq\; \kappa(|x^0|) \;+\; \alpha\left(\int_0^t \gamma(|u(s)|)\,ds\right)$$

for appropriate $\mathcal{K}_\infty$ functions (note the additional "α"). See (Angeli et al., 2000b) for more discussion on how this estimate is reached, as well as this result:

Theorem. *A system satisfies a weak integral to integral estimate if and only if it is iISS.*

Another interesting variant results by studying *mixed* integral/supremum estimates:

$$\underline{\alpha}(|x(t)|) \;\leq\; \beta(|x^0|,t) \;+\; \int_0^t \gamma_1(|u(s)|)\,ds \;+\; \gamma_2(\|u\|_\infty)$$

for suitable $\beta \in \mathcal{KL}$ and $\underline{\alpha}, \gamma_i \in \mathcal{K}_\infty$. This result is also from (Angeli et al., 2000b):

Theorem. *The system $\dot{x} = f(x,u)$ satisfies a mixed estimate if and only if it is iISS.*

We also remark a "separation principle" recently obtained for iISS. In (Angeli, Ingalls, Sontag, and Wang, Angeli et al.), a system is said to be *bounded energy converging state* (BECS) if it is forward complete and, for all trajectories,

$$\int_0^\infty \sigma(|u(s)|)\,ds < \infty \;\Rightarrow\; \liminf_{t\to\infty} |x(t)| = 0$$

(for suitable $\sigma \in \mathcal{KL}$). The authors then prove that a system is iISS if and only if it is BECS and the 0-system $\dot{x} = f(x,0)$ has the origin as an asymptotically stable equilibrium.

Input/Output Stability

The discussion so far (except for the application to chemical observers) has been exclusively for notions involving stability from inputs to internal states. We now turn to external stability.

For linear systems (1), *external stability* means that the transfer function is stable, or, in terms of a state-space realization, that an estimate as follows holds:

$$|y(t)| \;\leq\; \beta(t)|x^0| + \gamma\,\|u\|_\infty\;,$$

where γ is a constant and β converges to zero (β may be obtained from the restriction of A to a minimal subsystem). Observe that, even though we only require that y, not x, be "small" (relative to $\|u\|_\infty$), the initial internal states still affect the estimate in a "fading memory" manner, via the β term. (For example, in PID control, when considering the combination of plant, exosystem and controller, the overshoot of the regulated variable will be determined by the magnitude of the constant disturbance, and the initial state of the integrator.)

Under coordinate changes, and arguing just as earlier, external stability leads us to *input to output* stability (IOS) for systems with outputs $\dot{x} = f(x,u)$, $y = h(x)$. This is the property that, for some $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}_\infty$, the following estimate must hold along all solutions:

$$|y(t)| \;\leq\; \beta(|x^0|,t) \;+\; \sup_{s\in[0,t]} \gamma(|u(s)|)\,.$$

A dissipation (Lyapunov-) type characterization of this property is as follows. An *IOS-Lyapunov function* is a smooth $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ so that, for some $\alpha_i \in \mathcal{K}_\infty$, for all $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$:

$$\alpha_1(|h(x)|) \;\leq\; V(x) \;\leq\; \alpha_2(|x|)$$

and

$$V(x) > \alpha_3(|u|) \;\Rightarrow\; \nabla V(x)\, f(x,u) < 0\,.$$

For systems that are bounded-input bounded-state stable, we have (see (Sontag and Wang, Sontag and Wang)):

Theorem. *A system $\dot{x} = f(x,u)$, $y = h(x)$ is IOS if and only if it admits an IOS-Lyapunov function.*

One may re-interpret this result as the existence of a new output map $\widetilde{y} = \alpha_1^{-1}(V(x))$ which dominates the original output ($y \leq \widetilde{y}$) and which is monotonically decreasing (no overshoot) as long as inputs are small. This is, in fact, one generalization of a central argument used in regulator theory (Francis equations).

A "separation theorem" providing an asymptotic gain characterization of IOS, similar to that given earlier for ISS, can be found in (Ingalls et al., 2001). A version of the above theorem for systems with are not BIBS is given in (Ingalls and Wang, 2001).

Zero-Detectability: IOSS

Detectability is yet another property which is central to systems analysis. For linear systems (1), (zero-) detectability means that the unobservable part of the system is stable, i.e.,

$$y(t) = Cx(t) \equiv 0 \;\&\; u(t) \equiv 0 \;\Rightarrow\; x(t) \to 0 \text{ as } t \to \infty$$

or equivalently:

$$u(t) \to 0 \;\&\; y(t) \to 0 \;\Rightarrow\; x(t) \to 0$$

$$u \to 0 \;\longrightarrow\; \boxed{\Rightarrow\;\; x \to 0} \;\longrightarrow\; y \to 0$$

(see, for instance, the textbook (Sontag, 1998b)) and can be also expressed by means of an estimate of the following form:

$$|x(t)| \;\leq\; \beta(t)|x^0| + \gamma_1\,\|u\|_\infty + \gamma_2\,\|y\|_\infty$$

where the γ_i's are constants and β converges to zero (now β is obtained from a suitable matrix $A-LC$, where L is an observer gain) and the sup norms are interpreted as appying to restrictions to $[0,t]$.

Under cordinate changes, one is led to *input/output to state* stability (IOSS). This is the property defined by the requirement that an estimate of the following type hold along all trajectories:

$$|x(t)| \;\leq\; \beta(|x^0|,t) + \sup_{s\in[0,t]} \gamma(|u(s)|) + \sup_{s\in[0,t]} \gamma(|y(s)|)$$

(for some $\beta \in \mathcal{KL}$, $\gamma \in \mathcal{K}_\infty$). The terminology IOSS is self-explanatory: formally, there is "stability from the i/o data to the state".

Dissipation Characterization of IOSS

A smooth, proper, and positive definite $V : \mathbb{R}^n \to \mathbb{R}$ is an *IOSS-Lyapunov function* if, for some $\alpha_i \in \mathcal{K}_\infty$,

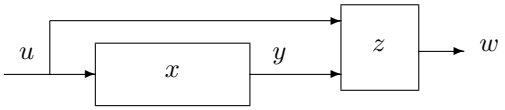
$$\nabla V(x)\, f(x,u) \;\leq\; -\alpha_1(|x|) + \alpha_2(|u|) + \alpha_3(|y|)$$

for all $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$.

This is from (Krichman, 2000) and (Krichman, Sontag, and Wang, Krichman et al.):

Theorem. *A system $\dot{x} = f(x,u)$, $y = h(x)$ is IOSS if and only if it admits an IOSS-Lyapunov function.*

As a corollary, IOSS is equivalent to the existence of a *norm-estimator*: driven by the i/o data generated by the original system, it estimates an upper bound on the internal state.



This is defined as a system $\dot{z} = g(z,u,y)$, $w = \ell(z)$, whose inputs are the i/o pairs of the original system, which is ISS with respect to u, y as inputs (so that there is robustness to signal errors), and, for some $\rho \in \mathcal{K}$ and $\beta \in \mathcal{KL}$,

$$|x(t)| \;\leq\; \beta(|x^0|+|z^0|,t) \;+\; \rho(|w(t)|) \quad \forall\, t \geq 0$$

for all initial states x^0 and z^0. (See the paper (Krichman, Sontag, and Wang, Krichman et al.) for the precise definition.)

An asymptotic gain characterization of IOSS also exists, see (Angeli, Ingalls, Sontag, and Wang, Angeli et al.).

Output to Input Stability and Minimum-Phase Systems

Recall that a linear system (1), let us say for simplicity single-input and single-output, is said to be *minimum-phase* if the inverse of its transfer function is stable, i.e. if all the zeroes of its transfer function have negative real part. The minimum-phase property is ubiquitous in control design, for instance because it allows to solve control problems by simple inversion; it is also needed for convergence of several adaptive control algorithms, and it allows stabilization by output feedback.

In the late 1980s, a notion of minimum-phase (and associated "zero dynamics") was introduced by Byrnes and Isidori (see (Byrnes and Isidori, 1988) as well as (Isidori, 1995)). This concept has proved very succesful in allowing the extension of many linear systems results to nonlinear systems. Basically, a minimum-phase system is one for which the zero-dynamics subsystem (which is obtained by clamping the output at zero) is GAS.

Often, however, an enhancement of this GAS property is needed, in effect impossing on the zero-dynamics an ISS property with respect to the output and its derivatives, see for instance (Praly and Jiang, 1993). The paper (Liberzon et al., 2000) showed that it is possible to define this enhancement directly, and with no recourse to "normal forms" or even zero-dynamics, by requiring an "output to input stability" property. A system is said to be *output-input stable* (OIS), or more precisely "derivatives of output to state and input stable" if, for some positive integer k, an estimate as follows:

$$|u(t)| + |x(t)| \leq \beta(|x^0|\,,t) + \gamma(\|y\|_\infty + \ldots \|y^{(k)}\|_\infty)$$

holds along all trajectories corresponding to smooth controls, for some $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}_\infty$ (the norms of y and its derivatives are understood as those of restrictions to $[0,t]$).

We refer the reader to (Liberzon et al., 2000) for details, as well as an application in adaptive control and the proof that the OIS property is equivalent (for scalar input and output real-analytic systems) to the existence of a relative degree plus an OSS property with respect to output derivatives.

Other Cascading Results

We remarked several times on the fact that cascades of ISS systems are ISS, and the role that such a property plays. It is easy to provide examples of the fact that the cascade of an iISS system $\dot{x} = f(x,z)$ with an ISS system $\dot{z} = f(z,u)$ need not be iISS, and this motivated work reported in (Arcak, Angeli, and Sontag, Arcak et al.) dealing with "matching" conditions which insure such cascade well-behavior, as well as the work reported in (Angeli et al., 2001) which deals with a notion of ISS with respect to input *derivatives* (there is a certain formal duality to the output to input stability property).

Closing Comments

The developments in ISS theory during the last decade have allowed a complete characterization of most of the

important properties identified so far (ISS itself, IOS, IOSS, iISS, etc). Nonetheless, the area remains very open, as major concepts still need clarification. Among the main questions are the need to further study and characterize *incremental* IOSS (not merely distinguishing from $x = 0$, but distinguishing every pair of states, as needed for observers), and the need to deal with feedback designs which provide an ISS property with respect to observation errors.

This brief survey has focused on basic theoretical constructs, instead of on applications. Let us turn now to some further references related to ISS-related theory as well as pointers to applications.

Textbooks and research monographs which make use of ISS and related concepts include (Freeman and Kokotovi'c, 1996; Isidori, 1999; Krstić and Deng, 1998; Krstić et al., 1995; Khalil, 1996; Sepulchre et al., 1997).

After the definition in (Sontag, 1989a) and the basic characterizations in (Sontag and Wang, 1995a), the main results on ISS are given in (Sontag and Wang, 1996). See also (Coron et al., 1995; Sussmann et al., 1994) for early uses of asymptotic gain notions. "Practical" ISS is equivalent to ISS with respect to compact attractors, see (Sontag and Wang, 1995b).

Several authors have pointed out that *time-varying* system versions of ISS are central to the analysis of asymptotic tracking problems, see e.g. (Tsinias and Karafyllis, 1999). In (Edwards et al., 2000), one can find further results on Lyapunov characterizations of the ISS property for time-varying (and in particular periodic) systems, as well as a small-gain theorem based on these ideas.

Perhaps the most interesting set of open problems concerns the construction of feedback laws that provide ISS stability with respect to observation errors. Actuator errors are far better understood (cf. (Sontag, 1989a)), but save for the case of special structures studied in (Freeman and Kokotovi'c, 1996), the one-dimensional case (see e.g. (Fah, 1999)) and the counterexample (Freeman, 1996), little is known of this fundamental question. Recent work analyzing the effect of small observation errors (see (Sontag, 1999)) might provide good pointers to useful directions of research (indeed, see (Liberzon, 2000) for some preliminary remarks in that direction). For special classes of systems, even output feedback ISS with respect to observation errors is possible, cf. (Nešić and Sontag, 1998).

Both ISS and iISS properties have been featured in the analysis of the performance of switching controllers, cf. (Hespanha and Morse, 1999a) and (Hespanha and Morse, 1999b).

Coprime factorizations are the basis of the parameterization of controllers in the Youla approach. As a matter of fact, as the paper's title indicates, their study was the original motivation for the introduction of the notion of ISS in (Sontag, 1989a). Some further work can be found in (Sontag, 1989b), see also (Fujimoto and Sugie, 1998), but much remains to be done.

There are now results on averaging for ISS systems, see (Nešić and Teel, Nešić and Teel), as well as on singular perturbations, see (Christofides and Teel, 1996).

Discrete-time ISS systems are studied in (Kazakos and Tsinias, 1994) and in (Jiang et al., 1999); the latter paper provides Lyapunov-like sufficient conditions and an ISS small-gain theorem, and more complete characterizations and extensions of many standard ISS results for continuous time systems are given in (Jiang and Wang, Jiang and Wang).

Discrete-time iISS systems are the subject of (Angeli, 1999b), which proves the very surprising result that, in the discrete-time case, iISS is actually no different than global asymptotic stability of the unforced system (this is very far from true in the continuous-time case, of course). In this context, of interest are also the relationships between the ISS property for a continuous-time system and its sampled versions. The result in (Teel et al., 1998) shows that ISS is recovered under sufficiently fast sampling; see also the technical estimates in (Nešić et al., 1999).

Stochastic ISS properties are treated in (Tsinias, 1998).

A very interesting area regards the combination of clf and ISS like-ideas, namely providing necessary and sufficient conditions, in terms of appropriate clf-like properties, for the existence of feedback laws (or more generally, dynamic feedback) such that the system $\dot{x} = f(x, d, u)$ becomes ISS (or iISS, etc) with respect to d, once that $u = k(x)$ is substituted. Notice that for systems with disturbances typically $f(0, d, 0)$ need not vanish (example: additive disturbances for linear systems), so this problem is qualitatively different from the robust-clf problem since uniform stabilization is not possible. There has been substantial work by many authors in this area; let us single out among them the work (Teel and Praly, 2000), which deals primarily with systems of the form $\dot{x} = f(x, d) + g(x)u$ (affine in control, and control vector fields are independent of disturbances) and with assigning precise upper bounds to the "nonlinear gain" obtained in terms of d, and (Deng and Krstić, 2000), which, for the class of systems that can be put in output-feedback form (controller canonical form with an added stochastic output injection term), produces, via appropriate clf's, stochastic ISS behavior ("NSS" = noise to state stability, meaning that solutions converge in probability to a residual set whose radius is proportional to bounds on covariances).

In connection with our example from tracking design for a robot, we mention here that the paper (Marino and Tomei, 1999) proposed the reformulation of tracking problems by means of the notion of input to state stability. The goal was to strengthen the robustness properties of tracking designs, and the notion of ISS

was instrumental in the precise characterization of performance. Incidentally, the same example was used, for a different purpose—namely, to illustrate a different nonlinear tracking design which produces ISS, as opposed to merely iISS, behavior—in the paper (Angeli, 1999a).

Neural-net control techniques using ISS are mentioned in (Sanchez and Perez, 1999).

A problem of decentralized robust output-feedback control with disturbance attenuation for a class of large-scale dynamic systems, achieving ISS and iISS properties, is studied in (Jiang et al., 1999).

Incremental ISS is the notion that estimates differences $|x_1(t) - x_2(t)|$ in terms of $\mathcal{KL}$ decay of differences of initial states, and differences of norms of inputs. It provides a way to formulate notions of sensitivity to initial conditions and controls (not local like Lyapunov exponents or as in (Lohmiller and Slotine, 1998), but of a more global character, see (Angeli, Angeli)); in particular when there are no inputs one obtains "incremental GAS", which can be completely characterized in Lyapunov terms using the result in (Lin et al., 1996), since it coincides with stability with respect to the diagonal of the system consisting of two parallel copies of the same system. This area is of interest, among other reasons, because of the possibility of its use in information transmission by synchronization of diffusively coupled dynamical systems ((Pogromsky et al., 1999)) in which the stability of the diagonal is indeed the behavior of interest.

Small-gain theorems for ISS and IOS notions originated with (Jiang et al., 1994); a purely operator version (cf. (Ingalls et al., 1999)) of the IOS small-gain theorem holds as well. There are ISS-small gain theorems for certain infinite dimensional classes of systems such as delay systems, see (Teel, 1998).

The notion of IOSS is called "detectability" in (Sontag, 1989b) (where it is phrased in input/output, as opposed to state space, terms, and applied to questions of parameterization of controllers) and was called "strong unboundedness observability" in (Jiang et al., 1994). IOSS and its incremental variant are very closely related to the OSS-type detectability notions pursued in (Krener, 1999); see also the emphasis on ISS guarantees for observers in (Marino et al., 1999). The use of ISS-like formalism for studying observers, and hence implicitly the IOSS property, has also appeared several times in other authors' work, such as the papers (Hu, 1991; Lu, 1995a; Pan et al., 1993).

It is worth pointing out that several authors had independently suggested that one should *define* "detectability" in dissipation terms. For example, in (Lu, 1995b), Equation 15, one finds detectability defined by the requirement that there should exist a differentiable storage function V satisfying our dissipation inequality but with the special choice $\alpha_3(r) := r^2$ (there were no inputs in the class of systems considered there). A variation of this is to weaken the dissipation inequality, to require merely

$$x \neq 0 \Rightarrow \nabla V(x)\, f(x,u) < \alpha_3(|y|)$$

(again, with no inputs), as done for instance in the definition of detectability given in (Morse, 1995). Observe that this represents a slight weakening of the ISS property, in so far as there is no "margin" of stability $-\alpha_1(|x|)$.

Norm-estimators are motivated by developments appeared in (Jiang and Praly, 1992) and (Praly and Wang, 1996).

The notion studied in (Shiriaev, 1998) is very close to the combination of IOSS and IOS being pursued in (Ingalls, Sontag, and Wang, Ingalls et al.).

Partial asymptotic stability for differential equations is a particular case of output stability (IOS when there are no inputs) in our sense; see (Vorotnikov, 1993) for a survey of the area, as well as the book (Rumyantsev and Oziraner, 1987), which contains a converse theorem for a restricted type of output stability. (We thank Anton Shiriaev for bringing this latter reference to our attention.)

References

Angeli, D., A Lyapunov approach to incremental stability properties. to appear.

Angeli, D., B. Ingalls, E. D. Sontag, and Y. Wang, Asymptotic Characterizations of IOSS. submitted.

Angeli, D., E. D. Sontag, and Y. Wang, "A characterization of integral input to state stability," *IEEE Trans. Auto. Cont.*, **45**, 1082–1097 (2000a).

Angeli, D., E. D. Sontag, and Y. Wang, "Further equivalences and semiglobal versions of integral input to state stability," *Dynamics and Control*, **10**, 127–149 (2000b).

Angeli, D., E. D. Sontag, and Y. Wang, A Note on input-to-state stability with input derivatives, In *Proc. Nonlinear Control System Design Symposium* (2001). to appear.

Angeli, D., "Input-to-state stability of PD-controlled robotic systems," *Automatica*, **35**, 1285–1290 (1999a).

Angeli, D., "Intrinsic robustness of global asymptotic stability," *Sys. Cont. Let.*, **38**, 297–307 (1999b).

Arcak, M., D. Angeli, and E. D. Sontag, Stabilization of cascades using integral input-to-state stability. submitted.

Byrnes, C. I. and A. Isidori, "Local stabilization of minimum phase nonlinear systems," *Sys. Cont. Let.*, **11**, 9–17 (1988).

Chaves, M. and E. D. Sontag, State-Estimators for chemical reaction networks of Feinberg-Horn-Jackson zero deficiency type. submitted.

Christofides, P. D. and A. R. Teel, "Singular perturbations and input-to-state stability," *IEEE Trans. Auto. Cont.*, **41**, 1645–1650 (1996).

Coron, J. M., L. Praly, and A. Teel, Feedback stabilization of nonlinear systems: sufficient conditions and Lyapunov and input-output techniques, In Isidori, A., editor, *Trends in Control*. Springer-Verlag, London (1995).

Deng, H. and M. Krstić, "Output-feedback stablization of stochastic nonlinear systems driven by noise of unknown covariance," *Sys. Cont. Let.*, **39**, 173–182 (2000).

Edwards, H., Y. Lin, and Y. Wang, On input-to-state stability for time varying nonlinear systems, In *Proc. 39th IEEE Conf. Decision and Control, Sydney, Dec. 2000*, pages 3501–3506. IEEE Publications (2000).

Fah, N. C. S., "Input-to-state stability with respect to measurement disturbances for one-dimensional systems," *Control, Optimisation and Calculus of Variations*, **4**, 99–121 (1999).

Feinberg, M., "Chemical reaction network structure and the stabiliy of complex isothermal reactors—I. The deficiency zero and deficiency one theorems," *Chem. Eng. Sci.*, **42**, 2229–2268 (1987). Review Article 25.

Feinberg, M., "The existence and uniqueness of steady states for a class of chemical reaction networks," *Archive for Rational Mechanics and Analysis*, **132**, 311–370 (1995).

Freeman, R. A. and P. V. Kokotovi'c, *Robust Nonlinear Control Design, State-Space and Lyapunov Techniques*. Birkhäuser, Boston (1996).

Freeman, R. A., "Global internal stabilizability does not imply global external stabilizability for small sensor disturbances," *IEEE Trans. Auto. Cont.*, **40**, 2119–2122 (1996).

Fujimoto, K. and T. Sugie, State-space characterization of Youla parametrization for nonlinear systems based on input-to-state stability, In *Proc. 37th IEEE Conf. Decision and Control*, pages 2479–2484, Tampa (1998).

Grune, L., F. R. Wirth, and E. D. Sontag, "Asymptotic stability equals exponential stability and ISS equals finite energy gain—if you twist your eyes," *Sys. Cont. Let.*, **38**, 127–134 (1999).

Hespanha, J. P. and A. S. Morse, "Certainty equivalence implies detectability," *Sys. Cont. Let.*, **36**, 1–13 (1999a).

Hespanha, J. P. and A. S. Morse, Supervisory control of integral-input-to-state stabilizing controllers, In *Proc. of the* 5^{th} *European Control Conference*, Karlsruhe (1999b).

Hu, X. M., "On state observers for nonlinear systems," *Sys. Cont. Let.*, **17**, 645–473 (1991).

Ingalls, B. and Y. Wang, On input-to-output stability for systems not uniformly bounded, In *Proc. IFAC Symposium on Nonlinear Control Systems (NOLCOS)*, St. Petersburg (2001). to appear.

Ingalls, B., E. D. Sontag, and Y. Wang. In preparation.

Ingalls, B., E. D. Sontag, and Y. Wang, Remarks on input to output stability, In *Proc. IEEE Conf. Decision and Control*, pages 1226–1231, Phoenix. IEEE Publications (1999).

Ingalls, B., E. D. Sontag, and Y. Wang, Generalizations of asymptotic gain characterizations of ISS to input-to-output stability, In *Proc. American Control Conf.* (2001). to appear.

Isidori, A., *Nonlinear Control Systems*. Springer-Verlag, London, third edition (1995).

Isidori, A., *Nonlinear Control Systems II*. Springer-Verlag, London (1999).

Jiang, Z.-P. and L. Praly, "Preliminary results about robust Lagrange stability in adaptive nonlinear regulation," *Int. J. of Adaptive Control and Signal Processing*, **6**, 285–307 (1992).

Jiang, Z.-P. and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*. to appear.

Jiang, Z.-P., A. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Mathematics of Control, Signals and Systems*, **7**, 95–120 (1994).

Jiang, Z.-P., F. Khorrami, and D. J. Hill, Decentralized output-feedback control with disturbance attenuation for large-scale nonlinear systems, In *Proc. 38th IEEE Conf. Decision and Control*, pages 3271–3276, Phoenix (1999).

Jiang, Z.-P., E. D. Sontag, and Y. Wang, Input-to-state stability for discrete-time nonlinear systems, In *Proc. 14th IFAC World Congress*, volume E, pages 277–282, Beijing (1999).

Kazakos, D. and J. Tsinias, "The input-to-state stability condition and global stabilization of discrete-time systems," *IEEE Trans. Auto. Cont.*, **39**, 2111–13 (1994).

Khalil, H. K., *Nonlinear Systems, Second Edition*. Prentice-Hall, Upper Saddle River, NJ (1996).

Krener, A. J., A Lyapunov theory of nonlinear observers, In *Stochastic analysis, control, optimization and applications*, pages 409–420. Birkhäuser, Boston (1999).

Krichman, M., E. D. Sontag, and Y. Wang, Input-output-to-state stability. submitted.

Krichman, M., *A Lyapunov approach to detectability of nonlinear systems*, PhD thesis, Rutgers University (2000). E. Sontag, advisor.

Krstić, M. and H. Deng, *Stabilization of Uncertain Nonlinear Systems*. Springer-Verlag, London (1998).

Krstić, M., I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear and Adaptive Control Design*. John Wiley & Sons, New York (1995).

Liberzon, D., A. S. Morse, and E. D. Sontag, "Output-input stability and minimum-phase nonlinear systems," *IEEE Trans. Auto. Cont.*, pages 2106–2111 (2000). Preliminary version appeared as "A new definition of the minimum-phase property for nonlinear systems, with an application to adaptive control", in *Proc. IEEE Conf. Decision and Control, Sydney, Dec. 2000*.

Liberzon, D., Nonlinear stabilization by hybrid quantized feedback, In Lynch, N. and B. H. Krogh, editors, *Proc. Third International Workshop on Hybrid Systems: Computation and Control*, volume 1790 of *Lecture Notes in Computer Science*, pages 243–257 (2000).

Lin, Y., E. D. Sontag, and Y. Wang, "A smooth converse Lyapunov theorem for robust stability," *SIAM J. Cont. Opt.*, **34**, 124–160 (1996).

Lohmiller, W. and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, **34**, 683–696 (1998).

Lu, W. M., "A class of globally stabilizing controllers for nonlinear systems," *Sys. Cont. Let.*, **25**, 13–19 (1995a).

Lu, W. M., "A state-space approach to parameterization of stabilizing controllers for nonlinear systems," *IEEE Trans. Auto. Cont.*, **40**, 1576–1588 (1995b).

Marino, R. and P. Tomei, "Nonlinear output feedback tracking with almost disturbance decoupling," *IEEE Trans. Auto. Cont.*, **44**, 18–28 (1999).

Marino, R., G. Santosuosso, and P. Tomei, Robust adaptive observers for nonlinear systems with bounded disturbances, In *Proc. 38th IEEE Conf. Decision and Control*, pages 5200–5205, Phoenix (1999).

McKeithan, T. W., "Kinetic proofreading in T-cell receptor signal transduction," *Proc. Natl. Acad. Sci. USA*, **92**, 5042–5046 (1995).

Morse, A. S., Control using logic-based switching, In Isidori, A., editor, *Trends in Control: A European Perspective*, pages 69–114. Springer, London (1995).

Nešić, D. and E. D. Sontag, "Input-to-state stabilization of linear systems with positive outputs," *Sys. Cont. Let.*, **35**, 245–255 (1998).

Nešić, D. and A. R. Teel, "Input-to-state stability for nonlinear time-varying systems via averaging," *Mathematics of Control, Signals and Systems*. to appear.

Nešić, D., A. R. Teel, and E. D. Sontag, "Formulas relating $\mathcal{KL}$ stability estimates of discrete-time and sampled-data nonlinear systems," *Sys. Cont. Let.*, **38**, 49–60 (1999).

Pan, D. J., Z. Z. Han, and Z. J. Zhang, "Bounded-input-bounded-output stabilization of nonlinear systems using state detectors," *Sys. Cont. Let.*, **21**, 189–198 (1993).

Pogromsky, A. Y., T. Glad, and H. Nijmeije, "On diffusion driven oscillations in coupled dynamical systems," *Int. J. Bifurcation Chaos*, **9**, 629–644 (1999).

Praly, P. and Z. P. Jiang, "Stabilization by output feedback for systems with ISS inverse dynamics," *Sys. Cont. Let.*, **21**, 19–33 (1993).

Praly, L. and Y. Wang, "Stabilization in spite of matched unmodelled dynamics and an equivalent definition of input-to-state

stability," *Mathematics of Control, Signals and Systems*, **9**, 1–33 (1996).

Rumyantsev, V. V. and A. S. Oziraner, *Stability and Stabilization of Motion with Respect to Part of the Variables (in Russian)*. Nauka, Moscow (1987).

Sanchez, E. N. and J. P. Perez, "Input-to-state stability (ISS) analysis for dynamic neural networks," *IEEE Trans. Circ. and Sys. I: Fund. Theory and Appl.*, **46**, 1395–1398 (1999).

Sepulchre, R., M. Jankovic, and P. V. Kokotović, *Constructive Nonlinear Control*. Springer (1997).

Shiriaev, A. S., The notion of V-detectability and stabilization of invariant sets of nonlinear systems, In *Proc. 37th IEEE Conf. Decision and Control*, pages 2509–2514, Tampa (1998).

Sontag, E. D., "Structure and stability of certain chemical networks and applications to the kinetic proofreading model of T-cell receptor signal transduction," *IEEE Trans. Auto. Cont.* Preprints available as math.DS/9912237 (1999) revised as math.DS/0002113 (2000) "in Los Alamos Archive, `http://arXiv.org`".

Sontag, E. D. and Y. Wang, "Lyapunov characterizations of input to output stability," *SIAM J. Cont. Opt.* to appear.

Sontag, E. D. and Y. Wang, "On characterizations of the input-to-state stability property," *Sys. Cont. Let.*, **24**, 351–359 (1995a).

Sontag, E. D. and Y. Wang, Various results concerning set input-to-state stability, In *Proc. IEEE Conf. Decision and Control, New Orleans, Dec. 1995*, pages 1330–1335. IEEE Publications (1995b).

Sontag, E. D. and Y. Wang, "New characterizations of input to state stability," *IEEE Trans. Auto. Cont.*, **41**, 1283–1294 (1996).

Sontag, E. D., "Smooth stabilization implies coprime factorization," *IEEE Trans. Auto. Cont.*, **34**, 435–443 (1989a).

Sontag, E. D., Some connections between stabilization and factorization, In *Proc. IEEE Conf. Decision and Control*, pages 990–995, Tampa. IEEE Publications (1989b).

Sontag, E. D., "Comments on integral variants of ISS," *Sys. Cont. Let.*, **34**, 93–100 (1998a).

Sontag, E. D., *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer-Verlag, New York, second edition (1998b).

Sontag, E. D., Stability and stabilization: Discontinuities and the effect of disturbances, In Clarke, F. H. and R. J. Stern, editors, *Nonlinear Analysis, Differential Equations and Control (Proc. NATO Advanced Study Institute, Montreal, Jul/Aug 1998)*, pages 551–598, Dordrecht. Kluwer (1999).

Sussmann, H. J., E. D. Sontag, and Y. Yang, "A general result on the stabilization of linear systems using bounded controls," *IEEE Trans. Auto. Cont.*, **39**, 2411–2425 (1994).

Teel, A. R. and L. Praly, "On assigning the derivative of a disturbance attenuation control," *Mathematics of Control, Signals and Systems*, **13**, 95–124 (2000).

Teel, A. and E. D. Sontag, "Changing supply functions in input/state stable systems," *IEEE Trans. Auto. Cont.*, **40**, 1476–1478 (1995).

Teel, A. R., D. Nešić, and P. V. Kokotović, A note on input-to-state stability of sampled-data nonlinear systems, In *Proc. 37th IEEE Conf. Decision and Control*, pages 2473–2479, Tampa (1998).

Teel, A. R., "Connections between Razumikhin-type theorems and the ISS nonlinear small gain theorem," *IEEE Trans. Auto. Cont.*, **43**, 960–964 (1998).

Tsinias, J. and I. Karafyllis, "ISS property for time-varying systems and application to partial-static feedback stabilization and asymptotic tracking," *IEEE Trans. Auto. Cont.*, **44**, 2173–2184 (1999).

Tsinias, J., "Stochastic input-to-state stability and applications to global feedback stabilization," *Int. J. Control*, **71**, 907–930 (1998).

Vorotnikov, V. I., "Stability and stabilization of motion: Research approaches, results, distinctive characteristics," *Automation and Remote Control*, **54**, 339–397 (1993).

Hybrid Systems in Process Control: Challenges, Methods and Limits

Stefan Kowalewski *
Corporate Research and Development
Robert Bosch GmbH †
Frankfurt, Germany

Abstract

Hybrid dynamics in process control arise for various reasons. The main source is the interaction between computer-realized discrete control functions and the continuous physico-chemical processes in the plant. But also physical or operational constraints lead to discrete phenomena in otherwise continuous processes. The modeling, the analysis and the design of hybrid control systems pose new problems for which the conventional systems and control theory is not an appropriate tool. The paper identifies these challenges using illustrative case studies and gives a survey of recent approaches to a formal treatment of hybrid systems. Particular emphasis is put on methods which were developed in computer science in the last years. We will discuss to what extent these techniques can support the analysis and design of hybrid systems in process control or automation and what the limits are.

Keywords

Hybrid systems, Hybrid automata, Safety analysis, Verification, Synthesis, Fault diagnosis

Introduction

The term "hybrid systems" was coined around ten years ago and refers to systems with both continuous and discrete dynamics. It is now a common research field in control theory and computer science with regular conference and workshop series. The purpose of this paper is to present a survey on approaches from this field and discuss their potential for solving problems in process control or automation. In particular, we will focus on models and methods from computer science research which do not receive much attention from the process control community.

In computer science, the treatment of hybrid systems belongs to the field of *formal methods*. The term formal method refers to approaches which systematically use means of description with rigorously and consistently defined syntax and semantics to represent design specifications or existing designs/plants for the purpose of either proving that certain design requirements hold or automatically synthesizing designs which provably meet these requirements. Examples of methods which in computer science are commonly agreed to be formal are *verification* or *synthesis*, whereas *simulation* (in the engineering sense of the word) or *optimization* usually are not regarded as being formal.

In the last years formal methods have found their way from academia to industry in several domains. Applications are reported, for instance, from the design of integrated circuits, communication protocols, embedded controllers for airplanes and cars, logic control of machine tools, and flexible manufacturing systems. The use of formal techniques in these fields is motivated by the need for discovery of faults in early design stages which helps to save costs and to achieve more reliable systems. However, this development does not seem to have reached the processing industries because from this domain only a few academic projects and hardly any industrial application of formal methods are known. The reason which is often given for this situation is that the first formal methods were applicable only to purely discrete problems (e.g., correctness of switching circuits) whereas problems in processing systems usually involve plants with continuous dynamics. Now that new formal methods have been developed for hybrid systems, it seems to be worthwhile to revisit the question of what can be achieved by formal methods in process control or automation. This is the motivation of this contribution.

The paper is organized as follows. In the following section we will present a pragmatic definition of a hybrid system and discuss why hybrid models arise. After that, the hybrid automaton model is introduced with the help of an example. In the section *Analysis of Hybrid Systems* reachability analysis is presented as the main analysis method. We discuss computational issues and the need for abstraction which arises as a consequence from the computational limitations. Based on these theoretical contemplations, the section *Hybrid Systems Problems* presents an analysis of the status and and the potential of formal methods for hybrid systems in the processing industries. Four possible application domains are identified, the current practice in these fields is described, existing academic initiatives for developing methods and tools are reported, and reasons are discussed why the practitioners are still reluctant to use them. A discussion concludes the paper.

What are Hybrid Systems?

The definition of a hybrid system as a system which combines continuous and discrete dynamics is a bit superficial. To be more precise, the term "hybrid systems"

*Robert Bosch GmbH, FV/SLD, P.O. Box 94 03 50, D-60461 Frankfurt, Germany. Email: stefan.kowalewski@de.bosch.com

†The results and opinions expressed in this paper were developed while the author was with the Department of Chemical Engineering at the University of Dortmund, Germany.

refers to models, not systems as such. A system is not hybrid by nature, but it becomes hybrid by modeling it this way. Whether it makes sense to build a hybrid model depends not only on the system, but also on the application and the purpose of the model. For example, it would be possible to describe the behavior of on/off valves by a continuous relation between the opening ratio and the control voltage. But in most situations it is sufficient to consider on/off valves as discrete switching elements. And if these valves interact with a process in which continuous phenomena are of interest, a hybrid model would be appropriate. In other words, it depends on the level of abstraction which is needed to solve a particular problem, whether the model should be discrete, continuous or hybrid.

As a consequence of this definition, hybrid systems arise in processing systems whenever both abstraction levels—continuous and discrete—have to be considered. The main source of such problems are constellations in which discrete (or logic) controllers interact with continuous physico-chemical processes. The need for the discrete abstraction level comes from the man-made control functions (but also from physical constraints), and the continuous level is necessary to describe the processes driven by natural laws. However, to stress again the fact that the appropriate model depends on the purpose: not every computer-realized controller has to be described by a discrete model, and not every chemical process needs a continuous model. If we consider continuous controllers being implemented on a digital computer, it usually is sufficient for the design to describe controller and plant as continuous systems and to abstract from the discrete implementation. On the other hand, for the analysis of a logic control function it is customary to use a discrete abstraction of the continuous process (e.g., only distinguishing between specified and unspecified process states).

The second example points to a main topic in hybrid systems research: abstraction methods. Often, it is either not possible or not appropriate to analyze a hybrid model. In this case it is often helpful to discretize the continuous part and solve the problem using discrete analysis techniques. So, hybrid systems research is not only concerned with hybrid models and their analysis, but also with the problem how to map hybrid problems into spaces where they can be solved better. This will be discussed in more detail in the section *Analysis of Hybrid Systems*. The section *Hybrid Systems Problems* will present examples for hybrid problems in the design and analysis of processing systems. There we will see that it is not only discrete logic but also the human distinction between certain modes (e.g., safe, unsafe, or tolerable process states; start-up, production, or shut-down phases; specified, unspecified, or possibly deviated process behaviors) which leads to hybrid models.

Modeling of Hybrid Systems

Two Different Starting Points and Directions

Modeling frameworks for hybrid systems were introduced independently in control theory and computer science. Since the original domains of interest in these two fields were on both ends of the hybrid dynamics spectrum—purely continuous dynamic systems in control theory, discrete state systems in computer science—the modeling approaches moved in opposite directions:

- In control theory, the starting point was continuous models which were then extended by discrete mechanisms like switching or resetting. The resulting models consist of differential equations, algebraic equations and/or inequalities with continuous and binary variables. The later are used to activate and deactivate terms by multiplication, e.g. to switch the right hand side of the state equation. This class of systems is often refered to as *switched continuous systems*.
- The computer scientists came the other way. They extended discrete formalisms (most prominently finite automata, but also Petri nets and logics) by continuous variables which evolve according to differential equations associated with discrete states. Discrete transitions can switch between continuous modes, and the continuous variables can be reset when a transition takes place. The resulting framework is called *hybrid automata*.

In principle, both modeling approaches are equivalent in the sense that the models are equally expressive[1]. In the following, we will focus our attention on the hybrid automata paradigm from computer science and discuss their potential in process control and automation. For more information about the "control theory approach" to hybrid systems the reader is referred to the literature, in particular to recent special issues on hybrid systems in various control journals (e.g., Antsaklis and Nerode, 1998; Morse et al., 1999; Antsaklis, 2000), or to Lemmon et al. (1999).

It should be mentioned that timed discrete event systems (i.e. models which consist of a discrete transition system and time as the only continuous variable) are often not regarded as hybrid systems but rather classified as discrete event systems. If this distinction is made, systems are called hybrid only when more complex differential equations than $\dot{x} = 1$ are involved. In this survey we will include the application of timed models as part of the general discussion of hybrid systems, because time is a continuous variable, and in many cases, timed models are a sufficient abstraction of hybrid dynamics.

[1] Of course, certain assumptions have to be made. For example, the value sets of the discrete variables in the switched continuous model have to be finite because the discrete state space of a hybrid automaton is finite by definition.

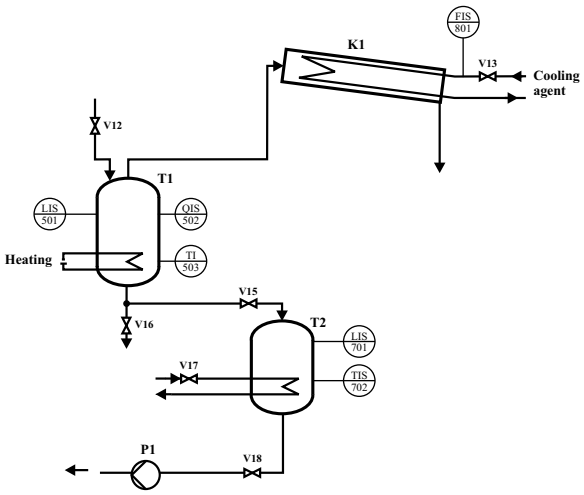


Figure 1: Example.

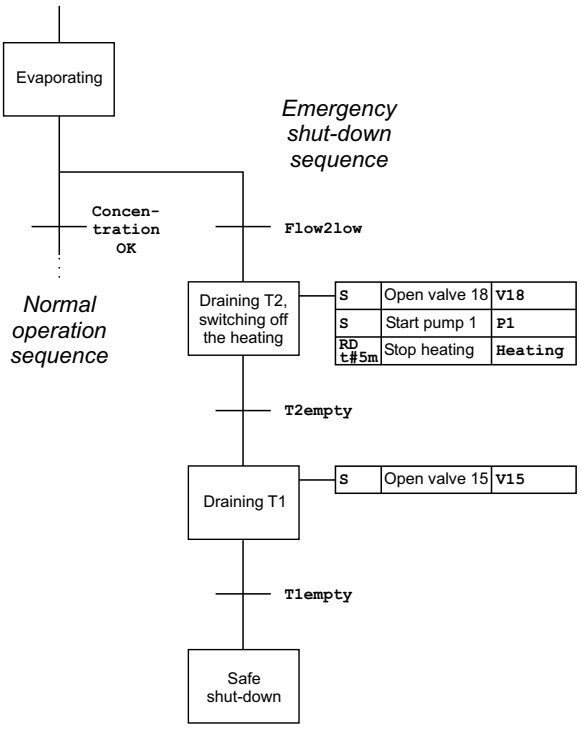


Figure 2: Sequential function chart.

An Example

To introduce hybrid automata we will use a simple example which illustrates that this modeling approach arises naturally for certain classes of process control or automation problems. The example is taken from Kowalewski et al. (2001) and has been used as a benchmark case for instance in the European Union research project *Verification of Hybrid Systems* (VHS, 2000).

Figure 1 shows the piping and instrumentation diagram of the example plant. It is a batch evaporator in which the following production sequence takes place. First, a solution is filled into tank T1 and the solvent is evaporated until a desired concentration of the dissolved substance is reached. During the evaporation stage, the condenser K1 is in operation and collects the steam coming from T1. When the desired concentration is reached, the material is drained from T1 into T2 as soon as T2 is available (i.e., emptied from the previous batch). A post-processing step then takes place in T2, before the material can be pumped out of T2 to a downstream part of the plant.

We focus our attention on the problem of an appropriate reaction of the controller to a cooling breakdown in the condenser. This failure will lead to a temperature and pressure increase in the condenser tube K1 and the evaporator tank T1, if the evaporation process is continued. It must be avoided that the pressure in K1 will rise above a dangerous upper limit. To achieve this, the heating in T1 has to be switched off before the safety pressure valve is triggered. This, in turn, causes a decrease of the temperature of the material in T1. When the temperature in T1 becomes too low, a crystallization effect leads to precipitation of solids, which spoils the batch. This, of course, is an undesired situation, too. Thus, the timespan between the cooling failure and switching-off of the heating is critical: it has to be short enough such that the pressure increase is limited but on the other hand as long as possible such that crystallization will not occur. Any given control program for this process has to be checked against this specification.

Figure 2 shows a suggestion for a controller which shall realize the required behavior. The representation in Figure 2 is a Sequential Function Chart (SFC) according to the IEC 1131 standard (IEC, 1992). In particular, action blocks are used to specify the control actions performed in each step. An action block consists of a qualifier, an action name, and a manipulated variable. The left branch of Figure 2 represents the normal operation sequence. In the undisturbed case, the system will leave the step *Evaporating* when the desired concentration is reached. The part branching to the right describes the control actions during an emergency shut-down as a reaction on a cooling breakdown. There is a waiting time implemented between the cooling failure and stopping the heater. As discussed above, the controller will open valve 18 and start pump 1 to drain T2 as soon as the flow of cooling agent is too low. The corresponding actions are labeled with the qualifier "S", which means that the variables V18 and P1 are set to TRUE and that they will remain TRUE after the step was left. The label "RD t#5m" on the third action in the step *Draining T2, switching off the heating* symbolizes that the heating is switched off with a delay of 5 minutes, even if the step has become inactive in the meantime. As soon as T2 is empty, the SFC will switch to the subsequent step and

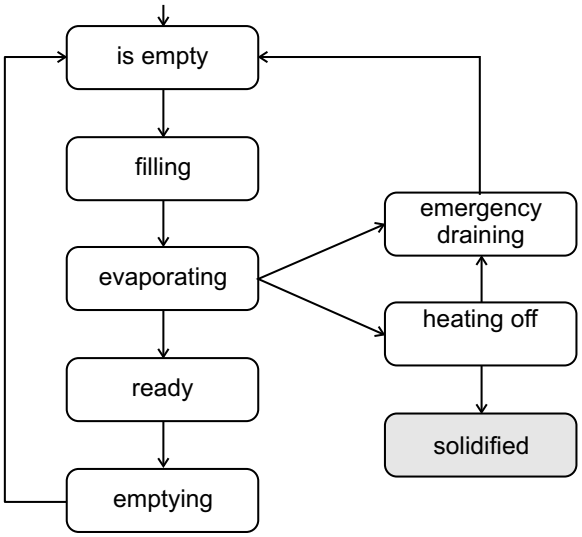


Figure 3: Finite automaton model of the evaporator tank T1.

open valve 15. When T1 is empty and the heating is switched off, the controller assumes the system to be in a state of safe shut-down.

Hybrid Automata

When we want to analyze whether the controller in the example from the previous subsection meets the requirements (i.e., the pressure will never become too high and crystallization will not occur), a model of the plant is needed which represents the relevant behavior. Obviously, a large part of this behavior is discrete because the interaction between controller and plant is specified in discrete terms. Figure 3 shows a corresponding finite automaton model of tank T1 of the evaporator. The discrete states represent the discrete steps during the evaporation cycle (*is empty, filling, evaporating, ready, and emptying*) as well as three important modes in the case of a cooling failure (*emergency draining, heating off, and solidified*). When the failure occurs, T1 will be in the state *evaporation*. The controller will then empty tank T2 and start the clock. This is not captured by the model of Figure 3 but happens in the environment (which would have to be modeled for a formal analysis, too). If T2 is empty before the waiting time is elapsed, the system will move to *emergency draining*, stop the heating, and safely go back to *is empty*. The more interesting case is when the waiting time is elapsed before T2 becomes empty. In the model of Figure 3 this means that the system moves to the state *heating off*. In this state the temperature in T1 decreases while the controller waits for T2 to become empty. Now the question is whether T2 will become empty before the temperature drops below the crystallization threshold. If this is the case, the transition to *emergency draining* will be taken and the system is safe. If not, the state *solidified* will be reached, representing that the crystallization has started.

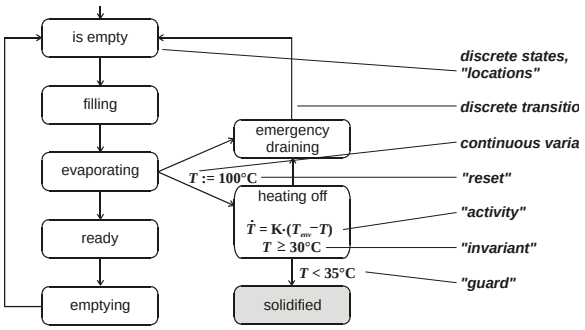


Figure 4: Example of a hybrid automaton.

Obviously, a purely discrete model like the one above is not able to answer this question. What is needed here is the information about how fast the temperature in T1 decreases. Thus, a hybrid model is necessary. Figure 4 shows how the additional information can be added to the discrete model by using a hybrid automaton (Alur et al., 1995). Roughly speaking, the hybrid automaton model complements finite automata by continuous variables. These variables can be reset by discrete transitions (here: the temperature is defined to be 100° Celcius when the heating is switched off). While the system is in a certain discrete state (called *location*), the continuous variables evolve according to differential equations, called *activities* (here: a linear first order ODE). Conditions can be formulated which have to be true while the system remains in a discrete state. They are called *invariants*. Finally, so-called *guards* represent conditions for transitions between discrete states. In the example the invariant and the guard together express that the transition to *solidified* will occur between 35° and 30° Celcius.

Figure 5 shows an important special class of hybrid automata, a so-called *timed automaton* (Alur and Dill, 1990). Here, the continuous variable is a *clock* which can only be reset to zero and the value of which increases by the rate of one. Such a description is useful for example when detailed continuous models are not available or not necessary, but information about process durations is known.

Analysis of Hybrid Systems

The major analysis procedure for hybrid automata is the *reachability analysis*. It answers the question whether for a given hybrid automaton a certain hybrid state (discrete location and region in the continuous space) is reachable from the initial state. This problem is so important because many problems can be reduced to a reachability problem. For instance in the example above, the question whether the controller is correct is mapped into the

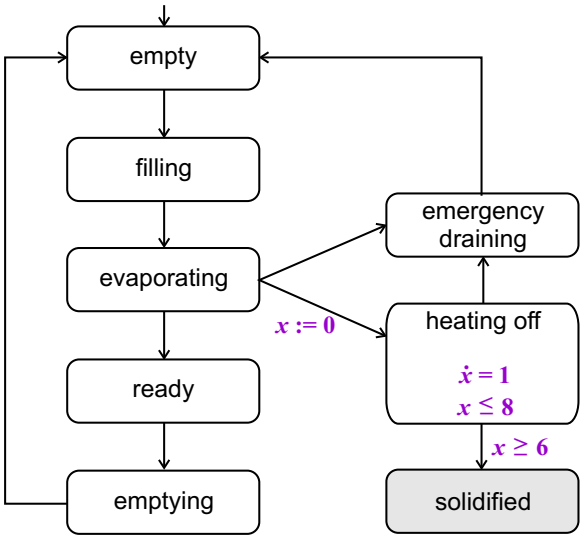


Figure 5: Example of a timed automaton.

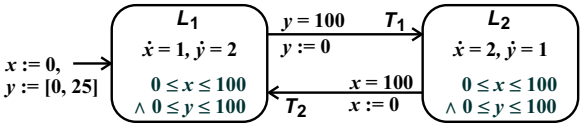


Figure 6: Example for reachability analysis.

problem whether the state *solidified* is reachable in the plant model.

However, the application of reachability analysis for hybrid automata is restricted due to computational issues. We will illustrate the basic computational problem by the small example in Figure 6 which is taken from Preußig (2000). It is a very simple hybrid automaton with only two locations and two continuous variables, x and y. In both locations, the invariants restrict the values of x and y to a square of the size $(0 \leq x \leq 100 \wedge 0 \leq y \leq 100)$. In location L_1, x grows by a rate of 1 and y by a rate of 2. In location L_2 it is vice versa. The transition T_1 can only be taken when the guard $y = 100$ is true. When it is taken, y is reset to 0. Transition T_2 is guarded by the condition $x = 100$, and the reset is $x := 0$. Since the invariants, guards and resets are linear expressions and the solutions of the differential equations for the activities are linear functions, this example belongs to the class of *linear hybrid automata* (Alur et al., 1995). It is the largest class for which tools for exact reachability analysis exist. The most prominent is *Hytech* (Henzinger et al., 1997). *Hytech* uses polyhedra as the data structure for representing, manipulating and storing regions in the continuous state space during the exploration. The basic reachability algorithm is the following:

0. Initialize hybrid automaton (location $l :=$ initial location, polyhedron $P :=$ initial continuous region P_0).
1. Intersect P with the invariant of l.
2. Let P increase with time according to the activity of location l.
3. Intersect P with the invariant of l.
4. Stop, if l was visited before with P.

   Else: For all transitions T from l do:

   5. Intersect P with the guard of T.
   6. Reset P according to the reset expression of T.
7. Set $l :=$ target location of T, go to step 1.

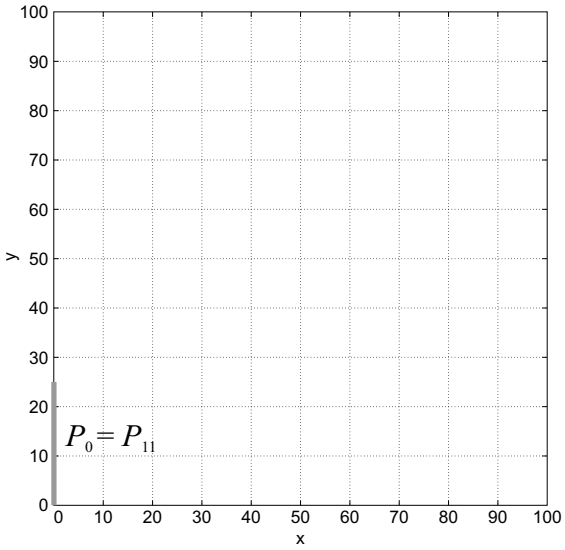


Figure 7: Situation after step 1 of the reachability analysis algorithm.

Figure 7 shows the result of the steps 0 and 1 of the algorithm: The initial values of x and y are $(x = 0 \wedge 0 \leq y \leq 25)$. Step 1 determines which of these values are possible in location L_1. In this case the set is unchanged and represented by the polyhedron P_{11} (the dark grey bar in Figure 7) The subscripts symbolize iteration 1 and step 1. In step 2, the region of the continuous state space is computed which can be reached while the system is in l, without considering the invariant (or, in other words, assuming that the system remains in L_1 forever). The result is P_{12} in Figure 8. Obviously, not all of this region is actually reachable, because the invariant will force the system to leave L_1 as soon as it is violated. Step 3 determines the part which is in accordance with the invariant, i.e. P_{13} in Figure 9.

At this stage, the algorithm would take the list of previously visited regions and check whether P_{13} (or a subset of it) had been computed for L_1 before. If this is case,

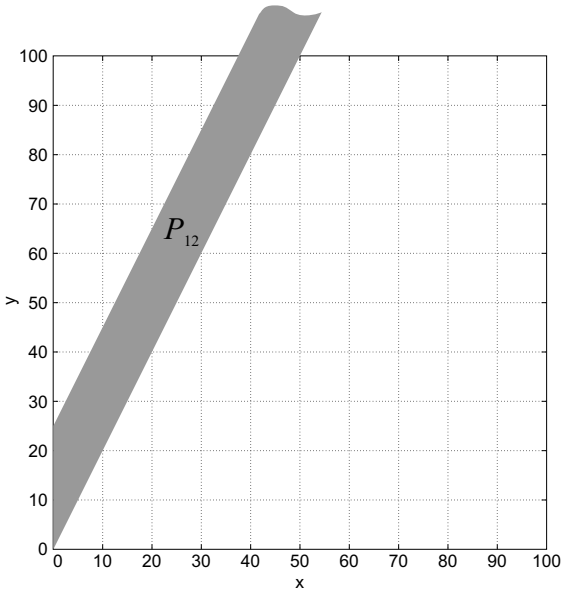


Figure 8: Result of step 2.

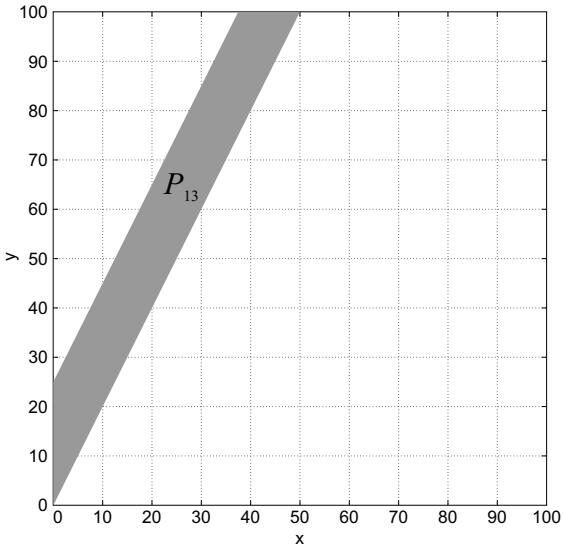


Figure 9: Result of step 3.

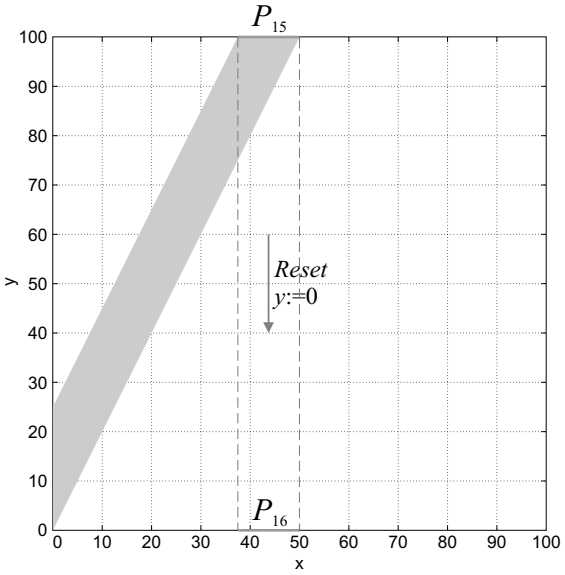


Figure 10: Results of step 5 and 6.

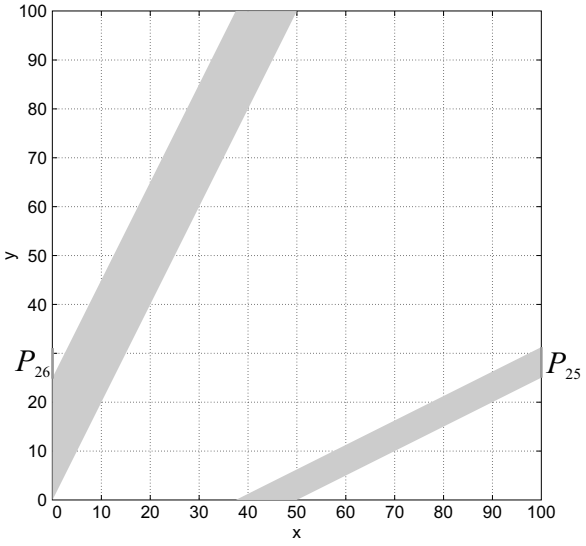


Figure 11: Result of iteration 2.

it would abort the current search branch. If not, the algorithm has to determine the possible transitions from L_1. In our example, there is only T_1. To find out whether it is possible, we have to check whether the guard can become true for the reachable values of x and y in L_1. This is computed in step 5 by the intersection. If it is empty, the transition is not possible. Here, the result is P_{15} in Figure 10. When the transition is taken, the reset leads to P_{16}. This is the possible region with which the system can enter location L_2. Now, the second iteration starts. Figure 11 shows the resulting polyhedra P_{25} and P_{26} (the light grey parts represent the already visited region).

If we continue the algorithm, it will become apparent that after each iteration the reachable set is increased by a polyhedron which becomes smaller and smaller. In fact, the length of the horizontal and vertical boundary lines is half of the length of the previous iteration. This converges asymptotically to the reachable set given in Figure 13. The consequence is that the algorithm will not terminate. This is not only true for this example—it was proven that reachability for linear hybrid automata is not decidable, which means that there does not exist

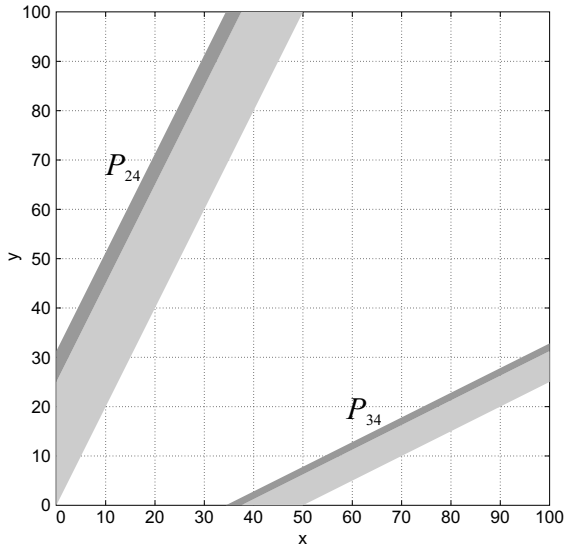


Figure 12: Results of iterations 3 and 4.

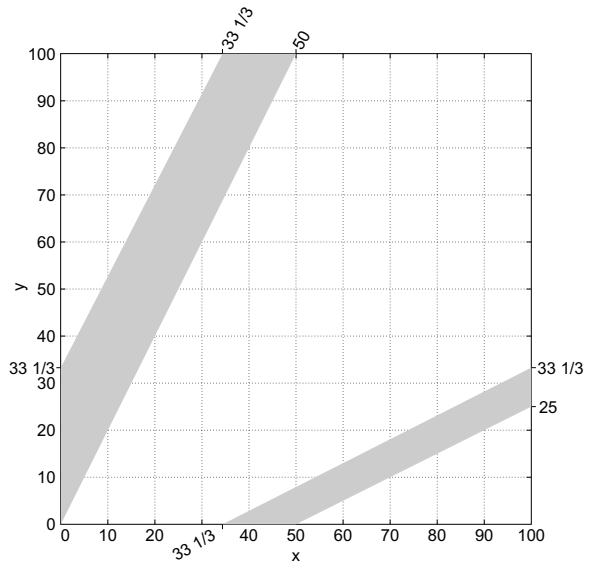


Figure 13: Reachability set.

an algorithm which will terminate for every linear hybrid automaton (Alur et al., 1995).

In practice, this undecidability result is not the only problem. Since *HyTech* is performing its computations by integer arithmetic, memory overflows for the integer variables are a common problem. A further shortcoming of this approach is the exponential complexity of the algorithm with the number of continuous variables (which adds to the discrete state space explosion problem). These shortcomings are the motivation for a very active area in hybrid systems research which is concerned with abstraction techniques for hybrid systems. The idea is to build a substitute model which is more abstract than the hybrid system under investigation (that is, it omits some details but includes the behavior of the original system), which belongs to a class of systems for which reachability is easier to solve (e.g., timed automata or purely discrete systems). A collection of papers following this approach can be found in Engell (2000). Alur et al. (2000) provided a survey on fundamental theoretical results in discrete abstractions of hybrid systems. A slightly different approach is not to do the abstraction for the whole dynamics first and then analyse reachability, but to apply approximation techniques during each single iteration step of the reachability analysis. Main representatives of this work are for instance Chutinan and Krogh (1999), Dang and Maler (1998), Greenstreet and Mitchell (1999), Raisch and O'Young (1998), and Preußig et al. (1999).

Hybrid Systems Problems

In this section we identify a list of process control problems (or in a wider sense, process automation problems) which pose challenges to formal methods for hybrid systems. The problems are presented in the order in which they arise during the life cycle of a processing plant. The project management for processing plants is built on the following life cycle model. Starting with a market analysis, the early phases take place in laboratories and pilot plants to increase the knowledge, identify the optimal chemical or physical processes for realising the desired product and estimate the economic prospects. After that, the so-called *basic engineering* begins during which the unit operations (e.g. reactions, separations) are fixed, the corresponding types of equipment are chosen, and the basic means for controlling the plant are specified. The major resulting document of this phase is the so-called *Piping and Instrumentation Diagram* (P&ID). This is a flowsheet representing all pieces of processing equipment like tanks, valves, reactors etc., all the pipes, the measurement points, as well as the continuous control loops and the safety trips (the latter two being represented roughly as connections between the initiating sensor and the corresponding actuator). Following the basic engineering is the *detail engineering* phase. Here, the information of the P&ID is specified and refined to obtain the necessary documents for *erection* (e.g., pipe routing, equipment sizes, choice of materials etc.) and *commissioning* (choice of sensors and control systems, configuration/programming of the controllers, design of operator panels, etc.). Often the boundaries between these phases are fuzzy. It is not uncommon, for example, that the programming of the control system has to be finished on site while the plant is being put into operation. After the commissioning, *operation and maintenance* is the next phase, and *decommissioning and dismantling* is regarded as a phase

of its own which finishes the life cycle of a processing plant. Refering to this life cycle model, the following tasks with the potential for the application of hybrid systems methods can be identified.

- safety analysis of plant instrumentations (*basic engineering*)
- design and analysis of discrete controllers, e.g., interlocks, trips, switching of continuous control modes, sequence control (*detail engineering, commissioning*)
- generation and analysis of control recipes for batch processes (*operation*)
- event-based fault diagnosis (*operation and maintenance*)

In the remainder of this section, the four tasks will be discussed in more detail.

Safety Analysis of Plant Instrumentations

Task Description. Many processing plants represent a potential danger for the life and health of the people working in the plant and those living in their vincinity and for the environment. Therefore, in most countries the national authorities demand a thorough and systematic analysis of the possible hazards and the safety concept of a new plant. This procedure has to take place at the end of the basic engineering phase because it is important to identify safety-related design faults in an early stage of the project, so that modifications can be done with little effort. The aim of the safety analysis is to discover potential sources of hazards and to check that the proposed safety devices are capable of handling these hazards in an appropriate way.

Clearly, many safety problems are of a hybrid nature. It often depends on the interaction between discrete safety devices (e.g., limit switches, relief valves) and the continuous dynamics (e.g., an exothermal reaction) whether dangerous situations can occur or whether the instrumentation can prevent them.

Current Practice. In practice, the safety analysis is carried out by a team of experts coming from different disciplines involved in the plant design. These experts get together and discuss the safety concept based on the P&ID diagram. They literally start at one end of the plant and move from one piece of equipment to the other. Disturbances and equipment failures are assumed and the consequences are estimated. The results are documented thoroughly. This procedure can take weeks.

To provide guidelines for these discussions, companies and authorities have developed a handful of systematic methods for safety analyses. Among those the Hazard and Operability Studies (HAZOP) methodology (Lawley, 1974) is now established as the most popular one. In HAZOP, a list of guidewords (e.g., "none", "more than", or "reversed") is applied to each piece of equipment and each relevant physical variable such that all possible deviations of the nominal values can be determined. However, even for a medium sized plant it is still impossible to consider all cases being brought up by the HAZOP method. The experts have to select what they think are the most important cases. This bears the danger of missing a possible hazard. Another problem is that hazard analyses are based on the "one failure at a time" assumption. This is also done to keep the number of considered cases manageable. However, hazards caused by combinations and sequences of failures will not be discovered this way. Finally, the consequences of the interaction between continuous and discrete dynamics can only be estimated. This can be problematic because responses of hybrid systems are much less predictable than purely continuous systems.

State of Research. Since at the time of the safety analysis the available information about the plant behavior is rough, a model-based analysis will have to build on qualitative models which abstract from the continuous dynamics. First approaches in this direction were presented by Vaidhyanathan and Venkatasubramanian (1995), Yang and Chung (1998) and Graf and Schmidt-Traub (1998). In all cases, the idea is to follow the HAZOP methodology and to provide an automated checker for the guideword questions which uses a quantitative plant model. In Vaidhyanathan and Venkatasubramanian (1995) and Yang and Chung (1998) the model is built on signed directed graphs, in Graf and Schmidt-Traub (1998) state transition systems are used which are specified and simulated using the tool *Statemate* (Harel and Naamad, 1996). In the latter case, the HAZOP method essentially is mapped to a reachability problem for discrete state transition systems. This points to the similarity of the safety analysis for qualitative models and formal verification of data processing systems (see next section). It can be expected that more approaches from this field will be applied to hazard identification and safety analysis in the future, including the application of the hybrid analysis method presented in the first part of this paper.

Discussion. From the description of the currently applied procedures for hazard and safety analysis it becomes clear that there is a potential for improving the current practice by applying formal methods in the sense of an automated, rigorous model-based safety analysis which could then overcome the problem of overlooked hazards as well as the single failure assumption.

The approaches mentioned in the previous section are promising because they provide a means to automatically check the consequences of large numbers of deviations or failures and combinations of them. Already during the modeling step, valuable insight will be gained about

the safety of the plant design. And when reachability analysis algorithms are applied, it is guaranteed that every hazard being considered in the model—explicitly or implicitly—will be discovered. However, before these approaches can be applied in practice, two problems have to be solved. First, the modeling effort is very high and for the practitioner it is often not apparent whether it will pay off during the analysis. Second, the automated analysis will result in a large number of scenarios leading to dangerous states differing only in small details which would have been considered as one hazard in a manual analysis. Also, due to the intrinsic nondeterminism of qualitative models, many "dangerous" trajectories are physically impossible. To overcome these problems, methods for filtering the analysis results before they are presented to the user have to be developed.

Design and Analysis of Discrete Controllers

Task Description. In the process industries, discrete controllers are realized by programmable logic controllers (PLCs) or distributed control systems (DCSs). They perform basic control functions, e.g., process supervision to avoid unwanted or even dangerous states of the process or damage to the equipment, sequence control, startup and regular or emergency shutdown procedures, switching the mode of operation of PID controllers, or supervision of sensor inputs and outputs to actuators. A significant part of this discrete control logic is critical for the safe operation of the process. The correctness of the logic control system often depends on the interaction between the discrete control function and the continuous process.

Current Practice. In practice, discrete control logic is still very seldom produced in a systematic manner. Usually, rough and often incomplete specifications lead to a first design and direct implementation which is then improved by testing onsite. The support of the software design is restricted to programming environments and a standard for programming languages and common standard software elements (IEC (1992), see Maler (1999) for a critical analysis). This situation seems to be inappropriate in particular for the citical parts of the logic control software.

State of Research. The research on a more systematic and reliable design for logic controllers in processing systems follows a number of approaches. Among those, the following three can be regarded as the most important methods with a formal basis: verification, synthesis, and code generation from a formal specification.

Formal Verification. The notion of formal verification originates from computer science where, in general terms, it means a mathematical proof that a model of an algorithm fulfills given formal properties. In the last decades, different representations and methods have been developed and, in the recent years, some of them have been applied successfully in the area of hardware and communication protocol design (see Clarke and Kurshan, 1996, for a survey). However, in the field of logic controllers for processing systems, in particular when continuous processes are involved as it is the case for most processing plants, formal verification is currently not applied in practice and only a few research projects are described in the literature (see below).

The pioneering work in the verification of logic controllers for industrial and in particular, chemical processing systems was done by Powers and his co-workers (1992) for control programs represented in Relay Ladder Logic. They applied the symbolic model checking method from Burch et al. (1992) in which the system that has to be verified is modeled as a finite state machine and the specifications of the desired behavior are represented by temporal logic expressions. Their approach has been extended to include plant models and recent work shows that the formal verification of logic controllers for processes of moderate size is possible (Probst et al., 1997). The continuous dynamics of subsystems of the plant are discretized in an elementary manner: The range of values of each continuous variable is partitioned into intervals and the model simply describes the possible transitions between these intervals. Timers are incorporated in the same qualitative fashion by neglecting the timer value and keeping only two states, running and elapsed. This suffices if checking real-time constraints is not required, e.g., whether a controller response to a plant event is fast enough to avoid unwanted process behavior.

In the last years, the model checking approach has been extended to incorporate real-time and hybrid specifications and analysis. This work is mostly based on the timed or hybrid automata model and it resulted in the development of analysis tools for such systems (Yovine, 1997; Larsen et al., 1997; Henzinger et al., 1997). Applications of this framework to the modeling and analysis of PLC programs are reported, for example, in Mader and Wupper (1999) and VHS (2000). In Kowalewski et al. (1999) an approach is presented which integrates a number of available model checking tools from computer science with an engineering-oriented modeling interface. For this purpose, a modular, block-oriented framework for modeling hybrid systems is introduced which is based on hybrid automata connected by certain discrete valued signals. It aims at applications where formal verification requires a quantitative analysis of the interaction between timers or threshold values in the logic control program and the plant dynamics. Further examples of verification approaches to PLC programs including time or hybrid aspects are Herrmann et al. (1998) and Heiner et al. (1999).

A recent development in the chemical engineering community is represented by approaches to apply tech-

niques from optimization, in particular mathematical programming, to verification problems. For instance, the same basic problem as in the approach by Moon et al. (1992), i.e. model checking of Relay Ladder Logic diagrams, has been treated by Park and Barton (1997) using these techniques. Here, the problem of checking whether a temporal logic formula holds for the model of the logic controller is mapped into the feasibility problem for a system of equalities and inequalities for binary variables. The latter is then solved by integer programming. Application results show that there exist examples for which this approach generates solutions faster than classical model checking. The problem of abstracting continuous or real-time information is solved in the same way as in the first approach.

The use of mathematical programming not only for checking the control logic but also for the analysis of switched continuous models has been suggested by Dimitriadis et al. (1996, 1997). The reachability problem is reformulated as an optimization problem in the discrete time domain which can be solved by mixed integer programming. Basically, the optimization determines the worst possible behavior, meaning that the system is most often in an undesired region of the continuous state space. The approach is general in the sense that it can be applied to hybrid systems as well as to purely discrete or purely continuous systems. Its strength lies in the ability to take advantage of well tested and efficient optimization procedures. A limitation is given by the fact that the size of the mixed integer program grows with the product of the number of discrete time steps and the number of equations and logical expressions describing the plant and the controller, respectively. A similar approach has been suggested by Bemporad and Morari (1999). Here, an iterative scheme is used to perform conventional reachability analysis. This scheme avoids setting up a huge one-step optimization problem which is most likely not tractable. The verification method is part of a comprehensive modeling and analysis approach to hybrid systems, including a scheme for model-predictive control which is reported in another paper of the CPC VI.

Synthesis. The synthesis of logic controllers is a transformation of the continuous control synthesis problem to a discrete event or hybrid setting: Given a model of the possible plant behavior for arbitrary control inputs and a specification of the desired process behavior, a controller has to be designed which guarantees that the closed-loop system satisfies the requirements. When the term synthesis is used in the research literature for this task, it is usually understood that the controller is derived from the plant model and the specifications *automatically* by an appropriate algorithm. Due to a traditional and deliberate mutual isolation of the control systems and computer science research communities, control engineers refer to the framework set up by Ramadge and Wonham (1989) whereas computer scientists trace the roots back to early game theory (Buechi and Landweber, 1969) when the origins of research on logic controller synthesis are concerned. In both cases, the problem formulation and algorithms are basically equivalent.

In the Ramadge and Wonham framework, the plant is modeled as a finite state machine in which the transitions represent discrete events. They either can be inhibited by the controller or are uncontrollable in the sense that their occurence is not under influence of the environment. The desired behavior can be specified by a set of acceptable event sequences, a so-called *target language* (Ramadge and Wonham, 1987b), or a set of undesired states which corresponds to the forbidden states specification mentioned in the previous section on formal verification (Ramadge and Wonham, 1987a). For both requirements, Ramagde and Wonham present synthesis algorithms with linear complexity refering to the size of the plant model. However, the size of the plant model itself increases exponentially with the number of components from which it is built.

There exist numerous extensions and derivations of this approach from which only a selection can be mentioned briefly here. While in the original approach the complete controller is computed off-line, Chung et al. (1992) introduce an on-line synthesis approach using a limited lookahead horizon. The fact that the controller can only inhibit plant events but not force them to occur is often regarded as a shortcoming of the Ramadge and Wonham framework. In Golaszewski and Ramadge (1987), Sanchez (1996), Kowalewski et al. (1996), and Krogh and Kowalewski (1996) synthesis concepts are investigated where the controller can pre-empt events by forcing other events. A further line of research is concerned with Petri nets as an alternative modeling paradigm. Examples are Hanisch et al. (1997), Li and Wonham (1993, 1994), and Chouika et al. (1998). A survey is presented in Holloway et al. (1997). Applications of formal synthesis approaches for process control problems can be found, for instance, in Tittus (1995) and Marikar et al. (1998).

An important extension is the incorporation of quantitative time, such that additional specifications for meeting deadlines or minimal state residence times can be formulated for the synthesis. This is useful for process control applications because timers are often used when information about the evolution of the continuous process dynamics is not available from measurements. Two main approaches can be distinguished: The first one is based on a discrete time axis and assumes that state transitions only occur synchronously to clock ticks. This leads back to the original untimed Ramadge and Wonham setting (Brandin and Wonham, 1994). The second approach is based on timed automata. Here the time axis is continuous and the problem is formulated as a

game between controller and plant (Asarin and Maler, 1999; Wong-Toi, 1997). Any winning strategy can serve as a specification for a correct controller.

As it was the case for the formal verification, the application of the discrete or timed synthesis methods to process systems requires the substitution of models with continuous dynamics by discrete or timed models. Systematic ways to obtain valid approximations for the synthesis of controllers are reported in Chutinan and Krogh (1999) and Raisch and O'Young (1998).

Code Generation from a Formal Specification. A more pragmatic approach to systematic design of logic controllers is represented by top-down design methods where formal specifications of the control code are automatically translated into control code. In principle, this procedure does not exclude a step in which the specification is checked against problem-specific requirements by formal verification before it is fed into the code generator, or a final proof that the generated program is in accordance with the specification. However, in the pertinent research known from the literature this is not yet done. The PLC code is either generated from the result of a synthesis (see above, Hanisch et al., 1997; Marikar et al., 1998) or the specification is set up by hand and analysed only with respect to general properties which are not specific to the particular control problem, like absence of deadlocks or reversibility of the initial state. The latter approach is mostly based on Petri net representations (Frey and Litz, 1999).

Discussion. It is safe to say that in the field of logic controllers for processing systems, in particular when continuous processes are involved as it is the case for most processing plants, the presented academic approaches for verification or synthesis are currently not applied in practice. There are several apparent reasons for this situation. The first one is that PLC or DCS software is usually developed by engineers and not by computer scientists. Thus, the developers are not familiar with the available formal methods and tools. Another reason is that most control software in this area is written by the PLC or DCS user for one particular application or a rather small number of similar plants. This is a different situation to hardware or protocol design where the software is part of a mass product and the verification effort pays off more easily.

Regarding the prospects of the verification approaches mentioned above, it has to be noted that the application scope is limited by the computational cost of the analysis and, again, the necessary effort to build the models. Therefore, more effort has to be invested mainly to increase the efficiency of verification algorithms and to support the modeling process, possibly by interfaces to already existing process information (P&ID, data sheets).

Generation and Analysis of Control Recipes for Batch Processes

Task Description. Batch processes follow a stepwise production sequence. While in continuous processes the input and output material is continuously flowing into and out of the plant, respectively, batch processes are characterized by the fact that discrete amounts of material, so-called *batches*, are processed one after the other (and possibly in parallel). Often, batch plants are designed for a flexible production of multiple products. Each product has its own processing sequence, the so-called *recipe*, which can be realized in the batch plant usually in more than one way. It is also often the case that multiple batches, possibly of different products, can be produced in parallel.

For the management and control of such plants, the concept of *recipe control* has established itself in industry and is described in the standard (ISA, 1995). The main idea is to assign a *basic recipe* to each product which specifies the necessary production steps and their sequence independently from the specific equipment available in the particular plant. Each step is described on the level of *process operations* like, for instance, mixing, heating, reaction, or separation. When the production management decides that a certain amount of a product has to be produced, the basic recipe is taken and to each process operation an appropriate plant unit (e.g., mixer, reactor, etc.) is assigned. This leads to the *control recipe*, which can be regarded as an executable control procedure for the realization of the basic recipe on a particular batch plant.

The problem arising in this procedure is that the assignment of specific equipment can lead to undesired consequences which are not easy to predict. While the eligiblity of plant units for certain operations (for instance with respect to size, available heating or cooling capacity, resistance of material etc.) can be checked statically, interference with other currently running recipes is much more difficult to analyse. In particular, it is possible that a ressource is assigned which will not be available when it is needed, because it is used by another recipe. While in manufacturing processes this situation usually only causes a delay, it can lead to the spoiling of the batch in a processing plant due to chemical and physical effects during the waiting period (think of cristallization). It is also possible that deadlocks occur due to bad equipment allocation.

Current Practice. Due to the broad use of the recipe control concept in industry, most DCS vendors offer software packages for the integration of recipe management and control into DCS applications for batch plants. These packages support editing of basic recipes, storage of recipes in data bases, and manual assignment of equipment from the controlled plant to steps of the recipes to generate control recipes. However, achieving

correctness of the control recipe in its environment (consisting of the plant and the already generated parallel control recipes) with respect to realizability and avoidance of deadlocks and critical waiting situations like the one described above, is left to the experience and intuition of the user.

State of Research. One solution to the problem described above is to simulate the execution of the control recipes on the corresponding plant. For this purpose, a discrete or hybrid simulator is necessary. While in principle any general purpose simulation package with discrete or hybrid capabilities can be used for this task, there are simulators available which specificly support simulation of recipe-driven batch plants. One example is the tool *BaSiP* (Fritz et al., 1998) which offers the possibility to specify recipes using standardized languages and to build the plant model from pre-defined blocks. For the simulation the user can choose between a discrete and a hybrid simulator or an interface to the commercial *gPROMS* package.

Most of the research in the operation of batch processes is concerned with optimization of production schedules. This problem is situated on a higher level than the one considered here. For the scheduling it is assumed that realizable control recipes are available and the task is to find optimal starting times and sequences to satisfy certain production goals. In this field mathematical programming has become one of the most popular tools. This line of research is mentioned here because it is possible to take operating requirements (like avoiding excessive waiting) into account as constraints in the optimization problem.

A lot of research activities addressing specificly the correctness and realizability of recipes in batch plants is built on timed Petri nets. Here, the pioneering work goes back to Hanisch (1992). Later approaches were reported, for example, in Tittus (1995). To check the correctness, the plant and the recipes are modeled by means of a Petri net and the available analysis techniques for this kind of model are used to determine, for example, deadlocks, reversibilty, or execution times of recipes. Meanwhile these approaches have been extended to incorporate the continuous aspect using so-called *hybrid Petri nets* (David and Alla, 1992).

A new approach to analyse or generate recipes takes advantage of the model checking tools for timed automata (see the previous section in this paper entitled *Design and analysis of discrete controllers* and also Fehnker (1999) and Niebert and Yovine (1999)). Here the idea is to model the recipe and the plant by nondeterministic timed automata. The composition will then yield a model representing any possible allocation of plant units to recipe steps and any conceivable sequence of recipe execution. Using the available reachability analysis algorithms, it is then possible to check whether certain states representing successful execution of a pre-defined number of recipes are reachable within a desired time. This approach can be used for optimization by an iterative procedure during which the desired time is decreased until no more realizable schedules can be found. However, these approaches are in their infancy and currently only applicable to downsized examples with a handful of recipes running in parallel.

Discussion. In the domain of batch processes, a lot of research is carried out and the extent to which research results have been transfered to industrial applications is larger than in the previous cases. However, this applies mostly to the optimization of schedules. The problem of finding a realizable and proper control recipe for a basic recipe and a plant is currently solved manually and not supported by formal methods. The two main reasons are the same as in the case of logic control design: modeling effort and tool performance. Usually, the companies already spent a lot of time to model the batch plant and specify the recipes. And since the mentioned simulation tools require different representations as their input, this means double effort for the industrial users. This is not regarded as worth it, in particular because experienced operators of batch plants often perform a rather good job in creating control recipes and schedules. However, considering the increasing complexity of batch plants, this situation may change and the incentive to automatically generate and analyze recipes and schedules may become stronger.

To make the academic tools more attractive to practioners, interfaces to model representations already existing for the DCS are certainly desirable. Also, the performance of the tools has to be increased. Here, the timed automata approach could gain importance if it will be possible to extend existing compositional reachability algorithms like Lind-Nielsen et al. (1998) to a timed setting.

On the methodological side, it could be fruitful to investigate the similarity between the generation of a control recipe from a basic recipe for a particular plant to the compilation of computer programs written in a high-level language to machine code executable on particular hardware. Concepts from computer science could then be used, for example, to analyze whether a control recipe is a valid "refinement" of the basic recipe, where refinement means adding more detailed information without violating the originally specified behavior.

Event Based Fault Diagnosis

Task Description. If the process variables deviate from their nominal values, usually a step-wise procedure starts before a safety device is actually triggered: First a warning is sent to the operator console, and if the value deviates further, an alarm is generated. This leaves some time to the operating personnel to identify the cause of

the fault and to get the process back into nominal behavior. In practice, discovering the cause of an alarm (or a warning) is based solely on the message associated to an alarm and on the experience and knowledge of the operators. In most cases, this is sufficient. However, it is possible that an alarm already represents an effect of a not directly determinable cause. And it is also often the case that one fault successively causes further deviations which results in a fast and long sequence of alarms on the operator screen. To determine the original fault and act correspondingly in a short time and under stress conditions may become an impossible task for the operator then.

Current Practice. The DCS supports the operating personnel in situations as described above only by identifying the plant unit and process variables which initiated the alarm, and possibly by displaying the current values or recent trends of related process variables. The analysis of these data is left to the operator.

State of Research. Methods for event-based fault diagnosis that automatically (and more or less instantly) determine the possible causing failure events are available from the field of discrete event systems research. Fundamental notions in this context are *observability*, *invertibility*, *testability*, and *diagnosability* of discrete event systems. A comprehensive, tool-supported approach for fault diagnosis of discrete event systems is presented in Sampaath et al. (1996) where a diagnoser can be generated which derives the set of the currently possible (and partly unobservable) plant behavior from the observable events.

Discussion. Tools to automatically determine causes of alarms or estimations of the current discrete plant state can be very helpful to minimize the time of plant shutdowns due to alarm trips and to reduce the risk of wrong interference of the operating personnel in an alarm situation. While these tools are available, their acceptance in industry can be achieved only if appropriate modeling environments will be integrated into the DCSs.

With respect to process systems, it is also desirable that the approaches are extended from a purely discrete to a timed or hybrid setting. This would make it possible to take the process dynamics into account and, for example, determine that certain variables change too fast or too slow by checking whether a limit switch event will occur before or after a certain time threshold.

Conclusions

This paper presented an introduction to formal methods for hybrid systems and discussed their potential with respect to problems in process control or automation. Four application fields were identified in the design and operation of processing systems where hybrid systems methods represent promising tools for a better support of the engineering process. The selection was made by the author, and it is likely that more tasks can be identified which can be supported in the same or similar way.

To summarize, it can be said that for each of these tasks appropriate methods and tools have been made available from academic researchers. However, they are currently not applied in industry. The reasons are the same for each task: First, the existence of the methods is unknown to many practitioners. And if a process engineer gets in touch with a formal method, the often inaccessible nature of its semantics and nomenclature will probably prevent her or him from applying it. If, however, this could not scare her or him away, the huge modeling effort and limited tool performance will destroy the remaining illusions. Of course, this picture is painted excessively bleak and there certainly are promising developments, but it clearly points to the open problems which have to be solved to make formal methods for hybrid systems a well established tool in the process industries.

Acknowledgments

The views presented in this survey were developed while I was a member of the Process Control Laboratory of the Chemical Engineering Department at the University of Dortmund. They are the result of many discussions with colleagues and partners in several research projects. I am grateful to Nanette Bauer, Paul Chung, Sebastian Engell, Holger Graf, Hans-Michael Hanisch, Oded Maler, Bruce Krogh, Yassine Lakhnech, Angelika Mader, Peter Niebert, Jörg Preußig, Olaf Stursberg, and Heinz Treseler who helped to understand the sometimes very different worlds of process engineering, logic control design and computer science. The research projects on this topic in which I could participate have been funded by the European Commission in the ESPRIT LTR project *Verification of Hybrid Systems (VHS)*, by the German Research Council (DFG) in the focussed research program *Analysis and Synthesis of Technical Systems with Continuous-Discrete Dynamics (KONDISK)* and the temporary graduate school ("Graduiertenkolleg") *Modelling and Model-Based Design of Complex Technical Systems*, and by the German Academic Exchange Service (DAAD) in the exchange programs *British-German Academic Research Collaboration (ARC)* with the British Council and *Project-related Exchange of Personnel* with the NSF.

References

Alur, R. and D. Dill, "A theory of timed automata," *Theoretical Computer Science*, **126**, 183–235 (1990).

Alur, R., C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P.-H. Ho, X. Nicollin, A. Olivero, J. Sifakis, and S. Yovine, "The

algorithmic analysis of hybrid systems," *Theoretical Computer Science*, **138**, 3–34 (1995).

Alur, R., T. A. Henzinger, G. Lafferiere, and G. J. Pappas, "Discrete Abstractions of Hybrid Systems," *Proceedings of the IEEE*, **88**(7), 971–984 (2000).

Antsaklis, P. and A. Nerode, editors, *Special Issue on Hybrid Control Systems*, volume 43 of *IEEE Trans. Auto. Cont.* (1998).

Antsaklis, P., editor, *Special Issue on Hybrid Systems: Theory and Applications*, volume 88, no. 7 of *Proceedings of the IEEE* (2000).

Asarin, E. and O. Maler, As soon as possible: time optimal control for timed automata, In Vaandrager, F. W. and J. H. van Schuppen, editors, *Hybrid Systems: Computation and Control, Proc. 2nd Int. Workshop, HSCC'99, Berg en Dal, The Netherlands, March 1999*, volume 1569 of *Lecture Notes in Computer Science*, pages 19–30. Springer (1999).

Bemporad, A. and M. Morari, Verification of hybrid systems using mathematical programming, In Vaandrager, F. W. and J. H. van Schuppen, editors, *Hybrid Systems: Computation and Control, Proc. 2nd Int. Workshop, HSCC'99, Berg en Dal, The Netherlands, March 1999*, Lecture Notes in Computer Science 1569, pages 31–45. Springer (1999).

Brandin, B. A. and W. M. Wonham, "Supervisory control of timed discrete event systems," *IEEE Trans. Auto. Cont.*, **39**, 329–342 (1994).

Buechi, J. R. and L. H. Landweber, "Solving sequential conditions by finite state operators," *Trans. AMS*, **138**, 295–311 (1969).

Burch, J. R., E. M. Clarke, K. L. McMillan, D. L. Dill, and L. J. Hwang, "Symbolic model checking: 10^{20} states and beyond," *Inform. and Comput.*, **98**(2), 142–170 (1992).

Chouika, M., B. Ober, and E. Schnieder, Model-based control synthesis for discrete event systems, In *Proc. IAESTED Int. Conf. on Modeling and Simulation, Pittsburgh, USA, 1998*, pages 276–280 (1998).

Chung, S. L., S. Lafortune, and F. Lin, "Limited lookahead policies in supervisory control of discrete event systems," *IEEE Trans. Auto. Cont.*, **37**(12), 1921–1935 (1992).

Chutinan, A. and B. H. Krogh, Computing approximating automata for a class of linear hybrid systems, In *Hybrid Systems V: Proc. Int. Workshop, Notre Dame, USA*, Lecture Notes in Computer Science 1567, pages 16–37. Springer (1999).

Clarke, E. M. and R. P. Kurshan, "Computer-aided verification," *IEEE Spectrum*, pages 61–67 (1996).

Dang, T. and O. Maler, Reachability Analysis via Face Lifting, In Henzinger, T. A. and S. Sastry, editors, *Hybrid Systems: Computation and Control, Proc. 1st Int. Workshop, HSCC'98, Berkeley, USA, March 1998*, Lecture Notes in Computer Science 1386, pages 96–109. Springer (1998).

David, R. and H. Alla, *Petri nets and Grafcet*. Prentice Hall, New York (1992).

Dimitriadis, V. D., N. Shah, and C. C. Pantelides, "A case study in hybrid process safety verification," *Comput. Chem. Eng.*, **20, Suppl.**, S503–S508 (1996).

Dimitriadis, V. D., N. Shah, and C. C. Pantelides, "Modelling and safety verification of dicrete/continuous processing systems," *AIChE J.*, **43**(4), 1041–1059 (1997).

Engell, S., editor, *Special Issue on Discrete Event Models of Continuous Systems*, volume 6, no. 1 of *Mathematical and Computer Modelling of Dynamical Systems* (2000).

Fehnker, A., Scheduling a steel plant with timed automata (1999). Technical Report CSI-R9910, Computing Science Institute Nijmegen.

Frey, G. and L. Litz, Verification and validation of control algorithms by coupling of interpreted Petri nets, In *Proc. IEEE SMC, October 1998, San Diego, USA*, pages 7–12 (1999).

Fritz, M., K. Preuß, and S. Engell, A framework for flexible simulation of batch plants, In Zaytoon, J., editor, *Proc. 3rd Int. Conf. ADPM'98, Reims, France, March 1998*, pages 263–270 (1998).

Golaszewski, C. H. and P. J. Ramadge, Control of discrete event processes with forced events, In *Proc. 26th Conf. Decision and Control*, pages 247–251 (1987).

Graf, H. and H. Schmidt-Traub, A model-based approach to process hazard identification, In *Proc. 13th Int. Congress of Chemical and Process Engineering, CHISA'98, Prague, 1998* (1998).

Greenstreet, M. and I. Mitchell, Reachability Analysis Using Polygonal Projections, In Vaandrager, F. W. and J. H. van Schuppen, editors, *Hybrid Systems: Computation and Control, Proc. 2nd Int. Workshop, HSCC'99, Berg en Dal, The Netherlands, March 1999*, Lecture Notes in Computer Science 1569, pages 103–116. Springer (1999).

Hanisch, H.-M., A. Lüder, and M. Rausch, "Controller synthesis for net condition/event systems with a solution to incomplete state observation," *Euro. J. Cont.*, **3**, 280–291 (1997).

Hanisch, H.-M., "Coordination control modeling in batch production systems by means of Petri nets," *Comput. Chem. Eng.*, **16**(1), 1–10 (1992).

Harel, D. and A. Naamad, "The STATEMATE Semantics of Statecharts," *ACM Transactions on Software Engineering and Technology*, **4**(5), 293–333 (1996).

Heiner, M., P. Deussen, and J. Spranger, "A case study in design and verification of manufacturing system control software with hierarchical Petri nets," *Advanced Manufacturing Technology*, **15**, 139–152 (1999).

Henzinger, T. A., P. S. Ho, and H. Wong-Toi, "HyTech: A model checker for hybrid systems," *Software Tools for Technology Transfer*, **1**(1,2), 110–122 (1997).

Herrmann, P., G. Graw, and H. Krumm, Compositional specification and structured verification of hybrid systems in cTLA, In *Proc. 1st IEEE int. Symposium on Object-Oriented Real-Time Distributed Computing, Kyoto, Japan* (1998).

Holloway, L. E., B. H. Krogh, and A. Giua, "A survey of petri net methods for controlled discrete event systems," *J. Disc. Event Dyn. Sys.*, **7**, 151–190 (1997).

IEC, International Standard 1131: Programmable logic controllers. Part 3: Languages, International Electrotechnical Commission, Geneva (1992).

ISA, International Standard S88.01: Batch Control, Part 1: Models and Terminology, Instrumentation Society of America (1995).

Kowalewski, S., H.-M. Hanisch, and U. Anderssohn, Logic controller synthesis for non c/u-partitionable automata with forbidden states, In *Preprints IFAC 13th World Congress, San Francisco, 1996, vol. J*, pages 359–364 (1996).

Kowalewski, S., S. Engell, J. Preußig, and O. Stursberg, "Verification of logic controllers for continuous plants using timed condition/event system models," *Automatica*, **35**(3), 505–518 (1999).

Kowalewski, S., O. Stursberg, and N. Bauer, "An Experimental Batch Plant as a Case Example for the Verification of Hybrid Systems," *Euro. J. Cont.* (2001). To appear.

Krogh, B. H. and S. Kowalewski, "State feedback control of condition/event systems," *Mathematical and Computer Modeling*, **23**(11/12), 161–174 (1996).

Larsen, K. G., P. Pettersson, and W. Yi, "UPPAAL in a nutshell," *Software Tools for Technology Transfer*, **1**(1,2), 134–152 (1997).

Lawley, H. G., "Operability studies and hazard analysis," *Chem. Eng. Prog.*, **70**, 105–116 (1974).

Lemmon, M., K. He, and I. Markovsky, "Supervisory hybrid systems," *IEEE Cont. Sys. Mag.*, **19**, 42–55 (1999).

Li, Y. and W. M. Wonham, "Control of vector discrete event systems: I. the base model," *IEEE Trans. Auto. Cont.*, **38**, 1214–1227 (1993).

Li, Y. and W. M. Wonham, "Control of vector discrete event systems: II. controller synthesis," *IEEE Trans. Auto. Cont.*, **39**, 512–531 (1994).

Lind-Nielsen, J., H. R. Andersen, G. Behrmann, H. Hulgaard, K. Kristoffersen, and K. G. Larsen, Verification of large state/event systems using compositionality and dependency analysis, In *Proc. TACAS'98*, Lecture Notes in Computer Science 1384, pages 201–216. Springer (1998).

Mader, A. and H. Wupper, Timed automaton models for simple programmable logic controllers, In *Proc. Euromicro Conf. on Real-Time Systems, York, UK, June 1999* (1999).

Maler, O., On the programming of industrial computers (1999). Report of the ESPRIT project VHS, see (VHS, 2000).

Marikar, M. T., G. E. Rotstein, A. Sanchez, and S. Macchietto, Computer aided analysis and synthesis of procedural controllers, In *Proc. Workshop on Discrete Event Systems (WODES'98), Cagliari, Italy, 1998*, pages 420–425. IEE (1998).

Moon, I., G. J. Powers, J. R. Burch, and E. M. Clarke, "Automatic verification of sequential control systems using temporal logic," *AIChE J.*, **38**(1), 67–75 (1992).

Morse, A. S., C. C. Pantelides, S. S. Sastry, and J. M. Schumacher, editors, *Special Issue on Hybrid Systems*, volume 35, issue 3 of *Automatica* (1999).

Niebert, P. and S. Yovine, Computing optimal operation schedules for multi batch operation of chemical plants (1999). Report of the ESPRIT project VHS, see (VHS, 2000).

Park, T. and P. I. Barton, "Implicit model checking of logic based control systems," *AIChE J.*, **43**(9), 2246–2260 (1997).

Preußig, J., O. Stursberg, and S. Kowalewski, Reachability Analysis of a Class of Switched Continuous Systems by Integrating Rectangular Approximation and Rectangular Analysis, In Vaandrager, F. W. and J. H. van Schuppen, editors, *Hybrid Systems: Computation and Control, Proc. 2nd Int. Workshop, HSCC'99, Berg en Dal, The Netherlands, March 1999*, Lecture Notes in Computer Science 1569, pages 209–222. Springer (1999).

Preußig, J., *Formale Überprüfung der Korrektheit von Steuerungen mittels rektangulärer Automaten*, PhD thesis, Department of Chemical Engineering, University of Dortmund, Germany (2000). (in German).

Probst, S. T., G. J. Powers, D. E. Long, and I. Moon, "Verification of a logically controlled solids transport system using symbolic model checking," *Comput. Chem. Eng.*, **21**(4), 417–429 (1997).

Raisch, J. and S. O'Young, "Discrete approximation and supervisory control of continuous systems," *IEEE Trans. Auto. Cont.*, **43**(4), 569–573 (1998).

Ramadge, P. J. and W. M. Wonham, "Modular feedback logic for discrete event systems," *SIAM J. Cont. Optim.*, **25**, 1202–1218 (1987a).

Ramadge, P. J. and W. M. Wonham, "Supervisory control of a class of discrete event processes," *SIAM J. Cont. Optim.*, **25**, 206–230 (1987b).

Ramadge, P. J. and W. M. Wonham, "The control of discrete event systems," *Proc. IEEE*, **77**, 81–98 (1989).

Sampaath, M., R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Failure diagnosis using discrete event models," *IEEE Trans. Cont. Sys. Tech.*, **4**(2), 105–124 (1996).

Sanchez, A., *Formal specification and synthesis of procedural controllers for process systems*, Lecture Notes in Control and Information Sciences 212. Springer (1996).

Tittus, M., *Control synthesis for batch processes*, PhD thesis, Chalmers University of Technology, Göteborg, Sweden (1995).

Vaidhyanathan, R. and V. Venkatasubramanian, "Digraph-based models for automated HAZOP analysis," *Reliability Engineering and Systems Safety*, **50**, 33–49 (1995).

VHS, ESPRIT project Verification of Hybrid Systems (2000). `http://www-verimag.imag.fr/VHS/main.html`.

Wong-Toi, H., The synthesis of controllers for linear hybrid automata, In *Proc. Conf. Decision and Control*. IEEE (1997).

Yang, S. and P. W. H. Chung, "Hazard analysis and support tool for computer-controlled processes," *J. Loss Prevention in the Process Industries*, **11**, 333–345 (1998).

Yovine, S., "KRONOS: a verification tool for real-time systems," *Software Tools for Technology Transfer*, **1**(1,2), 123–133 (1997).

Hybrid System Analysis and Control via Mixed Integer Optimization

Manfred Morari*
Automatic Control Laboratory,
ETH—Swiss Federal Institute of Technology,
CH-8092 Zurich, Switzerland

Abstract

The paper discusses a framework for modeling, analyzing and controlling systems whose behavior is governed by interdependent physical laws, logic rules, and operating constraints, denoted as *Mixed Logical Dynamical* (MLD) systems. They are described by linear dynamic equations subject to linear inequalities involving real and integer variables. MLD models are equivalent to various other system descriptions like Piece Wise Affine (PWA) systems and Linear Complementarity (LC) systems. They have the advantage, however, that all problems of system analysis (like controllability, observability, stability and verification) and all problems of synthesis (like controller design and filter design) can be readily expressed as mixed integer linear or quadratic programs, for which many commercial software packages exist.

In this paper we first recall how to derive MLD models and then illustrate their use in predictive control. Subsequently we define "verification" and show how verification algorithms can be used to solve a variety of practical problems like checking the correctness of an emergency shutdown procedure implemented on a PLC, or assessing the performance of a constrained MPC controller. The eventual practical success of these methods will depend on progress in the development of the various optimization packages so that problems of realistic size can be tackled.

Keywords

Hybrid systems, Predictive control, Dynamic models, Mixed-integer programming, Optimization problems

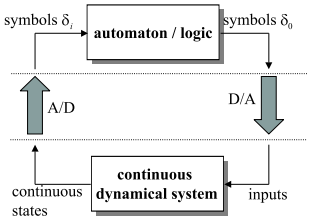
Introduction

The concept of a *model* of a system is traditionally associated with differential or difference equations, typically derived from physical laws governing the *dynamics* of the system under consideration. Consequently, most of the control theory and tools have been developed for such systems, in particular for systems whose evolution is described by smooth linear or nonlinear state transition functions. On the other hand, in many applications the system to be controlled is also constituted by parts described by *logic*, such as for instance on/off switches or valves, gears or speed selectors, evolutions dependent on if-then-else rules. Often, the control of these systems is left to schemes based on heuristic rules inferred by practical plant operation.

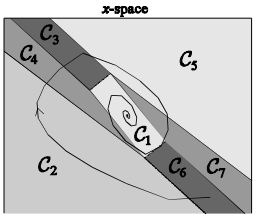
Recently, in the literature researchers started dealing with *hybrid systems*, namely hierarchical systems composed of dynamical components at the lower level, governed by upper level logical/discrete components (Grossmann et al., 1993; Branicky et al., 1998; Labinaz et al., 1997; Branicky, 1995). Hybrid systems arise in a large number of application areas, and are attracting increasing attention in both academic theory-oriented circles as well as in industry. Our interest is motivated by several clearly discernible trends in the process industries which point toward an extended need for new tools to design control and supervisory schemes for hybrid systems and to analyze their performance.

This paper discusses a framework for modeling, analyzing and controlling models of systems described by interacting physical laws, logical rules, and operating constraints. We will focus exclusively on discrete time models. We note, however, that interesting mathematical phenomena occurring in hybrid systems, such as Zeno behaviors (Johansson et al., 1999) do not exist in discrete-time. On the other hand, most of these phenomena are usually a consequence of the continuous-time switching model, rather than the real natural behavior. Our main motivation for concentrating on discrete-time models stems from the need to analyze these systems and to solve optimization problems, such as optimal control or scheduling problems, for which the continuous-time counterpart would not be easily computable.

Two main categories of hybrid systems were successfully adopted for analysis and synthesis purposes (Branicky, 1995): *hybrid control systems* (Alur et al., 1993;



(a) Hybrid control systems. Finite state machines and continuous dynamics interact through analog-to-digital (A/D) and D/A interfaces



(b) Piecewise affine systems

Figure 1: Hybrid models.

*phone +41-1-632 7626, fax +41-1-632 1211, email: morari@aut.ee.ethz.ch. This paper is based on the work of the Hybrid Systems Group at ETH currently including A. Bemporad, F. Borrelli, F. Cuzzola, G. Ferrari-Trecate, T. Geyer, D. Mignone and F. D. Torrisi

Asarin et al., 1995; Bemporad and Morari, 1999; Lygeros et al., 1996, 1999), which consist of the interaction between continuous dynamical systems and discrete/logic automata (Figure 1(a)), and *switched systems* (Branicky, 1998; Johansson and Rantzer, 1998; Sontag, 1981), where the state-space is partitioned into regions, each one being associated with a different continuous dynamics (Figure 1(b)).

Several modeling frameworks have been introduced for discrete-time hybrid systems. We will refer frequently to the *piecewise affine* (PWA) systems (Sontag, 1981, 1996).

$$\begin{aligned} x(t+1) \quad &= A_i x(t) + B_i u(t) + f_i \ , \\ &\text{for } x(t) \in \mathcal{C}_i \triangleq \{x : \ H_i x \leq K_i\} \end{aligned} \quad (1)$$

where $x \in X \subseteq \mathbb{R}^n$, $u \in \mathbb{R}^m$, $\{\mathcal{C}_i\}_{i=0}^{s-1}$ is a polyhedral partition of the sets of states X and f_i is a constant vector. Furthermore, we mention here linear complementarity (LC) systems (Heemels, 1999; van der Schaft and Schumacher, 1998; Heemels et al., 2000) and extended linear complementarity (ELC) systems (De Schutter and De Moor, 1999), max-min-plus-scaling (MMPS) systems (De Schutter and van den Boom, 2000), and mixed logical dynamical (MLD) systems (Bemporad and Morari, 1999). Each modeling framework has its advantages. For instance, stability criteria were formulated for PWA systems (Johansson and Rantzer, 1998), and control and verification techniques were proposed for MLD discrete-time hybrid models (Bemporad and Morari, 1999; Bemporad et al., 2000). In particular, MLD models were proven successful for recasting hybrid dynamical optimization problems into mixed-integer linear and quadratic programs, solvable via branch and bound techniques (Nemhauser and Wolsey, 1988). Recently, the equivalence of PWA, LC, ELC, MMPS, and MLD hybrid dynamical systems was proven constructively in (Heemels et al., 2001; Bemporad et al., 2000a). Thus the theoretical properties and tools can be easily transferred from one class to another.

Mixed Logical Dynamical (MLD) Systems

We briefly recall the *mixed logical dynamical* (MLD) system framework introduced in (Bemporad and Morari, 1999). MLD systems are hybrid (control) systems defined by the interaction of logic, finite state machines, and linear discrete-time systems, as shown in Figure 1(a). The ability to include constraints, constraint prioritization, and heuristics augment the expressiveness and generality of the MLD framework.

The MLD modeling framework relies on the idea of translating logic relations into mixed-integer linear inequalities (Bemporad and Morari, 1999; Cavalier et al., 1990; Raman and Grossmann, 1991; Tyler and Morari, 1999; Williams, 1977, 1993). By following standard notation, we adopt capital letters X_i to represent statements, e.g. "$x \geq 0$" or "Temperature is hot". X_i is commonly referred to as a *literal*, and has a *truth value* of either TRUE or FALSE. Boolean algebra enables statements to be combined in compound statements by means of *connectives*: "$\wedge$" (and), "$\vee$" (or), "$\sim$" (not), "$\rightarrow$" (implies), "$\leftrightarrow$" (if and only if), "$\oplus$" (exclusive or). Connectives satisfy several properties (see e.g. (Christiansen, 1997)), which can be used to transform compound statements into equivalent statements involving different connectives, and simplify complex statements. Correspondingly one can associate with a literal X_i a *logical variable* $\delta_i \in \{0,1\}$, which has a value of either 1 if $X_i =$TRUE, or 0 otherwise. A propositional logic problem, where a statement X_1 must be proved to be true given a set of (compound) statements involving literals $X_1, \ldots, X_n$, can be solved by means of a linear integer program by suitably translating the original compound statements into linear inequalities involving logical variables δ_i. In fact, the propositions and linear constraints reported in Table 1 can easily be seen to be equivalent.

These translation techniques can be adopted to model logical parts of processes and heuristic knowledge about plant operation as integer linear inequalities. The link between logic statements and continuous dynamical variables, in the form of logic statements derived from conditions on physical dynamics, is provided by properties (AD1), (DA1) in Table 1, and leads to *mixed-integer linear inequalities*, i.e., linear inequalities involving both *continuous variables* of $\mathbb{R}^n$ and logical (*indicator*) variables in $\{0,1\}$. Consider for instance the statement $X \triangleq [f_1(x) \leq 0]$ where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a linear (affine) function, assume that $x \in \mathcal{X}$, where $\mathcal{X} \subset \mathbb{R}^n$ is a given bounded set, and define

$$M_i \triangleq \max_{x \in \mathcal{X}} f_i(x), \ m_i \triangleq \min_{x \in \mathcal{X}} f_i(x), \ i = 1,2$$

Theoretically, an over[under]-estimate of M_i $[m_i]$ suffices for our purpose. By associating a binary variable δ with the literal X, one can transform $X \triangleq [f_1(x) \leq 0]$ into mixed integer inequalities as described in (AD1), Table 1, where ϵ is a small tolerance (typically the machine precision), beyond which the constraint is regarded as violated. Note that sometimes translations require the introduction of *auxiliary variables* (Williams, 1993, p. 178), for instance according to (DA1) a quantity which is either $f_1(x)$ is X is true, or $f_2(x)$, requires the introduction of a real variable z.

The rules of Table 1 can be generalized for relations involving an arbitrary number of discrete variables combined by arbitrary connectives. *Any* combinational relation of logical variables can in fact be translated to the conjunctive normal form (CNF)

$$P_1 \wedge P_2 \wedge \ldots \wedge P_m, \ P_i = Y_1^i \vee Y_2^i \vee \ldots \vee Y_{n_i}^i$$

where Y_j^i is a literal X or its inverse $\sim X$, by using standard methods (Bemporad et al., 2000b). As an example,

	Relation	Logic	(In)equalities
L1	**AND** ($\wedge$)	$X_1 \wedge X_2$	$\delta_1 = 1$, $\delta_2 = 1$
L2	**OR** ($\vee$)	$X_1 \vee X_2$	$\delta_1 + \delta_2 \geq 1$
L3	**NOT** ($\sim$)	$\sim X_1$	$\delta_1 = 0$
L4	**XOR** ($\oplus$)	$X_1 \oplus X_2$	$\delta_1 + \delta_2 = 1$
L5	**IMPLY** ($\rightarrow$)	$X_1 \rightarrow X_2$	$\delta_1 - \delta_2 \leq 0$
L6	**IFF** ($\leftrightarrow$)	$X_1 \leftrightarrow X_2$	$\delta_1 - \delta_2 = 0$
L7	**ASSIGNMENT** ($=, \leftrightarrow$)	$X_3 = X_1 \wedge X_2$ $X_3 \leftrightarrow X_1 \wedge X_2$	$\delta_1 + (1 - \delta_3) \geq 1$ $\delta_2 + (1 - \delta_3) \geq 1$ $(1 - \delta_1) + (1 - \delta_2) + \delta_3 \geq 1$
AD1	**EVENT**	$[f(x) \leq 0] \leftrightarrow [\delta = 1]$	$f(x) \leq M - M\delta$ $f(x) \geq \epsilon + (m - \epsilon)\delta$
DA1	**IF-THEN-ELSE** (**=Product**)	**IF** X **THEN** $z = f_1(x)$ **ELSE** $z = f_2(x)$ ($z = \delta f_1(x) + (1 - \delta) f_2(x)$)	$(m_2 - M_1)\delta + z \leq f_2(x)$ $(m_1 - M_2)\delta - z \leq -f_2(x)$ $(m_1 - M_2)(1 - \delta) + z \leq f_1(x)$ $(m_2 - M_1)(1 - \delta) - z \leq -f_1(x)$

Table 1: Basic conversion of logic relations into mixed-integer inequalities. Relations involving the inverted literals $\sim X$ can be obtained by substituting $(1 - \delta)$ for δ in the corresponding inequalities.

the statement (L7) $X_3 \leftrightarrow X_1 \wedge X_2$ is equivalent to

$$(\sim X_1 \vee \sim X_2 \vee X_3) \wedge (X_1 \vee \sim X_3) \wedge (X_2 \vee \sim X_3) \quad (2)$$

Subsequently, the CNF form can be translated (again, automatically and without introducing additional integer variables) into the set of integer linear inequalities

$$y_1^i + y_2^i + \ldots + y_{n_i}^i \geq 1, \ i = 1, \ldots, m,$$

where $y_j^i = \delta_j^i$ if $Y_j^i = X$ or $(1 - \delta_j^i)$ if $Y_j^i = \sim X$. For instance, it is immediate to check that the CNF (2) maps to the inequalities reported in Table 1(L7).

The state update law of finite state machines can be described by logic statements involving binary states, their time updates, and indicator variables. It is clear that the automatized translation mentioned above can be directly applied to translate automata into a set of linear integer inequalities. An example will be provided when modeling the PLC control code of the batch evaporator process benchmark.

By collecting the equalities and inequalities generated by the translation of the automata, analog-to-digital interface (AD1), digital-to-analog interface (DA2), and by including the linear dynamic difference equations, we can model the hybrid system depicted in Figure 1(a) as the Mixed Logical Dynamical (MLD) system

$$x(t+1) = \Phi x(t) + \mathcal{G}_1 u(t) + \mathcal{G}_2 \delta(t) + \mathcal{G}_3 z(t) \quad (3a)$$
$$y(t) = \mathcal{H} x(t) + \mathcal{D}_1 u(t) + \mathcal{D}_2 \delta(t) + \mathcal{D}_3 z(t) \quad (3b)$$
$$\mathcal{E}_2 \delta(t) + \mathcal{E}_3 z(t) \leq \mathcal{E}_1 u(t) + \mathcal{E}_4 x(t) + \mathcal{E}_5 \quad (3c)$$

where $x \in \mathbb{R}^{n_c} \times \{0,1\}^{n_\ell}$ is a vector of continuous and binary states, $u \in \mathbb{R}^{m_c} \times \{0,1\}^{m_\ell}$ are the inputs, $y \in \mathbb{R}^{p_c} \times \{0,1\}^{p_\ell}$ the outputs, $\delta \in \{0,1\}^{r_\ell}$, $z \in \mathbb{R}^{r_c}$ represent auxiliary binary and continuous variables respectively, which are introduced when transforming logic relations into mixed-integer linear inequalities, and Φ, $\mathcal{G}_1$, $\mathcal{G}_2$, $\mathcal{G}_3$, $\mathcal{H}$, $\mathcal{E}_1, \ldots, \mathcal{E}_5$ are matrices of suitable dimensions. The inequalities (3c) include the constraints obtained by the D/A, A/D, and automaton part of the system, as well as possible physical constraints on states and inputs. Despite the fact that the description (3) seems to be linear, clearly the nonlinearity is concentrated in the integrality constraints over binary variables. The following simple example illustrates the technique.

Example 1

Consider the following switched system

$$x(t+1) = \begin{cases} 0.8x(t) + u(t) & \text{if} \quad x(t) \geq 0 \\ -0.8x(t) + u(t) & \text{if} \quad x(t) < 0 \end{cases} \quad (4)$$

where $x(t) \in [-10, 10]$, and $u(t) \in [-1, 1]$. The condition $x(t) \geq 0$ can be associated with a binary variable $\delta(t)$ such that

$$[\delta(t) = 1] \ \leftrightarrow \ [x(t) \geq 0] \quad (5)$$

By using the transformation (AD1) in Table 1, Equation 5 can be expressed by the inequalities

$$-m\delta(t) \leq x(t) - m$$
$$-(M + \epsilon)\delta \leq -x - \epsilon$$

where $M = -m = 10$, and ϵ is a small positive scalar. Then (4) can be rewritten as

$$x(t+1) = 1.6\delta(t)x(t) - 0.8x(t) + u(t) \quad (6)$$

By defining a new variable $z(t) = \delta(t)x(t)$ which, by (DA1) in Table 1, can be expressed as

$$\begin{aligned} z(t) &\le M\delta(t) \\ z(t) &\ge m\delta(t) \\ z(t) &\le x(t) - m(1-\delta(t)) \\ z(t) &\ge x(t) - M(1-\delta(t)), \end{aligned}$$

the evolution of system (4) is ruled by the linear equation

$$x(t+1) = 1.6z(t) - 0.8x(t) + u(t)$$

subject to the linear constraints above.

We assume that system (3) is *completely well-posed* (Bemporad and Morari, 1999), which means that for all x, u within a bounded set the variables δ, z are uniquely determined, i.e., there exist functions F, G such that, at each time t, $\delta(t) = F(x(t), u(t))$, $z(t) = G(x(t), u(t))$[1]. Therefore the x- and y-trajectories exist and are uniquely determined by the initial state $x(0)$ and input signal u. In light of the transformations of Table 1, it is clear that the well-posedness assumption stated above is usually guaranteed by the procedure used to generate the linear inequalities (3c), and therefore this hypothesis is typically verified by MLD relations derived from modeling real-world plants. Nevertheless, a numerical test for well-posedness is reported in (Bemporad and Morari, 1999, Appendix 1).

Recently, in (Bemporad et al., 2000b) a declarative language for specifying hybrid control systems has been developed which fully automatizes the construction of the MLD matrices. Such a language (called HYSDEL, HYbrid System DEscription Language) will be used later to model the batch evaporator process benchmark. Throughout the paper, we will assume that both the PWA and the MLD forms are available. Their role is complementary and in our algorithms we use whichever one is more appropriate for a specific task.

Predictive Control of MLD Systems

The predictive control problem for MLD sytems can be defined formally as follows. Consider an equilibrium pair (x_e, u_e) Let the components $\delta_{e,i}$, $z_{e,j}$ correspond to desired steady-state values for the indefinite auxiliary variables. Let t be the current time, and $x(t)$ the current state. Consider the following optimal control problem

$$\begin{aligned} \min_{\{v_0^{T-1}\}} J(v_0^{T-1}, x(t)) \triangleq \sum_{k=0}^{T-1} &\|v(k) - u_e\|_{Q_1}^2 \\ + \|\delta(k|t) - \delta_e\|_{Q_2}^2 &+ \|z(k|t) - z_e\|_{Q_3}^2 \\ + \|x(k|t) - x_e\|_{Q_4}^2 &+ \|y(k|t) - y_e\|_{Q_5}^2 \end{aligned} \quad (7)$$

subject to

$$\begin{cases} x(T|t) &= x_e \\ x(k+1|t) &= Ax(k|t) + B_1 v(k) + B_2\delta(k|t) \\ &\quad + B_3 z(k|t) \\ y(k|t) &= Cx(k|t) + D_1 v(k) + D_2\delta(k|t) \\ &\quad + D_3 z(k|t) \\ E_2\delta(k|t) &+ E_3 z(k|t) \le E_1 v(k) + E_4 x(k|t) + E_5 \end{cases} \quad (8)$$

where $Q_1 = Q_1' > 0$, $Q_2 = Q_2' \ge 0$, $Q_3 = Q_3' \ge 0$, $Q_4 = Q_4' > 0$, $Q_5 = Q_5' \ge 0$, $x(k|t) \triangleq x(t+k, x(t), v_0^{k-1})$, and $\delta(k|t)$, $z(k|t)$, $y(k|t)$ are similarly defined. Assume for the moment that the optimal solution $\{v_t^*(k)\}_{k=0,\ldots,T-1}$ exists. According to the *receding horizon* philosophy, set

$$u(t) = v_t^*(0), \quad (9)$$

disregard the subsequent optimal inputs $v_t^*(1), \ldots, v_t^*(T-1)$, and repeat the whole optimization procedure at time $t+1$. The control law (7)–(9) will be referred to as the *Mixed Integer Predictive Control* (MIPC) law.

In principle all the design rules about parameter choices and theoretical results regarding stability developed for MPC over the last two decades can be applied here to MIPC after some adjustments. For instance, the number of control degrees of freedom can be reduced to $N_u < T$, by setting $u(k) \equiv u(N_u - 1)$, $\forall k = N_u, \ldots, T$. While this choice usually reduces the size of the optimization problem dramatically at the price of inferior performance, here the computational gain is only partial, since all the T $\delta(k|t)$ and $z(k|t)$ variables remain in the optimization.

Infinite horizon formulations are inappropriate for both practical and theoretical reasons. In fact, approximating the infinite horizon with a large T is computationally prohibitive, as the number of 0-1 variables involved in the MIQP depends linearly on T. Moreover, the quadratic term in δ might oscillate (for example, for a system which approaches the origin in a "switching" manner), and hence "good" (i.e., asymptotically stabilizing) input sequences might be ruled out by a corresponding infinite value of the performance index; it could even happen that no input sequence has finite cost.

Using an appropriate stability definition (Bemporad and Morari, 1999) have proven asymptotic stability of the MIPC scheme. In the typical fashion, the end point constraint was invoked in the Lyapunov arguments.

Despite the fact that very effective methods exist to compute the (global) optimal solution of the MIQP problem (7)–(9) (see Section below), in the worst-case the solution time depends exponentially on the number of integer variables. In principle, this might limit the scope of application of the proposed method to very slow systems, since for real-time implementation the sampling time should be large enough to allow the worst-case computation. However, the stability proof does not require that the evaluated control sequences $\{\mathcal{U}_t^*\}_{t=0}^{\infty}$ are global

[1] A more general definition of well-posedness, where only the components of δ and z entering (3a)–(3b) are required to be unique, is given in (Bemporad and Morari, 1999).

optima, but only that they lead to a decrease in the objective function. Thus the solver can be interrupted at any intermediate step to obtain a suboptimal solution $\mathcal{U}_{t+1}^*$ which satisfies the decrease condition. For instance, when Branch & Bound methods are used to solve the MIQP problem, the new control sequence $\mathcal{U}_t^*$ can be selected as the solution to a QP subproblem which is integer-feasible and has the lowest value. Obviously in this case the performance deteriorates.

Example 2

Consider the following system

$$\begin{cases} x(t+1) = 0.8 \begin{bmatrix} \cos\alpha(t) & -\sin\alpha(t) \\ \sin\alpha(t) & \cos\alpha(t) \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \\ y(t) = [1\ 0]x(t) \\ \alpha(t) = \begin{cases} \frac{\pi}{3} & \text{if} \quad [1\ 0]x(t) \geq 0 \\ -\frac{\pi}{3} & \text{if} \quad [1\ 0]x(t) < 0 \end{cases} \\ x(t) \in [-10, 10] \times [-10, 10] \\ u(t) \in [-1, 1] \end{cases} \tag{10}$$

By using auxiliary variables $z(t) \in \mathbb{R}^4$ and $\delta(t) \in \{0, 1\}$ such that $[\delta(t) = 1] \leftrightarrow [[1\ 0]x(t) \geq 0]$, Equation 10 can be rewritten in the form (3) as

$$\begin{aligned} x(t+1) = &\begin{bmatrix} I & I \end{bmatrix} z(t) \\ &\begin{bmatrix} 10 \\ -10-\epsilon \\ -M \\ -M \\ M \\ M \\ M \\ M \\ -M \\ -M \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \delta(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ I & 0 \\ -I & 0 \\ 0 & I \\ 0 & -I \\ I & 0 \\ -I & 0 \\ 0 & I \\ 0 & -I \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} z(t) \leq \\ &\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ B \\ -B \\ B \\ -B \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} u(t) + \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & \\ 0 & \\ 0 & \\ 0 & \\ A_1 & \\ -A_1 & \\ A_2 & \\ -A_2 & \\ I & \\ -I & \\ 0 & \\ 0 & \end{bmatrix} x(t) + \begin{bmatrix} 10 \\ -\epsilon \\ 0 \\ 0 \\ M \\ M \\ M \\ M \\ 0 \\ 0 \\ N \\ N \\ 1 \\ 1 \end{bmatrix} \end{aligned}$$

where $B = [0\ 1]'$, A_1, A_2 are obtained by (10) by setting respectively $\alpha = \frac{\pi}{3}, -\frac{\pi}{3}$, $M = 4(1+\sqrt{3})[1\ 1]' + B$, $N \triangleq 10[1\ 1]'$, and ϵ is a properly small positive scalar.

In order to stabilize the system to the origin, the feedback control law (7)–(9) is adopted, along with the parameters $T = 3$, $u_e = 0$, $\delta_e = 0$, $z_e = [0\ 0\ 0\ 0]'$,

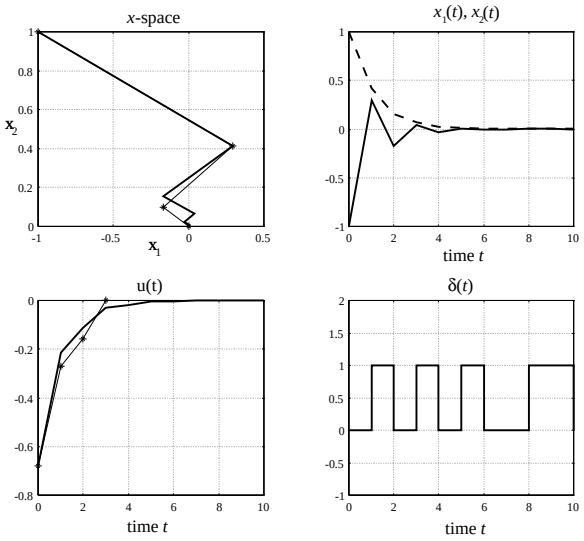


Figure 2: Closed-loop regulation problem for system (10). Closed-loop trajectories (thick lines) and optimal solution at $t = 0$ (thin lines, right plots).

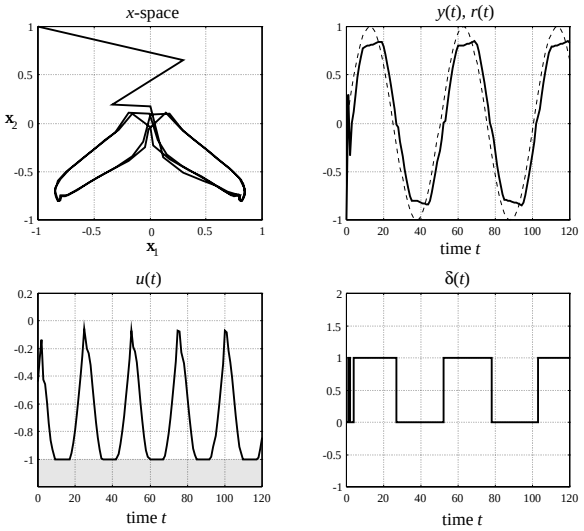


Figure 3: Closed-loop tracking problem for system (10), with $y(t) = x_1(t)$.

$x_e = [0\ 0]'$, $y_e = 0$, and the weights $Q_1 = 1$, $Q_2 = 0.01$, $Q_3 = 0.01 I_4$, $Q_4 = I_2$, $Q_5 = 0$. Figure 2 shows the resulting trajectories. The trajectories obtained at time $t = 0$ by solving the optimal control problem (7)–(8) are also reported in the right plots (thin lines). Consider now a desired reference $r(t) = \sin(t/8)$ for the output $y(t)$. We apply the same MIPC controller, with the exception of $Q_4 = 10^{-8} I_2$, $Q_5 = 1$. The steady-state parameters are selected as $y_e = r(t)$, and u_e, x_e, δ_e, z_e consistently. Figure 3 shows the resulting closed-loop trajectories. Notice that the constraint $-1 \leq u(t) \leq 1$ prevents the system from tracking the peaks of the sinusoid, and therefore

the output trajectory is chopped. □

MIQP Solvers

With the exception of particular structures, mixed-integer programming problems involving 0-1 variables are classified as $\mathcal{NP}$-complete, which means that in the worst case, the solution time grows exponentially with the problem size (Raman and Grossmann, 1991). Despite this combinatorial nature, several algorithmic approaches have been proposed and applied successfully to medium and large size application problems (Floudas, 1995), the four major ones being

- *Cutting plane methods*, where new constraints (or "cuts") are generated and added to reduce the feasible domain until a 0-1 optimal solution is found.
- *Decomposition methods*, where the mathematical structure of the models is exploited via variable partitioning, duality, and relaxation methods.
- *Logic-based methods*, where disjunctive constraints or symbolic inference techniques are utilized which can be expressed in terms of binary variables.
- *Branch and bound methods*, where the 0-1 combinations are explored through a binary tree, the feasible region is partitioned into sub-domains systematically, and valid upper and lower bounds are generated at different levels of the binary tree.

For MIQP problems, (Fletcher and Leyffer, 1998) indicate Generalized Benders' Decomposition (GBD) (Lazimy, 1985), Outer Approximation (OA), LP/QP based branch and bound, and Branch and Bound as the major solvers. See (Roschchin et al., 1987) for a review of these methods.

Several authors agree on the fact that branch and bound methods are the most successful for mixed integer quadratic programs. (Fletcher and Leyffer, 1998) report a numerical study which compares different approaches, and Branch and Bound is shown to be superior by an order of magnitude. While OA and GBD techniques can be attractive for general Mixed-Integer Nonlinear Problems (MINLP), for MIQP at each node the relaxed QP problem can be solved without approximations and reasonably quickly (for instance, the Hessian matrix of each relaxed QP is constant).

As described by (Fletcher and Leyffer, 1998), the Branch and Bound algorithm for MIQP consists of solving and generating new QP problems in accordance with a tree search, where the nodes of the tree correspond to QP subproblems. Branching is obtained by generating child-nodes from parent-nodes according to branching rules, which can be based for instance on a-priori specified priorities on integer variables, or on the amount by which the integer constraints are violated. Nodes are labeled as either pending, if the corresponding QP problem has not yet been solved, or fathomed, if the node has already been fully explored. The algorithm stops when all nodes have been fathomed. The success of the branch and bound algorithm relies on the fact that whole subtrees can be excluded from further exploration by fathoming the corresponding root nodes. This happens if the corresponding QP subproblem is either infeasible or an integer solution is obtained. In the second case, the corresponding value of the cost function serves as an upper bound on the optimal solution of the MIQP problem, and is used to further fathom other nodes having greater optimal value or lower bound.

Some of the simulation results reported in this paper have been obtained in Matlab by using the commercial Fortran package (Fletcher and Leyffer, 1994) as a MIQP solver. This package can handle both dense and sparse MIQP problems. The latter has proven to be particularly effective to solve most of the optimal control problems for MLD systems. In fact, the constraints have a triangular structure, and in addition most of the constraints generated by representation of logic facts involve only a few variables, which often leads to sparse matrices.

Explicit Computation of MPC Control Law

In (Bemporad et al., 2002) the authors presented a new approach to implement MPC, where all the computation effort is moved off-line. The idea is based on the observation that in (8) the state $x(0|t)$ can be considered a vector of parameters, and (7)–(9) as a multi-parametric Mixed Integer Quadratic Program (mp-MIQP). If the 1-norm is used in (7) instead of the 2-norm a multi-parametric Mixed Integer Linear Program (mp-MILP) results. An algorithm to solve mp-MILP problems was presented in (Acevedo and Pistikopoulos, 1997). Once the multi-parametric problem (7,8) has been solved off line, i.e., the solution $U_t^* = f(x(t))$ of (7,8) has been found, the model predictive controller is available explicitly. In (Acevedo and Pistikopoulos, 1997) the authors also show that the solution $U^* = f(x)$ of the mp-MILP problem is piecewise affine. Clearly, the same properties are inherited by the controller, i.e.,

$$u(t) = \begin{array}{l} F_i x(t) + g_i, \text{ for} \\ x(t) \in \mathcal{C}_i \triangleq \{x : \ H_i x \le S_i\}, \ i = 1, \dots, s \end{array} \tag{11}$$

where $\cup_{i=1}^{s} \mathcal{C}_i$ is the set of states for which a feasible solution to (7,8) exists. Therefore, the closed MPC loop is of the form (1), where $A_i = A + BF_i$, $f_i = Bg_i$, $B_i = 0$. Note that the form of the closed-loop MPC system remains PWA also when (i) the matrices A, B of the plant model are different from those used in the prediction model, and (ii) the plant model has a PWA form. Typically, the MPC law is designed based on a linear model obtained by linearizing the nonlinear model of the plant around some operating condition. When the nonlinear model can be approximated by a PWA system (e.g., through multiple linearizations at different

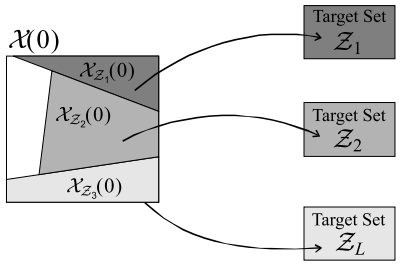


Figure 4: Reachability analysis problem.

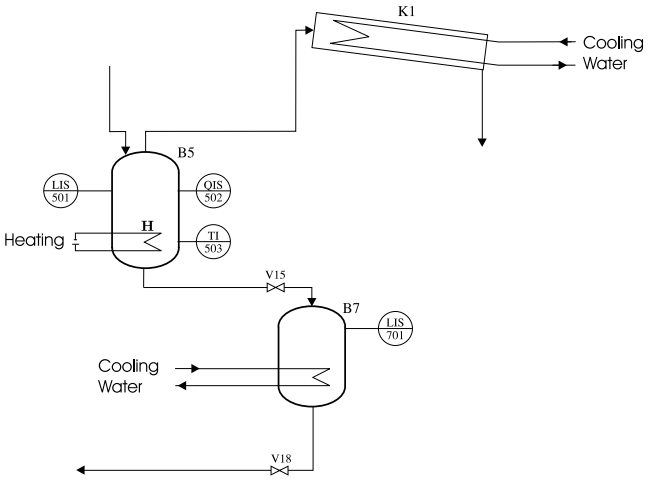


Figure 5: Flowchart of the benchmark evaporator system.

operating points or by approximating nonlinear static mappings by piecewise linear functions), the closed-loop consisting of the nonlinear plant model and the MPC controller can be approximated by a PWA system as well.

The explicit representation of the MPC controller discussed above is significant for several reasons. First of all, it gives some insight into the mathematical structure of the controller which is otherwise hidden behind the optimization formalism. Furthermore it offers an alternative route to an efficient controller implementation, opening up the route to use MPC in "fast" and "cheap" systems where the on-line solution of a QP and especially an MILP is prohibitive. Finally, the fact that we can represent the closed loop system in a PWA form allows us to apply new tools for performance analysis as discussed below.

Verification of MLD and PWA Systems

In this section we will briefly review the topic of verification and sketch the available algorithms. There are numerous practical applications of verification, two of which (checking the correctness of an emergency shut down control system and performance analysis of MPC) will be discussed in some detail.

The problem of verification of hybrid systems, or, in system theoretical words, the *reachability analysis* of hybrid systems can be defined as follows:

Reachability Analysis Problem. Given a hybrid system Σ (either in PWA form (1) or MLD (3)), a set of initial conditions $\mathcal{X}(0)$, a collection of disjoint target sets $\mathcal{Z}_1$, $\mathcal{Z}_2$, ..., $\mathcal{Z}_L$, a bounded set of inputs $\mathcal{U}$, and a time horizon $t \leq T_{\max}$, determine (*i*) if $\mathcal{Z}_j$ is reachable from $\mathcal{X}(0)$ within $t \leq T_{\max}$ steps for some input sequence $\{u(0), \ldots, u(t-1)\} \subseteq \mathcal{U}$; (*ii*) if yes, the subset of initial conditions $\mathcal{X}_{\mathcal{Z}_j}(0)$ of $\mathcal{X}(0)$ from which $\mathcal{Z}_j$ can be reached within $T_{\max}$ steps; (*iii*) for any $x_1 \in \mathcal{X}_{\mathcal{Z}_j}(0)$ and $x_2 \in \mathcal{Z}_j$, the input sequence $\{u(0), \ldots, u(t-1)\} \subseteq \mathcal{U}$, $t \leq T_{\max}$, which drives x_1 to x_2.

We will denote by $\mathcal{X}(t, \mathcal{X}(0))$ the *reach set* at time t starting from any $x \in \mathcal{X}(0)$ and by applying any input $u(k) \in \mathcal{U}$, $k \leq t-1$.

Although finite time reachability analysis cannot answer certain "liveness" questions (for instance, if $\mathcal{Z}_i$ will be ever reached), the reachability problem stated above is decidable. The reason for focusing on finite-time reachability is that the time-horizon $T_{\max}$ has a clear meaning, namely that states which are reachable in more than $T_{\max}$ steps are in practice unreachable. Many undecidable problems can be approximated by decidable ones which are equivalent from a practical point of view. The decidable algorithm shown in (Bemporad et al., 2000a) for observability analysis, and the decidable stability analysis proposed in (Bemporad et al., 2000) are other examples of such a philosophy. Nevertheless, the problem is $\mathcal{NP}$-hard.

Verification

Algorithm. Assume $\mathcal{X}(0) \subset C_i$ is a convex polyhedral set. With the algorithm introduced in (Bemporad et al., 2000) computing the evolution $\mathcal{X}(T_{\max}, \mathcal{X}(0))$ requires: (*i*) the reach set $\mathcal{X}(t, \mathcal{X}(0)) \bigcap \mathcal{C}_i$, i.e., the set of evolutions at time t in $\mathcal{C}_i$ from $\mathcal{X}(0)$; (*ii*) crossing detection of the guardlines $\mathcal{P}_h \triangleq \mathcal{X}(t, \mathcal{X}(0)) \bigcap \mathcal{C}_h \neq \emptyset$, $\forall h = 0, \ldots, i-1, i+1, \ldots, s-1$; (*iii*) elimination of redundant constraints and approximation of the polyhedral representation of the new regions $\mathcal{P}_h$ (approximation is desirable, as the number of facets of $\mathcal{P}_h$ can grow linearly with time); (*iv*) detection of emptiness of $\mathcal{X}(t, \mathcal{P}_h)$ (emptiness happens when all the evolutions have crossed the guardlines) and detection of $\mathcal{X}(t, \mathcal{P}_h) \subseteq \mathcal{Z}_j$, $j = 1, \ldots, L$ (these will be referred to as *fathoming* conditions), (*v*) detection of $\mathcal{X}(t, \mathcal{P}_h) \cap \mathcal{Z}_j \neq \emptyset$, $j = 1, \ldots, L$ (reachability detection).

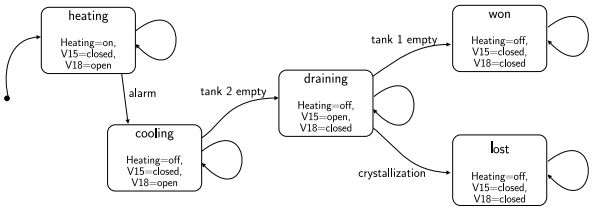


Figure 6: Model of the controller as finite state machine.

Batch Evaporator Process Benchmark. In this section we report on an application of the verification algorithm outlined in the previous section to the benchmark problem proposed within the ESPRIT-LTR Project 26270 VHS (Verification of Hybrid Systems)[2] as case study 1. The full system consists of an experimental batch plant (Kowalewski, 1998). In this paper we will focus only on the *evaporator system*, as proposed in (Kowalewski and Stursberg, 1998), which is schematically depicted in Figure 5. The considered subsystem consists of three parts: the upper tank 1 (labeled as B5 in Figure 5), the lower tank 2 (B7) and the condenser (K1). Tank 1 is equipped with a heating system (H), and is connected to tank 2 by a pipe and a valve (V15), while the outlet of tank 2 is controlled by valve V18. Both the heating system and the valves can only have two configurations: on (open) and off (closed). The levels h_1, h_2 of the solution in the two tanks, and the temperature T of tank 1 are detected by sensors. These provide four logic signals: tank 1 empty, tank 2 empty, alarm, and crystallization. A tank is considered as empty when its level is lower than 0.01 m.

During normal operation of the plant, an aqueous solution of salt enters tank 1 to be concentrated. The exiting steam flows through a condenser. When the concentration of salt has reached a certain level, the heating system is switched off, valve V15 is opened, and the solution flows to tank 2 for post-processing operations. The plant is designed in such a way that more than one batch can be produced at the same time, so that tank 1 and tank 2 can process different batches in parallel.

Here we want to analyze the exception handling needed when the condenser does not work properly. Suppose that for some reason (e.g. lack of cooling agent) the condenser malfunctions. In this case, the steam cannot be cooled down and the pressure rises in tank 1. The heating system should be switched off to prevent damages to the plant due to over-pressure. On the other hand, the temperature in tank 1 should not get lower than a critical temperature T_c, otherwise the salt may crystallize and expensive procedures would be needed to restore the original functionality.

The plant is controlled by a PLC (*programmable logic controller*). The finite state machine underlying the control code is described in Figure 6, where the event alarm occurs when $T \geq T_a$ and crystallization when $T \leq T_c$.

The control strategy can be explained as follows. When a malfunctioning of the condenser is detected, the controller enters the alarm mode and immediately opens valve V18 to empty tank 2. During this phase, the heating is still on (heating state). When the temperature in tank 1 reaches the alarm level T_a (alarm), the heating is switched off and the controller enters the state cooling. Finally, when tank 2 is empty, the controller gets in the state draining, where valve V18 is closed and valve V15 is opened, and the solution flows from the upper tank to the lower one. From draining, the controller can either switch to the state won if tank 1 gets empty, or to lost if the temperature in tank 1 gets lower than the critical value T_c.

The goal is to verify that the controller satisfies the following safety requirements: (i) if a malfunctioning in the cooling system of the condenser occurs, the heater must be turned off quickly enough to prevent damages to the condenser, (ii) the solution in tank 1 is drained to tank 2 before it eventually solidifies, (iii) when the valve 15 is open tank 2 is empty.

Certifying that the PLC code satisfies these specifications amounts to verify that from all the initial states in a given set $\mathcal{X}_0$ the system never reaches the state lost, or, equivalently, that the system always reaches the state won

Modeling the Evaporator Benchmark in MLD Form. In order to use the verification tools outlined above, we need to obtain a hybrid model of the batch evaporator process in MLD form. We consider the model described in (Stursberg, 1999), which only takes into account the heights h_1, h_2 and the temperature T of tank 1. The model is summarized in Table 2 (dynamic equations associated with each logical state), and is based on the following simplifying assumptions: (i) the pressure increase during the evaporation in the heating phase is neglected, (ii) the dynamics during the cooling phase is the same for $T \geqslant 373\ K$, (iii) average constants replace ranges of physical parameters.

After the piece-wise linear approximation of the square root relation (three sections) the model can be readily expressed in the HYSDEL language (see Appendix). The MLD model generated by the compiler includes three continuous states x_c, three logic states x_ℓ, 19 Boolean inputs δ and eight auxiliary variables z.

Verification Results. We aim at verifying that after an exception occurs, the PLC code based on the control logic of Figure 6 safely shuts down the plant to the won state for any initial condition $x(0) = \left[\begin{smallmatrix} x_c(0) \\ x_\ell(0) \end{smallmatrix}\right] \in \mathcal{X}_0 = \{T, h_1, h_2, x_\ell : T = 373, 0.2 \leq h_1 \leq 0.22, 0.28 \leq h_2 \leq 0.3, x_\ell = \left[\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right]\}$.

[2] http://www-verimag.imag.fr/VHS/

Logic state	Heating	V15	V18	Dynamic behavior
Heating	on	closed	open	$\begin{cases} \frac{\partial T}{\partial t} = k_3(q - k_4(T - t_e)) \\ \frac{\partial h_1}{\partial t} = 0 \\ \frac{\partial h_2}{\partial t} = -k2\sqrt{h_2} \end{cases}$
Cooling	off	closed	open	$\begin{cases} \frac{\partial T}{\partial t} = k_5(T - t_e) \\ \frac{\partial h_1}{\partial t} = 0 \\ \frac{\partial h_2}{\partial t} = -k2\sqrt{h_2} \end{cases}$
Draining	off	open	closed	$\begin{cases} \frac{\partial T}{\partial t} = k_5(T - t_e) \\ \frac{\partial h_1}{\partial t} = -k1\sqrt{h_1} \\ \frac{\partial h_2}{\partial t} = k2\sqrt{h_1} \end{cases}$
Won	off	closed	closed	$\begin{cases} \frac{\partial T}{\partial t} = k_5(T - t_e) \\ \frac{\partial h_1}{\partial t} = 0 \\ \frac{\partial h_2}{\partial t} = 0 \end{cases}$
Lost	off	closed	closed	$\begin{cases} \frac{\partial T}{\partial t} = k_5(T - t_e) \\ \frac{\partial h_1}{\partial t} = 0 \\ \frac{\partial h_2}{\partial t} = 0 \end{cases}$

Table 2: Hybrid model of the evaporator process.

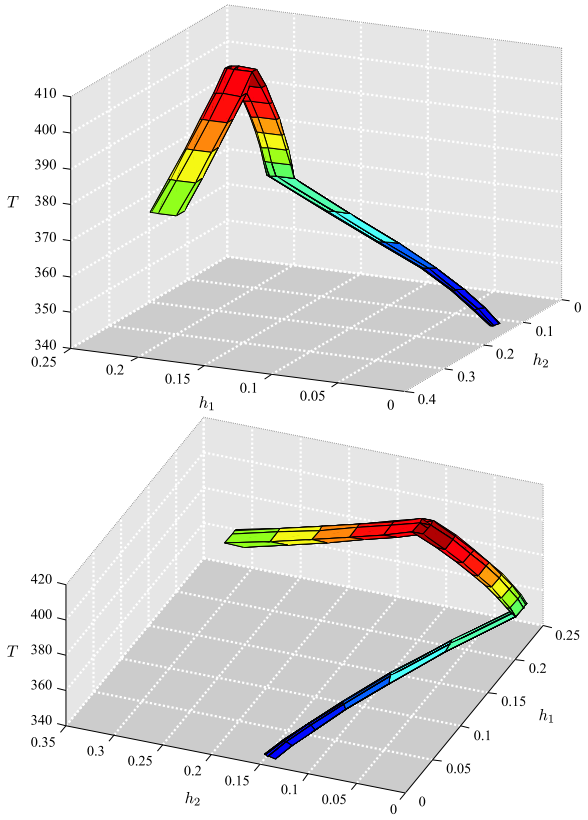


Figure 7: Set evolution from $\mathcal{X}_0$ to the target set $\mathcal{Z}_1$ (won) for $T_a = 391$ (same evolution, different perspectives).

To this end, we apply the verification algorithm presented above, and label as target set $\mathcal{Z}_1$ the set of *safe* states $\{x : x_\ell = \left[\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right]\}$ (won), and as target set $\mathcal{Z}_2$ the set of *unsafe* states $\{x : x_\ell = \left[\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right]\}$ (lost). The results of the algorithm are plotted in Figure 7, where the *set evolution* in the three-dimensional continuous state space h_1, h_2, T from the initial conditions $\mathcal{X}(0)$ is depicted from different view angles.

The tool can also easily perform parametric verification if the vector of parameters θ enters the model linearly, and its range is a polyhedral set Θ (e.g. Θ is an interval). Constant parameters can in fact be taken into account by augmenting the state vector to $\left[\begin{smallmatrix}x\\\theta\end{smallmatrix}\right]$, adding a constant dynamics $\theta(t+1) = \theta(t)$ for the additional state θ, and setting the set of initial conditions to $\mathcal{X}(0) \times \Theta$. Vice versa, varying parameters with unknown dynamics can be modeled as additional inputs to the system, i.e., as disturbances.

We use parametric verification for checking against variations of the alarm temperature T_a in the range 383 K$\leq T_a \leq$ 393 K. As T_a is a constant parameter of the PLC code, it is treated as an additional state.

The parametric verification produces the following result: for $T_a \geq 390.4902$ the controller drives the plant to the terminal state $\mathcal{Z}_1$ (won) for all the initial heights and temperature in $\mathcal{X}_0$. The parametric verification requires 82 s to build the graph of evolution on a PC Pentium II 400 MHz running interpreted Matlab code.

Performance Characterization

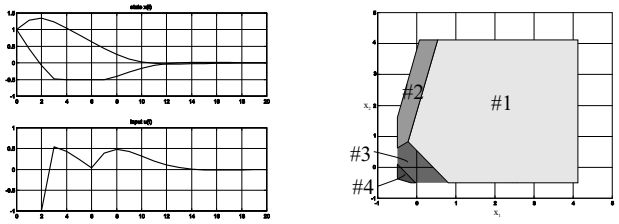
While there are many performance measures for linear systems (ranging from the traditional Integral-Square-Error to the modern H_{infty} criterion) the performance of systems with constraints under MPC is more difficult to characterize in a compact meaningful manner. Obviously, as a minimum requirement the closed loop MPC system must be stable. All the available MPC stability results hold when the associated optimization problem is feasible. Thus, a possible performance charcterization would be to determine that region of the state space for which all emanating trajectories lead to feasible optimization problems as they evolve. To arrive at a more quantitative measure we can define a region $\mathcal{C}_0$ around the origin and determine that region $\mathcal{D}_T(0)$ of the state space for which all emanating trajectories lead into $\mathcal{C}_0$ in T time steps. This problem can also be solved by the proposed verification algorithm as detailed below.

We aim at estimating the domain of attraction of the origin, and the set of initial conditions from which the state trajectory remains feasible for the constraints. As mentioned in the previous section, the nominal MPC closed-loop system is an autonomous PWA system. The origin belongs to the interior of one of the sets of the partition, namely the region where the LQ gain K is asymptotically stabilizing while fulfilling the constraints, which by convention will be referred to as $\mathcal{C}_0$. Denote by $\mathcal{D}_\infty(0) \subseteq \mathbb{R}^n$ the (unknown) domain of attraction of the origin. Given a bounded set $\mathcal{X}(0)$ of initial conditions, we want to characterize $\mathcal{D}_\infty(0) \bigcap \mathcal{X}(0)$.

By construction, the matrix A_0, associated with the region $\mathcal{C}_0$, is strictly Hurwitz and $f_0 = 0$ (in fact, in $\mathcal{C}_0$ the feedback gain is the unconstrained LQR gain $F_0 = K$, $g_0 = 0$ (Bemporad et al., 2002)). Then we can compute an invariant set in $\mathcal{C}_0$. In particular, we compute the *maximum output admissible set* (MOAS) $\mathcal{X}_\infty \subseteq \mathcal{C}_0$. $\mathcal{X}_\infty$ is the largest invariant set contained in $\mathcal{C}_0$, which by construction of $\mathcal{C}_0$ is compatible with the constraints $u_{\min} \leq Kx(t) \leq u_{\max}$, $x_{\min} \leq x(t) \leq x_{\max}$. By (Gilbert and Tan, 1991, Th.4.1), the MOAS is a polyhedron with a finite number of facets, and is computed through a finite number of linear programs (LP's) (Gilbert and Tan, 1991).

In order to circumvent the undecidability of stability, we give the following

Definition 1 *Consider the PWA system (1), and let the origin* $0 \in \mathring{\mathcal{C}}_0 \triangleq \{x : H_0 x < S_0\}$, *and* A_0 *be strictly Hurwitz. Let* $\mathcal{X}_\infty$ *be the maximum output admissible set (MOAS) in* $\mathcal{C}_0$, *which is an invariant for the linear system* $x(t+1) = A_0 x(t)$. *Let* T *be a finite time horizon. Then, the set* $\mathcal{X}(0) \subseteq \mathbb{R}^n$ *of initial conditions is said to belong to the* domain of attraction in T steps $\mathcal{D}_T(0)$ *of the origin if* $\forall x(0) \in \mathcal{X}(0)$ *the corresponding final state* $x(T) \in \mathcal{X}_\infty$.



(a) Closed-loop MPC. (b) Explicit solution.

Figure 8: Example 12.

Note that $\mathcal{D}_T(0) \subseteq \mathcal{D}_{T+1}(0) \subseteq \mathcal{D}_\infty(0)$, and $\mathcal{D}_T(0) \to \mathcal{D}_\infty(0)$ as $T \to \infty$. The horizon T is a practical information about the speed of convergence of the PWA system to the origin and thus about its dynamic performance.

Definition 2 *Consider the PWA system (1), and let* $\mathcal{X}_{\text{infeas}} \triangleq \mathbb{R}^n \setminus \cup_{i=1}^s \mathcal{C}_i$. *The set* $\mathcal{X}(0) \subseteq \mathbb{R}^n$ *of initial conditions is said to belong to the* domain of infeasibility in T steps $\mathcal{I}_T(0)$ *if* $\forall x(0) \in \mathcal{X}(0)$ *there exists* t, $0 \leq t \leq T$ *such that* $x(t) \in \mathcal{X}_{\text{infeas}}$.

In Definition (2), the set $\mathcal{X}_{\text{inf}}$ must be interpreted as a set of "very large" states. Although instability in T steps does not guarantee instability (for any finite T, a trajectory might reach $\mathcal{X}_{\text{inf}}$ and converge back to the origin), it has the practical meaning of labeling as "unstable" the trajectories whose magnitude is unacceptable, for instance because the PWA system is no longer valid as a model of the real system.

Given a set of initial conditions $\mathcal{X}(0)$, we aim at finding subsets of $\mathcal{X}(0)$ which are safely asymptotically stable $(\mathcal{X}(0) \bigcap \mathcal{D}_T(0))$, and subsets which lead to infeasibility in T steps $(\mathcal{X}(0) \bigcap \mathcal{I}_T(0))$. Subsets of $\mathcal{X}(0)$ leading to none of the two previous cases are labeled as *non-classifiable in* T *steps*. As we will use linear optimization tools, we assume that $\mathcal{X}(0)$ is a convex polyhedral set (or the union of convex polyhedral sets). Typically, non-classifiable subsets shrink and eventually disappear for increasing T.

An Example. Consider the system $y(t) = \frac{s+1}{s^2+s+2}u(t)$, and sample the dynamics with $T = 0.2$ s. The task is to regulate the system to the origin while fulfilling the constraints $-1 \leq u(t) \leq 1$ and $x(t) \geq \left[\begin{smallmatrix}-0.5\\-0.5\end{smallmatrix}\right]$. To this aim, we design an MPC controller based on the optimization problem

$$\begin{aligned} \min_{u_t, u_{t+1}} \quad & ||x_{t+2|t}||_P^2 + \sum_{k=0}^{1} ||x_{t+k|t}||^2 + .1||u_{t+k}||^2 \\ \text{subj. to} \quad & -2 \leq u_{t+k} \leq 2, \ k = 0, 1 \\ & x_{t+k|t} \geq x_{\min}, \ x_{\min} \triangleq \left[\begin{smallmatrix}-0.5\\-0.5\end{smallmatrix}\right], \ k = 0, 1 \end{aligned} \tag{12}$$

where P is the solution to the Riccati equation (in this example $Q = \left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$, $R = 0.1$, $N_u = N_y = N_c = 2$).

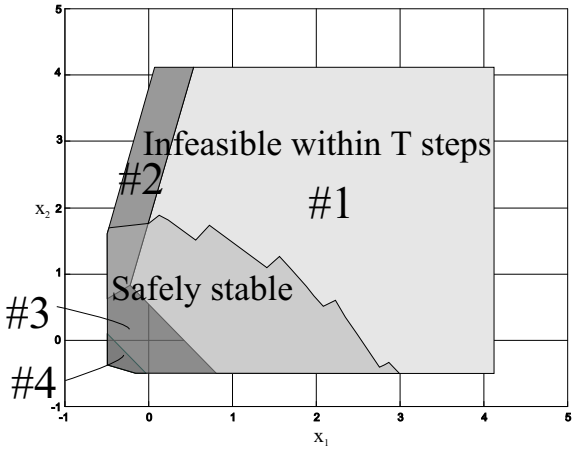


Figure 9: Partition of initial states into safely stable, and infeasible in $T = 20$ steps.

Note that this choice of P corresponds to setting $u_{t+k} = Kx_{t+k|t}$ for $k \geq 2$, where K is the LQR gain, and minimizes $\sum_{k=0}^{\infty} x'_{t+k|t}x_{t+k|t} + .01u^2_{t+k}$ with respect to u_t, u_{t+1}. The closed loop response from the initial condition $x(0) = [1\ 1]'$ is shown in Figure 8(a).

The solution to the mp-QP problem was computed by using the solver in (Bemporad et al., 2002) in 0.66 s on a PC Pentium III 650 MHz running Matlab 5.3, and the corresponding polyhedral partition of the state-space is depicted in Figure 8(b). The MPC law is linear in each one of the four depicted regions.

Region #3 corresponds to the unconstrained LQR controller, #1 and #4 to saturation at -1 and $+1$, respectively, and #2 is a transition region between LQR control and the saturation.

Note that the union of the regions depicted in Figure 8(b) should not be confused with the region of attraction of the MPC closed-loop. For instance, by starting at $x(0) = [3.5\ 0]'$ (for which a feasible solution exists), the MPC controller runs into infeasibility after $t = 5$ time steps.

The reachability analysis algorithm described above was applied to determine the set of safely stable initial states and states which are infeasible in $T = 20$ steps (Figure 9). The algorithm computes the graph of evolutions in 115 s on a Pentium II 400 running Matlab 5.3.

Conclusions

The paper argues that many unsolved problems of practical interest involve systems where dynamics and logic interact. A big subclass of such systems can be modeled in discrete time as Mixed Logic Dynamical (MLD) systems described by linear dynamic equations subject to linear inequalities involving real and integer variables. As an immediate benefit of the MLD description most system analysis and synthesis tasks can be cast as mixed integer optimization problems, for which many commercial solvers exist.

Our group has concentrated on this model paradigm and developed a wide variety of tools and techniques (only a small fraction of which was discussed in this paper), among them: HYSDEL, a modelling language for the specification of MLD systems and a compiler to generate the MLD models; a model predictive controller with an explicit representation where the optimization effort is entirely shifted off-line; an algorithm for the verification of MLD systems which is useful for a variety of tasks ranging from checking the correctness of PLC code to assessing the performance of MPC loops; several algorithms to analyze the observability of MLD systems essential for filter design, process monitoring and fault detection; several filtering algorithms based on the moving horizon idea with rather general convergence guarantees.

In collaboration with different companies we have applied the tools to a range of problems including traction control and gear shift/clutch operation on automotive vehicles, power management for electrical utilities, fault detection on a benchmark multi-tank system, optimal operation of a gas supply system, blood pressure control in anesthesia and analysis of an emergency shutdown system for a pilot batch plant.

All the investigated theoretical problems are "hard" in the mathematical sense (maybe all interesting problems are?), which implies—loosely speaking—that in the worst case the computational effort grows exponentially with the problem size. Thus the future challenge will be to develop approximate methods which provide good, if not optimal answers for problems with specific structures and where the computational effort grows only in a polynomial fashion. Otherwise the applicability of the developed tools will remain limited to small problems.

An extensive set of reports describing all aspects of our work on hybrid systems is available from our web site `http://control.ethz.ch`

Acknowledgments

The author wishes to thank Domenico Mignone who helped in the preparation of the paper.

References

Acevedo, J. and E. N. Pistikopoulos, "A Multiparametric Programming approach for linear process engineering problems under uncertainty," *Ind. Eng. Chem. Res.*, **36**, 717–728 (1997).

Alur, R., C. Courcoubetis, T. A. Henzinger, and P.-H. Ho, Hybrid Automata: an algorithmic approach to the specification and verification of hybrid systems, In R. L. Grossman, A. Nerode, A. P. R. and H. Rischel, editors, *Hybrid Systems*, volume 736 of *Lecture Notes in Computer Science*, pages 209–229. Springer Verlag (1993).

Asarin, A., O. Maler, and A. Pnueli, "On the Analysis of Dynam-

ical Systems having Piecewise-Constant Derivatives," *Theoretical Computer Science*, **138**, 35–65 (1995).

Bemporad, A. and M. Morari, "Control of Systems Integrating Logic, Dynamics, and Constraints," *Automatica*, **35**(3), 407–427 (1999).

Bemporad, A., P. Hertach, D. Mignone, M. Morari, and F. D. Torrisi, HYSDEL—Hybrid Systems Description Language, Technical Report AUT00-03, ETH Zurich (2000b).

Bemporad, A., G. Ferrari-Trecate, and M. Morari, "Observability and Controllability of Piecewise Affine and Hybrid Systems," *IEEE Trans. Auto. Cont.*, **45**(10), 1864–1876 (2000a).

Bemporad, A., F. D. Torrisi, and M. Morari, Optimization-Based Verification and Stability Characterization of Piecewise Affine and Hybrid Systems, In Krogh, B. and N. Lynch, editors, *Hybrid Systems: Computation and Control, Proceedings 3rd International Workshop on Hybrid Systems, Pittsburgh, PA, USA*, Lecture Notes in Computer Science. Springer Verlag (2000).

Bemporad, A., M. Morari, V. Dua, and E. N. Pistikopoulos, "The Explicit Linear Quadratic Regulator for Constrained Systems," *Automatica*, **38**(1), 3–20 (2002).

Branicky, M. S., V. S. Borkar, and S. K. Mitter, "A unified framework for hybrid control: model and optimal control theory," *IEEE Trans. Auto. Cont.*, **43**(1), 31–45 (1998).

Branicky, M. S., *Studies in Hybrid Systems: Modeling, Analysis, and Control*, PhD thesis, Massachussets Institute of Technology (1995).

Branicky, M. S., "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Trans. Auto. Cont.*, **43**(4), 475–482 (1998).

Cavalier, T. M., P. M. Pardalos, and A. L. Soyster, "Modeling and integer programming techniques applied to propositional calculus," *Comput. Oper. Res.*, **17**(6), 561–570 (1990).

Christiansen, D., *Electronics Engineers' Handbook, 4th edition.* IEEE Press/ McGraw Hill, Inc. (1997).

De Schutter, B. and B. De Moor, The Extended Linear Complementarity Problem and the Modeling and Analysis of Hybrid Systems, In Antsaklis, P., W. Kohn, M. Lemmon, A. Nerode, and S. Sastry, editors, *Hybrid Systems V*, volume 1567 of *Lecture Notes in Computer Science*, pages 70–85. Springer (1999).

De Schutter, B. and T. van den Boom, Model predictive control for max-plus-linear systems, In *Proc. American Contr. Conf.*, pages 4046–4050 (2000).

Fletcher, R. and S. Leyffer, A mixed integer quadratic programming package, Technical report, University of Dundee, Dept. of Mathematics, Scotland, UK (1994).

Fletcher, R. and S. Leyffer, "Numerical Experience with Lower Bounds for MIQP Branch-And-Bound," *SIAM J. Optim.*, **8**(2), 604–616 (1998). `http://epubs.siam.org/sam-bin/dbq/toclist/SIOPT`.

Floudas, C. A., *Nonlinear and Mixed-Integer Optimization.* Oxford University Press (1995).

Gilbert, E. G. and K. T. Tan, "Linear systems with state and control constraints: the theory and applications of maximal output admissible sets," *IEEE Trans. Auto. Cont.*, **36**(9), 1008–1020 (1991).

Grossmann, R. L., A. Nerode, A. P.Ravn, and H. R. (Eds.), *Hybrid Systems.* Springer Verlag, New York (1993). no. 736 in LCNS.

Heemels, W. P. M. H., J. M. Schumacher, and S. Weiland, "Linear Complementarity Systems," *SIAM J. Appl. Math.*, **60**(4), 1234–1269 (2000).

Heemels, W. P. M. H., B. De Schutter, and A. Bemporad, "Equivalence of Hybrid Dynamical Models," *Automatica*, **37**(7), 1085–1093 (2001).

Heemels, W. P. M. H., *Linear complementarity systems: a study in hybrid dynamics*, PhD thesis, Dept. of Electrical Engineering, Eindhoven University of Technology, The Netherlands (1999).

Johansson, M. and A. Rantzer, "Computation of Piecewise Quadratic Lyapunov Functions for Hybrid Systems," *IEEE Trans. Auto. Cont.*, **43**(4), 555–559 (1998).

Johansson, K. H., M. Egerstedt., J. Lygeros, and S. Sastry, "On the Regularization of Zeno hybrid automata," *Sys. Cont. Let.*, **38**, 141–150 (1999).

Kowalewski, S. and O. Stursberg, The Batch Evaporator: A Benchmark Example For Safety Analysis Of Processing Systems Under Logic Control, In *4th Int. Workshop on Discrete Event Systems (WODES 98)*, Cagliari (Italy) (1998).

Kowalewski, S., Description of VHS Case Study 1 "Experimental Batch Plant", `http://astwww.chemietechnik.uni-dortmund.de/\symbol{126}vhs/cs1descr.zi%p` (1998). Draft. University of Dortmund, Germany.

Labinaz, G., M. M. Bayoumi, and K. Rudie, "A Survey of Modeling and Control of Hybrid Systems," *Annual Reviews of Control*, **21**, 79–92 (1997).

Lazimy, R., "Improved algorithm for mixed-integer quadratic programs and a computational study," *Math Prog.*, **32**, 100–113 (1985).

Lygeros, J., D. N. Godbole, and S. Sastry, A game theoretic approach to hybrid system design, In Alur, R. and T. Henzinger, editors, *Hybrid Systems III*, volume 1066 of *Lecture Notes in Computer Science*, pages 1–12. Springer Verlag (1996).

Lygeros, J., C. Tomlin, and S. Sastry, "Controllers for reachability specifications for hybrid systems," *Automatica*, **35**(3), 349–370 (1999).

Nemhauser, G. L. and L. A. Wolsey, *Integer and Combinatorial Optimization.* Wiley (1988).

Raman, R. and I. E. Grossmann, "Relation between MILP modeling and logical inference for chemical process synthesis," *Comput. Chem. Eng.*, **15**(2), 73–84 (1991).

Roschchin, V. A., O. V. Volkovich, and I. V. Sergienko, "Models and methods of solution of quadratic integer programming problems," *Cybernetics*, **23**, 289–305 (1987).

Sontag, E. D., "Nonlinear Regulation: The Piecewise Linear Approach," *IEEE Trans. Auto. Cont.*, **26**(2), 346–358 (1981).

Sontag, E. D., Interconnected automata and linear systems: A theoretical framework in discrete-time, In Alur, R., T. A. Henzinger, and E. D. Sontag, editors, *Hybrid Systems III—Verification and Control*, number 1066 in Lecture Notes in Computer Science, pages 436–448. Springer-Verlag (1996).

Stursberg, O., The Batch-Evaporator Benchmark Example—A simplified formulation (1999). `http://astwww.chemietechnik.uni-dortmund.de/~olaf/`.

Tyler, M. L. and M. Morari, "Propositional logic in control and monitoring problems," *Automatica*, **35**(4), 565–582 (1999).

van der Schaft, A. J. and J. M. Schumacher, "Complementarity Modelling of Hybrid Systems," *IEEE Trans. Auto. Cont.*, **43**, 483–490 (1998).

Williams, H. P., "Logical problems and integer programming," *Bulletin of the Institute of Mathematics and Its Applications*, **13**, 18–20 (1977).

Williams, H. P., *Model Building in Mathematical Programming.* John Wiley & Sons, Third Edition (1993).

Appendix

HYSDEL (HYbrid Systems DEscription Language)

The derivation of the MLD model, i.e., a set of linear difference equations and mixed-integer linear inequalities, from an interconnection of components involving continuous systems and logic is an involved tedious task for all but the most simple example problems. Therefore we have developed a HYbrid Systems DEscription Language (HYSDEL) for the specification of such systems and a compiler which readily generates the equivalent MLD form. The HYSDEL compiler is available on-line at `http://control.ethz.ch/~hybrid/hysdel`. Thanks to the equivalence between the various hybrid system descriptions mentioned in the paper, the MLD form can be used as an intermediate step to obtain the corresponding PWA, LC, ELC, or MMPS counterpart.

In HYSDEL systems are viewed as the interconnection of basic objects. Each object admits an equivalent representation as linear mixed-integer equalities and inequalities. The basic objects are:

- **A/D (Analog-to-Digital) block.** Can extract logic facts from the activity level of linear thresholds.A graphical representation is provided in Figure 10.
- **D/A (Digital-to-Analog) block.** Is the counterpart of the A/D block. This block (see Figure 11) is able to associate with the output different linear expressions according to the different logic value of the input.
- **Automaton.** Evolves according to the logic part of the overall system input and the logic signals coming out from the A/D blocks and from other automata. As the schematic representation depicted in Figure 12 shows, the automaton typically makes its internal state available to other components.
- **Continuous dynamics.** Is a *Discrete-time Linear Time Invariant* (D-LTI) system and the different modes of operation are obtained by connecting (see Figure 13) the input to the output of a D/A block.

As an illustration the HYSDEL code for the batch evaporator example is shown in Figure 14.

Figure 10: A/D block—Continuous to logic conversion.

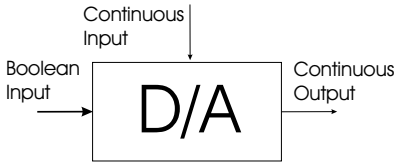


Figure 11: D/A block—Logic to continuous conversion.

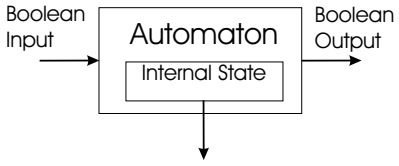


Figure 12: Automaton.

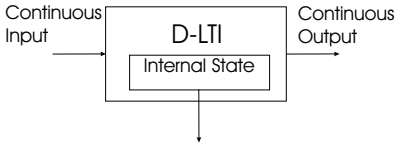


Figure 13: Discrete-time-invariant linear continuous dynamics.

```
/* VHS Esprit Project - Case Study 1 */

SYSTEM batchevaporator {
INTERFACE {
    STATE {
        REAL tmp,h1,h2,tal;
        BOOL p1,p2,p3; }
    OUTPUT {
        REAL outreal1;
        BOOL outbool1; }
    PARAMETER {
        REAL q = 5000;    /* kW     */

                (other parameters are omitted for brevity)

    }
}

IMPLEMENTATION {
    AUX {
        REAL zT, zh1a, zh1b, zh1c, zh2a, zh2b, zh2c, zh2d;

        BOOL ti1,ti2,l1,l2,h,v18,v15,d1;
        BOOL l1h1,l1h2; /* Linearization of sqrt */
        BOOL da1,da2,da3,da4,da5,da6,da7;  }
    LOGIC {
        h = ~p1&~p2&~p3;
        v18 = ~p1&~p2&~p3 | ~p1&~p2&p3;
        v15 = p1&~p2&~p3;
        da1=v18&~l1&l1h1;
        da2=v18&~l1&~l1h1;
        da3=~v18&~l1;
        da4=v15&~l2&l1h2;
        da5=~v15&~l2&l1h2;
        da6=v15&~l2&~l1h2;
        da7=~v15&~l2&~l1h2;  }
    AD {
        ti1 = -tmp+tal <= 0 [-Tmin+Talmax,-Tmax+Talmin,1e-2];
        ti2 = tmp-338 <= 0 [Tmax-338,Tmin-338,1e-2];
        l1 = h1-0.01 <= 0 [hmax-0.01,hmin-0.01,1e-6];
        l2 = h2-0.01 <= 0 [hmax-0.01,hmin-0.01,1e-6];
        l1h1 = h1-l1h1t <= 0 [hmax-l1h2t,hmin-l1h2t,1e-6];
        l1h2 = h2-l1h2t <= 0 [hmax-l1h2t,hmin-l1h2t,1e-6];   }
    DA {
        zT =   {IF h THEN atmp1*tmp+btmp1 [atmp1*Tmax+btmp1,atmp1*Tmin+btmp1,0]
                     ELSE atmp2*tmp+btmp2 [atmp2*Tmax+btmp2,atmp2*Tmin+btmp2,0] };
        zh1a = {IF da1 THEN ah1a*h1+bh1a [ah1a*hmax+bh1a,ah1a*hmin+bh1a,0]      };
        zh1b = {IF da2 THEN ah1b*h1+bh1b [ah1b*hmax+bh1b,ah1b*hmin+bh1b,0]      };
        zh1c = {IF da3 THEN h1 [hmax,hmin,0]                                    };
        zh2a = {IF da4 THEN ahh2a*h1+h2+bhh2a [ahh2a*hmax+hmax+bhh2a,ahh2a*hmin+hmin+bhh2a,0]};
        zh2b = {IF da5 THEN ah2a*h2+bh2a [ah2a*hmax+bh2a,ah2a*hmin+bh2a,0]      };
        zh2c = {IF da6 THEN ahh2b*h1+h2+bhh2b [ahh2b*hmax+hmax+bhh2b,ahh2b*hmin+hmin+bhh2b,0]};
        zh2d = {IF da7 THEN ah2b*h2+bh2b [ah2b*hmax+bh2b,ah2b*hmin+bh2b,0]      };   }
    CONTINUOUS {
        tmp =  zT;
        h1 = zh1a + zh1b + zh1c;
        h2 = zh2a + zh2b + zh2c + zh2d;
        tal = tal;  }
    AUTOMATA {
        p1= (~p1&~p2&p3&l2) | (p1&~p2&~p3&~ti2&~l1);
        p2= (p1&~p2&~p3&l1) | (p1&~p2&~p3&ti2) | (~p1&p2&~p3) | (~p1&p2&p3);
        p3= (~p1&~p2&~p3&ti1) | (~p1&~p2&p3&~l2) | (p1&~p2&~p3&ti2) | (~p1&p2&p3);   }
    MUST   {
        ~(ti1 & ti2);      /* Excludes combination ti1,ti2=11 */
        ~(p1 & (p2 | p3)); /* Excludes logical states 101,110,111 */
        ~l1h1|~l1;
        ~l1h2|~l2;  }
}
}
```

Figure 14: Example of HYSDEL code for the batch evaporator example.

Discrete Optimization Methods and their Role in the Integration of Planning and Scheduling

Ignacio E. Grossmann*, Susara A. van den Heever and Iiro Harjunkoski
Department of Chemical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

The need for improvement in process operations, logistics and supply chain management has created a great demand for the development of optimization models for planning and scheduling. In this paper we first review the major classes of planning and scheduling models that arise in process operations, and establish the underlying mathematical structure of these problems. As will be shown, the nature of these models is greatly affected by the time representation (discrete or continuous), and is often dominated by discrete decisions. We then briefly review the major recent developments in mixed-integer linear and nonlinear programming, disjunctive programming and constraint programming, as well as general decomposition techniques for solving these problems. We present a general formulation for integrating planning and scheduling to illustrate the models and methods discussed in this paper.

Keywords

Planning, Scheduling, Optimization, Mixed-integer programming

Introduction

The development of optimization models for planning and scheduling of chemical processes has received significant attention over the last 5-7 years. One major reason has been the realization by industry that large potential savings can be achieved by improving the logistics of manufacturing in chemical processes. Examples of savings include lower inventories, lower transition costs, and reduction in production shortfalls. The interest in planning and scheduling has further increased with industry's goal of improving the management and dynamics of their supply chains. Finally, major advances in large-scale computation and mathematical programming have promoted the interest in applying these techniques to planning and scheduling problems.

The goal of this paper is to provide an overview of the optimization based models for planning and scheduling, review the solution strategies and mathematical programming methods that are available for solving these problems, and propose a conceptual model for integrating planning and scheduling. Finally, we present three examples to illustrate the application of some of the techniques discussed in this paper.

Review on Planning and Scheduling

Both planning and scheduling deal with the allocation of available resources over time to perform a collection of tasks. In the context of process systems, planning and scheduling refer to the strategies of allocating equipment and utility or manpower resources over time to execute processing tasks required to manufacture one or several products. The difference between planning and scheduling is not always clear cut. However, in general the difference is that planning deals with longer time horizons (e.g. weeks, few months) and is largely concerned with high level decisions such as investment in new facilities and production levels. Scheduling on the other hand deals with shorter time horizons (e.g. days, few weeks) with the emphasis often being on the lower level decisions such as sequencing of operations. Also, in planning maximization of profit usually plays a major role, while in scheduling the emphasis tends to be on feasibility for fulfilling a given number of orders, or on completing the required tasks in the shortest time. Hence, economics tends to play a greater role in planning than in scheduling. It should be noted, however, that the distinction between planning and scheduling is becoming increasingly blurred by the capability of optimizing simultaneous planning and scheduling decisions, particularly in the context of supply chain optimization problems.

Planning

A detailed review of planning and scheduling is out of the scope of this paper. In this section we therefore focus on pointing the reader to some useful papers covering both specific planning problems and reviews on some problem classes, as well as a general discussion on the nature of planning problems. While no review alone covers all types of planning problems, some reviews can be found in the operations research literature for specific types of planning problems. Erengüç et al. (1999) review work on the integrated production and distribution planning of supply chains. They discuss the different stages of the supply chain, give some general formulations and critically evaluate the relevant literature from the operations research community. Other reviews discuss planning models for freight transportation (Crainic and Laporte, 1997), optimization methods for electric utility resource planning (Hobbs, 1995), and strategic facility

*Author to whom correspondence should be addressed. email: grossmann@cmu.edu.

location methods that consider either stochastic or dynamic problem characteristic (Owen and Daskin, 1998). In the chemical engineering literature, a review on literature for single- and multi-site planning and scheduling can be found in (Shah, 1998), while reviews on planning and scheduling literature for batch/semicontinuous plants can be found in Reklaitis (1991, 1992) and Rippin (1993).

Planning problems can mainly be categorized as strategic, tactical or operational, depending on the decisions involved and the time horizon under consideration. Strategic planning covers the longest time horizons in the range of one to several years and decisions cover the whole width of the organization, while focussing on major investments. Examples of strategic planning problems include facility location problems (e.g. Mazzola and Neebe, 1999), hydrocarbon well platform investment planning (e.g. Iyer et al., 1998; Van den Heever and Grossmann, 2000), and longterm planning of process networks (e.g. Sahinidis et al., 1989) where it is essential to consider the far future in making big investment decisions. Tactical planning typically covers the midterm horizon of between a few months to a year and decisions cover issues such as production, inventory and distribution. Midterm production planning or supply chain planning is a good example of tactical planning (e.g. Bok et al., 2000; McDonald and Karimi, 1997; Perea et al., 2000; Dimitriadis et al., 1997). Operational planning usually covers a horizon of one week to three months and involves decisions regarding the actual operations and resource allocation. Applications include the operational planning of utility systems (e.g. Iyer and Grossmann, 1998b) and the planning of refinery operations (e.g. Moro et al., 1998). On this level, planning decisions are often closely related to scheduling decisions and it becomes more important to integrate these. In the past planning and scheduling issues have mostly been addressed separately or sequentially for reasons of complexity, and only recently have simultaneous planning and scheduling approaches emerged. Birewar and Grossmann (1990) proposed a model for the simultaneous planning and scheduling of multipurpose batch plants, while Shah and Pantelides (1991) presented a model for simultaneous campaign formation and planning. Papageorgiou and Pantelides (1996a,b) address the issue in a two part article proposing a mathematical formulation and decomposition approach for integrated campaign planning and scheduling of multipurpose batch/semicontinuous plants. We address the integration of planning and scheduling in Section 4 through a generalized disjunctive model.

In terms of uncertainties, planning models have either a deterministic or stochastic nature. Deterministic models assume predictions for prices, demands and availabilities to be known with certainty. These models are often sufficient for short-term planning and scheduling, but when longer time horizons are considered incorporating uncertainty directly becomes more important. However, deterministic models are still useful even when uncertainty needs to be incorporated, since they can be used to analyze different scenarios for the uncertain parameters without the additional complexity associated with stochastic models. In addition, deterministic models form the basis for stochastic models that include uncertainty through scenarios. Stochastic models include uncertainty either by explicit probability distributions or by scenarios, and often require specialized solution methods due to their complexity. A vast number of articles have been published in the area of process planning under uncertainty and a complete list of all relevant ones cannot be provided within the scope of this paper. We therefore refer readers to some recent publications: Liu and Sahinidis (1996) proposed a two-stage stochastic programming approach for process planning under uncertainty. These authors first consider discrete random parameters and develop a Benders-based decomposition algorithm for the solution. They then continue to show the applicability of their approach to continuous random variables. Ierapetritou et al. (1996) discuss modeling issues in design and planning under uncertainty and propose a decomposition algorithm for a combined multiperiod/stochastic programming problem. Clay and Grossmann (1997) consider planning problems with uncertainty in both demands and cost coefficients, and represent these uncertainties by finite discrete probability distribution functions. They also propose an iterative aggregation/disaggregation algorithm that is highly parallel for the solution of this two-stage stochastic programming problem. Ahmed and Sahinidis (1998) propose a linear method of incorporating robustness of the second stage decisions into two-stage models for process planning under uncertainty and present a heuristic solution approach, but emphasize the need to exact solution methods for this model. Rather than using a stochastic optimization framework, Applequist et al. (2000) proposed a risk measure for supply chain optimization. Some significant progress has also been made in incorporating uncertainty into combined planning and scheduling models. Petkov and Maranas (1997) extend the combined planning and scheduling model first proposed by Birewar and Grossmann (1990) to include demand uncertainties, and express the stochastic elements in equivalent deterministic forms to yield solution times comparable to pure deterministic models. A scenario-based approach to incorporate uncertainty at the planning level for an online scheduler of a multiproduct batch plant was proposed by Sand et al. (2000). Their model has two hierarchical levels, where uncertainty at the planning level is incorporated explicitly in the upper level. While the above mentioned approaches show significant progress in incorporating uncertainty into large-scale planning and scheduling models, the characterization of uncertainties and development of efficient solution algorithms remain

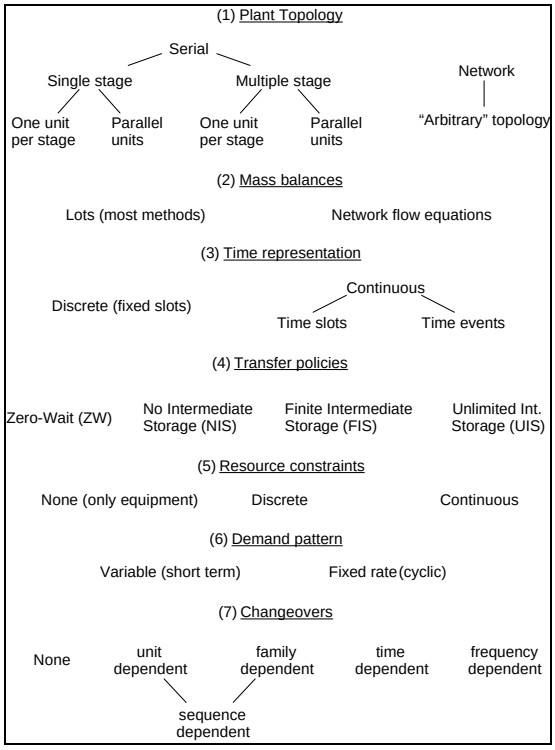


Figure 1: Classification and road map scheduling problems.

complex and challenging problems.

Scheduling

Comprehensive reviews on scheduling can be found in Rippin (1993) who addresses the general status of batch processing systems engineering with emphasis in design, planning and scheduling. Reklaitis (1991, 1992) presents a comprehensive review of scheduling and planning of batch process operations. His main focus is to describe the basic components of the scheduling problem and review the existing solution methods. Pekny and Zentner (1994) summarize the basic scheduling technology with association to the advances in computer technology. Grossmann et al. (1996) provide an overview of mixed integer optimization techniques for the design and scheduling of batch processes, with emphasis on general purpose methods for mixed integer linear (MILP) and mixed integer nonlinear (MINLP) problems. Pinto and Grossmann (1998) present a classification of scheduling problems, and characterize the major types of integer and mixed-integer constraints that arise for the assignment and sequencing decisions. Shah (1998) presents an overview of single and multisite scheduling methods, while Pekny and Reklaitis (1998) provide a review in terms of the computational complexity that is involved in scheduling problems.

A major difficulty that has been faced in the area of scheduling is that there is a great diversity of problems that have tended to prevent the development of unified solutions. To appreciate this issue, consider Figure 1 from Pinto and Grossmann (1998) that presents a road map for classifying scheduling problems. Equipment similarity and unit connectivity define the topology of the plant. In serial plants, products follow the same production path, therefore it is possible to recognize a specific direction in the plant floor. Networks of arbitrary topology tend to occur when products have low recipe similarity and/or when equipment is interconnected. Most methods do not handle mass balances explicitly; instead, production is represented by batches (or lots). Products follow a series of tasks, which are collections of elementary chemical and physical processing operations. Note the close relationship between the plant topology and the sequence of tasks for products: if all products follow the same sequence of tasks it is usually possible to define processing stages in the plant, defined as processing equipment that can perform the same operations. Moreover, lot sizes can be variables, such as in the case of the lot-sizing problem, or fixed parameters.

A major issue in modeling scheduling problems concerns the time domain representation. The most general is a continuous time domain representation that makes use of either time slots of variable length, or time events. If a discrete time representation is adopted, slots have equal and fixed duration. In this case there is the need to use a sufficiently large number of slots in order to have a suitable approximation of the original problem. An advantage, however, with discrete time domain is that it is much easier to handle resource constraints or track inventory levels. In continuous time formulations it is usually possible to postulate a much smaller number of time slots or time events reducing the problem size, although often at the expense of introducing nonlinearities in the model.

Another major issue in plant scheduling deals with the presence of intermediate storage. There are four different transfer policies: Zero-Wait (ZW), No-Intermediate-Storage (NIS), Finite-Intermediate-Storage (FIS) and Unlimited-Intermediate-Storage (UIS) (Ku et al., 1987). It is important to note that FIS corresponds to the most general case. Nevertheless, the main advantage of the remaining three cases is that there is no need to model inventory levels. In the scheduling of a process plant, processing tasks require utilities and manpower. Utilities may include, for example, steam, electricity, cooling water, etc. In some scheduling applications, apart from equipment, finite resources that are limited are required for these process tasks. Resource constrained scheduling problems are inherently difficult, due to the fact that besides the efficient allocation of units to meet product demands, it is also necessary to consider the feasible grouping of simultaneously executed tasks so as to utilize

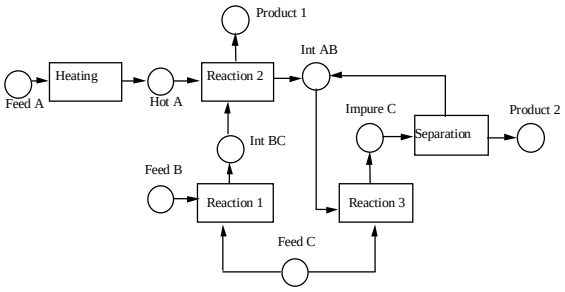


Figure 2: State-task network for numerical example.

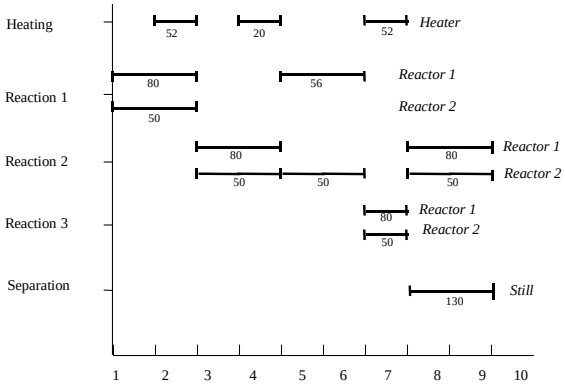


Figure 3: Optimal schedule for network in Figure 2.

resources within their availability limits.

Short term scheduling is relevant to plants that must satisfy individual customer orders with varying demand patterns. In this case, product requirements are given as a set of orders, where each order has associated with it a certain product, the amount and a due date. In contrast, cyclic scheduling is relevant for plants operating with a stable market in which the product demands are given as constant rates. This allows a more simplified plant operation in which the same production sequence is executed repeatedly with a fixed frequency. When switching between products, or even after one or more batches of the same product, units may require cleaning and setup for safety and/or product quality. Changeover requirements depend on the nature of the units and the products in the plant. Sequence dependent changeovers represent the most general and difficult situation, in which every pair of consecutive operations may give rise to different time and/or cost requirements. The need for unit setup may be expressed in terms of the frequency of utilization. For instance, a changeover may be needed after every batch or after a certain number of batches, regardless of the nature of the products. In the case of time dependent cleaning, there is a maximum time interval during which a unit may be utilized.

From all the scheduling models that have been proposed in the chemical engineering literature, the most general model is the one by Kondili et al. (1993), which addresses short term scheduling of batch operations. Major capabilities of this multiperiod MILP model include the following: (a) assignments of equipment to processing tasks need not be fixed, (b) variable size batches can be handled with the possibility of mixing and splitting, (c) different intermediate storage and transfer policies can be accommodated, as well as limitations of resources. In the work by Kondili et al. (1993) a major assumption is that the time domain can be discretized in intervals of equal size, which in practice often means having to perform some rounding to the original data. In addition, changeover times are usually neglected since they cannot be easily handled by this model. The key aspect in the MILP model by Kondili et al. (1993), is the state-task network (STN) representation. This network has two types of nodes: (a) state nodes that correspond to feeds, intermediates and final products; (b) task nodes that represent processing steps. Figure 2 presents an example of a state task network. It should be noted that the equipment is considered separately. In general it is assumed that each unit can perform several of the tasks in the STN network. The resulting MILP model determines the timing of the operations, assignments of equipment to operations, and flow of material through the network. The objective is to maximize a given profit function. Figure 3 shows the results of the optimal schedule of the example in Figure 2. It should be noted that the reformulation by Shah et al. (1993) led to a significant improvement in the LP relaxation of the MILP, with which fairly large problems can be solved. Furthermore, Pantelides (1994) proposed the Resource Task Network (RTN) representation, which leads to a more compact model than the STN, although it is actually equivalent. It is interesting to note that in the context of Figure 1, both the STN and RTN models can handle networks with arbitrary topology, can handle flow equations for the mass balances, are based on discrete time representation, can handle all types of transfers and resource constraints, and deal with short term variable demands. Continuous time versions of this model have been proposed for instance by Zhang and Sargent (1996), Mockus and Reklaitis (1996), and Ierapetritou and Floudas (1998).

Mathematical Programming

Planning and scheduling problems generally give rise to discrete/continuous optimization problems and we therefore find a discussion on the major mathematical programming techniques appropriate in the current context. When these optimization problems are represented in algebraic form, they correspond to mixed-integer optimization problems that have the following form:

$$\min Z = f(x, y) \qquad \text{(MIP)}$$

subject to

$$h(x,y) = 0$$
$$g(x,y) \leq 0$$
$$x \in X, \quad y \in \{0,1\}^m$$

where $f(x,y)$ is the objective function (e.g. cost), $h(x,y) = 0$ are the equations that describe the performance of the system (material balances, production rates), and $g(x,y) \leq 0$ are inequalities that define the specifications or constraints for feasible plans and schedules. The variables x are continuous and generally correspond to state variables, while y are the discrete variables, which generally are restricted to take 0-1 values to define for instance the assignments of equipment and sequencing of tasks. Problem (MIP) corresponds to a mixed-integer nonlinear program (MINLP) when any of the functions involved are nonlinear. If all functions are linear it corresponds to a mixed-integer linear program (MILP). If there are no 0-1 variables, the problem (MIP) reduces to a nonlinear program (NLP) or linear program (LP) depending on whether or not the functions are linear.

The formulation and solution of major types of mathematical programming problems can be effectively performed with modeling systems such as GAMS (Brooke et al., 1992), and AMPL (Fourer et al., 1992). While these require that the model be expressed explicitly in algebraic form, they have the advantage that they automatically interface with codes for solving the various types of problems. They also perform automatic differentiation and allow the use of indexed equations, with which large scale models can be readily generated. It should also be noted that these modeling systems are now widely available on desktop PCs. We review the major classes of mathematical programming models in the following paragraphs.

Linear and Mixed-Integer Programming

These are without a doubt the types of models that are most frequently encountered for planning and scheduling. The reason is that these models involve in most cases discrete time representations coupled with fairly simple performance models. While in the past most models were LPs, most of them are nowadays MILPs due to the discrete decisions that are involved in investment, expansion and operation for planning, and assignment and sequencing decisions for scheduling.

Mixed-integer linear programming problems have the general form:

$$\min Z = a^T x + b^T y \qquad \text{(MILP)}$$

subject to

$$Ax + By \leq d$$
$$x \geq 0, \quad y \in \{0,1\}^m$$

For the case when no discrete variables y are involved, the problem reduces to a linear programming (LP) problem. This is a special class of convex optimization problems for which the optimal solution lies at a vertex of the polytope defined by the inequalities . The solution of LP problems relies largely on the simplex algorithm (Chvatal, 1983; Saigal, 1995), although lately interior-point methods (Marsten et al., 1990) have received increased attention for solving very large problems because of their polynomial complexity. MILP methods rely largely on simplex LP-based branch and bound methods (Nemhauser and Wolsey, 1988) that consists of a tree enumeration in which LP subproblems are solved at each node, and eliminated based on bounding properties. These methods are being improved through cutting plane techniques (Balas et al., 1993), which produce tighter lower bounds for the optimum. LP and MILP codes are widely available. The best known include CPLEX (ILOG Inc., 2000), OSL (IBM, 1992) and XPRESS (Dash Associates, 1999), all which have achieved impressive improvements in their problem solving capabilities. It is worth noting that since MILP problems are NP-complete it is always possible to run into time limitations when solving problems with a large number of 0-1 variables, especially if the integrality gap (difference between optimal integer objective and optimal LP relaxation) is large.

Nonlinear Programming

NLP models have the advantage over LP models of being able to explicitly handle nonlinearities and are largely used for real-time optimization. These models only involve continuous variables and are fairly restrictive for planning and scheduling, although they are important subproblems in MINLPs. NLP problems correspond to continuous optimization problems that can be expressed as follows:

$$\min Z = f(x) \qquad \text{(NLP)}$$

subject to

$$h(x) = 0$$
$$g(x) \leq 0$$
$$x \in X$$

Provided the functions are continuous and differentiable, and certain constraint qualifications are met, a local optimum solution to problem (NLP) is given by the Karush-Kuhn-Tucker conditions (Minoux, 1983). The solution of NLP problems (Fletcher, 1987; Bazaraa et al., 1994), relies either on the successive quadratic programming algorithm (SQP) (Han, 1976; Powell, 1978; Schittkowski, 1981), or on the reduced gradient method (Murtagh and Saunders, 1978, 1982). Major codes include MINOS and CONOPT (Drud, 1992) for the reduced gradient method, and OPT (Vasantharajan et al., 1990) for the

SQP algorithm. These NLP methods are guaranteed to find the global optimum if the problem is convex (i.e. convex objective function and constraints). When the NLP is nonconvex a global optimum cannot be guaranteed. One option is to try to convexify the problem, usually through exponential transformations, although the number of cases where this is possible is rather small. Alternatively, one could use rigorous global optimization methods, which over the last few years have made significant advances. These methods assume that special structures are present in the problem, such as bilinear, linear fractional and concave separable functions. Although this may appear to be quite restrictive, Smith and Pantelides (1996) have shown that algebraic models are always reducible to these structures, provided they do not involve trigonometric functions. For a general review on global optimization see Horst and Tuy (1993), Horst and Pardalos (1995) and Floudas (2000). Recent developments in chemical engineering can be found in Grossmann (1996). Computer codes for global optimization still remain in the academic domain, and the best known are BARON by Sahinidis and Sahinids and Ryoo (1995), and α-BB by Floudas et al. (1996). It should also be noted that non-rigorous techniques which have also become popular, such as simulated annealing (Kirkpatrick et al., 1983) and genetic algorithms (Goldberg, 1989), do not make any assumptions on the functions, but cannot guarantee rigorous solutions in a finite amount of time. Furthermore, these methods do not formulate the problem as a mathematical program since they involve procedural search techniques that in turn require some type of discretization, and the violation of constraints is handled through ad-hoc penalty functions.

Mixed-integer Nonlinear Programming

MINLP models typically arise in planning and also in scheduling when using continuous time representations, particularly for cyclic policies and for nonlinear performance models. The most common form of MINLP problems is a special case of problem (MIP), in which the 0-1 variables are linear while the continuous variables are nonlinear:

$$\min Z = c^T y + f(x) \qquad \text{(MINLP)}$$

subject to

$$\begin{aligned} h(x) &= 0 \\ By + g(x) &\leq 0 \\ x \in X, \quad y &\in 0,1^m \end{aligned}$$

Major methods for MINLP problems include Branch and Bound (BB) (Gupta and Ravindran, 1985; Borchers and Mitchell, 1994; Stubbs and Mehrotra, 1996), which is a direct extension of the linear case, except that NLP subproblems are solved at each node. Generalized Benders Decomposition (GBD) (Benders, 1962; Geoffrion, 1972), and Outer-Approximation (OA) (Duran and Grossmann, 1986; Yuan et al., 1988; Fletcher and Leyffer, 1994; Ding-Mai and Sargent, 1992), are iterative methods that solve a sequence of alternate NLP subproblems with all the 0-1 variables fixed, and MILP master problems that predict lower bounds and new values for the 0-1 variables. The difference between the GBD and OA methods lies in the definition of the MILP master problem; the OA method uses accumulated linearizations of the functions, while GBD uses accumulated Lagrangean functions parametric in the 0-1 variables. The LP/NLP based branch and bound by Quesada and Grossmann (1992) essentially integrates both subproblems within one tree search, while the Extended Cutting Plane Method (ECP) (Westerlund and Pettersson, 1995) does not solve the NLP subproblems, and relies exclusively on successive linearizations. All these methods assume convexity to guarantee convergence to the global optimum. Nonrigorous methods for handling nonconvexities include the equality relaxation algorithm by Kocis and Grossmann (1987) and the augmented penalty version of it (Viswanathan and Grossmann, 1990). A review on these methods and how they relate to each other can be found in Grossmann and Kravanja (1997). The only commercial code for MINLP is DICOPT (OA-GAMS), although there are a number of academic versions (MINOPT by Floudas and co-workers, α-ECP by Westerlund and co-workers). (Tawarmalani and Sahinidis, 2000) have recently expanded their BARON global optimization code to MINLP problems through a number of extensions of the above methods.

Logic-based Optimization

In recent years a new trend that has emerged is to formulate and solve discrete/continuous optimization problems with logic-based optimization models and methods. These methods, which facilitate problem formulation and often reduce the combinatorial search, are starting to have a significant impact in planning and scheduling problems. The two major methods are Generalized Disjunctive Programming (GDP) (Raman and Grossmann, 1994) and Constraint Programming (Van Hentenryck, 1989).

Generalized Disjunctive Programming. The basic idea in GDP models is to use Boolean and continuous variables, and formulate the problem with an objective function, subject to three types of constraints: (a) global inequalities that are independent of discrete decisions; (b) disjunctions that are conditional constraints involving an OR operator; (c) pure logic constraints that involve only the Boolean variables. More specifically, the problem is given as follows:

$$\min Z = \sum_{k \in K} c_k + f(x) \qquad \text{(GDP)}$$

subject to

$$g(x) \le 0$$
$$\bigvee_{j \in I_k} \begin{bmatrix} y_{jk} \\ h_{jk}(x) \le 0 \\ c_k = \gamma_{jk} \end{bmatrix} \quad k \in K$$
$$\Omega(y) = True$$
$$x \in X, \quad y_{jk} \in \{True, False\}$$

where x are continuous variables and y are the Boolean variables. The objective function involves the term $f(x)$ for the continuous variables (e.g. operating cost) and the charges ck that depend on the discrete choices. The equalities/inequalities $g(x) \le 0$ must hold regardless of the discrete conditions, and $h_{jk}(x) \le 0$ are conditional constraints that must be satisfied when the corresponding Boolean variable y_{jk} is True for the j^{th} term of the k^{th} disjunction. The set I_k represents the number of choices for each disjunction defined in the set K. Also, the fixed charge c_k is assigned the value γ_{jk} for that same variable. Finally, the constraints $\Omega(y)$ involve logic propositions in sterms of Boolean variables.

Problem (GDP) represents an extension of disjunctive programming (Balas, 1985), which in the past has been used as a framework for deriving cutting planes for the algebraic problem (MIP). It is interesting to note that any GDP problem can be reformulated as an MIP problem, and vice-versa. It is more natural, however, to start with a GDP model, and reformulate it as an MIP problem. This is accomplished by reformulating the disjunctions using the convex hull transformation (Türkay and Grossmann, 1996) or with "big-M" constraints. The propositional logic statements are reformulated as linear inequalities (Raman and Grossmann, 1991, 1994). For the linear case of problem GDP, and when no logic constraints are involved, Beaumont (1991) proposed a branch and bound method that does not rely on 0-1 variables and branches directly on the equations of the disjunctions. This method was shown to outperform the solution of the alternative algebraic MILP models. Raman and Grossmann (1994) developed a branch and bound method for solving problem GDP in hybrid form, i.e. with disjunctions and mixed-integer constraints. For this they introduced the concept of "w-MIP representability" to denote those disjunctive constraints that can be transformed into mixed-integer form without loss in the quality of the relaxation. Hooker and Osorio (1997) developed a different branch and bound method which in a way is a generalization of Beaumont's method in that it does not introduce 0-1 variables, and addresses problems directly in the form of the GDP problem.

For the nonlinear case of problem (GDP), Lee and Grossmann (1999) have developed reformulations and algorithms that rely on obtaining the algebraic description of the convex hull of the nonlinear convex inequalities. The reformulations lead to tight MINLP problems, while

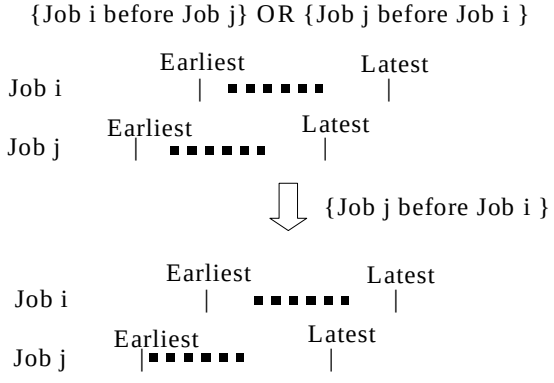


Figure 4: Edge finding technique for jobshop scheduling.

the algorithms generally involve branch and bound methods where branching is performed on disjunctions. For the case of process networks, Türkay and Grossmann (1996) proposed a logic-based Outer-Approximation algorithm. This algorithm consists of solving NLP subproblems in reduced space, in which constraints that do not apply in the disjunctions are disregarded, with which both the efficiency and robustness can be improved. In this method the MILP master problems correspond to the convex hull of the linearization of the nonlinear inequalities. Also, several NLP subproblems must be solved to initialize the master problem in order to cover all the terms in the disjunctions. Penalties can also be added to handle the effect of nonconvexities as in the method by Viswanathan and Grossmann (1990). The above methods have been implemented in the computer prototype LOGMIP, a GAMS-based computer code developed by Vecchietti and Grossmann (1997).

Constraint Programming. This area, which has emerged recently as a logic-based optimization tool, has proved to be particularly successful for certain types of scheduling problems. The basic idea in Constraint Programming (CP) (Van Hentenryck, 1989; Puget, 1994) is to use compact languages for expressing optimization problems in terms of variables that are continuous, integer, and/or Boolean, and constraints that can be expressed in algebraic form (e.g. $h(x) = 0$), as disjunctions (e.g. $[A_1 x \le b_1] \vee [A_2 x \le b_2]$), or as conditional logic statements (e.g. If $g(x) \le 0$ then $r(x) = 0$). In addition the language can support special implicit functions such as the all different $(x_1, x_2, \ldots, x_n)$ constraint for assigning different values to the integer variables $x_1, x_2, \ldots, x_n$. The language consists of C++ procedures, although the recent trend has been to provide higher level languages such as OPL. Other commercial CP software packages include ILOG Solver (ILOG Inc., 1999c), CHIP (Dincbas et al., 1988), ECLiPSe (Wallace et al., 1997), and Pro-

log IV. Rather than relying on traditional mathematical programming methods, CP relies on a tree search using implicit enumeration. The tree is normally enumerated with a depth first search in which the lower bound is given by partial solutions, and the upper bound by the best feasible solution. At each of the nodes in the tree search, constraint propagation is performed through domain reduction of the variables. This involves for instance the reduction of bounds in the case of continuous variables, and/or domains in the case of discrete variables. The former uses procedures for tightening bounds for linear and monotonic functions, while the latter is performed either by inference techniques, or by special procedures. A good example is the "edge-finding" method for jobshop scheduling. Figure 4 presents a simple example of such a method to resolve a disjunction about the relative processing of jobs i and j.

A General Disjunctive Model for the Integration of Planning and Scheduling

In the past, planning and scheduling models have largely been solved separately due to the complexity associated with including and solving both levels of decision making in one model. Only very recently have simultaneous planning and scheduling models emerged (e.g. Papageorgiou and Pantelides, 1996a; Birewar and Grossmann, 1990). While the advances have shown progress towards integration of planning and scheduling, these problems remain in general intractable. This is due to the size of the resulting problem, and the mismatch of the time scales in planning and scheduling. This indicates that there is a need to derive efficient models and algorithms for integrated planning and scheduling. In this section we present a model that reflects the hierarchy of decisions that can be potentially exploited for an efficient solution.

From the review in the previous section, it can be concluded that LP and MILP methods, which are extensively used in planning and scheduling, have become quite powerful. In addition, NLP methods are able to tackle increasingly larger problems and are being advanced by rigorous global optimization algorithms. Together, these developments facilitate faster solution of MINLPs. A new exciting direction is logic based optimization methods, such as Generalized Disjunctive Programming and Constraint Programming, which promise to facilitate problem formulation and improve the solution efficiency and robustness. In order to illustrate the use of logic based methods, we present in this section a general GDP model that also has the important feature of integrating planning and scheduling for process networks in a single model. As will be seen the model gives rise to a generalized disjunctive program that involves embedded disjunctions that reflect the hierarchical nature of decisions involved in the integration problem.

We use a discrete time representation for the planning and the scheduling time domains. Also, we assume that the scheduling model corresponds to the State Task Network (STN) (Kondili et al., 1993). Consider optimizing a given STN superstructure over a time horizon, H. Such a superstructure consists of a set of units, J, capable of performing a set of tasks, I. Feeds, intermediates and products are represented by the set of states, S. In order to integrate both planning and scheduling into the optimization model, H is divided into a number of planning periods, $t = 1, \ldots, T$, and a number of scheduling periods, $k = 1, \ldots, K$. The length of a planning period is typically in the order of weeks or a few months, while the length of a scheduling period is typically in the order of hours. We define the set $Int(t, k)$ to denote which of the scheduling periods, k, belong to planning period t. The complete set, parameter and variable definitions are as follows:

Sets:

S	set of states (feeds, intermediates, products)
I	set of tasks
J	set of units
T	set of time periods in the planning horizon
K	set of time periods in the scheduling horizon
$Int(t,k)$	set of scheduling time periods k belonging to planning time period t

Indices:

s	state in set S
i	task in set I
j	unit in set J
t	time period in set T
k	time period in set K

Parameters:

α_{jt}	variable expansion cost for unit j in period t
β_{jt}	fixed expansion cost for unit j in period t
γ_{jt}	fixed operating cost for unit j in period t
c^p_{st}	cost associated with state s in planning time period t
c^s_{sk}	cost associated with state s in scheduling time period k
c^r_{ijk}	cost associated with resource usage for task i on unit j in time period k
η_{ij}	fixed resource cost for task i on unit j
δ_{ij}	variable resource cost for task i on unit j
t^d_{ij}	delay time associated with task i on unit j

Variables:

Binary decision variables:

y_j	selection of investment in unit j
w_{jt}	operation of unit j in period t
z_{jt}	capacity expansion of unit j in period t
v_{ijk}	task i is performed on unit j in period k

Continuous decision variables:

Q_{jt}	capacity of unit j in period t
QE_{jt}	capacity expansion of unit j in period t
x_t	state variables in period t

$$\min Z = \sum_t \left[\sum_j (CO_{jt} + CE_{jt}) + \sum_s c^p_{st} x^p_{st} + \sum_{k \in Int(t,k)} \sum_i \sum_j c^r_{ijk} R_{ijk} + \sum_{k \in Int(t,k)} \sum_s c^s_{sk} x^s_{sk} \right] \quad (1)$$

subject to

$$g_t(x_t, x_{t-1}) \le a \qquad \forall t \quad (2)$$

$$f_{sk}(x^s_{sk}, x^s_{s,k-1}, x^s_{s,k-t^d}) \le b \qquad \forall s,k \quad (3)$$

$$\left[\begin{array}{c} y_j \\ \left[\begin{array}{c} w_{jt} \\ h_{jt}(Q_{jt}, x_t, x_{t-1}) \le d \ (5) \\ CO_{jt} = \gamma_{jt} \ (6) \\ \left[\begin{array}{c} z_{jt} \\ Q_{jt} = Q_{j,t-1} + QE_{jt} \ (9) \\ CE_{jt} = \alpha_{jt} QE_{jt} + \beta_{jt} \ (10) \end{array} \right] \vee \left[\begin{array}{c} -z_{jt} \\ Q_{jt} = Q_{j,t-1} \ (11) \\ CE_{jt} = 0 \ (12) \end{array} \right] \\ \left[\begin{array}{c} v_{ijk} \\ 0 < B_{ijk} \le Q_{jt} \ (13) \\ R_{ijk} = \eta_{ij} + \delta_{ij} B_{ijk} \ (14) \end{array} \right] \vee \left[\begin{array}{c} -v_{ijk} \\ B_{ijk} = 0 \ (15) \\ R_{ijk} = 0 \ (16) \end{array} \right] \forall k \in Int(t,k) \end{array} \right] \vee \left[\begin{array}{c} -w_{jt} \\ D^{jt} x_t = 0 \ (7) \\ CO_{jt} = 0 \ (8) \end{array} \right] \Bigg\} \forall t \\ \vee \left[\begin{array}{c} -y_j \\ D^{jt} x_t = 0 \ (4) \\ \forall t \end{array} \right] \forall j \end{array} \right]$$

$$y_j \Rightarrow \bigvee_{t=1}^{T} w_{jt} \quad \forall j,t \quad (17), \qquad w_{jt} \Rightarrow y_j \quad \forall j,t \quad (18)$$

$$w_{jt} \Rightarrow \bigvee_{t=1}^{t} z_{jt} \quad \forall j,t \quad (19), \qquad z_{jt} \Rightarrow w_{jt} \quad \forall j,t \quad (20)$$

$$w_{jt} \Rightarrow \bigvee_{i,k \in Int(t,k)}^{I} v_{ijk} \quad \forall j,t \quad (21), \qquad v_{ijk} \Rightarrow w_{jt} \quad \forall i,j,k \in Int(t,k) \quad (22)$$

$$\Omega_1(y) = True \quad (23), \qquad \Omega_2(v) = True \quad (24)$$

$$CE, CO, Q, QE, x, R, B \ge 0 \quad (25), \qquad y, w, z, v \in \{True, False\} \quad (26)$$

Figure 5: GDP model.

x^p_{st}	subset of state variables for state s in planning time period t
CO_{jt}	operating cost of unit j in period t
CE_{jt}	expansion cost of unit j in period t
x^s_{sk}	subset of state variables for state s in scheduling time period k
R_{ijk}	resource usage for task i on unit j in time period k
B_{ijk}	batch size for task i on unit j in period k

Based on the above definitions, the GDP model is as shown in Figure 5.

The objective (1) is to minimize costs over the whole time horizon, and includes operating costs, expansion costs, and costs associated with states over the planning period, as well as resource costs and costs associated with states over the scheduling period. Sales are included by assigning negative values to the appropriate cost coefficients. Global constraints valid for a particular planning period, such as mass balances over mixers, are represented by (2), while global constraints valid for a particular scheduling period, such as inventory constraints, are represented by (3). Note that both (2) and (3) may generally involve "pass-on" variables from the previous period, giving rise to linking constraints. In addition, the global scheduling constraints (3) may generally also involve a scheduling time delay, t^d, due to processing times, clean-up times, and changeover times.

Constraints (5–16) are grouped into a set of nested disjunctive constraints for each unit j. The outer disjunction represents the decision to include unit j in the design or not, which is a strategic planning decision. If unit j is included in the design, ($y_j = True$), then the set of constraints on the left hand side of the disjunction is applied, otherwise ($y_j = False$), a subset of state variables associated with unit j are set to zero for all periods through the matrix D^{jt} in (4). The middle disjunction represents the decision to operate unit j in planning period t or not and is only applied if $y_j = True$. If the unit j operates in period t, ($w_{jt} = True$), which can be interpreted as either an operational or tactical planning decision, then constraints (5) and (6), as well as the two remaining disjunctions representing expansion and scheduling decisions, are applied. (5) represents constraints that are valid for a given unit j in a particular planning period t, such as unit input-output relationships, while the operating cost of unit j in planning period t is calculated in

(6). If unit j does not operate in period t, $(w_{jt} = False)$, a subset of state variables and the operating cost associated with unit j are set to zero for period t through (7) and (8), respectively.

The two inner disjunctions are only applied if $w_{jt} = True$, and of these the first represents the decision to expand unit j in planning period t or not. If unit j is expanded in period t, $(z_{jt} = True)$, which is also a planning decision, constraints (9) and (10) are applied. (9) states that the capacity at the current period equals the capacity at the previous period plus the capacity expansion, while the expansion cost is calculated in (10). If the decision is not to expand unit j in period t $(z_{jt} = False)$, the capacity remains the same as in the previous period, and the expansion cost is set to zero (see (11) and (12)).

Unit specific scheduling decisions are represented by the second inner disjunction. As pointed out in the previous section there exists no real generalization of scheduling models. We therefore focus on the ideas from STN scheduling first proposed by Kondili et al. (1993), since this formulation can be applied to arbitrary network structures. Note that this inner disjunction is only applied for scheduling periods k within the planning period t, as denoted by the set $Int(t, k)$. This disjunction states that if task i is started on unit j in scheduling period k, $(v_{ijk} = True)$, then the batch size is limited by the unit capacity in (13) and the resource usage is calculated in (14). If task i is not started on unit j in period k, $(v_{ijk} = False)$, the starting batch size and resource usage are set to zero in (15) and (16) respectively.

Constraints (17) through (24) are logic propositions representing logical relationships between the discrete variables. (17) states that the inclusion of unit j in the design implies that it must be operated in at least one period t, while (18) states the converse, i.e. that operation of unit j in any period t implies the inclusion of unit j in the design. Similarly, constraint (19) states that operation of unit j in period t implies that it must have been expanded at least once in a previous period, while (20) states the converse that expansion of unit j in period t implies that it will also be operated in that period. Constraint (21) states that the operation of unit j in planning period t implies that at least one task i must be started on unit j in a scheduling period k belonging to planning period t. If a task i starts on unit j in a scheduling period k belonging to t, then unit j must be operated in period t as denoted by (22). Constraint set (23) represents logic propositions relating the discrete design variables, y, for the topology of the network (which combinations are permitted). The relationships among the discrete scheduling variables (v), for example the condition that starting one task on unit j in period k implies that no other task can be started on unit j until task i is finished, are represented by the constraint set (24). Finally, the domains of the variables are given in (25) and (26).

The above represents a conceptual model that integrates planning and scheduling decisions within one single formulation. One advantage of this GDP model is that special structures are revealed, for example the clear hierarchy of decisions from design, operation, and expansion of units to assignment of units and sequencing of tasks. This facilitates the development of tailored algorithms using techniques such as decomposition, as will be discussed in the next section. Furthermore, a GDP model allows the application of specialized logic-based methods that have the effect of reducing non-convexities, and yielding tighter relaxations and ultimately faster solutions. It is also important to note that by fixing some Boolean variables and eliminating subsets of disjunctions, the proposed model can easily be shown to reduce to specific forms of planning or scheduling problems.

Solution Strategies

While moderately sized planning and scheduling models as presented in sections 2 and 4 can be solved with the mathematical programming methods as discussed in section 3, larger problem instances, which are these days often required for accurate representation of the problem characteristics, require some type of decomposition, aggregation and/or the use of heuristics for their solution. In this section we review some of these approaches that are applicable to large-scale mixed integer linear or nonlinear problems in addition to the methods mentioned in Section 3.

Decomposition

When choosing a decomposition method it is important to consider how to exploit the structure of the model most efficiently and also to choose a degree of decomposition that allows solution in reasonable time while still finding an optimal or near-optimal solution. Several decomposition schemes have been proposed in the literature. Benders decomposition (Benders, 1962) and dual decomposition or Lagrangean relaxation (see e.g. Fisher, 1981) exploit the primal and dual structures of the model respectively. Cross decomposition (see e.g. Van Roy, 1983) exploits both the primal and dual structures and is applicable to models where both the primal and dual subproblems are easy to solve. Bilevel decomposition (e.g. Iyer and Grossmann, 1998a) exploits the structure of models that include different hierarchical levels, such as the hierarchy from design, to planning, to scheduling. Ruszczynski (1997) gives a comprehensive review on decomposition methods for stochastic problems, including cutting plane methods, augmented Lagrangean decomposition, splitting methods and nested decomposition. Below, we discuss only Lagrangean relaxation and bilevel decomposition in further detail, since an in depth discus-

sion of all decomposition methods is beyond the scope of this paper. The discussion on bilevel decomposition is motivated by its relevance to an example presented in the next section as well as to combined planning and scheduling models. The discussion on Lagrangean relaxation is motivated by its wide applicability to large-scale optimization models and its ease of implementation in practice.

Bilevel decomposition. One approach to exploit the hierarchical structure of combined design, planning and/or scheduling models is to decompose the model into an upper level problem at the higher hierarchical level, and a lower level problem at a lower hierarchical level. Iyer and Grossmann (1998a) proposed such a bilevel decomposition algorithm for an MILP design and planning problem, where the upper level involves mainly design decisions while the lower level involves mainly planning decisions. Van den Heever and Grossmann (1999) expanded this approach to MINLPs through the use of GDP. Consider an original model (P) where superscript d denotes design variables and superscript p denotes planning variables.

$$\min f(x^d, y^d, x^p, y^p) \tag{P}$$

subject to

$$h(x^d, y^d, x^p, y^p) \leq 0$$
$$x \in \Re, \quad y \in \{0,1\}$$

To derive the upper level design problem (DP), all the discrete planning variables are relaxed. This results in a much smaller number of nodes in the branch and bound search facilitating a faster solution. Also, some of the constraints and/or variables may be aggregated at this level, indicated by Λ.

$$\min f(x^d, y^d, x^p, y^p) \tag{DP}$$

subject to

$$\Lambda h(x^d, y^d, x^p, y^p) \leq 0$$
$$0 \leq y^p \leq 1$$
$$x \in \Re, \quad y^p \in \Re, \quad y^d \in \{0,1\}$$

After (DP) is solved, the discrete design variables are fixed (indicated by the bar on y_d) and the lower level planning problem (PP) is solved for the fixed design.

$$\min f(x^d, \bar{y}^d, x^p, y^p) \tag{PP}$$

subject to

$$h(x^d, \bar{y}^d, x^p, y^p) \leq 0$$
$$x \in \Re, \quad y^p \in \{0,1\}$$

Subproblems (DP) and (PP) are solved iteratively and design and integer cuts are added at each iteration to ensure an optimal solution. Note that even though both (DP) and (PP) are in a reduced space, both consider the design and planning model as a whole. A further benefit of this approach is that it significantly reduces the computational effort compared to solving the combined problem as a whole, while still guaranteeing the optimal solution to the original combined model in the convex case. Papageorgiou and Pantelides (1996b) proposed a similar decomposition approach for combined campaign planning and scheduling of multipurpose batch/semicontinuous plants. In their work, the upper level problem concerns mainly campaign planning decisions while the scheduling decisions are aggregated, and the lower level problem is solved with some of the campaign planning variables fixed. Again both levels consider the problem as a whole. In the experience of the authors, the bilevel decomposition approach works particularly well for large-scale industrial applications over a long time horizon, especially when combined with the aggregation of time periods as discussed in Example 1.

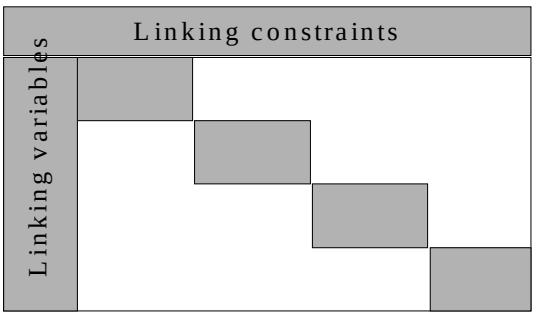


Figure 6: Block diagonal structure.

Lagrangean relaxation. This is an approach that is often applied to models with a block diagonal structure. In such models, distinct blocks of variables and constraints can be identified that are linked with a few "linking" constraints and variables (see Figure 6). Some applications include scenario decomposition for planning under uncertainty (Carøe and Schultz, 1999), unit commitment in power plants (Nowak and Römisch, 1998), midterm production planning (Gupta and Maranas, 1999), oilfield investment planning (Van den Heever et al., 2000) and combined transportation and scheduling (Equi et al., 1997), to name but a few.

Consider a model (L) that has been partitioned into blocks of constraints $p = 1, \ldots, P$ where the blocks are linked by a constraint set h:

$$\min \sum_p f_p(x_p) \tag{L}$$

subject to

$$g_p(x_p) \leq 0 \quad \forall p$$
$$h(x_1, \ldots, x_p) \leq b$$
$$x \in X$$

The basic idea behind Lagrangean relaxation as applied to the decomposition of block diagonal structures, is to dualize the linking constraint set, h, by removing it and replacing it with a penalty in the objective function involving the associated Lagrangean multipliers, λ, as seen in model (LR):

$$\min_x \sum_p f_p(x_p) + \lambda(h(x_1, \ldots, x_p) - b) \qquad \text{(LR)}$$

subject to

$$g_p(x_p) \leq 0 \quad \forall p$$
$$x \in X$$

Model (LR) is now decomposable into P subproblems and, for any choice of λ, also yields a lower bound to the optimal solution of (L) if the constraints are convex. The case where variables link the blocks can be dealt with by introducing duplicates for each linking variable, setting the duplicates equal, and dualizing this equality constraint. This is referred to as Lagrangean decomposition (Guignard and Kim, 1987). Obtaining the tightest lower bound to (L) requires the solution of the Lagrangean dual problem (LD):

$$\max_\lambda \min_x \sum_p f_p(x_p) + \lambda(h(x_1, \ldots, x_p) - b) \qquad \text{(LD)}$$

subject to

$$g_p(x_p) \leq 0 \quad \forall p$$
$$x \in X$$

If all the constraints are convex and all the variables are continuous, the optimum of (LD) will equal the optimum of (L). However, a duality gap might exist in the presence of integer variables or other non-convexities, which means that the optimal solution to the dual problem will be strictly less than the true optimum of (L). Guignard (1995) and Bazaraa et al. (1994) give comprehensive graphical interpretations of the duality gap in the case of integer variables and non-convex constraints respectively. Solving (LD) can be difficult to implement and time consuming, although Fisher (1981) reports on some algorithms for this purpose. A code for solving the dual was developed by Kiwiel (1993), but this code is not widely available to the best of our knowledge. Solving the dual to optimality is therefore often circumvented by using an iterative heuristic approach where (LR) is solved to generate lower bounds to (L) and a heuristic method is used to generate feasible solutions to (L) which are also upper bounds. λ is updated at each iteration with some updating rule, for example a subgradient method (see e.g. Fisher, 1981). This decomposition method reduces the computational effort by solving several subproblems instead of the original problem, and the associated algorithms lend themselves to parallelization to reduce the computational effort even more. For a thorough background on the application of Lagrangean relaxation, we refer the reader to Guignard (1995) and Fisher (1981, 1985).

Aggregation

For some models, decomposition alone is not enough to obtain a good solution in reasonable time, and some form of aggregation is required to further reduce the model size. Rogers et al. (1991) give a good review on the use of aggregation/disaggregation in optimization. These authors define the major components of this framework, namely aggregation analysis, disaggregation analysis and error analysis. The first component involves determining which elements of the model to combine into a single element and how to define the single element, while the second component conversely involves deriving a more refined model from the aggregate one. Error analysis determines the error introduced by aggregation and disaggregation. These three components can be addressed sequentially or iteratively to reduce the computational effort of solving the original problem, with the iterative approach aiming at decreasing the error at each iteration.

It should be noted that the solution to the aggregate formulation is not necessarily feasible for the disaggregate case. However, for certain models it may be possible to formulate the aggregation in such a way as to yield a strict bound to the original problem, and to guarantee feasibility for the disaggregate level, as shown by Iyer et al. (1998) for the aggregation of oil wells for oil production planning One approach to reduce the number of constraints is to linearly combine some of them into a surrogate constraint where the aggregation coefficients are modified iteratively (see e.g. Ermoliev et al., 1997). Wilkinson et al. (1996) use a constraint aggregation approach to solve a large-scale production and distribution planning problem for multiple production sites. In their work an upper level aggregate model is solved to set production targets and also yield a strict upper bound to the original problem, after which the detailed scheduling can be optimized for each site individually with fixed targets thus decreasing the computational effort significantly. Wilkinson (1996) proposed aggregate formulations for large-scale process scheduling problems using ideas of approximation of difference equations, as well as decomposition approaches for solving these models. In the case of multiperiod models, an approach that works well is to aggregate the time periods. This is especially true when the model involves two hierarchical

time levels, such as combined design and planning or combined planning and scheduling. Van den Heever and Grossmann (2000) combined the bilevel decomposition approach mentioned above with the aggregation of time periods by aggregating time in the upper level problem with subsequent disaggregation in the lower level planning problem. An additional subproblem is solved after each iteration to determine the best new aggregation scheme (which periods should be grouped together) and information from the aggregation subproblem is used at each iteration to eliminate variables in the lower level problem. It was found that the error introduced by the aggregation of the time periods was very small, mainly due to the optimal aggregation subproblem. Other aggregation schemes include the aggregation of products into families of similar products for the scheduling of multiproduct plants (Kondili et al., 1993). Where uncertainty is incorporated through a scenario-based model, scenario aggregation can speed up the solution time significantly. The scenario aggregation approach was applied to a mixed-integer linear multiproduct production planning problem by Jorsten and Leisten (1994) who exploited the coupling between continuous and integer planning variables to allow application of the scenario-aggregation algorithm originally proposed by Rockafeller and Wets (1991) for continuous models.

Apart from decomposition and aggregation techniques, some other heuristic approaches address the solution of large-scale planning and scheduling problems. One such heuristic is a capacity shifting heuristic presented by Ahmed and Sahinidis (2000) for a class of process planning problems. These authors show that the error of their heuristic algorithm vanishes asymptotically as the problem size increases. This is a very nice result, considering that the solution time increases exponentially with the number of time periods for an exact solution algorithm.

Examples

In this section we present three examples that illustrate some of the main points covered in this paper. Example 1 deals with a planning problem that gives rise to a large-scale multiperiod MINLP model, and that requires the use of a decomposition/aggregation strategy. Example 2 describes an MILP scheduling model for steel manufacturing that is also tackled through a special decomposition approach. Finally, example 3 describes a hybrid CP/MILP model for a parallel scheduling problem, which demonstrates the advantage of a combined approach as opposed to pure CP or MIP.

Example 1. Hydrocarbon Field Infrastructure Planning

The operation and investment planning involved in the design of hydrocarbon field infrastructures is a challenging problem that involves several complexities such as a long time horizon, nonlinear reservoir behavior, and complex fiscal rules leading to a multiperiod MINLP model with several discrete and continuous variables. In this example (for details see Van den Heever and Grossmann, 2000) we consider the design, planning and scheduling of an offshore oilfield infrastructure over a planning horizon of 6 years divided into 24 quarterly periods where decisions need to be made. The infrastructure under consideration consists of one Production Platform (PP), 2 Well Platforms (WP), 25 wells and connecting pipelines (see Figure 7). Each oilfield (F) consists of a number of reservoirs (R), while each reservoir in turn contains a number of potential locations for wells (W) to be drilled. Design decisions involve the capacities of the PPs and WPs, as well as decisions regarding which WPs to install over the whole operating horizon. Planning decisions involve the production profiles in each period, as well as decisions regarding when to install PPs and WPs included in the design, while scheduling decisions involve the selection and timing of drilling of the wells. This leads to an MINLP model with 9744 constraints, 5953 continuous variables, and 700 0-1 variables.

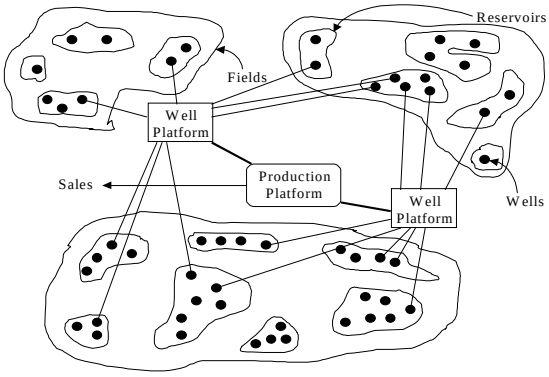


Figure 7: Configuration of fields, well platforms and production platforms.

An attempt to solve this model with a commercial package such as GAMS (Brooke et al., 1992) (using DICOPT (Viswanathan and Grossmann, 1990) with CPLEX 6.6 (ILOG Inc., 2000) for the MILPs and CONOPT2 (Drud, 1992) for the NLPs on an HP9000/C110 workstation), results in a solution time of 19386 CPU seconds. To overcome this long solution time, Van den Heever and Grossmann (2000) developed an iterative aggregation/disaggregation algorithm which solved the model in 1423 CPU seconds. This algorithm combines the concepts of bilevel decomposition, time aggregation and logic-based methods. The original design and planning problem is decomposed into an upper level design problem and a lower level planning problem. Both

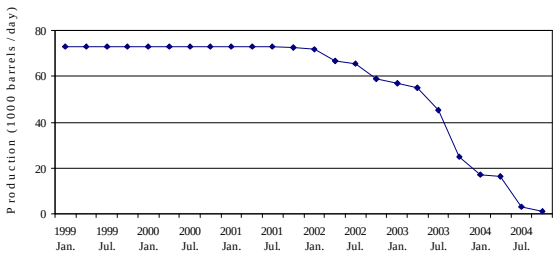


Figure 8: Production profile over 6 year horizon.

Item		Period Invested
PP		Jan. 1999
WP1		Jan. 1999
Reservoir	**Well**	
2	4	Jan. 1999
3	1	Jan. 1999
5	3	Jan. 1999
4	2	Apr. 1999
7	1	Jul. 1999
6	2	Oct. 1999
1	2	Jan. 2000
9	2	Jan. 2000
10	1	Jan. 2000

Table 1: The optimal investment plan.

subproblems are formulated as disjunctive models. The upper level design problem is solved in aggregate time, after which a design is fixed, time periods are disaggregated and the lower level planning problems is solved. This result is then used to determine a new time aggregation through a dynamic programming subproblem, integer cuts are added to the design problem, aggregation parameters are updated, and the iteration is repeated until the termination criteria are reached. Thus the application of combined decomposition and aggregation leads to an order of magnitude reduction in solution time, while the same optimal net present value of $68 million is found as with DICOPT. For this specific model, the large decrease in computational effort is mainly due to the aggregation/decomposition, while the disjunctive programming formulation contributed mainly towards reducing non-convexities due to zero flows and to the clarity of representation. However, for different planning models the disjunctive programming approach may reduce the computational effort significantly in addition to the benefits mentioned here, as shown by Van den Heever and Grossmann (1999) for the case of process network design and planning and the retrofit of batch plants.

Figure 8 shows the total oil production over the 6 year horizon, while Table 1 shows the optimal investment plan obtained. Note that only 9 of the 25 wells were chosen in the end. This solution resulted in savings in the order of millions of dollars compared to the heuristic method used in the oilfield industry that specify almost all the wells being drilled.

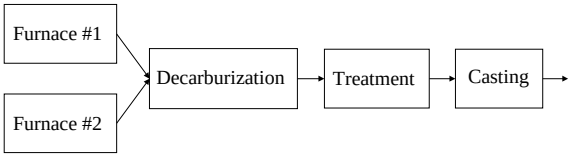


Figure 9: Processing steps in steel manufacturing.

In Van den Heever et al. (2000), the concept of hydrocarbon field infrastructure planning was expanded to include complex fiscal rules such as royalties, tariffs and taxes. This resulted in a model for which no solution could be found by GAMS in more than 5 days. To address this problem, a heuristic solution procedure based on Lagrangean decomposition was proposed that produces several good solutions in a day. This method can potentially be parallelized and combined with the aggregation of time periods to speed up the solution even more.

Example 2. An MILP Approach to Steel Manufacturing

In this section, an MILP approach to produce a production schedule for a steel-making process is discussed. As seen in Figure 9 the steel-making process consists of two furnaces, where the melt steel is combined with scrap and thereafter taken to decarburization and ladle treatment units. Finally the melt steel is solidified in a continuous caster under strictly constrained conditions. There are several complicating factors in the problem such as sequence dependent setup times, maintenance of equipment and production time limitations. Furthermore, temperature and purity issues are critical to the production that also includes low-carbon steel grades. Other problems arising from the metal chemistry as well as plant geometrics are not directly considered in the model. The production steps are illustrated in Figure 9, where the possible paths of the products are shown.

A continuous time representation for modeling a large-scale scheduling problem is applied. The proposed decomposition strategy consists of first positioning the orders into blocks, each of which is optimized as a job-shop scheduling problem. Next the blocks are optimally scheduled as a flowshop problem and finally an LP and/or MILP method is used to properly account for setup times and to optimize the allocation of some parallel equipment. This type of approach allows some of the more complicated constraints to be either isolated in a subproblem or inserted as parameters between two

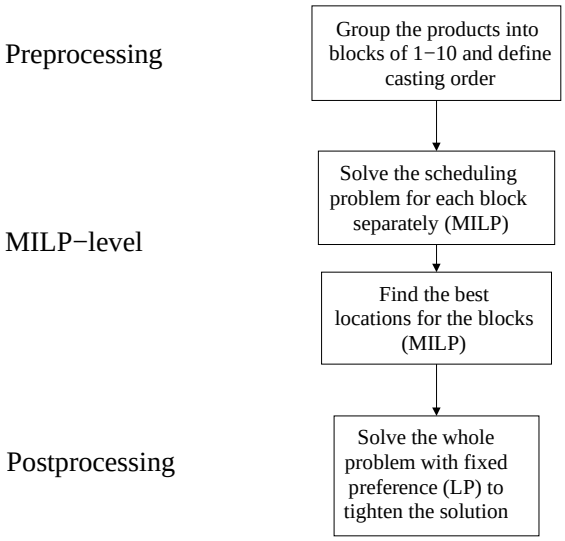


Figure 10: Solution strategy for steel scheduling.

solution steps. While this decomposition strategy is not guaranteed to yield the global optimal schedule it allows the solution of very large-scale problems. The strategy is illustrated in Figure 10.

The decomposition strategy is motivated by the fact that heats with similar product properties can be grouped into sequences in a preprocessing step. These sequences are then treated as independent blocks by first optimizing their internal production order and thereafter finding the optimal sequence between the blocks. In a postprocessing step, the heats are again treated as individual products with fixed ordering and the gaps, caused by the grouping, are closed by solving an LP-problem. The solution can furthermore be improved by solving a final MILP. This example forms an interesting approach where a problem is both decomposed and rejoined through optimization and thus one large MILP is replaced by a number of smaller and solvable subproblems. Also, the modularity of the procedure makes it possible to solve only parts of the problem when changes occur. Table 2 shows the results that were obtained from solving a one week problem containing 81 products. The optimization was performed on a Linux-platform using XPRESS-MP in GAMS.

The 81 product problem is not solvable with standard MILP methods, as seen from the failure in solving even the 10-product jobshop-example in reasonable time (the makespan is the best one obtained at 10,000 CPU seconds). The proposed strategy solves the complete problem in less than 20 CPU minutes. Even though the decomposition strategy is not expected to result into a global optimal solution, the maximum deviation of the makespan with respect to a theoretical optimum is only 2% and the makespan is reduced from one week to 5 days and 12 hours.

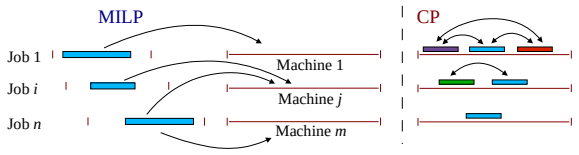


Figure 11: Scheduling of parallel machines.

Example 3. A Hybrid MILP/CP Approach for Parallel Scheduling

In this example, a strategy of decomposing a scheduling problem into a CP and MILP part is discussed. The basic idea is to combine the two methods such that their complementary strengths can be exploited. The problem is a single stage scheduling problem with parallel units reported by Jain and Grossmann (2000). In the decomposition strategy a relaxed MILP is solved at the master level and a feasibility subproblem is solved with CP. The relaxed MILP excludes the complicating constraints which in this case are the sequencing inequalities. Those are reformulated as a feasibility CP subproblem. The strategy consists of two main steps: the relaxed MILP is first solved to its global optimal solution to obtain a feasible assignment and thereafter a feasibility check is performed by solving a CP sequencing problem with fixed assignment (see Figure 11).

If the sequencing problem is infeasible, cuts are added to the next relaxed MILP to exclude the previous infeasible assignment. The procedure continues solving alternate CP and MILP problems until a feasible sequence is found. The communication between the solution steps is done through fixing assignment variables and generating integer cuts.

This strategy requires that the problem be decomposed into two subproblems of which the MILP (the assignment problem) provides a tight LP-relaxation and contains the objective function variables, and the CP (the sequencing problem) has no objective function variables and includes the constraints with poor relaxations. This ensures an efficient assignment of machines and that the first feasible sequence found is optimal. As can be seen in the following mathematical formulation, equivalence relations are also needed to join the variables in the two subproblems due to some structural differences. Here we will only present the main elements of the hybrid formulation. For more details, we refer to Jain and Grossmann (2000).

$$\min \sum_{i \in I} \sum_{m \in M} C_{im} x_{im} \tag{27}$$

Problem	Products	Groups	CPU-s	Makespan	Integer Gap
Proposed decomposition	81	19	1035.3	132h 6 min	2%
Conventional jobshop	10	1	>10,000	19h 7min	52%

Table 2: Comparison of results between proposed approached and conventional jobshop model.

Problem	1	2	3	4
Machines/Jobs	3/7	3/12	5/15	5/20
MILP	0.58	164.92	528.86	>40,000
CP	0.04	3.35	590.9	11666.4
HYBRID	0.49	5.27	0.56	35.64

Table 3: Computational results in parallel scheduling problem.

subject to

$$ts_i \geq r_i \quad \forall i \in I \tag{28}$$

$$ts_i \leq d_i - \sum_{m \in M} p_{im} x_{im} \quad \forall i \in I \tag{29}$$

$$\sum_{m \in M} x_{im} = 1 \quad \forall i \in I \tag{30}$$

$$\sum_{i \in I} x_{im} p_{im} \leq \max_i\{d_i\} - \min_i\{r_i\} \quad \forall m \in M \tag{31}$$

$$\text{if } (x_{im} = 1) \text{ then } (z_i = m) \quad \forall i \in I, m \in M \tag{32}$$

$$i.start \geq r_i \quad \forall i \in I \tag{33}$$

$$i.start \leq d_i - p_{z_i} \quad \forall i \in I \tag{34}$$

$$i.duration \leq p_{z_i} \quad \forall i \in I \tag{35}$$

$$i \text{ requires } t_{z_i} \quad \forall i \in I \tag{36}$$

$$ts_i \geq 0 \tag{37}$$

$$x_{im} \in \{0,1\} \quad \forall i \in I, m \in M \tag{38}$$

$$z_i \in M \quad \forall i \in I \tag{39}$$

$$i.start \in Z \quad \forall i \in I \tag{40}$$

$$i.duration \in Z \quad \forall i \in I \tag{41}$$

$$\sum_{i \in I} a^j_{im} x_{im} \leq \sum_{i \in I} a^j_{im} - 1 \quad \forall m \in M \tag{42}$$

The assignment MILP problem is defined by (27–31). The objective function (27) minimizes the processing costs for all jobs. The binary variable, x_{im}, equals one if job i is assigned to machine m, else it is zero. Constraint (28) and (29) ensure that processing of a job i starts after the release date and is completed before the duedate. Each job needs exactly one machine as is stated in (30) and the last MILP constraint (31) tightens the LP-relaxation. After solving the MILP problem the fixed assigments are transferred to the CP model using (30).

The sequencing CP problem, given in (33–36), is then solved separately for each machine. In the formulation, i is an activity (job), the start time of which is specified in (33) and (34) and duration in (35). It should be noted that it is possible to use a variable as an index in CP and p_{zi} refers to the processing time of job i in the assigned equipment. In constraint (36) the special construct requires enforces that job i needs a unary recourse from the set of resources t. If a machine cannot be scheduled, a cut of the form (42) is added to the next assignment MILP problem.

In Table 3, the problems are solved using complete MILP and CP formulations, as well as the hybrid model, with modified data given in Harjunkoski et al. (2000) where the original release dates, due dates and durations (Jain and Grossmann, 2000) have been arbitrarily changed and roughly multiplied by a factor of 10 to test the robustness of the method. The problems were solved on a Sun workstation with OPL Studio (ILOG Inc., 1999a) using CPLEX 6.5 and ILOG Solver (ILOG Inc., 1999c) and Scheduler (ILOG Inc., 1999b). In the following table the CPU times are listed for the four test problems.

In the hybrid formulation, most of the CPU time is consumed by the MILP. Even though CP overperforms the hybrid approach in the smallest problems the results clearly show that both CP and MILP suffer from combinatorial explosion, but the combination of these two methods performs very well even for the largest problem. It should be pointed out that the hybrid strategy does not compromise the global optimality.

Conclusions

This paper has presented an overview of planning and scheduling. It has been shown that these problems lead to discrete optimization models for which the associated mathematical programming problems correspond to integer programming problems, which can exhibit exponential behavior in their computation. Logic-based optimization techniques offer the potential of not only simplifying the formulations, but also decreasing the computational requirements. We have illustrated the use of logic based optimization as a modeling tool through a novel Generalized Disjunctive Program, that integrates planning and scheduling for process networks, and where the scheduling is represented with the STN model. We have also given a brief overview of decomposition strategies since these are essential in tackling large scale industrial problems. Finally, we have presented three examples, planning of oilfields, scheduling of steel manufacturing and scheduling of parallel machines, to illustrate the application of new techniques that are making possible the solution of problems that were essentially unsolvable a

few years ago.

While the integration of planning and scheduling remains a major challenge due to the potentially large size of the resulting optimization problem, another major challenge that has not been covered in this paper is the integration of planning and scheduling with control. This is essentially virgin territory in which very little work has been reported. The reader is referred for instance to the work by Bose and Pekny (2000), Perea et al. (2000), and Vargas-Villamil and Rivera (2000) who have addressed the incorporation of model predictive control in scheduling, and the dynamics and control of supply chains. At present, however, these works have addressed only specific applications due to the lack of a general framework, which possibly might be achieved through the use of hybrid systems (Kowalewski, 2001; Morari, 2001).

References

Ahmed, S. and N. V. Sahinidis, "Robust process planning under uncertainty," *Ind. Eng. Chem. Res.*, **37**, 1883–1892 (1998).

Ahmed, S. and N. V. Sahinidis, "Analytical investigations of the process planning problem," *Comput. Chem. Eng.*, **23**, 1605–1621 (2000).

Applequist, G., J. F. Pekny, and G. V. Reklaitis, "Risk and Uncertainty in Managing Chemical Manufacturing Supply Chains," *Comput. Chem. Eng.*, **24**(9-10), 2211–2222 (2000).

Balas, E., S. Ceria, and G. A. Cornuejols, "Lift-and-Project Cutting Plane Algorithm for Mixed 0-1 Programs," *Math. Prog.*, **58**, 295–324 (1993).

Balas, E., "Disjunctive Programming and a hierarchy of relaxations for discrete optimization problems," *SIAM J. Alg. Disc. Meth.*, **6**, 466–486 (1985).

Bazaraa, M. S., S. H. D., and S. C. M., *Nonlinear Programming*. John Wiley (1994).

Beaumont, N., "An Algorithm for Disjunctive Programs," *European Journal of Operational Research*, **48**, 362–371 (1991).

Benders, J. F., "Partitioning Procedures for Solving Mixed-variables Programming Problems," *Numerische Mathematik*, **4**, 238–252 (1962).

Birewar, D. B. and I. E. Grossmann, "Simultaneous Production Planning and Scheduling of Multiproduct Batch Plants," *Ind. Eng. Chem. Res.*, **29**, 570 (1990).

Bok, J.-K., I. E. Grossmann, and S. Park, "Supply chain optimization in continuous flexible process networks," *Ind. Eng. Chem. Res.*, **39**, 1279–1290 (2000).

Borchers, B. and J. E. Mitchell, "An Improved Branch and Bound Algorithm for Mixed Integer Nonlinear Programming," *Comput. Oper. Res.*, **21**, 359–367 (1994).

Bose, S. and J. F. Pekny, "A model predictive framework for planning and scheduling problems: a case study of consumer goods supply chain," *Comput. Chem. Eng.*, **24**, 329–335 (2000).

Brooke, A., D. Kendrick, and A. Meeraus, *GAMS—A User's Guide*. Scientific Press, Palo Alto (1992).

Carøe, C. C. and R. Schultz, "Dual Decomposition in Stochastic Integer Programming," *Operations Research Letters*, **24**(1-2), 37 (1999).

Chvatal, *Linear Programming*. Freeman (1983).

Clay, R. L. and I. E. Grossmann, "A Disaggregation algorithm for the optimization of stochastic planning models," *Comput. Chem. Eng.*, **21**, 751 (1997).

Crainic, T. G. and G. Laporte, "Planning models for freight transportation," *European Journal of Operational Research*, **97**, 409–438 (1997).

Dash Associates, *XPRESS-MP User Guide*. Dash Associates (1999).

Dimitriadis, A. D., N. Shah, and C. C. Pantelides, "RTN-Based Rolling Horizon Algorithms for Medium Term Scheduling of Multipurpose Plants," *Comput. Chem. Eng.*, **21**, S1061 (1997).

Dincbas, M., P. Van Hentenryck, H. Simonis, A. Aggoun, T. Graf, and F. Berthier, The constraint logic programming language CHIP, In *FGCS-88: Proceedings of International Conference on Fifth Generation Computer Systems*, pages 693–702, Tokyo (1988).

Ding-Mai and R. W. H. Sargent, A Combined SQP and Branch and Bound Algorithm for MINLP Optimization, Internal report, Centre for Process Systems Engineering, London (1992).

Drud, A., "CONOPT—A Large Scale GRG Code," *ORSA Journal on Computing*, **6** (1992).

Duran, M. A. and I. E. Grossmann, "An Outer-Approximation Algorithm for a Class of Mixed-integer Nonlinear Programs," *Math Prog.*, **36**, 307 (1986).

Equi, L., G. Gallo, S. Marziale, and A. Weintraub, "A combined transportation and scheduling problem," *European Journal of Operational Research*, **97**, 94–104 (1997).

Erengüç, S. S., N. C. Simpson, and A. J. Vakharia, "Integrated production/distribution planning in supply chains: An invited review," *European Journal of Operational Research*, **115**, 219–236 (1999).

Ermoliev, Y. M., A. V. Kryazhimskii, and A. Ruszcynski, "Constraint aggregation principle in convex optimization," *Math. Prog.*, **76**, 353–372 (1997).

Fisher, M. L., "The Lagrangean relaxation method for solving integer programming problems," *Management Science*, **27**(1), 1–17 (1981).

Fisher, M. L., "An applications oriented guide to lagrangean relaxation," *Interfaces*, **15**, 10–21 (1985).

Fletcher, R. and S. Leyffer, "Solving Mixed Integer Nonlinear Programs by Outer Approximation," *Math Prog.*, **66**, 327 (1994).

Fletcher, R., *Practical Methods of Optimization*. John Wiley & Sons, New York (1987).

Floudas, C. A., C. S. Adjiman, I. P. Androulakis, and C. D. Maranas, "A Global Optimization Method, (BB, for Process Design," *Comput. Chem. Eng.*, **20**, S419 (1996).

Floudas, C. A., Global Optimization in Design and Control of Chemical Process Systems, In *Preprints IFAC Symposium DYCOPS 5*, pages 167–176, Korfu (2000).

Fourer, R., D. M. Gay, and B. W. Kernighan, *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press, Belomont, CA (1992).

Geoffrion, A. M., "Generalized Benders Decomposition," *J. Optim. Theo. Appl.*, **10**(4), 237–260 (1972).

Goldberg, D. E., *Genetic Algorithms in Search, Optimisation and Machine Learning*. Addison-Wesley, Reading MA (1989).

Grossmann, I. E. and Z. Kravanja, Mixed-integer Nonlinear Programming: A Survey of Algorithms and Applications, In Biegler, Coleman, Conn, and Santosa, editors, *Large-Scale Optimization with Applications. Part II: Optimal Design and Control*, volume 93 of *The IMA Volumes in Mathematics and its Applications*, pages 73–100. Springer Verlag (1997).

Grossmann, I. E., J. Quesada, R. Raman, and V. Voudouris, Mixed Integer Optimization Techniques for the Design and Scheduling of Batch Processes, In Reklaitis, G. V., A. K. Sunol, D. W. T. Rippin, and O. Hortacsu, editors, *Batch Processing Systems Engineering*, pages 451–494. Springer-Verlag, Berlin (1996).

Grossmann, I. E., editor, *Global Optimization in Engineering Design*. Kluwer, Dordrecht (1996).

Guignard, M. and S. Kim, "Lagrangean decomposition: A model yielding stronger lagrangean bounds," *Math. Prog.*, **39**, 215–228 (1987).

Guignard, M., "agrangean relaxation: A short course," *Belgian Journal of OR: Special Issue Francoro*, **35**, 3 (1995).

Gupta, A. and C. D. Maranas, "A Hierarchical Lagrangean Relaxation Procedure for Solving Midterm Planning Problems," *Ind. Eng. Chem. Res.*, **38**, 1937–1947 (1999).

Gupta, O. K. and V. Ravindran, "Branch and Bound Experiments in Convex Nonlinear Integer Programming," *Management Science*, **31**(12), 1533–1546 (1985).

Han, S.-P., "Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems," *Math. Prog.*, **11**, 263–282 (1976).

Harjunkoski, I., V. Jain, and I. E. Grossmann, "Hybrid Mixed-Integer/Constraint Logic Programming Strategies for Solving Scheduling and Combinatorial Optimization Problems," *Comput. Chem. Eng.*, **24**, 337–343 (2000).

Hobbs, B. F., "Optimization methods for electric utility resources planning," *European Journal of Operational Research*, **83**, 1–20 (1995).

Hooker, J. N. and M. A. Osorio, "Mixed logic/linear programming," *Discrete Applied Mathematics* (1997).

Horst, R. and P. M. Pardalos, editors, *Handbook of Global Optimization*. Kluwer (1995).

Horst, R. and H. Tuy, *Global Optimization*. Springer-Verlag (1993).

IBM, "IBM Optimization Subroutine Library Guide and Reference," *IBM Systems Journal SC23-0519*, **31**(1) (1992).

Ierapetritou, M. and C. A. Floudas, "Effective continuous-time formulation for short-term scheduling," *Ind. Eng. Chem. Res.*, **37**, 4341–4359 (1998).

Ierapetritou, M. G., J. Avecedo, and E. N. Pistokopoulos, "An optimization approach for process engineering problems under uncertainty," *Comput. Chem. Eng.*, **20**(6-7), 703–709 (1996).

ILOG Inc., *ILOG OPL Studio 2.1. User's Manual*. ILOG Inc. (1999a).

ILOG Inc., *ILOG Scheduler 4.4. User's Manual*. ILOG Inc. (1999b).

ILOG Inc., *ILOG Solver 4.4. User's Manual*. ILOG Inc. (1999c).

ILOG Inc., *ILOG CPLEX 6.6. User's Manual*. ILOG Inc. (2000).

Iyer, R. R. and I. E. Grossmann, "A Bilevel Decomposition Algorithm for Longe-Range planning of process networks," *Ind. Eng. Chem. Res.*, **37**, 474 (1998a).

Iyer, R. R. and I. E. Grossmann, "Synthesis and operational planning of utility systems for multiperiod operation," *Comput. Chem. Eng.*, **22**, 979–993 (1998b).

Iyer, R. R., I. E. Grossmann, S. Vasantharajan, and A. S. Cullick, "Optimal planning and scheduling of offshore oil field infrastructure investment and operations," *Ind. Eng. Chem. Res.*, **37**, 1380 (1998).

Jain, V. and I. E. Grossmann, Algorithms for Hybrid MILP/CP Models for a Class of Optimization Problems, In *Presented at INFORMS, Paper SD32.1*, Salt Lake City (2000).

Jorsten, K. and R. Leisten, "Scenario aggregation in single-resource production planning models with uncertain demand," *Production Planning and Control*, **5**(3), 271–281 (1994).

Kirkpatrick, S., C. Gelatt, and M. Vecchi, "Optimization by Simulated Annealing," *Science*, **220**(4598), 671–680 (1983).

Kiwiel, K. C., *User's Guide for NOA 2.0/3.0: A Fortran package for convex nondifferentiable optimization*. Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland (1993).

Kocis, G. R. and I. E. Grossmann, "Relaxation Strategy for the Structural Optimization of Process Flow Sheets," *Ind. Eng. Chem. Res.*, **26**, 1869 (1987).

Kondili, E., P. C. C., and S. R. W. H., "A General Algorithm for Short-Term Scheduling of Batch Operations-I MILP Formulation," *Comput. Chem. Eng.*, **17**(2), 211–227 (1993).

Kowalewski, S., Hybrid Systems in Process Control: Challenges, Methods and Limits, In Rawlings, J. B., B. A. Ogunnaike, and J. W. Eaton, editors, *Chemical Process Control CPC VI*, Austin, TX. CACHE (2001).

Ku, H., D. Rajagopalan, and I. A. Karimi, "Scheduling in Batch Processes," *Chem. Eng. Prog.*, pages 35–45 (1987).

Lee, S. and I. E. Grossmann, Nonlinear Convex Hull Reformulations and Algorithms for Generalized Disjunctive Programming, Submitted for publication (1999).

Liu, M. L. and N. V. Sahinidis, "Optimization in Process planning under uncertainty," *Ind. Eng. Chem. Res.*, **35**, 4154–4165 (1996).

Marsten, R., S. M., J. Lustig, and D. Shanno, "Interior Point Methods for Linear Programming: Just Call Newton, Lagrange and Fiacco and McCormick!," *Interfaces*, **20**(4), 105–116 (1990).

Mazzola, J. B. and A. W. Neebe, "Lagrangian relaxation based solution procedures for a multiproduct capacitated facility location problem with choice of facility type," *European Journal of Operational Research*, **115**, 285–299 (1999).

McDonald, C. and I. A. Karimi, "Planning and Scheduling of Parallel Semicontinuous Processes 1. Production Planning," *Ind. Eng. Chem. Res.*, **36**, 2691 (1997).

Minoux, M., *Mathematical Programming: Theory and Algorithms*. John Wiley (1983).

Mockus, L. and G. V. Reklaitis, "Continuous time representation in batch/semi-continuous processes: randomized heuristic approach," *Comput. Chem. Eng.*, **20**, S1173–S1177 (1996).

Morari, M., Hybrid System Analysis and Control via Mixed Integer Optimization, In Rawlings, J. B., B. A. Ogunnaike, and J. W. Eaton, editors, *Chemical Process Control CPC VI*, Austin, TX. CACHE (2001).

Moro, L. F. L., A. C. Zanin, and J. M. Pinto, "A planning model for refinery diesel production," *Comput. Chem. Eng.*, **22**, S1039 (1998).

Murtagh, B. A. and M. A. Saunders, "Large-Scale Linearly Constrained Optimization," *Math. Prog.*, **14**, 41–72 (1978).

Murtagh, B. A. and M. A. Saunders, "A Projected Lagrangian Algorithm and its Implementation for Sparse Nonlinear Constraints," *Math. Prog. Study*, **16**, 84–117 (1982).

Nemhauser, G. L. and L. A. Wolsey, *Integer and Combinatorial Optimization*. Wiley-Interscience, New York (1988).

Nowak, M. P. and W. Römisch, Stochastic Lagrangean Relaxation applied to Power Scheduling in a Hydro-Thermal System under Uncertainty, Institut für Mathematik Humboldt-Universität Berlin, Preprint (1998).

Owen, S. H. and M. S. Daskin, "Strategic facility location: A review," *European Journal of Operational Research*, **111**, 423–447 (1998).

Pantelides, C. C., Unified frameworks for optimal process planning and scheduling, In Rippin, D. W. T., H. J. C., and J. F. Davis, editors, *Foundations of Computer Aided Process Operations*, pages 253–274. CACHE, Austin, TX (1994).

Papageorgiou, L. G. and C. C. Pantelides, "Optimal campaign planning/scheduling of multipurpose batch/semi-continuous plants, 1 Mathematical formulation," *Ind. Eng. Chem. Res.*, **35**, 488 (1996a).

Papageorgiou, L. G. and C. C. Pantelides, "Optimal campaign planning/scheduling of multipurpose batch/semi-continuous plants, 2 A mathematical decomposition approach," *Ind. Eng. Chem. Res.*, **35**, 510 (1996b).

Pekny, J. F. and G. V. Reklaitis, "Towards the Convergence of Theory and Practice: A Technology Guide for Scheduling/Planning Methodology," *AIChE Symp. Ser.*, **94**(320), 91–111 (1998).

Pekny, J. F. and M. G. Zentner, Learning to solve process scheduling problems: the role of rigorous knowledge acquisition frameworks, In Rippin, D. W. T., H. J. C., and J. F. Davis, editors, *Foundations of Computer Aided Process Operations*, pages 275–309. CACHE, Austin, TX (1994).

Perea, E., I. E. Grossmann, and E. Ydstie, "Towards the integration of dynamics and control for supply chain management," *Comput. Chem. Eng.*, **24**, 1143–1150 (2000).

Petkov, S. B. and C. D. Maranas, "Multiperiod planning and scheduling of multiproduct batch plants under demand uncertainty," *Ind. Eng. Chem. Res.*, **36**, 4864–4881 (1997).

Pinto, J. and I. E. Grossmann, "Assignment and Sequencing Models for the Scheduling of Chemical Processes," *Annals of Operations Research*, **81**, 433–466 (1998).

Powell, M. J. D., A Fast Algorithm for Nonlinearly Constrained Optimization Calculations, In Watson, G. A., editor, *Numerical Analysis, Dundee 1977*, volume 630 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin (1978).

Puget, J.-F., A C++ Implementation of CLP (Ilog Solver Collected Papers), Technical report, Ilog, France (1994).

Quesada, I. and I. E. Grossmann, "An LP/NLP Based Branch and Bound Algorithm for Convex MINLP Optimization Problems," *Comput. Chem. Eng.*, **16**, 937–947 (1992).

Raman, R. and I. E. Grossmann, "Relation Between MILP Modelling and Logical Inference for Chemical Process Synthesis," *Comput. Chem. Eng.*, **15**, 73 (1991).

Raman, R. and I. E. Grossmann, "Modelling and Computational Techniques for Logic Based Integer Programming," *Comput. Chem. Eng.*, **18**, 56 (1994).

Reklaitis, G. V., Perspectives on scheduling and planning of process operations, In *Presented at the Fourth international symposium on process systems engineering*, Montebello, Canada (1991).

Reklaitis, G. V., Overview of scheduling and planning of batch process operations, In *NATO Advanced Study Institute—Batch Process Systems Engineering*, Antalya, Turkey (1992).

Rippin, D. W. T., "atch process systems engineering: a retrospective and prospective review," *Comput. Chem. Eng.*, **17**, S1–S13 (1993).

Rockafeller, R. T. and R. J.-B. Wets, "Scenarios and policy aggregation in optimization under uncertainty," *Mathematics of Operations Research*, **16**(1), 119–147 (1991).

Rogers, D. F., R. D. Plante, R. T. Wong, and J. R. Evans, "Aggregation and disaggregation techniques and methodology in optimization," *Oper. Res.*, **39**, 553 (1991).

Ruszczynski, A., "Decomposition methods in stochastic programming," *Math. Prog.*, **79**, 333–353 (1997).

Sahinidis, N. V., I. E. Grossmann, R. E. Fornari, and M. Chathrathi, "Optimization Model for Long Range Planning in the Chemical Industry," *Comput. Chem. Eng.*, **13**(9), 1049–1063 (1989).

Sahinids, N. V. and H. S. Ryoo, "Global Optimization of Nonconvex NLP's and MINLP's with Applications in Process Design," *Comput. Chem. Eng.*, **19**, 551 (1995).

Saigal, *Linear Programming: A Modern Integrated Analysis*. Kluwer Academic Publishers (1995).

Sand, G., S. Engell, A. Markert, R. Schultz, and C. Schulz, "Approximation of an ideal online scheduler for a multiproduct batch plant," *Comput. Chem. Eng.*, **24**(2-7), 361–367 (2000).

Schittkowski, K., "The Nonlinear Programming Method of Wilson, Han, and Powell with an Augmented Lagrangian Type Line Search Function," *Numer. Math.*, **38**, 83–114 (1981).

Shah, N. and C. C. Pantelides, "Optimal Long-Term Campaign Planning and Design of Batch Operations," *Ind. Eng. Chem. Res.*, **30**, 2308–2321 (1991).

Shah, N., C. C. Pantelides, and R. W. H. Sargent, "A general algorithm for short-term scheduling of batch operations II. Computational issues," *Comput. Chem. Eng.*, **17**(2), 229–244 (1993).

Shah, N., "Single- and multisite planning and scheduling: Current status and future challenges," *AIChE Symp. Ser.*, **94**(320), 75 (1998).

Smith, E. and C. Pantelides, Global Optimisation of General Process Models, In Grossmann, I. E., editor, *Global Optimization in Engineering Design*, pages 355–386. Kluwer, Dordrecht (1996).

Stubbs, R. A. and S. Mehrotra, A Branch and Cut Method for 0-1 Mixed Convex Programming, In *Presented at INFORMS Meeting* (1996).

Tawarmalani, M. and N. V. Sahinidis, Global Optimization of Mixed Integer Nonlinear Programs: A Theoretical and Computational Study, Submitted for publication (2000).

Türkay, M. and I. E. Grossmann, "A Logic Based Outer-Approximation Algorithm for MINLP Optimization of Process Flowsheets," *Comput. Chem. Eng.*, **20**, 959–978 (1996).

Van den Heever, S. A. and I. E. Grossmann, "Disjunctive multiperiod optimization methods for design and planning of chemical process systems," *Comput. Chem. Eng.*, **23**, 1075–1095 (1999).

Van den Heever, S. A. and I. E. Grossmann, "An Iterative Aggregation/Disaggregation Approach for the Solution of a Mixed-Integer Nonlinear Oilfield Infrastructure Planning Model," *Ind. Eng. Chem. Res.*, **39**(6), 1955–1971 (2000).

Van den Heever, S. A., I. E. Grossmann, S. Vasantharajan, and K. Edwards, A Lagrangean Decomposition Approach for the Design and Planning of Offshore Hydrocarbon Field Infrastructures with Complex Economic Objectives, Submitted for publication (2000).

Van Hentenryck, P., *Constraint Satisfaction in Logic Programming*. MIT Press, Cambridge, MA (1989).

Van Roy, T. J., "Cross Decomposition for Mixed Integer Programming," *Math. Prog.*, **25**, 46–63 (1983).

Vargas-Villamil, F. D. and D. E. Rivera, "Multilayer optimization and scheduling using model predictive control: application to reentrant semiconductor manufacturing lines," *Comput. Chem. Eng.*, **24**, 2009–2021 (2000).

Vasantharajan, S., J. Viswanathan, and L. T. Biegler, "Reduced Successive Quadratic Programming Implementation for Large-Scale Optimization Problems with Smaller Degrees of Freedom," *Comput. Chem. Eng.*, **14**(8), 907–915 (1990).

Vecchietti, A. and I. E. Grossmann, "LOGMIP: A Discrete Continuous Nonlinear Optimizer," *Comput. Chem. Eng.*, **21**, S427–S43 (1997).

Viswanathan, J. and I. E. Grossmann, "A Combined Penalty Function and Outer-Approximation Method for MINLP Optimization," *Comput. Chem. Eng.*, **14**, 769 (1990).

Wallace, M., S. Novello, and J. Schimpf, "ECLiPSe: a Platform for Constraint Logic Programming," *ICL Systems Journal*, **12**(1), 159–200 (1997).

Westerlund, T. and F. Pettersson, "A Cutting Plane Method for Solving Convex MINLP Problems," *Comput. Chem. Eng.*, **19**, S131–S136 (1995).

Wilkinson, S. J., A. Cortier, N. Shah, and C. C. Pantelides, "Integrated production and distribution scheduling on a Europe-wide basis," *Comput. Chem. Eng.*, **20**, S1275–S1280 (1996).

Wilkinson, S. J., *Aggregate formulations for large-scale process scheduling problems*, PhD thesis, Imperial College, London (1996).

Yuan, X., S. Zhang, L. Piboleau, and S. Domenech, "Une Methode d'optimisation Nonlineare en Variables pour la Conception de Procedes," *Oper. Res.*, **22**, 331 (1988).

Zhang, X. and R. W. H. Sargent, "The optimal operation of mixed-production facilities-a general formulation and some approaches for the solution," *Comput. Chem. Eng.*, **20**, S897–S904 (1996).

Increasing Customer Value of Industrial Control Performance Monitoring—Honeywell's Experience

Lane Desborough* and Randy Miller†
Honeywell Hi-Spec Solutions
Thousand Oaks, CA 91320

Abstract

Within the process industries there is a significant installed base of regulatory and multivariable model predictive controllers. These controllers in many cases operate very poorly. This paper documents the current state of industrial controller performance, identifies the sources and ramifications of this poor performance, and discusses required attributes of a Process Control Monitoring System (PCMS). Finally, research directions are suggested.

Keywords

Performance assessment, Prioritization, Regulatory control, Control valve, PID control, Industrial survey, Stiction

Introduction

In an oil refinery, chemical plant, paper mill, or other continuous process industry facility there are typically between five hundred and five thousand regulatory controllers. As shown in Table 1, there are over eight thousand of these facilities in the United States alone (US Department of Energy, 1997).

There are somewhere between two thousand and three thousand multivariable model predictive control (MPC) applications installed world-wide, based on data from Qin and Badgwell (1997), with the market growing at a compound annual rate of approximately 18% (ARC Advisory Group, 1998, 2000b). Although use of MPC is now widespread, proportional-integral-derivative (PID) is by far the dominant feedback control algorithm. There are approximately three million regulatory controllers in the continuous process industries (based on data from Industrial Information Resources (1999); ARC Advisory Group (2000a) and an estimated ten thousand process control engineers (the latter estimate is based on data from Desborough et al. (2000) indicating the typical control engineer is responsible for between two and four hundred regulatory controllers).

When MPC is implemented, its manipulated variables are typically the setpoints of existing PID controllers. At the regulatory control level there has been little impact from other control algorithms. The importance of PID controllers certainly has not decreased with the wide adoption of MPC. Based on a survey of over eleven thousand controllers in the refining, chemicals and pulp and paper industries (Desborough et al., 2000), 97% of regulatory controllers utilize a PID feedback control algorithm.

Several trends are appearing that suggest the gap between desired and actual controller performance is widening:

- Competitive, environmental, and societal pressures are expected to require more changes in manufacturing facilities in the next 20–30 years than has occurred in the last 70 years (Katzer et al., 2000; American Petroleum Institute, 2000).
- When manufacturing sites are large enough to warrant a dedicated control engineer, their time is increasingly being diluted across implementing and maintaining advanced control technologies, display building, process historian support, and traditional PID controller maintenance.
- Process control application engineers often lack process control troubleshooting and time series / spectral analysis training and experience.
- Studies have shown that only about one third of industrial controllers provide an acceptable level of performance (Ender, 1993; Bialkowski, 1993). Furthermore, this performance has not improved in the past seven years (Miller, 2000), even though many academic performance measures have been developed in that time (Harris et al., 1999).

Facility Type	Total
Oil Refineries	246
Pulp and Paper Mills	584
Chemical Plants	2994
Power Generating Stations	3043
Primary Metal Industries	1453
Total	**8320**

Table 1: Continuous process manufacturing facilities in the United States.

Outline of the Paper

Practical control performance monitoring is a complex subject. In an attempt to explain the current state and articulate future research directions, a control metaphor has been adopted (Figure 1):

Minimize the deviation between measurements (current control performance) and setpoints (business ob-

*Lane.Desborough@Honeywell.com
†Randy.Miller@Honeywell.com

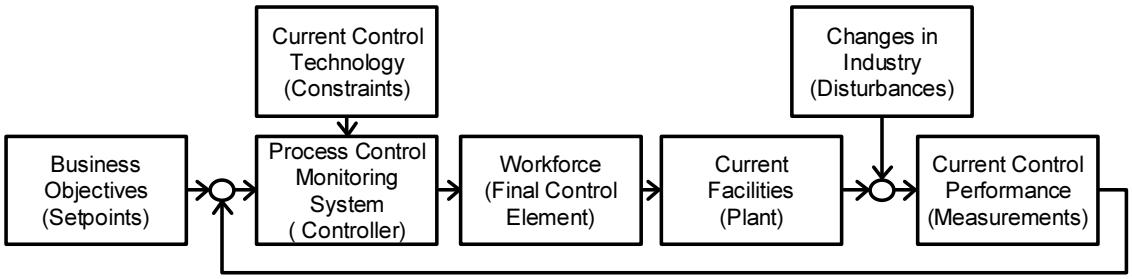


Figure 1: Control metaphor describing structure of the paper.

jectives) by implementing a controller (Process Control Monitoring System or PCMS) which is subject to constraints (current control technology). The PCMS changes the final control element (work activities of the control engineer) which in turn influences the plant (current facilities) and adapts to disturbances (changes in industry).

The outline of the paper is as follows:

- **Section 3: Current Control Performance (Measurements)**—the current control performance in industry is discussed based on a large worldwide sample of controllers.
- **Section 4: Business Objectives (Setpoints)**—the current business drivers within the continuous process industries are discussed.
- **Section 5: Current Control Technology (Constraints)**—the limitations of installed control systems and process models / testing are discussed.
- **Section 6: Workforce (Final Control Element)**—roles, responsibilities, and activities of industrial control engineers and other stakeholders are reviewed.
- **Section 7: Current Facilities (Plant)**—measurement types, facility uniqueness, and other issues are discussed.
- **Section 8: Changes in Industry (Disturbances)**—business, technology, people, and facilities factors expected to influence the direction of industrial control performance monitoring over the next decade are given.
- **Section 9: Process Control Monitoring System (Controller)**—the capabilities and characteristics of a Process Control Monitoring System (PCMS) are discussed.

Section 10 provides two industrial examples. In Section 11, research directions are suggested.

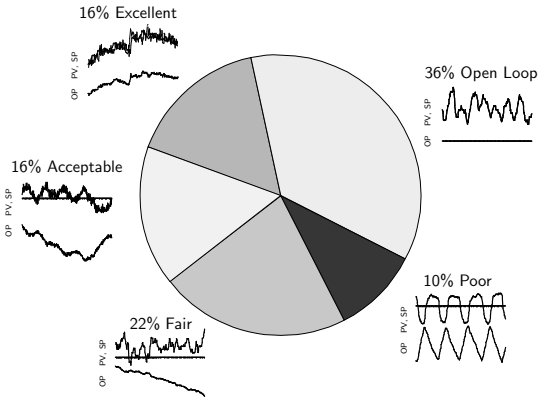


Figure 2: Global multi-industry performance demographics.

Current Control Performance (Measurements)

Performance demographics of twenty six thousand PID controllers collected over the last two years across a large cross sample of continuous process industries are shown in Figure 2 (Miller, 2000). An algorithm combining a minimum variance benchmark and an oscillation metric tuned for each measurement type (flow, pressure, level, etc.) was used to classify performance of each controller into one of five performance categories. These classifications were refined through extensive validation and industry feedback to reflect controller performance relative to practical expectations for each measurement type. Unacceptably sluggish or oscillatory controllers are generally classified as either "fair" or "poor" while controllers with minor performance deviations are classified as "acceptable" or "excellent". A level controller's performance is difficult to classify without knowing its objective—regulation, servo control, or most commonly surge attenuation. The above analysis assumes that level controllers have a surge attenuation objective, meaning they receive a "poor" classification if they transfer excessive variability to the manipulated variable (e.g. the flow out of the surge vessel). Controllers receive an "open

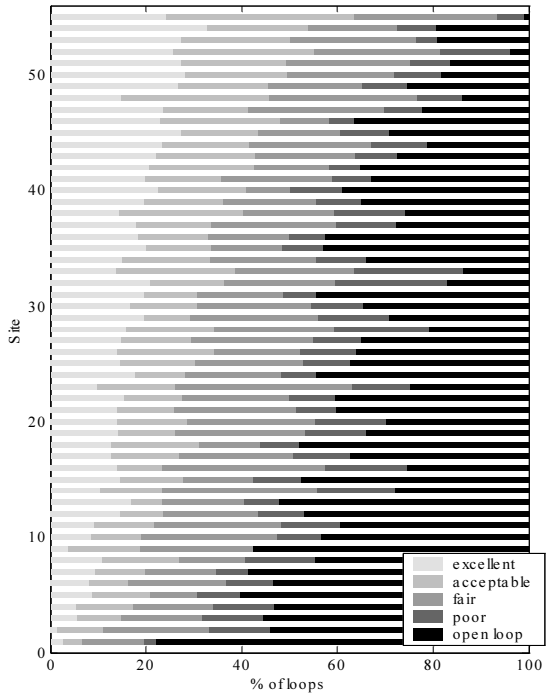


Figure 3: Site wide performance distribution.

loop" classification if they are in manual mode or the output is saturated (stuck at a limit) for more than 30% of the dataset (five thousand samples at the dominant time constant).

Only one third of the controllers were classified as acceptable performers and two thirds had significant improvement opportunity. Some controllers classified as open loop are truly in their normal mode, for example, a bypass flow controller used only during startup. However, many of the controllers under manual control are obsolete or cannot be closed due to an operability problem.

Business Objectives (Setpoint)

The major US process industries spend about thirty billion dollars annually on energy (see the Appendix) and over one hundred billion dollars on facility maintenance (Industrial Information Resources, 1999). Even a 1% improvement in either energy efficiency or improved controller maintenance direction represents hundreds of millions of dollars in savings to the process industries.

Businesses are measured by macroscopic metrics such as share price and customer orders. These are in turn affected by key performance indicators (KPI's) such as product quality, product consistency, throughput, energy efficiency, and lost time injuries. The majority of all business decisions in a continuous process facility are implemented by changing the signal to a control valve, almost always through the action of a regulatory controller. Thus regulatory control has a profound impact on key performance indicators and ultimately business value. Understanding the operational context of a particular controller is key to the success of a control performance monitoring work practice. Relating controller performance to KPI's requires a system-level view of regulatory control:

1. impact—does a particular subset of controllers impact bleach plant brightness more than others? Often these impacts are qualitative, descriptive, or immeasurable.
2. mode—is the facility in high production, startup, shutdown, or energy efficiency mode? Mode can often have profound impact on controller performance and vice-versa, as different procedures employ different controllers. As an example, MPC is not usually used in startup and shutdown mode because it often has a low turndown ratio.
3. grade—is the facility running heavy versus light crude or making newsprint instead of catalog paper? Differences in the active constraint set, objective function and process model from one grade to the next can significantly affect controller performance.
4. objective—does the tight tuning of level controllers in surge vessels accentuate rather than attenuate destabilizing unit-to-unit interactions? Controller objectives include servo control, regulatory control, constraint control, and surge attenuation.

The above-mentioned extrinsic effects of the controller are as important for a PCMS to address as the intrinsic controller performance itself. By tying individual controller performance to the effect that performance causes, the process control engineer can make an informed decision as to the priority of resolution. There will always be more work to be done than time available to do it.

Controller performance is often defined narrowly as the ability of the controller to transfer the proper amount of variability from the controlled variable (CV) to the manipulated variable (MV). While variability transfer is a very important contributor to a controller's performance, there are others as well:

- Alarms—almost every industrial PID controller or multivariable controller is configured with alarms to alert the operator when an unacceptable process deviation has occurred. Commonly configured alarms include process value high, low, rate of change, manipulated variable high, low, or frozen, and off normal control mode. These alarms are presented in a special alarm summary page on the control system's user interface, on panel-mounted enunciator boards, or as audible sirens or bells. Due to the ease with which alarms can be configured, there has been a tendency to build too many alarms, or

alarms with inappropriate limits. When a true incident occurs, an "alarm flood" is precipitated and the operator becomes unable to determine root cause and choose the correct path to resolution. Incidents traced to abnormal situations and the resulting alarm flood have resulted in over forty billion dollars in losses in the petrochemical industry alone (Campbell Brown, 1999). Measuring the number of "bad actors" or chattering alarms helps control engineers proactively manage and prioritize controller alarm performance.

- Interventions—process operators are responsible for the daily operation of the plant. Their principal means of effecting process change is to intervene in the operation of the MPC and regulatory controllers. Interventions include changing a controller's setpoint, changing its mode from automatic to manual, directly changing the output to the valve, or changing an MPC's constraint limits or cost function inputs. Operators spend their entire shift reacting to stimuli and making hundreds of interventions to the control system. These interventions can and do result in inappropriate variability transfer, often resulting in an easier to operate plant but one further from its economic optimum operating point. For instance, almost thirty percent of sampled PID controllers are in open loop, meaning the operator has intervened to remove any automatic control action. Some operating companies track and report operator interventions as an element of controller performance (Takada, 1998). The situation is equally acute in MPC, with as many as 30% of controllers inoperative and a similar number rendered effectively inoperative by the operator through clamped-down move limits and constraints.
- Configuration Changes—controller performance can be affected when a change is made in the feedback algorithm tuning, the transmitter, or the final control element. In one customer example (Desborough and Nordh, 1998), an environmental emissions team with a portable gas probe went from valve to valve, measuring for fugitive hydrocarbon emissions. Finding a leaky valve, they would tighten the actuator packing. Weeks later, the operator would complain to the control engineer about sluggishness and hysteresis (resulting in oscillations), and the control engineer would instruct the valve technician to loosen the actuator packing.

Most alarm, intervention, and configuration change events are recorded in the control system's event log, and are available for analysis.

Consider a typical scenario: an operator on the night shift makes a change in a controller's gain to improve the variability transfer performance while operating in maximum throughput mode (it's cooler at night so there are fewer cooling water temperature constraints). On the following day shift, the new operator, who has not been apprised of this tuning change and is now trying to operate the plant in energy conservation mode, acknowledges multiple alarms coming from the controller indicating high rate of change on the measured variable. He ultimately places the controller in manual so that its variability transfer problem is attenuated but in doing so sacrifices some energy efficiency. Through the remainder of his shift, he is forced to make multiple manual changes to the controller, which distracts him from his other duties. When the control engineer performs the troubleshooting activities surrounding why the day shift had difficulty running in energy conservation mode, five elements are involved: the energy conservation mode operating context, the variability transfer performance, the alarm performance, the operator intervention performance, and the configuration change management.

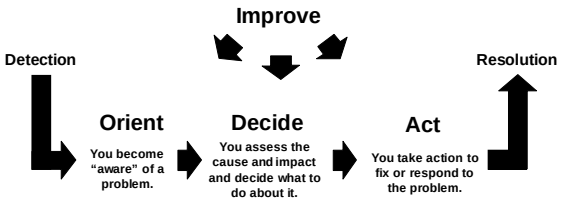


Figure 4: Decision support workflow.

Without an understanding of how the various controller performance measures (variability transfer, alarms, and operator interventions) relate to the business KPI's, the control engineer will not be able to focus their finite work effort on the most important problems, and instead will be forced to take subjective work direction from others who are more closely aligned with business performance.

Fighter pilots are taught to observe, orient, decide, and act—the so-called OODA Loop (Boyd, 1987). Similarly, the Six Sigma quality process teaches the DMAIC process improvement methodology: Define, Measure, Analyze, Improve and Control (Pyzdek, 2000). In oil refineries, paper mills, and other process industry facilities a similar workflow is followed by various stakeholders in controller performance (Figure 4). Managers, operators, process control engineers, and to a lesser extent maintenance technicians orient, decide, act, and improve controller performance:

- *Orient*—system-wide identification of specific problems, preferably automated "has the performance changed?"
- *Decide*—determine problem's causes / effects through analysis of facts / further investigation and decide on resolution "what should I do about the performance change?"

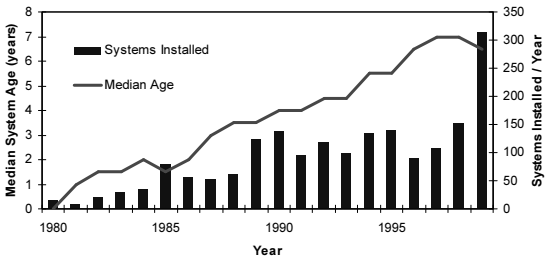


Figure 5: DCS installed systems and median age.

- *Act*—take action on problem through mitigation, investigation, or repair
- *Improve*—assess improvement in orient, decide and act processes

Although it is primarily the control engineer's job to orient, decide, and act on controller performance, often regulatory controls are not looked at unless there is a problem identified by the operator, or if it is a part of an MPC application. Control engineers are very busy with many responsibilities other than regulatory control.

Control Technology (Constraints)

In order to appreciate the issues surrounding practical industrial control performance monitoring, it is important to understand the system constraints present:

- Real time, high frequency time series data collection and automatic analysis is difficult and time consuming
- Legacy control systems weren't designed for performance monitoring hence many are not up to the task from a computing horsepower perspective
- Getting data from the legacy control system to a more powerful computing platform is limited by the available bandwidth
- Dynamic process models are unavailable for the vast majority of controllers, and would be prohibitively expensive to obtain
- The PID control algorithm dominates the continuous process industries

One of the biggest issues with practical controller performance assessment is data access and computing power. Based on a sampling of all US oil refinery and power plant distributed control systems, the median distributed control system (DCS) age is seven years and increasing (Figure 5) (Industrial Information Resources, 1999). Many plants have control systems which are fifteen years old.

The vast majority of distributed control systems operating in the world today were simply never designed to easily provide high frequency time series data (one sample per second) or perform complex calculations. Typically less than one thousand measurement parameters can be transferred per second to Windows-based computing platforms.

Dynamic process models are extremely expensive to obtain, either empirically or from first principles. Based on hundreds of Honeywell control projects, engineering costs typically range from \$250–\$1000 per single-input, single-output model. These costs include experimental design, plant testing, dynamic model identification, and model validation, but do not include software, hardware, or training. They also do not include the cost of process disruption as the plant is perturbed away from its economic operation conditions. About the only place the cost of dynamic modeling is ever warranted is during MPC implementation. Due to the significant costs involved, models exist for far less than one percent of all processes controlled by regulatory controllers. Even where these models have been created, they are typically very poorly documented or are out of date (models developed for MPC are the exception, as this is usually done as a well-documented project).

In a survey conducted by Honeywell (Desborough et al., 2000) of 11,600 regulatory controllers across eighteen facilities, the PID control algorithm was used almost exclusively. The site median for PID feedback control algorithm use was over 97% percent (Table 2).

There are at least three reasons for the predominance of the PID algorithm:

1. The PID algorithm works very well in the vast majority of applications. For the rare case of complex dynamics or significant time delays, other algorithms are occasionally used but it is more common to instead implement cascade control to facilitate dynamic decoupling.

2. The PID algorithm is easy to understand. A vast body of literature exists on PID implementation and tuning, and a number of software packages are available which facilitate PID tuning.

3. The PID algorithm is pre-programmed in every control system. Implementing a non-PID feedback control algorithm involves programming custom logic and could take as much as one hundred times the effort of implementing a PID algorithm, not counting the intangible lifecycle costs including documentation, support, and troubleshooting.

Commercially available multivariable model predictive controllers are implemented almost exclusively as constraint-pushing optimizers (see Sorensen and Cutler, 1998; Anderson et al., 1998; Hardin et al., 1995, for a representative sample). They tend to act more like dynamic optimizers than multivariable regulatory

	1st decile	average	median	9th decile
Feedback Control—PID	94.7%	97.3%	97.7%	99.7%
Feedback Control—non-PID	0.0%	1.7%	0.6%	4.8%
non- Feedback Control	0.0%	1.0%	0.6%	3.1%

Table 2: Use of PID in continuous manufacturing facilities world-wide.

People	Control Engineers spend a great deal of their time troubleshooting problems
	Control Engineers don't spend much time on regulatory control performance
	Instrument technicians don't use computers for passive data analysis
Process	Many problems with controller performance are due to external process problems
	MPC problems are usually caused by operability instead of model mismatch / tuning
	Problems are resolved via tuning only a small proportion of the time
	Operators are the control engineer's most important source of information
	Different groups have conflicting objectives which impact control performance, e.g. tighter valve packing reduces hydrocarbon emissions but increases stiction.
Tools	Information needed for diagnosis: • High frequency trend data • Process insight and other non-quantitative data
	Past internal attempts to develop a PCMS have failed

Table 3: Voice of the customer summary.

controllers and are rarely square. This has very important implications for control performance monitoring, as metrics commonly associated with controller performance such as minimum variance have virtually no relevance for a controller whose objective is not regulation, but constrained optimization. There are a surprisingly high number of these controllers operating so tightly constrained that the optimizer is ineffective and the system is essentially open-loop. In the experience of many users (Desborough and Nordh, 1998), performance of these controllers has more to do with the way the operator sets the various MV and CV constraints than the degree of MV / CV variability and variability transfer. This suggests a need for improved user interfaces, training, and diagnostics for operators so they won't constrain the controller so tightly.

Workforce (Final Control Element)

In April 1998, Honeywell visited eight customer sites around the world asking "What are the past, current, and future needs of those persons responsible for maintaining controllers in continuous process industry facilities such as pulp mills, oil refineries, and ethylene plants?" Over twenty managers, control engineers, and instrument technicians were interviewed. The "Voice of the Customer" (VOC) methodology (Burchill and Hepner Brodie, 1997) was followed. The VOC trip resulted in over 900 "voices" from customers. These were categorized and organized into a spreadsheet from which product requirements were identified (Desborough and Nordh, 1998). These are summarized in Table 3.

The most important and often heard requirements were to a) make the technology simple and easy to use, b) allow the user to find information quickly and easily, and c) be simple to setup and maintain.

Further, it was identified that the user interface should:

- incorporate time series trends to assist in the diagnosis of problems
- present a prioritized list of controllers that are not meeting performance criteria
- present information in a "push" fashion, for example a problem could be highlighted in an email to the control engineer, versus the engineering having to sort through additional reports.

In the fall of 1998, Honeywell sent out a 200 question survey and received approximately 35 responses from control engineers. In summary the following PCMS requirements were identified:

- Easy configuration
- PC platform
- Single page summaries
- Time series trends

- Client-server architecture
- Cost-effective
- Standard data formats
- PID instead of MPC performance monitoring initially
- On-demand analysis

Some additional results from the VOC trip and customer survey will now be discussed. Control engineers, operators, instrument technicians, and managers are usually co-located in the plant and have a good relationship. They have shared objectives of ensuring the safe and economic operation of the facility. Quite often these objectives are explicitly set through KPI's. As well, KPI's are often used directly as inputs to individual compensation.

A controller is a capital asset, and as with any capital asset it has a well-documented lifecycle within an organization starting with its purchase and ending with its disposal. The control engineer is involved with virtually every aspect of the controller's lifecycle. Delivery of the controller to a usable state is facilitated through HAZOP and engineering design, plant testing, model identification, commissioning, and operator training activities carried out by the control engineer. The current undergraduate control curriculum addresses some but not all of these activities—notable exceptions are troubleshooting, spectral analysis, statistics, and experimental design. Operators are the primary users of controllers. During a controller's useful life, the process control engineer plays a supporting role through the monitoring, diagnosis, and resolution of performance problems. Controllers are commissioned very infrequently, usually only during major plant expansions or during new plant construction. The control engineer thus spends the majority of their time monitoring and maintaining controllers and other applications resident on the DCS.

Control engineers have formal and on-the-job education in process control. Typically they divide their time between regulatory control troubleshooting and advanced control. They also maintain the alarm system used by operators. Their typical responsibilities are listed in Table 4 (Desborough et al., 2000). Informally, they can act as focal points for instrumentation, IT, operations, and process engineering. Their daily tasks are widely varied, generally consisting of meetings, troubleshooting, and new development. They use DCS engineering tools, multivariable control design tools, and standard office software such as Microsoft Word and Excel. They interact primarily with operators. Control engineers have survived decades of downsizing and outsourcing. They want to be rewarded by interesting work and are often frustrated by mundane operator-initiated troubleshooting tasks. MPC implementation and support are considered high-valued control engineer activities. Tuning skills required to maintain regulatory control are not perceived as unique or tremendously valuable. Control engineers are goal driven and usually have a very good understanding of the business and process objectives of the facility. Control engineers are the implementers and in many cases the maintainers of the controllers, but they are not the end users of the controllers that they implement—the operators are. It is important to note that many plants don't have a dedicated control engineer.

Operators are the users of the controllers. They are responsible for the day-to-day safe and economic operation of the facility. Operators control the plant, usually by changing modes or setpoints of regulatory and MPC controllers or giving instructions to outside operators, and they act as a focal point for anything which might affect the plant. They have a practical understanding of process operation, sometimes supplemented by formal education such as a two year technology certificate from a trade college. Operators interact with other operators, maintenance technicians, process engineers, and control engineers, typically via face-to-face discussions, but also via log books or other reporting mechanisms.

Instrument technicians are usually very responsive to the needs of operators and control engineers and spend most of their time maintaining the electronic and mechanical elements of the controller. Often they do not have access to PCs where they are doing their work (in the field). Their focus tends to be on the resolution rather than the identification of instrument problems. They are very task driven and only the best technicians have an understanding of the business and process objectives impacted by their work.

Managers are concerned with the economic operation of the facility. They have a wide range of experience, but are typically promoted from operations, maintenance, or engineering. They are responsible for making sure the people and groups within the facility work in a manner consistent with business objectives. They read reports and attend meetings. They use a variety of communication and office tools. They interact with control engineers, operators, and maintenance technicians.

Current Facilities (Plant)

Unlike a head position controller on a computer hard disk, where the same control algorithm and tuning parameters can be reused in thousands of identical units, every process in the continuous process industries is in some way unique and as a result every controller implementation is a custom activity. Nominally identical paper machines or ethylene furnaces will have subtly different dynamics, operating objectives, feedstocks, or product grades, requiring each controller to be individually commissioned and tuned to suit the particular business context. As a result there are few economies of scale.

	1st decile	average	median	9th decile
PID feedback control	152	332	289	576
non-PID feedback control	0	3.9	5.3	22.0
non-feedback control	0	2.5	2.0	7.7
Calculations / other	31	70	59	184
APC	1.5	6.3	5.0	13.5

Table 4: Control engineer responsibilities.

	1st decile	average	median	9th decile
Composition	0%	2%	3%	6%
Flow	22%	39%	37%	46%
Level	12%	20%	19%	30%
Pressure	16%	20%	20%	26%
Temperature	14%	19%	17%	36%

Table 5: Regulatory control measurement types.

Higher performance algorithms are rejected in lieu of the PID algorithm, which for reasons outlined in Section 5 is easier to implement and support.

Another reason the PID algorithm is so commonplace is that the vast majority of process measurements have fast dynamics with minimal process delay. One exception is composition control, which is often based on measurements from a slow analyzer with long delay such as a gas chromatograph. Table 5 shows a distribution of regulatory control measurement types from a sample of over ten thousand controllers at eighteen sites in multiple process industries (Desborough et al., 2000).

In the authors' experience, non-minimum phase processes are seldom encountered, usually on less than one in a hundred loops (typically boiler water level control or cold-hydrocarbon fed exothermic reactors). Even in a boiler where the boiler water level is subject to shrink-swell non-minimum phase behavior, the level controller is but one of the controllers required to operate the boiler effectively, and the other controllers do not exhibit non-minimum phase behavior.

Changes in Industry (Disturbances)

There are several profound changes that are expected to influence control performance expectations and challenge industry's ability to meet those expectations. The trend of tighter environmental regulations will continue, as evidence of the effect of CO2 emissions on climate grows (Meszler, 1999), creating new constraints layered on new economic objectives. As stated by the American Petroleum Institute (1999), "The petroleum industry of the future will be environmentally sound, energy-efficient, safe and simpler to operate. It will be completely automated, operate with minimal inventory, and use processes that are fundamentally understood." There will be higher expectations on control systems to reduce process variability that influence emissions and waste.

Fewer new plants are being built to meet increasing demand. For example, no new refineries were built in United States in the 1990's while capacity increased by 120,000 b/d in 1999 alone (Chang, 1999). The number of refineries in United States has actually decreased since the 1980's (American Petroleum Institute, 2000). Even the time between shutdowns is being challenged by developments in heat transfer fouling mitigation (Panchal and Ehr-Ping, 1998).

Competitive pressures, mergers and acquisitions have had the effect of increasing the responsibilities of each control engineer. The total number of engineers employed in the United States has been in decline since 1987 (`http://stats.bls.gov/oco/ocos027.htm`). Several sites that the authors communicate with are reporting skill shortages to meet the demand of maintaining MPC and regulatory control applications (Desborough and Nordh, 1998). To have any impact on industry, controller performance monitoring must be automated and intuitive to the average Bachelor's level engineer.

Instrumentation technology advancements will have a positive impact on variability and reliability. As instruments and valve positioners fail they are generally replaced with digital or smart devices that improve precision of control and provide self-diagnostic information.

One other important trend in the area of data access is the move to OPC (OLE for Process Control), an open standard for data access. OPC will make data access ubiquitous. All major control vendors have developed OPC support for their legacy systems and industrial users are aggressively moving to OPC (Studebaker, 1999).

Process Control Monitoring System (Controller)

The high-level goal of a Process Control Monitoring System (PCMS) is to provide plant control engineers with enhanced capabilities to identify problems for many controllers while minimizing additional effort or expense. By combining the computer's ability to rapidly gather and analyze large quantities of data with the control engineer's abilities to recognize patterns and understand complex relationships, controller performance can be improved while simultaneously freeing the control engineer to spend more time on high-valued activities.

A PCMS collects data, computes metrics, and presents these metrics in a form suitable for the user to take the appropriate action. Results from the Honeywell VOC trip and survey indicate that users seek the answers to two fundamental questions:

1. Has the controller performance changed?
2. What should I do about the performance change?

A PCMS must facilitate the orient-decide-act-improve workflow for business, operational, engineering, and maintenance stakeholders. It must include a mechanism for computing metrics and a framework for guiding users through appropriate diagnostic actions so they can choose the proper course of action. It should track changes in configuration and operations context.

Metrics

A metric is defined as a standard measure to assess performance. Controller performance metrics fall into three broad domains:

- Business metrics—is the controller meeting its business objectives? Examples include LP objective function for an MPC / optimizer, and product quality variation. These metrics help to indicate if the controller is meeting its business objectives.
- Operational metrics—is the controller being used effectively by the operator? Examples include uptime / service factor, and MV limit constraint shadow costs. These metrics help to indicate if the operator is interacting with the controller in a way that helps to fulfill the business objectives.
- Engineering metrics—is the controller meeting its engineering objectives? Examples include dynamic model accuracy and minimum variance benchmarks. These metrics help to diagnose engineering deficiencies within the controller.

Different users, because of their roles and responsibilities, have different metrics needs (Table 6). A PCMS must be able to meet the needs of each type of user, with a special emphasis on the control engineer.

According to Trimble (2000), there are two types of metrics:

- performance metrics—high-level measures of overall performance, usually focused on the effect of the asset's performance on the wider system or business. Business metrics and operational metrics tend to fall into this category. Performance metrics are primarily used to orient the user to the presence of a problem. It is preferable for these metrics to be computed and presented automatically so as not to burden users with additional tasks.

- diagnostic metrics—measures which indicate why performance is unacceptable, usually focused on the asset's internal workings. Engineering metrics tend to fall into this category. Diagnostic metrics help the user decide which action to take once a problem has been identified. These metrics quite often involve interactive data visualization, invasive testing, or other manual activities.

It was identified in Section 5 (Constraints) that the vast majority of controllers lack any kind of process model. Also, most control systems are poor providers of time series data and event data, making collection difficult and time consuming, and therefore expensive. Metrics that require special data may be extremely expensive to compute (Table 7).

Any metric that requires any kind of model or data which is difficult to obtain must have an informational benefit well in excess of the cost of model creation (invasive plant tests, model identification, documentation, etc) and collection of special data.

There are a number of criteria to consider when defining controller performance metrics. Trimble (2000) asserts that metrics must be SMART (Specific, Measurable, Actionable, Relevant, and Timely). Caplice and Sheffi (1994) similarly propose the following metric criteria: validity, robustness, usefulness, integration, economy, compatibility, level of detail, and behavioral soundness, which are described further in Table 8.

Metric Presentation

Due to the wide scope of responsibility for the control engineer and other stakeholders, it is important to collect and present metrics across a wide breadth of responsibility at the appropriate analysis depth, ranging from overall facility performance metrics to individual valve diagnostic metrics. A single metric with a narrow or shallow scope will not help users answer the basic questions posed at the beginning of this section and summarized in Table 9. Instead, a PCMS must contain an appropriate balance of detailed individual controller diagnostic metrics and overall performance metrics within a presentation environment which allows user to overview, zoom and filter, and finally obtain details on demand.

User	Business Metrics	Operational Metrics	Engineering Metrics
Manager	60%	30%	10%
Operator	10%	80%	10%
Control Engineer	10%	30%	60%

Table 6: Metrics needs by user.

Data Type	Comment
high frequency time series data	shorter than 5 second sampling period
event data	alarms, operator interventions, controller configuration changes
invasive process testing	designed experiments to obtain dynamic models
manually-entered data	configuration or economic data
continuous data collection	an automobile's odometer is only useful if it is collecting data all of the time

Table 7: Special data types.

Criterion	Description
Validity	The metric accurately captures the events and activities being measured and controls for any exogenous factors.
Robustness	The metric is interpreted similarly by all users, is comparable across time, location and organizations, and is repeatable.
Usefulness	The metric is readily understandable by the decision maker and provides a guide for action to be taken.
Integration	The metric includes all relevant aspects of the process and promotes coordination across functions and divisions.
Economy	The benefits of using the metric outweighs the costs of data collection, analysis, and reporting.
Compatibility	The metric is compatible with the existing information, material and cash flow systems in the organization.
Level of Detail	The metric provides a sufficient degree of granularity or aggregation for the user.
Behavioral Soundness	The metric minimizes incentives for counter-productive acts or game-playing and is presented in a useful form.

Table 8: Metrics criteria.

	Has the controller performance changed?	What should I do about the performance change?
Workflow:	Orient	Decide / Act
Information level:	Overview First	Details on Demand
Metric type:	Performance Metrics	Diagnostic Metrics
Data gathering	Automatic	Human-Facilitated
Result presentation:	Automatic Push	Manual Pull
Invasiveness	Non-invasive	Invasive
Breadth:	Wide	Narrow
Depth:	Shallow	Deep
Focus:	Broaden Focus	Narrow Focus

Table 9: Basic process control monitoring questions.

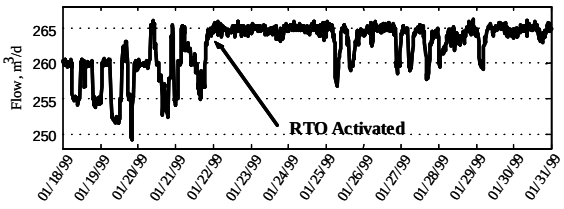


Figure 6: Hydrocracker unit throughput in response to control improvements.

Case Studies

BP Grangemouth Refinery, Scotland

Grangemouth Refinery is an integrated 210,000 BPD refinery complex. The hydrocracking unit of the refinery consists of four reactors: two series-flow reactors which hydrotreat the fresh feed, a first stage hydrocracking reactor which converts part of the feed, and a second stage hydrocracking reactor which receives feed from the fractionator bottoms/mild vacuum column (unconverted oil). The second stage reactor completes the conversion of feed to lighter material by cracking unconverted oil in the recycle feed stream. The hydrocracking reaction is highly exothermic. Both safety and throughput depend on tight temperature control to meet and push constraints.

A first generation MPC and a feed maximizing optimizer that were implemented in 1994 were upgraded to current technology in 1998. As is common in such upgrades, no new step tests or modeling were performed. During 1998 the feed source and product mix objectives shifted and a temperature oscillation manifested itself in the second stage reactor bed outlet temperatures. This cycle propagated throughout the entire recycle controller and ultimately to the main fractionator, mild vacuum column and the second stage reactor. Because of this cycle in reactor temperatures, the reactor weighted average bed temperature MPC was often fully constrained and thus could not achieve its objective. The feed maximizer could not run safely under these conditions, resulting in significant lost opportunity.

A complete step test and controller redesign were undesired due to the cost, workload of on-site staff, and the length of time required for a redesign. A vendor consultant, who is an expert on hydrocracking processes and MPC applications, examined active constraints, operating data, and MPC tuning closely. A restricted bump test was conducted to update a small fraction of the MPC models. A systematic controller assessment of the unit was also conducted that showed four key PID controllers were poorly tuned (Fedenczuk et al., 1998). In addition to PID tuning and updating a few models, several changes were made to temperature constraints and profiles. Figure 6 shows the increased throughput of the unit when the feed maximizer was enabled.

This is a typical MPC maintenance problem—complex and multifaceted, requiring a holistic diagnostic approach. Problems that are more subtle are often more difficult to diagnose. Reducing reliance on a human expert requires quantitative assessment of model adequacy, MPC constraints, MPC tuning, and alignment of MPC objectives with process objectives.

Engen Petroleum Durban Refinery

Engen Petroleum Durban refinery is a medium scale (nominal 100,000 BPD) refinery in South Africa. The only opportunity to service most control valves is during shutdowns, which are planned every two or three years. While it is important to correctly identify poor performing control valves that need maintenance, it is both expensive and time consuming to invasively test all control valves. To this end, a comparative test of a commercial non-invasive controller performance service and invasive valve tests was performed.

Results of non-invasive tests and invasive valve tests of seven problem controllers are listed in Table 10. Valve stick—slip is defined as the resolution in actual stem travel. Dead band is defined as the minimum change in the valve input signal before the stem will move. Non-invasive tests were based on qualitative pattern analysis of the process variable and controller output time series data in normal closed loop operation by a human expert. The Entech control valve dynamic specification (valve stick—slip plus deadband) cites a value of one percent as the threshold for nominal control valves (EnTech, 1998). Using this criterion, all valves in Table 10 fail the test and should be maintained. However, the validated problem resolution suggests that only three of the seven valves actually had a valve-limiting performance problem.

In the non-invasive assessment, only the valve stick-slip and dead band relevant to closed loop performance is significant, which proved to be more reliable in this comparison. Dependence on a human expert is however a strong condition that will be influenced by individual biases and experience.

New Research Directions

Many issues still need to be addressed before a reliable, comprehensive PCMS can be developed that meets the needs of the industrial user. The extent of industrial controller monitoring adoption will be strongly influenced by closing the gaps between the industrial user needs and the current state of the art (subject to the constraints already discussed).

First, general recommendations for performance metrics and diagnostic metrics will be given, covering a diverse set of new research directions. Next, a much smaller subset of the most important research directions will be addressed in greater detail.

Loop	Controller Behavior Problem	Non-invasive Valve Analysis	Invasive Test		Problem Resolution	Resolution Consistent With:	
			stick-slip	dead band		non-invasive test	invasive test
1	Oscillating	severe stiction	1.40%	2.33%	Valve	yes	yes
2	Low σ^2, not oscillating	valve OK	1.80%	1.39%	Tuning	yes	no
3	Oscillating	moderate stiction	0.40%	2.33%	Valve	yes	yes
4	Saw-tooth pattern	moderate stiction	0.50%	0.57%	Valve / Tuning	yes	yes
5	High σ^2, not oscillating	valve OK	1.50%	0.40%	Tuning	yes	no
6	Oscillating	valve OK	1.20%	0.30%	Tuning	yes	no
7	High σ^2, not oscillating	valve OK	0.30%	1.75%	Tuning	yes	no

Table 10: Comparison of invasive and non-invasive valve analysis.

General Research Directions

Performance Metrics. Performance metrics are designed to broaden rather than narrow the user's current focus, and help them orient to the presence of problems on controllers they wouldn't otherwise be examining. They are high-level measures of overall performance, usually focused on the effect of the asset's performance on the wider system or business. In general they require minimal user configuration effort. They are based on available data and their computation can be performed automatically. Their presentation should be automatic and intuitive to the average user. Performance metrics should ultimately help the user shift their current work activities to a more important area. A summary of recommended performance metric research directions is presented in Table 11.

Diagnostic Metrics. Diagnostic metrics are designed to narrow the user's current focus, and help them to decide which action to take to resolve a problem on a specific asset. They are often detailed measures of performance, usually focused on the asset's internal workings or inputs. They may require user configuration effort and quite often have a cost associated with them, either in terms of user effort or process disruption. They often require new data to be gathered, and computations and analysis usually require human intervention. Their presentation requires user interaction. Diagnostic metrics should ultimately help the user select the proper action to resolve a specific asset's problem. A summary of recommended diagnostic metric research directions is presented in Table 12.

Specific Research Directions

Knowledge Capture and Continuous Improvement. There is a need to establish a knowledge infrastructure founded on consistent models and representations of controller performance, much analogous to the financial community's standardized set of accounting metrics and practices:

- Benchmarking standards (facility-wide, MPC, regulatory control, valve)
- Alarm, operator intervention, and variability transfer performance tracking
- Probabilistic categorization of performance faults
- Normalization and scaling (for comparison to other controllers and protection of proprietary data)
- Weibull analysis—equipment failure; what drives controllers to fail and can this be generalized / predicted? (e.g. is there a relationship between positioner life and degree of oscillation?)
- Rigorous actuator nonlinear modeling

Automated Non-invasive Control Valve Stick-Slip Detection. Control valve problems account for about one third of the 32% of controllers classified as "poor" or "fair" in the industrial survey (Miller, 2000). Faults in control valves are often intermittent and are often misdiagnosed with simple minimum variance ratios and spectral analysis. An abundance of literature exists for invasive analysis of control valve performance that requires stroking the valve when either in-service or out-of-service (Fitzgerald, 1988, 1990; Ancrum, 1996a,b; Boyle, 1996; Wallen, 1997; Sharif and Grosvenor, 1998). With invasive tests, the amount of change in signal required to move the valve stem (stick) and the amount it moves when stem friction is overcome (slip) is easily quantified. However, except for the cross-correlation work of Horch (1998), no non-invasive methods have appeared in the literature. It is neither cost-effective nor practical to detect valve faults using invasive approaches across an entire site. A passive method that can reliably and automatically classify valve performance in closed loop is a desperately needed component in the orientation phase.

Scope	Research Direction
Facility-wide	1. Behavior clustering (common oscillations) 2. Automated model-free causal analysis 3. Performance change detection (changes from target, past history, industry norm) 4. Control performance analysis for specific processes: • sheet-forming processes (e.g. paper machine cross-direction / machine direction control) • processes with tight heat / material integration (e.g. ethylene plants) • processes with high degrees of government regulation (e.g. biotech, food and beverage, nuclear power, pharmaceuticals industries) 5. Business impact analysis of controller performance
MPC	1. Constraint and objective function representation 2. Operability (by operating shift, mode, grade, and objective)
Regulatory Control	1. Alarm, operator intervention, configuration, and variability transfer performance representation 2. Performance categorization based on measurement type / control objective (especially level control)
Valves	1. Non-invasive automated actuator stick-slip detection with only PV, SP, OP data

Table 11: Research Directions for performance metrics.

A valve technician usually carries out invasive valve analyses with an objective of comparing the open loop valve performance with the manufacturer's specifications. This is a valve-centric view of performance. The authors have noted several instances where the invasive analysis conflicts with a graphical analysis of the closed-loop time series. Quite often the valve is operating within its specification but still causes a significant stick-slip behavior in the process variable.

Several motivating examples are now given to illustrate a few validated patterns of control valve problems. The first example (Figure 7) shows an obvious valve-induced oscillation that is approximately symmetrical. The PV versus OP plot is characterized by a rectangular pattern tilted to the left. According to Horch (1998), valve-induced oscillations have a zero cross-correlation at zero lag while tuning-induced oscillations produce a minimum or maximum at zero lag. This first example has a cross-correlation of nearly zero at zero lag.

The second example (Figure 8) has an asymmetrical time series confounded by significant setpoint changes. A rectangular pattern tilted to the left with a shifting centroid can be seen in the PV versus OP plot. Here the cross-correlation is about half way between zero and the minimum, which is inconclusive.

In the third example (Figure 9), a large filter constant was applied in the DCS in an attempt to compensate for the valve behavior. The resulting pattern in the PV versus OP plot shows a wedge-shaped object tilted to the left. The cross-correlation is again inconclusive.

In the fourth and final example (Figure 10), an intermittent valve stick-slip can be seen where the slip is of different magnitudes. The patterns in the PV versus OP plot are still rectangular objects tilted to the left. Cross-correlation fails to identify this as a valve problem because the time series is not repeating. In each of these examples the trained human eye can quickly verify the existence of valve stick-slip.

Performance-Impact Prioritization. A controller is implemented with the objective of changing the final control element to ultimately achieve a business objective. Specifically, it causes the final control element to move from its current value to a value it wouldn't have otherwise pursued. By the same token, a PCMS (controller) has the same effect on the workforce (final control element)—causing the workforce to change their current set of work activities from what they otherwise would have done. The work activities of the control engineer should be prioritized based on making changes to the worst-performing controllers which have the greatest impact on business objectives.

Prioritizing controller repair activities based on performance alone could result in an economically unimpor-

Scope	Research Direction
Facility-wide	1. Multivariable time series-based model-free root cause analysis (e.g. subspace methods) 2. Time series feature extraction and pattern matching aligned along specific tuning and actuator failure modes using visual query language (VQL) techniques such as dynamic time warping (DTW) (Kassidas et al., 1998) 3. Causality analysis between infrequently collected data (e.g. lab samples, KPIs) and frequently collected data
MPC	1. Constraint handling ability 2. Identification of individual model(s) contributing to overall MPC instability using for example subspace methods, singular value analysis, or relative gain array (RGA) and its extensions 3. Dynamic model quality analysis (precision and accuracy, linearity, etc) 4. Inferred property bias, updating, and dynamic compensation problem diagnosis
Regulatory Control	1. Constraint handling (SISO MV saturation) 2. Oscillation characterization (waveforms, PV-OP bivariate analysis) 3. Identification of obvious tuning problems • Tight tuning causing oscillation • Loose tuning causing sluggishness • Inappropriate gain / integral / derivative values • Inappropriate PV filtering 4. Loop pairing analysis
Valves	1. Leveraging of new diagnostic information which is becoming available from smart valves 2. Trim wear detection 3. Low flow controllability characterization 4. Air supply problem detection

Table 12: Research Directions for diagnostic metrics.

tant controller being repaired before a better performing but economically important controller. Likewise prioritization based solely on impact will tend to narrow the focus to the small subset of controllers known (or rather perceived) to be economically important, which then receive attention whether their performance is poor or not. For example, a twenty-four inch gas flow service with a butterfly valve on an ethylene plant refrigeration system will have vastly different performance characteristics and business impact than a two-inch liquid flow service with a cage valve on a boiler condensate return line, yet if the repairs are prioritized based on performance measures alone, economic improvements may not be realized. If the objective of a PCMS is to change the priority of the control engineer's activities, then that priority should be based on metrics which consider both performance and impact.

Today, assessing controller impact requires expert knowledge of the process as well as qualitative experience with the controller's impact and interactions with other controllers and key business objectives. Only by understanding the actual role of each controller relative to unit and site objectives can controller impact be set. There is great benefit in simply combining today's available performance metrics with controller impact assessment. More work is needed, however, in the area of automatic detection of causal relationships between business objectives and individual controller performance so that performance-impact prioritization can be performed automatically.

Performance-impact prioritization is an important research area. Examples of possible applications include estimation of benefits or loss due to poor performance, shutdown planning, HAZOP assessments, and equipment reliability assessment.

Multivariate Assessment. Most multivariable control performance assessment research has focused on variability transfer performance (Harris et al., 1996; Huang and Shah, 1996, 1997). Multivariable variability

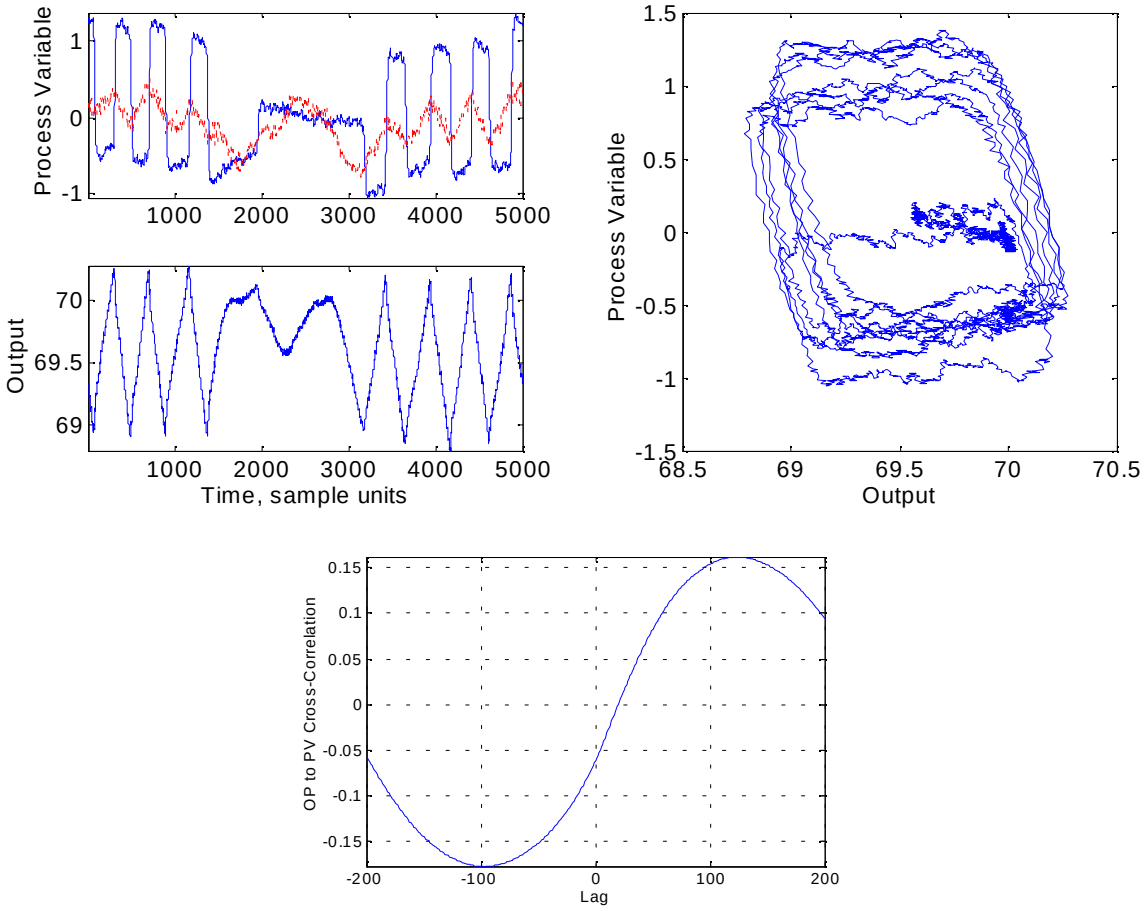


Figure 7: Control valve example 1.

transfer has turned out to be a complex subject and has not found its way to mainstream commercial controller monitoring applications. As stated in Section 5 (Constraints), most model predictive controllers are implemented as constraint-pushing optimizers. Regulation becomes an issue when the controller is fully unconstrained, which is rarely if ever the case. An area that has not been studied in much detail is multivariable control assessment in the context of economic optimization subject to constraints. Operators and engineers need better metrics to identify and diagnose MPC controllers that are failing to meet economic objectives in a safe manner (i.e. by satisfying mechanical and operability constraints). The primary methods at the operator's disposal to improve the economics of the controller are to change 1) the constraint limits and 2) the active set of controlled and manipulated variables. Performance and diagnostic metrics which help the operator decide when to make these changes would be of great value. The primary methods at the control engineer's disposal to improve the economics and dynamic operability of the controller are to change 1) the controller aggressiveness through tuning and 2) one or more of the dynamic models in the control matrix. Performance and diagnostic metrics which help the control engineer decide when to make these changes would be of great value.

Non-regulatory Objectives and Integrating Processes. About two thirds of level controllers have a surge attenuation objective. Failure to recognize the true objective of level controllers is common, often resulting in overly-aggressive tuning that propagates process variability downstream of the surge vessel. Most of the metrics available are either not appropriate or limited to non-integrating processes. In the authors' experience, there is currently a disproportionate fraction of assessment error in level control compared with other measurement types. Even if the operating objective and context of the controller is known, automated assessment of level controllers is challenging. In most facilities, vessel geometry is only documented on P&ID's, PFD's, or paper specification sheets—if at all. The effort of obtaining this information is non-trivial. A

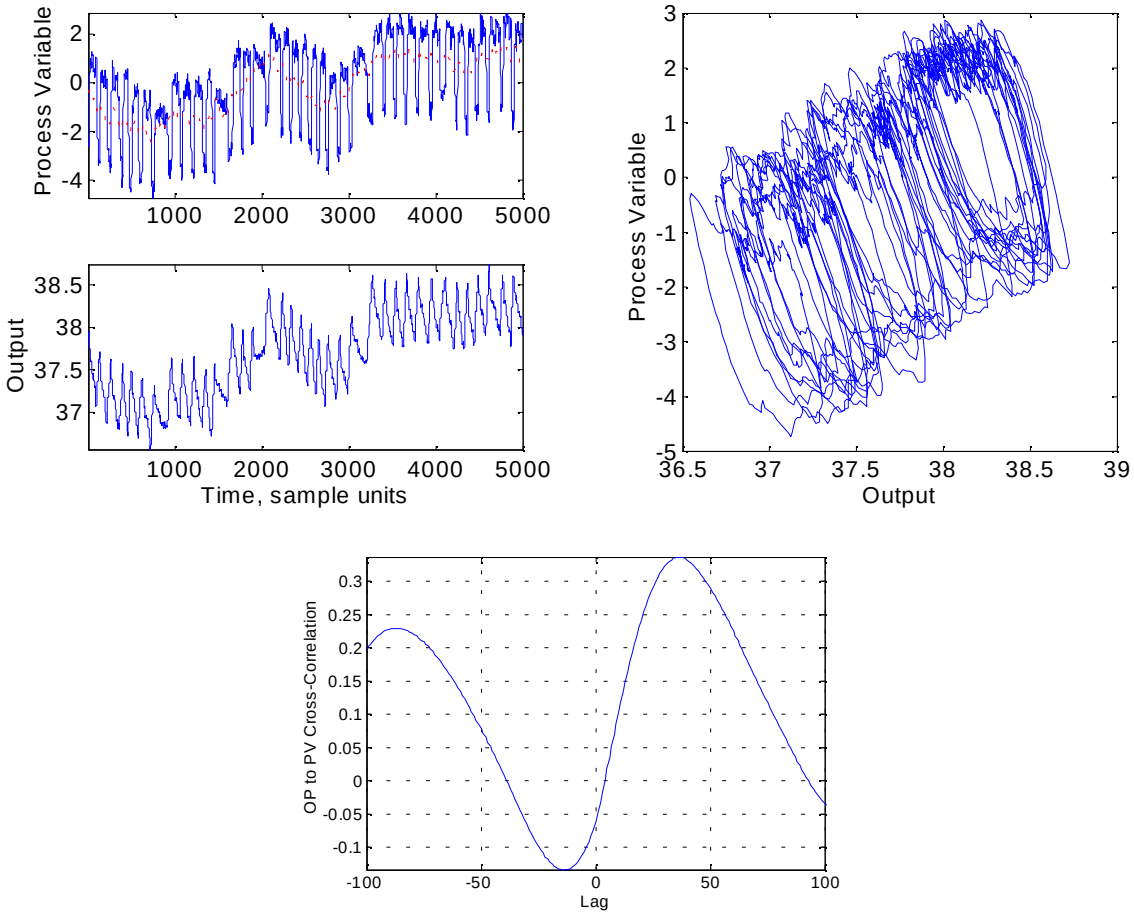


Figure 8: Control valve example 2.

level performance assessment solution that does not require a model or vessel geometry is far more likely to be adopted in industry. Research that specifically assesses level controllers and non-regulatory objectives is therefore of practical value.

Valve faults in level controllers are also very difficult to diagnose because the process is generally integrating. The same time series and PV versus OP patterns that clearly show valve problems for flow and pressure controllers are unclear in level processes (Figures 11 and 12).

Summary and Conclusions

Studies have shown that only about one third of industrial controllers provide an acceptable level of performance (Ender, 1993; Bialkowski, 1993). Furthermore, this performance has not improved in the past seven years (Miller, 2000), even though many academic performance measures have been developed in that time (Harris et al., 1999).

Over the past three years the authors have gathered input from hundreds of industrial practitioners of controller performance assessment and in many cases have directly observed their work practices and current Process Control Monitoring Systems. The authors have also developed a successful commercial PCMS designed to address the needs identified by industrial practitioners (Loop Scout™).

The current landscape of industrial process control contains some key considerations for developers of Process Control Monitoring Systems:

- Practicing control engineers desire a PCMS which is simple to setup, maintain, and use, and allows information to be found quickly and easily;
- Real time, high frequency time series data collection and automatic analysis is difficult and time consuming;
- Legacy control systems weren't designed for performance monitoring hence many are not up to the task from a computing horsepower perspective;
- Getting data from the legacy control system to a more powerful computing platform is limited by the

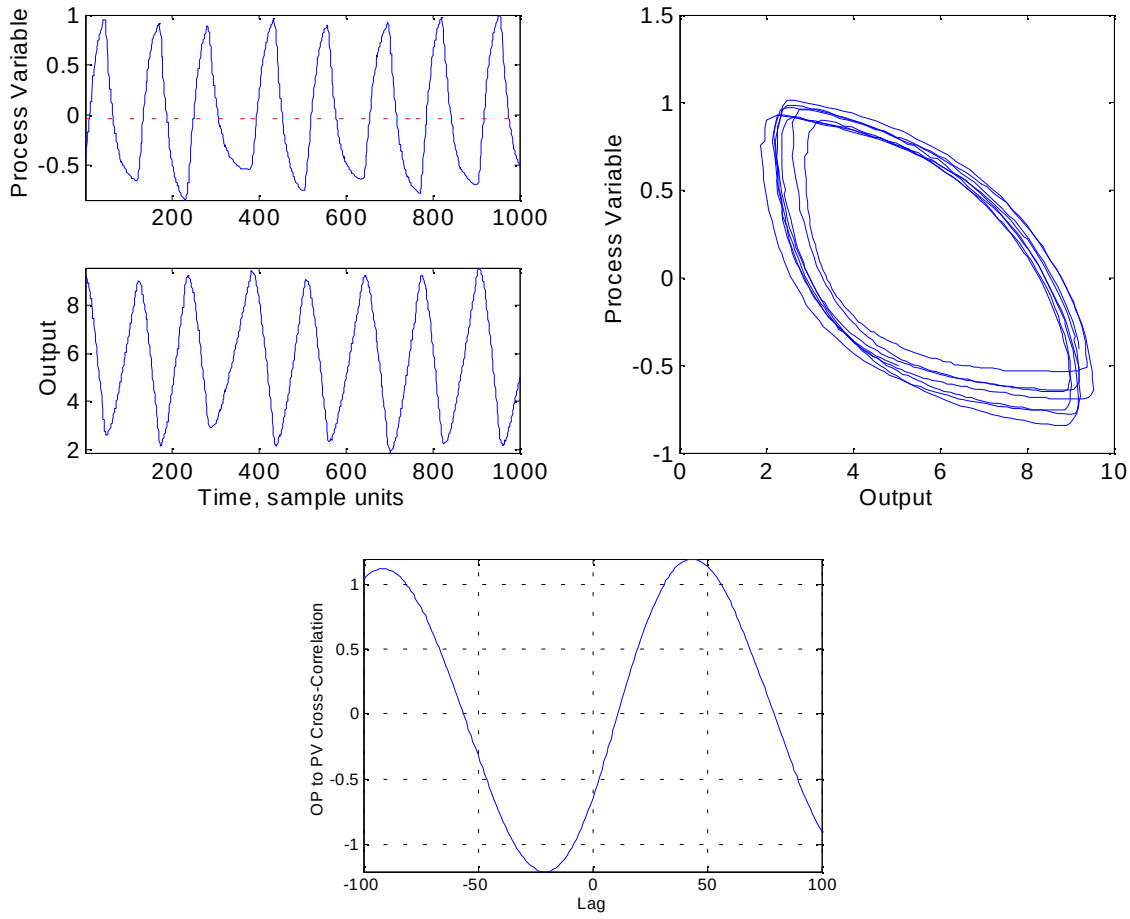


Figure 9: Control valve example 3.

available bandwidth;

- Dynamic process models are unavailable for the vast majority of controllers, and would be prohibitively expensive to obtain;
- Every process in the continuous process industries is in some way unique and as a result higher-performance algorithms are rejected in lieu of the PID algorithm which is easier to implement and support, and as a result is used 97% of the time;
- MPC is usually implemented with the objective of constraint-pushing optimization rather than multivariable regulation;
- MPC performance problems are usually caused by the way the controllers are operated;
- Typical MPC maintenance problems are complex and multifaceted, requiring a holistic diagnostic approach, often relying on process insight and other tacit knowledge.

Although there has been a great deal of academic work in the area of controller performance assessment (see Harris et al., 1999, and the references contained therein), there is still a great deal of work to be done. In particular, the following areas deserve special emphasis and consideration:

- A PCMS must facilitate the orient-decide-act-improve workflow for business, operational, engineering, and maintenance stakeholders, but especially the process control engineer;
- Of the four phases of the orient-decide-act-improve workflow, orientation has received the least amount of research attention but is actually the most important to the industrial process control engineer;
- A passive method that can reliably and automatically classify valve performance in closed loop is a desperately needed component in the orientation phase;
- More work is needed in the area of automatic detection of causal relationships between business objectives and individual controller performance so that performance-impact prioritization can be performed automatically;

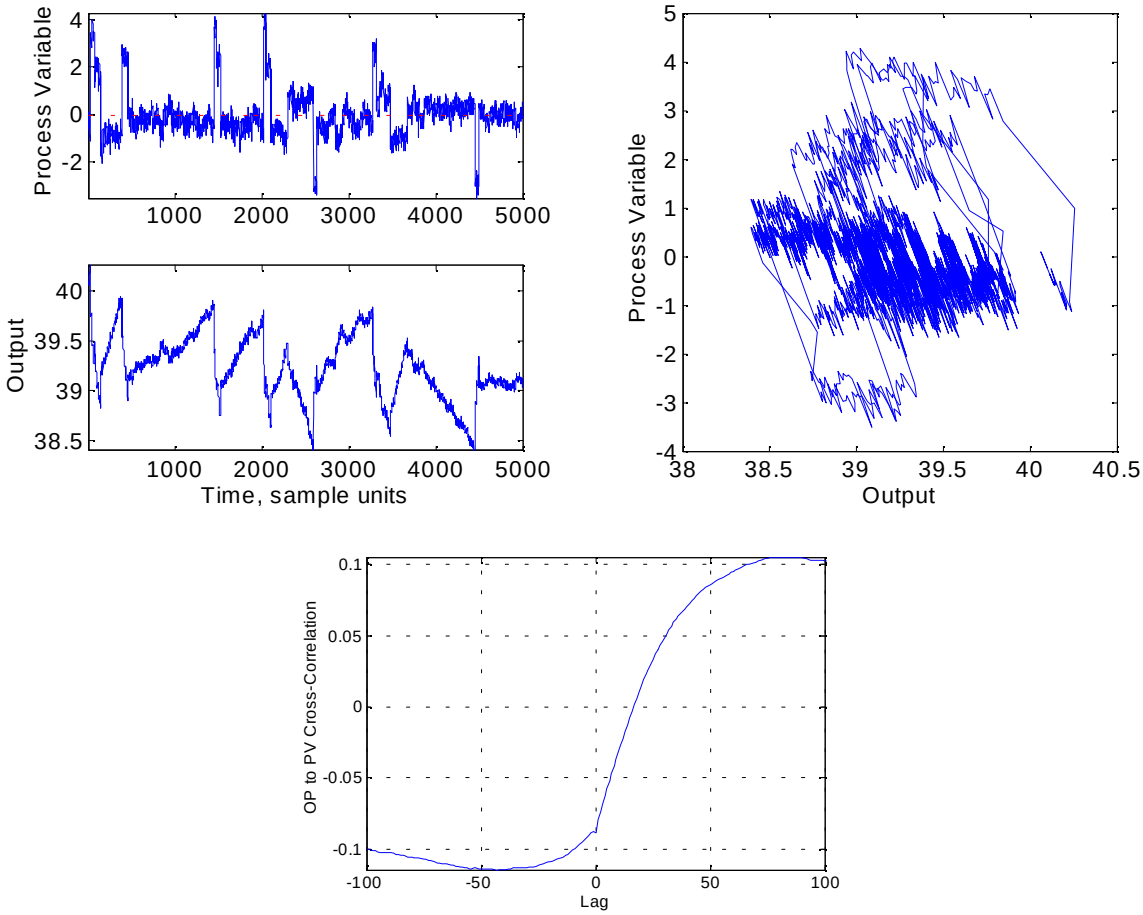


Figure 10: Control valve example 4.

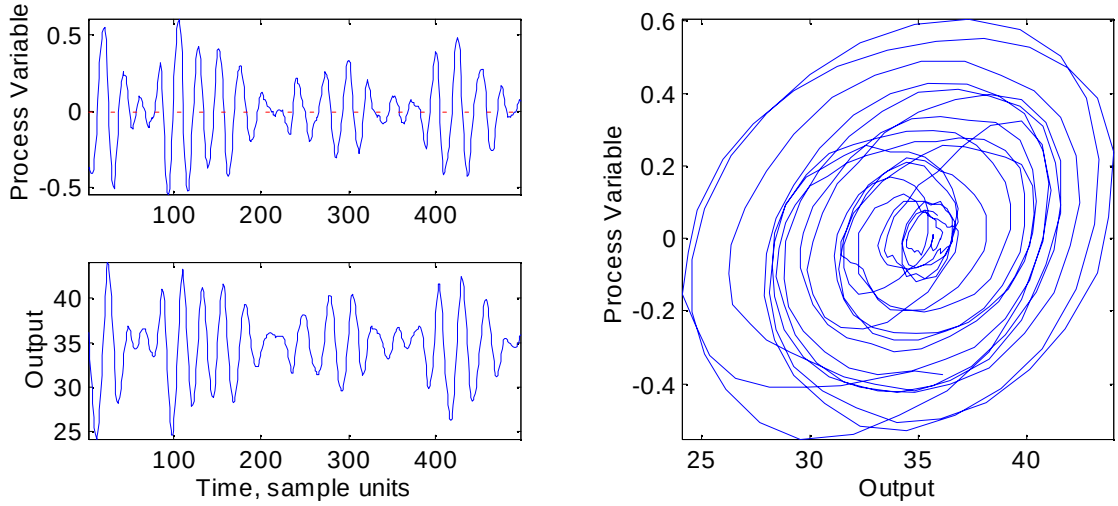


Figure 11: Level control example 1.

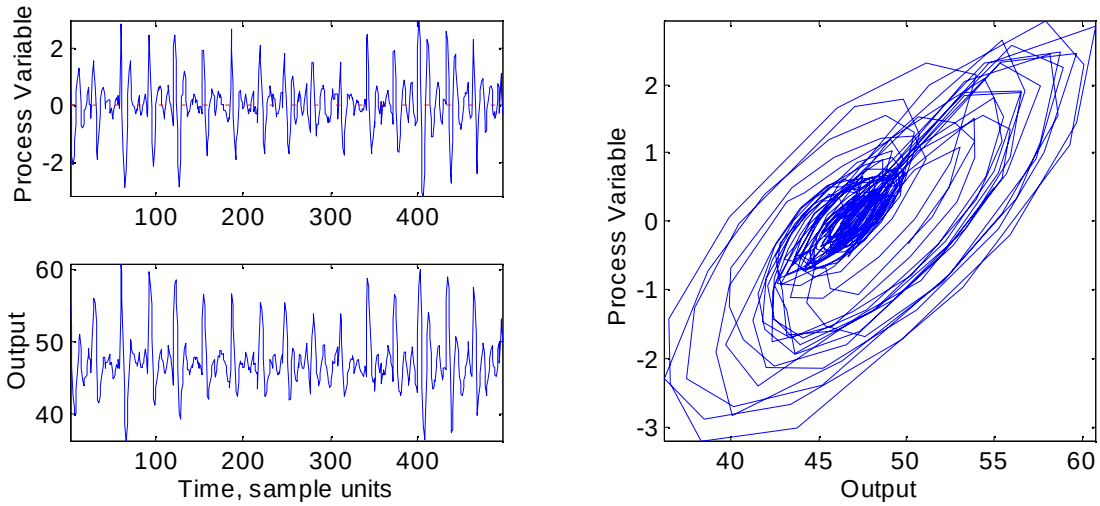


Figure 12: Level control example 4.

- Operators and engineers need better metrics to identify and diagnose MPC controllers that are failing to meet economic objectives in a safe manner (i.e. by satisfying mechanical and operability constraints);
- Research that specifically assesses level controllers and other controllers with non-regulatory objectives is required.

In summary, the industrial process control engineer is in an unenviable position. There will always be more work for them to do than time available to do it; time is their most precious resource. Process Control Monitoring Systems which automatically orient engineers to the likely location of the most economically important controller problems and then facilitate diagnosis and resolution of that controller's problems will play a vital role in increasing their effectiveness and hence their facility's effectiveness.

Acknowledgments

The authors wish to acknowledge the support of Honeywell Loop Scout™ customers, as well as Perry Nordh.

Glossary

DCS distributed control system

DTW dynamic time warping (Kassidas et al., 1998)

HAZOP Hazard and Operability Assessment

KPI key performance indicator

LP linear program

MV manipulated variable

MPC model predictive control

OLE object linking and embedding

OPC OLE for Process Control

OP output of controlled variable; signal sent to final control element (e.g. valve)

parameter an instance of a measurement associated with a point, e.g. TC101.PV or a configured attribute of that point, e.g. TC101.GAIN

PCMS Process Control Monitoring System

point database entity containing associated information about a controller, e.g. TC101

PFD process flow diagram

PID proportional, integral, derivative control algorithm

P&ID process and instrumentation diagram

PV process value of controlled variable—typically expressed in engineering units, e.g. kg/hr

RTO real time optimization

SP setpoint of controlled variable—typically expressed in engineering units, e.g. kg/hr

tagname see point

VOC voice of the customer

References

American Petroleum Institute, *Technology Vision 2020: A Technology Vision for the US*. API (1999).

American Petroleum Institute, *Draft Technology Roadmap for the Petroleum Industry*. API, Washington, DC (2000).

Ancrum, R., "Control Valve Diagnostics: Ready for Prime Time," *Control*, pages 60–62 (1996a).

Ancrum, R., "Control Valve Diagnostics: Ready for Prime Time Part 2," *Control*, pages 56–62 (1996b).

Anderson, J., A. Gokhale, and P. Mundy, Robust Technologies Reduce Implementation Costs & Maintenance Costs of a Multivariable Controller on a Crude Unit, In *Proc. Petrotech, Bahrain* (1998).

ARC Advisory Group, Simulation & Model-Based Control Software Global Outlook (1998).

ARC Advisory Group, PAS Worldwide Outlook (2000a).

ARC Advisory Group, Real Time Process Optimization and Training Worldwide Outlook (2000b).

Bialkowski, W. L., "Dreams Versus Reality: A View From Both Sides of the Gap," *Pulp and Paper Canada*, **94**(11), 19–27 (1993).

Boyd, J. R., A Discourse on Winning and Losing, A collection of unpublished briefings and essays. Air University Library, Document No. M-U 43947 (1987).

Boyle, S., "Test Systems for Evaluating the Dynamic Performance of Final Control Elements," *Proc. ISA*, pages 673–681 (1996).

Burchill, G. and C. Hepner Brodie, *Voices Into Choices: Acting on the Voice of the Customer*. Oriel Incorporated (1997).

Campbell Brown, D., "Alarm Management: A problem worth taking seriously," *Control Magazine* (1999).

Caplice, C. and Y. Sheffi, "A Review and Evaluation of Logistic Metrics," *The International Journal of Logistics Management*, **5**(2) (1994).

Chang, T., "Worldwide Refining Capacity Creeps Upward, Most Growth in Asia-Pacific," *Oil and Gas J.*, pages 41–90 (1999).

Desborough, L. D. and P. Nordh, Control Performance Monitoring Voice of the Customer Study, Honeywell Hi-Spec Solutions, unpublished manuscript (1998).

Desborough, L., R. Miller, and P. Nordh, Regulatory Control Survey, Honeywell, unpublished manuscript (2000).

Ender, D., "Process Control Performance: Not as good as You Think," *Control Engineering*, pages 180–190 (1993).

EnTech, *Control valve dynamic specification. Version 3.0*. EnTech Control Inc., Toronto, Canada (1998).

Fedenczuk, P., P. Fountain, and R. Miller, Loop Scout RPID and Profit Controller team up to produce significant benefits for BP, Honeywell IAC Users Group (1998).

Fitzgerald, W., "Automated Troubleshooting of Pneumatically Operated Control Valves," *Proc. ISA*, pages 43–52 (1988).

Fitzgerald, W., "Automated Control Valve Troubleshooting, the Key to Optimum Valve Performance," *Proc. ISA*, pages 75–98 (1990).

Hardin, M. B., R. Sharum, A. Joshi, and J. D. Jones, Rigorous Crude Unit Optimization at Conoco's Lake Charles Refinery, NPRA Computer Conference (1995).

Harris, T. J., F. Boudreau, and J. F. MacGregor, "Performance Assessment of Multivariate Feedback Controllers," *Automatica*, **32**(11), 1505–1518 (1996).

Harris, T. J., C. T. Seppala, and L. D. Desborough, "A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems," *J. Proc. Cont.*, **9**, 1–17 (1999).

Horch, A., *Extensions of a Static Performance Index and Detection of Static Friction in Process Controllers*, PhD thesis, Royal Institute of Technology (1998).

Huang, B. and S. L. Shah, Performance Limits: Practical Control Loop Performance Assessment, Presented at the 1996 AIChE Annual Meeting (1996).

Huang, B. and S. L. Shah, "Performance Assessment of Multivariate Control Loops on a Paper-Machine Headbox," *Can. J. Chem. Eng.*, **75**, 134–142 (1997).

Industrial Information Resources, PEC Report (1999).

Kassidas, A., P. A. Taylor, and J. F. MacGregor, "Off Line Diagnosis of Deterministic Faults Diagnosis in Continuous Dynamic Multivariate Processes Using Speech Recognition Methods," *J. Proc. Cont.*, **8**, 381–393 (1998).

Katzer, J. R., M. P. Ramage, and A. V. Sapre, "Petroleum Refining: Poised for Profound Changes," *Chem. Eng. Prog.*, pages 41–51 (2000).

Meszler, D., "Case Grows for Climate Change," *Chem and Eng. News*, page 16 (1999).

Miller, R., Loop Scout Regulatory Control Performance Study, Honeywell, unpublished manuscript (2000).

Panchal, C. B. and H. Ehr-Ping, Effects of Fouling Mitigation on the Energy Efficiency of Crude Oil Distillation, Presented at AIChE Spring meeting (1998).

Pyzdek, T., *The Complete Guide to Six Sigma*. Quality Press (2000).

Qin, S. J. and T. A. Badgwell, An overview of industrial model predictive control technology, In Kantor, J. C., C. E. García, and B. Carnahan, editors, *Proceedings of Chemical Process Control—V*, pages 232–256. CACHE, AIChE (1997).

Sharif, M. and R. Grosvenor, Fault Diagnosis in Industrial Control Valves and Actuators, In *IEEE Instrumentation and Measurement Technology Conference*, pages 770–778 (1998).

Sorensen, C. and C. R. Cutler, "LP integrates economics into dynamic matrix," *Hydrocarbon Process.*, **77**(9) (1998).

Studebaker, P., "Connectivity Reigns: Converging on the Next Millenium, Developments in Open Systems, Fieldbus, and Distributed Intelligence Promise New Architectures, Opportunities, and Challenges," *Control Magazine* (1999).

Takada, H., "Process Health Monitoring Applications in Chemical Facilities," *Insights 98* (1998).

Trimble, D., How to Measure Success: Uncovering The Secrets Of Effective Metrics, `http://www.prosci.com/metrics.htm` (2000).

US Department of Energy, Manufacturing Consumption of Energy Survey 1994, DOE/EIA-0512(94) (1997). `ftp://ftp.eia.doe.gov/pub/pdf/consumption/051294.pdf`.

Wallen, A., Valve Diagnostics and Automatic Tuning, In *Proc. ACC*, pages 2930–2934 (1997).

Appendix

Energy Savings from Improved Controller Performance

1. 1994 United States energy consumption statistics:

SIC code	Industry	Trillion BTU/yr	Number of Facilities
26	Paper and Allied Products	2665	584
28	Chemicals and Allied Products	5328	2994
2911	Petroleum Refining	6263	246
33	Primary Metal Industries	2462	1453
	Total	16718	5277

Source: US DOE, 1994 `http://www.eia.doe.gov/emeu/mecs/mecs94/consumption/mecs5.html`
A conservative estimate is that these industries in 1999 consumed 15×10^9 MBTUs of energy.

2. In 1996, on a dollars-per-million-Btu basis, petroleum was the most expensive fossil fuel ($3.16), natural gas was second ($2.64), and coal was least expensive ($1.29).
Source: `http://www.eia.doe.gov/neic/infosheets96/Infosheet96.html`
A conservative estimate is that energy in 1999 cost $2/MBTU.

3. It is very common to quote energy savings of 1–4% through implementation of advanced control and other process control technologies
Source: `http://www.foxboro.com/industries/gas/`
A conservative estimate is that improvement of existing controllers through enhanced Process Control Monitoring Systems could reduce energy costs in the process industries by 1%.

4. **Process Industry Energy Savings**
   = Energy Consumption × Energy Cost × Energy Savings from Improved Control
   = 15E9 MBTU/yr × $2/MBTU × 1%
   = **300 Million Dollars per Year**

Multivariate Controller Performance Analysis: Methods, Applications and Challenges

Sirish L. Shah*, Rohit Patwardhan† and Biao Huang
Department of Chemical and Materials Engineering
University of Alberta
Edmonton, Canada T6G 2G6

Abstract

This paper provides a tutorial introduction to the role of the time-delay or the interactor matrix in multivariate minimum variance control. Minimum variance control gives the lowest achievable output variance and thus serves as a useful benchmark for performance assessment. One of the major drawbacks of the multivariate minimum variance benchmark is the need for *a priori* knowledge of the multivariate time-delay matrix. A graphical method of multivariate performance assessment known as the Normalized Multivariate Impulse Response (NMIR), *that does not require knowledge of the interactor*, is proposed in this paper. The use of NMIR as a performance assessment tool is illustrated by application to two multivariate controllers. Two additional performance benchmarks are introduced as alternatives to the minimum variance benchmark, and their application is illustrated by a simulated example. A detailed performance evaluation of an industrial MPC controller is presented. The diagnosis steps in identifying the cause of poor performance, e.g. as due to model-plant mismatch, are illustrated on the same industrial case study.

Keywords

Multivariate minimum variance control, Time delay, Normalized multivariate impulse response, Model predictive control, Model-plant mismatch

Introduction

The area of performance assessment is concerned with the analysis of existing controllers. Performance assessment aims at evaluating controller performance from routine data. The field of controller performance assessment stems from the need for optimal operation of process units and from the need of getting value from immense volumes of archived process data. The field has matured to the point where several commercial algorithms and/or vendor services are available for process performance auditing or monitoring.

Conventionally the performance estimation procedure involves comparison of the existing controller with a theoretical benchmark such as the minimum variance controller (MVC). Harris (1989) and co-workers (1992; 1993) laid the theoretical foundations for performance assessment of single loop controllers from routine operating data. Time series analysis of the output error was used to determine the minimum variance control for the process. A comparison of the output variance term with the minimum achievable variance reveals how well the controller is doing currently. Subsequently Huang et al. (1996; 1997) and Harris et al. (1996) extended this idea to the multivariate case. In contrast to the minimum variance benchmark, Kozub and Garcia (1993), Kozub (1997) and Swanda and Seborg (1999) have proposed user defined benchmarks based on settling times, rise times, etc. Their work presents a more practical method of assessing controller performance. A suitable reference settling time or rise time for a process can often be chosen based on process knowledge.

The increasing acceptance of the idea of process and performance monitoring has also grown from the awareness that control software, and therefore the applications that arise from it, should be treated as capital assets and thus maintained, monitored and revisited routinely. Routine monitoring of controller performance ensures optimal operation of regulatory control layers and the higher level advanced process control (APC) applications. Model predictive control (MPC) is currently the main vehicle for implementing the higher level APC layer. The APC algorithms include a class of model based controllers which compute future control actions by minimizing a performance objective function over a finite prediction horizon. This family of controllers is truly multivariate in nature and has the ability to run the process close to its limits. It is for the above reasons that MPC has been widely accepted by the process industry. Various commercial versions of MPC have become the norm in industry for processes where interactions are of foremost importance and constraints have to be taken into account. Most commercial MPC controllers also include a linear programming stage that deals with steady-state optimization and constraint management. A schematic of a mature and advanced process control platform is shown in Figure 1. It is important to note that the bottom regulatory layer consisting mainly of PID loops forms the typical foundation of such a platform followed by the MPC layer. If the bottom layer does not perform and is not maintained regularly then it is futile to implement advanced control. In the same vein, if the MPC layer does not perform then the benefits of the higher level optimization layer, that may include real-time optimization, will not accrue.

The main contribution of this paper is in its general-

*sirish.shah@ualberta.ca
†Matrikon Consulting Inc., Suite 1800, 10405 Jasper Avenue, Edmonton, Canada , T5J 3N4

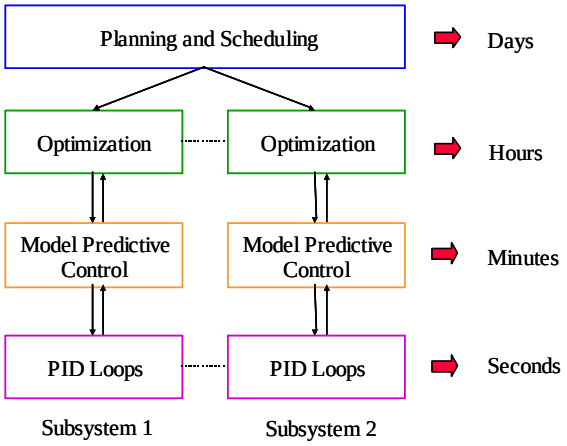


Figure 1: The control hierarchy.

ization of the univariate impulse response (between the process output and the whitened disturbance variable) plot to the multivariate case as the 'Normalized Multivariate Impulse Response' plot. A particular form of this plot, that does not require knowledge of the process time-delay matrix, is proposed here. Such a plot provides a graphical measure of the multivariate controller performance in terms of settling time, decay rates etc. Simple time and frequency domain measures such as multivariate autocorrelation and spectral plots are used to illustrate the interactions arising in multivariable control systems. Two relatively new multivariate performance evaluation ideas are also explored in detail: (1) the use of LQG as a benchmark based on the knowledge of the open loop process and noise models for the soft-constrained performance assessment problem (Huang and Shah, 1999) and (2) the use of the design performance as a benchmark (Patwardhan, 1999; Patwardhan et al., 2001). Both of these benchmarks can be applied to any type of controller. The LQG benchmark applies to all class of linear controllers, *irrespective of the controller objective function*, and is of use when input and/or output variance is of concern. The LQG benchmark represents the 'limits of performance' for a linear system, is more general and has the minimum variance as a special case. However, it needs a model of the linear process. The design objective function based approach can be applied to constrained MPC type controllers and is therefore a practical measure. However, it does not tell you how close the performance is relative to the lowest achievable limits. Issues related to the diagnosis of poor performance are discussed in the context of MPC controllers. Performance assessment of the general MPC is as yet an unresolved issue and presents a challenging research problem. A constrained MPC type controller is essentially a nonlinear controller, especially when operating at the constraints. Conventional MVC or linear controller benchmarking is infeasible and alternative techniques have to be developed. The development of new MPC performance monitoring tools thus represents an area of future challenges. The challenges associated with MPC performance evaluation are illustrated by considering an industrial case study of a 7×6 problem.

This paper is organized as follows. The next section provides a tutorial introduction to the concept of the time-delay matrix or the interactor. This is an important entity, particularly if one wants to evaluate MPC performance using multivariate minimum variance as a benchmark. The following two sections, respectively, discuss the tools required in the analysis of multivariate control loops such as the normalized multivariate impulse response, and alternative benchmarks for multivariate performance assessment. Applications are used to demonstrate the proposed techniques in each section. A discussion on the challenges in performance analysis and diagnosis and issues in MPC performance evaluation are outlined in the penultimate section, followed by a detailed industrial case study of an industrial MPC evaluation.

The Role of Delays for Univariate and Multivariate Processes

Time delays play a crucial role in performance assessment particularly when the minimum variance benchmark is used. The concept of the multivariate delay is explained below in a tutorial manner by first defining the univariate delay term and then generalizing this notion to the multivariate case.

Definition of a Delay Term for a Univariate Process:

The time-delay element in a univariate case is characterized by several different properties. For example, it represents the order of the first, non-zero (or *non-singular*) impulse response coefficient (also characterized by the number of infinite zeros of the numerator portion of the transfer function). It is important to fully understand the definition of a delay term for the univariate case in order to generalize the notion to a multivariate system. From a system theoretic point, the delay for a univariate system is characterized by the properties listed below. Consider a plant with the discrete transfer function or an impulse response model given by:

$$\begin{aligned} G(q^{-1}) &= \frac{q^{-d}B(q^{-1})}{A(q^{-1})} \\ &= 0q^{-1} + 0q^{-1} + \cdots + 0q^{-d+1} \\ &\quad + h_d q^{-d} + h_{d+1}q^{-d-1} + h_{d+2}q^{-d-2} \\ &\quad + \cdots \end{aligned}$$

The delay term for such a univariate system is defined by:

- the minimum integer r such that

$$\lim_{q^{-1}\to 0} q^r \left(\frac{q^{-d}B(q^{-1})}{A(q^{-1})}\right) = k \quad \neq 0$$

(i.e. a non-singular coefficient) which in the case considered above, for $r = d$ gives:

$$\lim_{q^{-1}\to 0} q^r \left(\frac{q^{-d}B(q^{-1})}{A(q^{-1})}\right) = h_d \quad \neq 0$$

Note that for the univariate case, the number of infinite zeros of the process as obtained by setting the numerator of the process transfer function to zero, i.e. $q^{-d}B(q^{-1}) = 0$, also yields d infinite zeros and n finite zeros given by the roots of $B(q^{-1}) = 0$.

- the static or steady-state value of the delay term should be equal to 1, i.e. at steady-state (when $q^{-1} = 1$), $q^{-d} = 1$.

Definition of a Delay Matrix for a Multivariate Process:

Analogous to the univariate case, it is possible to factorize the open-loop transfer matrix into two elements: the delay matrix, $D(q^{-1})^{-1}$ containing all the infinite zeros of the system, and the 'delay-free' transfer-function matrix, $T^*(q^{-1})$, containing all the finite zeros and poles.

$$\begin{aligned} T(q^{-1}) &= D(q^{-1})^{-1} \cdot T^*(q^{-1}) \\ &= H_1 q^{-1} + H_2 q^{-2} + \cdots \\ &\quad + H_d q^{-d} + H_{d+1} q^{-d-1} + \cdots \end{aligned}$$

where H_i are the impulse response or Markov matrices of the system parallel to the definition of the univariate delay, the multivariate delay matrix is defined by:

- Fewest number of linear combinations of the impulse response matrices that give a nonsingular matrix, i.e. (a finite and *nonsingular* matrix)

$$\lim_{q^{-1}\to 0} D(q^{-1})(D(q^{-1})^{-1} \cdot T^*(q^{-1})) = K$$

(a finite and nonsingular matrix)

Unlike the univariate case, a nonzero H_i may not necessarily indicate the delay order. Instead, for the multivariate case, it is the fewest linear combination of such non-zero H_i to give a non-singluar K that defines the delay matrix, $D(q^{-1})^{-1}$. Applying this idea to the univariate case will reveal that $K = h_d$, which is the first or leading non-zero coefficient of the impulse response or the Markov parameter of the scalar system. Such an interpretation makes the choice of $D(q^{-1})^{-1}$, as a multivariate generalization of the univariate delay term, a very meaningful one. Note that $\det(D(q)) = cq^m$, where m is the number of infinite zeros of the system and c is a constant. For the multivariate case the number of infinite zeros may not be related to the order d of the time-delay matrix.

- $D^T(q^{-1})D(q) = I$ (As compared to the univariate case where $q^{-d}q = 1$). This is known as the unitary interactor matrix. This unitary property preserves the spectrum of the signal, which leads us to the result that the variance of the actual output and the interactor filtered outputs are the same, i.e. $E(Y_t^T Y_t) = E(\tilde{Y}_t^T \tilde{Y}_t)$, where $\tilde{Y}_t = q^{-d}DY_t$ (see Huang and Shah, 1999).

Example. Consider the following transfer function matrix and its impulse response or Markov parameter model:

$$\begin{aligned} T(q^{-1}) &= \begin{bmatrix} \dfrac{q^{-2}}{1-q^{-1}} & \dfrac{q^{-3}}{1-2q^{-1}} \\ \dfrac{q^{-2}}{1-3q^{-1}} & \dfrac{q^{-3}}{1-4q^{-1}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} q^{-1} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} q^{-2} \\ &\quad + \begin{bmatrix} 1 & 1 \\ .33 & 1 \end{bmatrix} q^{-3} + \begin{bmatrix} 1 & 0.5 \\ 0.109 & 0.25 \end{bmatrix} q^{-4} \\ &\quad + \cdots \end{aligned}$$

Note that even though $H_2 \neq 0$, a linear combination of H_1 and H_2 does not yield a nonsingular matrix. In this example, at least three impulse response matrices are required to define the delay matrix for this system. The delay matrix that satisfies the properties listed above, is given by:

$$D(q^{-1})^{-1} = \begin{bmatrix} -0.707q^{-2} & -0.707q^{-3} \\ -0.707q^{-2} & 0.707q^{-3} \end{bmatrix}$$

The order of the delay is 3, i.e. a linear combination of at least 3 impulse response matrices is required to have a non-singular K.

Remark 1. The interactor matrix $D(q)$ can be one of the three forms as described in the sequel. If $D(q)$ is of the form: $D(q) = q^d I$, then the process is regarded as having a simple interactor matrix. If $D(q)$ is a diagonal matrix, i.e., $D(q) = diag(q^{d_1}, q^{d_2}, \ldots, q^{d_n})$, then the process is regarded as having a diagonal interactor matrix. Otherwise the process is considered to have a general interactor matrix.

To factor the general interactor matrix, *one needs to have a complete knowledge of the process transfer function or at least the first few Markov matrices of the multivariate system.* This is currently the main drawback in using this procedure. Huang et al. (1997) have provided a closed loop identification algorithm to estimate the first few Markov parameters of the multivariate system and thus compute the unitary interactor matrix. However,

this rank determination procedure is prone to errors as it requires one to check if a linear combination of matrices is of full rank or not. Ko and Edgar (2000) have also proposed the use of the first few Markov matrices for multivariate performance assessment based on the minimum variance benchmark. The factorization of the diagonal interactor matrix requires only time delays between each pair of the input and output variables. A diagonal interactor matrix by no means implies that the process has a diagonal transfer function matrix or that the process has no interaction. But the converse is true, i.e. a diagonal process transfer function matrix or a system with weakly interacting multivariate loops (a diagonally dominant system) does have a diagonal interactor matrix. In fact, experience has shown that many actual multivariable processes have the structure of the diagonal interactor, provided the input-output structuring has been done with proper engineering insight. This fact greatly simplifies performance assessment of the multivariate system.

The Multivariate Minimum Variance Benchmark

In a univariate system, the first d impulse response coefficients of the closed loop transfer function between the control error and the white noise disturbance term determine the minimum variance or the lowest achievable performance. In the same way, the first d impulse response matrices of the closed loop multivariate system are useful in determining the multivariate minimum variance benchmark, where d denotes the order of the interactor.

Performance assessment of univariate control loops is carried out, by comparing the actual output variance with the minimum achievable variance. The latter term is estimated by simple time series analysis of routine closed-loop operating data and knowledge of the process time delay. The estimation of the univariate minimum variance benchmark requires filtering and correlation analysis. This idea has been extended to multivariate control loop performance assessment and the multivariate filtering and correlation (FCOR) analysis algorithm has been developed as a natural extension of the univariate case (Huang et al., 1996, 1997; Huang and Shah, 1999). Harris et al. (1996) have also proposed a multivariate extension to their univariate performance assessment algorithm. Their extension requires a spectral factorization routine to compute the delay free part of the multivariate process and thus estimate the multivariate minimum variance or the lowest achievable variance for the process. The FCOR algortihm of Huang et al. (1996), on the other hand, is a term for term generalization of the univariate case to the multivariate case and also requires the knowledge of the multivariate time-delay or interactor matrix. Figure 2 summarizes the steps required in computing the mutivariate performance index. A quadratic measure of multivariate control loop performance is defined as:

$$J = E(Y_t - Y_t^{sp})^T (Y_t - Y_t^{sp})$$

(where Y_t represnts an n dimensional output vector).The lower bound or the quadratic measure of the multivariate control performance under minimum variance control is defined as

$$J_{\min} = E(Y_t - Y_t^{sp})^T (Y_t - Y_t^{sp})|_{mv}$$

It has been shown by Huang et al. (1997) that the lower bound of the performance measure $J_{\min}$can be estimated from routine operating data. In Huang et al. (1997), the multivariate performance index is defined as

$$\eta = \frac{J_{\min}}{J}$$

and is bounded by $0 \leq \eta \leq 1$. In practice, one may also be interested in knowing how each individual output (loop) of the multivariate system performs relative to multivariate minimum variance control. Performance indices of each individual output are defined as

$$\begin{bmatrix} \eta_{Y_1} \\ \vdots \\ \eta_{Y_n} \end{bmatrix} = \begin{bmatrix} \min(\sigma_{y_1}^2)/\sigma_{y_1}^2 \\ \vdots \\ \min(\sigma_{y_n}^2)/\sigma_{y_n}^2 \end{bmatrix} = diag(\tilde{\Sigma}_{mv}\tilde{\Sigma}_Y^{-1})$$

where $\tilde{\Sigma}_{mv} = diag(\Sigma_{mv})$ and $\tilde{\Sigma}_Y = diag(\Sigma_Y)$; Σ_Y is the variance matrix of the output Y_t and $\Sigma_{mv} = \min(\Sigma_Y)$ is the covariance matrix of the output Y_t under multivariate minimum variance control. It has been shown by Huang et al. (1997) that Σ_{mv} can also be estimated using routine operating data, and knowledge of the interactor matrix.

To summarize, a multivariate performance index is a single scalar measure of multivariate control loop performance relative to multivariate minimum variance control. Individual output performance indices indicate performance of each output relative to the loop's performance under multivariate minimum variance control. If a particular output index is smaller than other output indices, then some of the other loops may have to be de-tuned in order to improve this poorly tuned loop.

Alternative Methods for Performance Analysis of Multivariate Control Systems

Autocorrelation Function:

The autocorrelation function (ACF) plots may be used to analyze individual process variable performance. A typical example of the ACF plots for the two output variables of the simulated Wood-Berry (Wood and Berry, 1973) column control system is shown in Figure 3. A decentralized control system comprising two PI controllers was

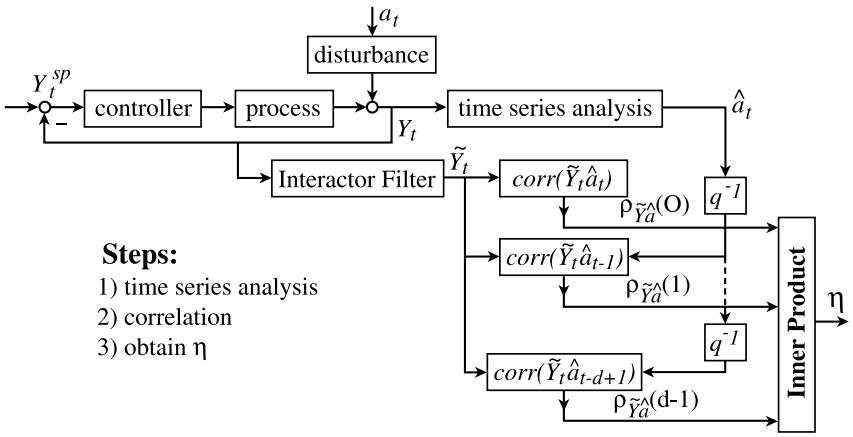


Figure 2: Schematic diagram of the multivariate FCOR algorithm.

used on the Wood-Berry column. The diagonal plots are autocorrelations of each output variable, while the off-diagonal plots are cross-correlation plots. The diagonal plots typically indicate how well each loop is tuned. For example, a slowly decaying autocorrelation function implies an under-tuned loop, and an oscillatory ACF typically implies an over-tuned loop. Off-diagonal plots can be used to trace the source of disturbance or the interaction between each process variables. Figure 3 clearly indicates that the first loop has relatively poor performance while the second loop has very fast decay dynamics and thus good performance. Interaction between the two loops can also be observed from the off-diagonal subplots. Note that the autocorrelation plot of the multivariate system is not necessarily symmetric.

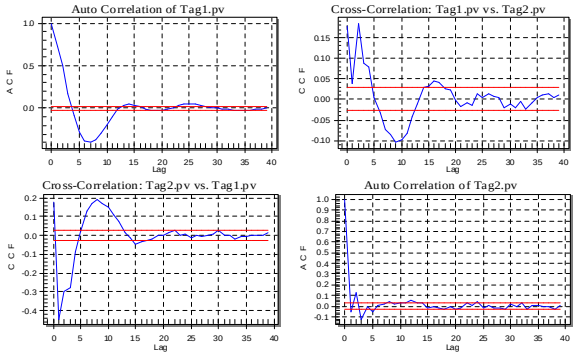


Figure 3: Autocorrelation function of multivariate process.

Spectral Analysis:

Frequency domain plots provide alternative indicators of control loop performance. They may be used to assess individual output dynamic behavior, interactions and effects of disturbances. For example, peaks in the diagonal plots typically imply oscillation of the variables due to an over-tuned controller or presence of oscillatory disturbances. Frequency domain plots also provide information on the frequency ranges over which the oscillations occur and the amplitude of the oscillations. Like time domain analysis, off-diagonal plots provide one with information on the correlation or interaction between the loops. Figure 4 is the power spectrum and the cross-power spectrum plot of the simulated Wood-Berry column. The first diagonal plot indicates that there is a clear mid-frequency harmonic in the 1st output. This could be due to an overtuned controller. Off-diagonal plots show a peak in the cross-spectrum at the same frequency. The poor performance of loop 1 can then be attributed to significant interaction effects from loop 2 to loop 1. In other words, the satisfactory or good performance of loop 2 could be at the expense of transmitting disturbances or upsets to loop 1 via the interaction term.. The power spectrum plots of a multivariate system are symmetric.

Normalized Multivariate Impulse Response (NMIR) Curve as an Alternative Measure of Performance:

As shown in Figure 2, the evaluation of the multivariate controller performance has to be undertaken on the interactor filtered output and not on the actual output. The reason for this is that the interactor filtered output vector, $\tilde{Y}_t = q^{-d}DY_t$, is a special linear combination of the actual output, lagged or otherwise, and this fictitious output preserves the spectral property of the system and facilitates simpler analysis of the multivariate minimum variance benchmark.

This new output ensures, as in the univariate case, that the closed loop output can be easily factored into two terms, a controller or feedback-invariant term and a second term that depends on the controller parameters. In the ensuing discussion, we consider an alter-

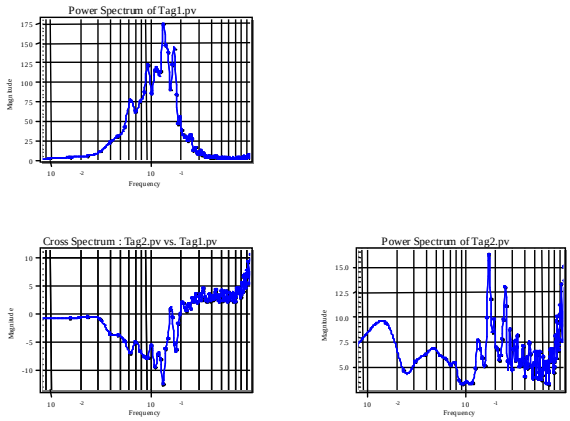


Figure 4: Frequencey response of multivariate process.

native graphical measure of multivariate performance as obtained from the interactor filtered output. So unless specified otherwise, the reader should assume that the operations elucidated below are on the interactor filtered output, Y_t.

An impulse response curve represents dynamic relationship between the whitened disturbance and the process output. This curve typically reflects how well the controller regulates stochastic disturbances. In the univariate case, the first d impulse response coefficients are feedback controller invariant, where d is the process time-delay. Therefore, if the loop is under minimum variance control, the impulse response coefficients should be zero after $d-1$ lags. The Normalized Multivariate Impulse Response (NMIR) curve reflects this idea. The first d NMIR coefficients are feedback controller invariant, where d is the order of the time-delay matrix or the interactor. If the loop is under multivariate minimum variance control, then the NMIR coefficients should decay to zero after $d-1$ lags. The sum of squares under the NMIR curve is equivalent to the trace of the covariance matrix of the data. In fact the NMIR is a graphical representation of the quadratic measure of the output variance as given by:

$$\begin{aligned} E(Y_t^T Y_t) &= E(\tilde{Y}_t^T \tilde{Y}_t) \\ &= tr(F_0 \Sigma_a F_0^T) + tr(F_1 \Sigma_a F_1^T) + \cdots \end{aligned}$$

where

$$\tilde{Y}_t = F_0 a_t + F_1 a_{t-1} + \cdots + F_{d-1} a_{t-d+1} + F_d a_{t-d} + \cdots$$

is an infinite series impulse response model of the interactor filtered output with respect to the whitened disturbance and matrices F_i represent the estimated Markov matrices of the closed loop multivariate system. In the new measure, the first NMIR coefficient is given by $\sqrt{tr(F_0 \Sigma_a F_0^T)}$, the second NMIR coefficient is given by $\sqrt{tr(F_1 \Sigma_a F_1^T)}$, and so on. The multivariate performance index is then equal to the ratio of the sum of the squares of the first d NMIR coefficients to the sum of squares of all NMIR coefficients (see top plot in Figure 5). Care has to be taken when interpreting the normalized impulse response curve. The NMIR represents a compressed scalar metric for a multi-dimensional system. It is a graphical representation of the weighted 2-norm multivariate impulse response matrix and provides a graphical interpretation of the multivariate performance index in much the same way as the univariate impulse response gives an indication of the level of damping afforded to a unit impulse disturbance.

The NMIR as outlined above and first described by Huang and Shah (1999) requires *a priori* knowledge of the interactor matrix. Since this NMIR curve is suitable for obtaining a graphical measure of the overall closed-loop response, we suggest an alternative measure that does not require knowledge of the interactor. We propose to use a similar normalized multivariate impulse curve without interactor filtering to serve a similar purpose. The NMIR *without interactor filtering* is calculated as before by computing the correlation coefficients between the pre-whitened disturbance and the actual output with lags $0, 1, 2, \ldots, d-1, d, d+1, \ldots$.

$$E(Y_t^T Y_t) = tr(E_0 \Sigma_a E_0^T) + tr(E_1 \Sigma_a E_1^T) + \cdots$$

where

$$Y_t = E_0 a_t + E_1 a_{t-1} + \cdots + E_{d-1} a_{t-d+1} + E_d a_{t-d} + \cdots$$

Note that unlike the original NMIR measure as proposed by Huang and Shah (1999), the new measure proposed here does *not* require interactor filtering of the output, i.e. an explicit knowledge of the interactor is not required in computing the new graphical and qualitative measure. From here onwards this new measure is denoted as NMIR$_{wof}$.

In the new measure, the first NMIR$_{wof}$ coefficient is given by $\sqrt{tr(E_0 \Sigma_a E_0^T)}$, the second NMIR$_{wof}$ coefficient is given by $\sqrt{tr(E_1 \Sigma_a E_1^T)}$, and so on. Note that the NMIR$_{wof}$ measure (without interactor filtering) is similar to NMIR with interactor filtering in the sense that both represent the closed-loop infinite series impulse response model of the output with respect to the whitened disturbance, one for the actual output and the other one for the interactor filtered output respectively. The newly proposed NMIR$_{wof}$ is physically interpretable, but does not have the property that the first d coefficients are feedback control-invariant. The main rationale for using the newly proposed NMIR$_{wof}$ is that the following two terms are asymptotically equal:

$$\begin{aligned} \lim_{n\to\infty} \Big\{ tr(E_0 \Sigma_a E_0^T) + tr(E_1 \Sigma_a E_1^T) + \cdots + tr(E_n \Sigma_a E_n^T) \Big\} = \\ \Big\{ tr(F_0 \Sigma_a F_0^T) + tr(F_1 \Sigma_a F_1^T) + \cdots + tr(F_n \Sigma_a F_n^T) \Big\} \end{aligned}$$

This follows from the equality: $E(Y_t^T Y_t) = E(\tilde{Y}_t^T \tilde{Y}_t)$ (Huang and Shah, 1999). It is then clear that the NMIR curves with and without filtering will coincide with each other for n sufficiently large. Alternately, the cumulative sum of the square of the impulse response coefficients can also be plotted and as per the above asymptotic equality, one would expect that the two terms will coincide for a sufficiently large n. These curves are reproduced here for the illustrative Wood-Berry column example. The ordinate in the bottom plot in Figure 5 gives the actual output variance when the curves converge for a sufficiently large n. Note that, unlike the NMIR curve with interactor filtering (solid line in Figure 5), the NMIR$_{wof}$ curve (dashed line in Figure 5) can *not* be used to calculate a numerical value of the performance index. However, it has the following important properties that are useful for assessment of multivariate processes:

1. The new NMIR$_{wof}$ represents the normalized impulse response from the white noise to the true output.
2. The new NMIR$_{wof}$ curve reflects the predictability of the disturbance in the original output. If the impulse response decays slowly, then this is clear indication of a highly predictable disturbance (e.g. an integrated white noise type disturbance) and relatively poor control. On the other hand a fast decaying impulse response is a clear graphical indication of a well-tuned multivariate system (Thornhill et al., 1999).
3. The new NMIR$_{wof}$ also provides a graphical measure of the overall multivariate system performance with information regarding settling time, oscillation, speed of response etc.

NMIRs with and without interactor filtering are calculated for the simulated Wood-Berry distillation column with two multiloop PI controllers. With sampling period 0.5 second, the interactor matrix of the process is found to have a diagonal structure and is given by

$$D(q) = \begin{bmatrix} q^3 & 0 \\ 0 & q^7 \end{bmatrix}$$

Since the order of the interactor is 7 sample units, the first 7 NMIR coefficients are feedback control invariant and depend solely on the disturbance dynamics and the interactor matrix. The sum of squares of these 7 coefficients is the variance achieved under multivariate minimum variance control. In fact as shown in Figure 6, the scalar multivariate measure of performance is equal to the sum of squares of the first 7 NMIR coefficients divided by the total sum of squares. Notice that for sufficiently large n or prediction horizon, the two curves, as expected, coincide with each other. In this simulation example, observe that the NMIR and the NMIR$_{wof}$ curves decay to zero relatively quickly after 7 sample units, indicating relatively good control performance.

Industrial MIMO Case study 1: Capacitance Drum Control Loops at Syncrude Canada Ltd. Capacitance drum control loops of Plant AB in Syncrude Canada Ltd. were analyzed for this study. The primary objective of Plant AB is to further reduce the water in the (Plant A) diluted bitumen product prior to it reaching the next plant (Plant B) storage tanks. As the grade of the feed entering plant A reduces, the water required to process the oilsands increases proportionately. A large portion of this excess water ends up in the Plant A froth feed tank and ultimately increases the % volume of water in the Plant A product. Aside from degrading the quality of the product, the increased volume of water means reduction in the amount of bitumen that can be piped to the diluted bitumen tanks. In addition, the higher water content means more of the chloride compound present in the oilsands is dissolved and finds its way to the diluted bitumen tanks and eventually to plant B. The higher chloride concentration increases the corrosion rate of equipment in the Upgrading units. Plant AB was developed as a means of reducing the water content, and ultimately, the chlorides sent to Upgrading. This reduction is achieved by centrifuging the Plant A Product.

All product from plant A is directed to the Plant AB feed storage tank. The IPS portion of the product is routed through 5 Cuno Filters prior to entering the Plant AB feed storage tank. Feed from the feed storage tank is then pumped through the feed pumps to the Alfa Laval centrifuges. The Alfa Laval centrifuges remove water and a small amount of solids from the feed. Each centrifuge has its own capacitance drum and product back pressure valve. This arrangement allows for individual centrifuge "E-Line" control and a greatly improved product quality. Heavy phase water from plant A is used as Process Water in plant AB.

The cap drum pressure controller controls the capacitance drum pressure by adjusting the nitrogen flow into the drum. The Cap Drum Primary level controller maintains the cap drum water level by adjusting water addition into the drum. Control of these two variables is essential to maintain the E-Line in the centrifuges. Currently these two loops are controlled by multiloop PID controllers. The two process variables, pressure and level, are highly interacting. The objective of the performance assessment is to evaluate the existing multiloop PID controllers' performance, and to identify opportunities, if any, to improve performance by implementing a multivariate controller

Discussion of Performance Analysis. Process data with a 5-second sampling interval are shown in Figure 7 . These are typical (representative) process data encoun-

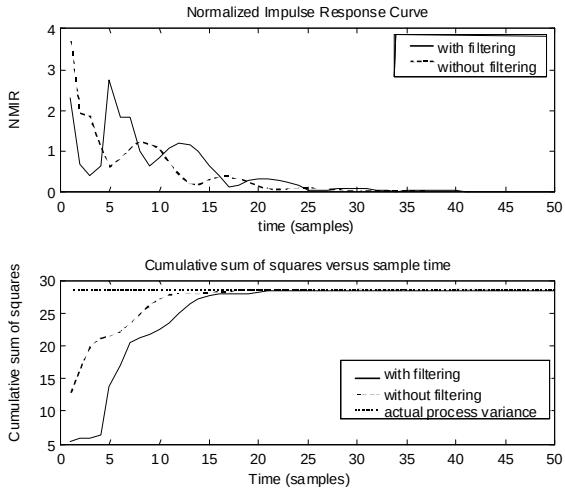


Figure 5: NMIR and NMIR$_{wof}$ curves (top plot) and the cumulative sum of squares plots of the impulse response coefficients (bottom plot).

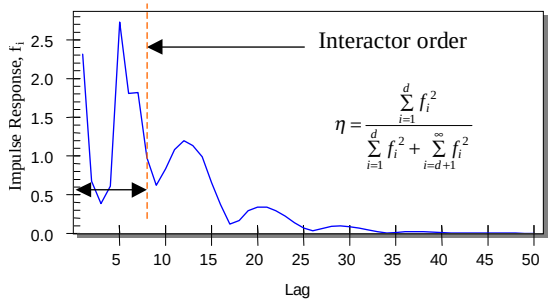


Figure 6: Normalized multivariate impulse response.

tered in this process. By assuming that both pressure and level loops have no time delay except for the delay induced by the zero-order-hold device, a scalar multivariate performance index was calculated as 0.022 and individual output indices are shown in Figure 8 . Based on these indices, one may conclude that controller performance is poor and may be improved significantly by re-tuning the existing controllers or re-designing a multivariate controller. However, since the exact time delays for these loops are unknown, further analysis of performance in both time domain and frequency domain is necessary. For example, the NMIR$_{wof}$ response shown in Figure 9 does indicate that the disturbance persists for about 50 samples before it is compensated by the feedback controllers. Overall the decay in the NMIR$_{wof}$ curve is rather slow indicating a predictable disturbance and generally ineffective regulatory control in dealing with such disturbances. This is equivalent to a settling time of 4 minutes for the overall system. To check which loop causes such long settling time, one can look at the auto- and cross-correlation plots.

The individual loop behavior can be observed from the auto and cross correlation plots shown in Figure 10 . It is observed that the pressure response does not settle down even after 40 samples. This is clearly unacceptable for a pressure loop. In addition some oscillatory response is observed in the level response as evident from the spectrum plot shown in Figure 11. Notice that the peak (oscillation) appears in both the pressure and level responses as well as in the cross spectrum plot. This indicates that both loops interact and oscillate at the same frequency. Thus, this analysis indicates that 1) the existing multiloop controller has relatively poor performance primarily due to the long settling time and oscillatory behavior or presence of oscillatory disturbances; 2) the two loops are strongly interacting and a multivariate controller may be able to improve performance significantly.

The final recommendation for this system was that performance of the two loops individually as well as a multivariate system is relatively poor. For predictable disturbances, there is insufficient integral action in the pressure loop resulting in a slowly decaying ACF plot as noticeable in the top left hand plot in Figure 10 . The performance is poor mainly due to interaction between the two loops. Because of the interaction, the multiloop retuning exercise may be futile. If however only a simple control solution is desired then the level loop can be detuned and the pressure loop can have larger gains with smaller integral action to reduce oscillations. If the system warrants, then a multivariate control loop could be designed.

As would be evident from the above discussion and case study, there remains much to be desired in obtaining practically meaningful measures of multivariate control performance. The minimum variance control is a useful benchmark as it requires little *a priori* information about the process. If however, more detailed performance measures are desired then, as would be expected, more process information is needed. For example, it would be desirable to include the control 'cost' or effort in the performance evaluation of a controller or answers to questions such as the following may be required: What is the best available control subject to soft constraints on the controller output variance. Two relatively new benchmarks are presented next as alternative measures of practical multivariate controller performance.

LQG Benchmarking

Preliminary results on the LQG benchmark as an alternative to the minimum variance benchmark were proposed in Huang and Shah (1999). These results are reviewed here and a new benchmark that takes the control cost into account is proposed. The main advantage of the minimum variance benchmark is that other than the time-delay, it requires little process information. On the other hand if one requires more information on controller performance such as how much can the output variance

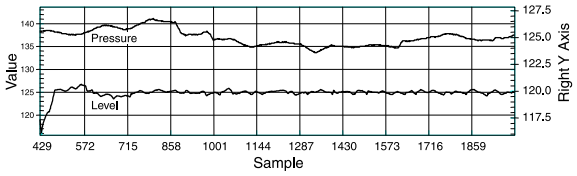


Figure 7: Pressure and level data with sampling interval 5 seconds.

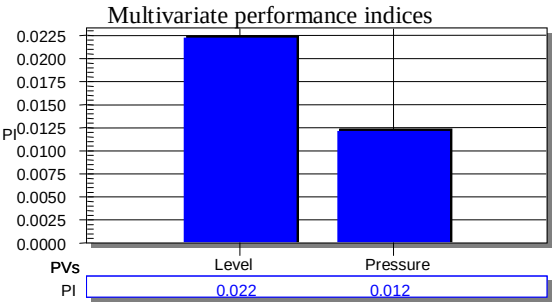


Figure 8: Individual output performance indices.

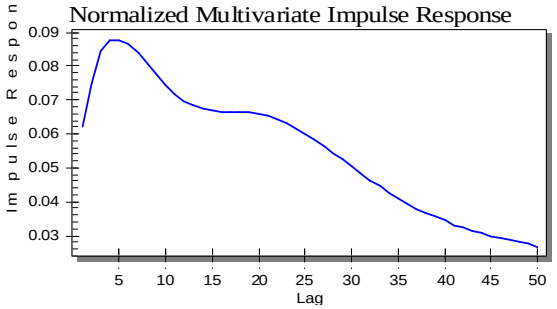


Figure 9: Normalized multivariate impulse response without interactor filtering.

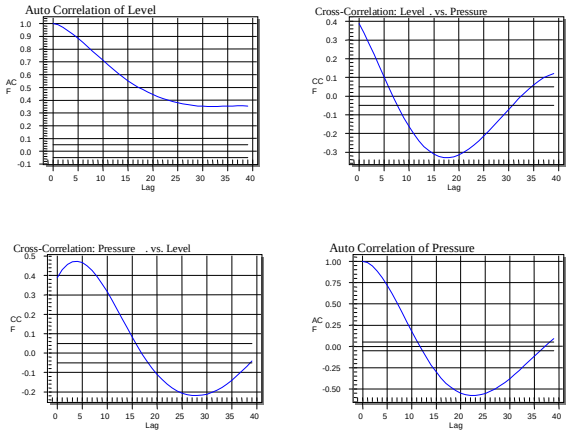


Figure 10: Auto and cross-correlation of process output.

be reduced without significantly affecting the controller output variance then one needs more information on the process. In short it is useful to have a benchmark that explicitly takes the control cost into account. The LQG cost function is one such benchmark. This benchmark does not require that an LQG controller be implemented for the given process. Rather the benchmark provides the 'limit of performance' for any linear controller in terms of the input and output variance. As remarked earlier, it is a general benchmark with the minimum variance as a special case. The only disadvantage is that the computation of the performance limit curve as shown in Figure 12 requires knowledge of the process model. For MPC type controllers these models may be readily available. Furthermore the benchmark cannot handle hard constraints, but it can be used to compare the performance of unconstrained and constrained controllers (see Figure 13).

In general, tighter quality specifications lead to smaller variations in the process output, but typically require more control effort. Consequently one may be interested in knowing how far away the control performance is from the "best" achievable performance with the same effort, i.e., in mathematical form the resolution of the following problem may be of interest:

> Given that $E(u^2) \leq \alpha$, what is the lowest achievable $E(y^2)$?

The solution is given by a tradeoff curve as shown in Figure 12. This curve can be obtained by solving the LQG problem (Kwakernaak and Sivan, 1972), where the LQG objective function is defined by:

$$J(\lambda) = E(y^2) + \lambda E(u^2)$$

By varying λ, various optimal solutions of $E(y^2)$ and $E(u^2)$ can be calculated. Thus a curve with the optimal output variance as ordinate, and the incremental manipulative variable variance as the abscissa can be plotted from these calculations. Boyd and Barratt (1991) have shown that any linear controller can only operate in the region above this curve. In this respect this curve defines the limit of performance of all linear controllers, as applied to a linear time-invariant process, including the minimum variance control law. If the process is modelled as an ARIMAX process then the resulting LQG benchmark curve due to the optimal controller will have an integrator built into it to asymptotically track and reject step type setpoints and disturbances respectively. Five optimal controllers may be identified from the tradeoff curve shown in Figure 12. They are explained as follows:

- Minimum cost control: This is an optimal controller identified at the left end of the tradeoff curve. The minimum cost controller is optimal in the sense that

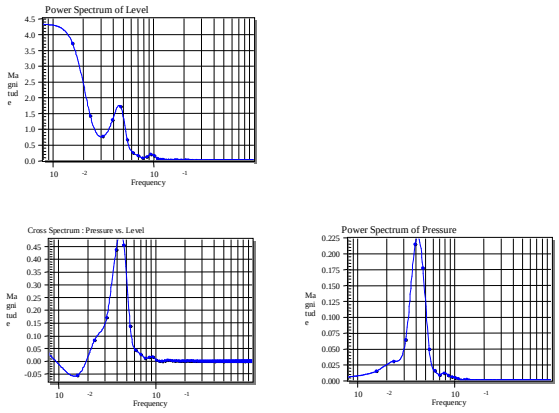


Figure 11: Frequency domain analysis of process output.

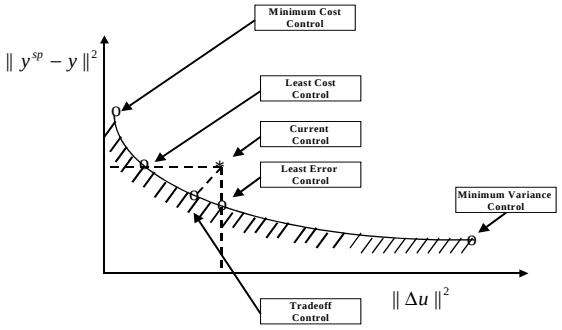


Figure 12: The LQG tradeoff curve with several optimal controllers.

it offers an offset-free control performance with the minimum possible control effort. It is worthwhile pointing out that this controller is different from the open-loop mode since an integral action is guaranteed to exist in this controller.

- Least cost control: This optimal controller offers the same output error as the current or existing controller but with the least control effort. So if the output variance is acceptable but actuator variance has to be reduced then this represents the lowest achievable manipulative action variance for the given output variance.
- Least error control: This optimal controller offers least output error for the same control effort as the existing controller. If the input variance is acceptable but the output variance has to be reduced then this represents the lowest achievable output variance for the given input variance.
- Tradeoff controller: This optimal controller can be identified by drawing the shortest line to the tradeoff curve from the existing controller; the intersection is the tradeoff control. Clearly, this tradeoff controller has performance between the least cost control and the least error control. It offers a tradeoff between reductions of the output error and the control effort.
- Minimum error (variance) control: This is an optimal controller identified at the right end of the tradeoff curve. The minimum error controller is optimal in the sense that it offers the minimum possible error. Note that this controller may be different from the traditional minimum variance controller due to the existence of integral action.

The challenges with respect to the LQG benchmark lie in the estimation of a reasonably accurate process model. The uncertainty in the estimated model then has to be 'mapped' onto the LQG curve, in which case it would become a fuzzy trade-off curve. Alternately the uncertainty region can be mapped into a region around the current performance of the controller relative to the LQG curve (see Patwardhan et al., 2000).

An Alternative Method for Multivariate Performance Assessment Using the Design Case as a Benchmark

An alternative approach is to evaluate the controller performance using a criterion commensurate with the actual design objective(s) of the controller and then compare the achieved performance. This idea is analogous to the method of Kammer et al. (1996), which was based on frequency domain comparison of the achieved and design objective functions for LQG. For a MPC controller with a quadratic objective function, the design requirements are quantified by:

$$\hat{J}(k) = \sum_{i=1}^{p} (y_{sp}(k+i|k) - \hat{y}(k+i|k))^T \Gamma_{k,i} (y_{sp}(k+i|k) - \hat{y}(k+i|k)) + \sum_{i=1}^{M-1} \Delta u(k+i-1)^T \Lambda \Delta u(k+i-1)$$

where

$\hat{y}(k+i|k)$ is the i-step ahead predictor of the outputs based on the process model

$y_{sp}(k+i|k)$ is the setpoint trajectory

$\Delta u(k+i-1)$ are the future moves of the inputs

$\Gamma_{k,i}$ are the output weightings that, in general, can depend upon the current time and the prediction horizon

Details of the model predictive control calculations can be found in any standard references (Garcia et al., 1989; Mayne et al., 2000; Qin and Badgwell, 1996). Here we restrict ourselves to the performance assessment aspects. The model predictive controller calculates the optimal

control moves by minimizing this objective function over the feasible control moves. If we denote the optimal control moves by $\Delta u^*(k)$, the optimal value of the design objective function is given by

$$\hat{J}^*(k) = \hat{J}(\Delta u^*(k))$$

The actual output may differ significantly from the predicted output due to inadequacy of the model structure, nonlinearities, modeling uncertainty etc. Thus the achieved objective function is given by:

$$\begin{aligned}\hat{J}(k) = &\sum_{i=1}^{p}(y_{sp}(k+i|k) - y(k+i|k))^T \\ &\Gamma_{k,i}(y_{sp}(k+i|k) - y(k+i|k)) \\ &+ \sum_{i=1}^{M-1} \Delta u(k+i-1)^T \Lambda \Delta u(k+i-1)\end{aligned}$$

where $y(k)$ and $\Delta u(k)$ denote the measured values of the outputs and inputs at corresponding sampling instants appropriately vectorized. The inputs will differ from the design value in part due to the receding horizon nature of the MPC control law. The value of the achieved objective function cannot be known *a priori*, but only p sampling instants later. A simple measure of performance can then be obtained by taking a ratio of the design and the achieved objective functions as:

$$\eta(k) = \frac{\hat{J}^*(k)}{J(k)}$$

This performance index will be equal to one when the achieved performance meets the design requirements. The advantage of using the design criterion for the purpose of performance assessment is that it is a measure of the deviation of the controller performance from the expected or design performance. Thus a low performance index truly indicates changes in the process or the presence of disturbances, resulting in sub-optimal control. The estimation of such an index does not involve any time series analysis or identification. The design objective is calculated by the controller at every instant and only the measured input and output data is needed to find the achieved performance. The above performance measure represents an instantaneous measure of performance and can be driven by the unmeasured disturbances. In order to get a better overall picture the following measure is recommended:

$$\alpha k = \frac{\sum_{i=1}^{k} \hat{J}^*(i)}{\sum_{i=1}^{k} J(k)}$$

$\alpha(k)$ is the ratio of the average design performance to the average achieved performance up to the current sampling instant. Thus $\alpha(k) = 1$ implies that the design performance is being achieved on an average. $\alpha(k) < 1$ means that the achieved performance is worse than the design. This alternative metric of multivariate controller performance has been applied towards the evaluation of a QDMC and another MPC controller. Further details on the evaluation of this algorithm can be found in (Patwardhan, 1999).

The motivation for a lumped performance index is that the MPC controllers in the dynamic sense, attempt to minimize a lumped performance objective. The lumped objective function and subsequently the performance index, therefore does reflect the true intentions of the controller. The motivation for this idea was to have a performance statistic for MPC that is commensurate with its constrained and time-varying nature. The idea of comparing design with achieved performance has been common place in the area of control relevant identification (also known as iterative identification and control, identification for control)—see the survey by Van den Hof and Schrama (1995). Performance degradation is measured as a deviation from design performance and becomes the motivation for re-design/re-identification.

Simulation Example: A Mixing Process. The above approach was applied to a simulation example. The system under consideration is a 2×2 mixing process. The controlled variables are temperature (y_1) and water level (y_2) and the manipulated inputs are inlet hot water (u_1) and inlet cold water (u_2) flow rates. The following model is available in discrete form,

$$P(z^{-1}) = \begin{bmatrix} \dfrac{0.025z^{-1}}{1-0.8607z^{-1}} & \dfrac{-0.1602z^{-1}}{1-0.8607z^{-1}} \\ \dfrac{0.2043z^{-1}}{1-0.9827z^{-1}} & \dfrac{0.2839z^{-1}}{1-0.9827z^{-1}} \end{bmatrix}$$

A MPC controller was used to control this process in the presence of unmeasured disturbances. The controller design parameters were:

$$p = 10, \quad , m = 2, \quad , \lambda = diag([1,4]), \quad \Gamma = diag([1,2])$$

White noise sequences at the input and output with covariance equal to 0.1I served as the unmeasured disturbances. First the LQG benchmark was found, and the performance of a constrained and unconstrained MPC was evaluated against this benchmark (see Figure 13). Constraints on input moves were artificially imposed in order to activate the constraints frequently. The unconstrained controller showed better performance, compared to the constrained controller, with respect to the LQG benchmark.

A plot of the LQG objective function compared to the achieved objective function is shown in Figure 14. A performance measure of $J_{lq}/J_{ach} = 0.467/0.71 = 0.6579$ was

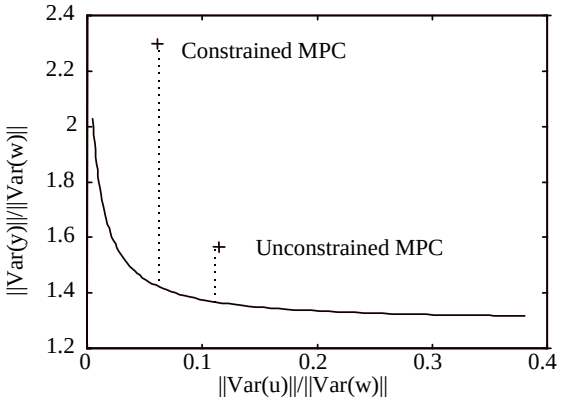


Figure 13: MPC performance assessment using LQG benchmark.

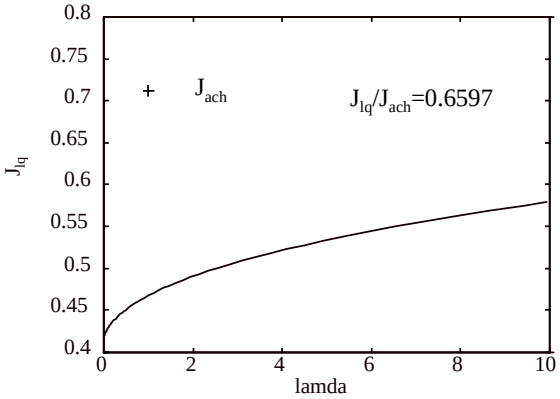


Figure 14: Comparison of the achieved performance with the LQG objective function.

	LQG	$\alpha(k)$
Unconstrained	0.6579	0.8426
Constrained	0.4708	1.00

Table 1: Effect of constraints on MPC performance.

obtained for the unconstrained controller. Performance assessment of the same controller using the design case benchmarking approach yields contrasting results (Table 1). For the unconstrained controller a performance index 0.8426 revealed satisfactory performance while the imposition of constraints led to a performance index of 1. The constrained controller showed improvement according to one benchmark and deterioration with respect to the LQG benchmark. The design case approach indicates that the controller is doing its best under the given constraints while the LQG approach which is based on comparison with an unconstrained controller shows a degradation in performance.

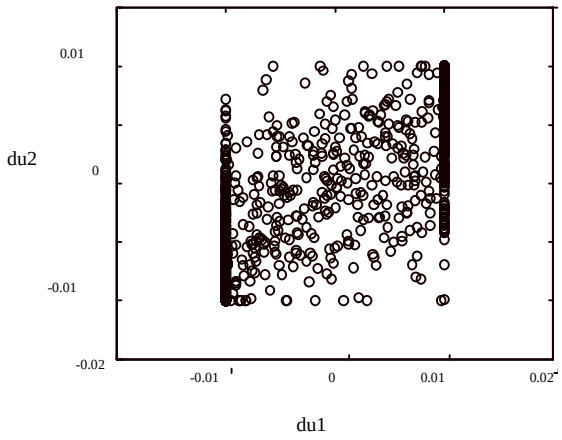


Figure 15: The input moves for the constrained controller during the regulatory run.

Figure 15 shows the input moves during the regulatory run for the constrained controller. The constraints are active for a large portion of the run and are limiting the performance of the controller in an absolute sense (LQG). On the other hand the controller cannot do any better due to design constraints as indicated by the design case benchmark.

Challenges in Performance Analysis and Diagnosis: General Comments and Issues in MPC Performance Evaluation

A single index or metric by itself may not provide all the information required to diagnose the cause of poor performance. Considerable insight can be obtained by carefully interpreting all the performance indices. For example, in addition to the minimum variance benchmark performance index, one should also look at the cross-correlation plots, normalized multivariate impulse response plots, spectrum analysis, etc., to determine causes of poor performance. As an example, if the process data is 'white' then the performance index will always be close to 1 irrespective of how large the variance is. On the other hand, if the data is highly correlated (highly predictable), then the performance index will be low irrespective of how small the output variance is. In this respect the performance index plus the impulse response or the auto-correlation plot would provide a complete picture of the root cause of the problem. (The auto-correlation plot would have yielded information on the predictability of the disturbance). In summary then, each index has its merits and its limitations. Therefore, one should not just rely on any one specific index. It would be more appropriate to check all relevant indices that reflect performance measures from different aspects.

As mentioned earlier, the multivariate extension of

the minimum variance benchmark requires a knowledge of the time-delay or the interactor *matrix*. This requirement of *a priori* information on the interactor has been regarded by many as impractical. However, from our experience this benchmark, when applied with care, can yield meaningful measures of controller performance. Yet, many outstanding issues remain open before one can confidently apply MIMO assessment techniques for a wide-class of MIMO systems. Some of the issues related to the evaluation of multivariate controllers are listed below:

1. To calculate a *general* interactor matrix, one needs to have more *a priori* information than just the time delays. However, experience has shown us that a significant number of MIMO processes do have the diagonal interactor structure. In fact, a properly designed MIMO control structure will most likely have a diagonal interactor structure (Huang and Shah, 1999). A diagonal interactor depends only on the time delays between the paired input and output variables.
2. Since models are available for all MPC based controllers, *a priori* knowledge of the time delay matrix is surely not an issue at all.
3. A more important issue in the analysis and diagnosis of control loops is the accuracy of the models and their variability over time. How does the model uncertainty affect the calculation of performance index? This question has not been answered so far. It is a common problem in both MIMO and SISO performance assessment. Therefore, one of the many outstanding issues remaining is the robustness of performance assessment, i.e. how to transfer the model uncertainty onto the uncertainty in the calculation of the performance index? This issue has been addressed to some extent in Patwardhan (1999; 2001), where SISO and MIMO examples, relating modeling uncertainty to uncertainty in performance measures, are given.

Industrial MPC is a combination of a dynamic part and a steady state part, which often comprises a linear programming (LP) step. The dynamic component consists of unconstrained minimization of a dynamic cost function, comprising of the predicted tracking errors and future input moves, familiar to academia. The steady state part focuses on obtaining economically optimal targets, which are then sent to the dynamic part for tracking as illustrated in Figure 16. This combination of the dynamic and steady state parts and constraint handling via the linear programming or the LP step renders the MPC system as a nonlinear multivariate system. Patwardhan et al. (1998) have illustrated the difficulties caused by the LP step on an industrial case study involving a demethanzier MPC (see Figure 17). In that particular

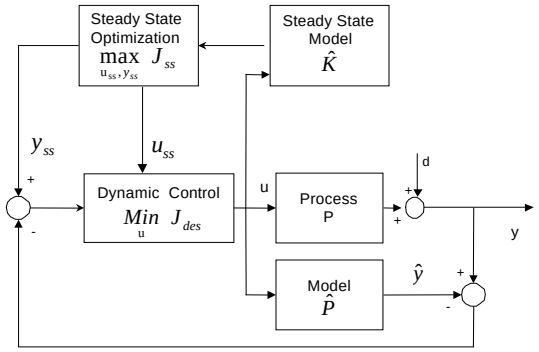


Figure 16: Schematic of a typical commercial MPC with a blended linear programming module that sets targets for the controlled and manipulative variables.

application as in a number of other MPC applications, when the LP stage is activated fairly routinely, significant correlation exists between the LP targets and the measurements that the LP relies upon. In such instances when the LP stage is activated at the same frequencies as the control frequencies, the controller structure is no longer linear. Situations such as these preclude the use of conventional performance assessment methods such the LQ or the minimum variance benchmark. We believe that the variable structure nature of industrial MPC can be captured by the objective function method since it takes into account the time varying nature of the MPC objective. Patwardhan (1999) has applied this method successfully on an industrial QDMC application. Even though one may argue that QDMC is devoid of the LP step, it is a variable structure MIMO controller that allows different inputs and outputs to swap into active and inactive states relative to the active constraint set. In this respect, the lumped objective function and subsequently the performance index proposed here does 'measure' the true intentions of the controller relative to the design case. It thus provides a useful performance metric. The only limitation being that the access to the actual design control objective has to be available in the MPC vendor software.

Establishing the root causes of performance degradation in industrial MPCs is indeed a challenging task. Potential factors include models, inadequately designed LP in that the LP operates at the control frequencies, inappropriate choice of weightings, ill-posed constraints, steady-state bias updates etc. In practice, these factors combine in varying degrees to give poor performance. Thus the issues and challenges related to the diagnosis aspects of MPC performance assessment are many. Some of these issues are listed below and one 'quantifiable' diagnosis issue related to model-plant mismatch is discussed. The diagnosis stage for poor performance involves a trial and error approach (Kesavan and Lee,

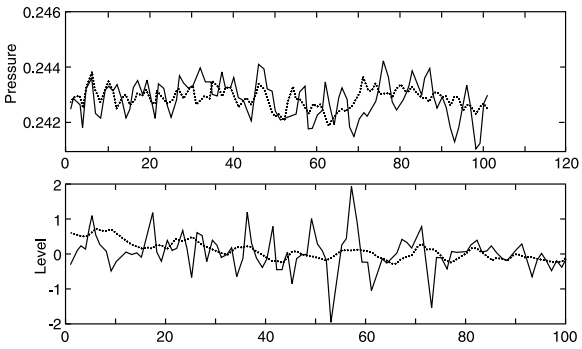


Figure 17: An example of the interaction between the steady state optimization and the dynamic layer in industrial MPC. Note that setpoints have higher variation compared to controlled variables!

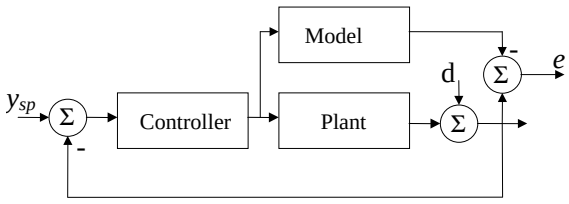


Figure 18: Schematic of a closed loop system in which the prediction error is monitored.

1997). For example, the diagnosis or the decision support system has to investigate the cause of poor performance as being due to:

- Poor or incorrect tuning.
- Incorrect controller configuration, e.g. choice of MVs may not be correct.
- Large disturbances, in which case the sources of measured disturbances have to be identified and potential feedforward control benefits should be investigated.
- Engineering redesign , e.g. is it possible to reduce process time delays?
- Model-plant mismatch in the case of MPC controllers and how does model uncertainity affect the calculation of the performance index (e.g. if the index has been obtained from an uncertain interactor).
- Poor choice of constraint variables and constrained values.

Some of the above referenced issues have been already dealt with in the literature, e.g. (ANOVA analysis to investigate the need for feedforward control by Desborough and Harris (1993), Vishnubhotla et al. (1997), and Stanfelj et al. (1993); others are open problems. The diagnosis issues related to the model-plant mismatch is briefly discussed below in a theoretical framework and illustrated on an industrial MPC evaluation case study that follows. A discussion of the poor performance diagnosis steps leading to guidelines for tuning and controller design issues is beyond the scope of this paper.

Model-Plant Mismatch

MPC controllers rely heavily on process models. In particular an accurate model is required if the process is to be regulated very tightly. On the other hand performance can be detuned in favour of robustness if an accurate model is not available or should the process change over a period of time. The extent of MPM can not be easily discerned by simply examining the closed loop prediction error. As shown in Figure 18, the prediction error under closed loop conditions is a function of the MPM, setpoint changes and measured and unmeasured disturbances. Thus the cause of large prediction errors may not necessarily be attributed to a large MPM. Consider Figure 18, where the prediction error is denoted as e.

Under open loop conditions, the prediction errors is: $e = (P - \hat{P})u + d$. Under closed loop control the prediction error expression is:

$$e = \left(\frac{(P-\hat{P})C}{1+CP}\right) y_{sp} + \left(\frac{1+C\hat{P}}{1+CP}\right) d$$

It is clear from the above expression that a large prediction error signal could be due to a large MPM term, or a large disturbance term or setpoint changes. Thus the question of attributing a large prediction error as being due to model-plant-mismatch needs careful scrutiny. Huang (2000) has studied the problem of detecting significant model plant mismatch or process parameter changes in the presence of disturbances.

Industrial MIMO Case study 2: Analysis of Cracking Furnace Under MPC Control

This section documents the results of the controller performance analysis carried out on an ethane cracking furnace. The control systems comprises of (1) a regulatory layer and (2) an advanced MPC control layer. The first pass of performance assessment revealed some poorly performing loops. Further analysis revealed that these loops were in fact well tuned but were being affected by high frequency disturbances and setpoint changes. The furnace MPC application considered here, however, is unlike conventional MPC applications. The steady state limits were set in such a way that the setpoints for the controlled variables were held constant, i.e. the focus of the evaluation was on the models and the tuning of the dynamic part.

The MPC layer displays satisfactory performance lev-

els when there are no rate changes. During rate changes, the MPC model over predicts thus causing poor performance. Re-identification of the model gains was found to be necessary to improve MPC performance.

Control Strategy Overview

The purpose of the Furnace MPC controllers is to maintain smooth operation, maximize throughput and minimize energy consumption to the furnaces while simultaneously honoring all constraints. There is one MPC controller per furnace. The conversions, the total dry feed, the wet feed bias, the steam/feed ratio and the oxygen to fuel ratio is controlled by MPC. These variables are manipulated by moving the north and south fuel gas duty setpoints, the north and south wet feed flow setpoints, the steam pressure setpoint, the fan speed controller setpoint and the induced fan draft. Thus there are 6 primary controlled variables and 7 manipulated variables. There are a number of secondary controlled variables, which MPC is required to maintain within a constraint region. These secondary CVs include valve constraints, constraints on critical variables such as the coil average temperatures (COTS). Considering the degrees of freedom (MVs) available, it may not always be possible to satisfy all the constraints. In such cases, a ranking mechanism decides what constraints are least important and could be let go.

A critical component of the MPC controllers is the model describing the relationships between the MVs and the CVs—primary as well as secondary. These models are developed on the basis of open loop tests. Step response curves are used to parameterize the models. These models are used by MPC to predict the future process response. A portion of the step response model matrix is shown in Figure 19.

Performance analysis of the furnace control loops was conducted in two stages. Phase one loop analysis was performed on the lower regulatory layer. With the exception of a few loops, the first pass of performance assessment revealed satisfactory performance of most loops. The higher level MPC performance assessment commenced next.

Multivariable Performance Assessment for MPC. Table 2 summarizes the performance statistics. A diagonal interactor was used, based on the knowledge of the process models.

The performance metrics indicated satisfactory performance on all the variables except for tags 4 and 6. The closed loop settling time is approximately the same as the open loop settling time, which indicates a conservatively tuned application. Part of the reason for the slow response in control of tags 4 and 6 could be the presence of a measurement delay. During January 2000, the service factor for MPC was low due to some communication issues, which have been since resolved. This meant that there was an opportunity to compare the furnace performance with and without MPC. Based on data from Jan 14-16 when MPC was shut off for part of the time, performance metrics were obtained to compare the two control systems—MPC and conventional PID controls. The statistics indicate that the overall control is only slightly better with MPC turned on.

Description	PI	Closed-Loop Settling Time (min)	Status
Tag 1	0.95	8	
Tag 2	0.94	1	
Tag 3	0.76	15	
Tag 4	0.44	15	LOW
Tag 5	0.71	15	
Tag 6	0.49	15	LOW
Tag 7	0.88	10	
Multivariable PI	**0.71**		

Table 2: Summary of MPC performance.

MPC Diagnostics

Is MPC doing its best? Can we improve the current performance levels of the furnace MPC controller? These questions lead us to two issues that are closely related to each other:

1. How good are the models used for predicting the process response?
2. How well tuned is the multivariable controller? "Tuning" includes a whole range of different of parameters—weightings, horizons, constraints, rankings ...

We will try to illustrate a case where the model predictions can mislead the controller and hence cause poor performance. This Furnace was showing poor MPC performance, especially during rate changes. This motivated us to look more closely at the model prediction accuracy.

Before evaluating the current predictions, we establish a baseline when the open loop tests were conducted. Figures 21–23 compare the conversion predictions for the open loop case as conducted before commissioning the MPC. The model accuracy is reasonable. The average prediction error, for this data set was 3.18.

$$\begin{matrix}\text{Average}\\ \text{Prediction}\\ \text{Error}\end{matrix} = \frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{N_y}\{y_i(k+1) - \hat{y}_i(k+1|k)\}^2$$

The predictions for other variables—COTs, Feed Flow and Bias, S/F ratio also fared well. The remaining variables are not shown for the sake of brevity.

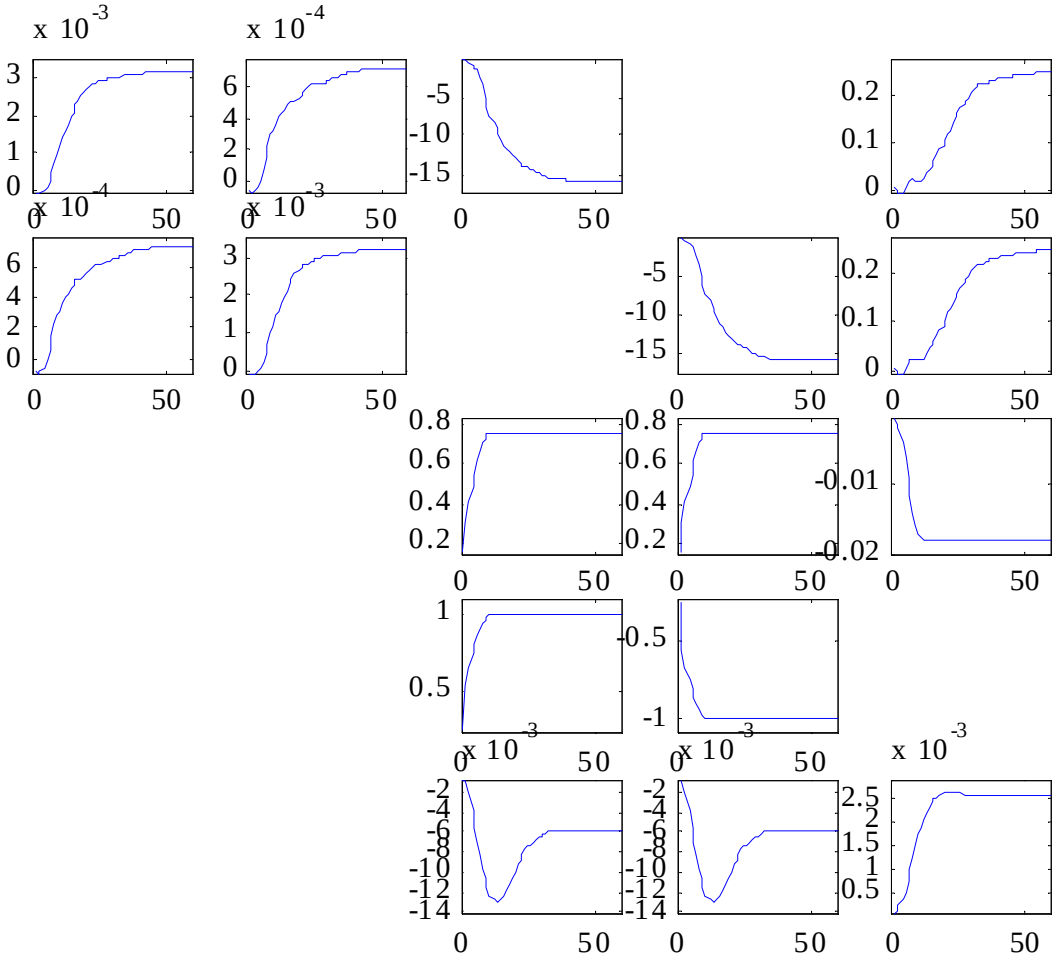


Figure 19: A portion of the step response models used by MPC.

The model accuracy can then be compared with the current predictions. There were two rate changes during this period, due to furnace decokes.

An average prediction error of 175.7 was observed during this period. The tags 3 to 6 predictions were much worse than the rest. Taking a closer look at the tags 3-6 predictions revealed that the models were over predicting by a factor of 2. Based on a combination of statistical analysis and process knowledge it was decided that the model gains were incorrect.

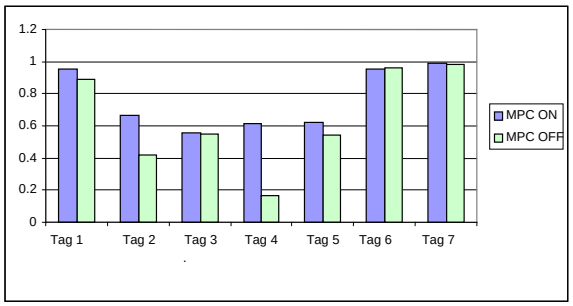


Figure 20: Comparison of performance, with and without MPC.

Improving MPC Performance

The main stumbling block to improving MPC performance was its model accuracy. The models causing the large mismatch were identified. One of the reasons for the model plant mismatch is the fact that open loop tests were carried out in a operating region which is quite different from the current operating conditions (higher rates, conversions, duties). Fairly routine plant test, were used to identify the suspected changes in steady state gains. These tests were conducted under closed loop conditions and the gain mismatch has now been fixed with satisfactory MPC performance. The newly identified gains were indeed found to be significantly different from the earlier gains.

To illustrate the effect of the MPM on cracking efficiency, conversion control on the 3 furnaces was com-

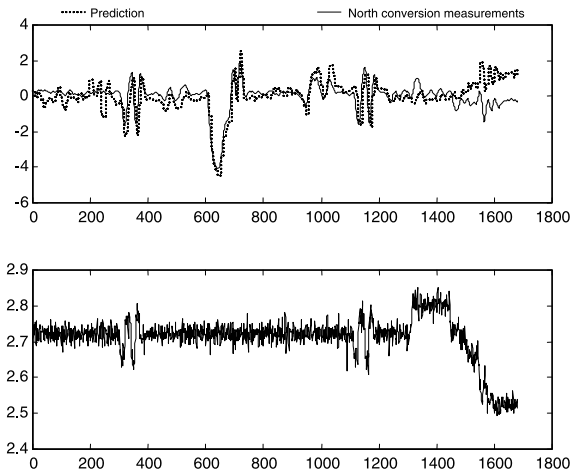


Figure 21: The predictions (green) and the North conversion measurements (blue) are shown in the top graph. The bottom graph shows the changes in fuel duty during this period.

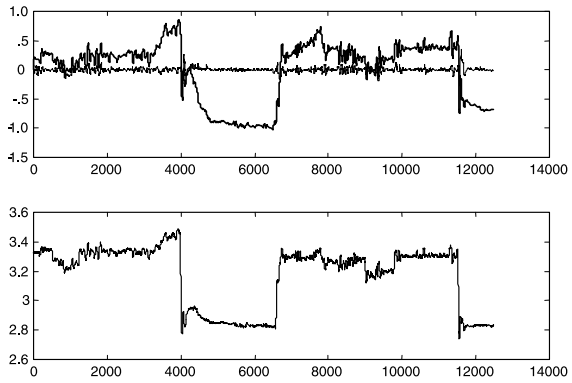


Figure 22: The scaled conversion predictions for Feb 11-15.

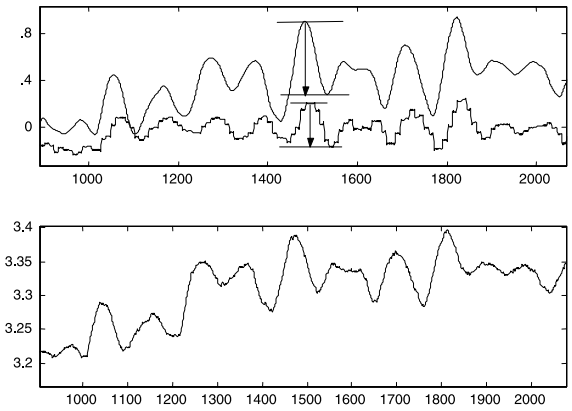


Figure 23: The scaled conversion predictions—apparent gain mismatch.

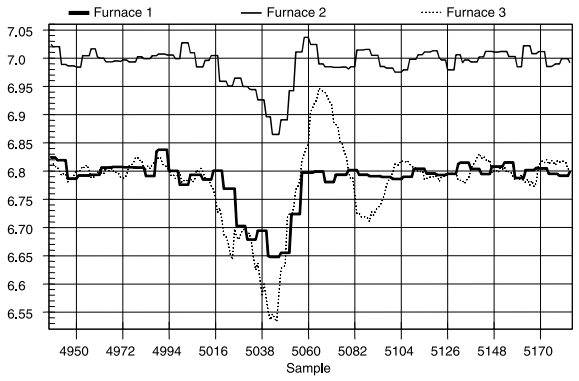


Figure 24: Comparison of Furnace control with updated models.

pared and is shown in Figure 24. Furnace 3 with the older models takes almost 45 minutes more to settle out, compared to furnaces 1 and 2, which have the updated models. The overshoot for Furnace 3 is 2.5%, as opposed to 1.5% on Furnace 1 and 2 (a significant 40% decrease). Thus the performance and subsequent diagnosis analysis of this MPC application illustrates the value in routine monitoring and maintenance of MPC applications.

Concluding Remarks

In summary, industrial control systems are designed and implemented or upgraded with a particular objective in mind. We hope that the new controller loop performance assessment methodology proposed in the literature and illustrated here, will eventually lead to automated and repeated monitoring of the design, tuning and upgrading of the control loops. Poor design, tuning or upgrading of the control loops will be detected, and repeated performance monitoring will indicate which loops should be re-tuned or which loops have not been effectively upgraded when changes in the disturbances, in the process or in the controller itself occur. Obviously better design, tuning and upgrading will mean that the process will operate at a point closer to the economic optimum, leading to energy savings, improved safety, efficient utilization of raw materials, higher product yields, and more consistent product qualities. Results from industrial applications have demonstrated the applicability of the multivariate performance assessment techniques in improving industrial process performance.

Several different measures of multivariate controller performance have been introduced in this paper and their applications and utility have been illustrated by simulation examples and industrial case studies. The multivariate minimum variance benchmark allows one to compare the actual output performance with the minimum achievable variance. However it requires knowledge

of the process time-delay matrix or the interactor. On the other hand the newly proposed NMIR_{wof} measure of performance provides a graphical 'metric' that requires little or no *a priori* information about the process and gives a graphical measure of mutivariate performance in terms of settling time, rate of decay etc.. The challenges related to MPC performance evaluation are illustrated by an industrial case study of an ethylene cracker. It is shown how routine monitoring of MPC applications can ensure good or 'optimal' control. The lumped objective function based method of monitoring MPC performance is shown to work well on the industrial case study. The study illustrates how controllers, whether in hardware or software form, should be treated like 'capital assets'; how there should be routine monitoring to ensure that they perform close to the economic optimum and that the benefits of good regulatory control will be achieved

Acknowledgments

The authors of this paper wish to express sincere gratitude and appreciation to the following persons who have provided assistance by participating in industrial evaluation studies at their respective work sites: Jim Kresta and Henry Tse (Syncrude Research), Janice Vegh-Morin and Harry Frangogiannopoulus (both of TDCC), Genichi Emoto (Mitsubishi Chemical Corporation, Japan) and Anand Vishnubhotla (Matrikon Consulting) Their feedback and industrial insight were helpful in enhancing the practicality and utility of the proposed methods to make it better suited to the process industry.

Financial support in the form of research grants from the Natural Science and Engineering Research Council of Canada and Equilon Enterprises (Shell-USA) is gratefully acknowledged.

References

Boyd, S. and C. Barratt, *Linear Control Design*. Prentice Hall, Englewood Cliffs, New Jersey (1991).

Desborough, L. and T. Harris, "Performance Assessment Measures for Univariate Feedback Control," *Can. J. Chem. Eng.*, **70**, 1186–1197 (1992).

Desborough, L. D. and T. J. Harris, "Performance Assessment Measures for Univariate Feedforward/Feedback Control," *Can. J. Chem. Eng.*, **71**, 605–616 (1993).

Garcia, C. E., D. M. Prett, and M. Morari, "Model Predictive Control: Theory and Practice—A Survey," *Automatica*, **25**(3), 335–348 (1989).

Harris, T. J., F. Boudreau, and J. F. MacGregor, "Performance Assessment of Multivariate Feedback Controllers," *Automatica*, **32**(11), 1505–1518 (1996).

Harris, T. J., "Assessment of Control Loop Performance," *Can. J. Chem. Eng.*, **67**, 856–861 (1989).

Huang, B. and S. L. Shah, *Performance Assessment of Control Loops*. Springer-Verlag, London (1999).

Huang, B., S. L. Shah, and K. K., How Good Is Your Controller? Application of Control Loop Performance Assessment Techniques to MIMO Processes, In *Proc. 13th IFAC World Congress*, volume M, pages 229–234, San Francisco (1996).

Huang, B., S. L. Shah, and E. K. Kwok, "Good, Bad or Optimal? Performance Assessment of Multivariable Processes," *Automatica*, **33**(6), 1175–1183 (1997).

Huang, B., S. L. Shah, and H. Fujii, "The Unitary Interactor Matrix and its Estimation Using Closed-Loop Data," *J. Proc. Cont.*, **7**(3), 195–207 (1997).

Huang, B., "Model Validation in the Presence of Time-variant Disturbance Dynamics," *Chem. Eng. Sci.*, **55**, 4583–4595 (2000).

Kammer, L. C., R. R. Bitmead, and P. L. Barlett, Signal-based testing of LQ-optimality of controllers, In *Proc. of the 35th IEEE CDC*, pages 3620–3624, Kobe, Japan (1996).

Kesavan, P. and J. H. Lee, "Diagnostic Tools for Multivariable Model-Based Control Systems," *Ind. Eng. Chem. Res.*, **36**, 2725–2738 (1997).

Ko, B. and T. F. Edgar, Performance Assessment of Multivariable Feedback Control Systems (2000). Submitted to *Automatica*.

Kozub, D. J. and C. E. Garcia, Monitoring and Diagnosis of Automated Controllers in the Chemical Process Industries, AIChE Annual Meeting, St. Louis (1993).

Kozub, D., Controller Performance Monitoring and Diagnosis Experiences and Challenges, In Kantor, J. C., C. E. García, and B. Carnahan, editors, *Chemical Process Control—V*, pages 156–164. CACHE (1997).

Kwakernaak, H. and R. Sivan, *Linear Optimal Control Systems*. John Wiley and Sons, New York (1972).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and Optimality," *Automatica*, **36**(6), 789–814 (2000).

Patwardhan, R. S. and S. L. Shah, Issues in Performance Diagnostics of Model-based Controllers, Accepted for publication, J. Proc. Control (2001).

Patwardhan, R. S., G. Emoto, and S. L. Shah, Model-based Predictive Controllers: An Industrial Case Study, AIChE Annual Meeting, Miami (1998).

Patwardhan, R., S. L. Shah, and K. Qi, Techniques for Assessment of Model Predictive Controllers, Submitted for publication (2000).

Patwardhan, R., S. L. Shah, and G. Emoto, Experiences in Performance Annalysis of Industrial Model Predictive Controllers, Submitted for publication (2001).

Patwardhan, R. S., *Studies in the Synthesis and Analysis of Model Predictive Controllers*, PhD thesis, University of Alberta (1999).

Qin, S. J. and T. A. Badgwell, An Overview of Industrial Model Predictive Control Technology, In Kantor, J. C., C. E. Garcia, and B. Carnahan, editors, *Fifth International Conference on Chemical Process Control—CPC V*, pages 232–256. American Institute of Chemical Engineers (1996).

Stanfelj, N., T. E. Marlin, and J. F. MacGregor, "Monitoring and Diagnosing Process Control Performance: The Single-Loop Case," *Ind. Eng. Chem. Res.*, **32**, 301–314 (1993).

Swanda, A. P. and D. E. Seborg, Controller Performance Assessment Based on Setpoint Response Data, In *Proceedings of the American Control Conference*, pages 3863–3867, San Diego, California (1999).

Thornhill, N. F., M. Oettinger, and P. Fedenczuk, "Refinery-wide control loop performance assessment," *J. Proc. Cont.*, **9**, 109–124 (1999).

Van den Hof, P. M. J. and R. J. P. Schrama, "Identification and control—closed-loop issues," *Automatica*, **31**(12), 1751–1770 (1995).

Vishnubhotla, A., S. L. Shah, and B. Huang, Feedback and Feedforward Performance Analysis of the Shell Industrial Closed Loop Data Set, In *Proc. IFAC Adchem 97*, pages 295–300, Banff, AB (1997).

Wood, R. K. and M. W. Berry, "Terminal Composition Control of a Binary Distillation Column," *Chem. Eng. Sci.*, **28**, 1707–1717 (1973).

Recent Developments in Controller Performance Monitoring and Assessment Techniques

Thomas J. Harris *
Department of Chemical Engineering
Queen's University
Kingston, ON, K7L 3N6, Canada

Christopher T. Seppala
Automation Engineering
Equilon Enterprises LLC
Houston, TX 77251

Abstract

In the past several years there has been considerable commercial and academic interest in methods for analyzing the performance of univariate and multivariate control systems. This focus is motivated by the importance that control systems have in enabling companies to achieve goals related to quality, safety and asset utilization. Control system performance cannot be adequately described by simple statistics, such as the mean and variance of manipulated and controlled variables, the percentage of time that constraints are satisfied, and the on-stream time. Although these are important performance measures, a comprehensive approach for controller performance monitoring usually includes the following elements: i) determination of the capability of the control system, ii) development of suitable statistics for monitoring the performance of the existing system, and iii) development of methods for diagnosing the underlying causes for changes in the performance of the control system. In this paper, recent developments related to these items will be reviewed for both univariate and multivariate systems. Some multivariate time series methods helpful in supporting these controller performance assessment techniques in practice will be discussed, and an industrial example will be provided. Finally, the future direction of commercial applications of controller performance assessment will be briefly discussed, as will the issue of whether controller performance assessment is destined to be offered as a product or a service.

Keywords

Controller performance, Minimum variance control, Multivariable systems, Subspace methods, Vector autoregressions

Introduction

Early interest in theory and methods for the on-line analysis of control systems can be traced to papers by Åström (1967) and De Vries and Wu (1978). Since that time work in this area has continued, with considerable development taking place during the 1990s. Reviews and critical analyses of several approaches for assessing control loop performance can be found in Huang and Shah (1999), Harris et al. (1999), and Qin (1998). Control system capability statistics based on the performance benchmark of minimum variance control for single-input single-output (SISO) systems were the initial underlying concept for much of this work. Regulation of stochastic and deterministic disturbances, setpoint tracking, extensions to multiple-input-single-output (MISO) systems (i.e., single output systems with feedforward variables) are readily accommodated in this framework, and these aspects have been described in the aforementioned references. Recently, Ko and Edgar (2000a) have extended these ideas to evaluate cascade control systems. Horch and Isaksson (1999) have proposed a modification to the basic performance measures to more closely connect the monitoring and control objectives. Thornhill et al. (1999) provide comprehensive guidelines for the application of control loop assessment and Miller and Desborough (2000) describe a commercial product/service for control loop assessment.

Extensions of minimum variance performance assessment techniques to multiple-input-multiple-output (MIMO) systems for general time delay systems was initially considered by Harris et al. (1996), and Huang et al. (1997b). The challenges in evaluating a multivariate control system (as opposed to analyzing single loops in a complex control system) are considerable. These challenges arise primarily from: i) the requirement for *a priori* knowledge of the time delay structure of the process, ii) the time-varying nature of control loops which arises from the constraint handling requirements of multivariate controllers, and iii) the requirement to use sophisticated identification methods to obtain meaningful estimates of the closed loop impulse weights. There are a considerable number of challenges, both theoretical and practical, in the assessment of multivariate schemes.

The purpose of this paper is to provide: i) a concise summary of results in univariate performance monitoring, ii) an overview of challenges and recent developments in the assessment of multivariate control systems, and iii) a brief discussion on the future direction of industrial applications of controller performance assessment, including the role of supporting technologies (including an industrial example), and the consumer's perspective on commercial performance assessment solutions—are they products or services?

Univariate Performance Assessment

Process Description

To introduce the concepts of control performance monitoring and assessment, consider a process whose behavior about a nominal operating point can be modeled by

*harrist@post.queensu.ca

a linear transfer function with an additive disturbance:

$$Y_t = \frac{\omega(q^{-1})q^{-b}}{\delta(q^{-1})}U_t + D_t \tag{1}$$

where Y_t denotes the difference between the process variable and a nominal operating point. U_t denotes the difference between the manipulated variable and its nominal value, and $\omega(q^{-1})$ and $\delta(q^{-1})$ are polynomials in the backshift operator, q^{-1}. b whole periods of delay elapse between making a change in the input and first observing its effect on the process output. The process disturbance, D_t, is represented by an Autoregressive-Integrated-Moving-Average (ARIMA) time series model of the form:

$$D_t = \frac{\theta(q^{-1})}{\nabla^d \phi(q^{-1})}a_t \tag{2}$$

where $\theta(q^{-1})$ and $\phi(q^{-1})$ are stable polynomials in the backshift operator, and ∇ is a shortcut notation for $(1 - q^{-1})$. The integer d denotes the degree of differencing ($0 \leq d \leq 2$ in most applications). a_t denotes a sequence of independently and identically distributed random variables with mean zero and variance σ_a^2. This disturbance structure is capable of modeling commonly occurring stochastic and deterministic disturbances.

The process is controlled by a linear feedback controller of the form:

$$U_t = G_c(q^{-1})(Y_{sp} - Y_t) \tag{3}$$

where $G_c(q^{-1})$ is the controller transfer function and Y_{sp} denotes the deviation of the setpoint from its reference value. We will assume that these values are equal; the general case is considered in Desborough and Harris (1992). With these assumptions, the closed loop is given by:

$$Y_t = \left(\frac{1}{1 + \frac{\omega(q^{-1})q^{-b}}{\delta(q^{-1})}G_c(q^{-1})} \right) D_t \tag{4}$$

Substituting Equation 2 for D_t in Equation 4 and simplifying allows the closed-loop to be written in rational transfer function form as follows:

$$Y_t = \frac{\alpha(q^{-1})}{\beta(q^{-1})}a_t = \psi(q^{-1})a_t \tag{5}$$

The closed-loop impulse response coefficients are given by:

$$\psi(q^{-1}) = 1 + \psi_1 q^{-1} + \psi_2 q^{-2} + \cdots \tag{6}$$

Convergence of the series in Equation 6 is guaranteed if the closed-loop is stable; the expansion is valid for computation of the impulse weights, ψ_j. Tyler and Morari (1996) present a useful discussion on the duality between the impulse weights ψ_j and other classic measures of controller performance including settling time, decay rate, and desired reference trajectories.

Minimum Variance Performance Bounds and Performance Measures

If one were to design a controller to minimize the variance of the output, the impulse response coefficients beyond the process deadtime, ψ_j, $j = b, b+1, \ldots$, would equal zero. The output variance would equal (Åström, 1967, 1970; Box and Jenkins, 1976):

$$\sigma_y^2 = \sigma_{mv}^2 = (1 + \psi_1^2 + \cdots + \psi_{b-1}^2)\sigma_a^2. \tag{7}$$

If the minimum variance performance fails to meet the controller's design objectives, then reductions in the output variance can only be achieved by modifying the process to change the disturbance characteristics or by reducing the deadtime. Because σ_{mv}^2 provides a fundamental lower bound on performance, simply retuning the controller, or implementing a more sophisticated linear controller with the same manipulated variable and control interval, will not reduce process variability. This bound depends only on the process delay and is otherwise independent of the dynamic characteristics of the controller.

Implementation of a minimum variance controller that achieves the bound described in Equation 7 requires that the polynomials $\omega(q^{-1})$ and $\delta(q^{-1})$ and be stable. When these conditions are not satisfied, it is still possible to design a controller that minimizes the variance of the output subject to stability of both the closed-loop and manipulated variable. The output variance will, by necessity, exceed that described by Equation 7. This topic is discussed further in a subsequent section.

Desborough and Harris (1992), Stanfelj et al. (1993), Kozub and Garcia (1993) and Kozub (1996) have introduced a number of performance indices to provide an indication of the departure of the current performance from minimum variance control. Typical performance measures are:

$$\xi(b) = \frac{\sigma_y^2}{\sigma_{mv}^2} \tag{8}$$

and

$$\begin{aligned} \eta(b) &= 1 - \frac{1 + \psi_1^2 + \cdots + \psi_{b-1}^2}{1 + \psi_1^2 + \cdots + \psi_{b-1}^2 + \psi_b^2 + \cdots} \\ &= 1 - \frac{\sigma_{mv}^2}{\sigma_y^2} \end{aligned} \tag{9}$$

where $\xi(b) \geq 1$ and $0 \leq \eta(b) \leq 1$. The performance index $\xi(b)$ corresponds to the ratio of the actual variance to that which could theoretically be achieved under minimum variance control. The normalized performance index, $\eta(b)$, is a number between 0 (minimum variance performance) and 1 (far from minimum variance performance) that reflects the inflation of the output variance over the theoretical minimum variance bound. As indicated in Desborough and Harris (1992), it is more useful

to replace σ_y^2 by the mean square error of y_t, thereby accounting for offset.

Note that the normalized performance index is independent of the magnitude of the disturbance driving force (a_t in Equation 2). It may happen that $\eta(b) = 0$, i.e., the system is operating at minimum variance control performance, yet σ_y^2 still exceeds process or product requirements. In this case the process—not the control system—is not capable.

Estimation of Minimum Variance Performance Bounds and Performance Measures

It is possible to calculate the minimum variance performance bound, σ_{mv}^2, by estimating a process plus disturbance model obtained from a designed experiment. This performance bound can then be used to determine the process capability. If the performance bound fails to meet process specifications, then the process modification remedies described above must be sought to rectify the situation. Such an approach would severely limit the usefulness of this bound because obtaining a process and disturbance model is labour intensive and intrusive, i.e., it requires a perturbation signal to be introduced. Furthermore, the disturbance structure may change during the period of data collection, potentially changing the results of the analysis.

Why then, has the minimum variance performance benchmark in Equation 7 proven to be so useful in practice? Its usefulness stems from two important properties:

1. **Autocorrelation Test**: Under minimum variance control, the autocorrelations of the y's are zero beyond lag $(b-1)$ since the closed-loop is a moving average process of order $(b-1)$, or an ARIMA$(0, 0, b-1)$ process (Åström, 1967, 1970; Box and Jenkins, 1976). Conversely, if any (stable) controller results in a closed-loop, which is an ARIMA$(0, 0, b-1)$ process, then the controller is a minimum variance controller. If the controller is unstable, except for the presence of p integrators, then the observed closed-loop may appear to be a moving average process of order less than $(b-1)$, (Foley and Harris, 1992). Except in these rare cases, the sample autocorrelation function of the y's, or a portmanteau test on the autocorrelations of y can be used to provide a simple, convenient, and useful method for testing whether any SISO controller is giving minimum variance performance (Harris, 1989; Stanfelj et al., 1993; Kozub and Garcia, 1993; Kozub, 1996).
2. **Invariance Property**: σ_{mv}^2 can, under mild conditions, be estimated from *routine operating data* when the time delay is known (Harris, 1989). It is straightforward to show that the first $(b-1)$ ψ_j coefficients of the closed-loop equal the first $(b-1)$ impulse coefficients of the disturbance transfer function. The remaining coefficients are functions of the controller, process, and disturbance transfer functions. Since the first $(b-1)$ ψ_j coefficients are not affected by any feedback controller they can collectively be interpreted as a system invariant (Harris, 1989; Tyler and Morari, 1996). These can be estimated by fitting a time series model to the closed loop error:

$$\alpha(q^{-1})(y_t - \bar{y}) = \beta(q^{-1})a_t \tag{10}$$

where $\alpha(q^{-1})$ and $\beta(q^{-1})$ are stable polynomials in the backshift operator, of order n_a and n_b, respectively. The term $\bar{y}$ accounts for non-zero mean data. The coefficients of the polynomials and their orders can be estimated using standard time series analysis techniques. Once these parameters have been estimated, the impulse weights are calculated by long division of $\alpha(q^{-1})$ into $\beta(q^{-1})$. Computational details, and variants, are discussed in Harris (1989), Desborough and Harris (1992) and Huang et al. (1997b).

It is important to note that the calculation of σ_a^2 and σ_{mv}^2 does not require separate identification of the process transfer function and disturbance transfer functions since $\eta(b)$ corresponds to the fraction of the output variance reduction that can be achieved by implementing a minimum variance controller. As a result of the above properties, σ_{mv}^2 and $\eta(b)$ can be estimated from routine operating data if the delay is known.

Exact distributional properties of the estimated performance indices are complicated, and not amenable to a closed-form solution. Desborough and Harris (1992) approximated first and second moments for the estimated performance indices and resorted to a normal theory to develop approximate confidence intervals. Asymptotically, the performance indices are ratios of correlated quadratic forms, and as such the distributions of the performance indices are non-symmetric. Refinements to the confidence intervals developed in Desborough and Harris (1992) can be obtained with little extra computational effort, by resorting to the extensive statistical literature on the distributional properties of quadratic forms (Harris, 2001).

Extensions and Modifications

The development thus far has been based on the simple process description given by Equation 4. Performance monitoring and assessment methods have been extended to include variable setpoints (Desborough and Harris, 1992), feedforward/feedback systems (Desborough and Harris, 1993; Stanfelj et al., 1993), processes with interventions (Harris et al., 1999), and cascade systems (Ko and Edgar, 2000a).

The performance bounds described above have been presented under idealized assumptions. The actual, as opposed to lower bound on performance, is also lim-

ited by the presence of non-invertible zeros, the requirement for smooth-manipulated variable movement, and the presence of hard constraints on the manipulated variable. A number of modifications have been proposed to accommodate these issues, and these will be reviewed in the following paragraphs.

Non-invertible systems. When the process transfer function is non-invertible, it is possible to design a modified minimum variance controller using spectral factorization methods (Bergh and MacGregor, 1987; Harris and MacGregor, 1987). This modified minimum variance controller has the lowest variance among all stable controllers. The following identities hold:

$$\sigma_a^2 \leq \sigma_{mv}^2 \leq \sigma_{mv^\star}^2 \leq \sigma_{\mathit{Åström}}^2 \tag{11}$$

where $\sigma_{mv^\star}^2$ denotes the variance of the modified minimum variance controller and $\sigma_{\mathit{Åström}}^2$ denotes the closed-loop variance of a simple pole placement algorithm proposed by (Åström, 1970). This latter controller is particularly easy to design; the limitation being that the non-invertible zeros of the process transfer function cannot be canceled. With this design the process output is a moving average process of order $(b-1+n^\star)$, where $n^\star$ is the number of zeros of $\omega(q^{-1})$, in q, outside the unit circle. When the location of the non-invertible zero is known, in addition to the time delay, $\sigma_{mv^\star}^2$ can be estimated from routine operating data (Harris et al., 1996; Tyler and Morari, 1995; Huang and Shah, 1999). These latter results use linear-quadratic-control theory to determine the achievable performance bound. The performance results can be sensitive to the location of the non-invertible zero (Tyler and Morari, 1995). Estimation of $\sigma_{mv^\star}^2$ requires considerably more process knowledge than is required to estimate σ_{mv}^2. Although not as rigorous, a number of alternate approaches, which retain the simplicity of the minimum variance bounds and calculations, can be used. These are discussed in subsequent sections. Recently, Ko and Edgar (2000c,b) have used fundamental results of Furuta and Wongsaisuwan (1993, 1995) to show how algorithms such as Dynamic Matrix Control (DMC) can be used to obtain several different performance bounds. This approach will be discussed further in the multivariate performance assessment section.

Excessive control action and robustness concerns. Minimum variance controllers may call for unacceptably large changes in manipulated variable action. This happens when the process is sampled "quickly" relative to its dominant time constant. In these circumstances minimum variance controllers (or deadbeat controllers) may be sensitive to process model mismatch (Åström, 1970; Bergh and MacGregor, 1987). In these instances, it has been found useful to modify the performance indices so that the latter more closely reflects the controller design requirements. Two modified controller performance indices have been proposed to deal with these issues: The extended horizon performance index and the user-defined benchmark performance index.

Extended-horizon performance index. Desborough and Harris (1992, 1993), Kozub (1996), Harris et al. (1996), and Thornhill et al. (1999) utilize an extended horizon performance index defined as:

$$\eta(b+h) = 1 - \frac{1+\psi_1^2+\cdots+\psi_{b-1}^2+\cdots+\psi_{b+h-1}^2}{1+\psi_1^2+\cdots+\psi_{b-1}^2+\psi_b^2+\cdots} \tag{12}$$

This normalized performance index gives the proportion of the variance arising from non-zero impulse coefficients ψ_j, $j > b+h$. $\eta(b+h)$ can also be interpreted as the square of the correlation between the current error and the least squares estimate of the prediction made $(b+h)$ control periods in the past (Harris et al., 1999). The extended horizon predictor closely matches control objectives of model based control strategies, such as Dynamic Matrix Control (DMC). It is important to note that when $h > 0$, the prediction error variance is affected by the structure and tuning of the feedback controller (in contrast to the case when $h = 0$). The use of the extended horizon performance index indirectly acknowledges the fact that minimum variance control may not be desirable or feasible. One obvious advantage of using $\eta(b+h)$ instead of $\eta(b)$ is that the former does not require a precise estimate of the process delay. Kozub (1996) and Thornhill et al. (1999) indicate that many problems in diagnosing the performance of controllers can be solved by estimating both $\eta(b)$ and $\eta(b+h)$.

User-defined benchmark performance index. Recently, Horch and Isaksson (1999) have introduced a normalized performance index:

$$\xi_{mod}(b) = \frac{\sigma_y^2}{\sigma_{mod}^2} \tag{13}$$

where:

$$\sigma_{mod}^2 = \left(1+\psi_1^2+\cdots+\psi_{b-1}^2+\psi_{b-1}^2\frac{\upsilon^2}{1-\upsilon^2}\right)\sigma_a^2 \tag{14}$$

and $0 \leq \upsilon < 1$. The motivation for this modified performance index is very simple; a minimum variance controller can be interpreted as a requirement that all of the closed-loop system poles be placed at the origin. If instead, one of the closed-loop poles is moved to a location specified by the designer, then the variance of the closed-loop is given by σ_{mod}^2 in Equation 13. Horch and Isaksson (1999) show that this design is equivalent to a requirement that the closed-loop have an exponential decay to target rather than the dead-beat response required of minimum variance control. With this interpretation, specification of υ is not difficult. Horch and Isaksson call the modified performance index a *user-defined benchmark*. They point out that the basic simplicity of the original performance index is retained, while offering greater flexibility. The authors do not require that

the controller be designed using this technique; rather they point out that the analysis of closed-loop data is facilitated by the choice of υ. Statistical properties of the performance index are proposed, and the relationship between the modified index and specifications on the autocorrelation function (suggested in Kozub and Garcia (1993) Huang and Shah (1998)) are also discussed.

Note the following properties of this modified controller performance index: i) if the process is operating at the desired user-defined benchmark, $\xi_{mod}(b) = 1$ ii) if the performance is "better" than the user-defined benchmark, $\xi_{mod}(b) < 1$, iii) if the process variance exceeds the user-defined benchmark, then $\xi_{mod}(b) > 1$, and iv) $\xi_{mod}(b) \geq 1 - \upsilon^2$. These properties provide a convenient "normalization" for the performance index.

Hard constraints. When the manipulated variable is at a hard constraint, the closed-loop is no longer described by Equation 4. However, it is possible to estimate σ^2_{mv} from routine data by including inputs and outputs in the time series model (Desborough and Harris, 1992). It is necessary to keep record of when the constraints are active, so that the model structure properly reflects the status of the control system. Manipulated variable constraints usually result in offset between controlled variables and their setpoint. Under such conditions, controller performance assessment can still be possible if the output(s) of interest are part of a multivariate predictive control scheme. A working solution in this case is to substitute the reachable target associated with the constrained output, say $Y^{\star}_{sp,t}$, for the setpoint in the calculation of the closed-loop error, i.e., $y_t = Y^{\star}_{sp,t} - Y_t$. The reachable target is internally calculated by the control algorithm and is simply a feasible value for the output of interest conditional on the active constraints. The estimates of σ^2_{mv} derived under such conditions may be suspect due to the influences of other input variables. In this situation, inspection of the closed-loop impulse response coefficients, which provide dynamic information on the output's tracking of the reachable target is recommended. In any case, when a controller is regularly switching between different sets of active constraints, benchmarking the dynamic performance may not be as important as monitoring how well the controller is meeting its overall design objectives, e.g., output prioritization and the distribution of offset.

Performance assessment with fixed controller structure. Most controllers employ a fixed structure, i.e., a Proportional-Integral-Derivative (PID) controller. It is of interest to develop performance monitoring and performance assessment methods for these widely used systems. Isaksson (1996), Ko and Edgar (1998) and Harris et al. (1999) have investigated these topics. Performance limitations arising from a fixed controller structure can only be determined if a process model is available. If opportunities for significant performance improvements are indicated using the minimum-variance methods, then one can determine the achievable limitations that arise from using a particular controller structure only by identifying a process and disturbance model. The use of previously identified models for assessment in a predictive control environment is discussed in a later section.

Detection of oscillations, valve stiction and other maladies. A number of researchers and practitioners have indicated that more realistic estimates of the achievable performance are obtained when one detects, diagnoses and "removes" the effect of oscillations (Owen et al., 1996; Owen, 1997; Horch, 2000). Methods for detecting oscillation and stiction are described in Hägglund (1995), Bittanti et al. (1997), Horch and Isaksson (1998, 1999), Seborg and Miao (1999), and Forsman (2000). Oscillations and valve stiction can be viewed as faults. There are other faults that beset control loops; the purpose of this paper is not to review this extensive literature (Isermann and Ballé, 1997). Rather, we indicate that automated procedures for control loop assessment using the methods proposed here, or descriptive statistics, must have proper data segmentation so that the presence of faults do not lead to improper interpretations or conclusions.

Nonlinear and time varying processes. In deriving the minimum variance controller, we assumed the process admits the description given in Equations 1 and 2. When the process is described by a nonlinear difference equation, either for the dynamics or disturbances, development of the nonlinear minimum variance controller may be very difficult or essentially impossible. This of course depends upon the structural form of the nonlinearity. For those descriptions which admit a nonlinear description and closed-form expressions for the minimum variance control law, it is possible to construct examples that show that the feedback invariance property does not exist. To ascertain performance bounds from routine operating data, one must assume that the process admits a local linear representation. The performance assessment results are "locally" valid. If changes in operating point cause changes in the process model, then the data must be properly segmented prior to analysis. Methods for detecting changes in model structure are discussed in Basseville (1998). If the disturbances are time-varying or consist of a mixture of stochastic and deterministic type disturbances, which is often the case, then the process description in Equations 1 and 2 must be expanded to account for this behavior. Again, methods for detecting these interventions must be part of the data analysis. The performance assessment techniques reviewed in this paper can then be applied to these types of processes (Harris et al., 1999).

Discussion

We point out that σ^2_{mv} may often not be a realizable performance bound due to the practical limitations de-

scribed above. It has been pointed out by a number of authors (Desborough and Harris, 1992; Huang et al., 1997b,c) that if $\sigma^2_{mv} < \sigma^2_y$ there *may* be opportunities to reduce the output variance. However, a diagnosis of the control system is required to investigate the cause(s) of variance inflation. If it suspected that non-invertible zeros or restrictions on the manipulated variables are limiting performance, then a process plus disturbance model must be identified to calculate $\sigma^2_{mv^\star}$, $\sigma^2_{\text{Åström}}$, or any other performance measure which requires knowledge of the process dynamics and disturbances. Alternatively, a number of the modified performance indices described above can be used to aid in the diagnosis of performance and detect changes in performance from a specified target value (Kozub, 1996).

Although the performance bounds and performance measures described in this section were originally introduced to ascertain how far the current performance was from minimum variance, they have found widespread use as a component of a more comprehensive performance monitoring and assessment methodology. Typically, industrial controller performance monitoring packages include some minimum variance-based performance statistics but also elementary descriptive statistics (such as mean, standard deviation, % uptime), histograms, power spectra, autocorrelation functions, impulse response functions and even non-linear valve diagnostics. Continuous performance monitoring applications also have significant information technology requirements such as access to historized data, dedicated servers, scheduling algorithms, and rule-based event notification and exception reporting (Jofriet et al., 1996). Guidelines for implementing univariate performance monitoring methods in practice are discussed in the references contained in Harris et al. (1999), Vishnubhotla et al. (1997), Thornhill et al. (1999), and Miller and Desborough (2000).

Multivariate Performance Assessment

The extension of performance assessment to multivariable systems has been studied by Harris et al. (1996), Huang et al. (1997a,b,c), and Huang and Shah (1998, 1999). Assessment of minimum variance performance bounds arising from deadtimes in MIMO systems requires knowledge of the interactor matrix. The interactor matrix allows a multivariate transfer function to be factored into two terms; one having its zeros located at infinity and another containing the finite zeros. To introduce this concept, consider a linear time-invariant process with n outputs and m inputs having transfer function $T(q^{-1})$. The interactor is a square matrix polynomial having the following properties (Dugard et al., 1984):

$$\lim_{q \to \infty} \xi(q) T(q^{-1}) = K \tag{15}$$

and

$$|\xi(q)| = q^B \tag{16}$$

where K is a non-singular matrix and B is the number of zeros of the transfer function located at infinity. In the univariate case, $\xi(q) = q^b$ and $B = b$. Other properties of the interactor matrix are discussed in Dugard et al. (1984), Goodwin and Sin (1984), Tsiligiannis and Svoronos (1988), Mutho and Ortega (1993), and Mutho (1995). It is important to note that the interactor matrix is not unique, and that it cannot always be constructed solely from knowledge of the delay structure. The interactor matrix can be constructed using linear algebra techniques from the process transfer function in the aforementioned references and Rogozinski et al. (1987). Huang et al. (1997a) have shown that the interactor matrix can be estimated from the Markov parameters of the process transfer function.

It is convenient to define the inverse-interactor matrix as follows:

$$\xi^{-1}(q^{-1}) = [\xi(q)]^{-1} = \xi_k q^{-k} + \cdots + \xi_d q^{-d} \tag{17}$$

where k is the minimum delay in the first row of the process transfer function, and d is not less than the maximum delay in the transfer function. Note that the bound on k shows that the interactor matrix is not unique; it can be altered by re-ordering the inputs and outputs.

Using the inverse interactor matrix, the process may be represented in right matrix fraction form as follows:

$$\begin{aligned} Y_t &= L(q^{-1}) R^{-1}(q^{-1}) U_t + D_t \\ &= \xi^{-1}(q^{-1}) \tilde{L}(q^{-1}) R^{-1}(q^{-1}) U_t + D_t \end{aligned} \tag{18}$$

where $\xi^{-1}(q^{-1})$ represents the inverse interactor matrix and D_t represents the process disturbance, which can often be modeled by a multivariate ARIMA process.

Once the interactor matrix is known, the multivariate extension of the univariate performance bounds can be established. Several methods can be used, all leading to equivalent results. Harris et al. (1996) define the performance bound:

$$\eta = 1 - \frac{E[Y^T_{mv} W Y_{mv}]}{E[Y^T_t W Y_t]} \tag{19}$$

where $E[\cdot]$ denotes mathematical expectation and $E[Y^T_{mv} W Y_{mv}]$ denotes the weighted multivariate minimum variance performance. W is a positive definite weighting matrix, which allows for differential weights on specific outputs. Determination of the multivariate minimum variance control performance requires that an all-pass representation of the interactor matrix be constructed. There are two general approaches for performing this: spectral factorization and construction of a spectral interactor.

Spectral Factorization

The spectral factor of the interactor matrix, $\gamma(q^{-1})$, is defined as the solution to the spectral factor equation:

$$\gamma^T(q)W\gamma(q^{-1}) = \xi^{-T}(q)W\xi^{-1}(q^{-1}). \tag{20}$$

Since $\xi^{-1}(q^{-1})$ is unimodular, the spectral factor $\gamma(q^{-1})$ is also unimodular. A property of a unimodular spectral factor is that $\gamma^{-1}(q^{-1})$ exists, and is a finite polynomial in q^{-1}. Efficient methods for the construction of the spectral factor involve solution of a bilinear set of equations, for which iterative application of Cholesky decompositions are very efficient (Kucera, 1979; Harris and Davis, 1992). These algorithms have quadratic convergence in a finite number of iterations when the polynomial matrix for which the spectral factor is to be obtained is unimodular.

Construction of the Spectral Interactor Matrix

Huang et al. (1997b) exploit the fact that the interactor matrix is not unique. They use a spectral interactor, $\tilde{\xi}^{-1}(q^{-1})$, introduced by Peng and Kinnaert (1992) and Bittanti et al. (1994) having the property:

$$\tilde{\xi}^{-T}(q)\tilde{\xi}^{-1}(q^{-1}) = I. \tag{21}$$

Linear algebra techniques can be used to construct the spectral interactor from the process transfer function or Markov parameters.

Once the all-pass filter representation has been obtained, it is possible to express the closed-loop system in the following form (Harris et al., 1996; Huang et al., 1997b):

$$S_t = \Psi_1(q^{-1})a_t + q^{-(d-1)}\Psi_2(q^{-1})a_t \tag{22}$$

where S_t is a filtered output, having the property that $E[S_t^T W S_t] = E[Y_t^T W Y_t]$. The terms $\Psi_1(q^{-1})a_t$ and $q^{-(d-1)}\Psi_2(q^{-1})a_t$ are uncorrelated. The first term on the right hand side of Equation 22 is a function only of the disturbance and the all-pass interactor matrix, and is otherwise independent of the dynamics of the process. This term is the multivariate equivalent of the system invariant, $(1 + \psi_1 q^{-1} + \cdots + \psi_{b-1}q^{-(b-1)})a_t$, encountered in univariate performance assessment. The term $\Psi_1(q^{-1})a_t$ represents the dynamics of the multivariate minimum variance controller. The second term in Equation 22 is a function of the controller, the process transfer function and the disturbances. In the derivation of Equation 22, it was assumed that a linear time-invariant controller was used.

Once the decomposition in Equation 22 has been affected, it is possible to calculate the performance index from Equation 19 as follows:

$$\eta = 1 - \frac{Tr\left(W\sum_{j=0}^{d-1}\Psi_j\Sigma_A\Psi_j^T\right)}{Tr\left(W\sum_{j=0}^{\infty}\Psi_j\Sigma_A\Psi_j^T\right)} \tag{23}$$

The two important properties encountered in the univariate case, namely the autocorrelation test and the invariance property, are also found in the multivariate extension. Once the spectral factor is obtained, one can also construct a portmanteau test for multivariate minimum variance control that is similar to the autocorrelation function (Harris et al., 1996). Performance bounds can be determined regardless of the number of inputs and outputs; there is no need that the process be "square". In the multivariate case it can also be shown that the minimum variance performance can be estimated from routine operating data if the interactor matrix is known, and there are several different methods for calculating the minimum variance performance bounds (Harris et al., 1996; Huang et al., 1997b). In the process of calculating the performance bound, it is necessary to fit a multivariate time series to the observations (when a linear controller is used). When constraints are active it is necessary to fit a predictive model to both the inputs and outputs. Haarsma and Nikolaou (2000) tested several identification methods in an application of multivariate performance assessment in the food processing industry. Other examples of the application of multivariate performance assessment and monitoring are given Harris et al. (1996), Huang et al. (1997a,b,c), Huang and Shah (1998, 1999), Miller and Huang (1997), and Huang et al. (2000).

The minimum variance controller described in Goodwin and Sin (1984) Dugard et al. (1984) is a sequential minimum variance controller that is dependent on the order of the inputs and outputs and choice of interactor representation. The construction of the all-pass filter representation of the interactor matrix leads to a "true" minimum variance controller, which is independent of these factors.

Remarks

1. Univariate and multivariate performance assessment are conceptually similar, however in the latter case knowledge of the time delay structure alone does not guarantee that the performance bounds can be calculated. Knowledge of the interactor matrix is an impediment to using multivariate techniques. Huang et al. (1997a) have shown that the interactor matrix may be calculated from the impulse coefficients of the process transfer function and have proposed a technique to estimate this from process data. This is akin to estimating the delay in an online fashion for SISO systems. However, this method requires that a dither signal be added to the process during the period of data collection. Furthermore, the method assumes that a linear, time-invariant controller be used during the period of data collection.

2. In the SISO case, an extended prediction horizon can be used for performance monitoring (Equa-

tion 12). This extended horizon serves two purposes: i) it provides an indication of the sensitivity of the performance index to the selection of delay, and ii) it indirectly addresses the issue that minimum variance control may not be the desired control objective. In multivariate analysis, one can also use a similar concept (Harris et al., 1996). Essentially, one replaces the interactor matrix by the term $\xi^{p+h}I$, where p is an estimate of the maximum order of the inverse-interactor matrix and $h > 0$. This approach does not enable calculation of the lower bound on performance; rather it is more useful in monitoring changes in the predictive structure of the process.

3. The performance bounds calculated using the interactor matrix are not restricted to those processes for which there are an equal number of inputs and outputs. In most cases, multivariable controllers are used where constraints are a factor. In these cases, the structure of the time-series model changes as constraints are engaged. One can adapt the structure of the time series model to reflect the evolving constraint set structure. There can be rather dramatic changes in the minimum variance performance bound when the set of constraint variables changes. The utility of the minimum variance performance bounds in these instances has yet to be determined. One can imagine that other performance measures may be more appropriate.
   Ko and Edgar (2000c,b) have proposed several methods to address performance assessment in the presence of constraints. Their work is based on the fundamental results of Furuta and Wongsaisuwan (1993, 1995) who show that a receding horizon controller, with input and output weightings, i.e., soft constraints, can be used to obtain the solution to the infinite horizon linear quadratic controller. In these papers, Furuta and Wongsaisuwan use the Markov parameters of the controller and disturbance (i.e., the impulse coefficients) to design the controller. Ko and Edgar (2000c,b) have used these results to provide a number of performance bounds. The method requires that an input/output model relating Y and U be available; a step response model used in the design of a predictive controller would suffice for this purpose. As usual, it is assumed that this model adequately describes the process. Given a record of $\{Y_t, U_t\}, t = 1..N$, the process disturbance can be reconstructed from the measurements as follows:

$$D_t = Y_t - T(q^{-1})U_t. \tag{24}$$

   A time series model is then fit to the D's. Once this time-series model has been determined, a number of performance bounds can be determined using the results of Furuta and Wongsaisuwan (1993, 1995). By applying zero weight to the inputs, Ko and Edgar (2000c,b) demonstrate that the minimum variance performance bound that corresponds to solution of the unconstrained linear quadratic minimum variance bound discussed in Harris et al. (1996) can be estimated. This bound equals the bound obtained from using the interactor matrix when the process transfer function has no non-invertible zeros, other than those associated with the time delay.
   By using a time series model for the disturbance, it is possible to simulate a generalized predictive controller, over the data set used to estimate the disturbance. By applying the same inputs and output constraints used in the actual controller, as well as the same prediction horizons for the inputs and outputs, an estimate of the performance using the identified disturbance structure is obtained. Recall that most receding horizon controllers assume that the disturbance is adequately modeled by a multivariate random walk. This bound correctly accounts for the presence of constraints.
   Both of these approaches, and other variations which can be derived from this approach, enable one to use a previously identified process model as part of the multivariate performance assessment process. A fundamental assumption is that this model is accurate, and that the disturbance model identified from Equation 24 has no model mismatch component.

4. When a more comprehensive model identification is undertaken it is possible to use more sophisticated performance measures. Kendra and Çinar (1997) have developed a frequency domain identification and performance assessment procedure for closed-loop multivariable systems. *A priori* information, such as design stage transfer function specifications, can be incorporated into the analysis. This allows model mismatch to be assessed, and makes possible comparisons of current operating performance to design specifications for the sensitivity and complementary sensitivity functions. External excitation must be provided in the form of a dither signal to enable identification of the sensitivity functions. Gustafsson and Graebe (1998) have developed a procedure to ascertain whether changes in closed-loop performance arise from changes in disturbance structure or changes in the process transfer function. A test signal must be applied for this analysis.

5. Intervention analysis provides a framework to incorporate variable setpoint changes, feedforward variables and deterministic disturbances in the univariate case. Analogous methods for the multivariable case have not been developed extensively.

One possible criticism of recent research developments in controller performance assessment is that too much

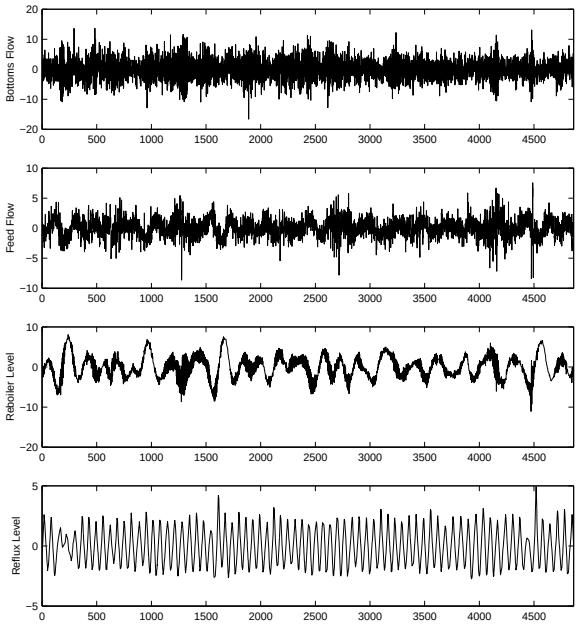


Figure 1: A time series plot showing the CV errors for the feed flow loop, the bottoms flow loop, the reboiler level loop, and the reflux level loop for the industrial distillation column. The trend plots show roughly five-thousand samples of one-minute data for each variable.

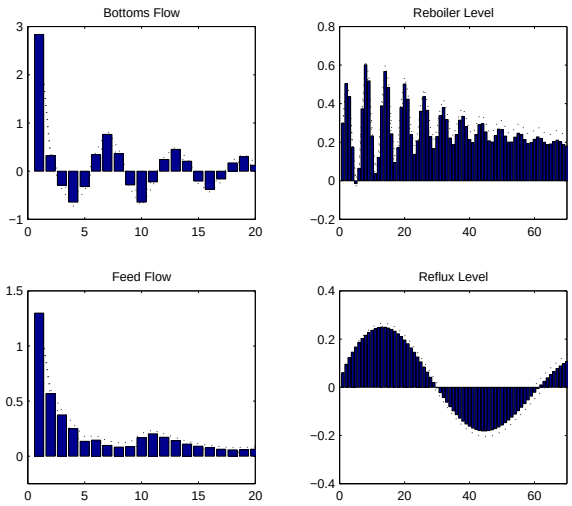


Figure 2: Univariate CV error impulse response functions. Clockwise, from the top left: the bottoms flow, the reboiler level, the feed flow, and the reflux drum level.

effort has been focused on estimating system invariants. While this is certainly one of the most interesting and challenging problems from an academic perspective, it is really only one of the many tools that an engineer might effectively use to monitor and/or analyze control system performance. The use of system invariants for multivariate performance assessment is a significant barrier to use due to the information requirements and the level of expertise needed to apply the methodology and interpret the results. Comprehensive methods for analyzing the interaction structure of the closed-loop are essential for diagnosing multivariate systems. In the next section we will demonstrate some analysis methods that can be derived using multivariable time series methods.

Example of Multivariate Process Analysis

In this section we will apply multivariate time series techniques to analyze an industrial data set. The objective of the analysis is to provide a qualitative and quantitative analysis of the closed-loop behavior. Our interest extends beyond the question as to whether or not the control system is operating at a desired performance benchmark.

Process Description

In this example, multi-output impulse response analysis will be used to study the dynamic relationships between four controlled variable (CV) error variables sampled from an industrial distillation column. The CV error vector was calculated from setpoint and output observations sampled at one-minute intervals from the following control loops on the column: the feed flow controller, the bottoms flow controller, the reboiler level controller, and the reflux drum level controller. Time series plots of these variables are shown in Figure 1. It is assumed that no prior information is available concerning the multivariable delay structure. Note, all of the modeling methods used in this section are standard results that have been adapted from the multivariate time series analysis literature (Hamilton, 1994; Lütkepohl, 1991).

Univariate Impulse Response Analysis

The first step was to estimate the closed-loop impulse weights, Equation 6, for each of the process variables. This was accomplished by fitting a univariate autoregressive model (using a least squares approach) to each variable and calculating the impulse weights by long division. The estimated impulse response plots are shown in Figure 2. The time horizon for the plots has been set to twenty minutes for the flow controllers (column one), and seventy minutes for the level controllers (column two).

Some deductions regarding the dynamic performance of each of these control loops can be made from the univariate impulse response functions. For example, strong cyclical behavior is observed in all the tracking error vari-

ables except for the feed flow. There is a one hour cycle in the reflux drum level, a seven minute cycle in the bottoms flow, and a seven minute cycle combined with another, slower, response in the reboiler level. The feed flow controller seems reasonably well tuned; it is free from overshoot or cycles, and damps out quickly. These are all valid observations, but no information regarding possible interactions between these loops can be made unless a multivariate analysis is performed. With the exception of the feed flow, none of the variables is close to their individual minimum variance performance bounds.

Multi-Output Impulse Response Analysis

An impulse response plot is simply a graphical representation of a time series model in moving average form. The displayed impulse response coefficients are the weights that describe the dynamic relationship between the input and the output. When the input is assumed to be a unit impulse, the impulse response plot shows the predicted output response. In the multi-output case, the $(i,j)^{th}$ entry in the $(n \times n)$ impulse response matrix gives the model weights between input driving force j and output i.

Note that in the current context of analyzing CV error dynamics, the multi-output impulse response estimates are based on routine operating data. In contrast, step response data to be used for identification is collected under experimental conditions where input variables are manipulated. So while models based on the latter approach can be considered causal, the same is not true for the former. If the underlying data has not been collected during an experiment, the tracking error impulse response plots simply help the analyst interpret the correlation structure between the tracking error trends, not the true causal relationships.

Modeling Control Error Trends—Vector Time Series Approach

Multi-output control error trends can be considered a group of univariate control error trends of equal length that all share the same time stamp. Rather than being a scalar at time t, a multi-output control error trend is an n dimensional vector at time t, with one element for each of the n controlled variable (CV) error trends. Define the following vector time series:

$$\mathbf{y}_t = \mathbf{Y}_{sp,t} - \mathbf{Y}_t \tag{25}$$

where $\mathbf{y}_t$, $\mathbf{Y}_{sp,t}$, and $\mathbf{Y}_t$ are vectors representing the control error, the output, and the setpoint, respectively. In practice, one would typically be working with $\mathbf{y}$, an $(N \times n)$ array of CV error data, based on N samples of n CV error trends.

Treating the dynamic analysis of multi-output control error trends as an endogenous estimation problem with no *a priori* information has been explored by Seppala (1999). Linear dynamic approximations of endogenous system behavior with no *a priori* information and no assigned input/output structure had been previously used in the field of applied econometrics. The simplest multivariate dynamic model that can represent $\mathbf{y}_t$ is a vector autoregressive (VAR) model which is written as follows:

$$\Phi(q^{-1})\mathbf{y}_t = \mathbf{a}_t \tag{26}$$

where $\mathbf{a}_t$ is a vector of driving forces, and $\Phi(q^{-1})$ is an autoregressive matrix polynomial defined as:

$$\Phi(q^{-1}) = I_n + \Phi_1 q^{-1} + \cdots + \Phi_p q^{-p} \tag{27}$$

where each Φ_i is an $(n \times n)$ coefficient matrix. The expanded form of Equation 26 is clearly analogous to a scalar autoregressive model; each variable in $\mathbf{y}_t$ is expressed as a function of lagged values of itself and the other $(n-1)$ variables in $\mathbf{y}_t$:

$$\mathbf{y}_t = -\Phi_1 \mathbf{y}_{t-1} - \cdots - \Phi_p \mathbf{y}_{t-p} + \mathbf{a}_t \tag{28}$$

where p is the autoregressive model order. The driving force covariance matrix, Σ_a, is given by:

$$\Sigma_a = E[\mathbf{a}_t \mathbf{a}_t^T]. \tag{29}$$

The diagonal elements of Σ_a are the driving force variances, and the off-diagonal elements are the driving force covariances.

To find the multi-output impulse responses, one proceeds in much the same fashion as in the univariate case. If the autoregressive matrix polynomial in Equation 27 is stable, then the VAR model for $\mathbf{y}_t$ may be expressed in vector moving average (VMA) form:

$$\begin{aligned} \mathbf{y}_t &= \Theta(q^{-1})\mathbf{a}_t \\ &= (1 + \Theta_1 q^{-1} + \cdots + \Theta_r q^{-r})\mathbf{a}_t \\ &= \sum_{i=0}^{r} \Theta_i \mathbf{a}_t - i \end{aligned} \tag{30}$$

where $\Theta(q^{-1})$ is the vector moving average matrix polynomial, defined such that $\Phi(q^{-1})\Theta(q^{-1}) = I_n$. The Θ_i's can be found using the recursion:

$$\Theta_i = \sum_{j=1}^{i} \Theta_{i-j}\Phi_j \tag{31}$$

where $i = 1, 2, \ldots,$ and $\Theta_0 = I_n$. The Θ_i coefficient matrices contain the multi-output impulse response coefficients.

The VMA model for $\mathbf{y}_t$ in Equation 30 is not unique; a property of many types of multivariate models. To illustrate the non-uniqueness property, consider any non-singular matrix P: the Θ_i matrices can be replaced by $\Psi = \Theta_i P$, and the driving forces can be replaced by $\mathbf{v}_t = P\mathbf{a}_t$, resulting in the following equivalent model for $\mathbf{y}_t$:

$$\mathbf{y}_t = \sum_{i=0}^{r} \Theta_i P P^{-1} \mathbf{a}_{t-i} = \sum_{i=0}^{r} \Psi_i \mathbf{v}_{t-i} \tag{32}$$

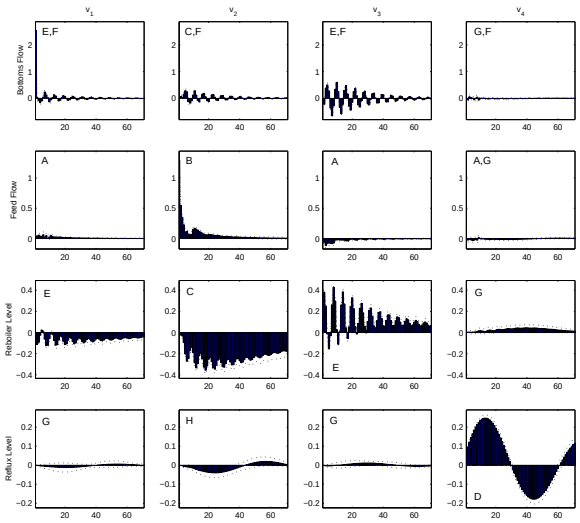


Figure 3: Multi-output CV error impulse response plot matrix for the four variable system consisting of the bottoms flow, the feed flow, the reboiler level and the reflux level. These responses were computed from an eighth order VAR model converted to orthogonal innovations form VMA via a Cholesky factorization of the covariance matrix. This plot shows the first seventy steps of the response.

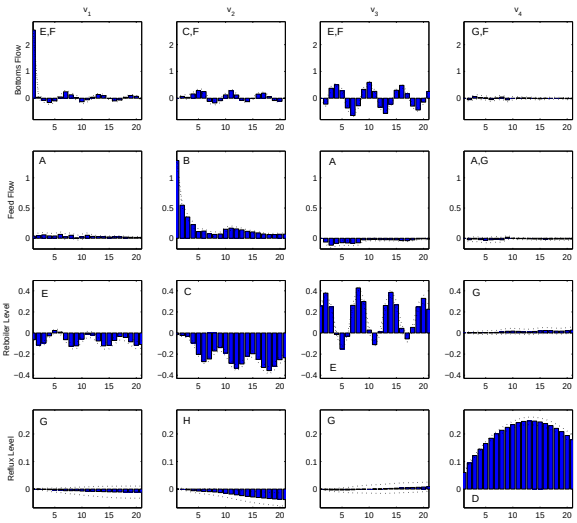


Figure 4: Multi-output CV error impulse response plot matrix for the four variable system consisting of the bottoms flow, the feed flow, the reboiler level and the reflux level. These responses were computed from an eighth order VAR model converted to orthogonal innovations form VMA via a Cholesky factorization of the covariance matrix. This plot shows the first twenty steps of the response.

The models in Equations 30 and 32 are equivalent in the sense that they produce identical estimates of the k-step ahead forecast error covariance. The non-uniqueness property of the VMA model can be used to choose a particular P that orthogonalizes the driving forces, thereby simplifying multi-output impulse response analysis and variance calculations. A common choice is to select P so that it is the Cholesky factor of the driving force covariance matrix, resulting in orthogonalized driving forces $\mathbf{v}_t$. When one interprets the impulse response coefficient matrices, the Ψ_i's, one can consider the effects of shocks to the driving force processes one-at-a-time because they are orthogonal. This topic and other methods for analyzing multi-output control systems are discussed in Seppala et al. (2001) and Seppala (1999).

Results of Multivariate Analysis

An eighth order VAR model was found to adequately model the control error data in Figure 1, i.e., Equation 28 was used with $n = 4$ and $p = 8$. Standard residual analysis showed the model to be adequate, and the residuals themselves to be nearly orthogonal. The impulse response form of the multivariate model for the distillation column data is shown in Figure 3 with a time horizon of seventy minutes, and in Figure 4 with a time horizon of twenty minutes. In these figures, the rows (from top to bottom) represent the four CV error variables (bottoms flow, feed flow, reboiler level, and reflux level) and the columns represent the system's driving forces. Because the driving forces are essentially orthogonal, the driving force in column i can be interpreted as a deviation from setpoint or disturbance in variable i, and the coefficients plotted in position (i, i) of the impulse response matrix can be considered the endogenous component of the response for variable i.

Non-significant interactions. The responses in the subplots labeled A and/or G in Figures 3 and 4 contain non-significant relationships because the 95% confidence intervals for the responses contain zero across the entire time horizon. This can be interpreted as a lack of significant correlation between a deviation from setpoint or a disturbance in variable i and the presence of any corresponding response in variable j. From Figures 3 and 4 it can be seen that the model shows non-significant interactions along the dynamic pathways from: i) the bottoms flow to the feed flow, ii) the bottoms flow to the reflux level, iii) the reboiler level to the feed flow, iv) the reboiler level to the reflux level, and v) the reflux level to any other system variable.

The feed flow error. Inspection of the second row of plots in Figures 3 and 4 reveals that only driving force v_2 (recall, v_2 is interpreted as an impulse-like upset in the feed flow) has a significant effect on the feed flow. This aspect of the model makes sense physically because in this example, only upstream properties affect the feed flow controller. If v_2 is the only driving force significantly

correlated with the feed flow control error, then the univariate and multivariate impulse response plots should be similar for the feed flow tracking error.

How do the other system variables respond to a shock in the feed flow loop? The plots labeled C show that after a couple of units of delay, the bottoms flow and the reboiler level both show significant responses with a strong cyclical component at seven minutes. The plot labeled H shows that a shock in the feed flow has a small, perhaps negligible, effect on the reflux level.

The reflux level error. The reflux level tracking error exhibits a dominant cycle with a period of about 60 minutes. Since all the responses labeled G in Figures 3 and 4 are non-significant, it can be deduced that the reflux level control error is unrelated to the other three tracking error variables over the period of data collection. Note also that the impulse response plot in pane D is essentially identical to the univariate impulse response plot for the reflux level in Figure 2. The multi-output impulse response plots in column four provide statistical evidence that errors in the reflux level controller are unrelated to dynamics in the other three loops during the period of data collection.

The reboiler level/bottoms flow pair. In Figures 3 and 4, the rows showing the responses for the bottoms flow (row 1) and the reboiler level (row 3) indicate that there is a strong seven minute cycle shared by both CV error variables in this cascaded pair. The multivariate model that has been estimated shows that an upset in any of the CV error variables except the reflux level error is related to this statistically significant response in the bottoms flow/reboiler level pair. As mentioned above, deductions regarding causality are out of the question, but the analysis shows that a common cycle exists between the reboiler level error and the bottoms flow error, and that the error in at least one external loop (the feed flow) is correlated with this pair.

Industrial Perspectives

At CPC V, controller performance monitoring was categorized as a new direction for academic research. Since then, there has been considerable research in this area, with a significant focus of this work directed towards developing controller performance monitoring (CPM) techniques for multivariable systems. Most of the work in this area has focused on developing multivariable system invariants, with more recent work addressing incorporation of constraints. There are other topics requiring attention, and we shall indicate a few that are of industrial interest.

As CPM matures as a technology, and as its acceptance becomes more widespread, the question of how to affect CPM solutions arises. Since 1996, a number of commercial products and services for CPM have appeared. There are an enormous number of challenges in developing, supporting and ensuring that these packages are used effectively. We will provide a short discussion of these challenges. Finally, a brief discussion on the relative merits of CPM solutions as vendor products or as vendor services is included.

New Areas for Research

Multivariate predictive controllers have an optimization layer in the form of a linear or quadratic program, and some plants have real time optimization (RTO) systems downloading targets to Multivariable Predictive Controllers (MPCs) or to the base level control system. What is there to be learned by monitoring the behavior of these targets? On the time scale of control systems, can this data be considered dynamic? RTO targets arrive on the order of hours and can be considered static; however, over-active optimization targets that cannot be considered static from the point of view of the control system have been observed. What effect does would the latter have on one's perception of control system performance?

The use of dynamic analysis of variance (ANOVA), i.e., studying the correlation and quantifying the variance propagation between key control system variables was used by Desborough and Harris (1993) to analyze multiple-input-single-output controlled systems. Ideally, one would use ANOVA methods to identify process variables that are chiefly responsible for variance inflation of key controlled variables. The technical challenges are well-known: causal ordering of upstream variables, the effects of feedback and recycle, collinearity of disturbances, and a dealing with the component of variance propagation due to invariants. Further study of this topic is warranted.

Several methods for modeling multivariate dynamic and/or static data have matured to the point where powerful software packages are now available for their application. A couple of examples well known in the control engineering community are: ADAPTX (Subspace ID), Simca-P (PCA/PLS), and the host of data analysis tools available for the Matlab$^{(R)}$ environment. In combination with process knowledge, these data analysis packages can be very useful for analysis and diagnosis. Note, however, that batches of control system data that have been gathered because they came from a previously identified problem area are good for analysis, but this does not count as monitoring. True performance monitoring requires constant, scheduled contact with the plant information system, and this has been known for some time. There are many practical challenges with real-time applications: data integrity, fault-detection, robust algorithms, data visualization and presentation, to name a few.

CPM in the Field: A Product or A Service?

Many practicing control specialists are now at least familiar with controller performance monitoring. A number of prototype industrial controller performance monitoring systems have been described in the open literature, Harris et al. (1999). Commercial products are available from Honeywell, Matrikon and Control Arts to name a few. When one considers purchasing CPM capabilities for a plant, besides the obvious issues of the level of technology required, the issue of whether to buy a CPM product, or a CPM service emerges. Using a CPM product would be like using any other piece of installed software; essentially one has access to on-line help and product support. Pursuing a CPM service could involve engineers visiting the site to perform control performance audits, service providers consulting on difficult CPM problems, or an electronic exchange of raw data for loop performance reports (Miller and Desborough, 2000).

CPM products and services will both be costly, and both will require support from the provider. Without trying to answer the question of which model is better, CPM as a product or CPM as a service, some of the important issues will be outlined below.

Whether CPM is used as a product or a service, proper training is required if plant operation going to benefit from CPM. One of the frustrations with applications of CPM, and other quality monitoring methods, is the level of training for individuals who are asked to use these methods. Although CPM technology is not as wide in scope or as complex as multivariate predictive control, the availability of training in the latter area far exceeds that which is available for CPM. As with most statistical methods, attention must be paid to the length of data and sampling interval used for analysis, the type of filtering used prior to analysis (such as compressed data) and other aspects of data integrity.

Process knowledge has long been known to be an essential ingredient to successful application of CPM in the field (Jofriet and Bialkowski, 1996; Haarsma and Nikolaou, 2000; Horch, 2000). In order to integrate CPM into engineers' work practices, regular hands-on experience with CPM is required to develop skills. CPM products are best suited for this, because the product becomes just another tool, one that does not rely on a third party to use successfully. A particular challenge is that advanced multivariate techniques, which require *a priori* structural information and advanced system identification techniques, may only be successful when used by experts in CPM. Widespread use by control engineers requires automation of most of the methodology, with an emphasis placed on interpretation and analysis. These requirements are not different than those encountered in applications of multivariate statistical process control. Finally, we note that control engineers working in environments where constraints on available funds, time, and support personnel are limiting, are the least likely to get involved in CPM. In this situation, if CPM is to be implemented at all, then the service model is probably more appropriate.

Conclusions

The use of controller performance monitoring and assessment tools in industrial settings has grown considerably in the past several years. Extensions and variations of minimum variance based methods have been used extensively, primarily due to ease of understanding, robust computational methods, and minimal requirements for *a priori* knowledge. Industrial versions of these packages are available as both products and services.

Since CPC V, there has been considerable development of the underlying theory for multivariable controller performance assessment methods. Two main approaches to multivariate controller performance assessment have emerged thus far. The first method requires the use of extensive *a priori* process knowledge. In particular, the use of previously identified process models, which enables constraint handling to be addressed in a logical and straightforward fashion. The outcome of enables one to ascertain performance bounds and thus to subsequently monitor changes in these bounds over time. Of course, the results of the analysis are interpreted presuming that the process model is correct. The second approach largely dispenses with the requirement for *a priori* knowledge. Empirical models are built and used to analyze the predictive structure of the data. In particular, process interactions, and variance-decompositions over time can be used to help diagnosis process interactions. With such an approach, one is not restricted to using time series models; many multivariate statistical methods can be used if they are modified to include lagged data to account for serial correlation. The two approaches share common features, and it is clear that they are non-trivial generalizations of the univariate measures. It remains to be shown whether the more demanding and complex multivariate methods can be successfully integrated into a plant-wide monitoring and assessment strategy.

Most approaches for performance assessment use data collected in a passive mode or use data generated when significant events occur (Isaksson et al., 2000; Stanfelj et al., 1993). This is one of the key attributes of the performance measures—one performs the analysis with representative data. When poor performance is detected, a combination of statistical tools and process knowledge is required to analyze and diagnose the underlying problems. The role of designed experiments, in either closed-loop or open-loop, to aid in the diagnosis and analysis is an area requiring attention. Preliminary results have been reported by Kendra and Çinar (1997) and Gustafs-

son and Graebe (1998). Since many of the proposed methods for accommodating constraints require that the process transfer function model be known, on-line and off-line methods for model validation are essential for these techniques to be used with confidence. The development of diagnostics for model-based control is an open area for research (Kesavan and Lee, 1997).

The focus of much of this paper has been on the performance measures themselves. Large-scale industrial applications require incorporation of such performance measures into a plant-wide monitoring and performance assessment package. Industrial experience indicates that many of the challenges to broader application of performance measures lie in the successful development and maintenance of such systems.

Acknowledgments

TJH would like to thank the Natural Sciences and Engineering Research Council of Canada and Equilon Enterprises LLC for their financial support.

References

Åström, K. J., "Computer Control of a Paper Machine—An Application of Linear Stochastic Control Theory," *IBM Journal, July 1967*, pages 389–404 (1967).

Åström, K. J., *Introduction to Stochastic Control Theory*. Academic Press, London (1970).

Basseville, M., "On-Board Component Fault Detection and Isolation Using the Statistical Local Approach," *Automatica*, **34**, 1391–1415 (1998).

Bergh, D. and J. F. MacGregor, "Constrained Minimum Variance Controllers: Internal Model Structure and Robustness Propperties," *Ind. Eng. Chem. Res.*, **26**, 1558–1564 (1987).

Bittanti, S., P. Colaneri, and M. F. Mongiovi, The Spectral Interactor Matrix for the Singular Riccatti Equation, In *Proc. 33rd CDC*, pages 2165–2169 (1994).

Bittanti, S., M. Campi, and S. M. Savaresi, "Unbiased Estimation of a Sinusoid in Colored Noise via Adapted Notch Filters," *Automatica*, **33**(2), 209–215 (1997).

Box, G. E. P. and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. Holden Day, San Francisco (1976).

De Vries, W. R. and M. Wu, "Evaluation of Process Control Effectiveness and Diagnosis of Variation in Paper Basis Weight via Multivariate Time-Series Analysis," *IEEE Trans. Auto. Cont.*, **23**, 702–708 (1978).

Desborough, L. D. and T. J. Harris, "Performance Assessment Measures for Univariate Feedback Control," *Can. J. Chem. Eng.*, **70**, 1186–1197 (1992).

Desborough, L. D. and T. J. Harris, "Performance Assessment Measures for Univariate Feedforward/Feedback Control," *Can. J. Chem. Eng.*, **71**, 605–616 (1993).

Dugard, L., G. C. Goodwin, and X. Xianya, "The Role of the Interactor Matrix in Multivariable Stochastic Adaptive Control," *Automatica*, **20**, 701–709 (1984).

Foley, M. and T. J. Harris, "Performance and Structure of H_∞ Controllers," *J. Optim. Cont. Appl. Meth.*, **13**, 1–28 (1992).

Forsman, K., On Detection and Classification of Valve Stiction, In *TAPPI Conf. on Process Control*, Williamsburg, VA (2000).

Furuta, K. and M. Wongsaisuwan, "Closed-Form Solutions to Discrete-Time LQ Optimal Control and Disturbance Attenuation," *Sys. Cont. Let.*, **20**, 427–437 (1993).

Furuta, K. and M. Wongsaisuwan, "Discrete-Time LQG Dynamic Controller Design Using Plant Markov Parameters," *Automatica*, **31**, 1317–1324 (1995).

Goodwin, G. C. and K. S. Sin, *Adaptive Filtering Prediction and Control*. Prentice-Hall, Englewood Cliffs, N. J. (1984).

Gustafsson, F. and S. F. Graebe, "Closed-Loop Performance Monitoring in the Presence of System Changes and Disturbances," *Automatica*, **34**, 1311–1326 (1998).

Hägglund, T., "A Control-Loop Performance Monitor," *Control Eng. Practice*, **5**(11), 1543–1551 (1995).

Haarsma, G. and M. Nikolaou, Multivariate Controller Performance Monitoring: Lessons from an Application to a Snack Food Process (2000). Submitted to *Journal of Process Control*.

Hamilton, J. D., *Time Series Analysis*. Princeton University Press, Princeton (1994).

Harris, T. J. and J. H. Davis, "An Iterative Method for Matrix Spectral Factorization," *SIAM J. Sci. Statist. Comput.*, **13**, 531–540 (1992).

Harris, T. J. and J. F. MacGregor, "Design of Multivariable Linear-Quadratic Controllers Using Transfer Functions," *AIChE J.*, **33**, 1481–1495 (1987).

Harris, T. J., F. Boudreau, and J. F. MacGregor, "Performance Assessment of Multivariable Feedback Controllers," *Automatica*, **32**, 1505–1518 (1996).

Harris, T. J., C. T. Seppala, and L. D. Desborough, "A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems," *J. Proc. Cont.*, **9**(1), 1–17 (1999).

Harris, T. J., "Assessment of Control Loop Performance," *Can. J. Chem. Eng.*, **67**, 856–861 (1989).

Harris, T. J., Confidence estimation for performance assessment measures (2001). In Preparation, Queen's University, Kingston.

Horch, A. and A. J. Isaksson, A method for detection of stiction in control valves, In *IFAC Workshop on On-Line-Fault Detection and Supervision in the Chemical Process Industry*, Lyon, France (1998).

Horch, A. and A. J. Isaksson, "A Modified Index for Control Performance Assessment," *J. Proc. Cont.*, **9**, 475–483 (1999).

Horch, A., *Condition Monitoring of Control Loops*, Ph. D. Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden (2000).

Huang, B. and S. L. Shah, "Practical Issues in Multivariable Feedback Control Performance Assessment," *J. Proc. Cont.*, **8**, 421–430 (1998).

Huang, B. and S. L. Shah, *Performance Assessment of Control Loops*. Springer-Verlag, London (1999).

Huang, B., S. L. Shah, and K. E. Kwok, "Good, Bad or Optimal? Performance Assessment of Multivariable Processes," *Automatica*, **33**(6), 1175–1183 (1997b).

Huang, B., S. L. Shah, K. E. Kwok, and J. Zurcher, "Performance Assessment of Multivariate Control Loops on a Paper-Machine Headbox," *Can. J. Chem. Eng.*, **75**, 134–142 (1997c).

Huang, B., S. L. Shah, and H. Fujii, "The Unitary Interactor Matrix and its Estimation Using Closed-Loop Data," *J. Proc. Cont.*, **7**(3), 195–207 (1997a).

Huang, B., S. L. Shah, and R. Miller, "Feedforward Plus Feedback Controller Performance Assessment of MIMO Systems," *IEEE Trans. Cont. Sys. Tech.*, **3**, 580–587 (2000).

Isaksson, A. J., A. Horch, and G. A. Dumont, Event-triggered deadtime estimation: A comparison of methods, In *Proc. Control Systems 2000*, pages 209–215, Victoria, BC (2000).

Isaksson, A. J., PID Controller Performance Assessment, In *Proc. Control Systems '96*, pages 187–194, Halifax, NS (1996).

Isermann, R. and P. Ballé, "Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes," *Control Eng. Practice*, **5**(5), 709–719 (1997).

Jofriet, P. and B. Bialkowski, Process Knowledge: The Key to On-Line Monitoring of Process Variability and Control Loop Performance, In *Proc. Control Systems 96*, pages 163–169, Halifax, NS (1996).

Jofriet, P., C. Seppala, M. Harvey, B. Surgenor, and T. Harris, "An Expert System for Control Loop Performance," *Pulp and Paper Canada*, **6**, 207–211 (1996).

Kendra, S. J. and A. Çinar, "Controller Performance Assessment by Frequency Domain Techniques," *J. Proc. Cont.*, **7**(3), 181–194 (1997).

Kesavan, P. and J. H. Lee, "Diagnostic Tools for Multivariable Model-Based Control Systems," *Ind. Eng. Chem. Res.*, **36**, 2725–2738 (1997).

Ko, B. and T. F. Edgar, Assessment of Achievable PI Control Performance for Linear Processes with Deadtime, In *Proc. American Control Conference*, pages 1548–1552, Philadelphia, PA (1998).

Ko, B. and T. F. Edgar, "Performance Assessment of Cascade Loops," *AIChE J.*, **46**, 281–291 (2000a).

Ko, B. and T. F. Edgar, Performance Assessment of Constrained Model Predictive Control Systems (2000b). Submitted to *AIChE Journal*.

Ko, B. and T. F. Edgar, Performance Assessment of Multivariable Feedback Control Systems (2000c). Submitted to *Automatica*.

Kozub, D. J. and C. E. Garcia, Monitoring and Diagnosis of Automated Controllers in the Chemical Process Industries, In *AIChE Proceedings*, St. Louis, MO (1993).

Kozub, D. J., Monitoring and Diagnosis of Chemical Processes with Automated Process Control, In *CPC V Proceedings*, Lake Tahoe (1996).

Kucera, V., *Discrete Linear Control: The Polynomial Equation Approach*. John Wiley & Sons, New York (1979).

Lütkepohl, H., *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin (1991).

Miller, R. M. and L. D. Desborough, "Web-enabled control loop assessment," *Chemical Engineering* (2000).

Miller, R. M. and B. Huang, Perspectives on Multivariate Feedforward/Feedback Controller Performance Measures for Process Diagnosis, In *Proc. IFAC Adchem 97*, pages 435–440, Banff, AB (1997).

Mutho, Y. and R. Ortega, "Interactor Structure Estimation for Adaptive Control of Discrete-Time Multivariable Nondecouplable Systems," *Automatica*, **29**, 635–647 (1993).

Mutho, Y., "A note on the invariant properties of ineractors," *Int. J. Control*, **62**, 1247–1252 (1995).

Owen, J. G., D. Read, H. Blekkenhorst, and A. A. Roche, A Mill Prototype for Automatic Monitoring of Control Loop Performance, In *Proc. Control Systems 96*, pages 171–178, Halifax, NS (1996).

Owen, J. G., Automatic control loop monitoring and diagnostics (1997). Patent Application. Int. Publ No WO 97/41494, Int. Appl. No PCT/CA97/00266.

Peng, Y. and M. Kinnaert, "Explicit Solution to the Singular LQ Regulation Problem," *IEEE Trans. Auto. Cont.*, **37**, 633–636 (1992).

Qin, S. J., "Control Performance Monitoring—A Review and Assessment," *Comput. Chem. Eng.*, **23**, 173–186 (1998).

Rogozinski, M. W., A. P. Paplinski, and M. J. Gibbard, "An Algorithm for Calculation of a Nilpotent Interactor Matrix for Linear Multivariable Systems," *IEEE Trans. Auto. Cont.*, **32**, 2234–2237 (1987).

Seborg, D. and T. Miao, Automatic detection of excessively oscillatory feedback control loops, In *IEEE Int. Conf. on Control Applications*, pages 359–364, Hawaii, USA (1999).

Seppala, C. T., T. J. Harris, and D. W. Bacon, Time Series Methods for Dynamic Analysis of Multiple Controlled Variables (2001). Submitted to *Journal of Process Control*.

Seppala, C. T., *Dynamic Analysis of Variance Methods for Monitoring Control System Performance*, Ph. D. Thesis, Queen's University, Canada (1999).

Stanfelj, N., T. E. Marlin, and J. F. MacGregor, "Monitoring and Diagnosing Process Control Performance: The Single-Loop Case," *Ind. Eng. Chem. Res.*, **32**, 301–314 (1993).

Thornhill, N. F., M. Oettinger, and P. Fedenczuk, "Refinery-wide control loop performance assessment," *J. Proc. Cont.*, **9**, 109–124 (1999).

Tsiligiannis, C. A. and S. A. Svoronos, "Dynamic Interactors in Multivariable Process Control—I: The General Time Delay Case," *Chem. Eng. Sci.*, **43**, 339–347 (1988).

Tyler, M. L. and M. Morari, Performance assessment for unstable and nonminimum-phased systems, In *Proc. IFAC Workshop on On-Line Fault Detection and Supervision in the Chemical Process Industries*, pages 200–205, Newcastle-upon-Tyne, UK (1995).

Tyler, M. L. and M. Morari, "Performance Monitoring of Control Systems Using Likelihood Methods," *Automatica*, **32**, 1145–1162 (1996).

Vishnubhotla, A., S. L. Shah, and B. Huang, Feedback and Feedforward Performance Analysis of the Shell Industrial Closed Loop Data Set, In *Proc. IFAC Adchem 97*, pages 295–300, Banff, AB (1997).

Simultaneous Design and Control Optimization under Uncertainty in Reaction/Separation Systems

Efstratios N. Pistikopoulos* and Vassilis Sakizlis
Centre for Process Systems Engineering
Department of Chemical Engineering
Imperial College, London SW7 2BY, U.K.

Abstract

In this paper, we overview recent advances towards the integration of process design, process control and process operability in separation and reaction/separation systems that were developed within our group at Imperial College. Based on novel mixed integer dynamic optimization algorithms, a simultaneous strategy is presented featuring high fidelity dynamic models, explicit consideration of structural process and control design aspects (such as number of trays, pairing of manipulated and controlled variables) through the introduction of 0-1 variables, and explicit consideration of time-varying disturbances and time-invariant uncertainties. The application of this strategy to two typical (a separation and a reactive separation) systems is discussed.

Keywords

Process design, Process control, Uncertainty, Operability, Mixed-integer dynamic optimization

Introduction

The need to consider operability issues at an early phase of process design is now becoming widely accepted in both academia and industry. As a result of this, in recent years, a number of methodologies and tools have been reported for taking account of the interactions between process design and process control, with well over fifty publications since 1982 plus several international workshops and dedicated conference sessions (see Van Schijndel and Pistikopoulos, 2000). Despite these developments, however, it is observable that a large proportion of the work in this field:

- has concentrated on the application of metrics (e.g., condition number) that provide some measure of a system's controllability, but may not relate directly and unambiguously to real performance requirements;
- relies on steady-state or simple, usually linear dynamic models for processes;
- does not account for the presence of both time-varying disturbances and time-invariant (or relatively slowly varying) uncertainties; and
- does not involve selection of the best process design and the best control scheme, taking into account both discrete and continuous decisions.

Van Schijndel and Pistikopoulos (2000) also put forward a number of key challenges that lie ahead in the area of Process Design for Operability. One such challenge is the need for a rigorous and efficient solution of the underlying optimization problem, which is at the heart of the mathematical representation of the simultaneous process and control design problem.

*Tel: +44 20 7594 6620. Fax: +44 20 7594 6606. E-mail: e.pistikopoulos@ic.ac.uk

The aim of this paper is to give a brief overview of some recent advances, towards this endeavor, and its application to typical separation and reactive separation problems, carried out whithin our group at Imperial College.

Simultaneous Design and Control Under Uncertainty Framework

As discussed in Van Schijndel and Pistikopoulos (2000), the problem of the integration of process design, process control and process operability can be conceptually posed as follows:

$$\text{minimize} \quad \text{Expected Total Annualized Cost} \qquad \text{(P)}$$

subject to

Differential-Algebraic Process Model
Inequality Path Constraints
Control Scheme Equations
Process Design Equations
Feasibility of Operation (over time)
Process Variability Constraints

To determine Process and Control Design

A general, algorithmic framework for solving (P) was proposed by Mohideen et al. (1996). Its steps, schematically shown in Figure 1, can be summarized as follows:

Step 1. Choose an initial set of scenarios for the uncertain parameters.

Step 2. For the current set of scenarios, determine the optimal process and control design by solving the (multi-period) mixed-integer dynamic optimization

(MIDO) problem:

$$\min_{\substack{\mathbf{d},\mathbf{y},\\ \mathbf{u_v^1},\mathbf{u_v^2},\ldots,\mathbf{u_v^{ns}}}} \sum_{i=1}^{ns} w_i.\phi\left(\dot{\mathbf{x}}^i(t_f),\mathbf{x}^i(t_f),\mathbf{z}^i(t_f),\mathbf{u_t}^i(t_f),\mathbf{u_v}^i,\boldsymbol{\nu}^i(t_f),\boldsymbol{\theta}^i,\mathbf{d},\mathbf{y},t_f\right) \quad (1)$$

subject to

$$\begin{aligned}\mathbf{f}\left(\dot{\mathbf{x}}^i(t),\mathbf{x}^i(t),\mathbf{z}^i(t),\mathbf{u_t}^i(t),\mathbf{u_v}^i,\boldsymbol{\nu}^i(t),\boldsymbol{\theta}^i,\mathbf{d},\mathbf{y}\right)&=\mathbf{0}\\ \mathbf{c}\left(\mathbf{x}^i(t),\mathbf{z}^i(t),\mathbf{u_t}^i(t),\mathbf{u_v}^i,\boldsymbol{\nu}^i(t),\boldsymbol{\theta}^i,\mathbf{d},\mathbf{y}\right)&=\mathbf{0}\\ \mathbf{g}\left(\dot{\mathbf{x}}^i(t),\mathbf{x}^i(t),\mathbf{z}^i(t),\mathbf{u_t}^i(t),\mathbf{u_v}^i,\boldsymbol{\nu}^i(t),\boldsymbol{\theta}^i,\mathbf{d},\mathbf{y}\right)&\le\mathbf{0}\\ i=1,\ldots,ns&\end{aligned}$$

where $\mathbf{d}$ includes the continuous process design variables and controllers' tuning parameters; $\mathbf{y}$ comprises the binary variables for the process and the control structure (corresponding to whether a manipulated variable is paired with a particular controlled variable or not); $\mathbf{u_v}$ is the set of time-invariant operating variables; i is the index set for the scenarios of the uncertain parameters $\boldsymbol{\theta}$; ns is the number of scenarios; $w_i,\ i=1,\ldots,ns$, are discrete probabilities for the selected scenarios ($\sum_{i=1}^{ns} w_i = 1$); ϕ is a cost function; $\mathbf{x}(t)$ is the vector of differential states; $\mathbf{z}(t)$ is the vector of algebraic variables; $\mathbf{u_t}(t)$ denotes the set of time-varying manipulated (control) variables; $\nu(t)$ represents the time-varying disturbances; $\mathbf{f}=\mathbf{0}$ and $\mathbf{c}=\mathbf{0}$ represent the differential and algebraic equations (DAEs), respectively, for the process and control system, for which consistent initial conditions are given; and $\mathbf{g}\le\mathbf{0}$ represents the set of constraints (end, point and path) that must be satisfied for feasible operation.

Step 3. Test the process and control design from Step 2 for feasibility over the whole ranges of the uncertain parameters by solving the dynamic feasibility test problem:

$$\chi(\mathbf{d},\mathbf{y}) = \max_{\boldsymbol{\theta}} \min_{\mathbf{u_v}} \max_{l\in L,\ t\in[0,t_f]} g_l(\cdot) \quad (2)$$

subject to

$$\begin{aligned}\mathbf{f}\left(\dot{\mathbf{x}}(t),\mathbf{x}(t),\mathbf{z}(t),\mathbf{u_t}(t),\mathbf{u_v},\boldsymbol{\nu}(t),\boldsymbol{\theta},\mathbf{d},\mathbf{y}\right)&=\mathbf{0}\\ \mathbf{c}\left(\mathbf{x}(t),\mathbf{x_a}(t),\mathbf{u_t}(t),\mathbf{u_v},\boldsymbol{\nu}(t),\boldsymbol{\theta},\mathbf{d},\mathbf{y}\right)&=\mathbf{0}\end{aligned}$$

If $\chi(\mathbf{d},\mathbf{y})\le 0$, feasible operation can be ensured dynamically for all values of $\boldsymbol{\theta}$ within the given ranges. In this case, the algorithm terminates; otherwise, the solution of Equation 2 identifies a critical scenario that is added to the current set of scenarios before returning to Step 2.

Remarks

1. If the active set formulation of Grossmann and Floudas (1987) is used to solve (2), as proposed by Dimitriadis and Pistikopoulos (1995) and Mohideen et al. (1996), then the problem, like (1), corresponds to a MIDO problem.
2. The formulation (P) is an exact closed-loop, dynamic analogue of the steady-state problem of optimal design with fixed degree of flexibility (Pistikopoulos and Grossmann, 1988).
3. The solution strategy shown in Figure 1 and described above, is a closed-loop dynamic analogue of the flexible design algorithm of Grossmann and coworkers (see Biegler et al., 1997, chapter 21).
4. Different control design criteria can be used for example, decentralized PI-control, as discussed in Mohideen *et al.*, multivariable PI-control as discussed in Kookos and Perkins (2000) or Q-parameterization methods, as discussed in Swartz et al. (2000).
5. To date, the framework has been applied to single- and double-effect (heat integrated) distillation systems (Mohideen et al., 1996), to rigorously modeled double-effect systems (Bansal et al., 2000c), and an industrial two column system (Ross et al., 1999), but with control model simplifications, fixed discrete decisions and simplification in the treatment of uncertainty.
6. The integrated design and control problem requires the solution of MIDO problem. Until recently, there were no reliable methods for dealing with such problems. Therefore, it is still imperative to develop rigorous theory and efficient methods to accomplish this.
7. The proposed decomposition scheme, as shown in Figure 1, requires the repetitive solution of two

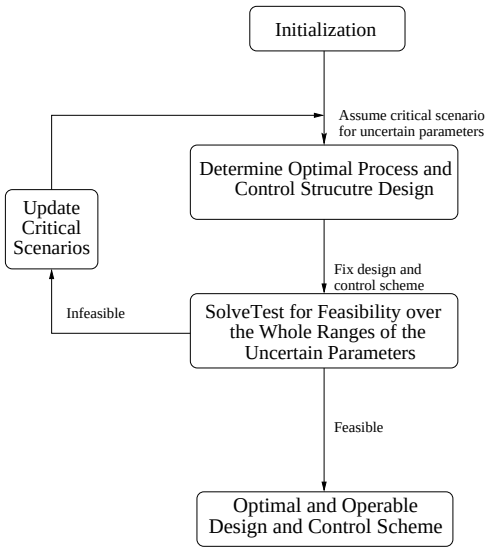


Figure 1: Decomposition algorithm of Mohideen et al. (1996).

MIDO problems in the design and the feasibility stage. It would be theoretically and computationally advantageous to avoid this iterative procedure, by solving in a single stage first, the feasibility problem and subsequently, the design problem, as it has been done for steady state systems by Bansal et al. (2000a). Currently, an endeavor is made towards adopting such a conceptual approach for the interactions of design and control under uncertainty and any progress in that area will be reported in the future.

In the next section an algorithm for solving mixed integer dynamic optimization problems is outlined. This algorithm is utilized in the simultaneous process and control design in the general case, where discrete decisions about the design and control structure are considered.

Mixed-Integer Dynamic Optimization (MIDO)

Optimal Control with the incorporation of binary variables, hence, Mixed Integer Dynamic Optimization (MIDO), plays a key role in methodologies that address the interactions of Design and Control (Mohideen et al., 1996; Schweiger and Floudas, 1997; Bahri et al., 1997; Kookos and Perkins, 2000). The simultaneous design and control framework described in the previous section involves the solution of MIDO problems in Steps 2 and 3. Moreover, MIDO is also encountered in several other modeling and optimization applications of chemical and process systems engineering. Avraam et al. (1999) used MIDO for addressing the issue of optimization on hybrid systems and recently, Barton et al. (2000) discuss the application of MIDO on the same area. Narraway and Perkins (1994) posed the Control Structure Selection problem in a Mixed Integer optimal control formulation. MIDO has also been employed for the design of batch / semi batch processes (Allgor and Barton, 1999; Barton et al., 1998; Sharif et al., 1998), dynamic optimization under uncertainty (Dimitriadis and Pistikopoulos, 1995; Samsatli et al., 1998) and for the reduction of kinetic mechanism models (Androulakis, 2000).

A number of algorithms have very recently started to appear in the open literature for solving MIDO problems. A common approach is to decompose directly the MIDO problem into a series of primal problems (upper bounds on the solution) and master problems (lower bounds on the solution). The primal problems correspond to continuous dynamic optimization problems where the values of the binary variables are fixed. These are commonly solved using control vector parameterization (CVP) techniques, where only the time-varying control variables are discretized. According to those techniques the differential system is initially integrated and then the gradients are calculated either via parameter perturbations or more accurately by integrating the sensitivity (Vassiliadis et al., 1994) or adjoint (Sargent and Sullivan, 1977) DAE system. The size of the sensitivity equations is proportional to the optimization parameters whereas the size of the adjoint system is approximately proportional to the number of constraints.

The MIDO algorithms that employ CVP for the primal problems mainly differ in how they construct the master problems, where the latter correspond to mixed-integer linear programs (MILPs) or non-linear programs (MINLPs) whose solutions give new sets of binary values for subsequent primal problems. Generalized Benders' Decomposition-based (GBD-based) approaches (Mohideen et al., 1997; Ross et al., 1998; Schweiger and Floudas, 1997), Outer Approximation-based (OA-based) approaches (Sharif et al., 1998), approaches based on "screening models" (Allgor and Barton, 1999) and "steady state models" (Kookos and Perkins, 2000) have been developed. These MIDO algorithms tend to depend on a particular type of method for integrating the DAE system in the primal problems and require the solution of a complex intermediate problem in order to construct the master problem. In the case of Allgor and Barton (1999) the method is case study-specific whereas the approach of Kookos and Perkins (2000) cannot in general be applied to intrinsic dynamic systems such as batch or semi-batch processes. In our approach, a variant of the Generalized Benders decomposition (Geoffrion, 1972; Floudas, 1995) method is employed for formulating the master problem. This is described next.

Generalized Benders Decomposition Approach for the Solution of MIDO Problems

Consider a general MIDO formulation:

$$\min_{u,d,y} \ \phi(\dot{x}(t_f), x(t_f), z(t_f), u(t_f), d, y, t_f) \tag{3}$$

subject to

$$\begin{aligned}
0 &= f(\dot{x}(t), x(t), z(t), u(t), d, y, t) \\
0 &= c(x(t), z(t), u(t), d, y, t) \\
0 &= r(x(t_0), \dot{x}(t_0), z(t_0), u(t_0), d, y, t_0) \\
0 &\geq g(\dot{x}(t), x(t), z(t), u(t), d, y, t) \\
0 &\geq q(\dot{x}(t_f), x(t_f), z(t_f), u(t_f), d, y, t_f) \\
t_o &\leq t \leq t_f
\end{aligned}$$

Here, $x \in \Re^{n_x}, z \in \Re^{n_z}$ are the vectors of the differential states and the algebraic variables respectively. The vectors $u \in \Re^{n_u}, d \in \Re^{n_d}$ represent the control and the time-invariant design variables, whereas $y \in \{0,1\}^{n_y}$ is the vector of the discrete binary variables. The functions f, c and r represent the differential equations, the algebraic equations and their initial conditions respectively. The objective function is denoted by ϕ and the path and

end point constraints by g and q respectively. The binary variables y participate only in a linear form in the objective function, the differential system and the constraints, since this is a necessary condition for applying GBD to a mixed integer optimization problem.

The *primal problem* is constructed by fixing the binaries to a specific value $y = y^k$. Then the problem given by Equation 3 becomes an optimal control problem that is solved with control vector parameterization. The control variables u are discretized to time-invariant parameters. From now on, the new total set of optimization variables will be denoted as v and includes the design and the parameterized controls $v = \{u_1, u_2, .., u_{N_u}, d\}$ $v \in \Re^{n_u \cdot N_u + n_d}$. The path constraints are converted to end-point constraints by introducing additional differential equations (e.g., Sargent and Sullivan, 1977) and state variables.

In GBD-based approaches the *master problem* is constructed using the dual information of the primal at the optimum solution. The *dual information* is embedded in the *Lagrange multipliers* μ of the constraints q and the *adjoint* time-dependent variables $\lambda(t), p(t)$ that are associated with the differential system of equations, i.e. f,c. Despite the fact that the Lagrange multipliers are calculated directly from the primal problem solution, the evaluation of the adjoint variables requires an extra integration of the so-called adjoint DAE system. This differential system has the form (Bryson and Ho, 1975; Vassiliadis, 1993):

$$\begin{aligned} -\frac{d\{[\frac{\partial f}{\partial \dot{x}}]^T \cdot \lambda(t)\}}{dt} &= -[\frac{\partial f}{\partial x}]^T \cdot \lambda(t) - [\frac{\partial c}{\partial x}]^T \cdot p(t) \\ 0 &= -[\frac{\partial f}{\partial z}]^T \cdot \lambda(t) - [\frac{\partial c}{\partial z}]^T \cdot p(t) \\ (\frac{\partial f}{\partial \dot{x}})^T_{t_f} \cdot \lambda(t_f) &= -\{(\frac{\partial \phi}{\partial x})^T_f + (\mu)^T \cdot (\frac{\partial q}{\partial x})^T_f \\ &\quad + [(\frac{\partial f}{\partial x})^T \, (\frac{\partial c}{\partial x})^T]_f \cdot \omega_f\} \end{aligned} \tag{4}$$

Equation 4 involves a backwards integration and can be computationally expensive. After the adjoint functions are calculated the master problem is constructed and has the following form:

$$\min_{y,\eta} \ \eta \tag{5}$$

subject to

$$\begin{aligned} \eta \geq \phi &+ (\mu^k)^T \cdot q \\ &+ (\omega_f^k)^T \cdot \begin{bmatrix} f \\ c \end{bmatrix}_f + (\omega_0^k)^T \begin{bmatrix} f \\ c \end{bmatrix}_o + (\rho^k)^T \cdot r \\ &+ \int_{t_o}^{t_f} \begin{bmatrix} \lambda^k(t) \\ p^k(t) \end{bmatrix}^T \cdot \begin{bmatrix} f \\ c \end{bmatrix}_t dt \\ & \qquad k = 1, K \quad k \in K \end{aligned}$$

ρ, ω_f and ω_0 are multipliers that are evaluated from the first order optimality conditions of the optimal control primal problem (Vassiliadis, 1993). The master problem is a relaxation of the equivalent to the MIDO, dual problem (Bazaraa et al., 1993) since the dual multipliers (Lagrange μ and adjoints λ, p) and the non-complicating continuous variables (x,z,v) remain fixed. The consecutive solutions of the master problem generate a series of supporting functions to the overall problem under several convexity assumptions (Floudas, 1995). If those assumptions do not hold the relax master problem might rule out parts of the feasible region where several local optima could lie decreasing the probability of detecting the global minimum. Nevertheless, the method ensures local optimality in the sense that when the integers are fixed the primal problem converges to a local solution in the space of continuous variables (primal ≡ valid upper bound).

The only variables that vary in the master problem are the binaries and the objective. The binaries participate in a linear form in the primal and master problems. As a result the master problem is an MILP and its solution apart from being lower bound to the MIDO problem also provides a new integer realization. If the lower bound evaluated at the master and the upper bound calculated in the primal cross then the solution is found and is equal to the upper bound, whereas if they do not cross the new integer set is augmented to the primal problem and the algorithm recommences.

The extra computationally demanding adjoint integration (Equation 4) limits the applicability of the method and renders the algorithm difficult to implement. Mohideen et al. (1997); Ross et al. (1998) employed a special numerical integration procedure for the primal dynamic optimization problem that brings some benefits in the adjoint calculation. However, these approaches restrict considerably the choice of primal solution techniques.

Recent developments in our group (Bansal et al., 2000b) show that the adjoint DAE system solution procedure can be eliminated by introducing an extra set of continuous optimization variables y_d, in the primal problem, that are fixed according to the equality constraint: $y_d - y^k = 0$. This gives rise to the following primal optimal control problem:

$$\min_{v,y_d} \ \phi(\dot{x}(t_f), x(t_f), z(t_f), v, y_d, t_f) \tag{6}$$

subject to

$$\begin{aligned} 0 &= f(\dot{x}(t), x(t), z(t), v, y_d, t) \\ 0 &= c(x(t), z(t), v, y_d, t) \\ 0 &= r(x(t_0), \dot{x}(t_0), z(t_0), v, y_d, t_0) \\ 0 &\geq q(\dot{x}(t_f), x(t_f), z(t_f), v, y_d, t_f) \\ 0 &= y_d - y^k \\ & \quad t_o \leq t \leq t_f \end{aligned}$$

The master problem is constructed in a similar mode:

$$\min_{y,\eta} \ \eta \tag{7}$$

subject to

$$\begin{aligned}\eta \geq\ & \phi(\dot{x}^k(t_f), x^k(t_f), z^k(t_f), v^k, y_d^k, t_f) \\ & + (\mu^k)^T \cdot q(\dot{x}^k(t_f), x^k(t_f), z^k(t_f), v^k, y_d^k, t_f) \\ & + (\omega_f^k)^T \cdot \begin{bmatrix} f \\ c \end{bmatrix}_f + (\omega_0^k)^T \begin{bmatrix} f \\ c \end{bmatrix}_o + (\rho^k)^T \cdot r \\ & + \int_{t_o}^{t_f} \begin{bmatrix} \lambda^k(t) \\ p^k(t) \end{bmatrix}^T \cdot \begin{bmatrix} f \\ c \end{bmatrix}_t dt \\ & + \Omega^k(y_d^k - y) \\ & \qquad\qquad k = 1, K \quad k \in K\end{aligned}$$

In the new primal formulation, Equation 6, the differential algebraic equations and the constraints do not include binaries any longer. Therefore at the master problem their associated terms are equal to zero due to the exact satisfaction of the DAE system and the complementarity conditions that apply to the constraints. By removing then those terms the master problem is simplified to the equation:

$$\min_{y,\eta} \ \eta \tag{8}$$

subject to

$$\begin{aligned}\eta \geq\ & \phi(\dot{x}^k(t_f), x^k(t_f), z^k(t_f), v^k, y_d^k, t_f) \\ & + \Omega^k(y_d^k - y) \\ & \qquad\qquad k = 1, K \quad k \in K\end{aligned}$$

In the modified equivalent master problem, Equation 8, all the terms are calculated at the solution of the primal problem and no adjoint calculations are required. Additionally, the formulation of the problem is considerably simplified compared to the original master problem structure of Equation 5.

However, the additional continuous optimization variables and the additional constraints may increase the computational effort for solving the primal problem while they may also introduce extra model complexity. Therefore, initially the primal is solved in its original form (Equation 3 with fixed binaries) and then a resolve session precedes the master problem where one additional optimization iteration is performed on the modified primal (Equation 6).

If the primal problem is infeasible the constraints are relaxed and a feasibility optimization problem is solved. The corresponding master problem is modified accordingly (Floudas, 1995; Mohideen, 1996). Integer cuts in the master problem formulation can also be included to exclude previous primal integer solutions.

The steps of the algorithm are briefly summarized below:

- Fix the values of the binary variables, $y = y^k$, and solve a standard dynamic optimization problem (Equation 3, kth primal problem). An upper bound, UB, on the solution to the MIDO problem is obtained from the minimum of all the primal solutions obtained so far.
- Re-solve the primal problem at the optimal solution (Equation 6) with additional constraints of the form $y_d - y^k = 0$, where y_d is a set of continuous search variables and y^k is the set of (complicating) binary variables. Convergence is achieved in one iteration. Obtain the Lagrange multipliers, Ω^k, corresponding to the new constraints.
- Construct the kth relaxed master problem from the kth primal solution, ϕ^k, and the Lagrange multipliers, Ω^k (Equation 8). This corresponds to the mixed-integer linear program (MILP) The solution of the master, η^k, gives a lower bound, LB, on the MIDO solution. If $UB - LB$ is less than a specified tolerance ϵ, or the master problem is infeasible, the algorithm terminates and the solution to the MIDO problem is given by UB. Otherwise, set $k = k + 1$, y^{k+1} equal to the integer solution of the master, and return to step 1.

The main advantage of the algorithm is that, even when the binary variables y participate within the DAE system (as they do for the distillation example presented later in this paper), the master problem does not require any direct dual information with respect to the DAE system and so no intermediate adjoint problem is required for its construction. The master problem, Equation 8, also has a very simple form compared to when adjoint variables are required (Mohideen et al., 1997; Schweiger and Floudas, 1997). Furthermore, the MIDO approach is independent of the type of method used for solving the dynamic optimization primal problems.

It should be noted, however, that since the algorithm is based GBD principles, shares the limitations of most decomposition methods. In particular, although a locally optimal solution is guaranteed when the integer variables are chosen as the complicating variables, the convexity conditions required for the algorithm to converge to the global optimum will not be satisfied by most process engineering problems. Investigation into the quality of solutions obtained from such MIDO algorithms is a current active research area.

Illustrative example

The application of the mentioned algorithm for solving Mixed Integer Dynamic Optimization problems is demonstrated through a process example that is taken from the MINOPT User's Guide (Schweiger et al., 1997). The case study examines a distillation column (Figure 2) that has a fixed number of trays and the objective is to determine the optimal feed location (discrete decision),

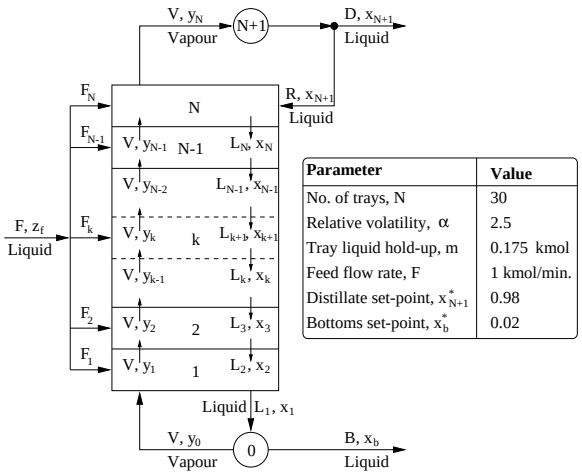


Figure 2: Illustrative example.

N^o Iterations	1	2	3	4
Primal				
Feed Loc.	20	26	24	25
V (kmol/min)	1.223	1.783	1.411	1.543
R (kmol/min)	0.684	1.243	0.871	1.002
ISE	0.1991	0.1827	0.183	*0.1817*
UB	0.1991	0.1827	0.1827	*0.1817*
Master				
Feed Loc.	26	24	25	23
LB	0.158	0.1813	0.1815	0.185
Equivalent Number of Simulations			298	

Table 1: Results on illustrative MIDO example.

vapor boil-up, reflux flow rate (continuous decisions) in order to minimize the integral square error (ISE) between the bottoms and distillate compositions and their set-points. Several model assumptions are made in the model, such as: (i) constant liquid hold ups, (ii) constant relative volatility, (iii) no pressure drops, (iv) negligible vapor hold-ups. The system is initially at steady state and the dynamics are caused by a stepwise variation in the feed composition. Additionally, two constraints are necessary to be satisfied at the end of the time horizon, these being on the top and bottoms compositions ($x_d \geq 0.98, x_b \leq 0.02$). The proposed algorithm was applied in this example (Bansal, 2000) and produced the same results as Schweiger et al. (1997) that appear in Table 1.

Two more algorithms for solving MIDO problems are discribed in the Appendix. The first one also employs the GBD principles and aims at reducing even further the computational requirements of the MIDO approach whereas the other is based on Outer approximation for dealing with the binary variables and is expected to converge in less iterations between primal and master problems. However, it should be noted that those algorithms are still under development from an implementation point of view, therefore, they have not been applied to the simultaneous process and control design problem.

Process Examples

Next, two examples are presented that illustrate the characteristics of the framework for the integration for process and control design. The first one includes process and control discrete decisions and for that reason it utilizes extensively the developed MIDO approach. In the second example discrete degrees of freedom are not considered.

Distillation System—(Benzene Toluene)

Here, an example is presented for demonstrating the features of the simultaneous process and control design framework and the utilization of mixed integer dynamic optimization within this framework. This example has been solved by Bansal et al. (2000b). The system under consideration, adapted from one presented by Viswanathan and Grossmann (1999), is shown in Figure 3. A mixture of benzene and toluene is to be separated into a top product with at least 98 mol% benzene and a bottom product with no more than 5 mol% toluene. The system is subject to uncertainty in the feed flow rate and the cooling water inlet temperature (where the latter can be described dynamically by a slow sinusoid representing diurnal, ambient variations), as well as a high-frequency sinusoidal disturbance in the feed composition. The objective is to design the distillation column and its required control scheme at minimum total annualized cost (comprising capital costs of the column, internals and exchangers, and operating costs of the steam and cooling water utilities), capable of feasible operation over the whole of a given time horizon, where feasibility is defined through the satisfaction of constraints such as product quality specifications; minimum column diameter requirement due to flooding; fractional entrainment limit; temperature driving forces in the reboiler and condenser; limit on the heat flux in the reboiler; limit on the cooling water outlet temperature; above atmospheric pressure operation for the column; limits on the liquid levels in the reflux drum and reboiler; and limits on the flow rates of steam and cooling water. Solution of the problem thus requires the determination of (i) the optimal process design, in terms of the number of trays and feed location (discrete decisions), and the column diameter, condenser and reboiler surface areas (continuous decisions); and (ii) the optimal control design, in terms of the pairings of manipulated and controlled variables (discrete decisions), and the tuning parameters for the given control structure (continuous decisions).

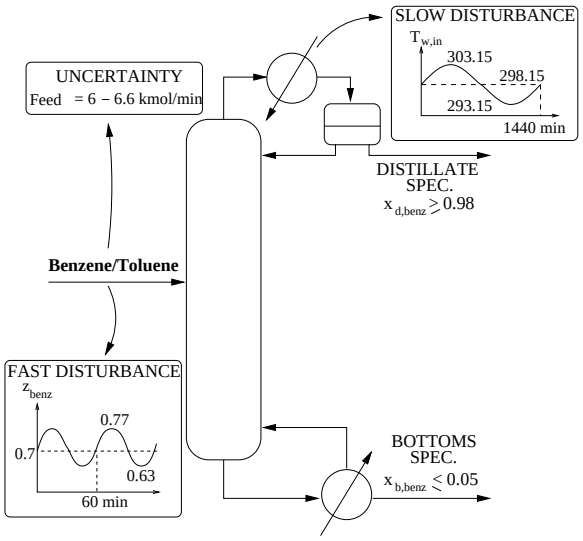


Figure 3: Distillation example.

Modeling Aspects. Due to the complexity and highly constrained nature of the problem described above, it is likely that a simplified dynamic model using "traditional" assumptions (such as constant molal overflow, relative volatility, liquid and vapor hold-ups) will be inadequate for realistically portraying the operability characteristics of the distillation system over time. A rigorous model is thus developed along similar lines to that used by Bansal et al. (2000c); however here, binary variables yf_k and yr_k are incorporated in order to account for the locations of the feed and reflux trays, respectively, where $yf_k = 1$ if all the feed enters tray k, and is zero otherwise, and $yr_k = 1$ if the reflux enters tray k, and is zero otherwise. This leads to a mixed-integer dynamic distillation model that is considerably more detailed than those that have already been reported (Mohideen et al., 1996; Schweiger and Floudas, 1997). The principal differential-algebraic equations (DAEs) for the trays are given below. A full list of nomenclature, values of the parameters, details of the DAEs for the reboiler, condenser and reflux drum, cost correlations for the objective function and inequality path constraints, can be found in Bansal (2000).

For $k = 1, ..., N$, where N is an upper bound on the number of trays required, and $i = 1, ..., NC$, where NC is the number of components:

Component molar balances:

$$\left(\sum_{k'=k}^{N} yr_{k'}\right) \cdot \frac{dM_{i,k}}{dt} = L_{k+1} \cdot x_{i,k+1} + V_{k-1} \cdot y_{i,k-1} + F_k \cdot z_{i,f} + R_k \cdot x_{i,d} - L_k \cdot x_{i,k} - V_k \cdot y_{i,k}.$$

Molar energy balances:

$$\left(\sum_{k'=k}^{N} yr_{k'}\right) \cdot \frac{dU_k}{dt} = L_{k+1} \cdot h^l_{k+1} + V_{k-1} \cdot h^v_{k-1} + F_k \cdot h_f + R_k \cdot h^l_d - L_k \cdot h^l_k - V_k \cdot h^v_k.$$

Component molar hold-ups:

$$M_{i,k} = M^l_k \cdot x_{i,k} + M^v_k \cdot y_{i,k}.$$

Molar energy hold-ups:

$$U_k = M^l_k \cdot h^l_k + M^v_k \cdot h^v_k - 0.1 \cdot P_k \cdot Vol_{tray}.$$

Volume constraints:

$$\frac{M^l_k}{\rho^l_k} + \frac{M^v_k}{\rho^v_k} = Vol_{tray}.$$

Definition of Murphree tray efficiencies:

$$y_{i,k} = y_{i,k-1} + Eff_{i,k} \cdot \left(y^*_{i,k} - y_{i,k-1}\right).$$

$$Eff_{i,k} = a_{i,k} + (1 - a_{i,k}) \cdot \sum_{k'=1}^{k} yr_{k'}.$$

Equilibrium vapor phase composition:

$$\Phi^v_{i,k} \cdot y^*_{i,k} = \Phi^l_{i,k} \cdot x_{i,k}.$$

Mole fractions normalization:

$$\sum_{i=1}^{NC} x_{i,k} = \sum_{i=1}^{NC} y_{i,k} = 1.$$

Liquid levels:

$$Level_k = \frac{M^l_k}{\rho^l_k \cdot A_{tray}}.$$

Liquid outlet flow rates:

$$L_k = 110.4 \cdot \rho^l_k \cdot Length_{weir} \cdot (Level_k - Height_{weir})^{1.5}.$$

Pressure driving force for vapor inlet:

$$P_{k-1} - P_k = 1e-5 \cdot \left(\sum_{k'=k}^{N} yr_{k'}\right) \cdot \left(\alpha \cdot vel^2_{k-1} \cdot \tilde{\rho}^v_{k-1} + \beta \cdot \tilde{\rho}^l_k \cdot g \cdot Level_k\right).$$

Vapor velocities:

$$vel_{k-1} = \frac{1}{60} \cdot \left(\frac{V_{k-1}}{\rho^v_{k-1} \cdot A_{holes}}\right).$$

Fractional entrainment for 80% flooding factor:

$$ent_k = 0.224e-02 + 2.377 \cdot exp\left(-9.394 \cdot FLV^{0.314}_k\right).$$

Sherwood flow parameter:

$$FLV_k = \frac{\tilde{L}_k}{\tilde{L}_k} \cdot \left(\frac{\tilde{\rho}_k^v}{\tilde{\rho}_k^l}\right)^{0.5}.$$

Flooding velocity:

$$vel_k^{flood} = 60 \cdot \left(\frac{\sigma_k^l}{20}\right)^{0.2} \cdot K1_k \cdot \left(\frac{\tilde{\rho}_k^l - \tilde{\rho}_k^v}{\tilde{\rho}_k^v}\right)^{0.5}.$$

Empirical coefficient:

$$K1_k = 0.0105 + 0.1496 \cdot Space^{0.755} \cdot exp\left(-1.463 FLV_k^{0.842}\right).$$

Minimum allowable column diameter and area:

$$D_{col,k}^{min} = \left(\frac{4 \cdot A_{col,k}^{min}}{\pi}\right)^{0.5}.$$

$$A_{net,k}^{min} = 0.9 \cdot A_{col,k}^{min}.$$

$$A_{net,k}^{min} = \frac{V_k}{\rho_k^v \cdot Floodfrac \cdot vel_k^{flood}}.$$

Feed and reflux flow rates to each tray:

$$F_k = F \cdot yf_k.$$
$$R_k = R \cdot yr_k.$$

Only one tray each receives feed and reflux; feed must enter below reflux:

$$\sum_{k=1}^{N} yf_k = \sum_{k=1}^{N} yr_k = 1.$$

$$yf_k - \sum_{k'=k}^{N} yr_{k'} \leq 0.$$

The complete distillation model constitutes a system of $[N(7NC+27)+15NC+56]$ DAEs in $[N(7NC+27)+15NC+64]$ variables (after specification of the feed and utilities' inlet conditions), of which $[N(NC+1)+3NC+5]$ are differential state variables. For the case study in this paper with $N = 30$ and $NC = 2$, there are 1316 DAEs in 1324 variables (101 states). The remaining eight variables consist of the three continuous design variables for optimization (column diameter, surface areas of the reboiler and the condenser), and the five manipulated variables (reflux, distillate, cooling water, steam and bottoms flow rates), whose values are determined by the tuning parameters of the control scheme used.

Application of the Framework.

Step 1. An initial set of two scenarios, $[6, 6.6]$, is chosen with weights $[0.75, 0.25]$. These correspond to the nominal and upper values, respectively, of the feed flow rate.

Step 2. Since the distillation column does not operate at very high purity, advanced control techniques are not required, and so multi-loop proportional-integral (PI) controllers are considered. For the purposes of this study, the control structure is considered to be a square system of measured and manipulated variables. The possible manipulated variables are: the reflux flow, R, the distillate flow, D, the cooling water flow, F_w, the steam flow, F_s and the bottoms flow, B. The set of the measured variables consists of: the distillate composition, x_d, the liquid level in the reflux drum, $Level_{drum}$, the pressure of the condenser, P_{cond} and the bottoms composition, x_b. The pairing between those variables, however, which is not known a priori, is treated as a discrete decision about the control design and is left to be determined through the optimization. One integer variable, y_k, is assigned to each possible control pairing and the modeling of the control structure selection is carrying out similarly to Narraway and Perkins (1994).

The MIDO problem (1) for this example consists of approximately 2700 DAEs and 216 inequality path constraints, with 85 binary search variables (thirty for the feed, thirty for the reflux location and twenty five for the control structure selection) and 18 continuous search variables (column diameter, surface areas of the reboiler and condenser, and gains, reset times and set-points for each of the five control loops). The problem was solved using the algorithm outlined in the section *Mixed Integer Dynamic Optimization*, with *gPROMS/gOPT* (PSE, 1999) used for solving the dynamic optimization primal problems and GAMS/CPLEX (Brooke et al., 1992) for the MILP master problems.

Step 3. In this example there are no time-invariant operating variables, and so the dynamic feasibility test, Equation 2, reduces to a conventional dynamic optimization problem with a single maximisation operator in the objective. Testing the design and control system resulting from Step 2 gives $\chi = 0$, indicating that there are no more critical scenarios, so the algorithm terminates.

Table 2 shows the iterations carried out between the Primal & the Master Problems. The economically optimal process and control design that gives feasible operation for all feed flow rates in the range 6-6.6 $kmol\ min^{-1}$ is summarized in Tables 3 and 4. Table 3 also compares the process design with the optimal steady-state, but dynamically inoperable, nominal and flexible designs. The latter was obtained through application of the analogous, steady-state algorithm to that described in §2 (Biegler et al., 1997). It can be seen that in order to accommodate feed flow rates above the nominal value of 6 $kmol\ min^{-1}$ requires more over-design when the dynamic behavior of the system is accounted for than when only steady-state effects are considered. This illustrates a weakness of considering design and control in a sequential manner.

Figures 4 and 5 show the dynamic simulations of the

Iteration Number	**1**	**2**	**3**	**4**	**5**
Primal Problem Solution					
Discrete decisions					
No. of Trays	25	24	23	22	24
Feed Tray	15	14	13	13	13
Control Scheme*	2	1	1	1	1
Process design					
D_{col} (m)	2.03	1.99	1.99	2.00	2.00
S_{reb} (m^2)	127.6	134.2	140.0	138.9	138.0
S_{cond} (m^2)	91.45	85.03	84.13	84.02	85.78
Controllers' gains					
$K_{\sqcup,1}$ $(x_{1,d})$	6.70	33.74	48.85	70.00	32.10
$K_{\sqcup,2}$ (L_d)	-105.0	-41.29	-18.64	-24.55	-25.39
$K_{3,3}$ $(P_c)^{\dagger}$	-28.00	-31.44	-29.24	-26.57	-36.16
$K_{\sqcup,4}$ $(x_{1,b})$	9.71	-2.22	-2.38	-3.37	-0.93
$K_{\sqcup,5}$ (L_0)	-1042	-600.0	-560.1	-580.5	-550.0
Reset times					
$\tau_{\sqcup,1}$	160.0	87.3	100.0	143.2	77.2
$\tau_{\sqcup,2}$	530.0	568.2	568.9	568.9	684.5
$\tau_{\sqcup,3}$	9935	3483	3615	5032	2809
$\tau_{\sqcup,4}$	2325	59.8	66.3	61.7	150.6
$\tau_{\sqcup,5}$	663.6	693.7	662.1	664.2	695.2
Set-points					
$set_{1,1}$	0.9883	0.9849	0.9843	0.9835	0.9853
$set_{1,2}$	0.5368	0.0668	0.0773	0.0746	0.0703
$set_{1,3}$	1.1944	1.2800	1.3022	1.3164	1.2694
$set_{1,4}$	0.0182	0.0223	0.0250	0.0293	0.0179
$set_{1,5}$	0.6002	0.8995	0.8994	0.8980	0.8995
Costs					
($\$100k\ yr^{-1}$)					
Capital	1.941	1.883	1.858	1.823	1.894
Operating(1)‡	6.367	6.268	6.287	6.334	6.269
Operating(2)§	7.220	7.122	7.136	7.194	7.097
Expected	8.521	8.364	8.357	8.372	8.370
UB	**8.521**	**8.364**	**8.357**	**8.357**	**8.357**
Master Problem Solution					
No. of Trays	24	23	22	24	22
Feed Tray	14	13	13	13	11
Control Scheme	1	1	1	1	1
LB	**8.242**	**8.282**	**8.341**	**8.355**	**8.357**
$UB - LB \leq 1e-4$	**No**	**No**	**No**	**No**	**Yes STOP**

*Control scheme 1: $R - x_{1,d}$, $D - Level_d$, $F_w - P_c$, $F_s - x_{1,b}$, $B - Level_0$.
*Control scheme 2: $R - x_{1,d}$, $D - Level_d$, $F_w - P_c$, $B - x_{1,b}$, $F_s - Level_0$.
†For $K_{3,3}$, the cooling water flow rate is scaled (0.01 F_w).
‡Nominal feed flow $F = 6kmolmin^{-1}$.
§Feed flow upper bound $F = 6.6kmolmin^{-1}$.

Table 2: Progress of iterations for the multi-period MIDO design and control problem.

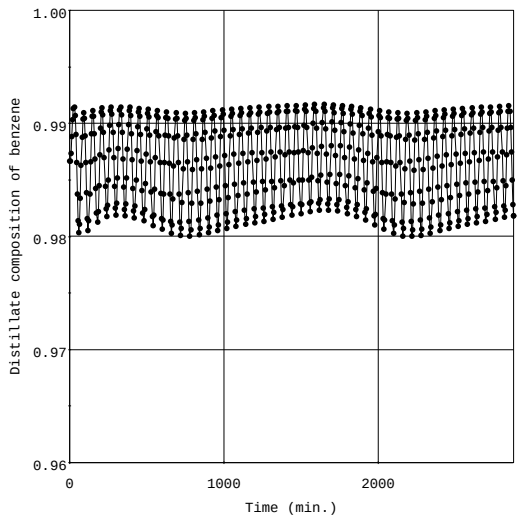


Figure 4: Controlled distillate composition at feed flow rate of 6.6 $kmol\ min^{-1}$.

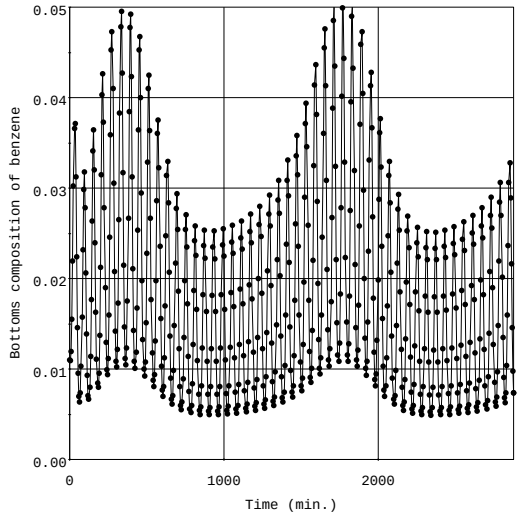


Figure 5: Controlled bottoms composition at feed flow rate of 6.6 $kmol\ min^{-1}$.

controlled compositions, that are given as part of the solution of the MIDO problem. Notice how one of the compositions, in this case the distillate, is tightly controlled relative to the other (in fact, the bottoms composition loop is open—see Table 4). This effect of controlling both compositions with one tight loop and one loose loop is due to the negative interaction of the two control loops, and is a common feature of distillation control reported in the literature (Kister, 1990).

Variable	SS nominal	SS flexible	Dynamic
No. of trays	23	23	26
Feed location	12	12	14
$D_{col}\ (m)$	1.82	1.91	1.99
$S_{reb}\ (m^2)$	113	116	134
$S_{cond}\ (m^2)$	83	83	88
Capital cost	169	175	195
Operating cost	591	607	641
Total ($\$\ k\ yr^{-1}$)	760	782	836

Table 3: Steady-state *vs.* dynamically operable design.

Loop	K	$\tau\ (min)$	Set-point
$R - x_d$	5.10	25.69	0.9867
$D - Level_{drum}$	-39.29	566.41	0.5187
$F_w - P_{cond}$	-44.81	7766.61	1.2183
$F_s - x_b$	0	250	0.0110
$B - Level_{reb}$	-501.93	663.96	0.8997

Table 4: Control design from simultaneous framework.

Reactive Distillation System—(Production of ethyl-acetate)

Here the problem that is considered is the production of ethyl acetate from the esterification of acetic acid and ethanol, as shown in Figure 6 (Georgiadis et al., 2000, 2001). The saturated liquid mixture is fed at a rate of 4885 mol/h in order to produce a top product with at least 0.52% ethyl acetate composition and a bottom product of no more than 0.26% ethyl acetate. Reaction takes place in all 13 trays of the column. The objective is then to design the column and the control scheme at minimum total cost, able to maintain feasible operation over a finite time horizon of interest (24 hours); subject to (i) high-frequency sinusoidal disturbances in the acetic acid inlet composition; (ii) "slow-moving" disturbance in the cooling water inlet temperature; (iii) product quality specifications; (iv) flooding, entrainment and minimum column diameter requirements; (v) thermodynamic feasibility constraints for the heat exchangers and (vi) operating pressure limits for the column.

The basis of the detailed model has been presented in our previous work (Schenk et al., 1999). The model includes details that are normally neglected, such as:

- Detailed flooding and entrainment calculations for each tray and 'subsequent' calculation of 'critical' points in the column and the minimum allowable column diameter.
- Equation for the pressure drop for each tray

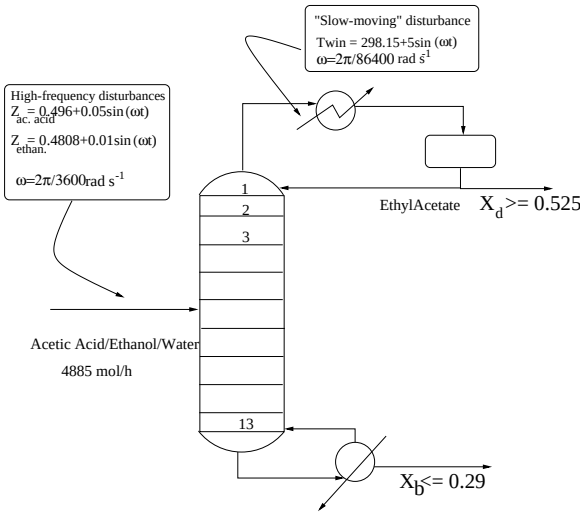


Figure 6: Schematic illustration of the reactive distillation system.

- Liquid hydraulics and liquid level on each tray and in the auxiliary units by using modified Francis weir formulae.
- The liquid-vapour equilibrium has been represented accurately using non-ideal models.

The model and its steady-state analogue have been implemented within gPROMS (PSE, 1999).

Sequential Design. A systematic sequential design and control approach is first carried out. The nominal and flexible steady-state designs are initially obtained via an optimization based approach. The nominal column design obtained is not feasible for the whole range of uncertain cooling water temperatures. The flexible design is obtained by applying a steady-state multiperiod approach that corresponds to equation (1). Three degrees of freedom (reflux ratio, steam flow rate and cooling water flow rate) can, in principle, be adjusted to offset the effects of the uncertainty. The following cases were considered (i) all three degrees of freedom allowed to vary ("best-case" design) and (ii) no degrees of freedom allowed to vary ("worst-case" design). The different optimal designs and resulting annual costs for the nominal and the two cases considered are shown in Table 5. Note that D refers to column diameter and S refers to the surface area of the heat exchange coil in the reboiler, Reb, or the condenser, Cond.

Both, the "best-case" and the "worst-case" flexible designs were dynamically tested in the presence of the sinusoidal feed composition disturbance and the "slow-moving" profile for the cooling water inlet temperature which ranges between suitable lower and upper bounds. As expected, there were a large number of constraints violations for both designs, and so they both require a

Design Variable	Nominal	Case 1	Case2
D (m)	6.09	6.09	6.12
$S_{Reb}(m^2)$	280	286	325
$S_{Cond}(m^2)$	417	458	498
Capital Cost ($\$\ myr^{-1}$)	0.45	0.46	0.47
Operating Cost ($\$\ myr^{-1}$)	3.95	3.99	4.35
Total Cost ($\$\ myr^{-1}$)	4.40	4.45	4.82

Table 5: Comparison of different designs for the reactive distillation example.

control scheme in order for feasibility to be maintained. The control structure considered here has been proven to be stable and exhibit satisfactory performance. The control loops are $(R - X_d)$, $(F_{stream} - X_b)$ and $(F_{water} - P_c)$ where R is the reflux ratio, F_{stream} is the steam flowrate and F_{water} the cooling water flowrate. Finally, x_d and x_b are the distillate and bottom compositions and the pressure in the condenser. The dynamic equations of the PI controllers are properly incorporated into the model. No set of controller's tuning parameters (gains, reset times, set-points and biases) could be found for either design that would enable all the system constraints to be satisfied over the entire time horizon. In particular the "best-case" design produced large constraint violations whereas the "worst-case" design exhibit the major operability bottleneck in the minimum column diameter requirements that are related to flooding. Then only the minimum column diameter was modified accordingly and a new steady-state flexible design was obtained.

The next step in the sequential design approach is to identify the optimal tuning of the controller's gains, reset times, set-points and biases keeping the modified "worse case" design fixed and optimizing the total annualized cost of the system over a fixed time-horizon. The operating variables obtained with that procedure along with the process design variables calculated in the previous step, i.e. column diameter, heat exchange areas comprise the results of the sequential design depicted in Table 6.

Simultaneous Process and Control Design. The sequential strategy outlined above has illustrated that interactions do exist between process design and process control. However, a more systematic approach for exploiting these interactions is to also include the process design variables as optimization variables whilst optimizing the controller settings. The framework presented in §2 is adopted. However, all the features of the general approach were not exploited fully, i.e. the design and control structural decisions remained fixed. The optimization variables are the design variables (column diameter and heat exchanger areas) and the gains, reset time, set points and biases of the controllers. The prob-

Quantity	Sequential	Simultaneous
D (m)	6.3	6.37
$S_{Reboiler}$ (m^2)	325	315
$S_{Condenser}$ (m^2)	498	425
PI1 Controller (top compositions control)		
Gain	747	181
Reset Time	1	0.59
Set-Point	0.54	0.527
PI2 Controller (bottom compositions control)		
Gain	1959	4573
Reset Time	12000	7294
Set Point	0.1	0.22
PI3 Controller (Condenser pressure control)		
Gain	-9500	-10400
Reset Time	0.603	1.32
Set-Point	1.023	1.023
Capital Cost ($ Million)	0.48	0.48
Operating Cost ($ Million)	4.37	4.17
Total Cost ($ Million)	4.85	4.63

Table 6: Results on the reactive distillation example.

lem is again solved as a large scale dynamic optimization problem with 12 optimization variables and a number of path and end-point inequality constraints describing the feasible operation of the process.

The optimal design, controller tuning parameters and associated costs are shown in Table 6 and compared with the results obtained using the sequential strategy with optimally tuned controllers. The simultaneous strategy has the same capital costs and lower operating costs leading to a 5% total annual cost savings ($220,000 per year). It is interesting to note that the simultaneous approach is able to give a fully operational system with an annual cost in between the costs of the "best-case" and "worst case" while the sequential approach gives a system which is more expensive than the "worst-case" optimal flexible design. This clearly demonstrates how a simultaneous approach can exploit the interactions between design and control to give process designs that are cheaper and more easily controlled than those found by even state-of the-art sequential approaches. It is also interesting to observe the different control settings adopted by the two approaches. As can be seen from Table 6 the simultaneous approach gives tighter bottom product composition that is closer to the constraint boundary; also tighter top product control and almost identical pressure control. The economic impact of this control action is the reduction in the operating costs.

Conclusions

This paper demonstrates the progress that has been made in simultaneous process and control design under uncertainty. A well-established decomposition framework for that purpose is reviewed. This framework in its general form requires the repetitive solution of mixed integer dynamic optimization problems. An algorithm for MIDO that has been recently developed by our group for that purpose is outlined.

Two process examples are considered that demonstrate the applicability and the benefits of the developed methods in the context of the integration of process design, process control and process operability. The first example considers the separation of a binary mixture and utilizes the novel MIDO algorithm for treating control, design and discrete structural decisions. The second example studies the process and control design of a reactive distillation system and does not take into account discrete decisions.

Future work will mainly focus on improving the convergence properties of the current MIDO methodologies and developing a more efficient single-stage approach for dealing with the inevitable presence of the uncertainty as opposed to the current decomposition (two-stage) approach. Also other control technology will be considered, such as multivariable PI controllers (Kookos and Perkins, 2000), as opposed to decentralized PI controllers that are almost exclusively used so far and in addition, an effort will be made to address synthesis issues that have not been fully considered within a mixed integer dynamic optimization framework.

Acknowledgments

The ideas described are based on the collaboration over a number of years of the corresponding author with Professor John Perkins, Dr. Jezri Mohideen, Dr. Vik Bansal, Dr. Rod Ross, Dr. Myrian Schenk, Dr. Michael Georgiadis and Dr. Jan van Schijndel.

Financial support from EU (HYSEP project), DETR/ETSU (The Process Design Toolbox Project), Shell, APCI and the Centre for Process System Engineering (IRC grant and IRC Industrial Consortium) is also acknowledged.

References

Allgor, R. J. and P. I. Barton, "Mixed—integer dynamic optimization I: problem formulation," *Comput. Chem. Eng.*, **23**, 567–584 (1999).

Androulakis, I. P., "Kinetic Mechanism Reduction Based on an Integer Programming Approach.," *AIChE J.*, **46**(2), 361–371 (2000).

Avraam, M. P., N. Shah, and C. C. Pantelides, "A Decomposition Algorithm for the Optimisation of Hybrid Dynamic Processes," *Comput. Chem. Eng.*, pages Suppl. S451–S454 (1999).

Bahri, P. A., J. A. Bandoni, and J. A. Romagnoli, "Integrated Flexibility and Controllability Analysis in Design of Chemical

Processes," *AIChE J.*, **43**(4), 997–1015 (1997).

Bansal, V., R. Ross, J. D. Perkins, and E. N. Pistikopoulos, "The interactions of design and control: double-effect distillation systems," *J. Proc. Cont.*, **10**, 219–227 (2000c).

Bansal, V., J. D. Perkins, and E. N. Pistikopoulos, "Parametric Programming Algorithms for the Flexibility Analysis and Design of Linear Process Systems," *AIChE J.*, **46**(2), 335–354 (2000a).

Bansal, V., J. D. Perkins, E. N. Pistikopoulos, R. Ross, and J. M. G. van Schijndel, "Simultaneous Design and Control Optimization under uncertainty," *Comput. Chem. Eng.*, **24**(2-7), 261–266 (2000b).

Bansal, V., *Analysis, Design and Control Optimization of Process Systems under Uncertainty*, PhD thesis, Imperial College, University of London (2000).

Barton, P. I., R. J. Allgor, W. F. Feehery, and S. Galan, "Dynamic Optimization in a Discontinuous World," *Ind. Eng. Chem. Res.*, **37**(3), 966–981 (1998).

Barton, P. I., J. R. Banga, and S. Galan, "Optimization of hybrid discrete/continuous systems," *Comput. Chem. Eng.*, **24**(9-10), 2171–2182 (2000).

Bazaraa, M. S., H. D. Sherali, and C. M. Shetty, *Nonlinear Programming. Theory and Applications*. Wiley, New York (1993).

Biegler, L. T., I. E. Grossmann, and A. W. Westerberg, *Systematic Methods of Chemical Process Design*. Prentice Hall (1997).

Brooke, A., D. Kendrick, and A. Meeraus, *GAMS Release 2.25: A User's Guide*. San Francisco: The Scientific Press (1992).

Bryson, A. E. and Y. Ho, *Applied Optimal Control*. Taylor & Francis, New York (1975).

Dimitriadis, V. D. and E. N. Pistikopoulos, "Flexibility analysis of dynamic systems," *Ind. Eng. Chem. Res.*, **34**, 4451–4462 (1995).

Duran, M. A. and I. E. Grossmann, "An outer approximation algorithm for a class of mixed integer nonlinear programs," *Math. Prog.*, **36**(3), 307–339 (1986).

Floudas, C. A., *Nonlinear and Mixed-Integer Optimization*. New York: Oxford University Press (1995).

Geoffrion, A. M., "Generalized Benders decomposition," *J. Optim. Theo. Appl.*, **10**, 237–260 (1972).

Georgiadis, M. C., R. Gani, and E. N. Pistikopoulos, A Synthesis Method for Hybrid Separation, Technical report, European Commission, Research Contract No: JOE3-CT97-0085 (2000).

Georgiadis, M. C., M. Schenk, R. Gani, and E. N. Pistikopoulos, "The Interactions of Design, Control and Operability in Reactive Distillation Systems," *Proc. of ESCAPE-11*, pages 997–1002 (2001). Denmark May 2001.

Grossmann, I. E. and C. A. Floudas, "Active constraint strategy for flexibility analysis in chemical processes," *Comput. Chem. Eng.*, **11**, 675–693 (1987).

Kister, H. Z., *Distillation Operation*. New York: McGraw-Hill (1990).

Kookos, I. K. and J. D. Perkins, An algorithm for simultaneous process design and control (2000). Submitted for publication to the Ind. Eng. Chem. Res.

Mohideen, M. J., J. D. Perkins, and E. N. Pistikopoulos, "Optimal design of dynamic systems under uncertainty," *AIChE J.*, **42**, 2251–2272 (1996).

Mohideen, M. J., J. D. Perkins, and E. N. Pistikopoulos, "Towards an efficient numerical procedure for mixed integer optimal control," *Comput. Chem. Eng.*, **21**, S457–S462 (1997).

Mohideen, J. M., *Optimal design of Dynamic Systems under Uncertainty*, PhD Dissertation, Imperial College of Science, Technology and Medicine, London (1996).

Narraway, L. and J. D. Perkins, "Selection of Process control Structure based on economics," *Comput. Chem. Eng.*, **18**, Suppl. S511–S515 (1994).

Pistikopoulos, E. N. and I. E. Grossmann, "Optimal Retrofit Design for Improving Process Flexibility in Linear Systems," *Comput. Chem. Eng.*, **12**(7), 719–731 (1988).

PSE, *gPROMS Advanced User's Guide*. Process Systems Enterprise Ltd., London (1999).

Ross, R., V. Bansal, J. D. Perkins, and E. N. Pistikopoulos, A Mixed-Integer Dynamic Optimization Approach to Simultaneous Design and Control, In *AIChE Annual Meeting*, Paper 220e, Miami Beach Florida, US. American Institute of Chemical Engineering (1998).

Ross, R., V. Bansal, J. D. Perkins, E. N. Pistikopoulos, G. L. M. Koot, and J. M. G. Schijndel, "Optimal design and control of an industrial distillation system," *Comput. Chem. Eng.*, **23**, S875–S878 (1999).

Sakizlis, V., V. Bansal, R. Ross, J. D. Perkins, and E. N. Pistikopoulos, "New Algorithms for Mixed-Integer Dynamic Optimization," *Proc. of ESCAPE-11*, pages 273–278 (2001). Denmark May 2001.

Samsatli, N. J., L. G. Papageorgiou, and N. Shah, "Robustness Metrics for Dynamic Optimization Models under Parameter Uncertainty," *AIChE J.*, **44**(9), 1993–2006 (1998).

Sargent, R. W. H. and G. R. Sullivan, Development of an Efficient Optimal Control Package, In *Proc. 8th IFIP conf. optimization techniques*, Wurzuburg. IFIP (1977).

Schenk, M., R. Gani, D. Bogle, and E. N. Pistikopoulos, "A hybrid modelling approach for separation systems involving distillation," *Chem. Eng. Res. Des.*, **77**(A6), 519–534 (1999).

Schweiger, C. A. and C. A. Floudas, Interaction of design and control: optimization with dynamic models, In Hager, W. W. and P. M. Pardalos, editors, *Optimal Control: Theory, Algorithms and Applications*, pages 388–435. Kluwer (1997).

Schweiger, C. A., A. Rojnuckarin, and C. A. Floudas, *MINOPT: User's Guide*. Princeton University, software version 2.0 edition (1997).

Sharif, M., N. Shah, and C. C. Pantelides, "On the Design of Multicomponent Batch Distillation Columns," *Comput. Chem. Eng.*, **22**, Suppl. S69–S76 (1998).

Swartz, C. L. E., J. D. Perkins, and E. N. Pistikopoulos, Incorporation of Controllability in Process Design Through Controller Parameterization, In *Process Control and Instrumentation 2000*, pages 49–54. University of Strathclyde (2000).

Van Schijndel, J. M. G. and E. N. Pistikopoulos, Towards the Integration of Process Design, Process Control & Process Operability—Current Status & Future Trends, In Malone, M. F., J. A. Trainham, and B. Carnahan, editors, *Foundations of Computer-Aided Process Design*, volume 96, pages 99–112, Breckenridge, Colorado. American Institute of Chemical Engineers (2000).

Vassiliadis, V. S., R. W. H. Sargent, and C. C. Pantelides, "Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints.," *Ind. Eng. Chem. Res.*, **33**(9), 2111–2122 (1994).

Vassiliadis, V., *Computational Solution of Dynamic Optimization Problems with General Differential—Algebraic Constraints*, PhD Dissertation, Imperial College of Science, Technology and Medicine, London (1993).

Viswanathan, J. and I. E. Grossmann, "A Compined penalty function and Outer Approximation method for MINLP Optimization," *Computers and Chemical Engineering*, **14**(7), 769–780 (1999).

Appendix

Adjoint Based Algorithm on Mixed Integer Dynamic Optimization

From the discussion on mixed integer dynamic optimization the conclusion drawn is that an efficient adaptation of Generalized Benders Decomposition in MIDO requires the simplification of the master problem construction and the reduction or possible elimination of the dual information calculations. An algorithm for achieving that is presented in the main document and here an alternative approach is outlined aiming at reducing the computational requirements of the linked subproblems. The details of this approach can be found in Sakizlis et al. (2001).

According to the original GBD-approach in MIDO (Schweiger and Floudas, 1997; Mohideen et al., 1997) after the solution of the primal optimal control problem an additional subproblem has to be solved that involves the backwards integration of the so-called adjoint DAE system, Equation 4. Since this can be computationally expensive, a method is developed for eliminating the extra calculations by adapting an adjoint-based approach for evaluating the gradients of the constraints and the objective function of the primal optimal control problem. This provides at the optimal solution of the primal a set of vectors of adjoint variables that are associated with the constraints and the objective function, denoted as $[\lambda_\phi(t)\ p_\phi(t)]$, $[\lambda_q(t)\ p_q(t)]$ respectively. Those adjoint functions are given by the same linear DAE system as the adjoint functions that are necessary for the master problem construction. However, $[\lambda_\phi(t)\ \ p_\phi(t)]$, $[\lambda_q(t)\ p_q(t)]$ are given by different final conditions as opposed to $[\lambda(t), p(t)]$. Their final conditions are:

$$[\frac{\partial f}{\partial \dot{x}}]^T_{t_f} \cdot \lambda_\phi(t_f) = -[(\frac{\partial \phi}{\partial x})^T_f + [(\frac{\partial f}{\partial x})^T\ (\frac{\partial c}{\partial x})^T]_f \cdot (\omega_f)_\phi]$$
$$[\frac{\partial f}{\partial \dot{x}}]^T_{t_f} \cdot \lambda_q(t_f) = -[(\frac{\partial q}{\partial x})^T_f + [(\frac{\partial f}{\partial x})^T\ (\frac{\partial c}{\partial x})^T]_f \cdot (\omega_f)_q]$$

The linear properties of the adjoint differential system and its boundary conditions, enable the evaluation of the adjoint variables required for the master problem $[\lambda, p]$ as a function of $[\lambda_\phi\ p_\phi]$, $[\lambda_q\ p_q]$ from the equation:

$$\begin{bmatrix} \lambda(t) \\ p(t) \end{bmatrix} = \begin{bmatrix} \lambda_\phi(t) \\ p_\phi(t) \end{bmatrix} + \mu^T \cdot \begin{bmatrix} \lambda_q(t) \\ p_q(t) \end{bmatrix} \tag{9}$$

Equation 9 can be proved using the transition matrix theory. Using that approach, the rigorous adjoint integration for the master problem derivation is not required any more after the primal has terminated and the derivation of the dual information is reduced exclusively to Equation 9.

Even if the easily obtained time-dependent functions $\lambda(t), p(t)$ are supplied to the master problem many calculations are still required due to the presence of the time integral in Equation 5 and the usually complicated non-linear functions involved in the DAE system. In order to simplify further the master problem equations f,c are decomposed in terms of the binary and continuous variables:

$$\begin{aligned} f &= f'(\dot{x}, x, z, v, t) + f^y(\dot{x}, x, z, v, t) \cdot y \\ c &= c'(x, z, v, t) + c^y(x, z, v, t) \cdot y \\ r &= r'(\dot{x}_o, x_o, z_o, v, t_o) + r^y(\dot{x}_o, x_o, z_o, v, t_o) \cdot y \end{aligned} \tag{10}$$

f^y, c^y, r^y are a matrices of dimensions $n_x \times n_y$, $n_z \times n_y$, $n_x \times n_y$ respectively. This separation is allowed, since the binaries participate in the DAE in a linear form (variant-2 of GBD). At the primal solution though:

$$f = 0 \ \ c = 0 \ \ r = 0 \tag{11}$$

So, $(f')^k$ can be written as:

$$(f')^k = -(f^y)^k \cdot y^k \tag{12}$$

Similarly for c', c^y, r', r^y. Finally we have:

$$\begin{aligned} f &= (f^y)^k \cdot (y - y^k) \\ c &= (c^y)^k \cdot (y - y^k) \\ r &= (r^y)^k \cdot (y - y^k) \end{aligned} \tag{13}$$

Once Equation 13 is substituted in Equation 5 the modified master problem becomes:

$$\min_{y,\eta}\ \eta \tag{14}$$

subject to

$$\begin{aligned} \eta \geq \phi + (\mu^k)^T \cdot q + \{(\omega_f^k)^T \cdot \begin{bmatrix} (f^y)^k \\ (c^y)^k \end{bmatrix}_f \\ + (\omega_0^k)^T \begin{bmatrix} (f^y)^k \\ (c^y)^k \end{bmatrix}_o + (\rho^k)^T \cdot (r^y)^k \\ + \int_{t_o}^{t_f} [(\lambda^k)^T\ (p^k)^T] \cdot \begin{bmatrix} (f^y)^k \\ (c^y)^k \end{bmatrix}_t dt\} \cdot (y - y^k) \end{aligned}$$

In this manner, the size of the master problem formulation is significantly reduced. The multiplier of the binary terms $y - y^k$ corresponds to a time dependent vector of size equal to the dimensions of the binaries n_y. This vector does not contain any integer terms and hence, it remains fixed throughout the master problem solution. In order to evaluate the components of that vector it suffices to transform the contained integrals to an ODE system of size n_y introducing differential states of zero initial conditions. Then by solving numerically the ODE system the construction of the master problem is completed. Alternatively, if the formulation of Equation 5 was retained, every equation that contains a binary term would have to be integrated. So the size of that corresponding ODE system would be of order of magnitude $O(n_x + n_z) >> n_y$.

From an implementation point of view, Equation 14 is easier to construct than Equation 5 because the matrices f^y, c^y, r^y are the Jacobians of the DAE with respect to the binaries and are simply generated (numerically or analytically) using well-established commercial codes or algorithms.

The master optimization problem consists of the binary variables and the continuous objective. The other variables are fixed. Since the optimization variables participate in a linear form, the problem is an MILP and it is solved with the current well-established methodologies (e.g. branch and bound algorithm).

The steps of the algorithm are summarized as follows:

1. Fix the values of the binary variables, $y = y^k$, and solve a standard dynamic optimization problem (k^{th} primal problem). An upper bound, UB, on the solution to the MIDO problem is obtained from the minimum of all the primal solutions obtained so far.

2. At the solution of the primal problem, using Equation 9, obtain the adjoint functions $\lambda(t), p(t)$.

3. Use the problem variables $x(t), z(t), v$, the adjoint functions $\lambda(t), p(t)$ and the Lagrange multipliers of the constraints μ to construct the k^{th} relaxed master problem, Equation 14, from the k^{th} primal solution. The Master problem corresponds to a Mixed -integer linear program (MILP), that its solution provides the lower bound, LB, on the MIDO solution. If UB-LB is less than a specified tolerance ϵ, or the master problem is infeasible, the algorithm terminates and the solution to the MIDO problem is given by UB. Otherwise, set $k = k + 1$ and y^{k+1} equal to the integer solution of the master problem and return to step 1.

This alternative algorithm eliminates completely the adjoint evaluation and does not require any resolve session after the primal problem is solved. It also manages to simplify considerably the master problem construction. Moreover, descpite the fact that it is restricted to using only an adjoint based gradient evaluation procedure for the primal optimal control problem it is not confined to a particular type of DAE integrator as in Mohideen et al. (1997); Ross et al. (1998).

An Outer Approximation Based Method for Mixed Integer Dynamic Optimization

The algorithms presented on Mixed Integer dynamic optimization are based on Generalized Benders decomposition for obtaining the lower bound to the problem. Therefore, the results that they will produce will be equivalent. Despite the benefits of GBD, that among others are the simple modeling of the discrete decisions and the straight forward formulation of the master problem, the lower bounds that are generated are relatively relaxed since in every master problem, only a single constraint is added to the iterative procedure. This can increase the number of the subsequent problem solutions deteriorating the convergence properties. A desired reduction in the primal-master iterations can be achieved by adapting another decomposition approach for constructing the master problem based on Outer Approximation (Duran and Grossmann, 1986). An outline of the concepts that enable the application of OA to MIDO is presented here.

The application of OA to an MINLP problem requires the participation of the integer variables in the equalities, inequalities and objective in a linear and separable form. The translation of this condition to MIDO makes imperative the removal of the binary variables from the DAE system. The reason being that even if the binaries participate linearly in the DAE their implicit contribution to the objective and the constraints is non-linear due to the non-linearities introduced by the dynamics. The removal of the binaries from the dynamic system is done in a way similar to the one presented in the main document. Namely, an extra set of continuous search variables y_d is introduced in the primal problem, that are fixed according to the double inequality constraint: $y_d - y^k \geq 0$,$y_d - y^k \leq 0$. This gives rise to the following primal optimal control problem:

$$\min_{v,y_d} \ \phi(\dot{x}(t_f), x(t_f), z(t_f), v, y_d, t_f) \tag{15}$$

subject to

$$\begin{aligned}
0 &= f(\dot{x}(t), x(t), z(t), v, y_d, t) \\
0 &= c(x(t), z(t), v, y_d, t) \\
0 &= r(x(t_0), \dot{x}(t_0), z(t_0), v, y_d, t_0) \\
0 &\geq q(\dot{x}(t_f), x(t_f), z(t_f), v, y_d, t_f) \\
0 &\leq y_d - y^k \\
0 &\geq y_d - y^k \\
t_o &\leq t \leq t_f
\end{aligned}$$

The master problem that aims to generate a lower bound, is constructed by linearizing the constraints and the objective around the primal optimal point only in the space of the search variables (v, y_d). The resultant master problem is:

$$\min_{y,y_d,v,\eta} \ \eta \tag{16}$$

subject to

$$\begin{aligned}
\eta &\geq \phi(\dot{x}^k(t_f), x^k(t_f), z^k(t_f), v^k, y_d^k, t_f) \\
&\quad + \frac{d\phi}{dv} \cdot (v - v^k) + \frac{d\phi}{dy_d} \cdot (y_d - y_d^k) \\
0 &\geq q(\dot{x}^k(t_f), x^k(t_f), z^k(t_f), v^k, y_d^k, t_f) \\
&\quad + \frac{dq}{dv} \cdot (v - v^k) + \frac{dq}{dy_d} \cdot (y_d - y_d^k) \\
0 &\leq y_d - y \\
0 &\geq y_d - y
\end{aligned}$$

In this master formulation in every k^{th} iteration a set of inequality constraints equal to the number of the original constraints plus one is added. Therefore, if the number of constraints is relatively high the lower bounds generated by OA will be tighter than the ones produced from GBD, hence the algorithm convergence will be achieved in less iterations.

A summary of the steps of the OA algorithm are presented below:

1. Fix the values of the binary variables, $y = y^k$, and solve a standard dynamic optimization problem (k^{th} primal problem). An upper bound, UB, on the solution to the MIDO problem is obtained from the minimum of all the primal solutions obtained so far.
2. At the solution of the primal problem add the extra set of continuous search variables y_d and the inequality constraints: $y_d - y^k \geq 0$,$y_d - y^k \leq 0$. Resolve the primal problem at the optimal solution. Convergence is achieved in one iteration and the gradients: $\frac{d\phi}{dy_d}, \frac{d\phi}{dy_d}, \frac{dq}{dv}$ and $\frac{dq}{dy_d}$ are evaluated via numerical integration of the sensitivity (Vassiliadis et al., 1994) or the adjoint DAE system (Sargent and Sullivan, 1977).
3. Use the problem continuous optimization variables and the corresponding gradients for formulating the master problem, Equation 16. The Master problem corresponds to a Mixed -integer linear program (MILP), that its solution provides the lower bound, LB, on the MIDO solution. If UB-LB is less than a specified tolerance ϵ, or the master problem is infeasible, the algorithm terminates and the solution to the MIDO problem is given by UB. Otherwise, set $k = k + 1$ and y^{k+1} equal to the integer solution of the master problem and return to step 1.

This presented algorithm for MIDO has the potential of providing tighter (higher) lower bounds to the overall problem thus accelerating the MIDO solution convergence.

Optimal Operation and Control of Simulated Moving Bed Chromatography: A Model-Based Approach

Karsten-U. Klatt*, Guido Dünnebier, Felix Hanisch and Sebastian Engell
Department of Chemical Engineering
University of Dortmund
D-44221 Dortmund, Germany

Abstract

Chromatographic separations are an expanding technology for the separation of Life Science products, such as pharmaceuticals, food, and fine chemicals. The simulated moving bed (SMB) process as a continuous chromatographic separation is an interesting alternative to conventional batch chromatography, and gained more and more impact recently. The SMB process is realized by connecting several single chromatographic columns in series. A countercurrent movement of the bed is approximated by a cyclic switching of the inlet and outlet ports in the direction of the fluid stream. Because of its complex dynamics, the optimal operation and automatic control of SMB processes is a challenging task. This contribution presents an integrated approach to the optimal operation and automatic control of SMB chromatographic separation processes. It is based on computationally efficient simulation models and combines techniques from mathematical optimization, parameter estimation and control theory. The overall concept and the realization of the elements are explained, and the efficiency of the proposed approach is shown in a simulation study for the separation of fructose and glucose on an 8-column SMB plant.

Keywords

Chromatographic separation, Simulated moving bed, Dynamic models, Optimization, Model-based control

Introduction

The chemical process industry is currently undergoing a substantial restructuring: the classical bulk business is more and more substituted by Life Science products with higher profit margins. In this area, particularly in the development and production of pharmaceuticals, it is of the utmost importance to be ahead of the competitors in the race to the market. This requires a detailed and integrated process design already in the product development phase. In this context, product separation and purification is the critical element in many cases.

Chromatographic processes provide a versatile tool for the separation of substances which have different adsorption affinities. They are especially suitable for temperature-sensitive compounds and substances with similar molecular structure and physico-chemical properties. Chromatography is well established in the field of the chemical analysis, but in recent years it gained more and more importance on the preparative scale as a highly efficient, highly selective separation process. Due to their origin and the close relation to the instruments from chemical analysis (i.e. HPLC and gas chromatographic analyzers), chromatographic separation processes are mainly operated in the classical batch elution mode. To improve the economic viability, a continuous countercurrent operation is often desirable, but the real countercurrent of solids—such as the adsorbent in chromatographic processes—leads to serious operating problems. Therefore, the simulated moving bed (SMB) process is an interesting alternative since it provides the advantages of a continuous countercurrent unit operation while avoiding the technical problems of a true moving bed.

The SMB process was first realized in the family of SORBEX processes by UOP (Broughton and Gerhold, 1961) and is increasingly used in a wide range of industries. Currently, the main applications of continuous chromatographic separations can be divided into two groups: the large-scale industrial production of relatively cheap specialty products, like xylene production or sugar separation, and the separation of high-value products in small amounts, which very often exhibit separation factors near unity (e.g. enantiomer separations in the pharmaceutical industry). The separation costs in both cases are very high in relation to the overall process costs and easily dominate those. An optimal design and operation might therefore be the only possibility to exploit the economic potential of the process and to make its application feasible.

In practice, the SMB process is nowadays mainly realized by connecting several single chromatographic columns in series. The countercurrent movement is then approximated by a cyclic switching of the feed stream and the inlet and outlet ports in the direction of the fluid flow. Thus, the process shows mixed continuous and discrete dynamics with complex interactions of the corresponding process parameters. If the SMB process is operated close to its economic optimum, high sensitivities to disturbances and changes in the operating parameters result. Furthermore, concentration measurements are expensive and can only be installed at the outlet of the separation columns. Therefore, the control of SMB chromatographic separation processes in order to ensure a safe and economical operation while guaranteeing the product specifications at any time is a challenging task.

*K.-U. Klatt and G. Dünnebier are presently with Bayer AG, D-51368 Leverkusen, Germany.

Currently, most processes are operated at some distance from the optimum to avoid off-spec production and to ensure sufficient robustness margins. In order to exploit the full economic potential of this increasingly applied technology, model-based optimization and automatic control of SMB processes are required.

Several publications on both process optimization and feedback control of SMB processes can be found in the literature but they predominantly do not treat optimization and control in an integrated manner. The purpose of this contribution is to propose such an integrated approach based on a rigorous dynamic process model. The overall concept and the realization of the elements are explained in the remainder of this paper. In each section, we review the state of the art and refer to related work of other authors. We start from a short description of chromatographic separations and SMB chromatography in particular, and then explain the generation and implementation of sufficiently accurate and computationally efficient process models, which are the essential prerequisite for model-based optimization and control. We proceed with the issue of determining the optimal operating regime of the process, followed by the description of the overall control concept and its components. The feasibility and the capabilities of the proposed approach are then demonstrated on an application example, the separation of fructose and glucose on an 8-column SMB laboratory plant. We finalize with some conclusions, highlighting unresolved issues and future research directions.

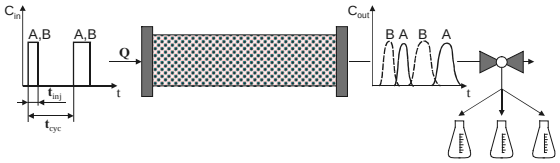


Figure 1: Batch elution chromatography.

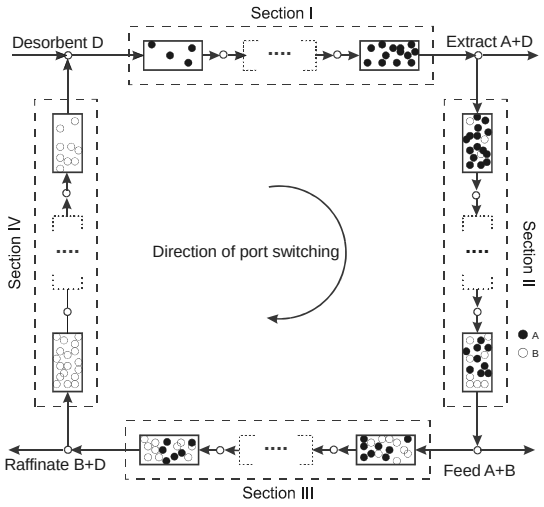


Figure 2: Simulated moving bed chromatography.

Process Description

Chromatography is a separation technique which is based on the preferential adsorption of one component. In adsorption, the solutes are transferred from a liquid or gas mixture to the surface of a solid adsorbent, where they are held by intramolecular forces. Desorption is the reverse process whereby the solute, called adsorbate, is removed from the surface of an adsorbent. By the use of a suitable stationary phase, components that are difficult to separate by other methods can be obtained in very high purities. In comparison to other thermal separation methods, e.g. distillation, less energy is consumed. Chromatography is particularly useful for the separation of temperature-sensitive components because it can often be performed at room temperature (Hashimoto et al., 1993; Adachi, 1994).

The classical implementations of chromatographic separations are batch processes in elution mode (see Figure 1). A feed pulse, containing the components to be separated, is injected into a chromatographic column filled with a suitable adsorbent, alternating with the supply of pure solvent. On its way along the column, the mixture is gradually separated and the products can be fractionated at the column outlet. One of the major drawbacks of this method is the high amount of solvent needed to perform the separation. This also leads to a high dilution of the products. During the migration of the components along the column, only a small part of the adsorbent is used for the separation. Another disadvantage is the batch operation mode of the process. In industrial applications, processes with continuous product streams are preferred.

These drawbacks led to the development of continuous countercurrent adsorption processes. The main advantage of such an arrangement is the countercurrent flow, as in heat exchangers or distillation columns, that maximizes the average driving force. Thus, the adsorbent is used more efficiently. However, the movement of the solid particles is very difficult to realize. One reason is the inevitable back-mixing of the solid that reduces the separation efficiency of the columns. Another problem is the abrasion of the particles which is caused by the movement.

The invention of the Simulated Moving Bed process overcame these difficulties, providing a profitable alternative mainly for the separation of binary mixtures. The countercurrent movement of the phases is approximated by sequentially switching the inlet and outlet valves of interconnected columns in the direction of the liquid flow. According to the position of the columns relative to the feed and the draw-off nodes, the process can be divided

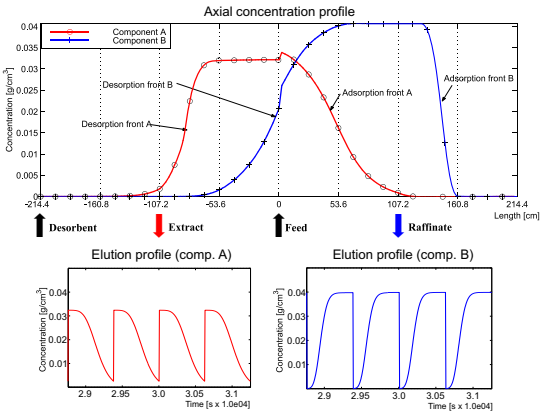


Figure 3: Cyclic steady state of the simulated moving bed process. Top: axial concentration profile (end of period), Bottom: elution profiles.

into four different sections (see Figure 2). The flow rates are different in every section and each section has a specific function in the separation of the mixture. The separation is performed in the two central sections where component B is desorbed and component A is adsorbed. The desorbent is used to regenerate the adsorbent by desorption of component A in the first section, and component B is adsorbed in the fourth section to regenerate the desorbent. The net flow rates of the components have different signs in the central sections II and III, thus component B is transported from the feed inlet upstream to the raffinate outlet with the fluid stream and component A is transported downstream to the extract outlet with the "solid stream".

The stationary operating regime of the SMB process is a cyclic steady state (CSS), in which in each section an identical transient takes place during each period between two valve switches. This periodic orbit is practically reached after a certain number of valve switches. The upper part of Figure 3 represents the axial concentration profile at the end of a switching period while operating in cyclic steady state. The resulting elution profiles below represent the time history of the product concentrations and highlight the periodic nature of the process dynamics.

Modeling and Simulation

Modeling of the SMB Process

The modeling and simulation of SMB processes has been a topic of intensive research in recent years. An overview can be found e.g. in Ruthven and Ching (1989), Ganetsos and Barker (1993), Zhong and Guiochon (1998) and Klatt (1999). The modeling approaches can be divided into two classes. In the first class, a rigorous SMB model is assembled from dynamic process models of the single chromatographic columns under explicit consideration of the cyclic switching operation. Alternatively, an equivalent solid velocity is deduced from the switching time and the balance equations for the corresponding true moving bed (TMB) are used.

By neglecting the cyclic port switching, the model is distinctly simplified, and can be solved very efficiently. It can be shown that the steady state solution of a detailed TMB model reproduces the concentration profile of a SMB model reasonably well in case of three or more columns per zone and linear adsorption behavior, which justifies the use of this type of model for the design of such units (Storti et al., 1988; Lu and Ching, 1997; Pais et al., 1998). However, many of the recent applications of the SMB process, especially in the area of fine chemicals and pharmaceuticals, are operated with less columns per zone and at higher concentrations with nonlinear adsorption equilibrium for economic reasons. In these cases, the accuracy of the TMB approximation becomes poor. Furthermore, only the dynamic SMB model correctly represents the complete process dynamics, which is essential for an optimization of the operating policy and for model-based control.

The rigorous dynamic SMB model is closely related to the real process and directly describes the column interconnection and the switching operation. It mainly consists of two parts: the node balances to describe the connection of the columns combined with the cyclic switching, and the dynamic simulation models of the single chromatographic columns. The node balances are used to calculate the inlet flows and inlet concentrations of the four zones of the process based on the mass balances at the corresponding nodes (Ruthven and Ching, 1989):
Desorbent node:

$$\begin{aligned} Q_{IV} + Q_D &= Q_I \\ c_{i,IV}^{out} Q_{IV} + c_{i,D} Q_D &= c_{i,I}^{in} Q_I \end{aligned} \tag{1}$$

Extract node:

$$\begin{aligned} Q_I - Q_{Ex} &= Q_{II} \\ c_{i,I}^{out} &= c_{i,II}^{in} = c_{i,Ex} \end{aligned} \tag{2}$$

Feed node:

$$\begin{aligned} Q_{II} + Q_F &= Q_{III} \\ c_{i,II}^{out} Q_{II} + c_{i,F} Q_F &= c_{i,III}^{in} Q_{III} \end{aligned} \tag{3}$$

Raffinate node:

$$\begin{aligned} Q_{III} - Q_{Raf} &= Q_{IV} \\ c_{i,III}^{out} &= c_{i,IV}^{in} = c_{i,Raf} \end{aligned} \tag{4}$$

with Q_i being the respective flow rate in each of the four zones, Q_D the desorbent flow rate, Q_F the feed flow rate, Q_{Ex} the extract flow rate, and Q_{Raf} the raffinate flow rate. The switching operation can, from a mathematical point of view, be represented by shifting the

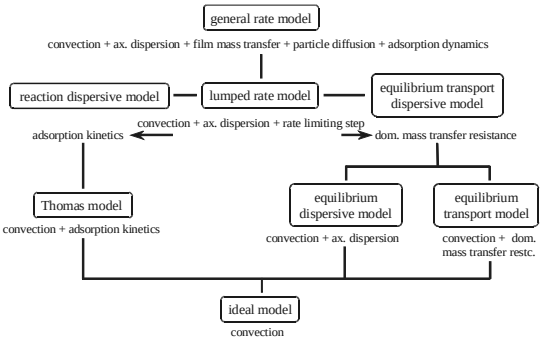


Figure 4: Classification of column models.

initial or boundary conditions for the single columns. Using the node model, the dynamic models of the single columns are interconnected. This modular approach allows the use of different column models which are appropriate for the problem at hand.

Modeling of a Chromatographic Column

Modeling and simulation of chromatographic separation columns has been a topic of research since the 1950s. An overview can be found in Guiochon et al. (1994), an interesting presentation of the phenomenological background is given by Tondeur (1995). Mostly, apart from few outdated approaches using a stage model, the model is formulated by a differential mass balance on a cross section of the chromatographic column. Many different modeling approaches can be found in the literature, and those can be classified by the phenomena which they include and by their level of complexity (see Figure 4).

The simpler modeling approaches in the bottom of Figure 4 can partly be solved analytically, and therefore they can be evaluated very efficiently (see e.g. Rhee et al., 1989; Helfferich and Whitley, 1996; Zhong and Guiochon, 1996; Dünnebier and Klatt, 1998; Dünnebier et al., 1998). However, the idealistic assumptions on which they are based are very unrealistic for most real systems. For optimal operation and control, a model which is both accurate and computationally efficient is essential. The more complex process models mainly require an appropriate numerical solution strategy. The models consist of a set of partial differential equations of the convection-diffusion (or hyperbolic-parabolic) type. Some properties of this type of equations, like shock layers and almost discontinuous solutions, make the application of many standard discretization procedures difficult. A lot of research has been devoted to the development of suitable spatial discretization schemes for chromatography column models in order to transform the PDEs to a set of ODEs (see e.g. Kaczmarski et al., 1997; Kaczmarski and Antos, 1996; Strube and Schmidt-Traub, 1996; Poulain and Finlayson, 1993; Ma and Guiochon, 1991; Spieker et al., 1998). Common to most of the known approaches is the need for large computational power which makes it difficult, even with modern computers, to perform simulations substantially faster than real time.

Therefore, the first objective of our research on model-based control of chromatography processes was the formulation and implementation of suitable process models. We followed a bottom up strategy, proceeding from the ideal model and increasing the complexity as far as necessary in order to achieve sufficient accuracy. From a mathematical point of view, it is useful to distinguish chromatographic processes by the type of adsorption isotherms, which describe the thermodynamic equilibrium of the separation system. Processes with a linear relation between the fluid phase concentration c_i and the solid phase concentration q_i (Henry's law)

$$q_i = K_{H,i} \cdot c_i \tag{5}$$

lead to systems of decoupled differential equations which are easier to solve than those with coupled nonlinear adsorption behavior, described for instance by competitive Langmuir isotherms

$$q_i = \frac{a_i}{1 + \sum_{j=1}^{n} b_j c_j} \cdot c_i \ . \tag{6}$$

Van Deemter et al. (1956) have shown that in case of a linear isotherm the effects of axial dispersion and mass transfer resistance are additive and can be incorporated into a single parameter, the apparent dispersion coefficient D_{ap}. This results in the following quasi-linear parabolic partial differential equation for the fluid phase concentration of each component

$$\gamma_i \frac{\partial c_i}{\partial t} + u_L \frac{\partial c_i}{\partial x} - D_{ap,i} \frac{\partial^2 c_i}{\partial x^2} = 0 \tag{7}$$

The parameter γ is defined as

$$\gamma_i = 1 + \frac{1-\varepsilon}{\varepsilon} K_{H,i} \qquad (i = A, B)$$

where ε represents the column void fraction, and the interstitial velocity u_L is assumed to be constant. Lapidus and Amundsen (1952) proposed a closed form solution of this type of equation for a set of general initial and boundary equations by double Laplace transform. From this, the dynamic SMB model can be generated by connecting the solutions for each single column by the respective node model (see Dünnebier et al., 1998, for details of the implementation). We denote this implementation as the DLI model (dispersive model for linear isotherms). It was shown that simulation times two orders of magnitude below real-time can be achieved while reproducing both the results obtained with more complex simulation models and experimental results very accurately.

In order to generate both an accurate and computationally efficient dynamic model also in the case of general nonlinear adsorption isotherms we first followed the same approach as in the linear case by analyzing a model where the non-idealities are lumped into a single parameter. However, a closed-form solution is no longer possible in the nonlinear case and the numerical solution using standard techniques for the spatial discretization did not improve the computational efficiency substantially. Fortunately, there exists a very effective numerical solution for the detailled general rate model

$$\begin{aligned}\frac{\partial c_i}{\partial t} - D_{ax}\frac{\partial^2 c_i}{\partial x^2} + u_L\frac{\partial c_i}{\partial x} + \frac{3(1-\varepsilon)k_{l,i}}{\varepsilon r_p}(c_i - c_{p,i}(r_p)) = 0\\ (1-\varepsilon_p)\frac{\partial q_i}{\partial t} + \varepsilon_p\frac{\partial c_{p,i}}{\partial t} - \varepsilon_p D_{p,i}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c_{p,i}}{\partial r}\right)\right] = 0\end{aligned} \tag{8}$$

with complex nonlinear isotherms proposed by (Gu, 1995). Here, D_{ax} represents the axial dispersion coefficient, $c_{p,i}$ the concentration within the particle pores, and q_i the solid phase concentration which is assumed to be in equilibrium with the pore-phase concentration. r_p denotes the particle radius, ε_p the particle porosity, $k_{l,i}$ the respective mass transfer coefficient, and $D_{p,i}$ the diffusion coefficient within the particle pores. A finite element formulation is used to discretize the fluid phase, and orthogonal collocation for the solid phase. We applied this formulation to SMB processes resulting in a superb accuracy and simulation times almost two orders of magnitude below real time (Dünnebier and Klatt, 2000). Due to the favorable numerical properties, this certain implementation of the complex general rate model in terms of computational efficiency even outperforms the state of the art simulation models for SMB processes (equilibrium transport dispersive model—second layer of complexity in Figure 4) which follow a linear driving force approach with a lumped mass transfer rate (e.g. Strube and Schmidt-Traub, 1996; Kaczmarski and Antos, 1996; Kaczmarski et al., 1997). Furthermore, in terms of physical consistency the general rate model is more exact, because the lumping of the different mass transfer phenomena is strictly valid only for systems with linear isotherms and incorrect for substances with large molecules (as they appear e.g. in bioseparations).

Optimal Operating Regime

State of the Art

Most of the known approaches for the determination of operating parameters for simulated moving bed separation processes are not based on mathematical optimization methods and rigorous dynamic process models. Two main approaches can be distinguished: The first is to derive short-cut design methodologies based on the equivalent TMB process. The second type of work uses heuristic strategies combined with experiments and dynamic simulation of the SMB model.

By transforming the switching time τ into an equivalent solid flow rate

$$Q_S = \frac{(1-\varepsilon)V_{col}}{\tau} \tag{9}$$

the operating parameters of a SMB process can be expressed in terms of the operating parameters of the corresponding TMB process. In case of the ideal model and linear adsorption isotherms according to Equation 5, the TMB model can be solved in closed form. On the basis of this solution Nicoud (1992) and Ruthven and Ching (1989) introduced new operating parameters β_i and stated bounds for which the desired separation can be achieved:

$$\begin{aligned}Q_F &= Q_S(K_{H,A}/\beta_{III} - K_{H,B}\beta_{II})\\ Q_{Ex} &= Q_S(K_{H,A}\beta_I - K_{H,B}\beta_{II})\\ Q_D &= Q_S(K_{H,A}\beta_I - K_{H,B}/\beta_{IV})\\ Q_{IV} &= Q_S(K_{H,B}/\beta_{IV} + \frac{\varepsilon}{1-\varepsilon})\\ 1 \le \beta_i \le \sqrt{\frac{K_{H,A}}{K_{H,B}}}, &\qquad K_{H,B} < K_{H,A}\end{aligned} \tag{10}$$

The β-variables were originally intended as slack variables to formulate the conditions for proper operation of the separation unit as a set of inequalities. The bounds on those variables result from retaining the adsorption and desorption fronts in the appropriate zone of the unit, i.e. the β's are safety factors for the ratio of the net mass flow rate of the solid and the liquid phase. A detailled exposition of this subject can be found in Zhong and Guiochon (1998).

As the objective function in this framework, the minimization of the specific desorbent consumption Q_D/Q_F or the maximization of the throughput Q_F can be used. Both objective functions lead to the same solution in the idealistic case, which is $Q_D = Q_F$ at the boundary of the feasible region with $\beta_i = 1$ and $Q_D/Q_F = 1$. The maximum feed inflow Q_F is bounded by the maximum allowed internal flow rates, which are limited by the pressure drop or the efficiency of the adsorbent. Storti et al. (1993), Mazzotti et al. (1997), and Migliorini et al. (1998) derived a graphical short-cut design methodology based on these ideas, the so-called *triangle theory* and extended the theory to systems with nonlinear adsorption isotherms. This methodology is currently state of the art and has been applied to a large number of separations. Due to the unrealistic assumptions of the ideal model, the triangle theory can only give initial guesses for a feasible operating point of the process because it does not permit a reliable prediction of the product purities which are the most important operating constraints.

To overcome the limitations of the ideal model, Ma and Wang (1997) and Wu et al. (1998) presented a stand-

ing wave design approach based on an equilibrium dispersive transport model of the TMB process. However, in case of systems with dominant mass transfer or diffusion effects the minimization of the desorbent consumption and the maximization of the throughput become conflicting. Additionally, as already mentioned, the quality of the prediction of the TMB model in general is only sufficient for a restricted range of applications. It is therefore necessary to develop design strategies based on rigorous dynamic SMB process models. The operating points determined with the triangle theory or the standing wave design can be used as initial guesses for the computation of an optimal operating point.

Strube et al. (1999) describe a heuristic design strategy based on the dynamic simulation of a SMB process model. The optimization objective is formulated in a set of competing and partly contradictory targets which are approached by parameter variation based on a heuristic strategy. The major advantage of this method is the use of a realistic process model, the disadvantage is the need for extensive manual simulation without any guarantee to determine the optimum, and the need for an experienced user.

From the shortcomings and the limitations of the previously described approaches, a list of desired properties for a model-based optimization strategy for simulated moving bed chromatographic processes can be stated as follows:

1. The algorithm has to be based on a realistic and efficient dynamic SMB process model.
2. The objectives for the optimization must be formulated in a single objective function avoiding competing and contradictory targets.
3. The product quality has to be included explicitly in the formulation since this is the most relevant constraint for the operation.
4. The strategy has to be based on a mathematical optimization procedure. This is the only way to ensure that, in combination with suitable initial guesses, a solution at least close to the optimum can be obtained in finite time and without requiring too much costly expertise.
5. The procedure should be as general as possible, so that it can be applied to a broad variety of SMB processes with linear or nonlinear adsorption equilibrium, any number of columns and any size of equipment.

A New Model-Based Optimization Strategy

To the best of our knowledge, there are only two approaches documented in the literature which treat the calculation of the optimal operating regime in a rigorous mathematical formulation: the strategy suggested by Kloppenburg and Gilles (1998), and the approach proposed in the sequel which is explained in detail in Dünnebier and Klatt (1999); Dünnebier et al. (2000).

We here consider the case where the plant design is fixed and the feed inflow is pre-specified, e.g. by the outflow of an upstream unit. In this case, the desorbent inflow Q_D constitutes the only variable contribution to the processing costs. By defining c_k as the axial concentration profile at the end of a switching period, and by expressing both the process dynamics between two switching operations and the switching operation itself by the operator Φ

$$c_{k+1} = \Phi(c_k) \tag{11}$$

we can write the following condition for the cyclic steady state:

$$\|\Phi(c_k) - c_k\| \leq \delta_{css} \tag{12}$$

The purity requirements for the products in the extract and raffinate stream are formulated as inequality constraints. Besides that, the efficiency and functionality of most adsorbents is only guaranteed up to a maximum interstitial velocity, which results in a constraint $Q_{\max}$ for the flow rate in the first section of the process. The optimization problem can then be stated as follows:

$$\min_{Q_j,\tau} \quad Q_D \tag{13}$$

subject to

$$\begin{gathered}
\|\Phi(c_k) - c_k\| \leq \delta_{css} \\
\int_0^\tau \frac{c_{A,Ex}(t)}{c_{A,Ex}(t) + c_{B,Ex}(t)} dt \geq Pur_{Ex,\min} \\
\int_0^\tau \frac{c_{B,Raf}(t)}{c_{A,Raf}(t) + c_{B,Raf}(t)} dt \geq Pur_{Raf,\min} \\
Q_I \leq Q_{\max}
\end{gathered}$$

Although the values for the optimization variables are constant during the switching periods, the optimization problem is inherently a dynamic one, because the operating regime of a SMB process is a periodic orbit and not a steady state. From a mathematical point of view, this is due to the dynamic nature of Equation 12 which is a system of partial differential equations with switching initial and boundary conditions, constituting the crucial constraint of the formulation given in Equation 13.

The natural choices of the degrees of freedom for the calculation of the optimal operating regime are the desorbent flow rate Q_D, the extract flow rate Q_{Ex}, the switching time τ and the recycle flow rate Q_{IV}. Due to the complex interactions, this results in a strongly coupled system dynamics, since each of the independent variables affects every zone of the process. Furthermore, it is impossible to formulate explicit and independent bounds on those variables, and the optimization problem is not well-conditioned from a numerical point of view.

We therefore exploit the results of the analysis for the ideal process model in case of linear isotherms. Equation 10 sets up a transformation where the "natural degrees of freedom" Q_D, Q_{Ex}, Q_{IV}, and τ are replaced by the variables β_i. Due to the physical meaning of the β_i (safety margins for stable process operation), the feasible region for the optimization problem is well-defined in terms of the bounds for the transformed variables given in Equation 10. Furthermore, due to the scaling effect and based on the feature, that any of the variables β_i mainly affects one zone of the process, the variable transformation results in a more favorably structured optimization problem. In case of nonlinear isotherms, the slope of the isotherm is concentration dependent and can no longer be represented by the constant Henry-coefficient K_H. In order to extend Equation 10 to the nonlinear case, we thus chose a reference concentration to formulate the transformation:

$$
\begin{aligned}
Q_S &= \frac{Q_F}{\frac{\partial g_A}{\partial c_A}(c_F)/\beta_{III} - \frac{\partial g_B}{\partial c_B}(c_F)\beta_{II}} = \frac{(1-\varepsilon_b)AL}{\tau} \\
Q_{Ex} &= Q_S\left(\frac{\partial g_A}{\partial c_A}(c_F)\beta_I - \frac{\partial g_B}{\partial c_B}(c_F)\beta_{II}\right) \\
Q_D &= Q_S\left(\frac{\partial g_A}{\partial c_A}(c_F)\beta_I - \frac{\partial g_B}{\partial c_B}(c_F)/\beta_{IV}\right) \\
Q_{IV} &= Q_S\left(\frac{\partial g_B}{\partial c_B}(c_F)/\beta_{IV} + \frac{\varepsilon}{1-\varepsilon}\right)
\end{aligned}
\tag{14}
$$

where

$$
g_i = \varepsilon_p c_{p,i} + (1-\varepsilon_p) q_i(c_A, c_B) \qquad (i = A, B)
$$

and describes the equilibrium isotherm. The feed concentration is a simple and suitable choice to calculate the slopes. The bounds on the variables β_i as given in Equation 10 are based on stability considerations for the ideal model and linear adsorption equilibrium and can therefore not simply be transferred to more complex systems. In case of nonlinear isotherms the bounds have to be relaxed suitably, and in terms of the triangle theory, the region limited by these bounds can be seen as a hull for the shaped triangle. However, the transformation Equation 14 helps to scale, structure and reduce the search space.

The crucial point in the solution of the optimization problem (13) is the efficient and reliable computation of the cyclic steady state. The solution of the associated fix-point problem given by Equation 12 corresponds to a mathematically similar formulation for pressure swing adsorption (PSA), the dynamics of which shows some similarities to SMB processes. Three main approaches have been identified and examined in the literature. Firstly, the solution can be achieved by direct dynamic simulation (Picard iteration). Alternatively, Equation 12 can be solved by a Quasi-Newton scheme (Unger et al., 1997; Kvamsdal and Hertzberg, 1997; Croft and Le Van, 1994), or by discretizing the equations in space and time and solving the resulting system of algebraic equations (Nilchan and Pantelides, 1998; Kloppenburg and Gilles, 1998).

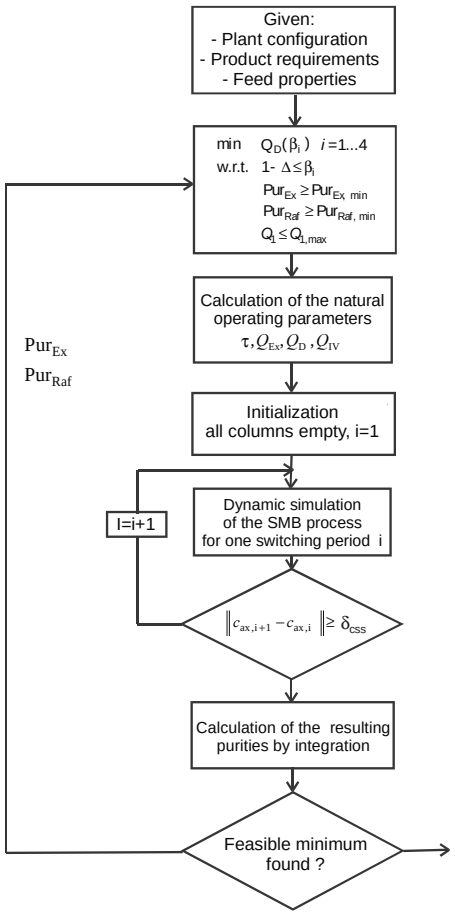


Figure 5: Optimization algorithm.

Even though there are many analogies between the SMB chromatography and PSA processes, the convergence of the direct dynamic simulation approach is much faster in the SMB case. For realistic SMB chromatography processes, the cyclic steady state is reached after a few hundred switching periods at the latest, whereas in the PSA case possibly several thousand cycles have to be evaluated. The inherent dynamics of the process are therefore much faster. On the other hand, especially for systems with nonlinear adsorption behavior and reaction kinetics, the system becomes very stiff and a very fine grid is needed, which makes the application of a global discretization approach difficult. We therefore choose the direct dynamic simulation as the most robust and efficient way for the calculation of the cyclic steady state within our optimization algorithm, which is schematically shown in Figure 5.

	Original	Run 1	Run 2
Pur_{Ex} [%]	99.5	99.5	98.0
Pur_{Raf} [%]	98.3	98.3	98.0
rel. Q_D [%]	100	**60.9**	**41.7**
Q_D [cm^3/s]	0.8275	0.5043	0.3452
Q_F [cm^3/s]	0.2500	0.2500	0.2500
Q_{Ex} [cm^3/s]	0.77989	0.4683	0.3489
Q_{IV} [cm^3/s]	0.2695	0.1489	0.1423
τ[s]	1337.4	2096.7	2675.2

Table 1: Optimization results for the separation of phenylalanine and tryptophan.

We follow a staged sequential approach for the solution of the dynamic optimization problem. The degrees of freedom β_i are chosen by the optimizer in an outer loop and are then transformed back into flow rates and switching times according to Equation 14. In the inner loop, the cyclic steady state is then calculated by direct dynamic simulation of the rigorous SMB process model. The purity constraints are evaluated by integration of the elution profiles for the cyclic steady state. The nonlinear program in the outer loop is small and can be solved by a standard SQP algorithm, while the required gradients are evaluated by perturbation methods. The results of the optimization are the optimal cyclic operating trajectory (CSS), the minimum desorbent inflow for the specified product purities, and the corresponding operating parameters.

The optimization algorithm was tested for a number of different separation systems with both linear and nonlinear adsorption equilibrium. One impressive example is shown in Table 1, the separation of phenylalanine and tryptophan. Estimated model parameters and the reference operating conditions were taken from Wu et al. (1998), where the original operating point was optimized by a standing wave design. Because of the nonlinear adsorption equilibrium (Langmuir isotherms), the general rate model was used within our optimization algorithm. The convergence to the cyclic steady state was achieved after 75 switching periods in the average, and the SQP solver in the outer loop converged after 12-15 steps depending on the initial point. This resulted in approx. 6-8 hours CPU time for each run on a 400 MHz PentiumII PC. In the first run, the purities for the operating point obtained by Wu et al. (1998) were taken as constraints. It can be seen, that compared to the operating regime determined by the standing wave design technique, the desorbent requirement was cut down by almost 40% using the proposed optimization approach. Furthermore, the purity requirements can be directly specified. The results for run 2 show the economical impact of a reduction of the purity specifications to 98% each.

Control Concept

In real applications, plant/model mismatch and disturbances will lead to more or less pronounced deviations from the optimal trajectory. However, the online optimization under real-time requirements is not possible with the computational power currently available. Therefore, a feedback control strategy based on suitable dynamic models and on-line measurement information is required in order to keep the process close to the optimal trajectory.

Only few publications can be found in the open literature which treat the automatic control of simulated moving bed chromatographic processes. Ando and Tanimura (1986), Cohen et al. (1997), Hotier and Nicoud (1996), and Hotier (1998) deal with the basic control of the internal flow rates, which itself is a difficult task and forms the basis for the more advanced control strategies. The concepts described in Holt (1995), Cansell et al. (1996), and Couenne et al. (1999) propose feedback control for certain operating variables (e.g. product purity, system yield) based on some concentration measurements. They are predominantly applied to the separation of aromatic hydrocarbons where on-line Raman spectroscopy (Marteau et al., 1994) can be utilized to measure the specific concentration of the compound at the outlet of the chromatographic columns. Those as well as the geometric nonlinear control concepts described in Kloppenburg and Gilles (1999) and Benthabet et al. (1997) are mainly based on a model for the corresponding true moving bed (TMB) process, where the cyclic port switching is neglected, and thus rely heavily on the applicability of the TMB model as a simplified model for the SMB process. This is particularly critical for SMB processes with a low number of columns—which are more and more utilized in industrial applications in order to reduce the investment costs—where the correspondence between the SMB dynamics and the TMB approximation may become poor. In a recent publication, Natarajan and Lee (2000) suggest to apply repetitive model predictive control on a reduced order linear state space model of the SMB process. They optimize the product yield for a given constraint on the purities using the reduced linear model for the control calculations. This approach definitely follows the right objective, i.e. the control of the process in the vicinity of an optimal operation point. However, in case of the SMB process, yield optimization even when considering the purity constraints does neither generically imply the economic optimum nor guarantee a stable process operation in the long run. Furthermore, the range of validity for the linear approximation and its impact on the control performance is not explicitly considered.

In order to overcome some of the shortcomings and limitations mentioned above, we proposed a two-layer control architecture which is shown schematically in Fig-

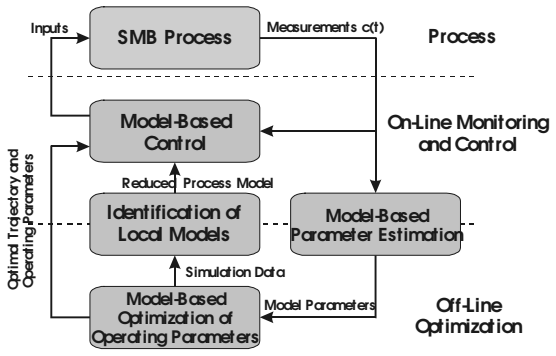


Figure 6: Control concept for SMB processes.

ure 6 (Klatt et al., 2000). The top layer features the off-line calculation of the optimal operating trajectory as described in the previous section, combined with an on-line estimation of the respective model parameters based on inline concentration measurements. The purpose of this estimation algorithm is twofold: It provides actual and reliable values for the model parameters, and together with the dynamic simulation model, it enables the monitoring of the complete axial concentration profile which is not directly measurable in the interior of the separation columns. If there is a too large discrepancy between the actual parameters and the parameter values used in the trajectory optimization (e.g. caused by aging of the adsorbent material), a new optimization run is initiated. The remaining control task then is to keep the process along the calculated nominal trajectory despite disturbances and plant/model mismatch, caused e.g. by small and non-persistent perturbations of the system parameters. This task is performed by the bottom layer where identification models based on simulation data of the rigorous process model along the optimal trajectory are combined with a suitable local controller. The realization of the remaining elements of the proposed control structure is explained below.

The model-based parameter estimation utilizes on-line measurement data from measurement devices located in the product outlets or in the connecting pipelines between the columns. Because of the high costs of online concentration measurements, measurements located after each single column will generally not be feasible. Typically, up to four measurement points can be found in real plants, if online measurement is employed at all. The devices are system specific and use either spectroscopic methods or combined measurement techniques to determine the concentration of each single species.

Starting from a complete set of model parameters determined in a priori experiments, the objective of the online estimation is to fit the model to the real process. The model parameters can be categorized in kinetic parameters (describing mass transfer, diffusion, dispersion) and adsorption parameters (constituting the adsorption isotherms). Due to the limited measurement information available, we solve a reduced parameter estimation problem, where one crucial parameter per class and component is adapted by optimizing a quadratic cost functional

$$\min J = \int_{t_1}^{t_2} (c_{i,meas}(t) - c_{i,sim}(t))^2 dt, \qquad i = A, B \tag{15}$$

where the measured outlet concentrations $c_{i,meas}$ are compared to the outlet concentrations $c_{i,sim}$ which are determined by the solution of the simulation model. The number and arrangement of the measurement devices depend on the mixture to be separated. In Zimmer et al. (1999) a measurement setup and estimation algorithm was proposed for systems with linear adsorption isotherms which is briefly sketched in the application example below.

The rigorous dynamic SMB model, consisting of PDEs (eqs. (7) and (8)) with switching initial and boundary conditions, is not well suited for a standard controller design. Thus, we base the trajectory control on a local model. In order to get rid of the hybrid system dynamics in the bottom layer of the control concept, we consider the reduced model as a discrete time model with the sampling interval equal to one switching period. The model predicts one characteristic parameter of the concentration profile per switching period, and therefore does not require the consideration of the discontinuities introduced by the switching operation.

Because of the favorable properties of the nonlinear transformation of the input space (14) which is utilized in the trajectory optimization, the variables β_i, resp. the deviations from their values on the nominal trajectory, are also chosen as inputs of the reduced model. This allows a variable switching time which is essential in case of inflow disturbances because feed flow-rate changes can be completely compensated only by a corresponding adaptation of the virtual solid stream. Furthermore, with an appropriate choice of outputs, the nonlinear transformation helps to create a nearly decoupled system dynamics, particularly for SMB separation processes with linear adsorption equilibrium.

One characteristic indicator of the separation performance of a SMB unit is the axial position of the adsorption and desorption fronts of the two components at a certain moment, e.g. at the end of a switching period (see Figure 3). They are referred to as front positions p_i in the sequel. Extensive simulation studies showed that the front positions do not only provide a suitable indicator for the long term stability of the system, but additionally exhibit an excellent correspondence with the product purities which are the relevant output variables. However, from a system dynamics point of view, the product concentrations are ill-suited as controlled variables due

to their strongly delayed response and their lack of sensitivity to changes in the manipulated variables. Thus, the deviations of the front positions from their nominal values on the optimized trajectory are chosen as outputs of the local model.

The axial concentration profile is not directly measurable, but assuming an on-line concentration measurement at the end of each column, an approximation of the complete axial concentration profile can be obtained by connecting all measurements of the elution profiles during one switching period. We call this representation the Assembled Elution Profile (AEP). If the system is at its cyclic steady state, the AEP can also be obtained by using the information of a single measurement over n_{col} switching periods. In case of a disturbed system, the AEP obtained from a system with less than n_{col} measurements becomes in some respect time variant, since the parts of the AEP are recorded at different instants during the transition, and the AEP constructed depends on the location of the measurements in the loop. Nevertheless, a reduction of the number of measurements below n_{col} is possible to a certain extent, and the effect on the AEP is comparable to a low pass filter. The required number and the exact location of the measurement devices depend on the specific system dynamics.

Three different methods for the calculation of the front positions from the AEP were evaluated:

I. Functional approximation of the fronts, followed by the calculation of the inflection point of this function.

II. Surface quadrature to calculate the center of gravity of the fronts.

III. Wavelet analysis to determine the inflection points of the fronts.

For method I, a variation of the Gauss error function proved to supply a good approximation for the fronts of SMB separation processes with linear and moderately nonlinear adsorption isotherms. The desorption fronts are approximated by

$$f(t) = \frac{1}{2}\left[1 + \frac{2}{\sqrt{\pi}}\int_0^{k(t-z)} e^{-\theta^2} d\theta\right], \qquad (16)$$

the adsorption fronts by

$$f(t) = \frac{1}{2}\left[1 - \frac{2}{\sqrt{\pi}}\int_0^{k(t-z)} e^{-\theta^2} d\theta\right]. \qquad (17)$$

The two parameters of the function, k and z are calculated by a least-squares fit to the available measurements which is updated in each switching period, and the inflection point of the function which represents the front position is determined analytically afterwards.

In the second method, an upper and lower bound for the concentration front is assumed first. Then the front position is defined as the point where $A_1 = A_2$ (see Figure 7). The integration is approximated by a summation over the available measurement points.

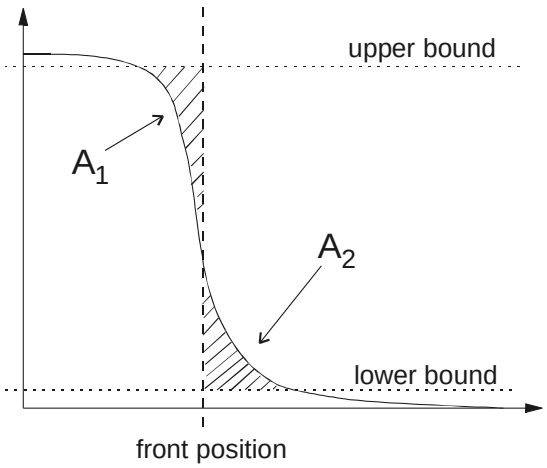


Figure 7: Evaluation by surface quadrature.

Wavelet analysis (III) is based on scale and position dependent coefficients of the form

$$C(a,k) = \int_{-\infty}^{\infty} f(t)\Psi(a,k,t)dt \qquad (18)$$

and allows the multi-scale analysis and representation of arbitrary time series. It also enables the direct identification of inflection points from noisy data (see, e.g. Daubechies, 1992). The method of choice mainly depends on the characteristics of each different separation process, i.e. the shape and the position of the adsorption and desorption fronts. The different methods have to be evaluated trading off the computational requirements for the calculation, the robustness against noise, and the physical meaning of the calculated positions.

Having defined the inputs and outputs, the structure of the local model and the input signal for the identification need to be specified. In our work, we utilize the MATLAB System Identification Toolbox (Ljung, 1995) to perform the necessary computations. To obtain the necessary data for identification, the rigorous simulation model is excited around the nominal operation regime with a random binary signal. For MIMO identification, non-correlated signals in the different channels are to be preferred, which can be achieved with a slight modification of the Pseudo Random Binary Signal, the Pseudo Random Multistep Signal (Isermann, 1992). We tested prediction error models as well as subspace state space identification methods and found both approaches suitable for the SMB separation processes we have investigated so far.

After identification of the local linear model any standard linear control design approach or linear model predictive control can be used to realize the controller on

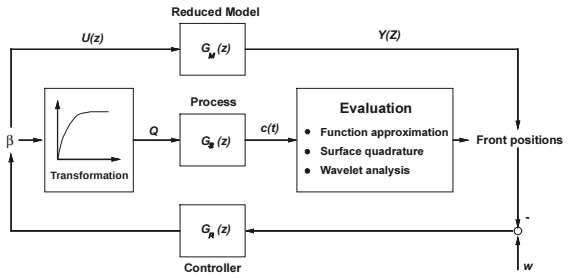


Figure 8: Structure of the trajectory control loop.

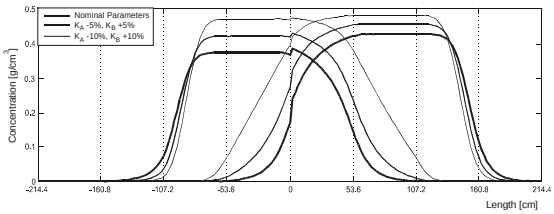


Figure 9: Axial profile (CSS, end of period) of the SMB fructose/glucose separation and its sensitivity to parameter perturbations.

System Parameters			
d_{col}	2.6 cm	L	53.6 cm
ε	0.38	d_p	0.0325 cm
ρ	1 g/cm^3	η	0.0058
K_A	0.54	K_B	0.28
Nominal Operating Parameters			
Concentration of feed c_F			0.5 g/cm^3
Feed flow rate Q_F			0.02 cm^3/s
Desorbent flow rate Q_D			0.0414 cm^3/s
Extract flow rate Q_{Ex}			0.0348 cm^3/s
Recycle flow rate Q_{IV}			0.0981 cm^3/s
Switching time τ			1552 s
Purity fructose			99.95%
Purity glucose			99.95%
β	[1.1371	1.0993	1.1164 1.1220]

Table 2: System and operating parameters for the separation of fuctose and glucose on an 8-column SMB laboratory plant.

the bottom layer. At a glance, the structure of the trajectory control loop is depicted in Figure 8. The variables represent deviations from the nominal trajectory: four input variables $u_i = \Delta\beta_i$, and deviations of the four front positions (normalized to the scaled length of all interconnected columns between 0 and 1) as outputs $y_i = \Delta p_i$.

Application Example

As an application example, we consider the separation of a fructose/glucose mixture on an 8-column laboratory scale SMB plant. Performing the proposed optimization algorithm, the plant was optimized for a product purity of 99.95% both in the extract and the raffinate stream. From a practitioner's point of view, this is an exceptionally high purity requirement for a sugar separation. Those are normally operated with a purity specification of at most 99%. However, the higher the purity specification the more challenging the control problem, because the system is operated much closer to the stability margins. Furthermore, in pharmaceutical separations the purity requirements are generally very high. Thus we decided to operate the sugar separation under these high purity requirements to demonstrate the feasibility of the proposed approach.

The system and operating parameters are shown in Table 2. The plant is operated at 60 °C. The liquid density can be considered as constant for the given feed concentration, and the adsorption isotherms are well described by Henry's law (5). Thus, the DLI model (7) is utilized within the optimization and control framework. The optimal operating trajectory and parameters were determined using the optimization algorithm described above. Here, convergence to the cyclic steady state was achieved after 65 switching periods on the average, and the SQP solver in the outer loop converged after 10-15 steps depending on the initial point. This resulted in approx. 4 hours CPU time on a 400 MHz PentiumII PC. Because of the complex hybrid system dynamics, steady state multiplicities are possible in principle. We therefore tested different initial points within the feasible region, but for the separation task at hand they all converged to the optimal solution reported in Table 2. However, this is of course system dependent and has to be thoroughly inspected for each individual separation task.

Figure 9 shows the axial concentration profile for the optimal operation mode and its sensitivity against variations of the adsorption isotherm parameters, caused e.g. by a varying feed quality. We here perturbed the system into the direction where the separation becomes more difficult, i.e. decreasing K_A and increasing K_B. Table 3 depicts the corresponding product purities. It is obvious, that the optimal operating regime is quite sensitive against parameter variations and thus the results of the trajectory optimization illustrated above essentially depend on a reliable determination of the model parameters.

In Zimmer et al. (1999) an on-line parameter estimation algorithm for systems with linear adsorption isotherms based on concentration measurements only in the product outlets was proposed. A combination

	Nominal	K_A -5% K_B +5%	K_A -10% K_B +10%
Extract	99.95%	99.50%	97.04%
Raffinate	99.95%	99.61%	96.71%

Table 3: Product purities corresponding to the concentration profiles shown in Figure 9.

of a polarmeter and a densimeter is used to determine the specific concentrations of fructose and glucose (Altenhöhner et al., 1997). Assuming the validity of the DLI model (7), the system has, in the case of a binary separation, four parameters describing the physical behavior, which are γ_i and $D_{ap,i} (i = A, B)$. γ incorporates the column void fraction ε and the respective adsorption parameter, the apparent dispersion coefficient D_{ap} lumps the kinetic effects. Thus, no further reduction of the parameter space is necessary in this case.

A linear adsorption isotherm implies that the components do not interact with regard to their adsorption behavior. Therefore, the two parameter sets $(\gamma_A, D_{ap,A})$ and $(\gamma_B, D_{ap,B})$ can be determined seperately by solving the least-squares problem posed by Equation 15. The simulated axial concentration profile $c_{i,sim}$ is determined by solving the PDE (7) for each column and each component. Unfortunately, the boundary and initial concentrations for each switching period are only partly available from the measurements. Therefore, a special approximation strategy was developed using the concentration of fructose and glucose in the extract and raffinate outflow in two consecutive switching periods. This parameter estimation additionally supplies an approximation of the non-measurable concentration profiles in the interior of the columns prior to the measurement. Because the measurement devices are located in the product outlets which are periodically switched downstream, the complete axial concentration profile can be reconstructed within a complete cycle (i.e. 8 switching periods).

The major task of the parameter estimation is to supply a set of actual model parameters which can be used for the trajectory optimization. Also they indicate if there is too large a deviation from the original set of parameters and thus a new optimization run has to be performed. However, as online optimization is not possible for the time being, smaller deviations and other disturbances have to be adjusted by the trajectory control loop.

For the fructose/glucose separation, a linear time invariant model of the ARX type was chosen as the local dynamic model:

$$\mathbf{A}(q)y(t) = \mathbf{B}(q)u(t) + e(t) \tag{19}$$

where $\mathbf{A}$ and $\mathbf{B}$ are 4×4 matrices containing polynomial functions of the discrete time shift operator q in

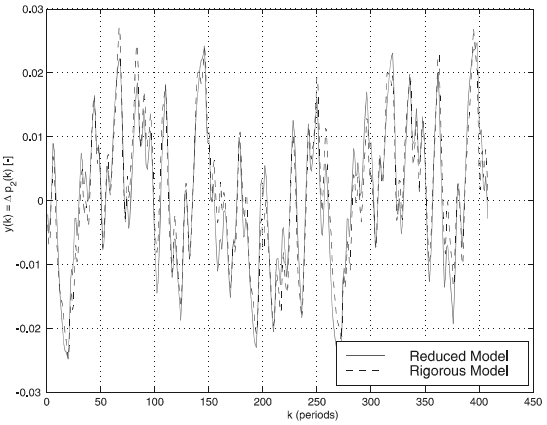


Figure 10: Model validation.

each element. The amplitude of the input signal $\Delta\beta_i$ for excitation of the rigorous DLI model was chosen to be 10% of the nominal β-value. This input range covers nearly the whole range of practically reasonable operating conditions. A first estimate of the necessary system order was obtained by the evaluation of step responses of the DLI model, which was then adapted based on the performance of the identified model. For the elements of the matrices $\mathbf{A}$ and $\mathbf{B}$, polynomials of at most second order proved to be sufficient. The performance of the model was tested with a second validation data set. The prediction of the front positions (desorption front of glucose (component B) as one example) of the rigorous and of the reduced model are compared in Figure 10. The prediction horizon is infinite here, i.e. the simulation models are reconciled only for initialization and then run independently. The approximation is very good, only in regions far from the nominal trajectory, where the assumption of linearity is violated, slight deviations occur.

As postulated, the system dynamics are diagonally dominant in the vicinity of the optimal operating trajectory. As mentioned above, in principle, any standard linear controller design method could be applied. For this example, we transformed the ARX model (19) to a z-domain transfer function representation and applied internal model control following the guidelines proposed in Zafiriou and Morari (1985) to design a discrete time SISO controller for each channel (i.e. the control of each single front position by manipulating the respective β-value). The design of the internal model controllers is very straightforward for the system at hand and allows the direct integration of the system dynamics into the controller.

In the conventional feedback representation, the internal model controller is of the form

$$C(z) = \frac{F(z)G^C(z)}{1 - F(z)G^C(z)G^M(z)} \tag{20}$$

where G^C is the internal model controller, G^M is the transfer function of the respective model, and F is a first order filter

$$F(z) = \frac{z(1-\alpha)}{z-\alpha} \tag{21}$$

where α has to be suitably chosen in the range $0 \leq \alpha < 1$. After a cancellation of numerically induced pole/zero pairs, we obtained the following transfer functions for the respective input-output channels

$$\begin{aligned} G_1^M &= \frac{0.0349 \cdot z^2}{(z-0.145)(z-0.8011)} \\ G_2^M &= \frac{0.0256 \cdot z^2}{(z-0.7058+0.1106j)(z-0.7058-0.1106j)} \\ G_3^M &= \frac{-0.0365 \cdot z^2}{(z-0.6967+0.2277j)(z-0.697-0.2277j)} \\ G_4^M &= \frac{-0.0143 \cdot z^2}{(z-0.5743)(z-0.8270)} \end{aligned} \tag{22}$$

In this case, direct inversion of G_M is possible and the internal model controller is realized by

$$G_i^C = \frac{1}{z} \cdot \frac{1}{G_i^M(z)} \tag{23}$$

The filter parameters were selected from simulation studies as

$$\alpha_1 = 0.4, \quad \alpha_2 = 0.4, \quad \alpha_3 = 0.6, \quad \alpha_4 = 0.6$$

and the range of the trajectory controller was chosen as a $\pm 10\%$ deviation from its value at the optimal operating regime for each of the β_i.

In the following, some simulation studies are presented in order to demonstrate the capabilities of the proposed trajectory control. For the simulation scenarios, the general rate mode according to Equation 8 is used to represent the real plant. This introduces a structural deviation between the DLI model on which the control design is based and the simulation model for the validation experiments. The scenarios depicted below were performed both without and with noise added to the concentration measurements. White noise with a standard deviation of 1% of the maximum concentration value was assumed.

Flow-Rate Disturbance

In this scenario it is assumed that a step disturbance in each of the internal flow rates occurs one at a time to analyze the effect on all front positions and possible coupling. Under practical considerations, these disturbances may be caused by an irregularity in the pumping and piping system or by an offset in the basic control layer. Each of the subplots in Figure 11 shows the manipulated variable and controlled variable moves resulting from the disturbance in the respective input channel,

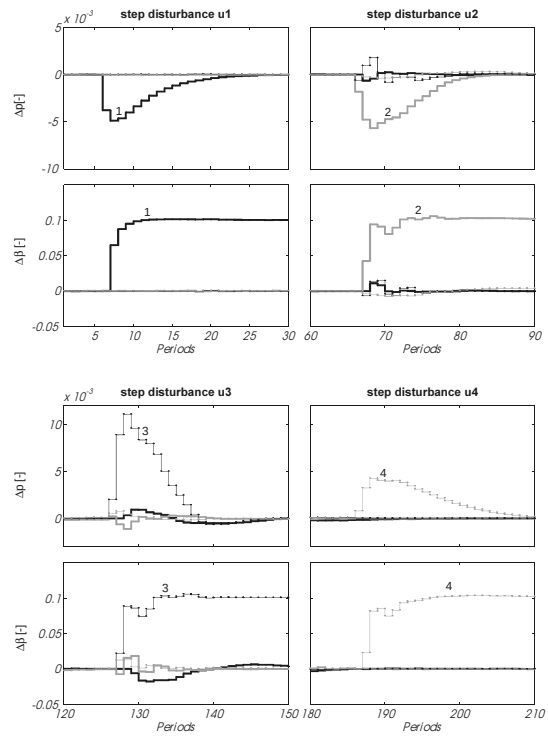


Figure 11: Closed loop reaction on input disturbance scenario.

where a 10% variation in terms of the β-variable corresponds to a 10-20% variation of the internal flow rates.

As can be clearly seen, the closed loop reacts in a mostly decoupled manner. The front position deviation is suppressed quickly by the controller, the maximum value of the deviation being about 0.01 in terms of scaled length. Though minor effects of coupling can be seen for disturbances in channels 2 and 3, the effects remain small. As expected, the results of the same sequence under noisy measurements show less smooth control actions and a slight wiggling of the front positions due to the noise's influence on the calculation of the inflection points. However, the results are principally the same with no major deterioration as compared to the noise-free setting. The corresponding plot is omitted here for the sake of brevity.

Feed Batch Change

In real-world applications a separation process follows prior production steps which might be continuous or discontinuous, a typical example in the Life Science context being a fermenter or a batch reactor. Thus a continuous SMB chromatography has to face changes in the feed batch quality. In Table 3 and in Figure 9 we already considered this scenario, which corresponds to a change of the characteristic adsorption parameters, and its im-

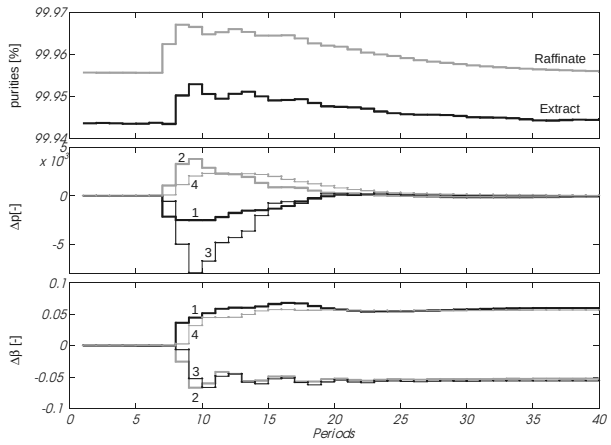


Figure 12: Closed loop response to change of feed batch.

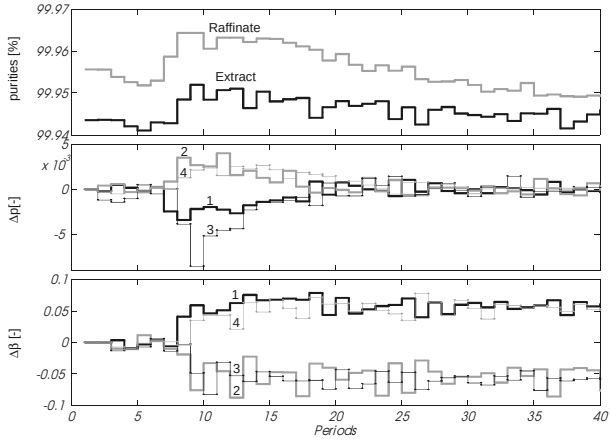


Figure 13: Closed loop response to change of feed batch in the presence of measurement noise.

pact on the axial concentration profile and the product purities within the cyclic steady state operating regime.

Starting form the optimal operating regime, after five periods we again perturbed the characteristic adsorption parameters by ±5%, but now with the feedback control loop closed. The results are shown in Figure 12 for the case of noise-free concentration measurements. The trajectory controller quickly suppresses the disturbance and drives the corresponding front positions (controlled variables) back to their nominal values without any offset. The axial concentration profile therefore does not change its position as in the scenario without feedback control, but is kept in the immediate vicinity of the optimal profile. As a result, the product concentrations which are the essential quality parameters are thus indirectly controlled and kept close to their optimal values. Figure 13 shows the same scenario under noisy measurements. The results are qualitatively the same and the effect on the product concentrations remains very small compared to the noise-free setting.

Conclusions and Further Research

Simulated moving bed chromatographic separation processes pose a challenging control problem because of their complex hybrid dynamics. In this contribution, we proposed an integrated, model-based approach for the optimal operation and advanced control of a SMB chromatographic process. To our knowledge, it is the first approach which is based solely on a rigorous dynamic process model and does not utilize the simplified TMB model in any of its elements. Furthermore, the proposed strategy explicitly aims at controlling the process close to its economic optimum.

The control architecture features two cascaded layers where in the top layer the optimal operating trajectory is calculated off-line by dynamic optimization. Due to the associated computational complexity, online optimization based on a rigorous dynamic process model is currently not possible. The respective model parameters are determined by an online estimation algorithm which, in addition to providing actual and reliable model parameters for the trajectory optimization, enables the online monitoring of the non-measurable concentration profile within the separation columns. If the discrepancy between the actual parameters and the parameter values on which the trajectory optimization was based on becomes significant, a new optimization run should be initiated. The task of the second layer of the control architecture is to keep the process along the optimized trajectory. This trajectory control is based on a local linear model which is identified from simulation data of the rigorous dynamic process model along the optimized trajectory.

The capabilities of the proposed control concept could be demonstrated in a set of simulation studies for the separation of fructose and glucose on an eight column SMB plant which is currently operated manually in the Plant Design Laboratory at the University of Dortmund. Due to the particular choice of the inputs and outputs, the dynamics of the local ARX model are approximately decoupled in the vicinity of the optimal operating trajectory, and thus a simple decentralized internal model controller could be used for the trajectory control. It performs satisfactorily, rejecting different types of disturbances both in a noise-free setting and with noisy measurements.

The overall control architecture is currently being implemented in an industrial standard control system. It consists of a decentralized control system (DCS) of the SIEMENS S7-400 series (CPU S7-414-2DP) and the

Windows Control Center (WinCC) as human machine interface. The algorithms and programs of the components of the control structure are integrated via the C-script interface, Global Script. Due to the fact, that the trajectory optimization is performed off-line, this task can be swapped out to another machine.

The experimental verification of the proposed approach on the laboratory SMB plant is planned for the near future, followed by the extension of the concept to a real pharmaceutical separation. However, for this latter task, some open issues have to be addressed because pharmaceutical separations may show a very complex nonlinear adsorption behavior and not all of the components of the proposed concept are currently completely adapted to this challenge. The off-line trajectory optimization can be performed for arbitrary complex isotherms, the successful extension to reactive adsorption processes was recently reported in Dünnebier et al. (2000). The online parameter estimation concept is currently realized only for systems with negligible coupling of the adsorption isotherms. This has to be extended in order to get a reliable online estimate for the interaction parameters of the adsorption isotherm in case those have a significant impact on the system dynamics. Due to the limited measurement information, a reduced parameter estimation problem has to be posed in the case of complex nonlinear isotherms, restricting the online adaptation to the most significant model parameters only. The design strategy for the inner control loop is also subject to revision because the range of validity for the local linear model may be too small for very nonlinear and strongly coupled systems. Currently, the combination of NARX models based on neural networks and nonlinear model predictive control as an alternative to realize the trajectory control loop for this type of separations is investigated. Furthermore, the correspondence between the characteristic points used for control and the product purities—which is excellent in case of the sugar separation example—has to be validated for different types of substances and the need for alternative calculation methods or even different characteristics to represent the respective fronts has to be inspected.

The above mentioned issues are addressed in current research projects and we expect a medium term solution for most of the problems. However, in the long run also the cascaded control structure is under discussion. A real time optimization based on a dynamic process model redundantizing the inner loop would be desirable. To achieve this goal, the computational efficiency of the dynamic optimization has to be enhanced by at least one order of magnitude. For this, both the numerical solution of the model equation as well as the calculation of the cyclic steady state have to be improved. Adaptive gridding for the solution of the model equations in combination with the direct solution of the sensitivity equations seems to be a promising approach and provides an interesting area for the cooperation between engineers and applied mathematicians.

Acknowledgments

We thank G. Zimmer and M. Turnu for their valuable contribution to the parameter estimation and reduced model identification issue, respectively. The financial support of the Bundesministerium fr Bildung und Forschung under grant number 03D0062B0 and of the Deutsche Forschungsgemeinschaft under grant number SCHM 808/5-1 is very gratefully acknowledged.

References

Adachi, S., "Simulated moving-bed chromatography for continuous separation of two components and its application to bioreactors.," *J. Chromatogr. A*, **658**, 271–282 (1994).

Altenhöhner, U., M. Meurer, J. Strube, and H. Schmidt-Traub, "Parameter estimation for the simulation of liquid chromatography," *J. Chromatogr. A*, **769**, 59–69 (1997).

Ando, M. and M. Tanimura, Method for controlling simulated moving bed system, US Pat. 4599115 (1986).

Benthabet, M., M. Bailly, and J. P. Corriou, Nonlinear control of a simulated moving bed, AIChE Annual Meeting, Sacramento (1997).

Broughton, D. B. and C. G. Gerhold, Continuous sorption processes employing fixed bed of sorbent and moving inlets and outlets, US Pat. 2985589 (1961).

Cansell, F., G. Hotier, P. Marteau, and N. Zanier, Method for regulating a process for the separation of isomers of aromatic hydrocarbons having from 8 to 10 carbon atoms, US Pat. 5569808 (1996).

Cohen, C., R. Jacob, G. B. du Columbier, and G. Hotier, Process for regulating at least one fluid flow circulating in a simulated moving bed chromatographic separation loop, US Pat. 5685992 (1997).

Couenne, N., P. Duchenne, G. Hotier, and D. Humeau, Method for controlling with precision a process for separating constituents of a mixture in a simulated moving bed separation system, US Pat. 5902486 (1999).

Croft, D. T. and M. D. Le Van, "Periodic states of adsorption cycles I. Direct determination and stability," *Chem. Eng. Sci.*, **49**, 1821–1829 (1994).

Daubechies, I., *Ten Lectures on Wavelets*. Society for industrial and applied mathematics, Philadelphia (1992).

Dünnebier, G. and K.-U. Klatt, Modeling of chromatographic separation processes using nonlinear wave theory, In Georgakis, C., editor, *Proc. of IFAC DYCOPS-5*, pages 521–526, New York. Elsevier Science (1998).

Dünnebier, G. and K.-U. Klatt, "Optimal operation of simulated moving bed chromatographic processes," *Comput. Chem. Eng.*, **23**, S189–S192 (1999).

Dünnebier, G. and K.-U. Klatt, "Modelling and simulation of nonlinear chromatographic separation processes: a comparison of different modelling approaches," *Chem. Eng. Sci.*, **55**, 373–380 (2000).

Dünnebier, G., I. Weirich, and K.-U. Klatt, "Computationally efficient dynamic modeling and simulation of simulated moving bed chromatographic processes with linear isotherms," *Chem. Eng. Sci.*, **53**, 2537–2546 (1998).

Dünnebier, G., J. Fricke, and K.-U. Klatt, "Optimal design and operation of simulated moving bed chromatographic reactors," *Ind. Eng. Chem. Res.*, **39**, 2290–2304 (2000).

Ganetsos, G. and P. M. Barker, *Preparative and Production Scale Chromatography*. Marcel Dekker, New York (1993).

Gu, T., *Mathematical modeling and scale up of liquid chromatography*. Springer, New York (1995).

Guiochon, G., S. G. Golshan-Shirazi, and A. M. Katti, *Fundamentals of Preparative and Nonlinear Chromatography*. Academic Press, Boston (1994).

Hashimoto, K., S. Adachi, and Y. Shirai, Development of new bioreactors of a simulated moving bed type, In Ganetsos, G. and P. E. Barker, editors, *Preparative and Production Scale Chromatography*. Marcel Dekker (1993).

Helfferich, F. and R. Whitley, "Non-linear waves in chromatography II: Wave interference and coherence in multicomponent systems," *J. Chromatogr. A*, **734**, 7–47 (1996).

Holt, R. E., Control process for simulated moving adsorbent bed separations, US Pat. 5457260 (1995).

Hotier, G. and R. M. Nicoud, Chromatographic simulated mobile bed separation process with dead volume correction using periodic desyncronization, US Pat. 5578215 (1996).

Hotier, G., Process for simulated moving bed separation with a constant recycle rate, US Pat. 5762806 (1998).

Isermann, R., *Identifikation dynamischer Systeme*. Springer, Heidelberg (1992).

Kaczmarski, K. and D. Antos, "Fast finite difference method for solving multicomponent adsorption-chromatography models," *Comput. Chem. Eng.*, **20**, 1271–1276 (1996).

Kaczmarski, K., M. Mazzotti, G. Storti, and M. Morbidelli, "Modeling fixed-bed adsorption columns through orthogonal collocation on moving finite elements," *Comput. Chem. Eng.*, **21**, 641–660 (1997).

Klatt, K.-U., G. Dünnebier, S. Engell, and F. Hanisch, "Model-based optimization and control of chromatographic processes," *Comput. Chem. Eng.*, **24**, 1119–1126 (2000).

Klatt, K.-U., "Modeling and efficient computational simulation of simulated moving bed chromatographic separation processes," *Chem.-Ing.-Tech.*, **71**, 555–566 (1999). (In German).

Kloppenburg, E. and E.-D. Gilles, "A new process management concept for chromatography with simulated countercurrent," *Chem.-Ing.-Tech.*, **70**, 1526–1529 (1998). (In German).

Kloppenburg, E. and E.-D. Gilles, "Automatic control of the simulated moving bed process for C8 aromatics separation using asymptotically exact input/output-linearization," *J. Proc. Cont.*, **9**, 41–50 (1999).

Kvamsdal, H. M. and T. Hertzberg, "Optimization of PSA systems—Studies on cyclic steady-state convergence," *Comput. Chem. Eng.*, **21**, 819–832 (1997).

Lapidus, L. and N. Amundsen, "Mathematics of adsorption in beds IV. The effect of longitudinal diffusion in ion exchange and chromatographic columns," *J. Phys. Chem*, **56**, 984–988 (1952).

Ljung, L., *MATLAB System identification toolbox handbook*. The Mathworks, Inc. (1995).

Lu, Z. P. and C. B. Ching, "Dynamics of simulated moving bed adsorptive separation processes," *Sep. Sci. Technol.*, **32**, 1993–2010 (1997).

Ma, Z. and G. Guiochon, "Application of orthogonal collocation on finite elements in the simulation of nonlinear chromatography," *Comput. Chem. Eng.*, **15**, 415–426 (1991).

Ma, Z. and N.-H. L. Wang, "Standing wave analysis of SMB chromatography: Linear systems," *AIChE J.*, **43**, 2488–2508 (1997).

Marteau, P. G. H., N. Zanier-Szydlowski, A. Aoufi, and F. Cansell, "Advanced control of C8 aromatics separation with real-time multipoint on-line Raman spectroscopy," *Process Control and Quality*, **6**, 133–140 (1994).

Mazzotti, M., G. Storti, and M. Morbidelli, "Optimal operation of simulated moving bed units for nonlinear chromatographic separations," *J. Chromatogr. A*, **769**, 3–24 (1997).

Migliorini, C., M. Mazzotti, and M. Morbidelli, "Continuous chromatographic separations through simulated moving bed under linear and nonlinear conditions," *J. Chromatogr. A*, **827**, 171–173 (1998).

Natarajan, S. and J. H. Lee, "Repetitive model predictive control applied to a simulated moving bed chromatography system," *Comput. Chem. Eng.*, **24**, 1127–1133 (2000).

Nicoud, R. M., "The simulated moving bed: A powerful chromatographic process," *Mag. Liquid Gas Chromatogr.*, **55**, 373–380 (1992).

Nilchan, S. and C. C. Pantelides, "On the optimisation of periodic adsorption processes," *Adsorption*, **4**, 113–148 (1998).

Pais, L. S., J. M. Loureiro, and A. E. Rodrigues, "Modeling strategies for enantiomeres separation by SMB chromatography," *AIChE J.*, **44**, 561–569 (1998).

Poulain, C. A. and B. A. Finlayson, "A comparison of numerical methods applied to nonlinear adsorption column," *Int. J. Numer. Methods Fluids*, **17**, 839–859 (1993).

Rhee, H., R. Aris, and N. Amundsen, *First oder partial differential equations, volume II: Theory and application of hyperbolic systems of quasilinear equations*. Prentice Hall, New York (1989).

Ruthven, D. M. and C. B. Ching, "Counter-current and simulated counter-current adsorption separation processes," *Chem. Eng. Sci.*, **44**, 1011–1038 (1989).

Spieker, A., E. Kloppenburg, and E.-D. Gilles, Computer modelling of chromatographic bioseparation, In Subramanian, G., editor, *Bioseparation and Bioprocessing*, pages 329–362. Wiley-VCH, Weinheim (1998).

Storti, G., M. Masi, R. Paludetto, M. Morbidelli, and S. Carra, "Adsorption separation processes: Countercurent and simulated countercurrent operations," *Comput. Chem. Eng.*, **12**, 475–482 (1988).

Storti, G., M. Mazzotti, M. Morbidelli, and S. Carra, "Robust design of binary countercurrent adsorption separation processes," *AIChE J.*, **39**, 471–492 (1993).

Strube, J. and H. Schmidt-Traub, "Dynamic simulation of simulated moving bed chromatographic processes," *Comput. Chem. Eng.*, **20**, S641–S646 (1996).

Strube, J., A. Jupke, A. Epping, H. Schmidt-Traub, M. Schulte, and M. Devant, "Design, optimization, and operation of SMB chromatography in the production of enantiomerically pure pharmaceuticals," *Chirality*, **11**, 440–450 (1999).

Tondeur, D., "Paradigms and paradoxes in modeling adsorption and chromatographic separations," *Ind. Eng. Chem. Res.*, **34**, 2782–2788 (1995).

Unger, J., G. Kolios, and G. Eigenberger, "On the efficient simulation and analysis of regenerative processes in cyclic operation," *Comput. Chem. Eng.*, **21**, S167–S172 (1997).

Van Deemter, J., F. Zuiderweg, and A. Klinkenberg, "Longitudinal Diffusion and resistance to mass transfer as causes of nonideality in chromatography," *Chem. Eng. Sci.*, **5**, 271–280 (1956).

Wu, D. J., Y. Xie, Z. Ma, and N.-H. L. Wang, "Design of simulated moving bed chromatography for amino acid separations," *Ind. Eng. Chem. Res.*, **37**, 4023–4035 (1998).

Zafiriou, E. and M. Morari, "Digital controllers for SISO systems: A review and a new algorithm," *Int. J. Control*, **42**, 855–876 (1985).

Zhong, G. and G. Guiochon, "Analytical solution for the linear ideal model of simulated moving bed chromatography," *Chem. Eng. Sci.*, **51**, 4307–4319 (1996).

Zhong, G. and G. Guiochon, "Fundamentals of simulated moving bed chromatography under linear conditions," *Adv. Chromatogr.*, **39**, 351–400 (1998).

Zimmer, G., G. Dünnebier, and K.-U. Klatt, On line parameter estimation and process monitoring of simulated moving bed chromatographic processes, In *Proceedings of the 1999 ECC, Karlsruhe, Germany* (1999).

Dynamic Optimization in the Batch Chemical Industry

D. Bonvin, B. Srinivasan and D. Ruppen
Institut d'Automatique, École Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
Computer-aided Process Engineering, Lonzagroup, CH-3930 Visp, Switzerland

Abstract

Dynamic optimization of batch processes has attracted more attention in recent years since, in the face of growing competition, it is a natural choice for reducing production costs, improving product quality, and meeting safety requirements and environmental regulations. Since the models currently available in industry are poor and carry a large amount of uncertainty, standard model-based optimization techniques are by and large ineffective, and the optimization methods need to rely more on measurements.

In this paper, various measurement-based optimization strategies reported in the literature are classified. A new framework is also presented, where important characteristics of the optimal solution that are invariant under uncertainty are identified and serve as references to a feedback control scheme. Thus, optimality is achieved by tracking, and no numerical optimization is required on-line. When only batch-end measurements are available, the proposed method leads naturally to an efficient batch-to-batch optimization scheme. The approach is illustrated via the simulation of a semi-batch reactor in the presence of uncertainty.

Keywords

Dynamic optimization, Optimal control, Batch chemical industry, On-line optimization, Batch-to-batch optimization, Run-to-run optimization

Introduction

Batch and semi-batch processes are of considerable importance in the chemical industry. A wide variety of specialty chemicals, pharmaceutical products, and certain types of polymers are manufactured in batch operations. Batch processes are typically used when the production volumes are low, when isolation is required for reasons of sterility or safety, and when frequent changeovers are necessary. With the recent trend in building small flexible plants that are close to the markets of consumption, there has been a renewed interest in batch processing (Macchietto, 1998).

From a process systems point of view, the key feature that differentiates continuous processes from batch and semi-batch processes is that the former have a steady state, whereas the latter are inherently time-varying in nature (Bonvin, 1998). This paper considers batch and semi-batch processes in the same manner and, thus herein, the term 'batch processes' includes semi-batch processes as well.

The operation of batch processes typically involves following recipes that have been developed in the laboratory. However, owing to differences in both equipment and scale, industrial production almost invariably necessitates modifications of these recipes in order to ensure productivity, safety, quality, and satisfaction of operational constraints (Wiederkehr, 1988). The 'educated trials' method that is often used for recipe adjustment is based on heuristics and results in conservative profiles. Conservatism is necessary here to guarantee feasibility despite process disturbances.

To shorten the time to market (by bypassing an elaborate scale-up process) and to reduce operational costs (by reducing the conservatism), an optimization approach is called for, especially one that can handle uncertainty explicitly. Operational decisions such as temperature or feed rate profiles are then determined from an optimization problem, where the objective is of economic nature and the various technical and operational constraints are considered explicitly. Furthermore, due to the repetitive nature of batch processes, these problems can also be addressed on a batch-to-batch basis.

The objectives of this paper are threefold: i) address the industrial practice prevailing in the batch specialty chemical industry and discuss the resulting optimization challenges, ii) review the dynamic optimization strategies available for batch processes, with an emphasis on measurement-based techniques, and iii) present a novel scheme that uses process measurements directly (i.e., without the often difficult step of model refinement) towards the goal of optimization. Accordingly, the paper has three major parts:

- *Industrial perspectives in batch processing:* The major operational challenges aim at speeding up process/product development, increasing the productivity, and satisfying safety and product quality requirements (Allgor et al., 1996). These tasks need to be performed in an environment characterized by a considerable amount of uncertainty and the presence of numerous operational and safety-related constraints. The measurements available could be used to help meet these challenges.

- *Optimization strategies for batch processes:* These are reviewed and classified according to: i) whether uncertainty is considered explicitly, ii) whether measurements are used, iii) whether a model is used to guide the optimization. The type of measurements

used for optimization (on-line, off-line) adds another dimension to the classification.

- *Invariant-based optimization scheme:* New insights into the optimal solution for a class of batch processes have led to an alternative way of dealing with uncertainty. It involves: i) the off-line characterization of the optimal solution using a simplified model, ii) the selection of signals that are derived from the conditions for optimality and are invariants to uncertainty, and iii) a model-free implementation by tracking these invariants using a limited number of measurements. This results in a model-free though measurement-based implementation that is quite robust towards uncertainty. An interesting feature of this framework is that it permits naturally to combine off-line data from previous batches with on-line data from the current batch.

The paper is organized as follows. The industrial perspectives in batch processing are presented first. The next section briefly reviews the optimization strategies available for batch processes and proposes a classification of the methods. The invariant-based optimization framework is then developed and an example is provided to illustrate the theoretical developments. Conclusions are drawn in the final section.

Industrial Perspectives in Batch Processing

It is difficult to address in generic terms the perspectives prevailing in the batch chemical industry since the processing environments and constraints differ considerably over the various activities (specialty chemicals, pharmaceuticals, agro and bio products, etc.). Thus, the situation specific to the production of intermediates in the specialty chemical industry will be emphasized in this section. The customer—typically an end-product manufacturer—often generates competition between several suppliers for the production of a new product. The suppliers need to investigate the synthesis route and design an appropriate production process. The competition forces the suppliers to come up, under considerable time pressure, with an attractive offer (price/kg) if they want to obtain the major share of the deal.

Batch processes are usually carried out in relatively standardized pieces of equipment whose operating conditions can be adjusted to accommodate a variety of products. The working environment that will be considered is that of multi-product plants. In a multi-product plant, a number of products are manufactured over a period of time, but at any given time, only one product is being made. The sequence of tasks to be carried out on each piece of equipment such as heating, cooling, reaction, distillation, crystallization, drying, etc. is pre-defined, and the equipment item in which each task is performed is also specified (Mauderli and Rippin, 1979).

Operational Objectives

The fundamental objective is of economic nature. The investment (in time, personnel, capital, etc.,) should pay off, as the invested capital has to compare favorably with other possible investments. This fundamental objective can in turn be expressed in terms of technical objectives and constraints, which are presented next.

- *Productivity:* This is the key word nowadays. However, high productivity requires stable production so as to reduce the amount of corrective manual operations that are costly in terms of production time and personnel. Reducing the time necessary for a given production is particularly interesting when the number of batches per shift can be increased. In multi-product plants, however, equipment constraints (bottlenecks) and logistic issues often limit productivity.
- *Product quality:* Quality is often impaired by the appearance of small amounts of undesired by-products. The presence of impurities (also due to recycled solvents) is very critical since it can turn an acceptable product into waste. Removing impurities is often not possible or can significantly reduce throughput. Also, from an operational, logistic and regulation point of view, it is often not possible to use blending operations in order to achieve the desired average quality.
  Reproducibility of final product composition despite disturbances and batch-to-batch variations is important when the process has to work closely to some quality limit (for example, when the quality limits are tight). Improving the selectivity of an already efficient process is often not seen as a critical factor. However, when the separation of an undesirable by-product is difficult, the selectivity objective may be quite important.
- *Safety aspects:* The safety aspects such as the avoidance of runaways are of course very important. Safety requirements can lead to highly conservative operation. Here, the real obstacle is the lack of on-line information. If information about the state of the process were available, the process engineer would know how to guarantee safety or react in the case of a latent problem. Thus, the difficulty results from a measurement limitation and not from a lack of operational knowledge.
- *Time-to-market:* The economic performance is strongly tied to the speed at which a new product/process can be developed. The product lifetime of specialty chemicals is typically shorter than for bulk chemicals. Since the production in campaigns reduces the time to learn, it is necessary to learn quickly and improve the productivity right away.

After a couple of years, a profitable new product may become a commodity (of much lesser value), for which the development of a second-generation process is often considered.
Nowadays, there is a trend in the specialty chemical industry to skip pilot plant investigations unless the process is difficult to scale up. The situation is somewhat different in pharmaceuticals production, where pilot plant investigations are systematically used since they also serve to produce the small 'first amounts' needed.

Industrial Practice

Though the problem of meeting the aforementioned objectives could be solved effectively as an optimization problem, there have been only a few attempts in industry to optimize operations through mathematical modeling and optimization techniques. The recipes are developed in the laboratory in such a way that they can be implemented safely in production. The operators then use heuristics gained from experience to adjust the process periodically (whenever this is allowed), which leads to slight improvements from batch to batch (Verwater-Lukszo, 1998). The stumbling blocks for the use of mathematical modeling and optimization techniques in industrial practice have been partly organizational and partly technical.

Organizational Issues. At the organizational level, the issues are as follows:

- *Registration:* Producers of active compounds in the food and pharmaceuticals areas have to pass through the process of registration with the Food and Drug Administration. Since this is a costly and time-consuming task, it is performed simultaneously with R&D for a new production process. Thus, the main operational parameters are fixed within specified limits at an early stage of the development. Since the specifications provided by the international standards of operation (GMP) are quite tight, there is very little room for maneuver left. It is important to stress that the registration is tied to both product *and* process.
- *Multi-step process:* In the R&D phase of a large multi-step process, different teams work on different processing steps. Often, each team tries to optimize its process subpart, thereby introducing a certain level of conservatism to account for uncertainty. Consequently, the resulting process is the sum of conservatively designed subparts, which often does not correspond to the optimum of the global process!
- *Role of control and mathematical optimization:* In many projects, control is still considered to be a standard task that has to be performed during the detailed engineering phase and not a part of the design phase of the process. It is like 'painting' a controller or an optimizer once the process has been built. At this late stage, there is so much conservatism and robustness in the system that it does not require a sophisticated control strategy. However, the performance may still be far from being optimal.

All these organizational problems can be resolved by resorting to 'global thinking'. It has become a challenge for both project leaders and plant managers to make chemists and engineers think and act in a global way. It is done through fostering interdisciplinary teamwork and simultaneous rather than sequential work for process research, development and production (R&D&P). The objective is a globally optimal process and not simply the juxtaposition of robust process subparts. Team work amounts to having R&D&P solutions worked out simultaneously by interdisciplinary teams consisting of a project leader, chemists, process engineers, production personnel and specialists for analytics, simulation, statistics, etc.

Technical Issues. The main technical issues relate to modeling and measurements, the presence of both uncertainty and constraints, and the proper use of the available degrees of freedom for process improvement. These are addressed next.

- *Modeling:* In the specialty chemical industry, molecules are typically more complex than in the commodity industry, which often results in complex reaction pathways. Thus, it is illusory to expect constructing detailed kinetic models. The development of such models may exceed one man-year, which is incompatible with the objectives of batch processing. So, what is often sought in batch processing, is simply the ability to predict the batch outcome from knowledge of its initial phase.
Modern software tools such as Aspen Plus, PRO/II, or gPROMs have found wide application to model continuous chemical processes (Marquardt, 1996; Pantelides and Britt, 1994). The situation is somewhat different in the batch specialty chemistry. Though batch-specific packages such as Batch Plus, BATCHFRAC, CHEMCAD, BatchCAD, or BaSYS are available, they are not generally applicable. Especially the two important unit operations, reaction and crystallization, still represent a considerable challenge to model at the industrial level.
For batch processes, modeling is often done empirically using input/output static models on the basis of statistical experimental designs. These include operational variables specified at the beginning of the batch and quality variables measured at the end

of the batch. Time-dependent variables are not considered beyond visual comparison of measured profiles. Sometimes the model is a set of simple linguistic rules based on experience, e.g. when 'low' then 'bad'. Occasionally, the model consists of a simple energy balance, or the main dynamics are expressed via a few ordinary differential equations. The modeling objective is not accuracy but rather the ability to semi-quantitatively describe the major tradeoffs present in the process such as the common one between quality and productivity in many transformation and separation processes. For example, an increase in reflux ratio improves distillate purity but reduces distillate flow rate; or a temperature increase can improve the yield at the expense of selectivity in a chemical reaction system.

- *Measurements:* Quality measurements are typically available at the end of the batch *via*, for example, off-line chromatographic methods (GC, HPLC, DC, IC). In addition, physical measurements such as temperature, flow, pressure, or pH may be available on-line during the course of the batch. However, they are rather unspecific with respect to the key variables (concentrations) of the chemical process. Other on-line measurements such as conductivity, viscosity, refractive index, torque, spectroscopy, and calorimetry are readily available in the laboratory, but rarely in production. Pseudo on-line GC and HPLC are less effective in batch processing than with continuous processes due to relatively longer measurement delays.
  On-line spectroscopy (FTIR, NIR, Raman) has opened up new possibilities for monitoring chemical processes (McLennan and Kowalski, 1995; Nichols, 1988). These techniques rely on multivariate calibration for accurate results, i.e., the spectral measurements need to be calibrated with respect to known samples containing all the absorbing species. Though on-line spectroscopy is getting more common in the laboratory, the transfer of many measurement systems from the laboratory to the plant is still a real challenge. For example, many processing steps deal with suspensions that lead to plugging and deposition problems. Even if these problems can be handled at the laboratory scale, they still represent formidable challenges at the production level. There is presently a strong push to develop and validate measurement techniques that can work equally well throughout the three levels of Research, Development and Production.
  When quality measurements are not directly available, state estimation (or soft sensing) is typically utilized. However, physical on-line measurements are often too unspecific for on-line state estimation in batch processes. Current practice indicates that there are very few applications of state estimation in the specialty chemistry. However, state estimation works well in fermentation processes due to the availability of additional physical measurements and the possibility to reconstruct concentrations without the use of kinetic models (Bastin and Dochain, 1990).
- *Uncertainty:* Uncertainty is widely present in the operation of batch processes. Firstly, it enters in the reactant quality (changes in feedstock), which is the main source of batch-to-batch variations. Secondly, uncertainty comes in the form of modeling errors (errors in model structure and parameters). These modeling errors can be fairly large since, according to the philosophy of batch processing, little time is available for the modeling task. Thirdly, process disturbances and measurement noise contribute to the uncertainty in process evolution (e.g. undetected failure of dosing systems; change in the 'quality' of utilities such as brine temperature, or of manual operations such as solid charge).
  Recipe modifications from one batch to the next to tackle uncertainty are common in the specialty chemical industry, but less so for the exclusive syntheses in agro and pharmaceuticals production. Uncertainty is typically handled through:
  - The choice of conservative operation such as extended reaction time, lower feed rate or temperature, the use of a slightly over-stoichiometric mixture in order to force the reaction to fully consume one reactant (Robust mode).
  - Feed stock analyses leading to appropriate adjustments of the recipe (Feedforward mode). Adjustment is usually done by scaling linearly certain variables such as the final time or the dilution, more rarely the feed rate or the temperature.
  - Rigorous quality checks through off-line analyses, or the use of standard measurements such as the temperature difference between jacket and reactor, leading to appropriate correction of the recipe (Feedback mode). For example, a terminal constraint can be met by successive addition of small quantities of feed towards the end of a batch to bring the reaction to the desired degree of completion (Meadows and Rawlings, 1991).

  The problem of scale-up can also be viewed as one of (model) uncertainty. The data available from laboratory studies do not quite extrapolate to the production level. Thus, when the strategies developed in the laboratory are used at the production level,

they do carry a fair amount of uncertainty. Furthermore, the pressure to reduce costs and to speed up process development calls for large scale-ups with a considerable amount of extrapolation. As a result, the proposed strategies can be rather conservative.

- *Constraints:* Industrial processing is naturally characterized by soft and hard constraints related to equipment and operational limitations and to safety aspects. In batch processing, there is the additional effect of terminal constraints (selectivity in reaction systems, purity in separation systems, admissible levels of impurities, etc.). Furthermore, in multiproduct batch production, the process has to fit in an existing plant. Thus, ensuring feasible operation comes before the issue of optimality, and process designers normally introduce sufficient conservatism in their design so as to guarantee feasibility even in the worst of conditions.
  The need to improve performance calls for a reduction of the conservatism that is introduced to handle uncertainty. Performance improvement can be obtained by operating closer to constraints, which can be achieved by measuring/estimating the process state with respect to these constraints. Riding along an operational constraint is often done when the constrained variable is directly implemented (such as maximum feed rate) or can be measured (such as a temperature).
- *Time-varying decisions:* Traditionally, chemists in the laboratories and operators in the plants were used to thinking in terms of constant values (experimental planning results in static maps between design variables and process performance). New sensors and increasing computing power (e.g. spectroscopic measurements, modern PLC systems) make on-line time-varying decisions possible. Along with these new time-dependent insights, the chemists in the laboratory start to vary process inputs as functions of time. The potential benefit of these additional degrees of freedom is paramount to using optimal control techniques. There are situations where variable input profiles can be of direct interest:
  - There may be a significant theoretical advantage of using a variable profile over the best constant profile (Rippin, 1983). The performance improvement can sometimes be considerable. In batch crystallization, for example, gains of up to 500% can be obtained by adjusting the temperature, the removal of solvent or the addition of a precipitation solvent as functions of time. Large gains are also possible in reactive semi-batch distillation.
  - It is more and more common to adjust the feed rate in semi-batch reactors so as to force the heat generation to match the cooling capacity of the jacket.

Industrial situation	Implications for optimization
Need to improve performance	Use optimization for computing time-dependent decisions
Absence of a reliable model	Use measurements for implementing optimal inputs
Few on-line measurements	Use off-line measurements in a batch-to-batch optimization scheme
Presence of uncertainty	Identify and track signals that are invariant to uncertainty
Operational and safety constraints	Track constraints so as to reduce conservatism

Table 1: Implications of the industrial situation regarding the choice of an appropriate optimization approach.

An interesting feature of batch processing is the fact that batch processes are repeated over time. Thus, the operation of the current batch can be improved by using the off-line measurements available from previous batches. The objective is then to get to the optimum over as few batches as possible. Also, with the tendency to skip pilot plant investigations whenever possible, this type of process improvement is of considerable interest for the initial batches of a new production campaign.

Implications for Optimization

The industrial situation, as far as technical issues are concerned, can be summarized as follows:

- There is an immediate need to improve the performance of batch processes.
- Models are poor, incomplete or nonexistent.
- On-line measurements are rare, and state estimation is difficult; however, off-line measurements can be made available if needed.
- There is considerable uncertainty (model inaccuracies, variations in feedstock, process disturbances).
- Several operational and safety constraints need to be met.

The implications of the current industrial situation regarding the choice of an appropriate optimization approach are presented in Table 1. The details will be clarified in the forthcoming sections. The main conclusion is that a framework that uses (preferably off-line) measurements rather than a model of the process for implementing the optimal inputs is indeed required.

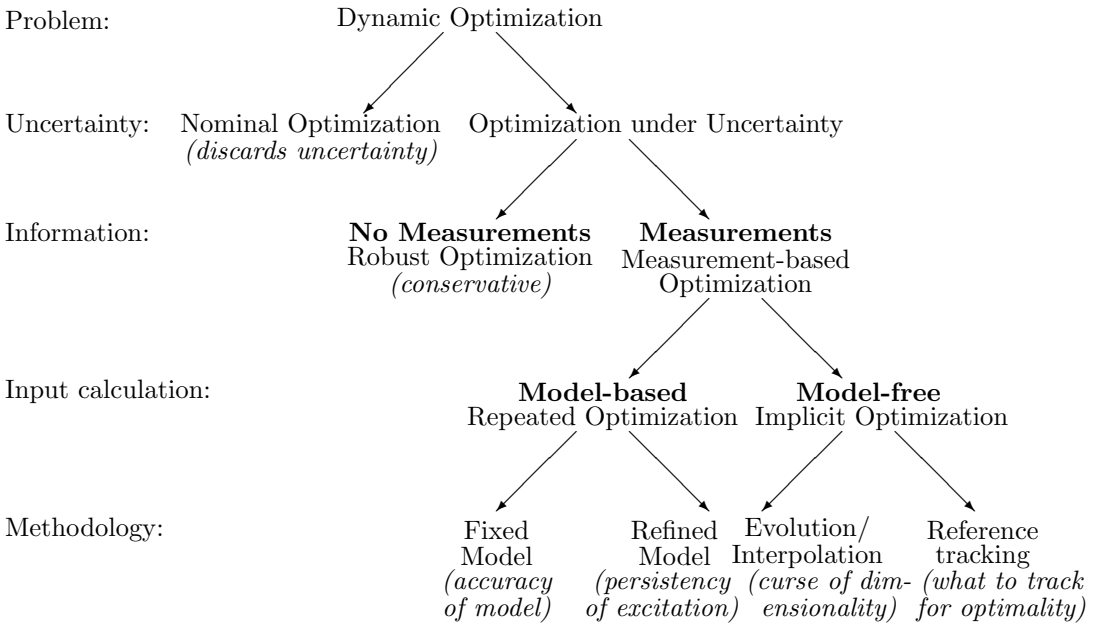


Figure 1: Dynamic optimization scenarios with, in parentheses, the corresponding major disadvantage.

Overview of Batch Process Optimization

The optimization of batch processes typically involves both dynamic and static constraints and falls under the class of *dynamic optimization*. Possible scenarios in dynamic optimization are depicted in Figure 1. The first level of classification depends on whether or not uncertainty (e.g., variations in initial conditions, unknown model parameters, or process disturbances) is considered. The standard approach is to discard uncertainty, leading to a nominal solution that may not even be feasible, let alone optimal, in the presence of uncertainty.

The second level concerns the type of information that can be used to combat uncertainty. If measurements are not available, a conservative stand is required. In contrast, conservatism can be reduced with the use of measurements.

In the next levels, the classification is based on how the measurements are used in order to guide the optimization. The calculation of inputs can be either model-based or model-free. In the model-based case, the type of model that is used (fixed or refined) affects the resulting methodology. If the input calculation is model-free, the current measurement is either compared to a reference or used for interpolation between pre-computed optimal values. The different scenarios are discussed in detail next.

Nominal Optimization

In nominal optimization, the uncertainty is simply discarded. A typical batch optimization problem consists of achieving a desired product quality at the most economical cost, or maximizing the product yield for a given batch time. The optimization can be stated mathematically as follows:

$$\min_{u(t)} \; J = \phi(x(t_f)) \tag{1}$$

subject to

$$\dot{x} = F(x,u), \qquad x(0) = x_0 \tag{2}$$

$$S(x,u) \leq 0, \qquad T(x(t_f)) \leq 0 \tag{3}$$

where J is the scalar performance index to be minimized, x the n-vector of states with known initial conditions x_0, u the m-vector of inputs, F a vector field describing the dynamics of the system, S the ζ-vector of path constraints (which include state constraints and input bounds), T the τ-vector of terminal constraints, ϕ a smooth scalar function representing the terminal cost, and t_f the final time.

The problem (1)–(3) is quite general. Even when an integral cost needs to be considered, i.e., $J = \bar{\phi}(x(t_f)) + \int_0^{t_f} L(x,u)dt$, with L a smooth function, it can be converted into the form (1)–(3) by the introduction of an additional state: $\dot{x}_{cost} = L(x,u)$, $x_{cost}(0) = 0$, and $J = \bar{\phi}(x(t_f)) + x_{cost}(t_f) = \phi(x(t_f))$. Let J^* be the optimal solution to (1)–(3). It is interesting to note that the minimum time problem with the additional constraint $\phi(x(t_f)) \leq J^*$, i.e,

$$\min_{t_f, u(t)} \; t_f \tag{4}$$

subject to

$$\dot{x} = F(x,u), \qquad x(0) = x_0 \tag{5}$$

$$S(x,u) \leq 0, \qquad T(x(t_f)) \leq 0 \tag{6}$$

$$\phi(x(t_f)) \leq J^* \tag{7}$$

will lead to exactly the same optimal inputs as (1)–(3), though the numerical conditioning of the two problems, (1)–(3) and (4)–(7), may differ considerably. The equivalence of solutions is verified using the necessary conditions of optimality. So, without loss of generality, the final time will be assumed fixed in this paper.

Application of Pontryagin's Maximum Principle (PMP) to (1)–(3) results in the following Hamiltonian (Bryson and Ho, 1975; Kirk, 1970):

$$H = \lambda^T\, F(x,u) + \mu^T\, S(x,u) \tag{8}$$

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x}, \quad \lambda^T(t_f) = \left.\frac{\partial \phi}{\partial x}\right|_{t_f} + \nu^T \left.\left(\frac{\partial T}{\partial x}\right)\right|_{t_f} \tag{9}$$

where $\lambda(t) \neq 0$ is the n-vector of adjoint states (Lagrange multipliers for the system equations), $\mu(t) \geq 0$ the ζ-vector of Lagrange multipliers for the path constraints, and $\nu \geq 0$ the τ-vector of Lagrange multipliers for the terminal constraints. The Lagrange multipliers μ and ν are nonzero when the corresponding constraints are active and zero otherwise so that $\mu^T S(x,u) = 0$ and $\nu^T T(x(t_f)) = 0$ always. The first-order necessary condition for optimality of the input u_i is:

$$H_{u_i} = \frac{\partial H}{\partial u_i} = \lambda^T \frac{\partial F}{\partial u_i} + \mu^T \frac{\partial S}{\partial u_i} = \lambda^T F_{u_i} + \mu^T S_{u_i} = 0. \tag{10}$$

Note that, in this paper, PMP will not be used to determine the optimal solution numerically since this procedure is well known to be ill-conditioned (Bryson, 1999). However, PMP will be used to decipher the characteristics of the optimal solution.

Robust Optimization

In the presence of uncertainty, the classical *open-loop* implementation of off-line calculated nominal optimal inputs may not lead to optimal performance. Moreover, constraint satisfaction, which becomes important in the presence of safety constraints, may not be guaranteed unless a conservative strategy is adopted. In general, it can be assumed that the model structure is known and the uncertainty concentrated in the model parameters and disturbances. Thus, in the uncertain scenario, the optimization can be formulated as follows:

$$\min_{u(t)} \; J = \phi(x(t_f)) \tag{11}$$

subject to

$$\dot{x} = F(x,\theta,u) + d, \qquad x(0) = x_0$$
$$S(x,u) \leq 0, \qquad T(x(t_f)) \leq 0$$

where θ is the vector of uncertain parameters, and $d(t)$ the unknown disturbance vector. In addition, the initial conditions x_0 could also be uncertain.

To solve this optimization problem, an approach referred to as *robust optimization*, where the uncertainty is taken into account explicitly, is proposed in the literature (Terwiesch et al., 1994). The uncertainty is dealt with by considering all (several) possible values for the uncertain parameters. The optimization is performed either by considering the worst-case scenario or in an expected sense. The solution obtained is conservative but corresponds to the *best* possibility when measurements are not used.

Measurement-based Optimization (MBO)

The conservatism described in the subsection above can be considerably reduced with the use of measurements, thereby leading to a better cost. Consider the optimization problem in the presence of uncertainty and measurements as described below:

$$\min_{u^k_{[t_l,t_f]}} \; J^k = \phi(x^k(t_f)) \tag{12}$$

subject to

$$\dot{x}^k = F(x^k,\theta,u^k) + d^k, \qquad x^k(0) = x^k_0$$
$$y^k = h(x^k,\theta) + v^k$$
$$S(x^k,u) \leq 0, \qquad T(x^k(t_f)) \leq 0$$

given

$$y^j(i),\ i = \{1,\cdots,N\}$$
$$\forall\, j = \{1,\cdots,k-1\},\ \text{and}$$
$$i = \{1,\cdots,l\}\ \text{for}\ j = k.$$

where $x^k(t)$ is the state vector, $u^k(t)$ the input vector, $d^k(t)$ the process disturbance, $v^k(t)$ the measurement noise, and J^k the cost function for the k^{th} batch. Let $y = h(x,\theta)$, a p-dimensional vector, be the combination of states that can be measured, $y^j(i)$ the i^{th} measurement taken during the j^{th} batch, and N the number of measurements within a batch. The objective is to utilize the measurements from the previous $(k-1)$ batches and the measurements up to the current time, t_l, of the k^{th} batch in order to tackle the uncertainty in θ and determine the optimal input policy for the remaining time interval $[t_l, t_f]$ of the k^{th} batch.

Role of the Model in the Calculation of the Inputs. Among the measurement-based optimization schemes, a classification can be done according to whether or not a model is used to guide the optimization.

Model-based input calculation: Repeated optimization. In optimization, the model of the system can be used to predict the evolution of the system, compute the cost sensitivity with respect to input variations so as to obtain search directions, and update the inputs towards the optimum. Measurements are then used to estimate the current states and parameters. As the estimation

and optimization tasks are typically repeated over time, this scheme is often referred to as repeated optimization. The model is either fixed or refined using measurements, the advantages and disadvantages of which are discussed next.

- *Fixed model*: If the model is not adjusted, it needs to be fairly accurate. This, however, is against the philosophy of the approach that assumes the presence of (considerable) uncertainty. If the uncertainty is only in the form of disturbances and not in the model parameters, it might be sufficient to use a fixed model. On the other hand, if the model is not accurate enough, the methodology will have difficulty converging to the optimal solution. Note that, since the measurements are used to estimate the states only (and not the parameters), there is no need for persistent inputs.
- *Refined model*: When model refinement is used, the need to start with an accurate model is alleviated, but it is necessary to excite appropriately the system for estimating the uncertain parameters. However, the optimal inputs may not provide sufficient excitation. On the other hand, if sufficiently exciting inputs are provided for parameter identification, the resulting solution may not be optimal. This leads to a conflict between the objectives of parameter estimation and optimization. This conflict has been studied in the adaptive control literature under the label *dual control problem* (Roberts and Williams, 1981; Wittenmark, 1995).

Model-free input calculation: Implicit optimization. In this approach, measurements are used directly to update the inputs towards the optimum, i.e., without using a model and explicit numerical optimization. However, a model might be used *a priori* to characterize the optimal solution. The classification here is based on whether the measurement is used for interpolation between precomputed optimal values or simply compared to a reference.

- *Evolution/interpolation*: The inputs are computed from past data and current measurements. If only batch-end measurements are used, the difference in cost between successive experimental runs can be used to provide the update direction for the inputs (evolutionary programming). The on-line version of this approach is based on the feedback optimal solution—the solution to the Hamilton-Jacobi-Bellman equation (Kirk, 1970)—being stored in one form or the other (e.g., neural network, look-up table, or dynamic programming).
  The main drawback of this approach is the curse of dimensionality. A large number of experimental runs are needed to converge to the optimum if only batch-end measurements are used. The use of the method with on-line measurements requires either a computationally expensive look-up table or a closed-form feedback law, which is analytically expensive or impossible to obtain in many cases.
- *Reference tracking*: The inputs are computed using feedback controllers that track appropriate references. The main question here is what references should be tracked. In most of the studies found in the literature, the references correspond to optimal trajectories computed off-line using a nominal model. Such an approach, however, is not guaranteed to be optimal in the presence of uncertainty. As will be explained later, this paper uses the concept of invariants to chose references, the tracking of which implies optimality.

Type of Measurements. The type of measurements (off-line or on-line) can add another level to the classification of optimization strategies.

- *Off-line measurements:* Off-line measurements include measurements taken at the end of the batch (batch-end measurements) and, possibly, off-line analysis of samples taken during the batch. These measurements cannot be used to improve the current batch but only subsequent ones. Off-line measurements are most common in industrial practice. They enable the set-up of a *batch-to-batch optimization* approach to account for parametric uncertainty by exploiting the fact that batch processes are typically repeated. Process knowledge obtained from previous batches is used to update the operating strategy of the current batch. This approach requires solving an optimization problem at the beginning of each batch. The objective is then to get to the optimum over a few batches.
- *On-line measurements:* When information is available during the course of the batch, an *on-line optimization* approach can be used. Measurements are used to compensate for uncertainty both within the batch and, possibly, also in a batch-to-batch manner. With this compensation, the variability caused by uncertainty is reduced to a large extent. Thus, it is possible to guarantee feasibility with smaller conservative margins which, in turn, helps improve the cost.

MBO *vs.* **MPC.** Model-predictive control (MPC), which has been well studied in the literature (see Rawlings et al. (1994); Qin and Badgwell (1997); Morari and Lee (1999) for surveys), has both some overlap and differences with MBO schemes that form the subject of this paper. MPC typically uses the repeated optimization approach to solve a control problem in an optimal manner. The major points that distinguish MBO from MPC are discussed next.

- *Goal and cost function:* The goal in MPC is control—choose the inputs to track given references—whereas the goal in MBO is optimization—maximize the yield of a product, minimize time for a given productivity, etc. In MPC, the control problem is formulated as an optimization with the cost function reflecting the quality of control, which typically is quadratic in nature, i.e., $J = \int_0^{t_f} \left(x^T Q\, x + u^T R\, u\right)\, dt$. In contrast, the cost function in MBO reflects the economic objective to optimize. However, once the optimization problem is formulated, similar tools are used for solution.
- *Continuous vs. batch processes:* MPC is oriented principally towards continuous processes. Stability is the main issue there and has been studied extensively in the MPC literature (Mayne et al., 2000). In contrast, MBO is oriented towards batch processes with a finite terminal time. Stability does not play a crucial role, and there is even a tendency to destabilize the system towards the end for the sake of optimality (the so-called batch kick). The important issues in MBO are feasibility and feedback optimality—how optimal is the operation in the presence of constraints and uncertainty. In contrast to MPC, MBO schemes can take advantage of the fact that batches are typically repeated. Run-to-run and implicit optimization schemes are thus particular to the MBO literature.
- *Role of constraints:* MBO typically has solutions that are on the constraints since the potential of optimization arises mainly from the presence of path and terminal constraints. Thus, it is important to go as close to the constraints as possible and, at the same time, guarantee feasibility. In contrast, though MPC has been designed to handle constraints, the typical problems considered in the framework of MPC try to avoid the solution being on the constraints by introducing a compromise between tracking performance and input effort.

Certain MBO schemes in the category of repeated optimization have been referred to as MPC schemes in the literature (Eaton and Rawlings, 1990; Meadows and Rawlings, 1991; Helbig et al., 1998; Lakshmanan and Arkun, 1999) and, thus, fall in the grey area between Batch MPC (Lee et al., 1999; Chin et al., 2000) and MBO for batch processes. It might be interesting to note that the search for optimality *via* tracking has also been studied for continuous processes. The terms "self-optimizing control", "feedback control", or "constraint control" are often used. The reader is referred to (Skogestad, 2000) for an overview of the work done in this area.

Classification of Measurement-based Optimization Methods. Only MBO methods (as opposed to MPC methods) that have been designed to deal explicitly with uncertainty in batch processing are considered in the classification. Table 2 illustrates the interplay between the type of measurements (off-line vs. on-line) and the role played by the model (model-based vs. model-free adaptation). Representative studies available in the literature are placed in the table.

Invariant-based Optimization

The idea of Invariant-Based Optimization (IBO) is to identify those important characteristics of the optimal solution that are invariant under uncertainty and provide them as references to a feedback control scheme. Thus, optimality is achieved by tracking these references without repeating numerical optimization. Also, the fact that batches are typically repeated over time can be used advantageously, thereby providing the possibility of on-line and/or batch-to-batch implementation. The methodology consists of:

1. a parsimonious state-dependent parameterization of the inputs,
2. the choice of signals that are invariant under uncertainty, and
3. the tracking of invariants using measurements.

These three steps are discussed in the following subsections.

Piecewise Analytic Characterization of the Optimal Solution

The parsimonious state-dependent parameterization arises from intrinsic characteristics of the optimal solution. The optimal solution is seen to possess the following properties (Visser et al., 2000):

- The inputs are in general discontinuous, but are analytic in between discontinuities. The time at which an input switches from one interval to another is called a *switching time*.
- Two types of arcs (constraint-seeking and compromise-seeking) are possible between switching instants. In a constraint-seeking arc, the input is determined by a path constraint, while in the other type of interval, the input lies in the interior of the feasible region (Palanki et al., 1993). The set of possible arcs is generically labelled $\eta(t)$.

The structure of the optimal solution is described by the type and sequence of arcs which can be obtained in three ways:

- educated guess by an experienced operator,
- piecewise analytical expressions for the optimal inputs,
- inspection of the solution obtained from numerical optimization.

Methodology	Batch-to-batch optimization (Off-line measurements)	On-line optimization (On-line measurements)
Model-based Fixed model	Zafiriou and Zhu (1990) Zafiriou et al. (1995) Dong et al. (1996)	Meadows and Rawlings (1991) Agarwal (1997) Abel et al. (2000)
Model-based Refined model	Filippi-Bossy et al. (1989) Marchal-Brassely et al. (1992) Rastogi et al. (1992) Fotopoulos et al. (1994) Le Lann et al. (1998) Ge et al. (2000) Martinez (2000)	Eaton and Rawlings (1990) Ruppen et al. (1998) Gattu and Zafiriou (1999) Noda et al. (2000) Lee et al. (2000)
Model-free Evolution Interpolation	Clarke-Pringle and MacGregor (1998)	Tsen et al. (1996) Rahman and Palanki (1996) Yabuki and MacGregor (1997) Krothapally et al. (1999) Schenker and Agarwal (2000)
Model-free Reference tracking	Scheid et al. (1999) Srinivasan et al. (2001)	Soroush and Kravaris (1992) Terwiesch and Agarwal (1994) Van Impe and Bastin (1995) de Buruaga et al. (1997) Ubrich et al. (1999) Fournier et al. (1999) Gentric et al. (1999) Lakshmanan and Arkun (1999) Visser et al. (2000)

Table 2: MBO methods specifically designed to compensate uncertainty.

In the most common third case, a simplified model of the process is used to compute a numerical solution in which the various arcs need to be identified. The exercise of obtaining the analytical expressions for the optimal inputs is undertaken only if the numerical solution cannot be interpreted easily. This analysis is, in general, quite involved and is only intended to provide insight into what constitutes the optimal solution rather than to implement the optimal solution. This analysis is discussed next.

Constraint-seeking *vs.* compromise-seeking arcs.

Constraint-seeking arc for u_i (Bryson and Ho, 1975). In this case, the input u_i is determined by an active constraint, say, $S_j(x,u) = 0$. Thus, $\mu_j \neq 0$. If $S_j(x,u)$ depends explicitly on u_i (e.g., in the case of input bounds), the computation of the optimal u_i is immediate. Otherwise, since $S_j(x,u) = 0$ over the entire interval under consideration, its time derivatives are also zero, $\frac{d^k}{dt^k} S_j(x,u) = 0$, for all k. Note that the time differentiation of $S_j(x,u)$ contains $\dot{x}$, i.e., the dynamics of the system. $S_j(x,u)$ can be differentiated with respect to time until u_i appears. The optimal input is computed from that time derivative of $S_j(x,u)$ where u_i appears. The computed input u_i is typically a function of the states of the system, thus providing a feedback law.

Compromise-seeking arc for u_i (Palanki et al., 1993). In this arc, none of the path constraints pertaining to the input u_i is active. The input is then determined from the necessary condition of optimality, i.e., $H_{u_i} = \lambda^T F_{u_i} = 0$. If F_{u_i} depends explicitly on u_i, the computation of the optimal u_i is immediate. Otherwise, since $H_{u_i} = 0$ over the entire interval, the time derivatives of H_{u_i} are also zero, $\frac{d^k}{dt^k} H_{u_i} = 0$, for all k. H_{u_i} can be differentiated with respect to time until u_i appears, from which the optimal input is computed. The computed input is a function of the states and might possibly also depend on the adjoint variables. If u_i does not appear at all in the time differentiations of H_{u_i}, then either no compromise-seeking arcs exist or the optimal input u_i is non-unique (Baumann, 1998).

The fact that the optimal solution is in the interior of the feasible region is the mathematical representation of the physical tradeoffs present in the system and affecting the cost. If there is no intrinsic tradeoff, the input u_i does not appear in the successive time differentiations of H_{u_i}. This forms an important subclass for practical applications. It guarantees that the optimal solution is always on the path constraints. This is the case in controllable linear systems, feedback-linearizable

systems, flat systems, and involutive-accessible systems, a category which encompasses many practical systems (Palanki et al., 1993; Benthack, 1997).

Constraint-seeking *vs.* compromise-seeking parameters.

Parsimonious input parameterization. The pieces described above are sequenced in an appropriate manner to obtain the optimal solution. The switching times and, possibly, a few variables that approximate the compromise-seeking arcs completely parameterize the inputs. The decision variables of the parameterization are labelled π. In comparison with piecewise constant or piecewise polynomial approximations, the parameterization proposed is exact and parsimonious.

Optimal choice of π. Once the inputs have been parameterized as $u(\pi, x, t)$, the optimization problem (1)–(3) can be written as:

$$\min_{\pi} \; J = \phi(x(t_f)) \tag{13}$$

subject to

$$\dot{x} = F(x, u(\pi, x, t)), \qquad x(0) = x_0 \tag{14}$$

$$T(x(t_f)) \leq 0 \tag{15}$$

Some of the parameters in π are determined by active terminal constraints (termed the *constraint-seeking parameters*) and some from sensitivities (termed the *compromise-seeking parameters*). Note the similarity with the classification of arcs for input u_i. Without loss of generality, let all τ terminal constraints be active at the optimum. Consequently, the number of decision variables arising from the parsimonious parameterization, n_π, needs to be larger than or equal to τ in order to satisfy all terminal constraints.

The idea is then to separate those variations in π that keep the terminal constraints active from those that do not affect the terminal constraints. For this, a transformation $\pi^T \rightarrow [\bar{\pi}^T \; \tilde{\pi}^T]$ is sought such that $\bar{\pi}$ is a τ-vector and $\tilde{\pi}$ a $(n_\pi - \tau)$-vector with $\frac{\partial T}{\partial \tilde{\pi}} = 0$. A linear transformation which satisfies these properties can always be found in the neighborhood of the optimum. Then, the necessary conditions for optimality of (13)–(15) are:

$$T = 0, \quad \frac{\partial \phi}{\partial \bar{\pi}} + \nu^T \frac{\partial T}{\partial \bar{\pi}} = 0, \quad \text{and} \quad \frac{\partial \phi}{\partial \tilde{\pi}} = 0. \tag{16}$$

The active constraints $T = 0$ determine the τ decision variables $\bar{\pi}$ while $\tilde{\pi}$ are determined from the sensitivities $\frac{\partial \phi}{\partial \tilde{\pi}} = 0$. Thus, $\bar{\pi}$ corresponds to the constraint-seeking parameters and $\tilde{\pi}$ to the compromise-seeking parameters. The Lagrange multipliers ν are calculated from $\frac{\partial \phi}{\partial \bar{\pi}} + \nu^T \frac{\partial T}{\partial \bar{\pi}} = 0$.

Signals Invariant under Uncertainty

In the presence of uncertainty, the numerical values of the optimal input u_i in the various arcs and the input parameters π might change considerably. However, what remains invariant under uncertainty is the fact that the necessary condition $H_{u_i} = 0$ has to be verified. $H_{u_i} = 0$ takes on different expressions for constraint-seeking and compromise-seeking arcs. To ease the development, it is assumed that the uncertainty is such that it does not affect the type and sequence of arcs nor the set of active terminal constraints.

Choice of invariants.

Choice of invariants for constraint-seeking and compromise-seeking arcs. A set of signals $I_i^\eta(t) = h_i^\eta(x(t), u(t), t)$, referred to as *invariants* for arcs, will be chosen such that optimality is achieved by tracking $I_{ref,i}^\eta = 0$. Note the dependence of h_i^η with respect to t, which indicates that h_i^η can be different in different intervals of the optimal solution.

Tracking $H_{u_i} = 0$ has different interpretations with respect to the two types of arcs. In the case of a constraint-seeking arc for u_i with the constraint S_j being active, $H_{u_i} = \lambda^T F_{u_i} + \mu_j \frac{\partial S_j}{\partial u_i} = 0$, with $\mu_j \neq 0$ and $\lambda^T F_{u_i} \neq 0$. The constraint has to be active for the sake of optimality since otherwise μ_j is zero and $H_{u_i} \neq 0$. Thus, the invariant along a constraint-seeking arc is $h_i^\eta(x, u, t) = S_j(x, u)$. For a compromise-seeking arc, $H_{u_i} = \lambda^T F_{u_i} = 0$. Therefore, the invariant is $h_i^\eta(x, u, t) = \lambda^T F_{u_i}(x, u)$.

Note that the element that remains invariant despite uncertainty is the fact that optimal operation corresponds to $I_{ref}^\eta = 0$. However, the uncertainty does have an influence on the value of $I^\eta(t)$, and the inputs need to be adapted in order to guarantee $I_{ref}^\eta = 0$.

Choice of invariants for constraint-seeking and compromise-seeking parameters. In addition to the choice of invariants for the various arcs, it is important to choose the invariants for the parameters π. Following similar arguments, a set of signals $I^\pi = h^\pi(x(t_f))$ can be constructed such that the optimum corresponds to $I_{ref}^\pi = 0$, also in the presence of uncertainty. Clearly, the invariants arise from the conditions of optimality. For the constraint-seeking parameters, they correspond to the terminal constraints $h^\pi(x(t_f)) = T(x(t_f))$ and, for the compromise-seeking parameters, to sensitivities $h^\pi(x(t_f)) = \frac{\partial \phi(x(t_f))}{\partial \tilde{\pi}}$.

To summarize, the invariants are as follows:

- For constraint-seeking arcs:

$$h_i^\eta(x, u, t) = S_j(x, u)$$

- For compromise-seeking arcs:

$$h_i^\eta(x, u, t) = \lambda^T F_{u_i}(x, u)$$

- For constraint-seeking parameters:

$$h^\pi(x(t_f)) = T(x(t_f))$$

- For compromise-seeking parameters:

$$h^{\pi}(x(t_f)) = \frac{\partial \phi(x(t_f))}{\partial \tilde{\pi}}$$

Sensitivity of the Cost. The sensitivity of the cost to non-optimal operation is in general much lower along a compromise-seeking arc than along a constraint-seeking arc. Consider the optimal input u_i determined by the path constraint S_j and the optimality condition $H_{u_i} = \lambda^T F_{u_i} + \mu_j \frac{\partial S_j}{\partial u_i} = 0$ with $\mu_j \neq 0$. If the input does not keep the constraint active, μ_j becomes zero. Thus, the change in cost is directly proportional to $\lambda^T F_{u_i}$, which is non-zero. In contrast, along a compromise-seeking arc, $H_{u_i} = \lambda^T F_{u_i} = 0$, and a small deviation of u_i from the optimal trajectory will result in a negligible loss in cost. Similarly, as seen from (16), the deviation in cost arising from the non-satisfaction of a terminal constraint is proportional to $\nu^T \frac{\partial T}{\partial \tilde{\pi}}$, whilst a small variation of $\tilde{\pi}$ cause negligible loss in cost.

In summary, it is far more important to track the path constraints S_j than the sensitivity $\lambda^T F_{u_i}$. Furthermore, it is far more important to track the terminal constraints T than the sensitivities $\frac{\partial \phi}{\partial \tilde{\pi}}$. Consequently, it is often sufficient in practical situations to focus attention only on constraint-seeking arcs and parameters.

Tracking of Invariants

The core idea of the optimization scheme is to use a model to determine the structure of the optimal inputs and measurements to update a few input parameters and the value of the inputs in some of the intervals. This way, the optimal inputs are determined directly from process measurements and not from a (possibly inaccurate) model.

Optimality despite uncertainty is approached by working close to the active constraints, i.e., where there is much to gain! Indeed, tracking path and terminal constraints is usually much more important than regulating sensitivities as was argued above. The structure given in Figure 2 is proposed to track the invariants by use of feedback (Srinivasan et al., 1997; Visser, 1999; Visser et al., 2000). The two major blocks are described below:

- *Analysis:* This level consists of a simplified (not necessarily accurate) model of the process. The tasks are as follows:
  1. The numerical optimizer solves the optimization problem using the simplified model and provides the nominal signals x^* and u^*.
  2. The type and sequence of arcs are deciphered from the numerical solution, leading to a parsimonious parameterization of the inputs (characterization of the optimal solution).
  3. The invariant functions $h^{\eta}(x(t), u(t), t)$ and $h^{\pi}(x(t_f))$ are obtained as proposed earlier. Note that the switching strategy is inherent in the choice of h^{η}.
- *Optimality through feedback:* The invariants $I^{\eta}_{ref} = 0$ and $I^{\pi}_{ref} = 0$ are tracked with the help of path and terminal feedback controllers, respectively. The trajectory generator computes the current inputs as a function of $\eta(t)$ and π.

The model-based 'Analysis' is normally carried out off-line once, and only the measurement-based 'Optimality through feedback' operates on-line during process runs. Thus, the implementation is *model-free* and *measurement-based*.

Practical Applicability of IBO

If the model and the true (unknown) system share the same input structure (type and sequence of arcs) and the same active terminal constraints, IBO will be capable of optimizing the true system. Thus, the value of IBO in practice will depend on both the robustness of the proposed input structure with respect to uncertainty (modeling errors and disturbances) and the ability to measure the path and terminal constraints. These issues are briefly discussed next.

- *Role of the model.* Although the implementation of optimal inputs in the invariant-based scheme is truly model-free, a model may still be needed at the analysis step. The role of the model is to determine the structure of the optimal inputs, i.e., the type and the sequence of arcs and the set of active constraints. The structure of the inputs is obtained either numerically or *via* educated guesses, with the proposed input structure being verified using PMP necessary conditions on the nominal model. The approach is directly applicable to large-scale industrial process, as long as the nominal model and the real system share the same input structure. So, in contrast to model-based approaches or what is sought for simulation purposes, there is no need for a detailed model or for accurate parameter values. The model simply needs to reflect the major tradeoffs specific to the optimization problem at hand. The parts of the model that do not address these effects can be discarded.
- *Construction of invariants from measurements.* Since I^{η} and I^{π} are not measured directly, they need to be reconstructed from the available measurements. In the case of constraint-seeking arcs and parameters, the invariants correspond to physical quantities (path or terminal constraints). Off-line measurements of terminal quantities are in general available. On-line measurements to meet path constraints might be more difficult to have. Three cases can be considered:

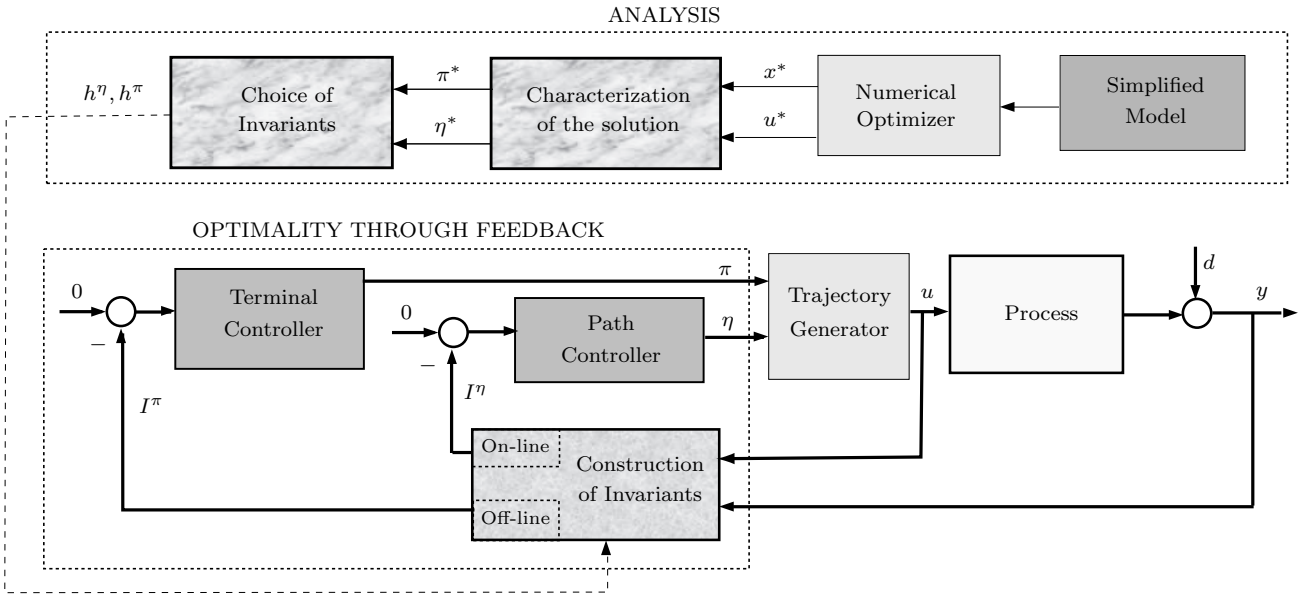


Figure 2: Invariant-based optimization.

1. The path constraint deals directly with a physical quantity that can be easily measured such as temperature or pressure.

2. The path constraint deals with a quantity that cannot be directly measured, but the constraint can be rewritten with respect to something that can be measured. This is, for example, possible when a heat removal constraint can be rewritten as a constraint on a cooling temperature.

3. Cases 1 and 2 do not apply, and some type of inference or state estimation is necessary to meet the constraint. This case is clearly more involved than the two preceding ones. The reconstruction problem is closely related to inferential control (Joseph and Brosilow, 1978; Doyle III, 1998). However, it may well happen that a conservative approach for meeting the path constraint (requiring easily-available or no measurements) is sufficient.

On the other hand, for compromise-seeking arcs and parameters, the invariants are sensitivities. For computation of sensitivities, either a model of the process or multiple process runs are required, which is typically more difficult. However, as discussed above, the sensitivity with respect to input variations in compromise-seeking arcs and parameters can often be neglected. In such a case, all compromise-seeking arcs and parameters are kept at their off-line determined values, and only the constraint-seeking arcs and parameters are adjusted.

- *Difference in time scale—on-line vs. off-line measurements.* In general, there is a difference in time scale between the path controller and the terminal controller. The path controller works within a batch using on-line measurements (running index is the batch time t) (Benthack, 1997). The terminal controller operates on a batch-to-batch basis using batch-end measurements (running index is the batch number k) (Srinivasan et al., 2001).
  If on-line measurements are not available, the path controller is inactive. If off-line measurements of the path constraint are available, it is possible to use the path controller in a batch-to-batch mode so that the system will be closer to the path constraint during the next batch (Moore, 1993). On the other hand, if it is possible to predict I^π from on-line measurements, it might be possible to use the terminal controller within the batch (Yabuki and MacGregor, 1997).
  The presence of disturbances influences both $\eta(t)$ and π. Disturbances affecting $\eta(t)$ within the batch are rejected by the path controller. However, the effect of any disturbance within the batch on π cannot be rejected since the terminal controller only works on a batch-to-batch basis. Constant disturbances (e.g. raw material variations) can be rejected from batch-to-batch by the terminal controller.

- *Backoff from constraints.* In the presence of disturbances and parametric uncertainty that cannot be

compensated for by feedback, the use of conservative margins, called backoffs, is inevitable to ensure feasibility of the optimization problem (Visser et al., 2000). The presence of measurement errors also necessitates a backoff. Based on an estimate of the uncertainty, the probability density function of the state variables can be calculated. The margins are then chosen such that the spread of the states remains within the feasible region with a certain confidence level. Note that the margins typically vary with time.

- *Reduction of backoff.* Due to the sensitivity reduction characteristic of feedback control, the conservatism can be reduced considerably in the proposed framework in comparison with the standard open-loop optimization schemes. The feedback parameters can be chosen so as to minimize the spread in the state variables resulting from uncertainty. The use of feedback becomes particularly important when the uncertainty tends to increase during a batch run. With reduced backoffs, the process can be driven closer to active constraints, thereby leading to improved performance.

Example—Semi-batch Reactor with Safety and Selectivity Constraints

Description of the Reaction System

- *Reaction:* $A + B \rightarrow C$, $2\,B \rightarrow D$.
- *Conditions:* Semi-batch, isothermal.
- *Objective:* Maximize the amount of C at a given final time.
- *Manipulated variable:* Feed rate of B.
- *Constraints:* Input bounds; limitation on the heat removal rate through the jacket; constraint on the amount of D produced.
- *Comments:* The reactor is kept isothermal by (say) adjusting the cooling temperature in the jacket, T_c. B is fed at the temperature $T_{in} = T$. To remain isothermal, the power generated by the reactions, q_{rx}, must be evacuated through the cooling jacket, i.e., $q_{rx} = UA(T - T_c)$. Thus, the heat removal constraint can be expressed in terms of a lower bound for the cooling temperature, T_{cmin}. The variables and parameters are described in the next subsection. Without any constraints, optimal operation would consist of adding the available B at initial time (i.e., batch mode). The presence of the heat removal constraints calls for semi-batch operation with constraint-seeking arcs. Furthermore, the constraint on the amount of D that can be produced imposes a compromise-seeking feeding of B in order to maximize C without violating the terminal constraint on D.

Problem Formulation

Variables and parameters: c_X: Concentration of species X, n_X: Number of moles of species X, V: Reactor volume, u: Feed rate of B, c_{Bin}: Inlet concentration of B, k_1, k_2: Kinetic parameters, ΔH_1, ΔH_2: Reaction enthalpies, T: Reactor temperature, T_c: Cooling temperature in the jacket, T_{in}: Feed temperature, U: Heat transfer coefficient, A: Surface for heat transfer, and q_{rx}: power produced by the reactions. The numerical values are given in Table 3.

Model equations:

$$\dot{c_A} = -k_1 c_A c_B - \frac{u}{V} c_A \tag{17}$$

$$\dot{c_B} = -k_1 c_A c_B - 2\,k_2 c_B^2 + \frac{u}{V}(c_{Bin} - c_B) \tag{18}$$

$$\dot{V} = u \tag{19}$$

with the initial conditions $c_A(0) = c_{A_o}$, $c_B(0) = c_{B_o}$, and $V = V_o$. Assuming $c_C(0) = c_D(0) = 0$, the concentration of the species C and D are given by $c_C = \frac{c_{A_o} V_o - c_A V}{V}$ and $c_D = \frac{(c_{B_o} V_o - c_B V) + c_{Bin}(V - V_o) - (c_{A_o} V_o - c_A V)}{2\,V}$. The power produced by the reactions and T_c are given by

$$q_{rx} = k_1\, c_A\, c_B(-\Delta H_1)V + 2\, k_2\, c_B^2(-\Delta H_2)V \tag{20}$$

$$T_c = T - \frac{q_{rx}}{U\,A} \tag{21}$$

Optimization problem:

$$\max_{u(t)} J = n_C(t_f) \tag{22}$$

subject to

$$\begin{gathered} (17) - (19) \\ T_c(t) \geq T_{cmin} \\ n_D(t_f) \leq n_{Df_{max}} \\ u_{min} \leq u \leq u_{max} \end{gathered}$$

Piecewise Analytic Characterization

Using Pontryagin's Maximum Principle, it can be shown that the competition between the two reactions results in a feed that reflects the optimal compromise between producing C and D. This compromise-seeking input can be calculated from the second time derivative of H_u as:

$$u_{comp} = \frac{V\, c_B(k_1 c_A(2\,c_{Bin} - c_B) + 4\,k_2 c_B c_{Bin})}{2\,c_{Bin}\,(c_{Bin} - c_B)} \tag{23}$$

The other possible arcs correspond to the input being determined by the constraints: (i) $u = u_{min}$, (ii) $u = u_{max}$, and (iii) $u = u_{path}$. The input u_{path} corresponds to riding along the path constraint $T_c = T_{cmin}$. The

k_1	0.75	l/(mol h)
k_2	0.014	l/(mol h)
ΔH_1	-7×10^4	J/mol
ΔH_2	-5×10^4	J/mol
c_{Bin}	10	mol/l
UA	8×10^5	J/Kh

u_{min}	0	l/h
u_{max}	100	l/h
T_{cmin}	10	°C
$n_{Df_{max}}$	5	mol

c_{Ao}	2	mol/l
c_{Bo}	0	mol/l
V_o	500	l
t_f	2.5	h

Table 3: Model parameters, operating bounds and initial conditions.

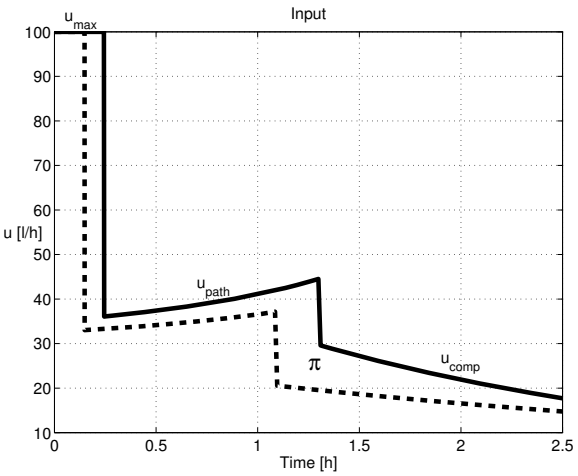


Figure 3: Nominal optimal input (solid) and Conservative optimal input (dotted).

input is obtained by differentiating the path constraint once with respect to time, i.e., from $\dot{T}_c = 0$:

$$u_{path} = \left.\frac{\mathcal{N}}{\mathcal{D}}\right|_{T_c=T_{cmin}} \tag{24}$$

$$\mathcal{N} = c_B V \Big(\Delta H_1 k_1 c_A (k_1 (c_A + c_B) + 2\, k_2 c_B) + 4\, \Delta H_2 k_2 c_B (k_1 c_A + 2\, k_2 c_B) \Big)$$

$$\mathcal{D} = \Delta H_1 k_1 c_A (c_{Bin} - c_B) + 4\, \Delta H_2 k_2 c_B (2\, c_{Bin} - c_B)$$

Sequence of arcs and parsimonious parameterization:

- The input is initially at the upper bound, $u = u_{max}$, in order to attain the path constraint as quickly as possible.
- Once T_c reaches T_{cmin}, $u = u_{path}$ is applied in order to keep $T_c = T_{cmin}$.
- The input switches to $u = u_{comp}$ at the time instant π so as to take advantage of the optimal compromise in order to maximize $n_c(t_f)$ and meet the terminal constraint $n_D(t_f) = n_{Df_{max}}$.

Since analytical expressions for the input in the various arcs exist, the optimal solution can be parameterized using only the switching time between the path constraint and the compromise-seeking arc. This parameter π is determined numerically so as to satisfy the terminal constraint $n_D(t_f) = n_{Df_{max}}$. The invariants I^η correspond to the input bound in the first interval, the path constraint in the second interval and the sensitivity $\lambda^T F_u$ for the compromise-seeking arc. For the switching time, the invariant is the terminal constraint itself. The optimal input is shown in Figure 3 with the optimal values $\pi = 1.31\,h$ and $J = 600.6$ mol.

Note that the input u_{path} given by (24) will keep the system on the path constraint once the path constraint $T_c = T_{cmin}$ is attained, but will not keep the path constraint active in the presence of uncertainty. The same can be said for u_{comp} in (23). Thus, the analytical expressions for u_{path} and u_{comp} will only be used for interpretation of the nominal optimal trajectory and not for implementation of the true optimal solution. At the implementation level, simple PI-controllers will be used.

Measurement-based Optimization

In practice, there can be considerable uncertainty both in the stoichiometric and kinetic models. This is reflected here as some uncertainty for the kinetic parameter k_1 in the interval $0.4 \le k_1 \le 1.2$ (The nominal value $k_1 = 0.75$ used in the simulation is assumed to be unknown). In order not to violate the constraints, a conservative feed profile (Figure 3) would have to be designed so that: i) the path constraint is not violated for $k_1 = k_{1max} = 1.2$, and ii) the terminal constraint is not violated for $k_1 = k_{1min} = 0.4$ (a smaller k_1 corresponds to more B in the reactor and thus to a higher production of D). So, the conservative profile would consist of computing u_{path} and the first switching instant using $k_1 = k_{1max}$, and adjusting π so that the terminal constraint is satisfied for $k_1 = k_{1min}$.

With respect to the measurements available, different optimization scenarios are considered:

1. *No measurements:* The conservative optimal feed profile defined above is applied open loop to the simulated nominal plant.

2. *Batch-end measurements:* Only the measurement of $n_D(t_f)$ is available and, thus, the switching time π is updated in a batch-to-batch manner. For the second interval, $u_{path} = u_{path}^{cons}$, the conservative value computed off-line using k_{1max} is applied. Due to the low sensitivity of the cost with respect to the fine shape of the input in the compromise-seeking

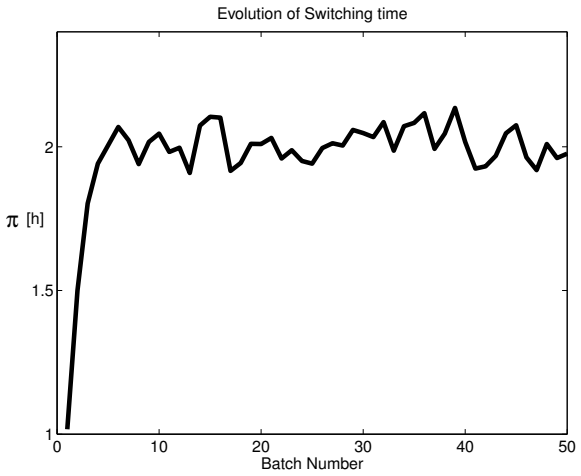


Figure 4: Evolution of switching time for one realization of the batch-to-batch optimization with only batch-end measurements (5% measurement noise).

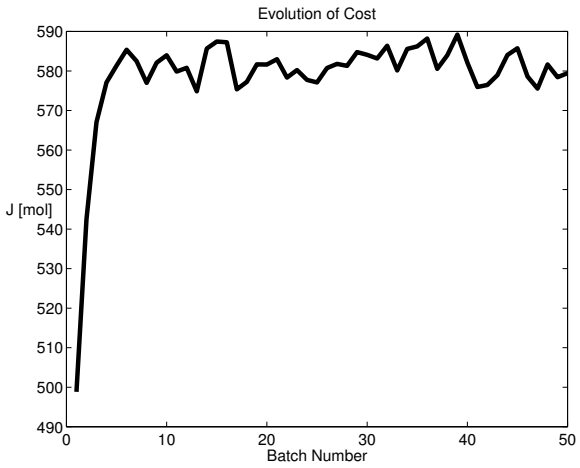


Figure 5: Evolution of cost for the same scenario as in Figure 4.

interval, the latter is approximated by the constant value $u_{comp} = 20$ (l/h).

3. *On-line and batch-end measurements:* On-line measurement of the cooling jacket temperature T_c is available. The path constraint is kept active using the feedback $u_{path}(t) = u_{path}^{cons} + k_p\,(T_{cmin} - T_c(t)) + k_i \int_0^t (T_{cmin} - T_c(t))dt$, where k_p and k_i are the parameters of a PI controller. In addition, the switching time π is updated in a batch-to-batch manner. As in Scenario 2, the compromise-seeking arc is approximated by $u_{comp} = 20$ (l/h).

The cases of both noise-free and noisy measurements (5% relative Gaussian measurement noise) are considered. The results are given in Table 4. If the measurements are noisy, a conservative margin (backoff) needs to be incorporated so as to guarantee feasibility. The backoffs are 0.25 mol for $n_{Df_{max}}$ and 1.25^{o}C for T_{cmin}.

It is seen that with only off-line (or batch-end) measurements, the terminal constraint can be satisfied by adapting the switching instant π. The evolutions of the switching instant and the cost for batch-to-batch optimization are shown in Figures 4 and 5. It can be seen that the solution gets close to the optimum within a few batches.

If, in addition, on-line measurements are available, the path constraint can be kept active as well. Thus, it is possible to get very close to the optimum by using measurements. The loss of 0.02% in the last noise-free scenario is due to the approximation of the compromise-seeking arc by the constant value $u_{comp} = 20$ (l/h).

Discussion

The model was only necessary to obtain the type and sequence of arcs: u_{max}, u_{path} and u_{comp}. As far as the implementation is concerned, $u_{max} = 100$ l/h is part of the problem formulation, u_{path} is determined by a PI-controller upon tracking T_{cmin}, and $u_{comp} = 20$ l/h is a constant-value approximation to the optimal profile computed off-line using the model. The switching time π between u_{path} and u_{comp} is adjusted in a run-to-run manner by a PI-controller in order to meet $n_{Df_{max}}$. The actual value of u_{comp} is of little relevance as any error in u_{comp} can be easily compensated for by an appropriate shift in π.

Assume that, in addition to the two modeled reactions, the true system also includes $B + C \rightarrow E$, $B \rightarrow F$. This would not affect the type and sequence of arcs (u_{max}, u_{path} and u_{comp}) since the two additional reactions are similar to the second reaction with respect to the effect of the input u, i.e., they consume B away from the first (desired) reaction. However, the two additional reactions are going to affect the two switching times t_1 and π as well as the numerical values of u_{path} and u_{comp}. The proposed scheme adjusts t_1 and u_{path} on-line (using the measurement of T_c) and π from batch-to-batch (using the measurement of $n_D(t_f)$). Since the cost is insensitive to the fine shape of u_{comp}, u_{comp}= 20 l/h is kept constant between π and t_f. Thus the proposed scheme would be equally applicable even in the presence of the two additional reactions.

In the formulation of the optimization problem it was assumed $T_{in} = T$. Even without this assumption, the proposed approach is applicable. The possibility of removing heat through temperature increase of the feed from T_{in} to T (so-called sensible heat) changes the heat

	Optimization Scenario	Terminal Constraint $n_D(t_f)$ mol ($n_{Df_{max}} = 5$ mol)	Path Constraint $\min_t T_c(t)\ ^oC$ ($T_{c,min} = 10^oC$)	Cost (mol)	Loss
1	Open-loop application of optimal conservative input	2.71	12.87	498.8	20%
2	Adaptation of π using off-line measurements (with 5% noise)	4.75	11.62	582.6	3%
	Adaptation of π using off-line measurements (no noise)	5.00	11.50	589.2	2%
3	Adaptation of $u_{path}(t)$ and π using on-line and off-line measurements (with 5% noise)	4.75	11.25	590.9	1.5%
	Adaptation of $u_{path}(t)$ and π using on-line and off-line measurements (no noise)	5.00	10.00	600.5	0.02%

Table 4: Invariant-based optimization. Results averaged over 100 noise realizations, each consisting of run-to-run adaptation over 50 batches.

removal constraint to:

$$q_{rx} - q_{in} \leq UA(T - T_{cmin}) \tag{25}$$

with q_{in} the rate of heat removal due to the feed of B. However, the implementation of the heat removal constraint remains unchanged as it concerns only the RHS of (25): $u_{path}(t)$ is determined as the output of a PI-controller designed to track T_{cmin}.

A final important remark: The model parameters given in Table 3 are not used for calculating the optimal feed rate. Only the off-line measurement of $n_D(t_f)$ and the on-line measurement of $T_c(t)$ are used to implement the proposed optimizing scheme.

Conclusions

This paper has addressed several optimization issues that directly affect the operation of batch processes. It is argued that process improvement is necessary for the economic well-being of many batch manufacturers. The industrial practice specific to the batch specialty chemistry is presented, with an emphasis on both organizational and technical problems. On the organizational side, the lack of global thinking in dealing with the individual steps of a complex process limits the potential for performance improvement. On the technical side, important limitations regarding both modeling and measurement aspects impair the use of optimization techniques. In addition, batch processes are characterized by a considerable amount of uncertainty and the presence of operational and safety constraints.

The lack of reliable models, together with the presence of uncertainty, has favored the investigation of process improvement *via* utilization of measurements (sometimes on-line, most often off-line). This paper has classified measurement-based optimization methods reported in the literature according to whether or not a model is used to guide the optimization and the type of measurements (on-line, off-line) available.

The major contribution towards process improvement of a constrained batch process is through operation on active constraints. Thus, a feedback-based framework has been proposed to keep the system 'close' to the active constraints. If only off-line measurements are available, this framework results in a batch-to batch optimization scheme with the objective to meet the terminal constraints within a few batches. If on-line measurements are available, the path constraints can also be kept active.

The proposed invariant-based optimization scheme addresses most of the requirements stemming from industrial practice and needs that were listed in Table 1. More specifically,

- it is aimed at process improvement via the use of time-dependent inputs,
- it is model-independent as far as implementation is concerned,
- if necessary, it uses only available off-line measurements,
- it is robust against uncertainty since signals that are invariant under uncertainty are tracked, and finally,

- it guarantees feasibility since the constraints are approached from the safe side.

The approach proposed is effective when the optimization potential stems mainly from meeting path and/or terminal constraints. Such is the case in most of the batch process optimization problems.

It is possible to perceive the proposed feedback-based optimization strategy from an industrial perspective. Classical PID control is the most popular technique used currently in industry, and trading it to attain optimality is unacceptable industrially. Therefore, in contrast to most model-based optimization studies, this work attempts to use feedback control for the sake of optimality. In this sense, the approach has great industrial potential and could help take optimization to the batch chemical industry.

References

Abel, O., A. Helbig, W. Marquardt, H. Zwick, and T. Daszkowski, "Productivity Optimization of an Industrial Semi-batch Polymerization Reactor under Safety Constraints," *J. Proc. Cont.*, **10**(4), 351–362 (2000).

Agarwal, M., "Feasibility of On-line Reoptimization in Batch Processes," *Chem. Eng. Commun.*, **158**, 19–29 (1997).

Allgor, R. J., M. D. Barrera, P. I. Barton, and L. B. Evans, "Optimal Batch Process Development," *Comput. Chem. Eng.*, **20**(6–7), 885–896 (1996).

Bastin, G. and D. Dochain, *On-line Estimation and Adaptive Control of Bioreactors*. Elsevier, Amsterdam (1990).

Baumann, T., *Infinite-order Singularity in Terminal-cost Optimal Control: Application to Robotic Manipulators*, PhD thesis 1778, Swiss Federal Institute of Technology, Lausanne, Switzerland (1998).

Benthack, C., *Feedback-Based Optimization of a Class of Constrained Nonlinear Systems: Application to a Biofilter*, PhD thesis 1717, Swiss Federal Institute of Technology, Lausanne, Switzerland (1997).

Bonvin, D., "Optimal Operation of Batch Reactors—A Personal View," *J. Proc. Cont.*, **8**(5–6), 355–368 (1998).

Bryson, A. E. and Y. C. Ho, *Applied Optimal Control*. Hemisphere, Washington DC (1975).

Bryson, A. E., *Dynamic Optimization*. Addison-Wesley, Menlo Park, California (1999).

Chin, I. S., K. S. Lee, and J. H. Lee, "A Technique for Integrated Quality Control, Profile Control, and Constraint Handling for Batch Processes," *Ind. Eng. Chem. Res.*, **39**, 693–705 (2000).

Clarke-Pringle, T. L. and J. F. MacGregor, "Optimization of Molecular Weight Distribution Using Batch-to-batch Adjustments," *Ind. Eng. Chem. Res.*, **37**, 3660–3669 (1998).

de Buruaga, Saenz, I., A. Echevarria, P. D. Armitage, J. C. de la Cal, J. R. Leiza, and J. M. Asua, "On-line Control of a Semi-batch Emulsion Polymerization Reactor Based on Calorimetry," *AIChE J.*, **43**(4), 1069–1081 (1997).

Dong, D., T. J. McAvoy, and E. Zafiriou, "Batch-to-batch Optimization using Neural Networks," *Ind. Eng. Chem. Res.*, **35**, 2269–2276 (1996).

Doyle III, F. J., "Nonlinear Inferential Control for Process Applications," *J. Proc. Cont.*, **8**(5–6), 339–353 (1998).

Eaton, J. W. and J. B. Rawlings, "Feedback Control of Nonlinear Processes using On-line Optimization Techniques," *Comput. Chem. Eng.*, **14**, 469–479 (1990).

Filippi-Bossy, C., J. Bordet, J. Villermaux, S. Marchal-Brassely, and C. Georgakis, "Batch Reactor Optimization by use of Tendency Models," *Comput. Chem. Eng.*, **13**, 35–47 (1989).

Fotopoulos, J., C. Georgakis, and H. G. Stenger, "Uncertainty Issues in the Modeling and Optimisation of Batch Reactors with Tendency Modeling," *Chem. Eng. Sci.*, **49**, 5533–5548 (1994).

Fournier, F., M. A. Latifi, and G. Valentin, "Methodology of Dynamic Optimization and Optimal Control of Batch Electrochemical Reactors," *Chem. Eng. Sci.*, **54**, 2707–2714 (1999).

Gattu, G. and E. Zafiriou, "A Methodology for On-line Setpoint Modification for Batch Reactor Control in the Presence of Modeling Error," *Chem. Eng. J.*, **75**(1), 21–29 (1999).

Ge, M., Q. G. Wang, M. S. Chin, T. H. Lee, C. C. Hang, and K. H. Teo, "An Effective Technique for Batch Process Optimization with Application to Crystallization," *Trans. IChemE*, **78A**, 99–106 (2000).

Gentric, C., F. Pla, M. A. Latifi, and J. P. Corriou, "Optimization and Nonlinear Control of a Batch Emulsion Polymerization," *Chem. Eng. J.*, **75**(1), 31–46 (1999).

Helbig, A., O. Abel, and W. Marquardt, Model Predictive Control for the On-line Optimization of semi-batch reactors, In *American Control Conference*, pages 1695–1699, Philadelphia, PA (1998).

Joseph, B. and C. Brosilow, "Inferential Control of Processes. III. Construction of Optimal and Suboptimal Dynamic Estimators," *AIChE J.*, **24**(3), 500–508 (1978).

Kirk, D. E., *Optimal Control Theory : An Introduction*. Prentice-Hall, London (1970).

Krothapally, M., B. Bennett, W. Finney, and S. Palanki, "Experimental Implementation of an On-line Optimization Scheme to Batch PMMA Synthesis," *ISA Trans.*, **38**, 185–198 (1999).

Lakshmanan, N. M. and Y. Arkun, "Estimation and Model Predictive Control of Nonlinear Batch Processes using Linear Parameter-varying models," *Int. J. Control*, **72**(7–8), 659–675 (1999).

Le Lann, M. V., M. Cabassud, and G. Casamatta, Modeling, Optimization, and Control of Batch Chemical Reactors in Fine Chemical Production, In *IFAC DYCOPS-5*, pages 751–760, Corfu, Greece (1998).

Lee, K. S., I. S. Chin, H. J. Lee, and J. H. Lee, "Model Predictive Control Technique Combined with Iterative Learning for Batch Processes," *AIChE J.*, **45**(10), 2175–2187 (1999).

Lee, J., K. S. Lee, J. H. Lee, and S. Park, An On-line Batch Span Minimization and Quality Control Strategy for Batch and Semi-batch Processes, In *IFAC ADCHEM'00*, pages 705–712, Pisa, Italy (2000).

Macchietto, S., Batch Process Engineering Revisited: Adding New Spice to Old Recipes, In *IFAC DYCOPS-5*, Corfu, Greece (1998).

Marchal-Brassely, S., J. Villermaux, J. L. Houzelot, and J. L. Barnay, "Optimal Operation of a Semi-batch Reactor by Self-adaptive Models for Temperature and Feedrate profiles," *Chem. Eng. Sci.*, **47**, 2445–2450 (1992).

Marquardt, W., "Trends in Computer-aided Modeling," *Comput. Chem. Eng.*, **20**, 591–609 (1996).

Martinez, E. C., "Batch Process Modeling for Optimization and Reinforcement Learning," *Comput. Chem. Eng.*, **24**, 1187–1193 (2000).

Mauderli, A. and D. W. T. Rippin, "Production Planning and Scheduling for Multi-purpose Batch Chemical Plants," *Comput. Chem. Eng.*, **3**, 199–206 (1979).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and Optimality," *Automatica*, **36**(6), 789–814 (2000).

McLennan, F. and B. Kowalski, editors, *Process Analytical Chemistry*. Blackie Academic and Professional, London (1995).

Meadows, E. S. and J. B. Rawlings, Model Identification and Control of a Semibatch Chemical Reactor, In *American Control Conference*, pages 249–255, Boston, MA (1991).

Moore, K. L., *Iterative Learning Control for Deterministic Systems*. Springer-Verlag, Advances in Industrial Control, London (1993).

Morari, M. and J. H. Lee, "Model Predictive Control: Past, Present, and Future," *Comput. Chem. Eng.*, **23**, 667–682 (1999).

Nichols, G. D., *On-line Process Analyzers*. John Wiley, New York (1988).

Noda, M., T. Chida, S. Hasebe, and I. Hashimoto, "On-line Optimization System of Pilot Scale Multi-effect Batch Distillation System," *Comput. Chem. Eng.*, **24**, 1577–1583 (2000).

Palanki, S., C. Kravaris, and H. Y. Wang, "Synthesis of State Feedback Laws for End-Point Optimization in Batch Processes," *Chem. Eng. Sci.*, **48**(1), 135–152 (1993).

Pantelides, C. C. and H. I. Britt, Multipurpose Process Modeling Environments, In *FOCAPD'94*, pages 128–141, Snowmass, CO (1994).

Qin, S. J. and T. A. Badgwell, An Overview of Industrial Model Predictive Technology, In *Chemical Process Control V*, pages 232–256, Tahoe City, CA (1997).

Rahman, S. and S. Palanki, "State Feedback Synthesis for On-Line Optimization in the Presence of Measurable Disturbances," *AIChE J.*, **42**, 2869–2882 (1996).

Rastogi, A., J. Fotopoulos, C. Georgakis, and H. G. Stenger, "The Identification of Kinetic Expressions and the Evolutionary Optimization of Specialty Chemical Batch Reactors using Tendency Models," *Chem. Eng. Sci.*, **47**(9-11), 2487–2492 (1992).

Rawlings, J. B., E. S. Meadows, and K. R. Muske, Nonlinear Model Predictive Control: A Tutorial and Survey, In *IFAC ADCHEM'94*, pages 185–197, Kyoto, Japan (1994).

Rippin, D. W. T., "Design and Operation of Multiproduct and Multipurpose Batch Chemical Plants: An Analysis of Problem Structure," *Comput. Chem. Eng.*, **7**, 463–481 (1983).

Roberts, P. D. and T. W. C. Williams, "On an Algorithm for Combined System Optimization and Parameter Estimation," *Automatica*, **17**, 199–209 (1981).

Ruppen, D., D. Bonvin, and D. W. T. Rippin, "Implementation of Adaptive Optimal Operation for a Semi-batch Reaction System," *Comput. Chem. Eng.*, **22**, 185–189 (1998).

Scheid, G. W., S. J. Qin, and T. J. Riley, Run-to-run Optimization, Monitoring, and Control on a Rapid Thermal Processor, In *AIChE Annual Meeting*, Dallas, TX (1999).

Schenker, B. and M. Agarwal, "On-line Optimized Feed switching in Semi-batch Reactors using Semi-empirical Dynamic models," *Control Eng. Practice*, **8**(12), 1393–1403 (2000).

Skogestad, S., "Plantwide Control: The Search for the Self-optimizing Control Structure," *J. Proc. Cont.*, **10**, 487–507 (2000).

Soroush, M. and C. Kravaris, "Nonlinear Control of a Batch Polymerization Reactor: An Experimental Study," *AIChE J.*, **38**(9), 1429–1448 (1992).

Srinivasan, B., E. Visser, and D. Bonvin, Optimization-based Control with Imposed Feedback Structures, In *IFAC ADCHEM'97*, pages 635–640, Banff, Canada (1997).

Srinivasan, B., C. J. Primus, D. Bonvin, and N. L. Ricker, "Run-to-run Optimization via Generalized Constraint Control," *Control Eng. Practice*, **9**(8), 911–919 (2001).

Terwiesch, P. and M. Agarwal, "On-line Correction of Predetermined Input Profiles for Batch Reactors," *Comput. Chem. Eng.*, **18**, S433–S437 (1994).

Terwiesch, P., M. Agarwal, and D. W. T. Rippin, "Batch Unit Optimization with Imperfect Modeling—A Survey," *J. Proc. Cont.*, **4**, 238–258 (1994).

Tsen, A. Y., S. S. Yang, D. S. H. Wong, and B. Joseph, "Predictive Control of Quality in a Batch Polymerization using a Hybrid Artificial Neural Network Model," *AIChE J.*, **42**, 435–455 (1996).

Ubrich, O., B. Srinivasan, F. Stossel, and D. Bonvin, Optimization of a Semi-batch Reaction System under Safety Constraints, In *European Control Conference*, pages F306.1–6, Karlsruhe, Germany (1999).

Van Impe, J. F. and G. Bastin, "Optimal Adaptive Control of Fedbatch Fermentation processes," *Control Eng. Practice*, **3**(7), 939–954 (1995).

Verwater-Lukszo, Z., "A Practical Approach to Recipe Improvement and Optimization in the Batch Processing Industry," *Comp. in Industry*, **36**, 279–300 (1998).

Visser, E., B. Srinivasan, S. Palanki, and D. Bonvin, "A Feedback-based Implementation Scheme for Batch Process Optimization," *J. Proc. Cont.*, **10**, 399–410 (2000).

Visser, E., *A Feedback-based Implementation Scheme for Batch Process Optimization*, PhD thesis 2097, Swiss Federal Institute of Technology, Lausanne, Switzerland (1999).

Wiederkehr, H., "Examples of Process Improvements in the Fine Chemicals Industry," *Comput. Chem. Eng.*, **43**, 1783–1791 (1988).

Wittenmark, B., Adaptive Dual Control Methods: An Overview, In *IFAC Symposium on Adaptive Syst. in Control and Signal Proc.*, pages 67–72, Budapest (1995).

Yabuki, Y. and J. F. MacGregor, Product Quality Control in Semi-batch Reactors using Mid-course Correction Policies, In *IFAC ADCHEM'97*, pages 189–194, Banff, Canada (1997).

Zafiriou, E. and J. M. Zhu, Optimal Control of Semi-batch Processes in the Presence of Modeling Error, In *American Control Conference*, pages 1644–1649, San Diego, CA (1990).

Zafiriou, E., H. W. Chiou, and R. A. Adomaitis, "Nonlinear Model-based Run-to-run Control for Rapid Thermal Processing with Unmeasured Variable Estimation," *Electrochem. Soc. Proc.*, **95**(4), 18–31 (1995).

Dynamics and Control of Cell Populations in Continuous Bioreactors

Prodromos Daoutidis *
Department of Chemical Engineering and Materials Science
University of Minnesota
Minneapolis, MN 55455

Michael A. Henson †
Department of Chemical Engineering
Louisiana State Unversity
Baton Rouge, LA 70803-7303

Abstract

Continuous bioreactors are critical unit operations in a wide variety of biotechnological processes. While they can be viewed as chemical reactors, bioreactors offer unique modeling and control challenges due to the complexity of the underlying biochemical reactions and the distributed properties of the cell population. The dynamic behavior of continuous bioreactors can be strongly affected by variations between individual cells that are captured only with cell population models. The objective of this paper is to outline recent progress in dynamic analysis and feedback control of continuous bioreactors described by cell population models. The industrially important process of continuous yeast production is used to illustrate various concepts. Future research problems in cell population modeling, dynamics and control are outlined to provide insights on the key challenges and opportunities in this emerging area.

Keywords

Biochemical reactors, Population balance models, Cell population dynamics, Nonlinear control

Introduction

Biochemical engineering is concerned with the industrial production of biologically based products such as foods and beverages, pharmaceuticals, commodity chemicals, specialty chemicals and agricultural chemicals. The biochemical manufacturing industry is growing rapidly due to dramatic advancements in biotechnology and the high value of biochemical products such as pharmaceuticals (Lee, 1992). Process control has played a rather limited role in the biochemical industry as the economic incentive for improved process operation often is dwarfed by costs associated with research and development. This situation is likely to change with the expiration of key patents and the continuing development of global competition. Another obstruction to process control has been the lack of on-line sensors for critical process variables (Pons, 1992). While this will remain an important issue for the forseeable future, recent advancements in biochemical measurement technology make the development of advanced process control systems a realistic goal. These trends suggest that biochemical processes will emerge as an important application area for control engineers.

A complete review of the modeling and control needs in the biochemical industry would require a lengthy book rather than a short paper. Therefore the scope of this paper is limited to continuous bioreactors used for the growth of microbial cell cultures important in the food and beverage, pharmaceutical and agricultural chemical industries. Other types of bioreactors (batch and semi-batch) and cell cultures (animal and plant) are not covered despite their industrial importance. The remainder of this section is used to provide an overview of continuous bioreactors with particular emphasis on the process modeling and control challenges.

Continuous Bioreactors

A typical biochemical process involves batch, semi-batch and/or continuous reactors in which raw materials are transformed into the desired biological products. In many applications, continuous bioreactors are preferred due to their ease of operation and higher productivity (Lee, 1992). A prototypical continuous stirred tank bioreactor (also known as a continuous fermentor) is depicted in Figure 1. Medium is supplied continuously to the reactor to sustain growth of the microbial cell population. The synthetic medium contains the substrate(s) metabolized by the cells during growth as well as other components such as mineral and salts required to replicate the natural growth environment. The culture is called aerobic if the biochemical reactions involved in cell growth require oxygen as a reactant. In this case, air must be supplied continuously to maintain the necessary dissolved oxygen concentration. By contrast, anaerobic cultures do not require oxygen for cell growth. For each microorganism there is a unique range of culture temperature and pH that support cell growth. A typical bioreactor has simple feedback control loops that maintain the temperature and pH at predetermined setpoints chosen to maximize cell growth (Pons, 1992).

An agitator is used to continuously mix the reactor contents. The agitator speed is chosen to provide satisfactory mixing while avoiding excessive shear forces that may damage cells (Lee, 1992). A stream is removed continuously from the reactor to achieve constant volume operation. The removal rate is characterized by the dilution rate, which is the reciprocal of the reactor residence time. The dilution rate is controlled by a simple feedback controller that manipulates the medium flow rate. The effluent stream contains unreacted substrate and biomass that is a complex mixture of cells and vari-

*daoutidi@cems.umn.edu
†henson@che.lsu.edu

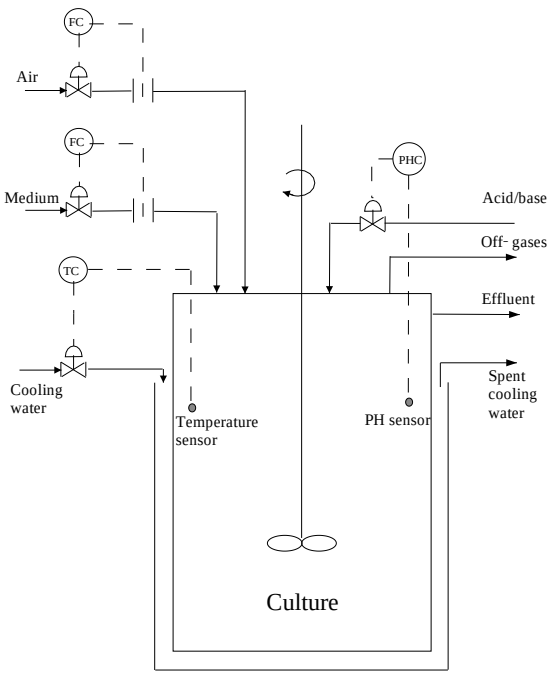


Figure 1: Continuous bioreactor.

ous metabolites produced by the cells. Desired products can be the cells themselves, one or more metabolites, or some combination of cells and metabolites. Products are separated from the other components via a series of recovery and purification operations (Lee, 1992). In addition to the liquid product stream, off-gases such as carbon dioxide may be produced as byproducts of the biochemical reactions.

Successful operation of a continuous bioreactor requires much more than simply supplying the necessary nutrients and extracting the desired products. Careful preparation of the growth medium is essential as microorganisms are strongly affected by changes in the culture environment. The microorganism must be innoculated in the reactor to initiate cell growth. Typically innoculation is achieved via a multistep procedure in which cells grown in a shake flask are transferred to increasingly larger volume bioreactors until the production bioreactor is innoculated (Lee, 1992). This procedure is necessary to achieve a sufficiently large cell population to sustain growth. A critical requirement is to maintain sterility of the medium and all processing equipment. Even a small amount of contamination can lead to complete loss of productivity and shutdown of the bioreactor. As a result of these complexities, effective operation of continuous bioreactors is a very challenging problem.

Opportunities for Process Modeling and Control

Mathematical modeling of cell growth kinetics in continuous bioreactors continues to be a major focus of biochemical engineering research (Nielsen and Villadsen, 1994). The potential impact of such models on bioprocess simulation, scale-up, optimization and control is significant. As compared to conventional chemical reactors, bioreactors are particularly difficult to model due to the complexity of the biochemical reactions, the unique characteristics of individual cells and the lack of measurements of key process variables. The consumption of substrates and production of metabolites results from hundreds of coupled biochemical reactions (Mauch et al., 1997). The identification and modeling of these complex reaction networks are very challenging problems usually not encountered in other chemical reaction systems. While it is convenient to view a microbial culture as a homogeneous mixture of identical cells, most cultures actually are comprised of a heterogeneous mixture of cells that differ with regard to size, mass and intracellular concentrations of proteins, DNA and other chemical components (Srienc and Dien, 1992). Accurate modeling of cell growth and product formation kinetics may require that individual cells be differentiated based on these characteristics. While on-line sensors for secondary variables such as carbon dioxide off-gas concentration are available, measurements of primary variables such as the biomass and product concentrations require expensive analytical equipment (Lee, 1992). Accurate and reliable measurement of these primary variables often is required to develop and validate mathematical models.

As shown in Figure 1, a typical control system for a continuous bioreactor consists of simple feedback control loops that regulate reactor residence time, temperature and pH. The control system is designed to supply the prescribed flow of nutrients while avoiding environmental conditions that adversely affect bioreactor productivity. With regards to key output variables such as the biomass and product concentrations, this is an open-loop control strategy based on the unrealistic assumption that unmeasured disturbances have a negligible effect on bioreactor performance. The development of closed-loop control strategies for reactor stabilization and biomass/product optimization would represent a major advance in the biochemical industry.

Overview of the Paper

The remainder of the paper is organized as follows. Section 2 contains an introduction to the mathematical modeling of continuous bioreactors with an emphasis on cell population models. Section 3 focuses on the dynamic behavior of cell population models with particular emphasis on yeast culture models. The design of feedback controllers using cell population models and the critical issue of on-line measurements are discussed in Section 4.

Finally our personal perspective on future research in cell population modeling, dynamics and control is presented in Section 5.

Mathematical Modeling of Cell Growth Dynamics

Classification and Overview

Mathematical models that describe cell growth processes can be classified into two broad categories:

- *Continuum or unsegregated models* which treat the cell population as a continuum or a lumped biophase, i.e. assume that it behaves as a homogeneous entity.
- *Corpuscular or segregated or cell population balance* models which account for the heterogeneous and distributed nature of cell growth, i.e. the fact that a cell population consists of individual cells.

Continuum models include compartmental (Roels, 1983) and detailed metabolic models (see e.g. Nielsen and Villadsen, 1992, and the references therein) which attempt to describe the influence of intracellular metabolism on cell growth, as well as cybernetic models (e.g. Kompala et al., 1986; Straight and Ramkrishna, 1994) which postulate the optimal nature of biomass growth and nutrient uptake in order to predict growth dynamics. The mathematical formulation of continuum models typically leads to a set of nonlinear ordinary differential equations, whereas corpuscular models typically consist of sets of first order partial integro-differential equations coupled with ordinary integro-differential equations that describe substrate consumption and/or product formation.

A second important classification of both continuum and corpuscular models is in *unstructured* and *structured* models. Structured continuum (structured or multivariable corpuscular) models account for the fact that the lumped biomass (single cell) is comprised of different chemical components, such as DNA, RNA, protein etc, while unstructured continuum and corpuscular models do not. Hence, structured corpuscular models not only account for the fact that cells within a population can behave differently, but they also account for chemical structure within a single cell. On the other hand, in structured continuum models, the chemical structure is included at the cell population level since the continuum approach does not distinguish between different cells.

Moreover, structured or unstructured cell population balance models are assorted in *single-staged* and *multi-staged* models depending on the number of cell cycle stages that are included in the cell population balance formulation. Finally, if the property or properties that are used to describe the intracellular structure obey the mass conservation law, then the cell population balance model is referred to as *mass structured*, whereas if age is used to differentiate each cell from other cells of the population, then the model is referred to as *age structured*.

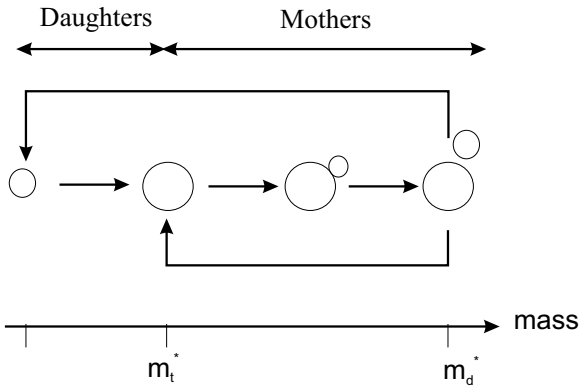


Figure 2: Simplified cell cycle for budding yeast.

Due to the level of detail built in their mathematical formulation, structured cell population balance models represent the most accurate way of describing the complicated phenomena associated with cell growth, nutrient uptake and product formation. Moreover, the mathematical formulation of such models naturally allows the incorporation of information about transition between successive cell cycle stages and partitioning of cellular material upon cell division. Furthermore, contrary to continuum models, which can predict only average population properties, cell population balance models are able to predict entire cell property distributions.

Baker's Yeast: A Motivating Example for Cell Population Balance Models

In what follows, we briefly discuss *Saccharomyces cerevisiae* as an illustrative example of a microorganism for which cell population balance models play a key role in the dynamic analysis and control of its cultures.

Saccharomyces cerevisiae is a key microorganism in the brewing, baking and genetic engineering industries. Also known as Baker's yeast, it has been widely studied due to its own importance as well as to understand the behavior of more complex cells present in plants and animals. It can be grown in aerated continuous bioreactors by feeding a nutrient stream containing glucose substrate. A variety of products including ethanol are produced.

A distinctive feature of yeast cells is that they divide via an asymmetric process known as budding (Hjortso and Nielsen, 1994). A simplified depiction of the yeast cell cycle is shown in Figure 2. The cell population is characterized in terms of daughter cells and mother cells. A daughter cell consumes substrate until it reaches a critical mass known as the transition mass. At this point, the daughter cell becomes a mother cell and a small bud attached to the cell begins to grow. Additional substrate

consumption increases the mass of the bud while the mass of the mother cell remains essentially constant. At a second critical mass known as the division mass, the bud separates from the mother cell producing a newborn daughter cell and a mother cell.

Many investigators have shown that continuous yeast cultures exhibit sustained oscillations under operating conditions encountered in industrial bioreactors (von Meyenburg, 1973; Parulekar et al., 1986). The intracellular mechanisms that cause these oscillations are controversial and have been a subject of three decades of intensive research. However, recent modeling and dynamical studies (Zhang et al., 2001) have established a strong dependence of the open-loop dynamics of yeast bioreactors on the initial cell mass *distribution*. Clearly, this dependence can only be captured (and accounted for in the controller design) by cell population balance models.

Furthermore, the existence of two distinct stages in yeast cultures (budded and unbudded) and the fact that key products of interest (such as ethanol) have been shown to be produced preferentially during the second part of the cell cycle (Alberghina et al., 1991; Frykman, 1999) suggest the use of two-staged population balance models for predicting and controlling the production of such products. This is also the case in other cell cultures, e.g. in murine hybridoma cells where the antibody secretion rates have been found to be much higher in the late stages of the cell cycle (Kromenaker and Srienc, 1994). This type of information is simply not present in continuum models, whereas it can be naturally incorporated in multi-staged cell population balance models.

Mathematical Formulation of Structured Cell Population Balance Models

In this section we briefly describe the mathematical formulation of cell population balance models in continuous bioreactors such as the one depicted in Figure 1. Each individual cell in the population of cells contained in the bioreactor is assumed to comprise of r biochemical components (DNA, RNA, protein etc.), with different cells containing different quantities of these components. The vector $x = [x_1, x_2, \cdots, x_r]$ with elements the amounts of these components in each cell is called the physiological state vector of the cell. The physiological state space is expressed as $G = [x_{n,min}, x_{max}]$, where $x_{n,min}, x_{max}$ denote the vectors containing the minimum and maximum, respectively, values for the amounts of the r biomass components of the newborn cells. Finally, x_{min} denotes the vector containing the minimum values of the amounts of the r biomass components of the dividing cells. For the sake of simplicity and without loss of generality, it is often assumed that the minimum and maximum values of the quantities of all biomass components are $x_{n,min} = x_{min} = 0$ and $x_{max} = 1$, respectively.

The state of the entire population is described by a time-dependent function $N(x,t)$, such that $N(x,t)dx$ represents the number of cells per unit of biovolume that at time t have physiological state representation between x and $x + dx$. The total number of cells per unit of biovolume (cell density) and the concentration of the i-th biomass component are respectively obtained from the zeroth and first moments of the state distribution function:

$$N_t(t) = \int_{x_{n,min}}^{x_{max}} N(x,t)\, dx \tag{1}$$

$$N_{b,i}(t) = \int_{x_{n,min}}^{x_{max}} x_i N(x,t)\, dx, \qquad i = 1, \ldots, r \tag{2}$$

The sum from 1 to r of all expressions defined in Equation 2 yields the total biomass concentration at time t. Finally, S denotes the substrate concentration vector (assuming s substrates), S_f the feed substrate concentration vector and D the dilution rate.

A cell population balance model includes information about nutrient uptake, growth, division and birth at the single-cell level. These processes are mathematically described by a set of functions known as intrinsic physiological state functions which, in general, depend on the physiological state of the cell x and the state of the substrate environment S. Specifically, the nutrient consumption is characterized by the s-dimensional consumption rate vector $q(x,S)$ whose elements express the single-cell rate of consumption of each substrate. The growth process is represented by the r-dimensional growth rate vector $r(x,S)$ whose elements express the single-cell rate of increase in the amount of the each cellular component. The cell division is described by the division rate $\Gamma(x,S)$. Finally, the birth process is described by the partition probability density function $p(x,y,S)$, which expresses the probability that a mother cell with physiological state vector y gives birth to a daughter cell with physiological state vector x; this function must satisfy the normalization condition:

$$\int_{x_{min}}^{x_{max}} p(x,y,S)\, dx = 1 \tag{3}$$

which guarantees that it is a probability density function. It should also be such that the amount of each one of the r biochemical components is conserved at cell division. In particular, since no daughter cell can have greater amounts of any component than the dividing cell from which it originates, the partitioning function p should be zero for all daughter cell states that are greater than the states of the corresponding mother cell, i.e.:

$$p(x,y,S) = 0, \qquad \forall x_i > y_i, \quad i = 1, \ldots, r \tag{4}$$

Finally, the probability of a dividing cell with physiological state vector y to produce a daughter cell of state x must be equal to the probability of producing a daughter cell of state $y - x$, i.e.

$$p(x,y,S) = p(y - x, y, S) \tag{5}$$

For simplicity, it is also assumed that the bioreactor operates in conditions under which the cell death rate is negligible (quite common in practice).

The Cell Population Balance Equation. Under the assumptions and the process description presented above, the dynamics of the state distribution function $N(x,t)$ are described by the general cell population balance equation (Fredrickson et al., 1967; Ramkrishna, 1979):

$$\frac{\partial N(x,t)}{\partial t} + \nabla_x \cdot [r(x,S)N(x,t)] + \Gamma(x,S)N(x,t) + DN(x,t) = 2\int_x^{x_{max}} \Gamma(y,S)p(x,y,S)N(y,t)\,dy \quad (6)$$

subject to the initial condition:

$$N(x,0) = N_0(x) \quad (7)$$

The first term in Equation 6 denotes accumulation. The second term accounts for the loss of cells with the physiological state vector representation x due to the fact that they grow into bigger cells. The third term represents loss of cells with physiological state vector x due to division leading to the birth of smaller cells. The fourth term is the dilution term describing the rate by which cells exit the reactor. The source term in the right-hand side is the rate of birth of cells with the physiological state vector x originating from the division of all bigger cells. The integration in this term is performed in all r dimensions of the physiological state space and has a lower limit of x due to the fact that cells of physiological state x can not be born from cells with amounts of biochemical components less than x. Moreover, the factor of two multiplying the integral birth term accounts for the fact that each division event leads to the production of two daughter cells. For a detailed discussion on the statistical foundation of the above model and the detailed assumptions made for its derivation, the reader is referred to Fredrickson et al. (1967).

Boundary Conditions. Besides an initial condition, appropriate boundary conditions for the first order partial differential equation in Equation 6 are required. Defining the boundary B of the physiological state space G as the set of points where at least one of the r biochemical biomass components obtains either its maximum or minimum value, the boundary conditions can be mathematically expressed as (Eakman et al., 1966; Fredrickson et al., 1967):

$$r(x,S)N(x,t) = 0, \qquad \forall x \in B \quad (8)$$

These conditions (often referred to as regularity conditions), essentially specify the boundary of the physiological state space (and hence can be more accurately thought of as 'containment' conditions (Fredrickson and Mantzaris, 2002)) and have been the subject of considerable discussion in the literature. They can be derived from balances for the cell density and the concentrations of the biomass components (see e.g. Mantzaris et al., 2001a), and essentially force the solution of the cell population balance equation to satisfy two facts imposed by the physics of the problem: a) that cell growth does not affect the number of cells, and b) that cell division preserves biomass.

The Dynamics of the Substrate Concentrations. The cell population balance equation is coupled with the equations describing the dynamics of the substrate concentrations:

$$\frac{dS}{dt} = D(S_f - S) - \int_{x_{n,min}}^{x_{max}} q(x,S)N(x,t)\,dx \quad (9)$$

subject to the initial conditions:

$$S(0) = S_0 \quad (10)$$

The integral term in the above mass balance represents the rate of loss of substrate leading to cell growth. In the case where a single rate limiting substrate is present the above set of equations reduces to a single equation.

Notice that the coupling between the cell population balance equation and the ordinary integro-differential equations shown above occurs through the dependence of the intrinsic physiological state functions on the concentrations of the substrates. This coupling is the only source of nonlinearity in the model. If the assumption of constant substrate concentrations is made (not a reasonable one in most cases of practical interest), then the model consists only of the cell population balance model and is linear.

Unstructured Cell Population Balance Models. A common simplification to the general structured cell population balance model presented above concerns the case of a single physiological state x, usually the cell mass m; this is quite meaningful in bioreactors where cell growth and division are strongly dependent on cell mass. In this case, the growth rate vector becomes a scalar and the resulting unstructured cell population model takes the form:

$$\frac{\partial N(m,t)}{\partial t} + \frac{\partial [r(m,S)N(m,t)]}{\partial m} + \Gamma(m,S)N(m,t) + DN(m,t) = 2\int_m^{m_{max}} \Gamma(m',S)p(m,m',S)N(m',t)\,dm' \quad (11)$$

with the integral in the birth term being a one-dimensional one, and the mass state space defined as $M = [0,\ m_{max}]$. The containment boundary conditions take the form:

$$r(0,S)N(0,t) = r(m_{max},S)N(m_{max},t) = 0 \quad (12)$$

and the substrate balance equations become:

$$\frac{dS}{dt} = D(S_f - S) - \int_0^{m_{max}} q(m,S)N(m,t)\,dm \quad (13)$$

Multi-Staged Cell Population Balance Models. Consider now the case where the cells grow in two distinct stages (e.g. budded and unbudded in the case of yeast), with stage 1 cells being born through the division of stage 2 cells, and stage 2 cells being formed through the transition of stage 1 cells to stage 2. Let $N_1(x,t)dx$ and $N_2(x,t)dx$ denote the number of cells per unit of biovolume in stages 1 and 2, respectively, which at time t have physiological state between x and $x+dx$. Let also $r_1(x,S), r_2(x,S)$ denote the corresponding growth rate vectors, $\Gamma_1(x,S), \Gamma_2(x,S)$ the transition rates from stage 1 to stage 2 and from stage 2 to stage 1, respectively, and $p(x,y,S)$ the partition probability density function. Then, the dynamics of the two subpopulations are described by the following coupled set of cell population balance equations:

$$\begin{aligned}\frac{\partial N_1(x,t)}{\partial t} &+ \nabla_x \cdot [r_1(x,S)N_1(x,t)] \\ &+ \Gamma_1(x,S)N_1(x,t) + DN_1(x,t) \\ &= 2\int_x^{x_{max}} \Gamma_2(y,S)p(x,y,S)N_2(y,t)\,dy \quad (14)\end{aligned}$$

$$\begin{aligned}\frac{\partial N_2(x,t)}{\partial t} &+ \nabla_x \cdot [r_2(x,S)N_2(x,t)] \\ &+ \Gamma_2(x,S)N_2(x,t) + DN_2(x,t) \\ &= \Gamma_1(x,S)N_1(x,t) \quad (15)\end{aligned}$$

Note that the above equations are coupled through their corresponding source terms appearing in the right-hand sides.

The balances on the substrates in this case take the form:

$$\begin{aligned}\frac{dS}{dt} = D(S_f - S) &- \int_{x_{n,min}}^{x_{max}} q_1(x,S)N_1(x,t)\,dx \\ &- \int_{x_{n,min}}^{x_{max}} q_2(x,S)N_2(x,t)\,dx \quad (16)\end{aligned}$$

where $q_1(x,S), q_2(x,S)$ denote the corresponding substrate consumption rates in the two stages. Appropriate initial and containment boundary conditions (see e.g. Mantzaris et al., 2002) complete the model formulation. The incorporation of multiple cell cycle stages in the mathematical model can also be performed in a similar way (Hatzis et al., 1995).

Numerical Solution

Despite the generality, accuracy, and predictive power of cell population balance models, and the fact that they have been formulated since the mid 60s, their use for design, optimization and control of bioprocesses has been sparse. One major obstacle to this end is that cell population balance models require information at the single-cell level; in particular, they require the knowledge of the intrinsic physiological state functions (single-cell growth rates, single-cell stage-to-stage transition rates and partitioning function). The experimental determination of these functions is hard, mainly due to the fact that it requires measurements at the single-cell level. However, the evolution of flow cytometry (Srienc, 1993) has contributed significantly in providing a basis for obtaining information at the single cell level. The analysis of such information with inverse population balance modeling techniques (Ramkrishna, 1994) has enabled, in certain cases, the determination of the intrinsic physiological state functions (Srienc, 1999).

A second obstacle towards the practical application of cell population balance models is the fact that owing to their complex mathematical nature (first order partial integro-differential equations, coupled in a nonlinear fashion with ordinary integro-differential equations), the development of numerical algorithms for the accurate approximation of their solution is a challenging task.

Several studies have addressed the numerical solution of age structured cell population balances (Hjortso and Bailey, 1983; Hjortso and Nielsen, 1994, 1995; Kim, 1996; Kim and Park, 1995a,b; Kostova, 1990, 1988; Kurtz et al., 1998). However, age structured models are limited by the fact that age is very difficult to measure experimentally in microbial populations.

On the other hand, some properties of cells, such as volume, total protein content, DNA content can be measured even at the single-cell level. Therefore, the use of such properties in the formulation of mass structured models is quite meaningful. Liou et al. (1997) developed analytical solutions of mass structured and age-mass structured cell population balances in the case of some simple single-cell growth rate expressions. Reports on the numerical solution of more general mass structured models had been sparse until recently. Subramanian and Ramkrishna (1971) employed a combination of the weighted residual method and the successive approximation method. However, this approach is limited to the case of linear growth rate where the cell population balance and the substrate concentration equation can be decoupled. Sulsky (1994) addressed a specific nonlinear mass structured population balance model, with the use of classical finite difference schemes as well as the method of characteristics. However, the model under consideration did not include changes in the environmental conditions, which when incorporated in the mathematical formulation can dramatically alter the dynamics of the cell population as well as the behavior of numerical schemes. Godin et al. (1999a,b) and Zhu et al. (2000; 2001) employed finite element techniques for the solution of the

problem under conditions of changing substrate concentration. Finally, Mantzaris et al. (1999) proposed a finite difference technique applicable to problems with changing substrate concentration and various sets of physiological state functions.

The above reports focused on unstructured models, which do not incorporate any internal chemical structure of the single cell. Mantzaris et al. (2001a,b,c) have recently developed several finite difference, spectral and finite element algorithms for the solution of structured cell population balance models, and evaluated these algorithms in terms of numerical stability, accuracy and computational speed. These algorithms are quite general in the sense that they are not limited by the choice of the physiological state functions, and can be applied for any number of substrates and constant or changing environmental conditions. With small modifications, they can also be applied in the case of multi-staged cell population balance models (Mantzaris et al., 2002).

In conclusion, the recent studies on the numerical solution of cell population balance models have led to a variety of algorithms that can be used to efficiently obtain accurate solutions of these models and hence facilitate their use in optimization and control.

Cell Population Dynamics

The dynamics of continuous bioreactors are important for simulation and control of industrial bioprocesses. Bioreactors can exhibit complex dynamic behavior due to nonlinearities associated with cell growth and division processes. Unlike most other types of chemical reactors, these nonlinear dynamics are not caused by the nonlinear dependence of reaction rates on temperature. Indeed continuous bioreactors operated at constant temperature can exhibit nonlinear behavior such as multiple steady states and limit cycles (Hjortso and Bailey, 1983; Hjortso and Nielsen, 1995). While cell metabolism certainly plays an important role, the observed nonlinear behavior is partially attributable to complex interactions between the cell population and the culture environment (Eakman et al., 1966; Subramanian and Ramkrishna, 1971). Consequently the study of cell population dynamics has considerable theoretical and practical significance. The objective of this section is to provide a brief introduction to the control relevant dynamics of cell population models with particular emphasis on limit cycle behavior in continuous yeast bioreactors.

Steady-State and Periodic Solutions

A rigorous dynamic analysis of the general cell population model (6)–(10) is very difficult due to the complexity of the model equations. The problem can be simplified by considering only a single physiological state (x) and a single rate limiting substrate (S). In this case the unstructured cell population model can be written as in (11)–(13). It is well known that this model can exhibit both steady-state and periodic solutions for specific forms of the physiological state functions.

The first problem considered is existence and stability of steady-state solutions. As can be seen from (11)–(13), a solution that exists for all values of dilution rate (D) and feed substrate concentration (S_f) is the so-called washout steady state: $N(m) = 0\ \forall m$, $S = S_f$. This corresponds to the highly undesirable situation where substrate is fed to the reactor but biomass is not produced. Stability of the washout steady state usually can be characterized in terms of a critical dilution rate (D_c) that is a complex function of the physiological state functions and parameter values. For $D < D_c$ the washout steady state is unstable while for $D \geq D_c$ it is stable. Hence there is a tradeoff between reactor stability (low D) and reactor throughput (high D). Clearly the most important requirement of any bioreactor control system is to avoid washout and maintain bioreactor productivity.

Non-trivial steady-state solutions of cell population models are more difficult to analyze. Closed-form solutions can be obtained using the method of characteristics if restrictive assumptions are imposed on the cell cycle and/or the culture environment (Hjortso, 1996; Hjortso and Nielsen, 1995). This approach has been used to analyze local stability of steady-state solutions for an age structured cell population model (Hjortso and Bailey, 1983). A more practical approach for local stability analysis involves spatial discretization of the cell population model to obtain a coupled set of nonlinear ordinary differential equations in time (Zhang et al., 2001). Steady-state solutions are calculated by solving the nonlinear algebraic equations which comprise the steady-state version of the discretized model. Local stability of a steady-state solution is analyzed by linearizing the discretized model about the steady-state operating point and computing the eigenvalues of the linearized model. Non-local stability analysis typically requires dynamic simulation of the discretized model. A secondary control objective may be stabilization of a particular cell mass distribution $N(m)$ that optimizes the steady-state production of certain products.

Experimental studies with different microorganisms have shown that continuous bioreactors can exhibit stable periodic solutions which are observable as sustained oscillations in measured variables (Daugulis et al., 1997; Jones, 1995; von Meyenburg, 1973). Several investigators have shown that cell population models are capable of generating such periodic solutions (Bellgardt, 1994; Hjortso and Nielsen, 1995; Zhu et al., 2000). Closed-form representation of periodic solutions have been derived directly from cell population models under certain simplifying assumptions (Hjortso and Nielsen, 1995). We are not aware of any analysis results concerning the stability of such periodic solutions. As discussed below for yeast bioreactors, periodic solutions usually are located

by dynamic simulation of a spatially discretized model. Another possible control objective is creation of periodic solutions that lead to increased production of certain products as compared to that achievable under steady-state conditions (Hjortso, 1996).

Dynamics of Continuous Yeast Bioreactors

Cell population dynamics play a key role in the sustained oscillations observed in continuous bioreactors producing Baker's yeast. Fundamental understanding of these dynamics could lead to important advances in yeast production processes and provide key insights into the cellular behavior of more complex cells present in plants and animals. Several investigators (Munch et al., 1992; Strassle et al., 1989) have shown that the appearance of sustained oscillations is related to the formation of distinct cell subpopulations via a mechanism known as cell cycle synchrony. A synchronized culture is recognized by well defined peaks in the cell distribution that correspond to large groups of cells that collectively pass through the cell cycle.

Recently it has been proposed that continuous yeast bioreactors can exhibit a stable steady state and a stable limit cycle at the same operating conditions as a consequence of cell cycle synchrony (Zhang et al., 2001). Experimental data that support this claim are shown in Figure 3 where the carbon dioxide off-gas concentration is used as a representative output signal. The experimental protocol used involves careful manipulation of the dilution rate to establish different initial conditions for the cell distribution. An initial condition corresponding to a synchronized cell population results in convergence to a stable limit cycle (top plot). By contrast, a steady-state solution appears to be attained when the initial cell population is less synchronized (bottom plot).

The experimental data in Figure 3 show that the open-loop dynamics of yeast bioreactors are strongly dependent on the initial condition of the cell distribution. At a particular value of the dilution rate there appears to be two stable solutions, each with a domain of attraction that is a complex function of the initial cell distribution. This interpretation provides a rational explanation for the observation that oscillations appear and disappear without measurable changes in external inputs such as dilution rate and feed substrate concentration (Parulekar et al., 1986). Moreover this nonlinear behavior would be fundamentally different than that observed in other particulate processes such as emulsion polymerization reactors (Rawlings and Ray, 1987) and solution crystallizers (Witkowski and Rawlings, 1987) that exhibit sustained oscillations as the steady-state solution becomes unstable.

A more detailed understanding of the nonlinear dynamics leading to sustained oscillations can be obtained via bifurcation analysis (Kuznetsov, 1995). A bifurcation represents a fundamental change in the qualitative

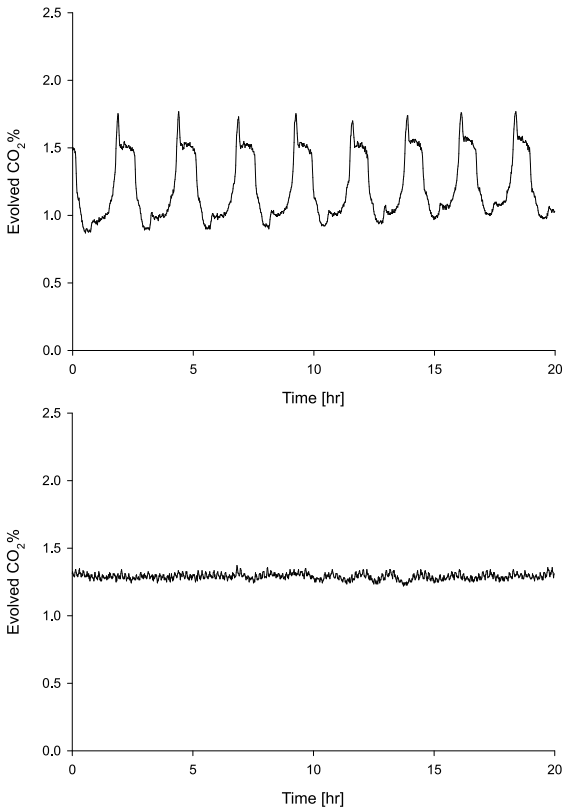


Figure 3: Multiple stable solutions for a yeast bioreactor.

behavior of a nonlinear system as a parameter is varied. The most common example is the Hopf bifurcation where the steady state becomes unstable and a stable limit cycle is created (Kuznetsov, 1995).

We have performed bifurcation analysis using the unstructured cell population model (11)–(13) with specific forms of the physiological state functions (Zhang et al., 2001). The single cell growth rate is modeled as:

$$r(m, S') = \frac{\mu_m S'}{K_m + S'} \tag{17}$$

where μ_m and K_m are constants. The effective substrate concentration S' is a filtered version of the actual substrate concentration (S) and accounts for the lagged response of cells to environmental changes. The growth rate function models the tendency of cells to reach a maximum growth rate (μ_m) at large substrate concentrations (substrate inhibition). The division rate function is modeled as:

$$\Gamma(m, S') = \begin{cases} 0 & m \leq m_t^* + m_o \\ \gamma \exp\left[-\epsilon(m - m_d^*)^2\right] & m \in [m_t^* + m_o, m_d^*] \\ \gamma & m \geq m_d^* \end{cases} \tag{18}$$

where m_t^* is the transition mass (see Figure 2), m_o is the

additional mass that mother cells must gain before division is possible, ϵ and γ are constants and m_d^* is the mass at which the division rate reaches its maximum value γ. This function models the tendency of cells to divide near the division mass (m_d^*). The cell cycle parameters (m_t^*, m_d^*) are functions of S' as discussed below.

The partition probability density function has the form:

$$p(m, m', S') = A\exp[-\beta(m - m_t^*)^2] + A\exp[-\beta(m - m' + m_t^*)^2] \quad (19)$$

where $m < m'$ and $m' > m_t^* + m_o$; the function is identically zero otherwise. Here A and β are constants. This function yields two Gaussian peaks in the cell mass distribution, one centered at m_t^* (corresponding to mother cells) and one centered at a location in the mass domain that is determined by mass conservation (corresponding to daughter cells). The substrate consumption rate is modeled as:

$$q(m, S') = \frac{1}{Y} r(m, S') \quad (20)$$

where Y is a constant yield parameter. The substrate dependence of the cell cycle parameters is modeled as:

$$m_t^*(S') = \begin{cases} m_{t0} + K_t(S_l - S_h) & S' < S_l \\ m_{t0} + K_t(S' - S_h) & S' \in [S_l, S_h] \\ m_{t0} & S' > S_h \end{cases} \quad (21)$$

$$m_d^*(S') = \begin{cases} m_{d0} + K_d(S_l - S_h) & S' < S_l \\ m_{d0} + K_d(S' - S_h) & S' \in [S_l, S_h] \\ m_{d0} & S' > S_h \end{cases} \quad (22)$$

where S_l, S_h, m_{t0}, m_{d0}, K_t and K_d are constants. As found experimentally (Alberghina et al., 1991), both m_t^* and m_d^* are increasing functions of the substrate concentration. Numerical values of the model parameters are presented elsewhere (Zhang et al., 2001).

The dilution rate (D) is chosen as the bifurcation parameter. Stability of steady-state solutions is determined by checking the eigenvalues of the Jacobian linearization. Stable limit cycles are located using a combination of dynamic simulation and continuation calculations (Kuznetsov, 1995). The resulting bifurcation diagram is shown in Figure 4 where the zeroth moment of the cell mass distribution (here denoted as m_0) is used as a representative output variable. As observed experimentally (Beuse et al., 1998), a stable steady state (+) is obtained for low and high ranges of the dilution rate. A Hopf bifurcation occurs at $D = 0.21$ h^{-1} that results in the appearance of a stable limit cycle with sustained oscillations of the magnitude indicated (*). A second Hopf bifurcation occurs at $D = 0.27$ h^{-1} where the limit cycle disappears. An unstable steady state (o) is observed at the intermediate dilution rates that support stable limit cycles.

The bifurcation diagram appears to be inconsistent with the experimental data in Figure 3 which indicate the coexistence of stable steady state and stable periodic solutions. However the model predicts that destabilization of the steady state occurs very slowly due to the small real parts of the eigenvalues that cross the imaginary axis due to the first Hopf bifurcation. We conjecture that this behavior is not observed in Figure 3 due to the relatively short duration of the experimental test. This hypothesis is supported by the small oscillations that are visible in the "stationary" response. This subtle dynamic behavior cannot be captured by unsegregated models such as cybernetic models (Jones and Kompala, 1999).

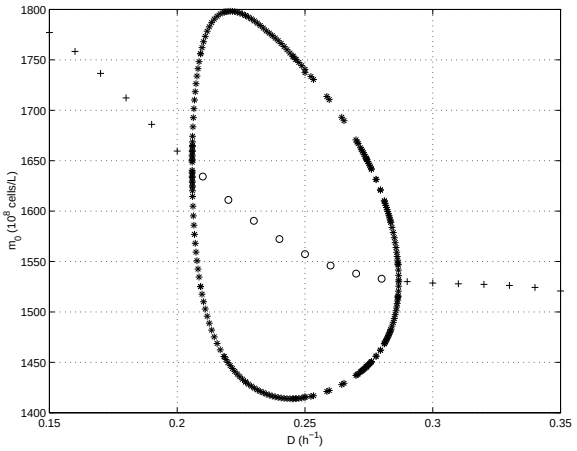


Figure 4: Bifurcation diagram for a yeast bioreactor model.

Feedback Control of Cell Populations

Feedback control is necessary to ensure satisfactory performance of continuous bioreactors in the presence of external disturbances and/or changes in the operational requirements. As depicted in Figure 1, a typical bioreactor control system consists of simple regulatory loops for residence time, temperature and pH designed to maintain the bioreactor at environmental conditions which promote cell growth. Such simple schemes do not allow direct control of variables such as the biomass concentration that determine profitability of the bioprocess. During the last decade substantial effort has gone into developing more advanced (nonlinear or adaptive) control strategies for continuous bioreactors (Bastin and Dochain, 1990; Hoo and Kantor, 1986; El Moubaraki et al., 1993; Pons, 1992; Kurtz et al., 2000). These efforts are based on continuum models that neglect the distributed nature of the cell population. As such, they rely on measurement and control of 'average' properties of the cell populations in the bioreactor. A typical control strategy along these lines involves manipulation of the dilution rate or the feed substrate concentration to maximize biomass concentration (Henson and Seborg, 1992; Proll and Karim, 1994).

The last decade has also seen the evolution of experimental techniques, specifically flow cytometry and cell staining techniques (Srienc, 1993), which enable the measurement of entire cell property distributions. Flow cytometers measure the frequency of fluorescence in the cell population, and hence can differentiate cells with respect to naturally fluorescent protein content (e.g. the green fluorescent protein, Gfp), or other variables (e.g. DNA content) after appropriate staining. When interfaced with proper flow injection systems (Zhao et al., 1999) they provide a powerful experimental tool for on-line monitoring and control. These advances in instrumentation and measuring devices, together with the advances in the numerical solution of cell population balance models outlined earlier, provide strong motivation to explore more advanced control strategies that utilize cell population balance models to address a wider range of control objectives, e.g. control of cell property distributions and/or cell cycle characteristics.

This realization has motivated research in our groups on the development of control strategies based on cell mass population balance models that address a variety of control objectives. Specifically, in Zhang et al. (2001) the problem of attenuating open-loop oscillations observed in yeast bioreactors was addressed, through the design of a feedback linearizing controller that manipulates the dilution rate and controls the zeroth moment of the cell mass distribution. The design of distributed feedback linearizing control laws that manipulate the dilution rate to influence the zeroth and first moment of cell mass distributions was also addressed in Mantzaris et al. (1998). Zhu et al. (2000) addressed the attenuation of oscillations as well as the induction of oscillations in yeast bioreactors, using a linear model predictive control strategy that manipulates the dilution rate and the feed substrate concentration. Finally, Mantzaris et al. (1999) addressed the productivity control for a cell culture where the desired product is produced only in the second stage of the cell cycle, using a feedback linearizing control strategy that manipulates the feed substrate concentration. In what follows, we briefly outline the last two control studies in order to demonstrate the potential advantages of using cell population models as the basis for controller design. More details on the controller design and additional simulation results are available in the original references.

Oscillation Attenuation in Yeast Bioreactors

The first case study addresses the control of oscillations in continuous yeast bioreactors. The motivation is provided by the fact that open-loop oscillations can adversely affect bioreactor stability and productivity, in which case it is imperative that they be attenuated. In other cases, inducing stable oscillatory behavior may lead to increased production of target metabolites preferentially produced during part of the cell cycle (Hjortso, 1996). Below we outline a linear model predictive control (LMPC) strategy which is well suited for both of the above control problems.

The controller design model is generated from the cell population model (11)–(13) and the physiological state functions (17)–(20). The population model is discretized in the mass domain using orthogonal collocation on finite elements, linearized about the desired steady-state operating point and discretized with a sampling time Δt = 0.1 hr. The resulting linear model has the form:

$$x(k+1) = Ax(k) + Bu(k) \tag{23}$$

where: x is the state vector which includes the cell mass density N_j at each collocation point j and the substrate concentration S; and u is the input vector comprised of the dilution rate D and feed substrate concentration S_f. It is assumed that the cell mass distribution can be measured via flow cytometry or reconstructed from on-line measurements of the particle size distribution (Yamashit et al., 1993).

The controller design model is completed by defining the output vector. An obvious approach is to choose the discretized cell mass densities N_j as the controlled outputs. This method is problematic because: (*i*) the resulting control problem is highly non-square (2 inputs, 113 outputs); (*ii*) cell mass densities at nearby collocation points are strongly collinear; and (*iii*) the substrate concentration must be controlled to avoid washout. We have found that good closed-loop performance can be obtained by controlling a subset of the cell mass densities and the substrate concentration:

$$y(k) = \left[\begin{array}{cccc} N_{j_1}(k) & \dots N_{j_p}(k) & S(k) \end{array}\right]^T = Cx(k) \tag{24}$$

where the indices $\{j_1, \dots, j_p\}$ denote the collocation points where the associated cell mass density is used as a controlled output. In the subsequent simulations, the outputs are chosen as the boundary points of the finite elements. This results in a much lower dimensional problem with 14 outputs.

The LMPC controller is formulated as an infinite horizon open-loop optimal control problem:

$$\begin{aligned} \min_{U_M(k)} \sum_{j=0}^{\infty} \{ & [y(k+j|k) - y_s]^T Q [y(k+j|k) - y_s] \\ & + [u(k+j|k) - u_s]^T R [u(k+j|k) - u_s] \\ & + \Delta u^T (k+j|k) S \Delta u(k+j|k) \} \end{aligned} \tag{25}$$

where: $y(k+j|k)$ and $u(k+j|k)$ are predicted values of the outputs and inputs, respectively; y_s and u_s are target values for the outputs and inputs, respectively; and $\Delta u(k) = u(k) - u(k-1)$. The decision variables are current and future values of the inputs: $U_M(k) = [u(k|k) \ \dots u(k+M-1|k)]$, where M is the control horizon. The inputs are subject to constraints that are determined by operational limitations: $u_{min} \le u \le u_{max}$.

The resulting problem can be reformulated as a finite horizon problem and solved using standard quadratic programming software (Muske and Rawlings, 1993). The control horizon is chosen as $M = 5$ and the weighting matrices (Q, R, S) are chosen by trial-and-error to provide acceptable closed-loop performance.

Figure 5 shows the ability of the LMPC controller to stabilize an oscillating bioreactor at a desired steady-state operating point. The zeroth-order moment of the cell mass distribution and the substrate concentration (S) are shown as representative output variables. The initial cell mass distribution $N(m, 0)$ corresponds to a stable periodic solution, while the setpoint vector is obtained from the discretized cell mass distribution at the desired steady-state operating point. The solid line is the LMPC response and the dashed line is the open-loop response in the absence of feedback control. The LMPC controller effectively attenuates the oscillations while generating reasonable control actions. The evolution of the cell mass distribution (here denoted as W) also is shown in Figure 5. The initial distribution is highly synchronized with two cell subpopulations that produce sustained oscillations. The controller achieves the desired steady-state distribution by dispersing the subpopulations.

Figure 6 shows the ability of the LMPC controller to create a desired periodic solution. The initial mass number distribution corresponds to the steady-state solution in Figure 5. The stable periodic solution shown as the open-loop response in Figure 5 is used to define a time-varying setpoint trajectory to be tracked. The controller stabilizes the periodic solution by generating oscillatory input moves. Although not shown, it is interesting to note that the oscillations are sustained with the same period when the controller is switched off at the end of the simulation and the bioreactor runs under open-loop conditions. The evolution of the cell mass distribution also is shown in Figure 6. Note that the oscillating dynamics are accompanied by the appearance of two distinct cell subpopulations. These results suggest that feedback control strategies which provide direct control of the cell distribution have the potential to enhance the stability and productivity of continuous yeast bioreactors.

Productivity Control in Two-staged Cell Growth

In the second case study, we address the problem of controlling the productivity with respect to a desired product in a continuous bioreactor where the cells grow in two stages, with the desired product being formed only in the second stage (this is consistent with the discussion in page 4 on the behavior of many microorganisms including yeast). Individual cells are differentiated with respect to their mass, m (or any other variable that is conserved). There is a single substrate with concentration S, whereas the product concentration is denoted by P.

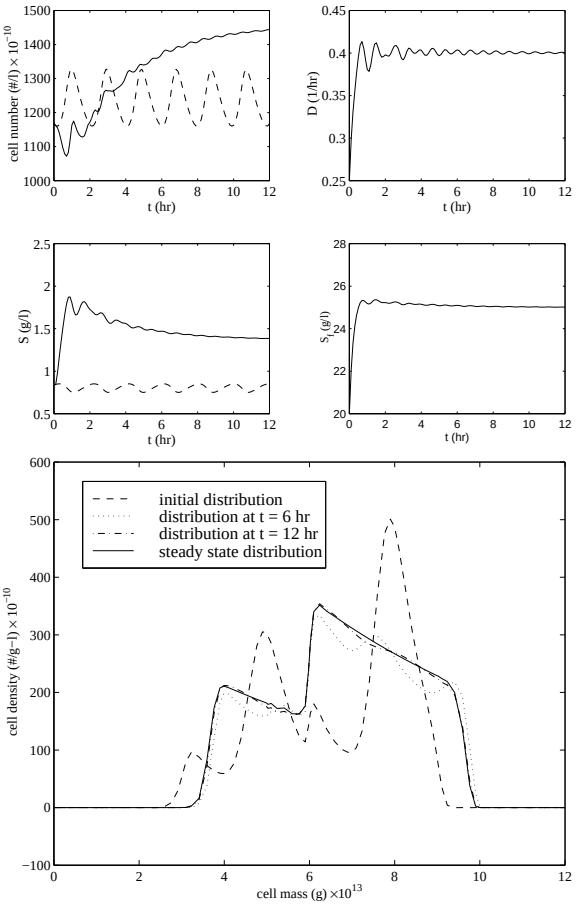


Figure 5: Oscillation attentuation with LMPC control.

It is assumed that the single-cell growth rates are linear with respect to cell mass and exhibit substrate inhibition during both stages of the cell cycle, whereas during the second stage the growth rate exhibits product inhibition as well (this is a rather standard assumption, see e.g. Henson and Seborg, 1992). Specifically, the single-cell growth rates are expressed as:

$$r_1(m, S) = \frac{\mu_m S}{K_m + S + \frac{S^2}{K_i}} m \tag{26}$$

$$r_2(m, S, P) = \frac{\mu_m \left(1 - \frac{P}{P_m}\right) S}{K_m + S + \frac{S^2}{K_i}} m \tag{27}$$

where μ_m, P_m, K_m, K_i are the maximum specific growth rate, the product and substrate saturation constants and the substrate inhibition constant, respectively.

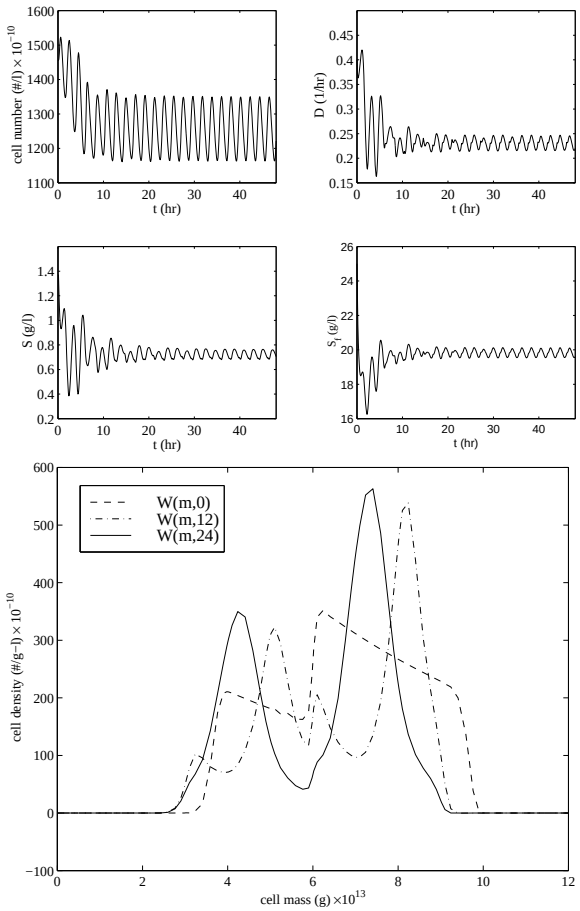


Figure 6: Oscillation stabilization with LMPC control.

The transition rates from stage 1 to stage 2, and from stage 2 to stage 1 are modeled as follows (Fredrickson et al., 1967):

$$\Gamma_1(m,S) = \frac{f_1(m)}{1-\int_0^m f_1(m')\,dm'} r_1(m,S) \tag{28}$$

$$\Gamma_2(m,S,P) = \frac{f_2(m)}{1-\int_0^m f_2(m')\,dm'} r_2(m,S,P) \tag{29}$$

where $f_1(m), f_2(m)$ are the transition probability density functions which are assumed to depend only on mass; these functions are taken to be Gaussian distributions with mean values μ_{f_1}, μ_{f_2} and standard deviations $\sigma_{f_1}, \sigma_{f_2}$, respectively.

The partitioning function is assumed to be independent of the substrate and product concentrations, and is taken to be a symmetric beta distribution with a parameter q:

$$p(m,m') = \frac{1}{B(q,q)} \frac{1}{m'} \left(\frac{m}{m'}\right)^{q-1} \left(1-\frac{m}{m'}\right)^{q-1} \tag{30}$$

The substrate consumption rates during the two stages of the cell cycle are expressed as:

$$q_1(m,S) = \frac{1}{Y_1} r_1(m,S) \tag{31}$$

$$q_2(m,S,P) = \frac{1}{Y_2} r_2(m,S,P) \tag{32}$$

where Y_1, Y_2 denote constant yield coefficients. Finally, the product formation rate is expressed as:

$$r_p(m,S,P) = a(\mu_2(S,P)+b)m \tag{33}$$

where a, b are constants. The parameter values used can be found in Mantzaris et al. (1999).

The dynamic model of the reactor consists of the two cell population balance equations for stages 1 and 2, the substrate balance and the product balance:

$$\begin{aligned} &\frac{\partial N_1(m,t)}{\partial t} + \frac{\partial [r_1(m,S)N_1(m,t)]}{\partial m} \\ &\qquad + \Gamma_1(m,S)N_1(m,t) + DN_1(m,t) \\ &= 2\int_m^{m_{max}} \Gamma_2(m,S,P)p(m,m',S)N_2(m,t)\,dm' \end{aligned} \tag{34}$$

$$\begin{aligned} &\frac{\partial N_2(m,t)}{\partial t} + \frac{\partial [r_2(m,S,P)N_2(m,t)]}{\partial m} \\ &\qquad + \Gamma_2(m,S,P)N_2(m,t) + DN_2(m,t) \\ &\qquad = \Gamma_1(m,S,P)N_1(m,t) \end{aligned} \tag{35}$$

$$\begin{aligned} \frac{dS}{dt} &= D(S_f - S) - \frac{1}{Y_1}\int_0^{m_{max}} r_1(m,S)N_1(m,t)\,dm \\ &\quad - \frac{1}{Y_2}\int_0^{m_{max}} r_2(m,S,P)N_2(m,t)\,dm \end{aligned} \tag{36}$$

$$\frac{dP}{dt} = -DP + \int_0^{m_{max}} r_p(m,S,P)N_2(m,t)\,dm \tag{37}$$

with the following boundary conditions for the two population balance equations (see Mantzaris et al., 1999, for their derivation):

$$\int_0^{m_{max}} \frac{\partial [r_1(m,S)N_1(m,t)]}{\partial m}\,dm = 0 \tag{38}$$

$$\int_0^{m_{max}} \frac{\partial [r_2(m,S,P)N_2(m,t)]}{\partial m}\,dm = 0 \tag{39}$$

For the numerical solution and controller design, the two cell population balance equations were discretized in the mass space using a Galerkin spectral method. Specifically, the stage 1 and 2 mass distributions were expanded as follows:

$$\begin{aligned} N_1(m,t) &= \Sigma_{i=1}^{\infty} a_i(t)\phi_i(m) \\ N_2(m,t) &= \Sigma_{i=1}^{\infty} b_i(t)\phi_i(m) \end{aligned} \tag{40}$$

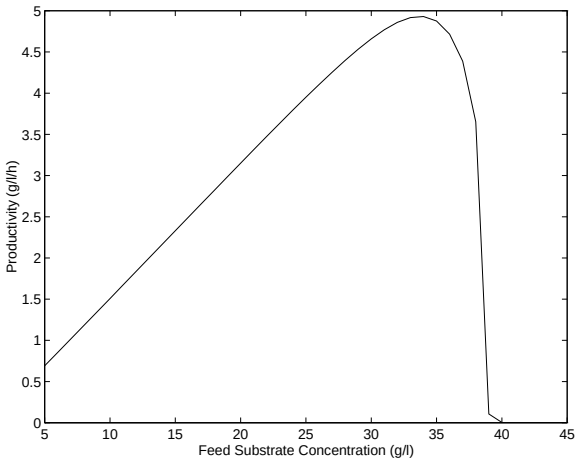


Figure 7: Steady state productivity vs. feed substrate concentration.

where $\phi_i = \sqrt{2}sin(i\pi m)$ and a_i, b_i denote the time dependent coefficients of the sine expansions (it can be easily verified that with these basis functions the boundary conditions are satisfied). Following the usual procedure of substituting the expansions to the partial differential equations and taking the inner product with the adjoint functions, an infinite set of ODEs for the time varying coefficients can be obtained. To obtain a finite-dimensional approximation of the infinite dimensional model, the infinite series expansion was truncated to include $n = 20$ terms.

Open-loop Behavior. Initially, the effect of the dilution rate D and the feed substrate concentration S_f on the steady state productivity DP was analyzed. The analysis established that there is a pair of these operating parameters, $D = 0.18h^{-1}$ and $S_f = 34g/l$, where the productivity is maximized (the maximum is approximately $DP_{max} = 4.93g/l/h$). Figure 7 shows a plot of the productivity as a function of the feed substrate concentration at steady state, for $D = 0.18h^{-1}$. The occurrence of a maximum in the productivity is consistent with the results obtained in Henson and Seborg (1992) which considered a continuum model consisting of biomass, substrate and product balances, with the same functions and parameters for cell growth, substrate consumption and product formation. This behavior is also indicative of the nonlinearity of the system and motivates the design of nonlinear controllers to maintain the productivity close to its maximum.

Nonlinear Productivity Control. The control study focused on controlling the productivity $y = DP$ close to its maximum value $y_{sp} = 4.93g/l/h$ using the feed substrate concentration S_f as manipulated input. The dilution rate was fixed at $D = 0.18h^{-1}$, and hence the control strategy essentially aimed at maximizing the product concentration.

The finite-dimensional approximation of the process model was used as the basis for the controller design. Specifically, the relative degree of this model was found to be two, as long as the reactor operates away from washout conditions and from a manifold which intersects the equilibrium curve of Figure 7 approximately at the maximum productivity (note that due to the complexity of the model one has to rely on numerical approximations for the above observations). Also, the zero dynamics with respect to the output variable was numerically found to be locally asymptotically stable at the setpoint conditions.

A nonlinear state feedback controller was designed that induces the following linear response in the closed-loop system:

$$\gamma_2 \frac{d^2y}{dt^2} + \gamma_1 \frac{dy}{dt} + y = y_{sp} \tag{41}$$

For comparison purposes, the approximate finite dimensional model was also linearized (around the steady state conditions corresponding to the setpoint), and a linear state feedback controller was designed on the basis of the resulting linear model to enforce the same behavior as above.

Figure 8 shows the results of a representative simulation run for $\gamma_1 = \gamma_2 = 10$. The initial cell mass distributions in the two stages, and the substrate and product concentrations were obtained from the steady state corresponding to $S_f = 25g/l$. The corresponding initial productivity was $y = 3.95g/l/h$, which is approximately 20% smaller than the setpoint value. Notice that the nonlinear controller induces the desired closed-loop response and smoothly brings the system to the desired setpoint. On the other hand, the output response under the linear controller is considerably more sluggish, taking almost twice as long to bring the productivity to its setpoint. Further, the manipulated input for the linear controller exhibits a much larger peak in the initial part of the response. For larger deviations from the setpoint, the linear controller led to closed-loop instability, whereas the nonlinear controller continued to perform very satisfactorily (successful results were obtained with the initial productivity being as much as seven times smaller than the maximum productivity).

A second simulation run addressed the disturbance rejection properties of the two controllers. Specifically, the actual value of the maximum specific growth rate was assumed to be approximately 10% smaller than the nominal one. The productivity setpoint was chosen as $y_{sp} = 4.43g/l/h$ which corresponds to the approximate maximum productivity for the actual value of the maximum specific growth rate. Integral action was incorporated in both controllers. Figure 9 shows the controlled output and manipulated input responses. The nonlinear controller exhibits again a faster response, with smaller

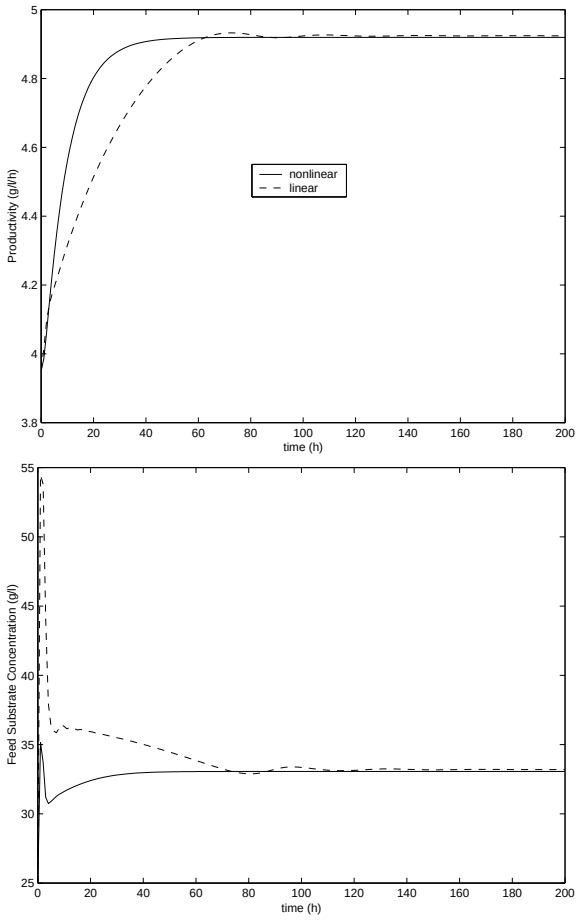


Figure 8: Closed-loop responses under linear and nonlinear controllers—setpoint tracking.

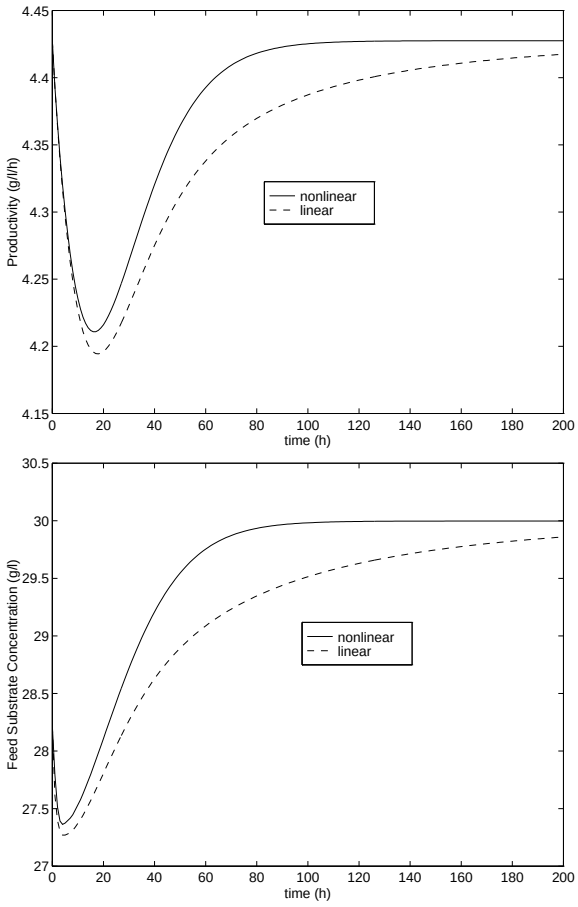


Figure 9: Closed-loop responses under linear and nonlinear controllers—disturbance rejection.

deviations from the setpoint compared to the linear controller.

Concluding Remarks and Future Research

Cell growth systems are characterized by overwhelming complexity and variety. The dynamics of such systems can be described at various levels of detail. In this paper, we focused on continuous reactors used for growth of microbial populations, and presented an overview of recent results in dynamical analysis and control which account directly for the heterogeneous nature of cell populations. These results illustrate the feasibility and advantages of using cell population balance models as the basis for feedback control of bioreactors, and in our view, make a clear case for further research towards the development and practical application of such bioreactor control strategies. The opportunities for scientific and engineering contributions, from a systems and control perspective, in this direction are numerous. Some of these are outlined below.

Cell Population Modeling

The development and validation of cell population balance models for cultures of specific microorganisms (such as yeast), through the combination of flow cytometric measurements and inverse population balance modeling techniques is a key research goal. Analyzing the effect of environmental parameters such pH, temperature etc. on these functions may also enable using such parameters as additional manipulated variables to achieve a broader range of control objectives, e.g. simultaneously controlling multiple moments of cell property distributions. Finally, the development of structured cell population balance models for specific cultures is the ultimate goal in this direction; such models, although invariably complex, can further broaden the range of control objectives that can be achieved, e.g. controlling metabolic pathways at such a distributed level.

Dynamic Analysis of Cell Population Models

Despite the fact that cell population balance models have been available for over thirty years, the literature on their dynamics is very sparse. Important open theoreti-

cal questions include the existence and stability of both steady-state and periodic solutions. Addressing these questions, either on the basis of the PDE models themselves or on the basis of ODE approximations obtained from spatial discretizations is a key research task. Dynamic simulation is another important tool to address these problems and to investigate other control relevant dynamics such as bifurcations between solutions. Unstructured models are a reasonable starting point for such studies owing to their ability to capture population dynamics with minimal complexity. Analysis of structured models is a considerably more difficult problem but offers the potential to enhance understanding of the complex interactions between individual cell metabolism and cell population dynamics.

Cell Distribution Control

Further automating and refining on-line flow cytometry and cell staining techniques, to be able to obtain rapid and robust measurements of desired cell property distributions, is clearly an essential task towards the practical application of control laws derived on the basis of cell population balance models. At the level of controller design, translating general operational objectives into specific control objectives involving the measured distribution properties, and evaluating a wide variety of controller design methods with regard to their suitability for these objectives are clearly important tasks that need to be addressed for specific cultures. Integrating sensor development and control algorithm development through laboratory experiments is also necessary to be able to prototype control systems suitable for industrial applications.

In closing we note that much like the results presented in this paper are the outcome of fruitful collaborations between biochemical engineers and control engineers in our respective groups, the future research goals outlined above can be most effectively pursued through such cross-disciplinary collaborations.

Acknowledgments

We would like to acknowledge Prof. Friedrich Srienc (Minnesota), Prof. Nikolaos Mantzaris (Rice), and Prof. Martin Hjortso (LSU) for their contributions in the research that generated the results outlined in the paper, as well as for their comments and suggestions on the paper itself. We would also like to acknowledge Prof. Ioannis Kevrekidis and Krishnan Sankaranarayanan (Princeton) for their assistance with the bifurcations calculations.

References

Alberghina, L., B. M. Ranzi, D. Porro, and E. Martegani, "Flow Cytometry and Cell Cycle Kinetics in Continuous and Fed-Batch Fermentations of Budding Yeast," *Biotechnol. Prog.*, **7**, 299–304 (1991).

Bastin, G. and D. Dochain, *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier Science Publishers (1990).

Bellgardt, K.-H., "Analysis of Synchronous Growth of Baker's Yeast: Part I: Development of a Theoretical Model for Sustained Oscillations," *J. Biotech.*, **35**, 19–33 (1994).

Beuse, M., R. Bartling, A. Kopmann, H. Diekmann, and M. Thoma, "Effect of Dilution Rate on the Mode of Oscillation in Continuous Culture of *Saccharomyces cerevisiae*," *J. Biotech.*, **61**, 15–31 (1998).

Daugulis, A. J., P. J. McLellan, and J. Li, "Experimental Investigation and Modeling of Oscillatory Behavior in the Continuous Culture of *Zymomonas mobilis*," *Biotech. Bioeng.*, **56**, 99–105 (1997).

Eakman, J. M., A. G. Fredrickson, and H. M. Tsuchiya, "Statistics and Dynamics of Microbial Cell Populations," *Chem. Eng. Prog.*, **62**, 37–49 (1966).

El Moubaraki, J., G. Bastin, and L. Levine, "Nonlinear Control of Biotechnological Processes with Growth-Production Decoupling," *Math. Biosciences*, **116**, 21–44 (1993).

Fredrickson, A. G. and N. V. Mantzaris, "A New Set of Population Balance Equations for Microbial and Cell Cultures," *Chem. Eng. Sci.* (2002). In press.

Fredrickson, A. G., D. Ramkrishna, and H. M. Tsuchyia, "Statistics and Dynamics of Procaryotic Cell Populations," *Math. Biosciences*, **1**, 327–374 (1967).

Frykman, S., *Single Cell Secretion Rates of Saccharomyces cerevisiae*, PhD thesis, University of Minnesota (1999).

Godin, F. B., D. G. Cooper, and A. D. Rey, "Development and Solution of a Cell Mass Population Balance Model Applied to the SCF Process," *Chem. Eng. Sci.*, **54**, 565–578 (1999a).

Godin, F. B., D. G. Cooper, and A. D. Rey, "Numerical Methods for a Population-balance Model of a Periodic Fermentation Process," *AIChE J.*, **45**, 1359–1364 (1999b).

Hatzis, C., F. Srienc, and A. G. Fredrickson, "Multistaged Corpuscular Models of Microbial Growth: Monte Carlo Simulations," *Biosystems*, **36**, 19–35 (1995).

Henson, M. A. and D. E. Seborg, "Nonlinear Control Strategies for Continuous Fermentors," *Chem. Eng. Sci.*, **47**, 821–835 (1992).

Hjortso, M. A. and J. E. Bailey, "Transient Responses of Budding Yeast Populations," *Math. Biosciences*, **63**, 121–148 (1983).

Hjortso, M. A. and J. Nielsen, "A Conceptual Model of Autonomous Oscillations in Microbial Cultures," *Chem. Eng. Sci.*, **49**, 1083–1095 (1994).

Hjortso, M. A. and J. Nielsen, "Population Balance Models of Autonomous Microbial Oscillations," *J. Biotech.*, **42**, 255–269 (1995).

Hjortso, M. A., "Population Balance Models of Autonomous Periodic Dynamics in Microbial Cultures: Their Use in Process Optimization," *Can. J. Chem. Eng.*, **74**, 612–620 (1996).

Hoo, K. A. and J. C. Kantor, "Global Linearization and Control of a Mixed-Culture Bioreactor with Competition and External Inhibition," *Math. Biosciences*, **82**, 43–62 (1986).

Jones, K. D. and D. S. Kompala, "Cybernetic Modeling of the Growth Dynamics of *Saccharomyces cerevisiae* in Batch and Continuous Cultures," *J. Biotech.*, **71**, 105–131 (1999).

Jones, K. D., *Oscillations in Continuous Cultures of Microorganisms: Criteria for Utility of Mathematical Models*, PhD thesis, University of Colorado, Boulder, CO (1995).

Kim, M.-Y. and E. J. Park, "Mixed Approximation of a Population Diffusion Equation," *Comp. Math. Appl.*, **30**, 5–17 (1995a).

Kim, M.-Y. and E. J. Park, "An Upwind Scheme for a Nonlinear Model in Age-Structured Population Dynamics," *Comp. Math. Appl.*, **30**, 95–103 (1995b).

Kim, M.-Y., "Galerkin Methods for a Model of Population Dynamics with Nonlinear Diffusion," *Numerical Methods for Partial Differential Equations*, **12**, 59–73 (1996).

Kompala, D., N. Jansen, G. Tsao, and D. Ramkrishna, "Investigation of Bacterial Growth on Mixed Substrates: Experimental Evaluation of Cybernetic Models," *Biotech. Bioeng.*, **28**, 1044–1055 (1986).

Kostova, T. V., "Numerical Solutions of a Hyperbolic Differential-Integral Equation," *Comp. Math. Appl.*, **15**, 427–436 (1988).

Kostova, T. V., "Numerical Solutions to Equations Modelling Nonlinearly Interacting Age-Dependent Populations," *Comp. Math. Appl.*, **19**, 95–103 (1990).

Kromenaker, S. J. and F. Srienc, "Effect of Lactic Acid on the Kinetics of Growth and Antibody Production in a Murine Hybridoma: Secretion Patterns During the Cell Cycle," *J. Biotech.*, **34**, 13–34 (1994).

Kurtz, M. J., G.-Y. Zhu, A. Zamamiri, M. A. Henson, and M. A. Hjortso, "Control of Oscillating Microbial Cultures Described by Population Balance Models," *Ind. Eng. Chem. Res.*, **37**, 4059–4070 (1998).

Kurtz, M. J., M. A. Henson, and M. A. Hjortso, "Nonlinear Control of Competitive Mixed-Culture Bioreactors via Specific Cell Adhesion," *Can. J. Chem. Eng.*, **78**, 237–247 (2000).

Kuznetsov, Y. A., *Elements of Applied Bifurcation Theory*. Springer-Verlag, New York, NY (1995).

Lee, J. M., *Biochemical Engineering*. Prentice-Hall, Englewood Cliffs, NJ (1992).

Liou, J. J., F. Srienc, and A. G. Fredrickson, "Solutions of Population Balance Models based on a Successive Generation Approach," *Chem. Eng. Sci.*, **9**, 1529–1540 (1997).

Mantzaris, N. V., F. Srienc, and P. Daoutidis, Control of Cell Mass Distribution in Continuous Bioreactors Using Population Balance Models, In *Proceedings of DYCOPS-6*, pages 602–607, Corfu, Greece (1998).

Mantzaris, N. V., J. J. Liou, P. Daoutidis, and F. Srienc, "Numerical Solution of a Mass Structured Cell Population Balance Model in an Environment of Changing Substrate Concentration," *J. Biotech.*, **71**, 157–174 (1999).

Mantzaris, N. V., P. Daoutidis, and F. Srienc, "Numerical Solution of Multivariable Cell Population Balance Models. I: Finite Difference Methods," *Comput. Chem. Eng.*, **25**, 1411–1440 (2001a).

Mantzaris, N. V., P. Daoutidis, and F. Srienc, "Numerical Solution of Multivariable Cell Population Balance Models. II: Spectral Methods," *Comput. Chem. Eng.*, **25**, 1441–1462 (2001b).

Mantzaris, N. V., P. Daoutidis, and F. Srienc, "Numerical Solution of Multivariable Cell Population Balance Models. III: Finite Element Methods," *Comput. Chem. Eng.*, **25**, 1463–1481 (2001c).

Mantzaris, N. V., F. Srienc, and P. Daoutidis, "Nonlinear Productivity Control using a Multi-Staged Cell Population Balance Model," *Chem. Eng. Sci.*, **57**, 1–14 (2002).

Mauch, K., S. Arnold, and M. Reuss, "Dynamic Sensitivity Analysis for Metabolic Systems," *Chem. Eng. Sci.*, **52**, 2589–2598 (1997).

Munch, T., B. Sonnleitner, and A. Fiechter, "New Insights into the Synchronization Mechanism with Forced Synchronous Cultures of *Saccharomyces cerevisiae*," *J. Biotech.*, **24**, 299–313 (1992).

Muske, K. R. and J. B. Rawlings, "Model Predictive Control with Linear Models," *AIChE J.*, **39**, 262–287 (1993).

Nielsen, J. and J. Villadsen, "Modelling of Microbial Kinetics," *Chem. Eng. Sci.*, **47**, 4225–4270 (1992).

Nielsen, J. and J. Villadsen, *Bioreaction Engineering Principles*. Plenum Press, New York, NY (1994).

Parulekar, S. J., G. B. Semones, M. J. Rolf, J. C. Lievense, and H. C. Lim, "Induction and Elimination of Oscillations in Continuous Cultures of *Saccharomyces Cerevisiae*," *Biotech. Bioeng.*, **28**, 700–710 (1986).

Pons, M.-N., *Bioprocess Monitoring and Control*. Hanser Publishers, New York (1992).

Proll, T. and N. M. Karim, "Nonlinear Control of a Bioreactor Model using Exact and I/O Linearization," *Int. J. Control*, **60**, 499–519 (1994).

Ramkrishna, D., "Statistical Models of Cell Populations," *Adv. Bioch. Eng.*, **11**, 1–47 (1979).

Ramkrishna, D., "Toward a Self-Similar Theory of Microbial Populations," *Biotech. Bioeng.*, **43**, 138–148 (1994).

Rawlings, J. B. and W. H. Ray, "Stability of Continuous Emulsion Polymerization Reactors: A Detailed Model Analysis," *Chem. Eng. Sci.*, **42**, 2767–2777 (1987).

Roels, J. A., *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press, Amsterdam, New York, Oxford (1983).

Srienc, F. and B. S. Dien, "Kinetics of the Cell Cycle of *Saccharomyces cerevisiae*," *Annals of the New York Academy of Sciences*, pages 59–71 (1992).

Srienc, F., Flow Cytometry in Biotechnology: Potential and Limitations, In *Proceedings of ECB6*, pages 13–17 (1993).

Srienc, F., "Cytometric Data as the Basis for Rigorous Models of Cell Population Dynamics," *J. Biotech.*, **71**, 233–238 (1999).

Straight, J. V. and D. Ramkrishna, "Cybernetic Modeling and Regulation of Metabolic Pathways. Growth on Complementary Nutrients," *Biotechnol. Prog.*, **10**, 574–587 (1994).

Strassle, C., B. Sonnleitner, and A. Fiechter, "A Predictive Model for the Spontaneous Synchronization of *Saccharomyces cerevisiae* Grown in Continuous Culture. II. Experimental Verification," *J. Biotech.*, **9**, 191–208 (1989).

Subramanian, G. and D. Ramkrishna, "On the Solution of Statistical Models of Cell Populations," *Math. Biosciences*, **10**, 1–23 (1971).

Sulsky, D., "Numerical Solution of Structured Population Balance Models. II Mass Structure," *J. Math. Biol.*, **32**, 491–514 (1994).

von Meyenburg, H. K., Stable Synchrony Oscillations in Continuous Culture of *Saccharomyces cerevisiae* under Glucose Limitation, In Chance, B., E. K. Pye, A. K. Shosh, and B. Hess, editors, *Biological and Biochemical Oscillators*, pages 411–417. Academic Press, New York, NY (1973).

Witkowski, W. R. and J. B. Rawlings, Modelling and Control of Crystallizers, In *Proceedings of American Control Conference*, pages 1400–1405, Minneapolis, MN (1987).

Yamashit, Y., M. Kuwashim, T. Nonaka, and M. Suzuki, "On-Line Measurement of Cell-Size Distribution and Concentration of Yeast by Image-Processing," *J. Chem. Eng. Japan*, **26**, 615–619 (1993).

Zhang, Y., A. M. Zamamiri, M. A. Henson, and M. A. Hjortso, "Cell Population Models for Bifurcation Analysis and Nonlinear Control of Continuous Yeast Bioreactors," *J. Proc. Cont.* (2001). In press.

Zhao, R., A. Natarajan, and F. Srienc, "A Flow Injection Flow Cytometry System for On-Line Monitoring of Bioreactors," *Biotech. Bioeng.*, **62**, 609–617 (1999).

Zhu, G.-Y., A. M. Zamamiri, M. A. Henson, and M. A. Hjortso, "Model Predictive Control of Continuous Yeast Bioreactors Using Cell Population Models," *Chem. Eng. Sci.*, **55**, 6155–6167 (2000).

Control of Product Quality in Polymerization Processes

Francis J. Doyle III
Department of Chemical Engineering
University of Delaware
Newark, DE 19716

Masoud Soroush
Department of Chemical Engineering
Drexel University
Philadelphia, PA 19104

Cajetan Cordeiro
Air Products and Chemicals, Inc.
Allentown, PA 18195-1501

Abstract

The increasingly aggressive global competition for the production of higher quality polymer products at lower costs, along with a general trend away from new capital investments in the U.S., has placed considerable pressure on the process engineers in the U.S. to operate the existing polymer plants more efficiently and to use the same plant for the production of many different polymer products. The more efficient operation has been realized by better process control and monitoring while the available polymer product-quality sensors have been inadequate. Although many product quality indices cannot be measured readily, they can be estimated/inferred in real time from the readily available measurements, allowing for inferential control of the polymer product quality. This paper presents a survey of the issues in controlling and monitoring plant-product quality indices such as molecular weight, copolymer composition, and particle size distributions in polymerization reactors. Examples will be given to illustrate some of the methods surveyed.

Keywords

Particle size distribution, Process control, Inferential control, Batch control, Population balance models, Polymerization reactor control, Multi-rate measurements, Polymer product quality, Polymerization reactor monitoring

Introduction

A polymer product is composed of macromolecules with different molecular weights, and the processability and subsequent utility of a polymer product depends strongly on the macromolecule distributions, such as molecular weight distribution (MWD), copolymer composition distribution (CCD) [in copolymerization], and particle size distribution (PSD) [in emulsion polymerization]. For instance, in coatings, film formation, film strength, and gloss depend on the MWD, CCD, and PSD of the polymer. Since the distributions are influenced greatly by the polymerization reactor operating conditions, the production of a high quality polymer requires effective monitoring and control of the operating conditions (Congalidis and Richards, 1998; Ogunnaike, 1995). The effective monitoring and control can be realized only when sufficient frequent information on the distributions is available.

Polymerization reactors are a class of processes in which many essential process variables related to product quality cannot be measured or can be measured at low sampling rates and with significant time delays. The lack of readily-available, frequent measurements from which polymer properties can be inferred, has motivated a considerable research effort in the following research directions:

- The development of new on-line sensors [lists of many of the currently-available on-line sensors are provided in (Ray, 1986; Chien and Penlidis, 1990)].
- The development of qualitative and quantitative relations between easier-to-measure quality indices such as density, viscosity and refractive index, and more-difficult-to-measure quality indices such as conversion and average molecular weights (Kiparissides et al., 1980; Schork and Ray, 1983; Canegallo et al., 1993; Soroush and Kravaris, 1994; Ohshima et al., 1995; Ohshima and Tomita, 1995).
- The development of state estimators that are capable of estimating unmeasurable polymer properties from readily available measurements. The availability of sufficiently-accurate, first-principles, mathematical models of many polymerization reactors has made possible the development of reliable state estimators for the reactors (Jo and Bankoff, 1996; Ellis et al., 1988; Kim and Choi, 1991; Kozub and MacGregor, 1992; Ogunnaike, 1994).

Product Quality in Polymerization Processes

Product quality is a much more complex issue in polymerization than in more conventional short chain reactions (Ray, 1986). Because the molecular structure of the polymer is so sensitive to reactor operating conditions, upsets in feed conditions, mixing, reactor temperature and so on can change significantly critical molecular properties such as molecular weight distribution, copolymer composition distribution, copolymer chain sequence distribution, stereoregularity, and degree of chain branching.

The properties of a polymer product, such as the mechanical properties and the characteristics in molding, having strong correlation with the molecular weight distribution (MWD) of the polymer. Nunes et al. (1982) found that thermal properties, stress-strain properties, impact resistance, strength and hardness of films of polymethyl methacrylate and polystyrene were all improved

by narrowing MWD. It is also generally said that the polymer of long chain length gives superior mechanical properties to polymer products but has insufficient molding characteristics. Then the molding characteristics can be improved by blending short chain length polymer into this long chain length polymer, while the good mechanical characteristics are kept. That is, the broader MWD can be obtained by this blending. Therefore, the development of the methodology for adjusting MWD during the reaction to suitable one according to its use is desired, especially in producing high quality polymers.

Schoonbrood et al. (1995) studied the influence of copolymer composition and microstructure on the mechanical bulk properties of styrene-methyl acrylate copolymers. They found that copolymer composition drift has an influence on polymer mechanical properties such as Young's modulus, maximum stress, and elongation at break. In the case of copolymers that are homogeneous with respect to chemical composition, (a) maximum stress and elongation at break depend on the molecular weight distribution, (b) Young's modulus is independent of copolymer composition and molecular weight distribution in the ranges studied, and (c) maximum stress and elongation at break weakly depend on the copolymer composition. In the case of copolymers that are heterogeneous with respect to chemical composition, copolymer microstructure affects strongly Young's modulus, maximum stress, and elongation at break.

In paints and coatings, molecular weight, composition, and functional group distributions all play a key role in polymer performance. For solution viscosity reasons, narrow molecular weight distribution is useful, but not every paint or coating benefits from it. It depends on the application. For example, air-dry paints benefit from very broad molecular weight distribution (Grady, 2000).

For processing and end-use performance of latex coatings, it is often advantageous to produce a latex with high solids content while maintaining viscosity within acceptable limits. Latex particle size and particle size distribution directly affect the relationship between solids volume fraction and rheological properties. The influence of monodisperse latex particles on latex viscosity is described by the Dougherty-Krieger equation (Krieger and Dougherty, 1959),

$$\eta_r = \left(1 - \frac{\phi}{\phi_m}\right)^{-2.5\phi_m}$$

where η_r is the ratio of emulsion viscosity to that of the pure fluid (water for instance), ϕ is the volume fraction of solids, and ϕ_m is the maximum volume fraction of latex particles. For polydisperse systems, it has long been established that blends of different size particles yield viscosities which are lower than the viscosities of any of the monodisperse particles used to make nthe blend (for equivalent solids concentrations). Eveson and coworkers (1951) suggested that a particle suspension with a bimodal distribution can be regarded as a system in which the larger particles are suspended in a continuous phase formed by suspension of the smaller particles in the fluid medium. In other words, a suspension of smaller particles behaves essentially as a fluid toward the larger particles. Farris (1968) extended this line of reasoning to a multimodal blend of particle sizes with any number of modes and Parkinson et al. (1970) to a continuous particle size distribution. In both cases, successive application of a monodisperse expression for relative viscosity to particles of increasing size in a blend yielded an expression for relative viscosity of the form

$$\eta_r = \prod_i \left(1 - \frac{\phi_i}{\phi_{m,i}}\right)^{-k\phi_{m,i}}$$

where the ϕ_i are volume fractions of particles of a given size in the particle size distribution. Although this expressions is not directly applicable to the prediction of viscosity for continuous latex distributions, the reasoning behind its derivation suggests that control of the particle size distribution would be an appropriate approach to targeting desired latex rheological properties. In contrast, the more common approach of controlling moments of the distribution, is only indirectly related to target properties.

Classification of Variables in a Polymerization Plant

A customer evaluates the quality of a polymer product on the basis of indices, end-product quality indices, that are usually different from the product quality indices, plant-product quality indices, known to the plant process engineer. The end-product quality indices are related to the final use of the polymer product and usually cannot be measured in real time because of the complicated and slow measurement techniques needed or simply the inability to measure the quality indices until the final polymer product is formulated and used. On the other hand, polymer plants are operated at desired conditions by setting and regulating the plant variables (such as pressures, temperatures, and flow rates) that are measured readily on-line. We will refer to these readily measurable variables as basic plant variables to distinguish them from the plant-product quality indices and the end-product quality indices (end-use properties). These differences lead us to categorize variables in a polymer plant into the following three classes, intersections of which may not be null:

- Basic plant variables,
- Plant-product quality indices,
- End-product quality indices.

Basic plant variables that can be measured readily on-line and whose values are set by the process engineer to

operate the plant at desirable operating conditions. Examples of the basic process variables are temperatures, pressures, liquid levels, flow rates, and feed compositions.

Plant-product quality indices are usually monitored by the process engineer to ennnsure proper operation of the plant. Measurements of these indices are rarely available on-line and are usually obtained by laboratory sample analyses. Examples of these indices are viscosity, melt viscosity, density, copolymer composition distribution, molecular weight distribution, melt index, copolymer chain sequence distribution, stereoregularity, particle size distribution, porosity, surface area, and degree of chain branching.

End-product quality indices, often referred to as customer specifications or end-use properties, quantify the quality of the final product. These indices are usually "abstract" (to the plant process engineer), and their relations to the plant-product quality indices are complex and not well-understood. In many cases, the relations are known qualitatively on the basis of experience. It is important also to note that in cases where the relationships between these end-use properties and the plant-product quality indices are known, they are not "one-to-one". The end-product quality indices are rarely measured off-line in the plant because the measurements usually cannot be made until the final polymer product is formulated and used. Furthermore, many of these end-use properties (such as "softness", "blockiness", and "color") are "categorical" but not quantifiable in numerical form at the present time. Examples of the end-product quality indices (customer specifications) are adhesive strength, impact strength, hardness, elastic modulus, flow properties (film blowing, molding, etc.) strength, stress crack resistance, color, clarity, corrosion resistance, abrasion resistance, density, temperature stability, plasticity uptake, spray drying characteristics, and coating and adhesion properties. More examples of the end-product quality indices can be found in (Nunes et al., 1982; Ray, 1986; Dimitratos et al., 1994).

One of the greatest difficulties in achieving quality control of polymer end-products is our poor understanding of the quantitative relationship between (a) the end-product quality indices and (b) the plant-product quality indices and the basic plant variables. The actual customer specifications are in terms of the end-product quality indices. Since the quantitative relationship is the least understood area in polymerization reaction engineering, it is very hard to calculate the values of plant-product quality indices that corresponds to the actual customer specifications.

Mathematical Modeling

A major objective of polymerization reaction engineering has been to understand how reaction mechanism, the physical transport phenomena (e.g. mass and heat transfer, mixing), reactor type and operating conditions affect the plant-product quality indices. As discussed in (Ray, 1991), various chemical and physical phenomena occurring in a polymer reactor can be classified into the following three levels of modeling:

1. Microscale chemical kinetic modeling: Polymer reactions occur at the microscale. If the elementary reaction steps of a polymerization mechanism are known, the distributions can be calculated in terms of the kinetic rate constants and the concentration of the reactants. The available mathematical models are statistical or are based on detailed species conservation methods. The most powerful approach to modeling polymerization kinetics is the detailed species balance method. Using the conservation laws of mass, one can derive an infinite set of equations for the species present in the reaction mixture.
2. Mesoscale physical/transport modeling: At this scale, interphase heat and mass transfer, intraphase heat and mass transfer, interphase equilibrium, micromixing, polymer particle size distribution, and particle morphology play important roles and further influence the polymer properties. For example, diffusion-controlled free-radical polymerizations are manifestations of mesoscale mass transfer phenomena. For a comprehensive list of results available in this area, the reader can refer to the excellent review paper by (Kiparissides, 1996).
3. Macroscale dynamic reactor modeling: At the macroscale, one has to deal with the development of models describing the macromixing phenomena in the reactor, the overall mass and energy balances, particle population balances, the heat and mass transfer from the reactor as well as the reactor dynamics and control.

Population Balance Model

Population balance model descriptions have found a wide range of application in distributed process systems including crystallization, precipitation, and polymerization. An excellent treatment of the theoretical aspects of the subject is given in (Ramkrishna, 2000). In this paper, we focus on the application of population balance models to a specific sub-class of polymerization systems—particle size distributions in an emulsion system. Within this class, there are two general categories of behaviors: zero-one and pseudo-bulk systems. When conditions are such that the rate of radical-radical bimolecular termination within a latex particle is extremely fast relative to the rate of radical entry into particles, evolution of the latex particle size distribution can be modeled as a zero-one system (Gilbert, 1995). This model considers latex particles containing either zero or one radical at a given instant. The reasoning behind this model is that a particle will flip between two states, the zero and one

radical states, each time a radical enters the particle (or exits).

Latex particle size, monomer type and concentration, are among several key factors which strongly influence whether a system approaches zero-one kinetics. For example, termination of radicals within small particles is rapid because diffusion distance of the radical reaction centers is small. Moreover, radical entry rates, according to the Smoluchowski relation, $k_e = 4\pi r_s N_A D_w$ (r_s is swollen particle radius) decrease with decreasing particle size. In fact, many systems (styrene for example) approach zero-one kinetics during early stages of particle nucleation and growth when the size of particles is small.

To model a zero-one system, the particle size population is divided into a population containing zero radicals, $n_0(r)$ and a population containing one radical, $n_1(r)$. The one radical population is further divided into a population containing a polymer radical, $n_1(r)$, which would not readily diffuse out of the particle due to its size, and a population containing a monomer radical formed from chain transfer reactions, $n_1^m(r)$, which presumably can readily exit particles.

The population, $n_1(r)dr$, represents the moles(or number) of polymer particles per liter of water with unswollen particle radii between r and $r + dr$ at time t. Population balance equations for a batch reactor are given by:

$$\begin{aligned}\frac{\partial n_0(r,t)}{\partial t} &= \rho(r)\left[n_1(r) + n_1^m(r) - n_0(r)\right] + k_o(r)\cdot n_1^m(r) \\ &+ \int_{r_{nuc}}^{r/2^{1/3}} \frac{r^2 B(r-r',r')}{(r^3-r'^3)^{2/3}}\left[n_0(r')n_0(r-r') + \right. \\ &\qquad \left. n_1(r')n_1(r'-r)\right]dr' \\ &- n_0(r)\int_{r_{nuc}}^{r_\infty} B(r,r')\left[n_0(r') + n_1(r')\right]dr'\end{aligned}$$

$$\begin{aligned}\frac{\partial n_1(r,t)}{\partial t} &= \rho_{init}(r)\cdot n_0(r) \\ &- \rho(r)\cdot n_1(r) - k_{tr}C_p n_1(r) + k_p C_p n_1^m \\ &+ \int_{r_{nuc}}^{(r^3-r_{nuc}^3)^{1/3}} \frac{r^2 B(r-r',r')}{(r^3-r'^3)^{2/3}} n_0(r')n_1(r-r')dr' \\ &- n_1(r)\int_{r_{nuc}}^{r_\infty} B(r,r')\left[n_0(r') + n_1(r')\right]dr' \\ &- \frac{\partial}{\partial r}\left[G(r)\cdot n_1(r)\right] + \left[k^w_{p,j_{crit}-1}C_W\left[IM_{j_{crit}-1}\right]\right. \\ &\qquad \left. + \sum_{i=z}^{j_{crit}-1} k_{em,i}C_{micelle}\left[IM_i\right]\right]\delta(r-r_{nuc})\end{aligned}$$

Taking the first equation as an example, the terms premultiplied by $\rho(r)$ represent radical entry into particles, the next term represents radical desorption from particles, and the integral terms represent coagulation. The second equation has an additional term that represents new particle formation by micellular and homogeneous nucleation mechanisms.

Pseudo-bulk systems are characterized by slow radical termination within the particles relative to rapid entry of polymer radicals and re-entry of exited monomer radicals. In such systems, particles can contain more than one radical at a given instant. Moreover, particles with zero, 1, 2, ... radicals, switch identities(number of radicals) so rapidly that the evolution of the particle size distribution can be described by a single type of particle with an average number of radicals, $\bar{n}(r)$. Again, latex particle size, monomer type and concentration are factors which strongly influence whether a system approaches pseudo-bulk kinetics. Specifically, larger particle sizes increase radical entry rates and decrease radical termination and desorption rates, all of which favor pseudo-bulk kinetics.

The gel effect also decreases radical termination rates. This phenomenon is operative in many polymer systems when monomer concentration in particles is low relative to polymer concentration due to monomer depletion. Monomer depleted conditions often occur at the end of a batch when particle sizes are large. Therefore, the gel effect often coincides with large particle size and can be an additional factor which pushes a system towards pseudo-bulk kinetics.

A particle size distribution model for a pseudo-bulk system is given by:

$$\begin{aligned}\frac{\partial n(r,t)}{\partial t} &= \frac{1}{2}\int_{r_{nuc}}^{r-r_{nuc}} B(r-r',r')n(r')n(r-r')dr' \\ &- n(r)\int_{r_{nuc}}^{r_\infty} B(r,r')n(r')dr' - \frac{\partial}{\partial r}\left[G(r)\cdot n(r)\right] \\ &+ \left[k^w_{p,j_{crit}-1}C_W\left[IM_{j_{crit}-1}\right]\right. \\ &\qquad \left. + \sum_{i=z}^{j_{crit}-1} k_{em,i}C_{micelle}\left[IM_i\right]\right]\delta(r-r_{nuc})\end{aligned}$$

where particle growth rate is given by

$$G(r) = \frac{k_p C_p \bar{n}(r)\rho_p w_m}{4\pi r^2 N_a}$$

k_p is the propagation rate constant, C_p the monomer concentration in polymer particles, and $\bar{n}(r)$ the average radical concentration.

In summary, at least for a batch polymerization, zero-one kinetics is expected to be operative at early stages of polymerization when particle size is small and particles are rich in monomer whereas pseudo-bulk kinetics is favored in the latter stages of polymerization when particle size is large and the gel effect is strong. Of course neither of these models treats the more complicated intermediate case wherein particles can contain 0, 1, 2, and

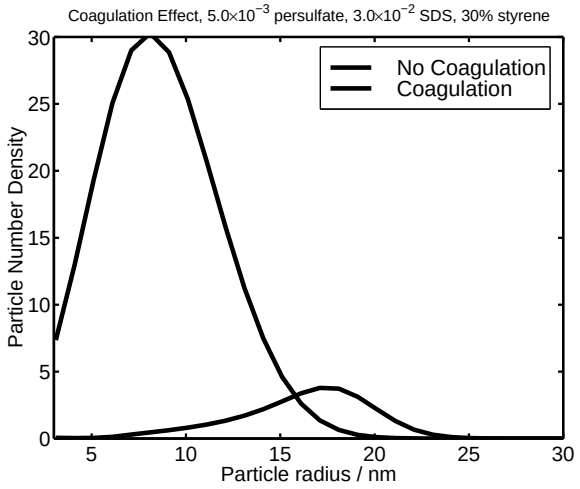


Figure 1: Simulation of zero-one model with and without coagulation 10 minutes after inception of polymerization.

3 radicals, for example, but for model based control purposes, we anticipate that the zero-one and pseudo-bulk models may be combined in a manner that adequately handles the intermediate case as well. Details of both these models can be found in (Gilbert, 1995).

Coagulation Coefficients. Coagulation is an extremely important phenomena in emulsion polymerization, having a large impact on latex particle numbers and particle size distributions. Actually, colloidal particles are thermodynamically unstable. Surface tension between particle and bulk phase dictates that free energy decreases upon particle coagulation due to decreased surface area (Atkins, 1978):

$$dG = \gamma dA$$
$$dG < 0 \qquad \text{for} \quad dA < 0$$

Colloidal stability is a consequence of kinetics. The electrostatic charge on the surface of a surfactant stabilized colloidal particle represents a significant energy barrier to coagulation between two approaching particles. If this barrier is large, very few particles will have sufficient kinetic energy to exceed the barrier and coagulate. Quantitatively, this potential energy barrier can be calculated using DLVO theory. This is based on calculating the total potential energy of interaction between two particles as the sum of a van der Waals attractive potential and an electrostatic repulsion potential (Ottewill, 1982). Details can be found in (Coen et al., 1998b) and the final result is

$$B(r_i, r_j) = \frac{2k_BT}{3\eta W_{ij}}\left(2 + \frac{r_{si}}{r_{sj}} + \frac{r_{sj}}{r_{si}}\right)$$
$$W_{ij} = \frac{r_{si} + r_{sj}}{4\kappa r_{si} r_{sj}} \exp\left(\frac{\Phi_{\max}}{k_BT}\right)$$

where W_{ij} is the Fuchs stability ratio and $\Phi_{\max}$ the maximum potential energy with respect to particle separation distance. One important characteristic of coagulation coefficients is that values increase exponentially as particle size is decreased. Also, for the smallest particles, the coagulation coefficient is nearly independent of the radius of the other particles with which the small particles coagulate. This feature was exploited for determining coagulation coefficients in our simulations by approximating the coagulation coefficients as depending only on the radius of the smaller of any two given particles coagulating. Figure 1 is a comparison of two simulated particle size distributions under identical operation conditions but with and without coagulation respectively. This figure highlights the influence that coagulation has on particle numbers and distributions during particle nucleation.

Calculation of coagulation coefficients in an on-line control application is problematic because calculation of $\Phi_{\max}$ is an extremum problem over particle separation distance and therefore, would be an embedded optimization problem. Another limitation of the DLVO theory is it does not account for shear effects which can markedly alter charge distributions surrounding particles and cause greatly accelerated coagulation rates. These problems are still open issues and feasible solutions will likely involve significant empiricism.

Numerical Solution

Approaches to solving population balance equations found in the literature can be generally classified into one of three distinct methods. Orthogonal collocation on fixed and moving finite elements has been used by several authors (Dafniotis, 1996; Rawlings and Ray, 1988a,b) to solve population balance models. This moving finite element method overcomes some of the numerical instabilities and inaccuracies associated with more conventional techniques such as finite differences.

Dafniotis (1996) describes the numerical problems associated with the solution of hyperbolic PDE's in terms of wave theory. If the desired solution is expressed in terms of a Fourier series, the behavior of a solution can be examined by examining individual components of the series, i.e., sine-cosine waves that propagate with specific phase and amplitude. Numerical operators such as finite differences, do not preserve the correct phase and amplitude. Particularly, errors associated with the phase are sometimes observed as spurious oscillations; referred to as dispersion. Also, when the amplitude of the numerical wave is damped relative to the exact solution

wave, discontinuities (for example boundary conditions) are smeared; referred to as numerical diffusion. Finally, hyperbolic partial differential equations often describe the propagation of near-shocks, or sharp wave fronts which require adequate resolution in the region of the shock.

In Dafniotis (1996), a moving finite element method(MFEM) is presented for solving the population balance equations for emulsion polymerization and to address some of the inherent numerical problems mentioned above. This method is based on the MFEM developed by Sereno and coworkers (1991).

A less sophisticated but much easier method to set up is the finite difference method. Here, partial derivatives in the population balance equations are approximated by finite differences. Gilbert (1997) has applied this method to modeling particle size distributions for emulsion polymerization systems.

A third method for solving population balance equations is sometimes referred to as methods of classes. Here, the distribution is discretized into classes of particles defined by finite particle size intervals. Mathematically, this involves transforming the partial differential equations from a differential to integrated form over small intervals. The presumed advantages of this approach include transformation of integral terms into more easily evaluated summation terms, elimination of partial differential terms with respect to particle size by forcing the discretization grid to move with particle growth rate, and the ability to coarsen the grid and still preserve key properties of the distribution such as moments (Kumar and Ramkrishna, 1996, 1997).

State Estimation

The inadequacy of frequent measurements related to the plant-product quality indices in polymer processes, has motivated the use of state estimators in controlling and monitoring the indices. The availability of sufficiently-accurate, first-principles, mathematical models for many polymerization reactors has made possible the development of the state estimators/observers. An estimator, which is designed on the basis of the process model, estimates unmeasured process variables from current and past process measurements. State estimators have also been used for sensor/plant fault detection and data reconciliation.

A major characteristic of polymerization reactors is their complex nonlinear behavior. Phenomena such as multiple steady states in continuouns stirred tank reactors, parametric sensitivity, and limit cycles are manifestations of the complex nonlinearity. Thus, reliable state estimation in polymerization reactors requires nonlinear models that can capture the complex nonlinear behavior. Motivated by the need for nonlinear state estimation, since the 1970s nonlinear state estimators have been used for polymerization reactors (Jo and Bankoff, 1996; Schuler and Suzhen, 1985; Ellis et al., 1988; Adebekun and Schork, 1989; Kim and Choi, 1991; Kozub and MacGregor, 1992; van Dootingh et al., 1992; Ogunnaike, 1994; Robertson et al., 1993; Liotta et al., 1997; Tatiraju et al., 1998a, 1999). In most of these studies, extended Kalman filters (EKF's) have been used for state estimation.

Multi-Rate State Estimation

In polymerization reactors, most of essential measurements related to plant-product quality indices, such as the leading moments of a MWD obtained by a gel permeation chromatograph (GPC), are available at low sampling rates and with considerable time-delays. On the other hand, measurements of basic plant variables such as temperatures, pressures, and densities are usually available at high sampling rates and with almost no delays. Because the plant product quality indices are usually not observable from the frequent, delay-free measurements alone, one has to design a multi-rate estimator/observer (i.e. one that uses both the frequent and infrequent measurements), to provide reliable estimates of the states, especially in the presence of model-plant mismatch and measurement noise. Multi-rate state estimation in polymerization processes has received considerable attention (Elicabe and Meira, 1988; Ellis et al., 1988; Dimitratos et al., 1989; Kim and Choi, 1991; Ogunnaike, 1994; Liotta et al., 1997; Mutha et al., 1997; Tatiraju et al., 1998a, 1999). Multi-rate EKF's have been used in most of these studies. For example, Ellis et al. (1988) used a multi-rate EKF to estimate the unmeasurable process states continuously from the frequently available measurements of temperature and density and the infrequent and delayed measurements of the average molecular weights (obtained by a gel permeation chromatograph). Mutha et al. (1997) proposed the use of a fixed-lag smoothing algorithm for multi-rate state estimation in a polymerization reactor. Tatiraju et al. (1998a, 1999) developed a method of multi-rate nonlinear state estimation and applied it to a solution polymerization reactor with fast measurements of reactor temperature, jacket temperature and density, and slow measurements of the zeroth, first and second moments of the polymer molecular weight distribution.

Inferential Control of Polymerization Reactors

An inferential control system has been defined conventionally as one that requires an estimated or inferred value of a controlled output to calculate the value of a manipulated input (Joseph and Brosilow, 1978a,b; Seborg et al., 1989; Marlin, 1995). Inferential control has application in processes in which (a) measurement of a controlled variable is not available frequently enough or

quickly enough to be used for feedback control, and (b) there are readily-available process measurements from which the value of the controlled variable can be inferred or estimated. The design of an inferential control system consists of two steps: (i) the synthesis of a controller assuming that all the controlled outputs are measured readily, and (ii) the design of an estimator to estimate the controlled outputs that are not measured readily. An inferential control system should ensure zero offset for all the controlled outputs. One of the industries that has benefitted greatly from inferential control is the polymer industry, where frequent measurements of even plant-product quality indices are rare. Extensive reviews of recent advances in inferential control can be found in (Doyle III, 1998; Soroush, 1998b).

Multi-Rate Control

In the polymer industry, there are many processes wherein the choice of sampling rate is limited by the availability of the output measurement. For example, composition analyzers such as gas chromatographs have a cycle time of say 5 to 10 minutes compared to a desired control interval of say 0.1 to 1 minute. If the control interval is increased to match the availability of measurements then control performance deteriorates significantly. In addition to the slow measurements (which are available at different low sampling rates and are delayed), there are usually process variables such as temperature and pressure that can be measured at high sampling rates and with almost no time delays, leading to multi-rate control problems. Successful recent implementations of multi-rate control on polymerization processes include (Ellis et al., 1994; Ogunnaike, 1994; Ohshima et al., 1994; Sriniwas et al., 1995; Crowley and Choi, 1996; Niemiec and Kravaris, 1997; Tatiraju et al., 1998b).

In the polymer industry, the problem of multi-rate control has been addressed by the following control methods:

- Cascade control
- Decentralized control
- State-estimator-based control

Cascade control systems have been used successfully to control essential variables whose measurements are not available frequently. The slave controller regulates a set of basic plant variables and adjust the manipulated inputs of process, while the master controller regulates the essential variables (usually plant-product quality indices) whose measurements are infrequent and delayed and calculates the set-points of the slave controller. While the inner loop is executed at the high rate at which the basic plant variables are measured, the outer loop is executed at the low rate at which the essential variables are measured (whenever these measurements are available). An advantage of this multi-rate control structure is its control system integrity in the face of any unforeseeable further delay in the essential slow measurements; whether the slow measurements arrive or not, the inner loop is always in place (Lee and Morari, 1990; Lee et al., 1992).

Decentralized control systems have been used in the polymer industry to control process variables whose measurements are available at different rates. As in every decentralized control system design, first the manipulated inputs and the controlled outputs should be paired. A single-input single-output controller is then designed for each pair. Each of the SISO controllers is executed (takes action) at the rate at which the measurements of the corresponding controlled output are available. An advantage of these multi-rate control structures is also its control system integrity in the face of any unforeseeable further delay in the slow measurements; whether the slow measurements arrive or not, the "fast" loops are always in place.

State-estimator-based multi-rate control systems include a state estimator which estimates frequently all state variables of the process from the available, fast and slow measurements. The frequent measurements and estimates are then fed to a single-rate controller as if the process has only single-rate measurements. In contrast to the first two classes of the multi-rate control systems that can be non-model-based, the last class of the multi-rate systems have to be model-based (since they include estimators).

Batch Control Issues

Model-based Optimization Approaches to Batch Polymerization Control

In the case of batch systems, one can formulate a classical optimal control problem in an effort to control the endpoint properties of the batch. In a number of studies, this is implemented in a receding horizon framework, yielding a so-called Model Predictive Controller (MPC). MPC utilizes a process model to compute a future open-loop control sequence which optimizes an objective function, given past and current information of the system. The first control move is implemented and the optimization problem is re-solved at the next sampling time as updated information becomes available.

Applications of MPC to semi-batch polymerization systems include (Russell et al., 1997), where *linear* MPC was applied to a Nylon system using empirical models for quality control. The primary modification to the MPC algorithm was the use of a shrinking horizon, originally proposed in (Joseph and Hanratty, 1993). A similar formulation of MPC was adopted by Georgakis and co-workers (Liotta et al., 1997). In their work, a *nonlinear* formulation was proposed; however, they employed a "least-squares"-like analytical solution to the unconstrained problem. Another notable citation is the work

in (Ettedgui et al., 1997), where a fed-batch reactor is studied (in simulation) for the application of nonlinear model-based estimation and predictive control. In that case, the sequential solution and optimization technique of (Wright and Edgar, 1994) was employed.

In general, MPC is posed as an on-line optimization problem, typically requiring the solution of a constrained linear, quadratic, or nonlinear programming problem. The generalized optimization problem considered can be expressed as:

$$\min_{\underline{u}}[\max \Phi_i(\underline{x},\underline{u})] \qquad i = 1, n_{obj}$$

subject to

$$\begin{aligned} \dot{\underline{x}} &= f(\underline{x},\underline{u}) \\ g_i(\underline{u}) &\le 0 \qquad i = 1, n_{ineq} \\ h_i(\underline{u}) &= 0 \qquad i = 1, n_{eq} \end{aligned}$$

Here, the vector $\underline{u}$ contains the values for the sequence of manipulated variable moves over the batch cycle (*e.g.*, surfactant feed), n_{obj} denotes the number of terms considered in the objective function, and the constraints describe the process model and corresponding operating constraints. Several forms of the objective function can be considered. The following 1-norm type objective could be considered:

$$\Phi = \sum_{i=1}^{N_E}\sum_{j=1}^{N_J} \frac{\left|n_{ij} - n_{ij}^{target}\right|}{n_{scale}}$$

Here, n_{scale} is a factor used to scale the objective function values. A 2-norm objective can also be formulated;

$$\Phi = \sum_{i=1}^{N_E}\sum_{j=1}^{N_J} \left[\frac{(n_{ij} - n_{ij}^{target})}{n_{scale}}\right]^2$$

Many interesting polymer products have corresponding distributions that are multi-modal in nature. These can be produced, for example, by multiple surfactant additions, sufficiently separated in time. Therefore, a potentially effective objective definition is to actually define multiple objective functions, each tied to a particular distribution mode, and perform a min-max optimization. For bimodal distributions, the objective can be expressed as follows:

$$\Phi_1 = \sum_{i=1}^{N_1}\sum_{j=1}^{N_J} \left[\frac{(n_{ij} - n_{ij}^{target})}{n_{scale}}\right]^2$$

$$\Phi_2 = \sum_{i=N_1+1}^{N_E}\sum_{j=1}^{N_J} \left[\frac{(n_{ij} - n_{ij}^{target})}{n_{scale}}\right]^2$$

$$\min_{u_i, i=1,11} \Phi = \max(\Phi_1, \Phi_2)$$

Here, N_1 represents the number of finite elements spanning the lower particle size mode of the distribution.

Case Study I: Control of an Emulsion Polymerization Reactor

The control of particle size distribution as an end-objective in emulsion polymerization control is well motivated in industrial practice, and has been well documented in the literature (see, for example, the recent review by Congalidis and Richards, 1998). The authors pointed out that "on-line control not only of average polymer properties but also of polymer distributions such as the particle size... will become important". They continue: "The instrumentation and control methodologies that will need be deployed to meet these needs is a challenging and vibrant area of investigation for academic researchers and industrial practitioners alike."

In this section, we present a case study for the optimal control of particle size distribution in a semi-batch styrene polymerization reactor. The isothermal model of Coen et al. (1998a) incorporates current theory on particle nucleation, growth and coalescence mechanisms for styrene at $50°C$ and serves as the modeling basis for this study. This model is a *zero-one* model, referring to the assumption of instantaneous bimolecular radical termination in polymer particles, which gives rise to a mixed population of particles containing either one or zero radicals. Parts of the model that deviate from (Coen et al., 1998a) are described in detail in (Crowley et al., 2000).

The basic control problem is defined as the achievement of a target PSD at the end of the batch. For this study, we consider the manipulation of both surfactant feed rate (and/or concentration), and initiator feed rate. On-line measurement of the full PSD, for example by light scattering, is assumed for this study.

Open-loop Optimization of PSD

Two distinct variables related to surfactant concentrations were optimized to match a target PSD. The first variable considered is surfactant feed rate. Specifically, the optimization routine calculates a sequence of 10 surfactant feed flow rates(zero-order hold), each with a sample hold time of 3 minutes, up to a final time of 30 minutes. The target PSD was generated by simulating the model up to 30 minutes, with values for the 10 surfactant feed flow rates that yield a bimodal distribution. Control trajectories were calculated by defining an objective function in terms of simulated PSD deviations from a target distribution, and minimizing the objective function using the sequential quadratic programming algorithm, FSQP.

In a different approach, we considered use of free surfactant concentration, rather than surfactant feed rate, as the control variable. The reason behind this choice is that free surfactant concentration above the cmc is the essential driving force for particle nucleation. The free surfactant profiles consist of a sequence of first-order holds (i.e. piecewise linear), with each hold spanning a 3

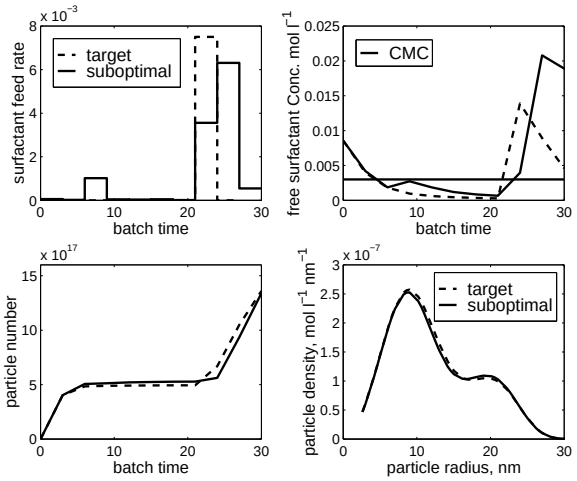


Figure 2: Optimization of surfactant feed rate sequence for target PSD; cmc= 3×10^{-3} mol/L, $[I]_0 =$ 0.01 mol/L, $[M]_0 = 2.59$ mol/L.

minute interval. As in the previous case, 10 holds were used to span a 30 minute control horizon. For this formulation of the optimization problem, the decision variables are 11 free surfactant concentration nodes, spaced at 3 minute intervals. Free surfactant concentration values between any two neighboring nodes are calculated simply by linear interpolation between the two nodal values. Linked in this way, the nodes form a continuous, though non-smooth, free surfactant profile.

The first optimization case involves computation of a sequence of surfactant feed flow rates which drive the system to a target PSD. The target distribution was generated by fixing the initial surfactant concentration at time= 0 and simulating the system out to 30 minutes with one large surfactant addition at 21 minutes. This addition is shown in Figure 2. Surfactant addition at 21 minutes produces a shoulder in the distribution at the end of the 30 minute simulation, with a peak height located at a particle size of about 20 nm. The solid stair case profile for surfactant feed rate in this figure represents the optimization solution. The 2-norm objective was used in this case. Although the optimizer appears to drive the simulated PSD to the target, as seen in this figure, the final free surfactant concentration for the optimized case is much larger than that of the target case. With disparity in surfactant levels at 30 minutes, the two distributions would diverge beyond this time. These differences can be improved through the use of input blocking (Crowley et al., 2000).

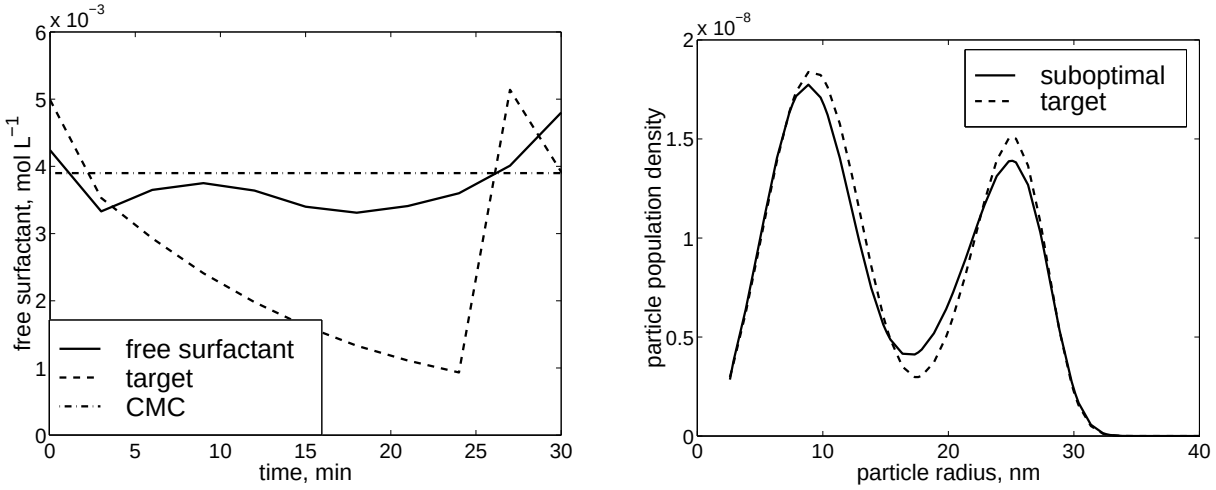
The optimal solution is quite sensitive to the choice of initial conditions, as well as the particular blocking formulation. Consequently, a second control strategy was considered. Because particle nucleation is dependent on the free surfactant level relative to the critical micelle concentration, a free surfactant trajectory is more closely related to the physical phenomena than surfactant feed flow rates. Of course, ultimately the feed flow rates are manipulated. A free surfactant trajectory represents a higher control level in a cascade configuration, with feed rates being the lowest level control variable. A useful property arises from this form. An intuitive initialization for optimization computations is to set all free surfactant concentrations in the control vector to the critical micelle concentration. Sensitivity of particle nucleation rate to surfactant perturbations is greatest at the CMC.

Local minima are often a cause of poor performance in gradient based optimization. For a bimodal distribution, it is possible to obtain suboptimal PSD solutions that are trapped in a local minimum because surfactant perturbations which decrease target offset for one of the modes may increase offset for the other mode. As an attempt to partially decouple this kind of interaction, we formulated a multi-function objective wherein each function is tied to a particular mode of the distribution. As described above, the resulting optimization is a min-max problem with two objective functions for a bimodal distribution. Figure 3 depicts the result of this case. When compared to the PSD offsets seen using the 2-norm (Crowley et al., 2000), the target offset for the min-max optimization is lower. Another advantage of this approach is that the optimization is "well-behaved" in the sense that offset of each mode from the target is balanced due to the structure of the objective function.

Batch-to-Batch Studies

In this section, a model refinement optimization technique is described using information from historical batch data of emulsion polymerization in conjunction with a fundamental first-principles model to determine the operating conditions necessary for a desired product quality.

The desired product quality must be a grade of product produced in the same range of operating conditions used in the historical batch data. As a result of this condition, the desired PSD must have a similar character to known PSDs produced by the process. A fundamental first-principles model of emulsion polymerization which accounts for polymer particle nucleation and polymer propagation exists. While these phenomena are well understood, emulsion polymerization presents a challenging process to control due to difficulty in modeling complex behavior such as particle aggregation and the significant nonlinear behavior involved in particle formation. The method described in this section seeks to combine this first-principles model of emulsion polymerization with an MPLS model to find the optimal control input sequence needed to achieve a desired product quality in a semi-batch emulsion polymerization reaction. In the experiments conducted in this section, the process variables are the surfactant and initiator feed rate inputs at distinct



(a) Free surfactant profiles.

(b) Particle size distributions.

Figure 3: Synthesis of optimal free surfactant profile using min-max objective. Target and suboptimal distributions shown are at a batch time of 30 minutes. Initial reactant conditions are as follows: cmc= 0.0039, $[I]_0 = 0.005$ mol/L, $[M]_0 = 2.59$ mol/L.

time intervals, and the product quality is determined by the particle size distribution (PSD) at the final time of the experiment.

The design of inputs for the emulsion polymerization process is performed on an off-line batch-to-batch basis. The hybrid model, which combines a first-principles model with MPLS, is referred to as the design model. The MPLS model is used to approximate the difference between the PSD the first-principles model calculates and the PSD obtained from executing a given input sequence in an actual semi-batch reactor, thus capturing the effect of phenomena for which the first-principles model does not take into account. The optimization will be performed by minimizing the sum of the squared residual error between the PSD yielded by the design model and the target PSD. The design model is a function of the states of the system and the control inputs. A calibration set of data from historical batches having a set of operating conditions similar to that of the desired PSD will be required to begin the optimization procedure. The PSDs in the historical batch data and the target PSD are created from a virtual process model which simulates the actual semi-batch reaction in a plant. This virtual process model is structurally different from the first-principles model in the design model because it accounts for aggregation and has four parameter values that are varied from the design model values by 5%.

The hybrid design model is potentially useful because the MPLS model will account for the phenomena such as particle aggregation which are not included in the first-principles model within the design model. Furthermore, the MPLS model can be adjusted to account for noise that may occur during PSD measurement. In addition, the MPLS model calibration set can be refined once a significant amount of batches have been designed by using only a subset of historical batches in the MPLS model, and including only those batches most recently designed. A pure MPLS model could also be attempted, but important information can be extracted from the first-principles model, so one would not want to abandon it altogether. The first-principles model in the design model could be useful for optimizing control inputs for new processes with little batch history, and for optimizing control inputs for new grades with no batch history. The efficacy of using a first-principles/MPLS hybrid model in achieving a target PSD is investigated in the following work.

While MPCA is used to compress information in the process variables to low-dimensional spaces that describe the historical batch operation, MPLS reduces both the process variables and product quality variables to low dimensional spaces, and attempts to find a correlation between these low-dimensional spaces (Neogi and Schlags, 1998). PLS attempts to maximize covariance, which means that it focuses on the variance of X that is more predictive for the product quality Y, rather than focusing on the variance of X only (Nomikos and MacGregor, 1995). In this investigation, the MATLAB PLS Toolbox 2.0 by Eigenvector, Inc is used to perform PLS analyses.

Before beginning the optimization algorithm, it is necessary to generate historic batch data and a target PSD with the use of the process model. The target PSD must

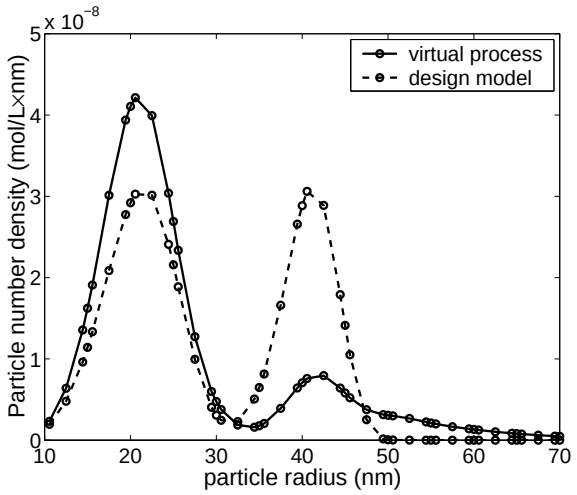


Figure 4: PSD for virtual process and design model.

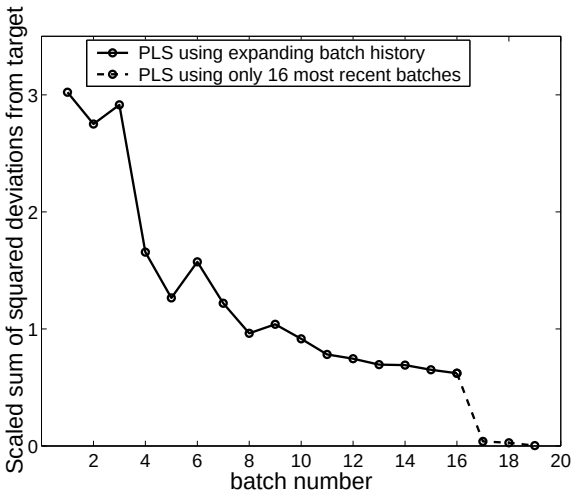


Figure 5: Improvement of PSD tracking error over the course of 20 successive batches.

be created within the same range of operating conditions used to generate the batch history, so that it is a different grade of product, within the same product family. The target PSD and historic PSDs are vectors containing the 48 particle number densities yi corresponding to 48 discrete particle radii r_i . A control input sequence of length 20 corresponding to the surfactant and initiator feed inputs at 10 evenly spaced time steps for a 100 minute batch reaction are used to create the PSDs.

The first step in the optimization algorithm is to obtain the scaling data and PLS regression matrix for the current batch history (X and Y blocks) in MATLAB.

The second step in the optimization is to design an input sequence using the following hybrid design model:

$$y^{design}(t = t_f) = y^{fp}(x, u, t) + y^{resid}(x, u)$$

where

$$y^{fp}(x, u, t_f) = \int_0^{t_f} f(x, u)dt$$

In this equation, t_f is the duration time of the semi-batch reaction, x and u are vectors containing the states of the system and the control inputs of the system, respectively y^{resid} is an approximation of the residual PSD between the process model PSD and the PSD from the first-principles model. $y^{resid}(x, u)$is calculated using the regression matrix calculated with the current batch history. The first iteration of the optimization will require an initial value for the control input sequence u. The optimization is performed by making small perturbations in u for each iteration. The direction of the perturbations are determined by the Jacobian matrix calculated from the previous iteration of the optimization. The size of the perturbation in u is fixed prior to optimization. The following objective function is minimized by the optimization procedure to determine control input sequence u to be used in the next batch:

$$\sum_{i=1}^{n_y} \left(y_i^{target} - y_i^{design}\right)^2$$

The optimized input sequence is then executed in the virtual process model, which simulates a batch reaction within an actual plant. Representative results of this study are shown below in Figures 4 and 5. The first figure shows the mismatch between the simulated process and the fundamental model used for PSD control. The overall trend along the batch history is mapped in the form of integral squared error in Figure 5. Clearly, the algorithm "learns" as the batches proceed.

Case Study II: Control of a Solution Polymerization Reactor

In this section, we present a comparative study of multi-rate control of a jacketed polymerization reactor in which free-radical solution polymerization of styrene takes place. A multi-rate control system consisting of the multi-rate nonlinear state estimator of (Tatiraju et al., 1999) and a mixed error- and state-feedback controller, is used. The performance of the multi-rate nonlinear control system is shown and compared with those of a multi-rate, PI, parallel cascade, control system and a multi-rate, PI, completely decentralized, control system.

Polymerization Process and the Control Problem

The reactor is a 3 m^3, jacketed, continuous, stirred tank reactor in which free-radical solution polymerization of styrene takes place. The solvent and initiator are benzene and azo-bis-iso-butyro-nitrile, respectively. The reactor has three feed streams: a pure monomer stream

at a volumetric flow rate of F_m, a pure solvent stream at a volumetric flow rate of F_s, and an initiator stream, which includes solvent and initiator, at a volumetric flow rate of F_i. The volume of the reacting mixture inside the reactor is constant.

We use the same dynamic model described in (Tatiraju et al., 1999) to represent the reactor. The model has the form:

$$\begin{aligned}
\frac{dC_i}{dt} &= -\left[\frac{F_t}{V} + k_i\right] C_i + \frac{F_i C_{i_i}}{V} \\
\frac{dC_s}{dt} &= -\frac{F_t C_s}{V} + \frac{F_i C_{s_i} + F_s C_{s_s}}{V} \\
\frac{d\lambda_0}{dt} &= -\frac{F_t \lambda_0}{V} + f_3(C_i, C_s, C_m, T) \\
\frac{d\lambda_1}{dt} &= -\frac{F_t \lambda_1}{V} + f_4(C_i, C_s, C_m, T) \\
\frac{d\lambda_2}{dt} &= -\frac{F_t \lambda_2}{V} + f_5(C_i, C_s, C_m, T) \\
\frac{dC_m}{dt} &= f_6(C_i, C_m, T) + \frac{F_m C_{m_m} - F_t C_m}{V} \\
\frac{dT}{dt} &= f_7(C_i, C_m, T, T_j) + \frac{F_t(T_{in} - T)}{V} \\
\frac{dT_j}{dt} &= f_8(T, T_j) + \alpha Q
\end{aligned} \tag{1}$$

where $F_t = F_i + F_m + F_s$, V is the volume of the reacting mixture inside the reactor, C_m is concentration of the monomer in the reactor outlet stream, C_i is concentration of the initiator in the reactor outlet stream, C_s is concentration of the solvent in the reactor outlet stream, T is the reactor temperature, T_j is the jacket temperature, Q is the rate of heat input to the reactor jacket, and λ_0, λ_1 and λ_2 are the zeroth, first, and second moments of the MWD of the polymer product, respectively. The functions $f_1, \cdots, f_8$ and the parameter values of the reactor model are given in (Tatiraju et al., 1999); for brevity they are not given here. The first-principles mathematical model of the process described by (1) is used to represent the actual process.

The number-average and weight-average molecular weights of the polymer product (denoted by M_n and M_w respectively) are related to the moments according to

$$M_n = \frac{\lambda_1}{\lambda_0}, \qquad M_w = \frac{\lambda_2}{\lambda_1}$$

The reacting mixture density, reactor temperature (T), and jacket temperature (T_j) are assumed to be measured on-line once every 30 seconds and with almost no time delays. The monomer conversion (and thereby the monomer concentration, C_m) can be inferred from the density measurement, and thus can be calculated on-line. The zeroth, first, and second moments of the MWD of the polymer product are assumed to be measured at sampling periods of 3 hours and with time delays of 1 hour. The rate of heat input to the reactor jacket, Q, and the flow rate of the initiator feed stream, F_i, can be set arbitrarily on-line within the following ranges: $-20 \le Q \le 50 \;\; kJ.s^{-1}$ and $0 \le F_i \le 3.0 \times 10^{-5} \;\; m^3.s^{-1}$.

The control problem is to maintain the weight-average molecular weight of the polymer, M_w, and the reactor temperature, T, at desired values by manipulating the rate of heat input to the reactor jacket, Q, and the flow rate of the initiator feed stream, F_i.

Multi-Rate Nonlinear Control System

State Feedback Synthesis. With M_w and T as controlled outputs ($y_1 = M_w$ and $y_2 = T$), and F_i and Q as manipulated inputs ($u_1 = F_i$ and $u_2 = Q$), relative orders (degrees) of the process $r_1 = 2$ and $r_2 = 1$, and the characteristic (decoupling) matrix of the process is generically singular. Because of this generic singularity, we request a state feedback that induces two completely decoupled, 2nd-order, process output responses of the form:

$$\beta_{12}\frac{d^2 M_w}{dt^2} + \beta_{11}\frac{dM_w}{dt} + M_w = M_{w_{sp}} \tag{2}$$

$$\beta_{22}\frac{d^2 T}{dt^2} + \beta_{21}\frac{dT}{dt} + T = T_{sp} \tag{3}$$

where β_{12}, β_{11}, β_{22} and β_{21} are positive adjustable scalar parameters, $M_{w_{sp}}$ is the weight-average molecular weight set-point, and T_{sp} is the reactor temperature set-point. Substituting for the time derivatives from the process model in the preceding two equations, we obtain two identities of the forms

$$\begin{aligned}
\phi_1(x, u_1) &= \frac{1}{\beta_{12}} M_{w_{sp}}, \\
\phi_2(x, u_1, u_2, \dot{u}_1) &= \frac{1}{\beta_{22}} T_{sp}
\end{aligned} \tag{4}$$

where x is the vector of the state variables of the reactor. Let us represent the solution for $u = [u_1 \;\; u_2]^T$ of the constrained minimization problem

$$\min_u \left\{ \left[\phi_1(x, u_1) - \frac{M_{w_{sp}}}{\beta_{12}}\right]^2 + \left[\phi_2(x, u_1, u_2, 0) - \frac{T_{sp}}{\beta_{22}}\right]^2 \right\}$$

subject to

$$\begin{aligned}
0 \le u_1 \le 3.0 \times 10^{-5} \;\; m^3.s^{-1} \\
-20 \le u_2 \le 50 \;\; kJ.s^{-1},
\end{aligned}$$

by

$$u = \Psi(x, M_{w_{sp}}, T_{sp}) \tag{5}$$

Using the identities of (4), we add integral action to the state feedback of (5) (see Soroush, 1998a, for the details),

leading to the mixed error- and state-feedback controller

$$\begin{aligned}
\dot{\eta}_1 &= \eta_2 \\
\dot{\eta}_2 &= -\frac{1}{\beta_{12}}\eta_1 - \frac{\beta_{11}}{\beta_{12}}\eta_2 + \phi_1(x, u_1) \\
\dot{\eta}_3 &= \eta_4 \\
\dot{\eta}_4 &= -\frac{1}{\beta_{22}}\eta_2 - \frac{\beta_{21}}{\beta_{22}}\eta_4 + \phi_2(x, u_1, u_2, 0) \\
u &= \Psi(x, e_1 + \eta_1, e_2 + \eta_3)
\end{aligned} \tag{6}$$

where $e_1 = M_{w_{sp}} - M_w$ and $e_2 = T_{sp} - T$. The controller of (6) has integral action and inherently includes optimal windup and directionality compensators (Soroush, 1998a).

Multi-Rate State Observer. Application of the multi-rate nonlinear state estimation method described in (Tatiraju et al., 1999) to this polymerization reactor leads to the following reduced-order, multi-rate, nonlinear, state estimator:

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \\ \dot{z}_5 \end{bmatrix} = \begin{bmatrix} -\left[\frac{F_t}{V} + k_i\right]\hat{C}_i + \frac{F_i C_{i_i}}{V} \\ -\frac{F_t \hat{C}_s}{V} + \frac{F_i C_{s_i} + F_s C_{s_s}}{V} \\ -\frac{F_t \hat{\lambda}_0}{V} + f_3(\hat{C}_i, \hat{C}_s, C_m, T) \\ -\frac{F_t \hat{\lambda}_1}{V} + f_4(\hat{C}_i, \hat{C}_s, C_m, T) \\ -\frac{F_t \hat{\lambda}_2}{V} + f_5(\hat{C}_i, \hat{C}_s, C_m, T) \end{bmatrix}$$
$$- K \begin{bmatrix} f_6(\hat{C}_i, C_m, T) + \gamma_1 \\ f_7(\hat{C}_i, C_m, T, T_j) + \gamma_2 \\ f_8(T, T_j) + \alpha Q \end{bmatrix} + L \begin{bmatrix} \lambda_0^* - \hat{\lambda}_0 \\ \lambda_1^* - \hat{\lambda}_1 \\ \lambda_2^* - \hat{\lambda}_2 \end{bmatrix} \tag{7}$$

$$\begin{aligned}
\hat{C}_i &= z_1 + K_{11} C_m + K_{12} T + K_{13} T_j \\
\hat{C}_s &= z_2 + K_{21} C_m + K_{22} T + K_{23} T_j \\
\hat{\lambda}_0 &= z_3 + K_{31} C_m + K_{32} T + K_{33} T_j \\
\hat{\lambda}_1 &= z_4 + K_{41} C_m + K_{42} T + K_{43} T_j \\
\hat{\lambda}_2 &= z_5 + K_{51} C_m + K_{52} T + K_{53} T_j
\end{aligned}$$

where $L = [L_{ij}]$ and $K = [K_{ij}]$ are the estimator gains,

$$\gamma_1 = \frac{F_m C_{m_m} - (F_i + F_m + F_s) C_m}{V}$$

$$\gamma_2 = \frac{(F_i + F_m + F_s)(T_{in} - T)}{V}$$

$\lambda_0^*(t)$, $\lambda_1^*(t)$ and $\lambda_2^*(t)$ are the predicted present values of the infrequent measurable outputs, each of which is obtained by fitting a least-squared-error line to the most recent, three measurements of the moment. These linear regressions are always carried out except when only one measurement of each slow measurable output is available. In this case, the predicted present value of each slow measurable output is set equal to the single available measurement. The estimator initial conditions are the same as those in (Tatiraju et al., 1999). The multi-rate state estimator of (7) can be written in the compact form

$$\begin{aligned}
\dot{z} &= F(z, \tilde{y}, Y^*, u) \\
\hat{x} &= Q(z, \tilde{y})
\end{aligned} \tag{8}$$

where $Y^* = [\lambda_0^* \ \lambda_1^* \ \lambda_2^*]^T$ and $\tilde{y} = [C_m \ T \ T_j]^T$.

Multi-Rate Nonlinear Control System. The use of the mixed error- and state-feedback controller of (6), together with the multi-rate state estimator of (7), leads to a multi-rate nonlinear control system of the form:

$$\begin{aligned}
\dot{z} &= F(z, \tilde{y}, Y^*, u) \\
\dot{\eta}_1 &= \eta_2 \\
\dot{\eta}_2 &= -\frac{1}{\beta_{12}}\eta_1 - \frac{\beta_{11}}{\beta_{12}}\eta_2 + \phi_1(\hat{x}, u_1) \\
\dot{\eta}_3 &= \eta_4 \\
\dot{\eta}_4 &= -\frac{1}{\beta_{22}}\eta_2 - \frac{\beta_{21}}{\beta_{22}}\eta_4 + \phi_2(\hat{x}, u_1, u_2, 0) \\
u &= \Psi(\hat{x}, \tilde{e}_1 + \eta_1, e_2 + \eta_3) \\
\hat{x} &= Q(z, \tilde{y})
\end{aligned} \tag{9}$$

where $\tilde{e}_1 = M_{w_{sp}} - \hat{\lambda}_2/\hat{\lambda}_1$.

Multi-Rate Cascade and Decentralized Control Systems

We compare the performance of the multi-rate nonlinear control system of (9) with a multi-rate, PI, cascade, control system and a multi-rate, PI, completely decentralized, control system.

The multi-rate, PI, cascade, control system consists of two PI controllers. The master PI controller regulates the weight-average molecular weight by manipulating the reactor temperature set-point. The master controller is executed once every three hours, since the average molecular weight measurements are available at that low rate. The slave PI controller regulates the reactor temperature by manipulating the rate of heat input to the reactor jacket. The slave controller is executed at a much faster rate (once every 30 seconds). The control system has the form

$$\begin{aligned}
\xi_1(k+1) &= \left[1 - \frac{\Delta t}{\tau_{I_1}}\right]\xi_1(k) + \frac{\Delta t}{k_{c_1}}\left[T_{sp}(k) - T_{sp_{ss}}\right] \\
\dot{\xi}_2(t) &= -\frac{1}{\tau_{I_2}}\xi_2(t) + \frac{1}{k_{c_2}}\left[Q(t) - Q_{ss}\right] \\
T_{sp}(k) &= \mathrm{sat}_T\left\{T_{sp_{ss}} + k_{c_1}\left[\bar{e}_1(k) + \frac{1}{\tau_{I_1}}\xi_1(k)\right]\right\} \\
Q(t) &= \mathrm{sat}_Q\left\{Q_{ss} + k_{c_2}\left[\bar{e}_2(t) + \frac{1}{\tau_{I_2}}\xi_2(t)\right]\right\}
\end{aligned} \tag{10}$$

with $\xi_1(0) = M_w(0)$, $\xi_2(0) = T(0)$, $k_{c_1} = -4.0 \times 10^{-5}$, $k_{c_2} = 1.0 \times 10^{-4}$, $\tau_{I_1} = 1.0 \times 10^6$ s, and $\tau_{I_2} = 1.0 \times 10^6$ s, where $\bar{e}_1(k) = M_{w_{sp}}(t) - M_w(k)$, $\bar{e}_2(t) = T_{sp}(t) - T(t)$, and $\Delta t = 1.08 \times 10^4$ s.

The multi-rate, PI, completely decentralized, control system consists of two completely-decentralized PI controllers. One of the PI controllers regulates the weight-average molecular weight by manipulating the flow rate of the initiator feed stream, and the other regulates the reactor temperature by manipulating the rate of heat input to the reactor jacket. While the first PI controller is executed once every three hours (sampling rate of the average molecular weight), the second controller is executed once every 30 seconds (sampling rate of the reactor temperature). The control system has the form

$$\begin{aligned} \xi_1(k+1) &= \left[1 - \frac{\Delta t}{\tau_{I_1}}\right] \xi_1(k) + \frac{\Delta t}{k_{c_1}} \left[F_i(k) - F_{i_{ss}}\right] \\ F_i(k) &= \text{sat}_{F_i} \left\{ F_{i_{ss}} + k_{c_1} \left[\bar{e}_1(k) + \frac{1}{\tau_{I_1}} \xi_1(k) \right] \right\} \\ \dot{\xi}_2(t) &= -\frac{1}{\tau_{I_2}} \xi_2(t) + \frac{1}{k_{c_2}} \left[Q(t) - Q_{ss}\right], \\ Q(t) &= \text{sat}_Q \left\{ Q_{ss} + k_{c_2} \left[\bar{e}_2(t) + \frac{1}{\tau_{I_2}} \xi_2(t) \right] \right\} \end{aligned} \tag{11}$$

with $\xi_1(0) = M_w(0)$, $\xi_2(0) = T(0)$, $k_{c_1} = -5.0 \times 10^{-7}$, $k_{c_2} = 2.5 \times 10^{-4}$, $\tau_{I_1} = 1.0 \times 10^7$ s, and $\tau_{I_2} = 1.0 \times 10^6$ s.

Each of the preceding PI control systems includes two PI controllers with windup compensators (Soroush, 1998a).

Simulation Results

The performance of the multi-rate nonlinear controller is evaluated by simulating the following two cases: (a) when there is no measurement noise or model-plant mismatch (nominal case); and (b) when there are measurement noise and model-plant mismatch (non-nominal case). For each case, the performance of the multi-rate nonlinear control (NC) system is compared with those of the cascade control (CC) system and the decentralized control (DC) system. Measurement noise is introduced by adding a white noise signal to each of the moments calculated by the process model. Each of the noise signals is a 10% deviation from the value of the moment at that particular time. Model error is simulated by adding a 10% error in the propagation step rate constant. The following values of tunable parameters are used for the nonlinear controller:

$$K_{11} = 1.0, \quad K_{12} = 0.0, \quad K_{13} = 1.0$$

$$K_{i1} = K_{i2} = K_{i3} = 0.0, \quad i = 2, \ldots, 5,$$

$$L_{i1} = L_{i2} = L_{i3} = 0.0, \quad i = 1, \ldots, 2$$

$$L_{ij} = 0.0, \quad i \neq j, \; i = 3, \ldots, 5; \; j = 1, \ldots, 3$$

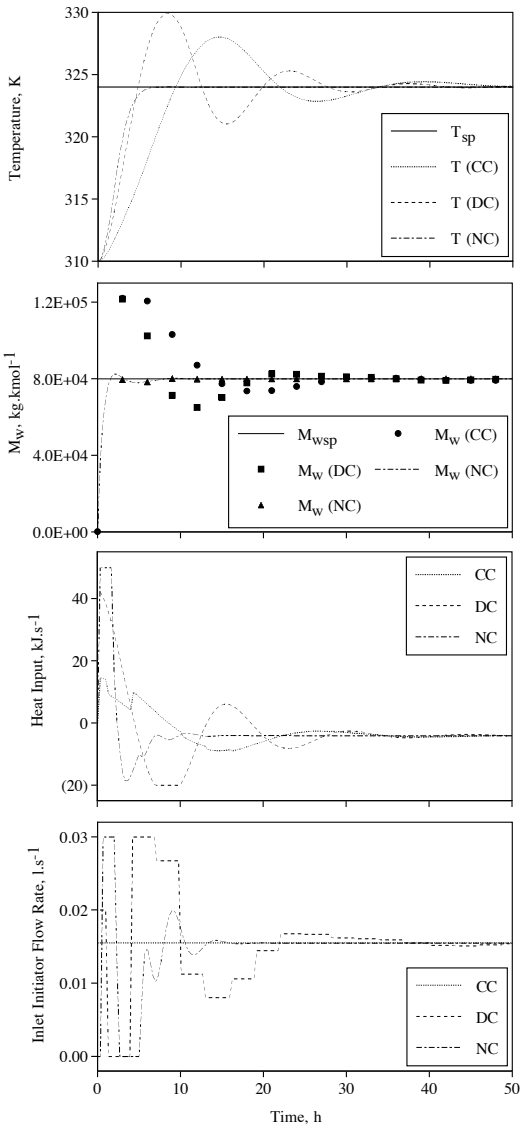


Figure 6: Controlled outputs and manipulated inputs under the multi-rate control systems (nominal case).

The temperature set-point $T_{sp} = 324.0$ K, and the weight-average molecular weight set-point $M_{w_{sp}} = 80{,}000$ $kg.kmol^{-1}$.

Figure 6 shows the profiles of the controlled outputs and the manipulated inputs for the nominal case under the three multi-rate control systems. The solid line represents the set point. For the M_w graph, the dashed line stands for the continuous estimates of M_w obtained by using the estimator of the multi-rate NC system, while the bullets stand for the infrequent and delayed "measurements" of the average-molecular weight. For this case, $L_{31} = 1.0 \times 10^{-4}, L_{42} = 1.0 \times 10^{-9}$, and $L_{53} = 1.0 \times 10^{-8}$. The benefit of using the multi-rate

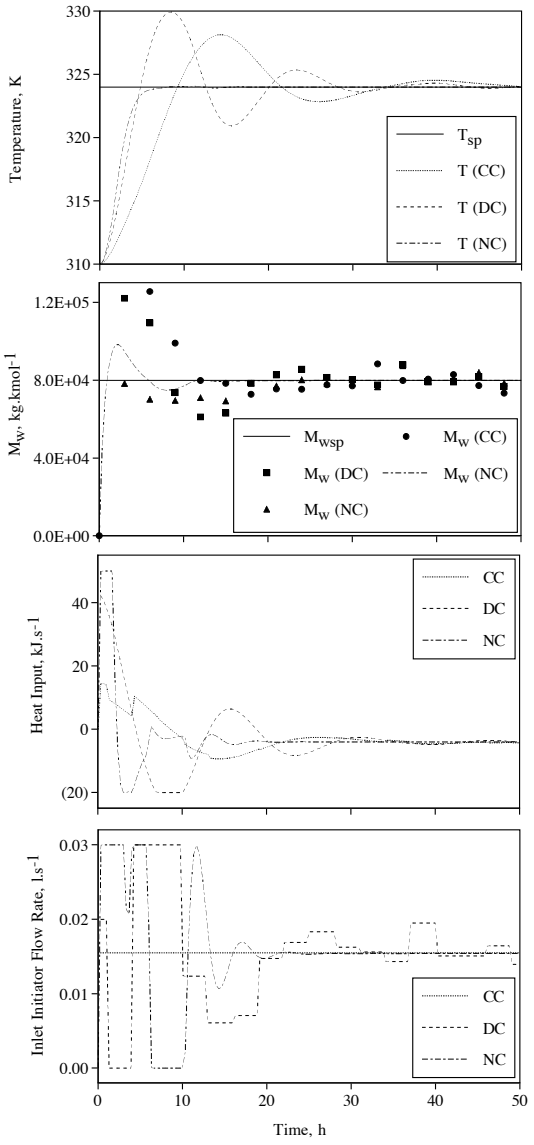


Figure 7: Controlled outputs and manipulated inputs under the multi-rate control systems (non-nominal case).

NC system is obvious in Figure 6. Not only under the NC system are the process output responses faster, but also they have smaller overshoots in the controlled variables. Under the CC and DC systems it takes more than 40 hours to take both the temperature and average molecular weight to their respective set-points, while under the NC system it takes less than 5 hours. The CC system regulates M_w only by manipulating Q, but the DC system regulates both M_w and T by manipulating Q and F_i. From the manipulated input graphs we can see that the NC system makes the most optimal use of manipulated variables. Under the CC and DC systems the manipulated inputs hardly reach the constraints, but still the CC and DC systems cannot be tuned to be more aggressive because the overshoots start increasing. Figure 7 depicts the profiles of controlled outputs and manipulated inputs for the non-nominal case. For this case, $L_{31} = 1.0\times10^{-4}$, $L_{42} = 2.0\times10^{-7}$, and $L_{53} = 9.5\times10^{-6}$. Again, the NC system shows a much superior performance as compared to the two PI control systems.

Conclusions

In this paper a brief survery of the recent advances that have led to improvement in polymer product quality was presented. Further improvement in polymer product quality requires solving challenging design, control, and monitoring problems that still exist in polymer processes.

Lack of sufficient controllability is a barrier to better product quality control in some polymer processes. In many polymer processes, better product quality requires minimizing/maximizing several product quality indices simultaneously. This multi-objective requirement may result in narrow ranges of process trajectories, putting a premium on the controllability of the process. For instance, in coatings, the product's composition, molecular weight, and particle size distributions should be maintained simultaneously in limited ranges to ensure the coating has a desired level of film formation, film strength, and gloss.

In many batch processes, product quality suffers from batch-to-batch inconsistency. There is a trend towards products with specific performance, which have higher value to a formulator or end-user. Furthermore, many of the current processes result in products with a broad inter-batch variance of molecular and physical characteristics, which in turn result in broad variance of performance. Blending of these batches usually lowers the average performance of the product lots. Segregation of "off-spec" product results in higher costs which may not be transferable to the customer.

Our understanding of the relationships among the basic plant variables, plant-product quality indices, and end-product quality indices is mostly empirical and qualitative. Polymer product development in the absence of qualitative relationships between the recipe, process and the final performance requires long times. Experimental techniques have been used to develop relationships that hold for the range of the experimental parameters studied. These products and processes therefore do not readily lend themselves to optimization, either in terms of productivity or reduction in variance. Having the ability to develop these relationships on a more fundamental basis will allow products to be developed in shorter times.

Acknowledgments

Financial support from the National Science Foundation through the grant CTS-9703278 is gratefully acknowledged by M. Soroush. The first author would like to acknowledge the financial support of NSF (BES 9896061) as well as the valuable contributions to this work by Tim Crowley, Charles Immanuel, and Chris Harrison.

References

Adebekun, D. K. and F. J. Schork, "Continuous Solution Polymerization Reactor Control 2: Estimation and Nonlinear Reference Control during Methyl Methacrylate Polymerization," *Ind. Eng. Chem. Res.*, **28**, 1846 (1989).

Atkins, P. W., *Physical Chemistry*. W. H. Freeman, second edition edition (1978).

Canegallo, S., G. Storti, M. Morbidelli, and S. Carra, "Densitometry for On-line Conversion Monitoring in Emulsion Homo- and Co-polymerization," *J. App. Poly. Sci.*, **47**, 961 (1993).

Chien, D. C. H. and A. Penlidis, "On-line Sensors for Polymerization Reactors," *JMS-Rev. Macromol. Chem. Phys.*, **C30(1)**, 1 (1990).

Coen, E. M., R. G. Gilbert, B. R. Morrison, H. Leube, and S. Peach, "Modeling particle size distributions and secondary particle formation in emulsion polymerisation," *Polymer*, **39**(26), 7099–7112 (1998a).

Coen, E. M., R. G. Gilbert, B. R. Morrison, H. Leube, and S. Peach, "Modelling particle size distributions and secondary particle formation in emulsion polymerisation," *Polymer*, **39**(26), 7099–7112 (1998b).

Congalidis, J. P. and J. R. Richards, "Process Control of Polymerization Reactors: An Industrial Perspective," *Polymer Reaction Eng.*, **6**(2), 71–111 (1998).

Crowley, T. J. and K. Choi, "On-line Monitoring and Control of a Batch Polymerization Reactor," *J. Proc. Cont.*, **6**, 119 (1996).

Crowley, T. J., E. S. Meadows, E. Kostoulas, and F. J. Doyle III, "Control of Particle Size Distribution Described by a Population Balance Model of Semibatch Emulsion Polymerization," *J. Proc. Cont.*, **10**, 419–432 (2000).

Dafniotis, P., *Modelling of emulsion copolymerization reactors operating below the critical micelle concentration*, PhD thesis, University of Wisconsin-Madision (1996).

Dimitratos, J., C. Georgakis, M. S. El-Aasser, and A. Klein, "Dynamic Modeling and State Estimation for an Emulsion Copolymerization Reactor," *Comput. Chem. Eng.*, **13**(1/2), 21–33 (1989).

Dimitratos, J., G. Elicabe, and C. Georgakis, "Control of Emulsion Polymerization Reactors," *AIChE J.*, **40**(12), 1993–2021 (1994).

Doyle III, F. J., "Nonlinear Inferential Control for Process Applications," *J. Proc. Cont.*, **8**, 339–353 (1998).

Elicabe, G. E. and G. R. Meira, "Estimation and Control in Polymerization Reactors. a Review," *Poly. Eng. & Sci.*, **28**, 121 (1988).

Ellis, M., T. W. Taylor, V. Gonzalez, and K. F. Jensen, "Estimation of the Molecular Weight Distribution in Batch Polymerization," *AIChE J.*, **34**, 1341–1353 (1988).

Ellis, M. F., T. W. Taylor, V. Gonzalez, and K. F. Jensen, "Estimation of the Molecular Weight Distribution in Batch Polymerization," *AIChE J.*, **34**, 1341 (1994).

Ettedgui, B., M. Cabassud, M.-V. Le Lann, N. L. Ricker, and G. Casamatta, NMPC-based Thermal Regulation of a Fed-Batch Chemical Reactor Incorporating Parameter Estimation, In *IFAC Symposium on Advanced Control of Chemical Processes*, pages 365–370, Banff, Canada (1997).

Evenson, G. F., S. G. Ward, and R. L. Whitmore, "Theory of size distributions: Paints, coals, greases,etc. Anomalous viscosity in model suspensions," *Discussions f the Faraday Society*, **11**, 11–14 (1951).

Farris, R. J., "Prediction of the viscosity of multimodel suspensions from unimodal viscosity data," *Trans. Soc. Rheology*, **12**(2), 281–301 (1968).

Gilbert, R. G., *Emulsion polymerization: a mechanistic approach*. Academic Press (1995).

Gilbert, R. G., Modelling Rates and Molar Mass Distributions, In Lovell, P. A. and M. S. El-Aasser, editors, *Emulsion Polymerization and Emulsion Polymers*, chapter 5, pages 165–203. John Wiley and Sons (1997).

Grady, M., Personal communication (2000).

Jo, J. H. and S. G. Bankoff, "Digital Monitoring and Estimation of Polymerization Reactors," *AIChE J.*, **22**, 361 (1996).

Joseph, B. and C. B. Brosilow, "Inferential Control of Processes: 1. Steady State Analysis and Design," *AIChE J.*, **24**, 485–492 (1978a).

Joseph, B. and C. B. Brosilow, "Inferential Control of Processes: 3. Construction of Optimal and Suboptimal Dynamic Estimators," *AIChE J.*, **24**, 500–509 (1978b).

Joseph, B. and F. W. Hanratty, "Predictive Control of Quality in a Batch Manufacturing Process Using Artificial Neural Networks," *Ind. Eng. Chem. Res.*, pages 1951–1961 (1993).

Kim, K. J. and K. Y. Choi, "On-line Estimation and Control of a Continuous Stirred Tank Polymerization Reactor," *J. Proc. Cont.*, **1**, 96 (1991).

Kiparissides, C., J. F. MacGregor, S. Singh, and A. E. Hamielec, "Continuous Emulsion Polymerization of Vinyl Acetate. Part III: Detection of Reactor Performance by Turbidity Spectra and Liquid Exclusion Chromatography," *Can. J. Chem. Eng.*, **58**, 65 (1980).

Kiparissides, C., "Polymerization Reactor Modeling: a Review of Recent Developments and Future Directions," *Chem. Eng. Sci.*, **51**, 1637 (1996).

Kozub, D. J. and J. F. MacGregor, "State Estimation for Semi-Batch Polymerization Reactors," *Chem. Eng. Sci.*, **47**, 1047–1062 (1992).

Krieger, I. M. and T. J. Dougherty, "A mechanism for non-Newtonian flow in suspensions of rigid spheres," *Trans. Soc. Rheology*, **3**, 137 (1959).

Kumar, S. and D. Ramkrishna, "On the solutions of population balance equations by discretization—I. A fixed pivot technique," *Chem. Eng. Sci.*, **51**(8), 1311–1332 (1996).

Kumar, S. and D. Ramkrishna, "On the solution of population balance equations by discretization–III. Nucleation, growth and aggregation of particles," *Comput. Chem. Eng.*, **52**(24), 4659–4679 (1997).

Lee, J. H. and M. Morari, "Robust Inferential Control of Multi-rate Sampled-data Systems," *Chem. Eng. Sci.*, **47**, 865 (1990).

Lee, J. H., M. S. Gelormino, and M. Morari, "Model Predictive Control of Multi-Rate Sampled-Data Systems: A State-Space Approach," *Int. J. Control*, **55**(1), 153–191 (1992).

Liotta, V., C. Georgakis, and M. S. El-Aasser, Real-time Estimation and Control of Particle Size in Semi-Batch Emulsion Polymerization, In *Proc. American Control Conf.*, pages 1172–1176, Albuquerque, NM (1997).

Marlin, T. E., *Process Control: Designing Processes and Control Systems for Dynamic Performance*. McGraw-Hill, Inc. (1995).

Mutha, R. K., W. R. Cluett, and A. Penlidis, "On-Line Nonlinear Model-Based Estimation and Control of a Polymer Reactor," *AIChE J.*, **43**(11), 3042–3058 (1997).

Neogi, D. and C. E. Schlags, "Multivariable statistical analysis of an emulsion batch process," *Ind. Eng. Chem. Res.*, **37**, 3971–3979 (1998).

Niemiec, M. and C. Kravaris, Nonlinear Multirate Model Algorithmic Control and its Application to a Polymerization Reactor, In *AIChE Annual Meeting* (1997).

Nomikos, P. and J. F. MacGregor, "Multi-way partial least squares in monitoring batch processes," *Chemometr. Intell. Lab.*, **30**, 97–108 (1995).

Nunes, R. W., J. R. Martin, and J. F. Johnson, "Influence of Molecular Weight and Molecular Weight Distribution on Mechanical Properties of Polymers," *Poly. Eng. & Sci.*, **4**, 205 (1982).

Ogunnaike, B. A., "On-Line Modelling and Predictive Control of an Industrial Terpolymerization Reactor," *Int. J. Control*, **59**(3), 711–729 (1994).

Ogunnaike, B. A., A Contemporary Industrial Perspective on Process Control Theory and Practicce, In *DYCORD '95* (1995).

Ohshima, M. and S. Tomita, Model-based and Neural-net-based On-line Quality Inference System for Polymerization Processes, In *AIChE Annual Meeting* (1995).

Ohshima, M., I. Hashimoto, H. Ohno, M. Takeda, T. Yoneyama, and F. Gotoh, "Multirate Multivariable Model Predictive Control and its Application to a Polymerization Reactor," *Int. J. Control*, **59**, 731 (1994).

Ohshima, M., A. Koulouris, S. Tomita, and G. Stephanopoulos, Wave-Net Based On-Line Quality Inference System for Polymerization Processes, In *4th IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes*, page 275 (1995).

Ottewill, R. H., *The stability and instability of polymer latices*, chapter 1. Academic Press (1982).

Parkinson, C., S. Masumoto, and P. Sherman, "The influence of particle-size distribution on the apparent viscosity of non-newtonian dispersed systems," *J. Colloid and Interface Sci.*, **33**(1), 151–161 (1970).

Ramkrishna, D., *Population Balances*. Academic Press, San Diego (2000).

Rawlings, J. B. and W. H. Ray, "The Modeling of Batch and Continuous Emulsion Polymerization Reactors. Part I: Model Formulation and Sensitivity to Parameters," *Poly. Eng. Sci.*, **28**, 237–256 (1988a).

Rawlings, J. B. and W. H. Ray, "The Modeling of Batch and Continuous Emulsion Polymerization Reactors. Part II: Comparison With Experimental Data From Continuous Stirred Tank Reactors," *Poly. Eng. Sci.*, **28**, 257–274 (1988b).

Ray, W. H., "Polymerization Reactor Control," *IEEE Cont. Sys. Mag.*, **6**((4)S), 3–8 (1986).

Ray, W. H., "Modeling of Addition Polymerization Processes—Free Radical, Ionic, Group Transfer, and Ziegler-Natta Kinetics," *Can. J. Chem. Eng.*, **69**, 626 (1991).

Robertson, D. G., P. Kesavan, and J. H. Lee, Detection and Estimation of Randomly Occurring Deterministic Disturbances, In *Proc. American Control Conf.*, page 4453, Seattle, WA (1993).

Russell, S. A., P. Kesavan, J. H. Lee, and B. A. Ogunnaike, Recursive Data-Based Prediction and Control of Product Quality for Batch and Semi-batch Processes Applied to a Nylon 6,6 Autoclave, In *AIChE Annual Meeting, paper no. 196* (1997).

Schoonbrood, H. A. S., H. M. G. Brouns, H. A. Thijssen, A. M. van Herk, and L. German, "The Effect of Composition Drift and Copolymer Microstructure on Mechanical Bulk Properties of Styrene-Methyl Acrylate Emulsion Copolymers," *Macromol. Symp.*, **92**, 133 (1995).

Schork, F. J. and W. H. Ray, "On-line Measurement of Surface Tension and Density with Application to Emulsion Polymerization," *J. Appl. Poly. Sci.*, **25**, 407 (1983).

Schuler, H. and Z. Suzhen, "Real-time Estimation of Chain Length Distribution in a Polymerization Reactor," *Chem. Eng. Sci.*, **40**, 1891 (1985).

Seborg, D. E., T. F. Edgar, and D. A. Mellichamp, *Process Dynamics and Control*. John Wiley and Sons, Inc., New York, NY (1989).

Sereno, C., A. Rodrigues, and J. Villadsen, "The moving finite element method with polynomial approximation of any degree," *Comput. Chem. Eng.*, **15**(1), 25–33 (1991).

Soroush, M. and C. Kravaris, "Nonlinear Control of a Polymerization CSTR with Singular Characteristic Matrix," *AIChE J.*, **40**, 980 (1994).

Soroush, M., Directionality and Windup Compensation in Nonlinear Model-Based Control, In Berber, R. and C. Kravaris, editors, *Nonlinear Model-Based Process Control*, volume 353 of *NATO ASI Series E*, page 173. Kluwer Academic Publishers, Dordrecht (1998a).

Soroush, M., "State and Parameter Estimations and their Applications in Process Control," *Comput. Chem. Eng.*, **23**, 229–245 (1998b).

Sriniwas, G. R., Y. Arkun, and F. J. Schork, "Estimation and Control of an Alpha-olefin Polymerization Reactor," *J. Proc. Cont.*, **5**, 303 (1995).

Tatiraju, S., N. Zambare, M. Soroush, and B. A. Ogunnaike, Multirate Control of a Polymerization Reactor, In *AIChE Spring Meeting* (1998b).

Tatiraju, S., M. Soroush, and B. A. Ogunnaike, Multi-Rate Nonlinear State Estimation in a Polymerization Reactor, In *Proc. American Control Conf.*, page 3165, Philadelphia, PA (1998a).

Tatiraju, S., M. Soroush, and B. Ogunnaike, "Multi-rate Nonlinear State Estimation with Application to a Polymerization Reactor," *AIChE J.*, **45**(4), 769 (1999).

van Dootingh, M., F. Viel, D. Rakotopara, J. P. Gauthier, and P. Hobbes, "Nonlinear deterministic observer for state estimation: Application to a continuous free radical polymerization reactor," *Comput. Chem. Eng.*, **16**, 777–791 (1992).

Wright, G. T. and T. F. Edgar, "Nonlinear Model Predictive Control of a Fixed-Bed Water-Gas Shift Reactor: An Experimental Study," *Comput. Chem. Eng.*, **18**(2), 83–102 (1994).

Particle Size and Shape Control in Crystallization Processes

Richard D. Braatz*
Department of Chemical Engineering
University of Illinois
Urbana, IL 61801, U.S.A.

Shinji Hasebe†
Department of Chemical Engineering
Kyoto University
Kyoto 606-8501, Japan

Abstract

Crystallization from solution is an industrially important unit operation due to its ability to provide high purity separations. The control of the crystal size distribution can be critically important for efficient downstream operations such as filtration and drying, and product effectiveness (e.g., bioavailability, tablet stability). This paper provides an overview on recent developments in the control of industrial crystallization processes. This includes descriptions of recent activities in the modeling, parameter estimation, state estimation, analysis, simulation, optimization, control, and design of crystallization processes.

Keywords

Crystallization, Particulate processes, Population balance models, Distributed parameter systems, Optimization, Control, Estimation, Simulation

Introduction

Crystallization from solution is an industrially important unit operation due to its ability to provide high purity separations. For efficient downstream operations (such as filtration and drying) and product effectiveness (e.g., bioavailability, tablet stability), the control of the crystal size distribution can be critically important. Also important is crystal purity and the crystal shape. The purity of the crystals is especially important in the food industry, where the crystals are consumed by humans. The crystal size and shape affects the dissolution rate, which is important for crystals that dissolve during final use (Winn and Doherty, 2000). In the pharmaceutical industry, the relative impact of drug benefit versus adverse side effects can depend on the dissolution rate. Control of crystal size and shape can enable the optimization of the dissolution rate to maximize the benefit while minimizing the side effects. For crystals used in photography, the size and shape uniformity is the principle concern of the customer (Miller and Rawlings, 1994). Poor control of crystal size and shape can result in unacceptably long filtration or drying times, or in extra processing steps, such as recrystallization or milling. Purity is especially important in the pharmaceutical industries, in which the crystals will be consumed.

The fundamental driving force for crystallization from solution is the difference between the chemical potential of the supersaturated solution and that of the solid crystal face (Kim and Myerson, 1996; Mullin and Sohnel, 1977). It is common to simplify this by representing the nucleation and growth kinetics in terms of the supersaturation, which is the difference between the solute concentration and the saturated solute concentration. Figure 1 is a schematic of a batch crystallizer where the supersaturation is caused by decreasing the temperature. Another method of supersaturation creation is by adding a solvent for which the solute has a lower solubility—this is often called *drowning out*, or *antisolvent addition*. Yet another common method of creating supersaturation is evaporation.

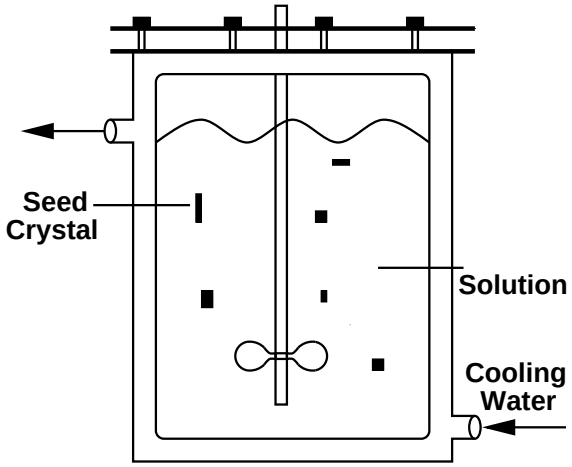


Figure 1: Schematic of a batch cooling crystallizer.

The challenges in crystallization processing are significant. First, there are significant uncertainties associated with their kinetics. Part of the difficulty is that the kinetic parameters can be highly sensitive to small concentrations of contaminating chemicals (Shangfeng et al., 1999), which can result in kinetic parameters that vary with time. Another significant source of uncertainty in industrial crystallizers is associated with mixing. Although crystallization models usually assume perfect mixing, this assumption is rarely true for an industrial-scale crystallizer.

Crystallization processes are highly nonlinear, and are modeled by coupled nonlinear algebraic integro-partial differential equations. For the case of distribution in shape as well as overall size, there are at least three independent variables in the equations (Ma et al., 1999). Simulating these equations can be challenging because

*braatz@uiuc.edu
†hasebe@cheme.kyoto-u.ac.jp

the crystal size distribution can be extremely sharp in practice, and can span many orders of magnitude in crystal length scale (0.01 nm to 200 μm) and time scale (20 μs to 100 min).

Another challenge in crystallization is associated with sensor limitations. The states in a crystallizer include the temperature, the solute concentration, and the crystal size and shape distribution. The solute concentration must be measured very accurately to specify the nucleation and growth kinetics. Because the kinetics are functions of the difference between two solute concentrations (one of these being the saturated solute concentration), the error in the difference is much larger than the error in a single solute concentration measurement. Obtaining an accurate measurement of the full crystal size distribution (CSD) is even more challenging. Hence it is desirable to estimate the states from the noisy measurements that are available.

The last review of efforts to control crystallization processes was published in 1993 (Rawlings et al., 1993). This paper reviews efforts towards the control of crystallization processes, focusing mostly on results since 1993. The next section describes the current status of sensor technologies for crystallization processes. This is followed by descriptions on crystallization modeling including model structure determination and parameter estimation. Activities in state estimation are reviewed. Investigations into the stability of continuous crystallization processes, and the robustness analysis of batch crystallization processes are described. The simulation techniques of method of moments, weighted residuals, discretized population balances, and Monte Carlo simulation are reviewed. On-line and off-line approaches to optimizing and controlling crystallization processes are reviewed, including a discussion of efforts to relate process design to process control. The paper concludes with some predictions on where future efforts are headed.

Measurements

Measurements of both solute concentration and the crystal size distribution are necessary for effective estimation and control.

Solute Concentration Measurement

The nucleation and growth rates are strongly dependent on the solute concentration, making its measurement necessary for estimating kinetic parameters, and highly useful for feedback control. One technique is to measure the refractive index (Helt and Larson, 1977; Mullin and Leci, 1972; Nyvlt et al., 1994; Sikdar and Randolph, 1976). Although this method can work when there is significant change in the refraction index with solute concentration, the method is sensitive to ambient light and air bubbles.

Another approach to obtaining solute concentration measurements is to sample the crystal slurry, filter out the crystals, and then measure the density of the liquid phase. This procedure has been demonstrated on-line for the cooling crystallization of potassium nitrate in water (Matthews III, 1997; Miller and Rawlings, 1994; Riebel et al., 1990; Redman and Rohani, 1994; Redman et al., 1997). The use of an external sampling loop can lead to operational difficulties such as clogging of the screen used to filter out the crystals, and to fluctuations in temperature in the sampling loop. This latter problem is especially important for crystals with a small metastable zone width, where a slight reduction in temperature can cause crystals to nucleate in the densitometer, leading to inaccurate solute concentration measurements.

In the crystallization of electrolytes, the solute concentration can be estimated by placing a conductivity probe in the crystal slurry (David et al., 1991; Franck et al., 1988; Garcia et al., 1999). While avoiding the operational problems associated with sampling, conductivity measurement has its own issues. It has been difficult to apply this technique to batch cooling crystallization processes, because conductivity strongly depends on temperature. Hlozny et al. (1992) and Nyvlt et al. (1994) extended the measurement technique so that temperature effect can be taken into consideration. This technique has been successfully applied to a batch cooling crystallizer (e.g., Jagadesh et al. (1996)).

An indirect method of determining the solute concentration is to use calorimetry, in which the measurements of temperature and flow rates are combined with a dynamic energy balance of the crystallizer (Fevotte and Klein, 1994, 1995, 1996). This approach has been demonstrated for the batch crystallization of adipic acid in water (Monnier et al., 1997). Solute concentration estimates determined from calorimetry can be expected to drift as the crystallization progresses.

Kuehberger and Mersmann (1997b) developed a special device for measurement of supersaturation. When the mother liquor contacts with a cold metal plate, crystals deposit on the surface. According to the amount of the solute deposited, the temperature of the metal plate rises due to the heat of crystallization. Thus, by detecting the temperature rise of the plate, the solute concentration can be estimated.

Loeffelmann and Mersmann (1999) suggested using the difference in electromagnetic properties of the crystal and solution. In this approach, a cooling plate is equipped with electrodes and the impedance between the electrodes is measured. The temperature of the plate is gradually lowered. When deposition of the crystals on the plate is detected by a change in impedance, the temperature is recorded and used to calculate the supersaturation. As an alternative to the impedance, they also suggested using attenuation ratio and phase change of electro-acoustic waves in order to detect crystal deposition.

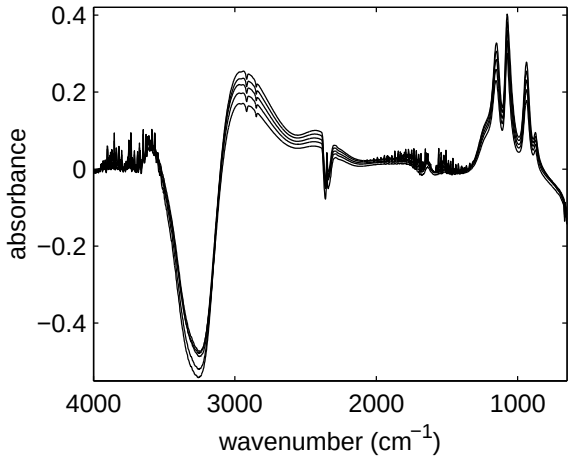


Figure 2: ATR-FTIR spectra for dihydrogen phosphate in water for five solute concentrations.

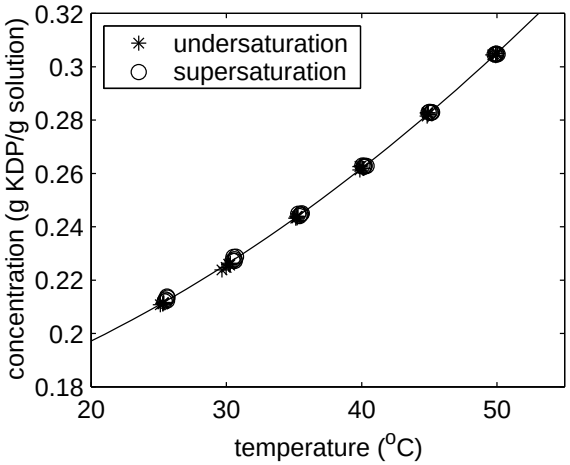


Figure 3: Solubility curve constructed using chemometric model and ATR-FTIR probe inserted into dense crystal slurry.

A limitation to the aforementioned methods for supersaturation measurement is the inability to track the concentrations of multiple dissolved species or multiple solvents.

The feasibility of attenuated total reflection (ATR) Fourier transform infrared (FTIR) spectroscopy for the in situ measurement of solution concentration in dense crystal slurries has been demonstrated (Dunuwila et al., 1994; Dunuwila and Berglund, 1997; Groen and Roberts, 1999; Lewiner et al., 1999, 2001). In ATR-FTIR spectroscopy, the infrared spectrum is characteristic of the vibrational structure of the substance in immediate contact with the ATR immersion probe (e.g., see Figure 2). A crystal in the ATR probe is selected so that the depth of penetration of the infrared energy field into the solution is smaller than the liquid phase barrier between the probe and solid crystal particles. Hence when the ATR probe is inserted into a crystal slurry, the substance in immediate contact with the probe will be the liquid solution of the slurry, with negligible interference from the solid crystals. That the crystals do not significantly affect the infrared spectra collected using the ATR probe has been verified experimentally (Dunuwila et al., 1994; Dunuwila and Berglund, 1997). The combination of ATR-FTIR with advanced chemometrics analysis can measure solute concentrations with accuracy as high as ±0.1 wt% in dense crystal slurries (Togkalidou et al., 2000) (see Figure 3). A significant advantage of ATR-FTIR spectroscopy over most other methods for solute concentration measurement is the ability to provide simultaneous measurement of multiple chemical species.

On-line Crystal Size Distribution Measurement

Several CSD sensors have become available. One is a Coulter Counter (Allen, 1990), which electronically counts particles as the crystal slurry passes through an orifice. Coulter Counters have small flow orifices that are prone to clogging, especially for high density crystal slurries, and may require grounding of the fluid to reduce background noise (Rovang and Randolph, 1980).

The forward light scattering approach is to direct a laser beam through a sample cell, and collect the light scattered through the cell. These instruments, such as the Malvern or the Microtrac particle sizers, can give useful CSD measurement for slurries with low solids density (Eek, 1995; Eek and Dijkstra, 1995; Randolph et al., 1981). Information on particle shape can be determined by examining the light intensity variations (Heffels et al., 1994). This shape information can be used to correct the particle size determination using commercial laser diffraction instrumentation (Heffels et al., 1996). The CSD in dense crystal slurries can be addressed by an automatic sampling and dilution unit (Jager et al., 1987); however, it is challenging to collect a representative sample from an industrial-scale crystallizer and ensure that the temperature is constant enough so that the sample remains representative.

The transmittance, which is the fraction of light that passes through the solution, can be measured either using a light scattering instrument or a spectrophotometer. The projected area of the crystals can be computed from the transmittance. For dense crystal slurries, the transmittance is essentially zero, and no useful information is obtained. The transmittance measurement has been used for the estimation of kinetic parameters for the crystallization of naphthalene in toluene (Witkowski, 1990; Witkowski et al., 1990), potassium nitrate in water (Miller, 1993; Miller and Rawlings, 1994), and a photochemical in heptane (Matthews III, 1997; Matthews and Rawlings, 1998).

An alternative light scattering approach is based on

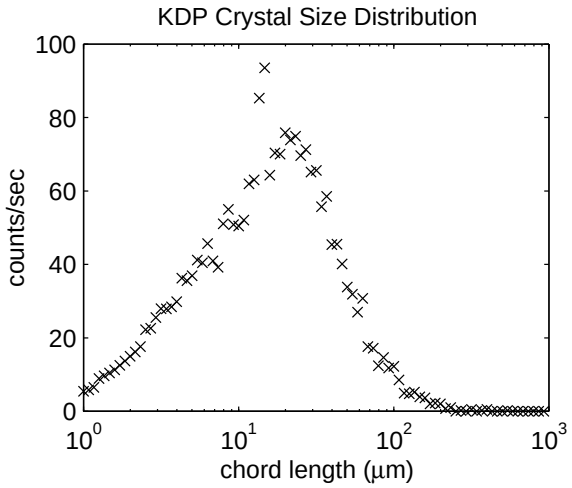


Figure 4: Chord length distribution of KDP collected from a Lasentec FBRM M400L.

inserting a probe directly in the crystallizer, focusing a laser beam forward through a window in the probe tip, and collecting the laser light scattered back to the probe. This approach can measure CSD information even for dense crystal slurries. One of the first commercial instruments of this type, the Par-Tec 100 analyzer, has been used to estimate kinetic parameters for the crystallization of adipic acid in water (Monnier et al., 1996) as well as in feedback control (Redman et al., 1997). Several publications describe applications of updated versions of the instrument, referred to as Lasentec *Focused Beam Reflectance Measurement* (FBRM) (Farrell and Tsai, 1995; Barrett and Glennon, 1999; Ma et al., 1999; Tahti et al., 1999; Togkalidou et al., 2001), which are rugged enough to be implemented on industrial crystallizers.

Like any laser-based method applied to a crystal slurry, a transformation is required to relate the collected laser light to the crystal size distribution. The FBRM instrument measures the chord length distribution (e.g., see Figure 4) as the laser beam emitted from the sensor randomly crosses two edges of a particle, with this distance being the chord length. There have been efforts to relate the chord length distribution to the particle size distribution, both by the Lasentec company and by some independent researchers (Becker, 2001; Clark and Turton, 1988; Han and Kim, 1993; Liu et al., 1998; Worlitschek and Mazzotti, 2000; Simmons et al., 1999; Tadayyon and Rohani, 1998). Chemometrics methods have been used to relate the chord length distribution to other variables, such as filtration resistance (Johnson et al., 1997; Togkalidou et al., 2001).

A weakness of the aforementioned CSD sensors is that the distribution of crystal shape cannot be directly determined. For example, a collection of rod-like crystals are characterized mathematically by a two-dimensional distribution (one dimension being the length, and the other dimension being the breadth), but the light scattering instruments only provide one-dimensional distributions. It is impossible to uniquely determine a two-dimensional distribution from a one-dimensional distribution. The shape information is "averaged out" to obtain a one-dimensional distribution.

Another instrument that has become available recently is the Lasentec Particle and Vision Measurement (PVM) system, in which pictures are taken of the crystals in solution using a probe inserted directly into the dense crystal slurry (Lasentec, 1997). This video microscope can collect 10-30 pictures a second, providing two-dimensional snapshots of the crystals in real time. On-line video microscopy can image crystals as small as 5-15 microns (Lasentec, 1997; Pacek et al., 1994), not as small as obtained by laser scattering instruments. Also, the quality of the images limits the ability of imaging software to automatically identify individual particles (e.g., see Figure 5), and quantify the characteristics of these particles (e.g., maximum axis, minimum axis, aspect ratio). On-line video microscopy has the advantage that the crystals are directly observed, allowing shape information to be obtained. Also, the PVM in particular is a rugged instrument suitable for use in industrial applications. The main use of on-line video microscopy today is for qualitative troubleshooting, with some researchers working on how to use the images for quantitative prediction (Baier and Widmer, 2000; Braatz et al., 2000b). One approach is to use multiway principal component analysis, where features are tracked in the space of principal components (Bharati and MacGregor, 1998). An alternative is to take moments of the images and then to use principal components analysis to relate the image moments to characteristics of the crystals (Braatz et al., 2000b). Given the importance of crystal shape in applications, and that progress becomes easier as computers continue to increase in speed, it seems likely that quantitative predictions will become available.

An alternative approach to on-line video microscopy is to remove slurry from a sampling stream and flow it as a thin film over the focal region of an ordinary light microscope (Eek, 1995; Puel et al., 1997; Rawlings and Patience, 1999). A disadvantage of this approach is the requirement of having a sampling stream in which the crystals may not be representative of what is in the crystallizer. A strong advantage of this approach is that the contrast between crystals and background can be made much sharper, and the number of overlapping crystals can be reduced. The images are sufficiently clean that standard image analysis algorithms can be used (Rawlings and Patience, 1999).

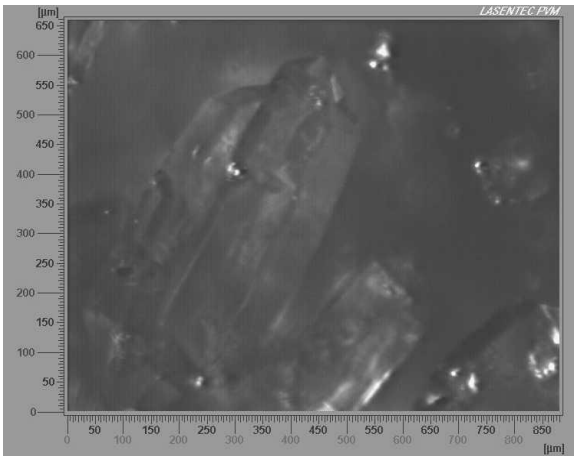


Figure 5: Image of KDP crystals in solution collected from a Lasentec PVM 700L.

Modeling

Model Structure Determination

Derivation of optimal operation patterns or design of control system should be carried out using a model that describes the behavior of the process with sufficient accuracy. The process models can be classified into two groups: physical models, which can be obtained using first principles, and models identified from time series data. Regardless of the type of the model, the model building involves determination of the model structure and estimation of parameters appearing in the model equations. We will discuss the problems of model structure determination in the first half of this section, which is followed by a discussion of some other topics.

There are several papers that discuss methods for modeling the dynamic behavior of the crystallizer using time-series data. Bravi et al. (1995) identified a dynamic model of a continuous crystallizer using artificial neural networks. Time series data generated in a series of simulations using gPROMS were used to train the neural network so that the fines slurry density is predicted from the feed flow rate, the flow rate recirculated through the fines dissolver, and the inlet concentration. The dynamics of the input and output signals were examined to determine the sampling frequency and the number of past data points to use for prediction of the behavior of the fines slurry density.

Rohani et al. (1999a) also conducted a simulation study on the modeling of the crystallizer behavior using ARX and neural network models. The identification data were generated from simulation using a model of a continuous cooling crystallizer that has been previously verified with experimental data. The fines dissolution rate, the clear liquor advance flow rate, and the crystallizer temperature were used as inputs, and output variables were three variables related to crystal size distribution, the purity, and the magma density of the product stream. With the view of using the model in model predictive control (MPC), the prediction performance of each model was examined in terms of 1 step ahead and 50 step ahead predictions.

When a crystallizer is subjected to an excitation signal such as a pseudo-random binary sequence (PRBS), the state variables may deviate far from the steady state due to the strong nonlinearity of the crystallizer. Thus, in order to build a linear dynamic model around the steady state, it is advantageous to collect time series data when the process is under feedback control. Eek et al. (1996) applied closed loop identification techniques to the modeling of the continuous crystallization of ammonium sulfate. Two different closed loop identification methods, direct identification (Soderstrom and Stoica, 1989) and a two-step method (van den Hof and Schrama, 1993), were applied to build three-input three-output models that predict the fines concentration, crystal mean size, and the magma density based on the measurements of fines flow rate, product flow rate, and total heat input. Canonical observability form was the assumed model structure, and the data for identification were obtained by exciting the process using by generalized binary noise sequence (Tullken, 1991). The prediction performance of the resulting model was as good as that of first principle models.

When building ARX or neural network models, the model structure determination and parameter estimation are carried out rather simultaneously. On the other hand, when a first principle model is to be constructed, before estimating parameters it is necessary to determine the model structure suited for the description of the crystallizer dynamics. Therefore, the process of building a first principle model is more complicated than building ARX or neural network models. However, first principle models have several advantages over ARX and neural network models. In particular, the operating range of conditions where a first principle model can provide accurate predictions is wider than for ARX or neural network models. This is because the first principle model incorporates the physical properties of the crystallizer into the model. When the control system is not performing satisfactorily, the first principle model can be used to analyze the cause behind this. Also, a first principle model enables an examination of the relationship between the design and controllability of the process.

When building a physical model of a batch crystallizer, it is usually assumed that the slurry in the crystallizer is perfectly mixed and the spatial distributions of the CSD and supersaturation are negligible. In this case, when modeling a batch crystallizer, it is only necessary to identify the crystal growth rate and nucleation rate (in the absence of agglomeration, attrition, or similar phenomena). On the other hand, the dynamic behavior of

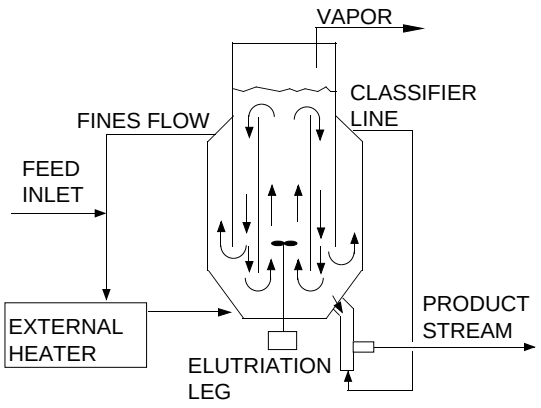


Figure 6: Schematic of a continuous DTB crystallizer.

an industrial continuous crystallizer is influenced by the spatial distribution of CSD and supersaturation. This is because the volume of continuous crystallizer is usually very large. It is said that the continuous crystallization is not economically advantageous over batch crystallization when the production rate is below 500 kg/hr (Wey, 1993). Figure 6 shows the schematic diagram of a continuous DTB evaporative crystallizer. A portion of fines is extracted through the settling zone around the baffle and fed to the external heater, in which the slurry is heated to dissolve crystals. This serves to reduce the total number of crystals and increases the product mean size. Only the crystals large enough to fall against the upward flow in the elutriation leg flow into the product stream. The CSD and supersaturation at each peripheral device are different from those in the crystallization vessel.

Jager et al. (1991) built a model of a continuous crystallizer in such a way that the spatial distribution of supersaturation can be taken into consideration. In their model, the crystallizer is divided into three sections. The changes in CSD and supersaturation in a lump of fluid element is calculated on the assumption that the lump circulates through the three sections in turn. They confirmed that the spatial distribution of supersaturation influences the dynamic behavior of a large scale crystallizer. The idea of circulating lump of fluid element was also applied to simulation of a batch crystallizer by Bohlin and Rasmuson (1996). They concluded that the non-ideal mixing of fluid does not have a strong influence on the performance of batch crystallizers.

Kramer et al. (1996) suggested using a compartmental approach for modeling crystallization processes. In this approach, the crystallizer vessel is divided into smaller parts, each of which is assumed to be well-mixed. Bermingham et al. (1998) presented a heuristic approach for derivation of a compartmental model for a continuous crystallizer. The information needed in this approach can be obtained by making use of computational fluid dynamics (CFD). They suggested that the compartments should be chosen in such a way that the gradients of supersaturation and energy dissipation rate in one compartment are sufficiently small. Sotowa et al. (2000) proposed a method for deriving an expression for classification of crystals between compartments based on the behavior of particles simulated using CFD.

Kramer et al. (1999) showed an approach to the design of a large scale crystallizer using the compartmental model. First, kinetic parameters were identified by conducting an experimental study using a 22-liter crystallizer. The resulting rate expressions were incorporated into compartment models, and the performance for each of the design alternatives were compared. The comparison between the model predictions and the experimental results was presented by Neumann et al. (1999).

In an evaporative crystallizer, the external heater, which provides the energy needed to evaporate solvent, serves to reduce the number of fines by dissolution. In previous studies on the dynamics and control of continuous crystallizers, it was commonly assumed that the fines entering the external heater dissolve completely, and thus the stream recycled from the heater to the vessel does not contain any crystals. However, in practice the fines do not dissolve completely in the external heater, since the residence time of the slurry within the external heater is very short. Naito et al. (1999) built a model of an external heater by taking account of the finite dissolution rate of crystals. The model was derived using a dissolution rate equation that was identified using experimental data. They demonstrated that the behavior of the crystallizer is strongly influenced by the degree of fines dissolution in the external heater.

There are two main approaches to simulating changes in the distribution of crystal shape in a crystallizer. One way is to model shape dynamics as changes in the crystal length-to-width ratio (Matthews and Rawlings, 1998). The alternative is to model the entire n-dimensional crystal size distribution, where n is the number of independent growth faces (Togkalidou and Braatz, 2000). For example, $n = 2$ for modeling rod-like crystals, where one dimension is the crystal length and the other is the crystal width (equal to breadth). Both methods introduce additional model parameters to be estimated.

Although various models have been used in the study of crystallizer control and operation, agglomeration is not usually taken into consideration in the population balance equation. Jagadesh et al. (1996) demonstrated that large crystals can be obtained by loading sufficiently large amount of seed crystals in batch crystallization. They argued that this is because newly born nuclei rapidly agglomerate with other crystals when the number of seed crystals is large. In order to derive an op-

timal policy for such an operation, the model must take agglomeration into consideration. However, the population balance equation with agglomeration is more challenging to analyze, and the identification of the agglomeration rate is more difficult. Recent developments in the analysis of crystallization with agglomeration should be incorporated into dynamic studies.

While many physical models are available for primary nucleation (this is nucleation directly from solution), most industrial crystallizers are seeded, with most of the nucleation occurring from particle-particle and particle-impeller collisions (this is called *secondary nucleation*). A physical model of secondary nucleation rate has been presented by Gahn and Mersmann (1997, 1999a,b). In this model, the number and distribution of secondary nuclei arising from attrition are predicted based on the frequency of crystal collision with the impeller, and the internal stress distribution at the time of each collision. This model of nucleation kinetics would represent a significant advance in crystallization modeling if the proposed model is verified by experimental data.

Progress in the computer technology has enabled the numerical simulation of slurry flow in a crystallizer with complex internal structure (see Simulation Section). CFD simulation provides hints for defining compartmental models, as well as information on particle segregation in the vessel. However, a limited number of studies have been conducted on the use of CFD results in the modeling of crystallizer behavior (Wei and Garside, 1997). More intensive effort should be conducted to enhance the progress in the modeling methodologies.

Estimation of Kinetic Parameters

Once the model structure is specified, the modeling problem is reduced to a parameter estimation problem. In crystallization processes, the most important parameters to be estimated are those related to the kinetics of crystal growth and crystal formation/depletion. In order to build a first principle model of a crystallizer, it is necessary to express each of the rates as a function of the operating conditions and state of the slurry.

When a crystal grows, solute in the bulk is transported to the crystal surface by the concentration gradient and then it is integrated onto the crystal surface. When rate equations for surface integration and mass transfer are combined in series, a more complicated growth rate equation is obtained. It is common in the study of the operation and control of crystallizers to use the empirical power-law expression to describe the crystal growth rate:

$$G = k_g \Delta C^g \tag{1}$$

where G is the crystal growth rate, ΔC is the supersaturation, and k_g and g are parameters that need to be estimated. When the growth rate is assumed to be size-dependent, Equation 1 is modified:

$$G = k_g \Delta C^g G_x(x) \tag{2}$$

where $G_x(x)$ represents variation of growth rate by the crystal size, x. When the effect of temperature on the growth rate cannot be neglected, k_g is assumed to be an Arrhenius type function of temperature.

As mentioned earlier, nucleation can be classified into two types. *Primary nucleation* takes place when the supersaturation is high. In the secondary nucleation, fine fragments of crystalline substance arise as a result of attrition or breakage of crystals which are already present in the liquor. These fragments then grow to be larger crystals. An empirical rate expression for nucleation is:

$$B_1 = k_n \Delta C^n \tag{3}$$

where B_1 represents the number of primary nuclei that arise per unit time. Experimental data must be used to estimate the values of k_n and n. A commonly used rate equation for secondary nucleation is

$$B_2 = k_b M_T^j \Delta C^b \tag{4}$$

where B_2 is the occurrence rate of secondary nuclei, and M_T is the magma density. In this case, three parameters (k_b, j, b) need to be identified using experimental data. Again, the parameters k_n and k_b are usually assumed to be an Arrhenius type function of temperature, when the temperature effect on the nucleation rate needs to be taken into consideration.

For estimation of the values of the parameters in the crystal growth and nucleation rates, the following approach is most commonly used. When a continuous MSMPR (mixed-suspension mixed-product removal) crystallizer is at a static steady state, the relationship between the crystal size, x, and the population density of the crystal, $n(x)$, can be written as:

$$n(x) = n_0 \exp\left(-\frac{x}{G\tau}\right) \tag{5}$$

where n_0 is the population density of nuclei, and τ is the mean residence time. As can be seen from this equation, the population density distribution appears as a straight line on the semi-log plot. The slope and the intercept can be used to calculate the growth rate and the nucleation rate at the operating condition. When data points are collected at various operating conditions, the parameters in Equations 1, 3, and 4 can be identified. There are many other approaches to parameter identification that are based on the population balance equation. Their details can be found elsewhere (e.g., Tavare, 1995).

There have been several studies in which the parameter estimation problem is solved as an optimization problem. Dash and Rohani (1993) used Gauss-Newton method to solve the optimization problem. The parameters to be estimated are the five parameters that appear in the rate equations. The objective function was

defined as the weighted sum of squared errors of the initial solute concentration and the CSD at four instances during batch operation. Qiu and Rasmuson (1994) estimated five parameters in the growth and nucleation rate equations. In the experiments, the solute concentration was measured every five minutes by the density method, and sieve analysis data of the final crystal size distribution was recorded. The optimization algorithm employed was a combination of Gauss-Newton and quasi-Newton methods. Farrell and Tsai (1994) also measured time series data of supersaturation and final crystal size distribution, and used the prediction errors as the objective function. They applied reparametrization to avoid appearance of ill-conditioned Hessian in the optimization problem.

When estimating the parameters of a real process, it is important to assess the reliability of the estimates. Miller and Rawlings (1994) quantified the reliability of the estimates in terms of a confidence ellipsoid and in terms of confidence intervals. Using time series of solute concentration and transmittance data, it was shown that accurate nucleation and growth parameters could be obtained with as little as two batch crystallization experiments. Closely related work showed that appropriate selection of the seed distribution used in the batch experiments results in parameter estimates of higher accuracy (Chung et al., 2000).

Usually studies on the parameter estimation of crystallization processes aim at estimating parameters in the crystal growth and nucleation rate equations. However, in order to derive a model that describes the behavior of an industrial crystallizers, there are many other parameters that need to be estimated. Particularly, a model of large scale continuous crystallizers contains many parameters whose values depend on the structure and dimensions. For a 970-liter continuous crystallizer, Eek et al. (1995a) constructed a model containing fifteen parameters. Six out of fifteen parameter values were fixed using the results from preliminary experiments. The remaining nine parameter values were determined using a parameter estimation algorithm. The objective function was the prediction error of the light intensity data measured by the detector rings of the Malvern particle sizer. Since the number of parameters is large, it is important to evaluate the reliability of the estimates. In their study, the standard deviations of the estimates obtained using different experimental runs were used to assess the magnitude of the parameter uncertainties.

The advantage of constructing model based on bench scale crystallizers is that the experiments are relatively cheap. However, the slurry in an industrial scale crystallizer is not perfectly mixed. Thus, predictions based on a model constructed from bench scale experiments may not agree with the process measurements for an industrial scale crystallizer, when the model is derived by assuming perfect mixing within the crystallizer. In building a model of an industrial process, it would be best to estimate the values of the model parameters using both process measurements (because this is based on the real process) and experimental data obtained from bench scale apparatus (since these measurements are cheap). It is desirable to establish a unified modeling methodology by exploiting various existing modeling techniques such as experimental design and CFD.

State Estimation

The state of a crystallizer is characterized by the crystal size distribution (which can include distributions in shape, age, purity, or other variables), supersaturation, and temperature. Since the volume of a bench scale batch crystallizer is relatively small, the slurry in the vessel is usually assumed to be perfectly mixed. In this case, the state variables depend only on time. In industrial scale crystallization processes, the spatial distribution of state variables cannot be neglected in some cases. Also, even when perfect mixing can be assumed, the crystal size distribution in the crystallizer vessel is significantly different from that in the fines recirculation loop or that in the elutriation leg. Such a continuous crystallizer can be modeled by combination of several compartments in which homogeneity of the state variables can be assumed. In a real crystallization process, only a limited number of state variables can be measured on-line. If values of the unmeasured state variables are required for optimal operation or control, they must be estimated using an state estimator.

For successful state estimation by observers or Kalman filters, it is necessary to use a model that describes the process dynamics with sufficient accuracy. However, since the crystallization process is a distributed parameter system, a population balance equation cannot be directly used in state estimation, instead it is approximated by a finite order system. The moment method is one of the most commonly used model reduction techniques. When a closed set of ordinary differential equations is obtained by the moment method, a low order linear state space model can be easily obtained by linearization. The state space model can be used for designing observers and state feedback controllers (Tsuruoka and Randolph, 1987).

Hashemi and Epstein (1982) discussed the controllability and observability of a crystallizer using a moment model which took the energy balance into account. The flow rate, concentration, and temperature of the feed stream were taken as inputs, and the zeroth moment, third-order moment, solute concentration, and temperature in the vessel were used as measurable variables. The condition numbers of the controllability and observability matrices were used as a measure of the controllability and observability. It was shown that both controllability and observability improve when the crystallizer is oper-

ated at an operating condition with high supersaturation.

Chiu and Christofides (1999) presented a framework for controlling general nonlinear distributed parameter systems. In this article, they suggested using a Luenberger-type nonlinear observer for state estimation, but the systematic procedure to determine the observer gain was not clearly shown.

Eek et al. (1995b), Eek (1995), and de Wolf et al. (1989) derived a 100th order state space model by discretizing the population balance equation using finite differences. Eek et al. (1995b) and Eek (1995) calculated the steady state Kalman gain using the linearized state space model, by assuming that the laser scattering intensity data is the measurable variable. When implementing an observer to their experimental rig, a nonlinear model was used in the observer, which takes nonlinearity of the process into consideration. Estimates of the crystal size distribution were used to calculate mean size, which showed good agreement with the experimentally measured values.

An increase in the number of measurable variables facilitates state estimation. However, as mentioned in the Introduction, the number of measurable variables is limited. In particular, the measurement of the supersaturation can be challenging, though this is the key state variable that is the driving forces for crystal nucleation and growth. Thus, intensive studies have been executed to develop hardware and soft sensors to measure supersaturation (see Measurements).

Analysis

Crystallization processes can illustrate some interesting dynamical behavior, including a high sensitivity to parameter variations. This section focuses on investigations into the analysis of dynamical behavior in continuous and batch crystallization processes.

Analysis in Continuous Crystallization

Continuous crystallization processes can demonstrate undesirable oscillations in the crystal size distribution, even in open loop. A primary cause of the oscillations is by product being removed that has a different population density than the average population density—this is referred to as *classification*. Industrial crystallizers are often designed to remove and dissolve the smallest crystals (fines), and to preferentially remove the larger crystals as product. Both of these practices increase the tendency for the crystal size distribution to oscillate (Ishii and Randolph, 1980; Randolph, 1980; Randolph et al., 1973, 1977). Oscillations can also be caused by a high order relationship between the nucleation rate and the supersaturation (Randolph and Larson, 1988; Sherwin et al., 1967), but this is probably not the most common cause of the oscillations observed in practice (Randolph, 1980).

Most stability analyses are based on linearized stability analysis for a single continuous crystallizer, usually by determining the localized stability of the moment equations, or by calculating the spectrum of the linearized integro-differential operator (Buyevich et al., 1991a; Witkowski and Rawlings, 1987). These instabilities are characterized as Hopf bifurcations. Conditions have been derived for which a crystallization process can exhibit multiple steady states (Lakatos, 1996; Tavare, 1989; Tavare and Garside, 1985). Other investigations have studied the dynamic behavior under forced oscillations (Buyevich et al., 1991b), where it is possible to obtain more complex dynamic phenomena such as resonance horns, quasi-periodic oscillations, and chaos (Lakatos and Blickle, 1995). These studies suggest that it may be wise in practice to suppress oscillatory disturbances (for example, in the feed conditions) to limit the complexity of dynamical behavior exhibited in the crystallizer. The nonlinear dynamical behavior of a cascade of well-mixed crystallizers has also been investigated (Natalukha, 1996). Nonlinear stability analyses are supported by visualization software that produces phase portraits based on the simulation program (Epstein and Sowul, 1980; Lakatos and Blickle, 1995; Witkowski and Rawlings, 1987).

Oscillations can be reduced by manipulation of the bulk throughput rate (Lei et al., 1971) or the fines destruction flowrate (Beckman and Randolph, 1977). The main difficulty with implementing these early schemes was the lack of measurements of the crystal size distribution (Randolph, 1980). Modern instrumentation makes such schemes implementable (see Measurements section).

Analysis in Batch Crystallization

Stability in a strict mathematical sense is not an issue in batch or semibatch crystallization processes, since the states of such a process cannot blow up in finite time. On the other hand, having consistent product quality during parameter variations or disturbances is a concern. Miller and Rawlings (1994) provided a clear analysis of the effect of model uncertainties on the product quality in batch crystallizers. Matthews et al. (1996) investigated the sensitivity of the optimal supersaturation profiles to seed loading, profile duration, and the difference in supersaturation orders for nucleation and growth. Also, the singular value decomposition was used to calculate a vector of perturbations in the model parameters that has the strongest effect on the supersaturation profile. Several researchers have shown that the quality of the product crystals is sensitive to the performance of the tracking control to the optimal temperature profile (Chianese et al., 1984; Bohlin and Rasmuson, 1992; Ma et al., 1999a).

An approach was developed that quantifies the im-

pact of such variations on the product quality without exhaustive simulation of all possible process conditions (Ma et al., 1999a; Ma and Braatz, 2001). The knowledge of the worst-case model parameters can be used to determine where experimental effort should be focused to improve model accuracy. The robustness analysis with regard to control implementation uncertainties can guide the selection of the control instrumentation, by determining where high precision sensing and actuation are required. The computation of the worst-case external disturbances determines which disturbances significantly affect the product quality and should be suppressed by redesign of the process or feedback control. The approach was applied to batch crystallization simulations, including to the multidimensional growth of crystals used in nonlinear optics applications, where the nominal parameters and uncertainties were quantified from experimental data (Ma et al., 1999a,b; Ma and Braatz, 2001). Robustness estimates were provided with reasonable computational requirements. It was found that a temperature deviation of $\pm 0.1K$ from the optimal profile could result in substantial reductions in the product quality.

Simulation

A significant roadblock to the development of estimation and control strategies for crystallization processes, especially for crystals that change shape during the growth process, is the lack of efficient simulation schemes for the population balance equations. Many simulation studies on crystal growth have been directed toward the solution of the population balance equation for unidirectional crystal growth:

$$\frac{\partial f}{\partial t} + \frac{\partial \{G[c(t), T(t), r] f\}}{\partial r} = h(r, t) \qquad (6)$$

where $f(r, t)$ is the crystal size distribution, t is time, r is the internal spatial coordinate (e.g., crystal size), c is the solute concentration, T is the temperature, G is the growth function, and h is the crystal creation/depletion function. This equation is augmented with associated algebraic and/or integro-differential equations to describe the energy balance, aggregation, breakage, growth, and nucleation phenomena. Simulating these equations is challenging because the crystal size distribution can be extremely sharp in practice, and can span many orders of magnitude in crystal length scale (0.01 nm to 200 μm) and time scale (20 μs to 100 min).

Several numerical techniques have been proposed (Ramkrishna, 1985). The techniques can be separated into four broad categories:

1. method of moments, in which only lower order moments of the crystal size distribution are simulated, and unknown parameters of an assumed distribution are fitted to the computed moments (Hulburt and Katz, 1964)
2. weighted residuals/orthogonal collocation methods, in which the solution is approximated as linear combinations of basis functions (Singh and Ramkrishna, 1977)
3. finite difference methods/discretized population balances, in which (6) is replaced by difference schemes (Kumar and Ramkrishna, 1996a)
4. Monte Carlo simulation, in which the histories of individual particles are tracked, each exhibiting random behavior in accordance with a probabilistic model (Maisels et al., 1999; Shah et al., 1977; Song and Qiu, 1999).

The advantage of the method of moments is that only a small number of ordinary differential equations needs to be solved when the moments are closed (that is, form a finite number of equations describing the lower order moments which are not a function of the higher order moments). A weakness of the method of moments is that the moment equations are not closed for most processes, leading to an infinite number of coupled ordinary differential equations to solve. Another weakness is that, even when the moment equations are closed, the numerical errors in a fitted assumed distribution can be arbitrarily large if the assumed distribution does not accurately parameterize the true distribution. Hence a *general* numerical solution of the population balance equation cannot be developed based on the method of moments. However, the method of moments does apply to many well-mixed batch and continuous crystallizers with nucleation and growth. These assumptions can be reasonable in bench scale crystallizers such as used in teaching laboratories (Braatz et al., 2000a). The method of moments is also useful for testing the accuracy of more sophisticated numerical simulation codes.

In the application of the method of weighted residuals to the population balance equation, the population density is approximated by a linear combination of user-specified time-independent basis functions with time-dependent weighting factors. The basis functions are selected so that the population density can be well approximated with only a finite number of terms. The linear combination of basis functions is substituted into the population balance equation, and ordinary differential equations for the coefficients are derived with the intent to minimize the error (or residual) in the population balance equation. The system of ordinary differential equations can be solved using any standard solver (Barton et al., 1998). A fast numerical algorithm results when only a small number of terms are needed in the expansion, which has been demonstrated for some crystallizers (Rawlings et al., 1992; Witkowski and Rawlings, 1987). The primary weakness of the method of weighted residuals is that basis functions that work well for one

type of crystallization process may not work well for another, which makes it difficult to derive a general fast algorithm for crystallization simulation using this method. This also applies to orthogonal collocation, which is essentially a class of weighted residual algorithms. Reviews of early work on the method of weighted residuals are available (Ramkrishna, 1985; Rawlings et al., 1993), including summaries of algorithms that combine orthogonal collocation with finite elements (Gelbard and Seinfeld, 1978).

Several discretizations of the population balance equation have been investigated and have been applied to various particulate systems (Gelbard et al., 1980; Hounslow, 1990; Hounslow et al., 1988; Marchal et al., 1988; Muhr et al., 1996). This includes an application to the simulation of a crystallization process in which the crystals have two characteristic growth axes, so that changes in the crystal shape distribution are simulated (Puel et al., 1997). Many of these algorithms were formulated with the intent to conserve moments of the computed population density. Different algorithms conserve different moments, and several choices of discretization points have been investigated (Batterham et al., 1981; Kumar and Ramkrishna, 1996b; Litster et al., 1995). Kumar and Ramkrishna (1996a) provide a critical review of these algorithms, including pointing out technical errors in some of the papers. Various numerical problems can occur when performing direct discretizations of the population balance equations. An approach that removes these problems is to combine the discretization with the method of characteristics (Kumar and Ramkrishna, 1997; Sotowa et al., 2000), which has been applied to particulate processes with pure growth, simultaneous aggregation and growth, and simultaneous nucleation and growth (Kumar and Ramkrishna, 1997).

High resolution finite difference schemes also avoid the numerical problems typically associated with discretizing population balance equations (Ma et al., 2001). The high resolution methods are able to obtain second-order accuracy without the undesirable oscillations that can occur with naive second-order methods. A high resolution method that exploits sparsity and efficiently manages memory resulted in a highly accurate dynamic simulation of the multidimensional crystal size distribution for a system with an extremely sharp distribution (see Figure 7), with the entire computation time in less than 10 minutes on a workstation. This was a simulation of a batch crystallizer which produced prism-like crystals with two characteristic length scales and nonlinear nucleation and growth rates. Numerical analysis indicates that the method can allow a coarse time discretization, which is one of the main reasons for the short computation times (Ma et al., 2001).

The use of computational fluid dynamics (CFD) codes is suitable for the simulation of crystallizers that are not perfectly mixed, since in this case the simulation

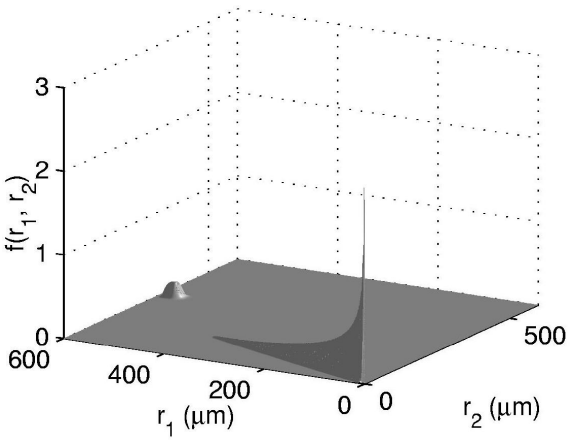


Figure 7: Population density function for rod-like crystals produced by nucleation and growth.

is best handled by solving the complete transport equations (Sha et al., 1999). CFD codes use either finite elements or finite volume methods, in which the conservation equations are applied directly to subregions to obtain numerical values for the variables of importance (Koenig, 1998). While such codes should probably be applied in the design of any industrial scale crystallizer, the computations are rather intensive for such simulations to be used for the development of estimation and control algorithms.

Monte Carlo methods are especially suitable for simulating stochastic population balance equations, and for especially complex systems (Ramkrishna, 1985). The number of papers applying Monte Carlo techniques has rapidly grown in recent years. Processes that have been simulated include:

1. a continuous crystallizer with size-dependent growth rate (Lim et al., 1998),
2. protein crystal growth (Durbin and Feher, 1991), including the case where both monomers and aggregates attach to the crystal surface (Ke et al., 1998; Strom and Bennema, 1997)
3. imperfectly mixed draft tube baffled and forced circulation crystallizers (Lim et al., 1999)
4. a crystallizer with attrition, in which there is a distribution of volumetric shape factors (Lim et al., 1999)
5. crystallizers with simultaneous growth rate dispersion and aggregation (Van Peborgh Gooch and Hounslow, 1996; Van Peborgh Gooch et al., 1996)
6. continuous crystallization of sodium chloride (Sen Gupta and Dutta, 1990b) and sucrose (Sen Gupta and Dutta, 1990a)

An advantage of Monte Carlo methods is that such code is relatively easy to write. A disadvantage of Monte Carlo methods is that they can be rather computationally expensive, which is a drawback when incorporating such models into estimation and control algorithms. Also, the main capabilities provided by Monte Carlo methods—the ability to handle nearly arbitrary stochastic phenomena and to handle extremely complex systems—may not be needed for most industrial scale crystallizers. The measurement noise is probably larger than other stochastic phenomena for most industrial scale crystallizers (Rawlings et al., 1993), in which case an adequate model can be obtained by appending additive stochastic variables to the results of a deterministic population balance equation simulation. Recent papers have shown that non-Monte Carlo simulation techniques (such as method of moments and finite differences) can be applied to more complex multidimensional crystallization processes, without requiring a significant increase in algorithm complexity (Ma et al., 1999, 2001; Togkalidou and Braatz, 2000).

Optimal Operation

The quality of crystals is determined by various factors including mean size, crystal size and shape distribution, and purity. An optimal operation problem of a crystallizer is formulated as a problem of finding the operating conditions that optimize an objective function defined by these factors. This section focuses on such optimization problems for an existing crystallizer, and the optimal sizing problem of crystallizers will not be dealt with.

The study of optimal operating conditions for crystallizers was initiated by Ajinkya and Ray (1974). Subsequent studies on the optimal operation of crystallizers have usually focused on maximizing the crystal mean size or minimizing the coefficient of variation. The reason for this is as follows. The crystals produced in a crystallizer are separated from the mother liquor in its downstream processes in order to avoid both inclusion of mother liquor as impurities and formation of bonding between crystals. When the crystal mean size is small, a large amount of mother liquor is retained between crystals due to capillary attraction. In order to facilitate filtration and drying operation, it is desirable to produce large and mono-dispersed crystals.

Usually, formation of an excessive number of nuclei results in smaller crystals with broad distribution. Thus, it is important to operate the crystallizer in such a way that unnecessary nucleation can be minimized.

Most of the studies on the optimal operation of batch crystallizers focus on the derivation of the optimal temperature profile in batch cooling crystallization. When a batch cooling crystallizer is operated in a natural cooling mode, the slurry temperature decays just like the step response for a first order system. This means that at the beginning of the operation, the reduction in temperature is very quick, and a large number of nuclei arise as a result of high supersaturation. On the other hand, temperature changes only slowly near the end of operation, and the nucleation rate becomes small. The first discussion on the temperature profile that increases the mean size is conducted by Mullin and Nyvlt (1971). They suggested that in order to produce large crystals, the crystallizer temperature should be changed in such a manner that the nucleation rate remains constant during the whole operation (Mullin and Nyvlt, 1971; Jones and Mullin, 1974). Rohani and Bourne (1990b) presented a simple method to calculate the temperature profile that makes nucleation rate or supersaturation constant during the operation. A method for calculating the optimal supersaturation level in batch crystallization is suggested by Mersmann (1995) and Kuehberger and Mersmann (1997a).

In the aforementioned studies, the optimal temperature trajectory has been derived on the assumption that constant nucleation rate or constant supersaturation is optimal. On the other hand, there are studies in which the temperature profile is obtained as the solution of an optimization problem using the quality of product crystals as the objective function. Jones (1974) employed the maximum principle to obtain the temperature profile that maximizes the size of the seed crystals at the final time. Chang and Epstein (1982) employed the mean size, total crystal volume, and the variance of distribution as the objective function, and the optimal temperature profile for each objective is calculated using the maximum principle.

In recent studies, the problem of deriving the optimal temperature profile is formulated as a nonlinear optimization problem, which is then solved using general purpose optimization algorithm. In Miller and Rawlings (1994), the temperature profile that maximizes the ratio of final seed mass to mass of nucleated crystals was obtained using a successive quadratic programming (SQP) code. Lang et al. (1999) used collocation on finite elements and reduced SQP to obtain the optimal temperature profile of a cooling medium that maximizes the mean size of product crystals produced by an unseeded batch crystallization.

In cooling crystallization, the temperature profile is often taken as the only optimization variable, although there are many other factors that determines the quality of the product crystals. Chianese et al. (1984) examined the impact of various parameters associated with the operation of batch crystallizers on the crystal size distribution of the product. They showed that the agitation rate, mass of seed crystals, and the distribution of seed crystals as well as temperature profile are important parameters that strongly affect the crystal size distribution of products.

There are several studies that report the influence of

the mass and size of seed crystals on the product quality. Moore (1994) argued that, to have consistent crystal product quality in batch crystallizers, the operating condition should be determined so that three-σ variation in the initial conditions should not affect the final product qualities. In such an operation, the seed mass is at least 0.5-2.0% of the product mass, which is much greater than the seed mass employed in some conventional crystallizer operations.

Jagadesh et al. (1996) showed that when a sufficiently large amount of seed crystals is loaded, a large and monosized crystals are obtained even under natural cooling. They also presented a "seed chart" that can be used to analyze experimental data to find the seed mass above which formation of secondary nuclei becomes essentially negligible. Doki et al. (1999) experimentally verified their approach using a 600-liter pilot scale crystallizer.

Chung et al. (1999) formulated an optimal operation problem in which the seed mass, the mean size of seed crystals, the width of the seed crystal size distribution, and the temperature profile were decision variables. Three objective functions were studied: the mean size of product crystals, the ratio of standard deviation to mean size, and the ratio of nucleated crystal mass to seed crystal mass at the end of operation. The optimal solution for each objective function was calculated using SQP. A parametric analysis showed the significant importance of optimization of the seed distribution for a wide range of possible nucleation and growth kinetics.

Agitation affects the secondary nucleation rate as well as the degree of mixing. Currently, there is no general model that predicts the effects of agitation on the secondary nucleation rate and the crystal growth rate, as it depends on the dimension of the vessel and impeller. As a result, despite its strong influence on the product qualities (Chianese et al., 1984), the agitation rate was not treated as an optimization variable in almost all previous studies. It is desirable that the impact of agitation on the product quality should be modeled, so that the agitation rate can be included in the decision variables in the optimization problem.

As discussed in the Analysis section, the optimal operating condition derived from an off-line calculation may not be the true optimal profile due to the uncertainties in the model. Also, if there is an error in tracking the optimal profile, the resulting product quality becomes different from the optimal. Ge et al. (2000) focused on the problem of plant-model mismatch and errors in the initial condition, and suggested an optimization method called iterative dynamical optimization, in which the operation profile is modified from batch to batch to improve the performance. In this approach, the operation data of previous batch runs are used to derive the plant model, and then the temperature profile is updated by solving optimization problem using conjugate gradient method.

Under the presence of disturbance, modeling error, or tracking error, the states of the crystallizer do not follow the optimal path. One way to address this problem is to incorporate robustness into the computation of the optimal path (Ma and Braatz, 2000). However, the performance of this approach will be limited by the chosen measured variables and the use of open loop optimization. Another way to address this problem is to choose another measurable variable as the controlled variable in the tracking control. As mentioned previously, the realized temperature profile has a strong influence on the quality of the product crystals. If the variation in product quality due to the modeling or tracking errors can be reduced by choosing a variable other than temperature as the controlled variable, an alternative configuration of the tracking control system should be studied. If such a variable is not directly measurable, a state estimation algorithm should also be developed.

Yet another approach is on-line optimization (Eaton and Rawlings, 1990; Rawlings et al., 1993). If the optimal profile is recomputed at regular intervals based on the state variables at each instance, the effects of various disturbances and uncertainties can be reduced. To carry out dynamic optimization using a physical model, estimates of the state variables must be known at each instance. Thus, the on-line optimization system should consist of the following subsystems (Noda et al., 2000):

1. A subsystem that estimates the current values of the state variables from past and present measurements.
2. A subsystem that derives the optimal trajectory from the current time.
3. A subsystem that controls the state variables according to the optimal path calculated by subsystem 2.
4. A subsystem that modifies the process model using the prediction and measurements.

There is a limited number of studies on the optimal operation of continuous crystallizers. This is probably due to the sustained oscillation of the crystal size distribution which can be observed in the operation of many continuous crystallization processes. Due to this phenomenon, research on the operation of continuous crystallizers have focused on development of stabilizing controllers. Recently, Hasebe et al. (1999) discussed the optimal operation of a continuous DTB crystallizer. The objective of their optimization was the maximization of the production rate of crystals which are larger than a specified size.

Under sustained oscillation, three different types of operation can be readily identified. The first type is to maintain the manipulated variables at the optimal value, while the crystal size distribution is allowed to oscillate. The second type is to periodically change the manipulated variables according to the optimal patterns. In the third type of operation, the oscillation is suppressed

using a stabilizing controller, and the crystallizer is operated at the optimal static steady state. Once a model of the crystallizer is obtained, optimal operating conditions for the first and the third types of operations can be easily obtained by solving a constrained optimization problem. However, it is difficult to derive the optimal manipulation pattern in the second type of operation. This is because the period of oscillation depends on the manipulation pattern, and the period of oscillation under optimal condition is not known before the optimization calculation. To overcome the difficulty, the manipulated variables can be defined as functions of the state variables rather than as functions of time (Hasebe et al., 1999). In their study, one cyclic period of an oscillatory variable was divided into eight phases according to the sign of the gradient and the value of the variable. It was assumed that the manipulated variable takes a unique value in each phase. With this technique, the problem of finding the optimal operation pattern is converted into a problem of finding eight optimal parameters, and the converted problem can be solved using standard optimization algorithms. The result of the optimization shows that a greater amount of large crystals can be obtained by changing the manipulated variable according to the optimal pattern, as compared with the case where the manipulated variables are maintained at the optimal values. It is extremely difficult to suppress oscillation of crystal size distribution in an industrial scale continuous crystallizer. However, the optimization result also shows that the production rate of large crystals increases dramatically, if the behavior of the crystal size distribution can be stabilized.

Sotowa et al. (1999a) demonstrated that the "ease of control" varies greatly with the operating condition. This study suggests that, when deriving the optimal operating condition for a continuous crystallizer, it is important to take controllability issues into consideration. Problems related to the interaction between the design and control are discussed in a later section.

Control

The focus of this section is on feedback control.

Early investigations in crystallization control were directed towards the stabilization of oscillations or other fluctuations in continuous crystallizers (Beckman and Randolph, 1977; Lei et al., 1971). An experimental study showed that fluctuations in the CSD can be reduced by feedback control, by measuring the crystal size distribution in the fines stream and manipulating the fraction of fines recycled back to the crystallizer (Randolph et al., 1987). A more recent study has shown that changes in the operating condition for a crystallizer can greatly affect the ability of a conventional controller to stabilize open loop oscillations (Sotowa et al., 1999a,b).

Many industrial jacketed batch crystallizers use PI control to follow a specified temperature trajectory, with the manipulated variable being the setpoint to a lower level control loop on the flow to the jacket. When manipulating a fines dissolution rate, a self-tuning controller outperformed a PI controller for a potash alum batch crystallizer (Rohani and Bourne, 1990a). Model predictive control has been used to follow a desired temperature trajectory in a batch jacketed crystallizer, using the temperature of the incoming water to the jacket as the manipulated variable (Matthews III, 1997; Miller, 1993). The jacket water temperature was used as a setpoint to a PID slave controller that adjusted a 3-way valve that blended hot water and cold water streams. A nonlinear model predictive control algorithm was applied to an experimental crystallization apparatus with two inputs and two outputs (Eek, 1995; Eek et al., 1995b). More recently, a multivariable nonlinear model predictive controller has been applied to a KCl cooling crystallizer (Rohani et al., 1999b).

Non-MPC nonlinear feedback control algorithms have been applied to crystallization processes (Chidambaram and Malleswararao, 1990). Some recent efforts have been directed towards taking model uncertainty into account in the feedback controller design procedure. One approach is to combine an extended Luenberger-type observer with a state feedback controller designed by geometric control methods (Chiu and Christofides, 1999). Associated analysis indicates that the nonlinear controller possesses robust stability with respect to sufficiently fast stable unmodeled dynamics. Simulations indicated improved closed loop performance compared to linear PI controller. A related strategy using Lyapunov's direct method explicitly handles time-varying uncertain variables, provided that any unmodeled dynamics are stable and sufficiently fast (Chiu and Christofides, 2000).

An alternative approach, which couples geometric control with bilinear matrix inequalities, allows the direct optimization of robust performance (Togkalidou and Braatz, 2000; Van Antwerp et al., 1997, 1999). In contrast to most approaches to robust nonlinear control, this approach introduces no conservatism during the controller synthesis procedure. Also, no prior limitations are required regarding the speed of the unmodeled dynamics; instead, engineering intuition is incorporated into weights which bound the unmodeled dynamics, similarly as to the linear time invariant case (Morari and Zafiriou, 1989; Skogestad and Postlethwaite, 1996). Application to a crystallization process demonstrated robustness to a wide range of nonlinear and time-varying perturbations (Braatz et al., 2000b; Togkalidou and Braatz, 2000).

Interaction Between Design and Control

A method for designing a crystallizer with a given production rate has been studied for a long time, and summaries of the results are available (Bennet, 1993; Tavare,

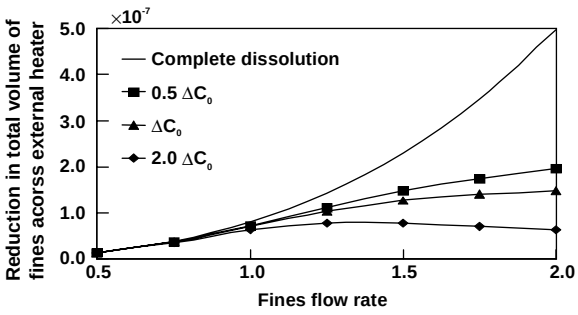


Figure 8: Influence of supersaturation on dissolution of fines.

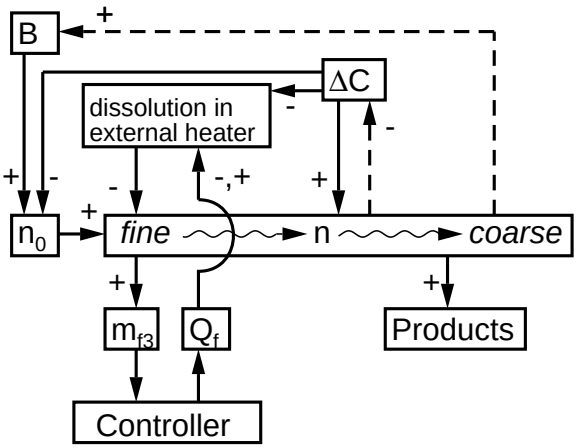


Figure 9: Interaction among state variables in the control system.

1995). This section does not deal with such design theories but focuses on the interaction between the design and control of crystallization processes.

Sustained oscillation is a commonly observed phenomenon in the operation of continuous crystallization processes. It is widely accepted by researchers in the field of crystallizer control that, in order to suppress the oscillation, the fines flow rate to the external heater should be manipulated in such a way that the nuclei population density or the concentration of fine crystals becomes constant. In most of the studies on the control of continuous crystallizers, the discussion on the controller performance and stability has been conducted on the assumption that the fines entering the external heater dissolve completely. However, in a real crystallization process, the fines do not always dissolve completely, because the residence time of the slurry in the external heater is very short. Naito et al. (1999) developed a model of an external heater by taking the dissolution rate of fines into consideration. Using the model, they have shown that the degree of dissolution depends on the fines flow rate, but this relationship is strongly affected by the supersaturation and the residence time in the external heater. Figure 8 shows a relationship between the fines flow rate and the reduction in total crystal volume across the external heater. When the dissolution rate is taken into consideration, the total volume of the dissolved fines is significantly smaller, as compared with that obtained with the assumption of complete dissolution of fines. Also, when the supersaturation increases from its nominal value (ΔC_0) by a factor of two, the amount of dissolution does not increase monotonically with increase in the fines flow rate. As a result, if the setpoint of the fines flow rate to the external heater is inappropriately determined, it is very difficult to regulate the fines concentration by the fines flow rate. By adjusting the total length and diameter of tubes in the heater, it is possible to change the residence time of the slurry in the external heater without changing the amount of heat input. Sotowa et al. (1999) demonstrated that the controller performance can be improved significantly, if the external heater is designed in such a way that the residence time of the slurry is sufficiently long.

At the design stage, the controllability assessment can be easily carried out, given indices for evaluating how easily the designed crystallizer can be controlled. Hashemi and Epstein (1982) used the condition number of the controllability matrix as such an index. A controllability study for a general class of systems described by population balance equations is discussed by Semino and Ray (1995). In their study, a crystallizer was taken as an example process, and the inlet solute concentration is taken as the manipulated variable. However, their result cannot be easily applied to a real crystallizer, because their analysis was carried out on the assumption that the crystal growth rate takes a constant value regardless of the supersaturation. Mathematical treatment of the problem becomes complicated if the crystal growth rate depends on the supersaturation.

Assume that, for a continuous crystallizer, we adopt a controller that regulates the fines concentration in the vessel by manipulating the fines flow rate to the external heater. Figure 9 shows the qualitative relationship among the state variables of the control system. The crystal population density is represented by an oblong rectangle in the middle of the figure, as it is a distributed variable over the crystal size. The population density of nuclei is denoted by n_0. Two positive feedback effects, which are responsible for the sustained oscillation of crystal size distribution, can be observed in the diagram. One is the feedback effect arising from variations in the nucleation rate, B. The other is due to variations in the supersaturation, ΔC. On the other hand, the controller adjusts the amount of fines dissolution by manipulating the fines flow rate so as to counteract the positive feedback effects. The time constants of these positive feedback loops are very long as compared with

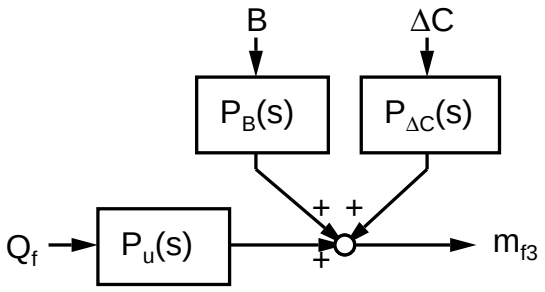


Figure 10: Block diagram of the fines subprocess.

the time constant of the controller. Thus, in order to examine the short term effects of the supersaturation, the nucleation rate, and the fines flow rate on the fines concentration, a hypothetical process is defined by removing the two broken arrows from Figure 9. In the following the hypothetical process will be referred to as the *fines subprocess* (Sotowa et al., 1999a,b).

Figure 10 shows a block diagram of a linear model that describes the behavior of the fines subprocess near the steady state. It is expected that oscillation of the crystal size distribution can be suppressed, if the fines flow rate can be manipulated in such a way that the fines concentration is kept constant regardless of the variations in supersaturation and nucleation rate. Such control action can exist only when the influence of the fines flow rate on the fines concentration is stronger than those of supersaturation and nucleation rate. Sotowa et al. (1999a,b) proposed the following indices which evaluate the relative strength of the influence of the fines flow rate on the fines concentration, as compared with those of supersaturation and nucleation rate:

$$R_{\Delta C} = \frac{P_u(0)}{P_{\Delta C}(0)} \tag{7}$$

$$R_B = \frac{P_u(0)}{P_B(0)} \tag{8}$$

Indices defined by Equations 7 and 8 are used as a measure of ease of control. When the absolute values of these indices are small, it is difficult to stabilize the crystallizer. It should be noted that these indices are only qualitative measures of the ease of control.

Once a dynamic model of the process is obtained, the steady state gains of the transfer functions in Figure 10 can be easily calculated, since the fines subprocess is a stable process. The validity of these indices as measures of the ease of control has been verified by simulation studies (Sotowa et al., 1999a,b). Hamamura et al. (2000) used these indices to study the interaction between the design and control of a continuous crystallizer. In their study, the cross sectional area of the annular settling zone and the volume of the external heater were taken as design variables. Through the study using the indices, they derived a design condition at which a great amount of large crystals can be produced and the crystallizer can be easily controlled. This idea can be extended to a general design problem of a crystallizer. For example, when the equipment cost function is given, the design condition can be obtained as a solution of a multi-objective optimization problem which derives the relationship between the ease of control and equipment cost.

Conclusions and Future Directions

The control of crystallization processes has been an active research area in recent years. Now is a good time to be working in this area, as advances in measurement technologies and computing power are removing the main factors that limited progress in the 1970s-1980s. Also, the need for improved control of crystallization processes has increased in recent years. Increased global competition has refocused efforts in optimizing industrial processes in general. Crystallization processes are often the least optimized in a chemical facility, and hence have the most to gain by optimization. Before a process can be optimized, however, its behavior must be understood. Also, the pharmaceuticals industry is continuing to grow faster than other segments of the process industries, and most pharmaceuticals must undergo multiple crystallization steps before arriving at the final product. This has increased the relative importance of crystallization processes within the process industries. Although not discussed in this paper, the development of some pharmaceuticals has been slowed by difficulties in crystallizing proteins for the determination of structure (which is then related to function). Process control engineers could make significant contributions in this area of crystallization.

Several trends in crystallization research can be identified. As discussed in the Measurements section, extracting accurate size shape information from in-process sensors is a very challenging theoretical problem. Substantial research is needed in this area, with proposed solutions likely to include the merging of digital imaging and laser backscattering information. The papers in the Measurements section and the book by van de Hulst (1981) form a starting point for such investigations.

Additives are additional solutes, usually at low concentrations, that can change the crystal shape. Many scientists and engineers have studied the effect of additives on crystal shape, and have proposed mechanisms for how the additives affect the crystal growth process (Dirksen and Ring, 1991). An exciting recent development is that simulation models are becoming available for predicting the effect of solvent type and additive concentrations on the crystal shape (Winn and Doherty, 2000). Industrial demand for such models ensures that this area will receive a significant amount of attention in future years. While there have been some successes, work is needed to validate the model predictions for more crystal sys-

tems. This will likely result in improvements in the assumptions underlying the simulation algorithms, including better models for the interactions between the solute and solvent molecules. Such simulation models will enable the selection of solvents and additives to give a desired crystal shape, and may someday reduce the amount of experimental data needed to identify models for relating the manipulated variables to the shape distribution in industrial crystallizers.

It is expected that it will become increasingly common to study crystallizers in which the fluid is not perfectly mixed (the case in practice). One approach is to model the crystallizer as an interconnection of perfectly mixed crystallizers (Bohlin and Rasmuson, 1996; Kramer et al., 1996; Sha and Palosaari, 2000), as was discussed in the Modeling section. This is already easily feasible with modern computing power. Others have started to apply full-blown computational fluid dynamics codes (Sha et al., 1999; Wei and Garside, 1997). It is expected that stochastic modeling techniques will receive greater attention in future years. Significant effort is expected over the next decade to incorporate the understanding obtained by these more complex simulation codes into estimation and control algorithms.

Another trend is that advanced control algorithms are beginning to be applied to crystallization processes. Crystallization processes have all the characteristics that makes an interesting control problem—partial differential equations, nonlinear dynamics, significant uncertainties, unmeasured state variables, significant disturbances and sensor noise, etc. Crystallization is among those processes that can benefit from advanced process control. Crystallization processes pose a rich array of control problems that are expected to keep process control engineers engaged for some time.

Acknowledgments

The first author gratefully acknowledges DuPont, Merck, and the National Center for Supercomputing Applications for support, and to Mitsuko Fujiwara, David Ma, and Timokleia Togkalidou for preparing some of the figures. The second author gratefully acknowledges the financial support by the Japan Society for the Promotion of Science under the project number JSPS-RFTF96R14301. The second author wishes to thank Dr. Ken-Ichiro Sotowa for useful discussions.

References

Ajinkya, M. B. and W. H. Ray, "On the optimal operation of crystallization processes," *Chem. Eng. Commun.*, **1**, 181–186 (1974).

Allen, T., *Particle Size Measurement*. Chapman and Hall, London, 4th edition (1990).

Baier, F. and F. Widmer, Measurement of the bubble size distribution in a gas-liquid-contactor using the PVM and image analysis, Technical report, Institute of Process Engineering, Swiss Federal Institute (ETH), Zurich, Switzerland (2000).

Barrett, P. and B. Glennon, "In-line FBRM monitoring of particle size in dilute agitated suspensions," *Part. Part. Syst. Charact.*, **16**(5), 207–211 (1999).

Barton, P. I., R. J. Allgor, W. F. Feehery, and S. Galan, "Dynamic optimization in a discontinuous world," *Ind. Eng. Chem. Res.*, **37**(3), 966–981 (1998).

Batterham, R. J., J. S. Hall, and G. Barton, Pelletizing Kinetics and Simulation of Full-scale balling circuits, In *Proc. of the 3rd Int. Symp. on Agglomeration*, page A136, Nurnberg, Germany (1981).

Becker, R., Attributes and limitations of FBRM and PVM for scale-up and troubleshooting of crystallization processes, In *11th Annual Meeting of the Association for Crystallization Technology*, Northbrook, Illinois (2001).

Beckman, J. R. and A. D. Randolph, "Crystal size distribution dynamics in a classified crystallizer: Part II. Simulated control of crystal size distribution," *AIChE J.*, **23**(4), 510–520 (1977).

Bennet, R. C., Crystallizer Selection and Design, In *Handbook of Industrial Crystallization*, chapter 5, page 103. Butterworth-Heineman (1993).

Bermingham, S. K., H. J. M. Kramer, and G. M. van Rosmalen, "Towards On-Scale Crystallizer Design Using Compartmental Models," *Comput. Chem. Eng.*, **22, Suppl.**, S335–S362 (1998).

Bharati, M. H. and J. F. MacGregor, "Multivariate image analysis for real-time process monitoring and control," *Ind. Eng. Chem. Res.*, **37**(12), 4715–4724 (1998).

Bohlin, M. and A. C. Rasmuson, "Application of Controlled Cooling and Seeding in Batch Crystallization," *Can. J. Chem. Eng.*, **70**, 120–126 (1992).

Bohlin, M. and A. C. Rasmuson, "Importance of Macromixing in Batch Cooling Crystallization," *AIChE J.*, **42**(3), 691–699 (1996).

Braatz, R. D., D. L. Ma, T. Togkalidou, M. Fujiwara, S. D. Patel, and D. K. Tafti, Modeling and control of multidimensional crystallization, In *AIChE Annual Meeting*, Los Angelos, California (2000b). Paper 253h.

Braatz, R. D., M. Fujiwara, T. Togkalidou, D. L. Ma, S. D. Patel, E. M. Tsui, and C. G. Lentz, Teaching the design of particulate processes, In *Proc. of the Educational Topical Conference, AIChE Annual Meeting*, Los Angeles, CA (2000a). Paper 60b.

Bravi, M., A. Chianese, and J. Zhang, Dynamic Process Modelling of a Cooling Crystallizer Using Locally Recurrent Neural Networks, In *Proc. of the 3rd European Control Conf.*, pages 2432–2437, Rome, Italy (1995).

Buyevich, Y. A., V. V. Mansurov, and I. A. Natalukha, "Instability and unsteady processes of the bulk continuous crystallization-I. Linear stability analysis," *Chem. Eng. Sci.*, **46**(10), 2573–2578 (1991a).

Buyevich, Y. A., V. V. Mansurov, and I. A. Natalukha, "Instability and unsteady processes of the bulk continuous crystallization-II. Non-linear periodic regimes," *Chem. Eng. Sci.*, **46**(10), 2579–2588 (1991b).

Chang, C.-T. and M. A. F. Epstein, "Identification of Batch Crystallization Control Strategies Using Characteristic Curves," *AIChE Symp. Ser.*, pages 69–75 (1982).

Chianese, A., S. D. Cave, and B. Mazzarotta, Investigation on Some Operating Factors Influencing Batch Cooling Crystallization, In Jancic, S. J. and E. J. de Jong, editors, *Industrial Crystallization 84*, pages 443–446, Amsterdam. Elsevier (1984).

Chidambaram, M. and Y. S. N. Malleswararao, "Robust control of high-order oscillations in continuous isothermal crystallizers," *Hungarian J. of Industrial Chemistry*, **18**, 91–100 (1990).

Chiu, T. and P. D. Christofides, "Nonlinear Control of Particulate Processes," *AIChE J.*, **45**(6), 1279–1297 (1999).

Chiu, T. and P. D. Christofides, "Robust control of particulate processes using uncertain population balances," *AIChE J.*, **46**(2), 1279–1297 (2000).

Chung, S. H., D. M. Ma, and R. D. Braatz, "Optimal Seeding in Batch Crystallization," *Can. J. Chem. Eng.*, **77**, 590–596 (1999).

Chung, S. H., D. L. Ma, and R. D. Braatz, "Optimal Model-based Experimental Design in Batch Crystallization," *Chemometrics and Intelligent Laboratory Systems*, **50**, 83–90 (2000).

Clark, N. N. and R. Turton, "Chord length distributions related to bubble size distributions in multiphase flows," *Int. J. of Multiphase Flow*, **14**(4), 413–424 (1988).

Dash, S. R. and S. Rohani, "Iterative Parameter Estimation for Extraction of Crystallization Kinetics of Potassium Chloride from Batch experiments," *Can. J. Chem. Eng.*, **71**, 539–548 (1993).

David, R., J. Villermaux, P. Marchal, and J.-P. Klein, "Crystallization and precipitation engineering—IV kinetic model of adipic acid crystallization," *Chem. Eng. Sci.*, **46**(4), 1129–1136 (1991).

de Wolf, S., J. Jager, H. J. M. Kramer, R. Eek, and O. H. Bosgra, Derivation of State-Space Model of Continuous Crystallizers, In *Preprints of the IFAC Symp. on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes*, pages 107–114, Maastricht, The Netherlands (1989).

Dirksen, J. A. and T. A. Ring, "Fundamentals of crystallization: kinetic effects on particle size distributions and morphology," *Chem. Eng. Sci.*, **46**(10), 2389–2427 (1991).

Doki, N., N. Kubota, A. Sato, M. Yokota, O. Hamada, and F. Masumi, "Scaleup Experiments on Seeded Batch Cooling Crystallization of Potassium Alum," *AIChE J.*, **45**(12), 2527–2540 (1999).

Dunuwila, D. D. and K. A. Berglund, "ATR FTIR spectroscopy for in situ measurement of supersaturation," *J. Crystal Growth*, **179**, 185–193 (1997).

Dunuwila, D. D., L. B. Carroll II, and K. A. Berglund, "An Investigation of the Applicability of Attenuated Total Reflection Infrared Spectroscopy For Measurement of Solubility and Supersaturation of Aqueous Citric Acid Solutions," *J. Crystal Growth*, **137**, 561–568 (1994).

Durbin, S. D. and G. Feher, "Simulation of lysozyme crystal growth by the Monte Carlo method," *J. Crystal Growth*, **110**, 41–51 (1991).

Eaton, J. W. and J. B. Rawlings, "Feedback Control of Chemical Processes Using On-Line Optimization Techniques," *Comput. Chem. Eng.*, **14**(4/5), 469–479 (1990).

Eek, R. A. and S. Dijkstra, "Design and experimental evaluation of a state estimator for a crystallization process," *Ind. Eng. Chem. Res.*, **34**, 567–574 (1995).

Eek, R. A., H. A. A. Pouw, and O. H. Bosgra, "Design and Experimental Evaluation of Stabilizing Feedback Controllers for Continuous Crystallizers," *Powder Technology*, **82**, 21–35 (1995b).

Eek, R. A., S. Dijkstra, and G. M. van Rosmalen, "Dynamic Modeling of Suspension Crystallizers, Using Experimental Data," *AIChE J.*, **41**(3), 571–584 (1995a).

Eek, R. A., J. A. Both, and P. M. J. Van den Hof, "Closed-loop Identification of a Continuous Crystallization Process," *AIChE J.*, **42**(3), 767–776 (1996).

Eek, R. A., *Control and Dynamic Modelling of Industrial Suspension Crystallizers*, PhD thesis, Delft University of Technology (1995).

Epstein, M. A. F. and L. Sowul, Phase space analysis of limit-cycle development in CMSMPR crystallizers using three-dimensional computer graphics, In Randolph, A. D., editor, *Design, Control, and Analysis of Crystallization Processes*, volume 76, pages 6–17. AIChE, New York (1980). AIChE Symposium Series.

Farrell, R. J. and Y.-C. Tsai, "Modeling, Simulation and Kinetic Parameter Estimation in Batch Crystallization processes," *AIChE J.*, **40**(4), 586–593 (1994).

Farrell, R. J. and Y.-C. Tsai, "Nonlinear controller for batch crystallization: development and experimental demonstration," *AIChE J.*, **41**(10), 2318–2321 (1995).

Fevotte, G. and J. P. Klein, "Application of On-line Calorimetry to the Advanced Control of Batch Crystallizers," *Chem. Eng. Sci.*, **49**, 1323–1336 (1994).

Fevotte, G. and J. P. Klein, "Crystallization calorimetry for the assessment of batch seeding and cooling policies," *Chem. Eng. J.*, **59**, 143–152 (1995).

Fevotte, G. and J. P. Klein, "A New Policy for the Estimation of the Course of Supersaturation in Batch Crystallization," *Can. J. Chem. Eng.*, **74**, 372–384 (1996).

Franck, R., R. David, J. Villermaux, and J. P. Klein, "Crystallization and precipitation engineering—II. Chemical reaction engineering approach to salicyclic acid precipitation: Modelling of batch kinetics and application to continuous operation," *Chem. Eng. Sci.*, **43**(1), 69–77 (1988).

Gahn, C. and A. Mersmann, "Theoretical Prediction and Experimental Determination of Attrition Rates," *Chem. Eng. Res. Dev.*, **75**, 125–131 (1997).

Gahn, C. and A. Mersmann, "Brittle Fracture in Crystallization Processes Part A. Attrition and Abrasion of Brittle Solids," *Chem. Eng. Sci.*, **54**, 1283–1292 (1999a).

Gahn, C. and A. Mersmann, "Brittle Fracture in Crystallization Processes Part B. Growth of Fragments and Scale-Up of Suspension Crystallizers," *Chem. Eng. Sci.*, **54**, 1273–1282 (1999b).

Garcia, E., S. Veesler, R. Biostelle, and C. Hoff, "Crystallization and dissolution of pharmaceutical compounds—an experimental approach," *J. Crystal Growth*, **198–199**, 1360–1364 (1999).

Ge, M., Q. G. Wang, M. S. Chiu, T. H. Lee, C. C. Hang, and K. H. Teo, "An Effective Technique for Batch Process Optimization With Application to Crystallization," *Chem. Eng. Res. Des.*, **78, Part A**, 99–106 (2000).

Gelbard, F. and J. H. Seinfeld, "Numerical solution of the dynamic equation for particulate systems," *J. Comp. Phys.*, **28**, 357–375 (1978).

Gelbard, F., Y. Tambour, and J. H. Seinfeld, "Sectional representations for simulating aerosol dynamics," *J. Colloid and Interface Sci.*, **76**(2), 541–556 (1980).

Groen, H. and K. J. Roberts, Application of ATR FTIR spectroscopy for on-line determination of solute concentration and solution supersaturation, In *Proc. of the 14th Int. Symp. on Industrial Crystallization*, Cambridge (1999).

Hamamura, T., K.-I. Sotowa, S. Hasebe, and I. Hashimoto, "The Effects of Design Variables on the Stabilizing Control of Continuous DTB Crystallizers," *Comput. Chem. Eng.*, **24**(2-7), 917–923 (2000).

Han, J. H. and S. D. Kim, "Bubble chord length distribution in three-phase fluidized beds," *Chem. Eng. Sci.*, **48**(6), 1033–1039 (1993).

Hasebe, S., T. Hamamura, K. Naito, K.-I. Sotowa, M. K. ano, I. Hashimoto, H. Betsuyaku, and H. Takeda, Optimal Operation and Control of a Continuous DTB Crystallizer, In *Preprints of the 14th IFAC World Congress*, volume N, pages 511–516, Beijing, China. IFAC (1999).

Hashemi, R. and M. A. F. Epstein, "Observability and Controllability Considerations in Crystallization Process Design," *Chem. Eng. Prog. Symp. Ser.*, **78**(215), 81–89 (1982).

Heffels, C. M. G., D. Heitzmann, E. D. Hirleman, and B. Scarlett, "The use of azimuthal intensity variations in diffraction patterns for particle shape characterization," *Part. Part. Syst. Charact.*, **11**, 194–199 (1994).

Heffels, C. M. G., P. J. T. Verheijen, D. Heitzmann, and B. Scarlett, "Correction of the effect of particle shape on the size distribution measured with a laser diffraction experiment," *Part. Part. Syst. Charact.*, **13**, 271–279 (1996).

Helt, J. E. and M. A. Larson, "Effect of temperature on the crystallization of potassium nitrate by direct measurement of supersaturation," *AIChE J.*, **23**(6), 822–830 (1977).

Hlozny, L., A. Sato, and N. Kubota, "On-Line Measurement of Supersaturation During Batch Cooling Crystallization of Ammonium Alum," *J. Chemical Engineering of Japan*, **25**(5), 604–606 (1992).

Hounslow, M. J., R. L. Ryall, and V. R. Marshall, "A discretized population balance for nucleation, growth, and aggregation," *AIChE J.*, **34**(11), 1821–1832 (1988).

Hounslow, M. J., "A discretized population balance for continous systems at steady-state," *AIChE J.*, **36**(1), 106–116 (1990).

Hulburt, H. M. and S. Katz, "Some problems in particle technology," *Chem. Eng. Sci.*, **19**, 555–574 (1964).

Ishii, T. and A. D. Randolph, "Stability of the high yield MSMPR crystallizer with size-dependent growth rate," *AIChE J.*, **26**(3), 507–510 (1980).

Jagadesh, D., N. Kubota, M. Yokota, A. Sato, and N. S. Tavare, "Large and Mono-sized Product Crystals From Natural Cooling Mode Batch Crystallizer," *J. Chemical Engineering of Japan*, **29**(5), 865–873 (1996).

Jager, J., S. de Wolf, W. Klapwijk, and E. J. de Jong, A new design for on-line product-slurry measurements, In Nyvlt, J. and S. Zacek, editors, *Industrial Crystallization 87*, Chemical Engineering Progress Symposium Series, pages 415–418, Bechyne, Czechoslovakia. Elsevier Science (1987).

Jager, J., H. J. M. Kramer, B. Scarlett, E. J. de Jong, and S. de Wolf, "Effect of Scale and Operation on CSD dynamics in Evaporative Crystallizers," *AIChE J.*, **37**(2), 182–192 (1991).

Johnson, B. K., C. Szeto, O. Davidson, and A. Andrews, Optimization of Pharmaceutical Batch Crystallization for Filtration and Scale-Up, In *AIChE J.* (1997). Paper 16a.

Jones, A. G. and J. W. Mullin, "Programmed Cooling Crystallization of Potassium Sulphate Solutions," *Chem. Eng. Sci.*, **29**, 105–118 (1974).

Jones, A. G., "Optimal Operation of a Batch Cooling Crystallizer," *Chem. Eng. Sci.*, **29**, 1075–1087 (1974).

Ke, S. C., L. J. De Lucas, and J. G. Harrison, "Computer simulation of protein crystal growth using aggregates as the growth unit," *J. Physics D—Applied Physics*, **31**, 1064–1070 (1998).

Kim, S. and A. S. Myerson, "Metastable solution thermodynamic properties and crystal growth kinetics," *Ind. Eng. Chem. Res.*, **35**, 1078–1084 (1996).

Koenig, H. A., *Modern Computational Methods*. Taylor & Francis, Philadelphia, PA (1998).

Kramer, H. J. M., J. W. Dijkstra, A. M. Neumann, R. O. Meadhra, and G. M. van Rosmalen, "Modelling of Industrial Crystallizers, A Compartmental Approach Using a Dynamic Flow-Sheeting Tool," *J. Crystal Growth*, **166**, 1084–1088 (1996).

Kramer, H. J. M., S. K. Bermingham, and G. M. van Rosmalen, "Design of Industrial Crystallizers for a Given Product Quality," *J. Crystal Growth*, **198/199**, 729–737 (1999).

Kuehberger, M. and A. Mersmann, How to Meet Product Requirements During Cooling Crystallization by Control of Supersaturation, In *AIChE Annual Meeting*, Los Angeles (1997a).

Kuehberger, M. and A. Mersmann, "Improved Product Quality at a Cooling Crystallization Process by Measurement and Control of supersaturation," *Chem. Eng. Res. Dev.*, **75**(Part A), 213–218 (1997b).

Kumar, S. and D. Ramkrishna, "On the solution of population balance equations by discretization—I. a fixed pivot technique," *Chem. Eng. Sci.*, **51**(8), 1311–1332 (1996a).

Kumar, S. and D. Ramkrishna, "On the solution of population balance equations by discretization—II. a moving pivot technique," *Chem. Eng. Sci.*, **51**(8), 1333–1342 (1996b).

Kumar, S. and D. Ramkrishna, "On the solution of population balance equations by discretization—III. nucleation, growth and aggregation of particles," *Chem. Eng. Sci.*, **52**(24), 4659–4679 (1997).

Lakatos, B. G. and T. Blickle, "Nonlinear dynamics of isothermal CMSMPR crystallizers: A simulation study," *Comput. Chem. Eng.*, **19**(Suppl.), S501–S506 (1995).

Lakatos, B. G., "Uniqueness and multiplicity in isothermal CMSMPR crystallizers," *AIChE J.*, **42**(1), 285–289 (1996).

Lang, Y., A. M. Cervantes, and L. T. Biegler, "Dynamic Optimization of a Batch Cooling Crystallization Process," *Ind. Eng. Chem. Res.*, **38**, 1469–1477 (1999).

Lasentec, Lasentec (1997). brochure.

Lei, S.-J., R. Shinnar, and S. Katz, The regulation of a continuous crystallizer with fines trap, In *Crystallization from Solution: Factors Influencing Size Distributions*, volume 67 of *Chemical Engineering Progress Symposium Series*, pages 129–143 (1971).

Lewiner, F., G. Fevotte, J. P. Klein, and G. Pfefer, Application of in situ ATR FTIR to the on line monitoring of batch crystallisation with aggolmeration, In *Proc. of the 14th Int. Symp. on Industrial Crystallization*, Cambridge (1999).

Lewiner, F., J. P. Klein, F. Puel, F. Conesa, and G. Fevotte, "On-line ATR FTIR measurement of supersaturation during solution crystallization processes. Calibration and applications on three solute/solvent systems," *Chem. Eng. Sci.* (2001). In press.

Lim, K. C., M. A. Hashim, and B. Sen Gupta, "Monte Carlo simulation of transient crystal size distribution in a continuous crystallizer using the ASL model," *Cryst. Res. Tech.*, **33**, 249–255 (1998).

Lim, K. C., M. A. Hashim, and B. Sen Gupta, "Effect of volume shape factor on the crystal size distribution of fragments due to attrition," *Cryst. Res. Tech.*, **34**, 491–502 (1999).

Lim, K. C., M. A. Hashim, and B. Sen Gupta, "Transient crystal size distribution in DTB and FC crystallizers," *Cryst. Res. Tech.*, **34**, 1055–1064 (1999).

Litster, J. D., D. J. Smit, and M. J. Hounslow, "Adjustable discretized population balance for growth and aggregation," *AIChE J.*, **41**(3), 591–603 (1995).

Liu, W., N. N. Clark, and A. I. Karamavruc, "Relationship between bubble size distributions and chord-length distribution in heterogeneously bubbling systems," *Chem. Eng. Sci.*, **53**(6), 1267–1276 (1998).

Loeffelmann, M. and A. Mersmann, Innovative Supersaturation Sensor for Industrial Crystallization, In *14th Int. Symp. on Industrial Crystallization*, Cambridge, UK (1999).

Ma, D. L. and R. D. Braatz, Robust batch control of crystallization processes, In *Proc. of the American Control Conf.*, pages 1737–1741, Piscataway, New Jersey. IEEE Press (2000).

Ma, D. L. and R. D. Braatz, "Worst-case Analysis of Finite-time Control Policies," *IEEE Trans. Cont. Sys. Tech.*, **9**(3) (2001). In press.

Ma, D. L., T. Togkalidou, and R. D. Braatz, Multidimensional crystal growth from solution, In *AIChE Annual Meeting* (1999). Paper 03A02.

Ma, D. L., S. H. Chung, and R. D. Braatz, "Worst-case Performance Analysis of Optimal Batch Control Trajectories," *AIChE J.*, **45**(7), 1469–1476 (1999a).

Ma, D. L., S. H. Chung, and R. D. Braatz, Worst-case Performance Analysis of Optimal Batch Control Trajectories, In *Proc. of the European Control Conf.*, Germany. IFAC (1999b). Paper F1011-2.

Ma, D. L., D. K. Tafti, and R. D. Braatz, Simulation of multidimensional crystallization processes, In preparation (2001).

Maisels, A., F. E. Kruis, and H. Fissan, "Direct Monte Carlo simulations of coagulation and aggregation," *J. Aerosol Science*, **30**, S417–S418 (1999).

Marchal, P., R. David, J. P. Klein, and J. Villermaux, "Crystallization and precipitation engineering. I: An efficient method for solving population balance in crystallization with agglomeration," *Chem. Eng. Sci.*, **43**(1), 59–67 (1988).

Matthews, H. B. and J. B. Rawlings, "Batch Crystallization of a Photochemical: Modeling, Control and Filtration," *AIChE J.*, **44**(5), 1119–1127 (1998).

Matthews, H. B., S. M. Miller, and J. B. Rawlings, "Model Identification for Crystallization: Theory and Experimental Verification," *Powder Technology*, **88**, 227–235 (1996).

Matthews III, H. B., *Model Identification and Control of Batch Crystallization for an Industrial Chemical System*, PhD thesis, University of Wisconsin, Madison (1997).

Mersmann, A., *Crystallization Technology Handbook*. Marcel Dekker, New York (1995).

Miller, S. M. and J. B. Rawlings, "Model Identification and Control Strategy for Batch Cooling Crystallizers," *AIChE J.*, **40**(8), 1312–1327 (1994).

Miller, S. M., *Modelling and Quality Control Strategies for Batch Cooling Crystallizers*, PhD thesis, Univ. of Texas at Austin (1993).

Monnier, O., J.-P. Klein, C. Hoff, and B. Ratsimba, "Particle size determination by laser reflection: methodology and problems," *Particle & Particle Systems Characterization*, **13**, 10–17 (1996).

Monnier, O., G. Fevotte, C. Hoff, and J. P. Klein, "Model identification of batch cooling crystallizations through calorimetry and image analysis," *Chem. Eng. Sci.*, **52**(7), 1125–1139 (1997).

Moore, W. P., "Optimize Batch Crystallization," *Chem. Eng. Prog.*, **90**(9), 73–79 (1994).

Morari, M. and E. Zafiriou, *Robust Process Control*. Prentice-Hall, Englewood Cliffs, New Jersey (1989).

Muhr, H., R. David, J. Villermaux, and P. H. Jezequel, "Crystallization and precipitation engineering—VI. Solving population balance in the case of the precipitation of silver bromide crystals with high primary nucleation rates by using the first order upwind differentiation," *Chem. Eng. Sci.*, **51**(2), 309–319 (1996).

Mullin, J. W. and C. J. Leci, "Desupersaturation of seeded citric acid solutions in a stirred vessel," *AIChE Symp. Ser.*, **68**, 8–20 (1972).

Mullin, J. W. and J. Nyvlt, "Programmed Cooling of Batch Crystallizers," *Chem. Eng. Sci.*, **26**, 369–377 (1971).

Mullin, J. W. and O. Sohnel, "Expressions of supersaturation in crystallization studies," *Chem. Eng. Sci.*, **32**, 683–686 (1977).

Naito, K., K.-I. Sotowa, M. Kano, S. Hasebe, and I. Hashimoto, "Stabilizing control of Continuous DTB Crystallizer—Influence of Undissolved Fine Crystals in External Heater," *Kagaku Kogaku Ronbunshu*, **25**(1), 51–58 (1999). in Japanese.

Natalukha, I. A., "Unstable regimes of continuous crystallization in a cascade of well-mixed vessels," *Chem. Eng. Sci.*, **51**(8), 1181–1185 (1996).

Neumann, A. M., S. K. Bermingham, and H. J. M. Kramer, Modeling Industrial Crystallizers of Different Scale and Type, In *14th Int. Symp. on Industrial Crystallization*, Cambridge, UK (1999).

Noda, M., T. Chida, S. Hasebe, and I. Hashimoto, "On-Line Optimization System of Pilot Scale Multi-Effect Batch Distillation System," *Comput. Chem. Eng.*, **24**(2-7), 1577–1583 (2000).

Nyvlt, J., M. Karel, and S. Pisarik, "Measurement of supersaturation," *Cryst. Res. Tech.*, **29**, 409–415 (1994).

Pacek, A. W., I. P. T. Moore, A. W. Nienow, and R. V. Calabrese, "Video technique for measuring dynamics of liquid-liquid dispersion during phase inversion," *AIChE J.*, **40**, 1940–1949 (1994).

Puel, F., P. Marchal, and J. Klein, "Habit transient analysis in industrial crystallization using two dimensional crystal size technique," *Chem. Eng. Res. Dev.*, **75**, 193–205 (1997).

Qiu, Y. and A. C. Rasmuson, "Estimation of Crystallization Kinetics from Batch Cooling Experiments," *AIChE J.*, **40**(5), 799–812 (1994).

Ramkrishna, D., "The status of population balances," *Rev. Chem. Eng.*, **3**(1), 49–95 (1985).

Randolph, A. D. and M. A. Larson, *Theory of Particulate Processes*. Academic Press, 2nd edition (1988).

Randolph, A. D., G. L. Beer, and J. P. Keener, "Stability of the class II classified product crystallizer with fines removal," *AIChE J.*, **19**(6), 1140–1149 (1973).

Randolph, A. D., J. R. Beckman, and Z. I. Kraljevich, "Crystal size distribution dynamics in a classified crystallizer: Part I. Experimental and theoretical study of cycling in a potassium chloride crystallizer," *AIChE J.*, **23**(4), 500–510 (1977).

Randolph, A. D., E. T. White, and C.-C. D. Low, "On-line measurement of fine-crystal response to crystallizer disturbances," *Ind. Eng. Chem. Proc. Des. Dev.*, **20**, 496–503 (1981).

Randolph, A. D., L. Chen, and A. Tavana, "Feedback control of CSD in a KCl crystallizer with a fines dissolver," *AIChE J.*, **33**(4), 583–591 (1987).

Randolph, A. D., CSD dynamics, stability, and control (a review paper), In Randolph, A. D., editor, *Design, Control, and Analysis of Crystallization Processes*, volume 76, pages 1–5. AIChE, New York (1980). AIChE Symposium Series.

Rawlings, J. B. and D. B. Patience, On-line monitoring and control of crystal size and shape, In *Annual Meeting of the International Fine Particle Research Institute*, Somerset, New Jersey (1999).

Rawlings, J. B., W. R. Witkowski, and J. W. Eaton, "Modelling and control of crystallizers," *Powder Technology*, **69**, 3–9 (1992).

Rawlings, J. B., S. M. Miller, and W. R. Witkowski, "Model Identification and Control of Solution Crystallization Processes: A Review," *Ind. Eng. Chem. Res.*, **32**, 1275–1296 (1993).

Redman, T. P. and S. Rohani, "On-line determination of supersaturation of a KCl-NaCl aqueous solution based on density measurement," *Can. J. Chem. Eng.*, **72**, 64–71 (1994).

Redman, T., S. Rohani, and G. Strathdee, "Control of the crystal mean size in a pilot plant potash crystallizer," *Chem. Eng. Res. Dev.*, **75**, 183–192 (1997).

Riebel, U., V. Kofler, and F. Loffler, Control of supersaturation in instationary suspension crystallization, In Mersmann, A., editor, *Industrial Crystallization 90*, pages 595–599. Plenum Press, New York (1990).

Rohani, S. and J. R. Bourne, "Self-tuning control of crystal size distribution in a cooling batch crystallizer," *Chem. Eng. Sci.*, **45**(12), 3457–3466 (1990a).

Rohani, S. and J. R. Bourne, "A Simplified Approach to the Operation of a Batch Crystallizer," *Can. J. Chem. Eng.*, **68**, 799–805 (1990b).

Rohani, S., M. Haeri, and H. C. Wood, "Modeling and Control of a Continuous Crystallization Process Part 1. Linear and Non-Linear Modeling," *Comput. Chem. Eng.*, **23**, 263–277 (1999a).

Rohani, S., M. Haeri, and H. C. Wood, "Modeling and control of a continuous crystallization process. Part 2. Model predictive control," *Comput. Chem. Eng.*, **23**, 279–286 (1999b).

Rovang, R. D. and A. D. Randolph, On-line particle size analysis in the fines loop of a KCl crystallizer, In Randolph, A. D., editor, *Design, Control, and Analysis of Crystallization Processes*, volume 76, pages 18–26. AIChE, New York (1980). AIChE Symposium Series.

Semino, D. and W. H. Ray, "Control of Systems Described by Population Balance Equations—I. Controllability Analysis," *Chem. Eng. Sci.*, **50**(11), 1805–1824 (1995).

Sen Gupta, B. and T. K. Dutta, "Growth rate dispersion in a continuous sucrose crystallizer and its effect on CSD," *Chem. Eng. Technol.*, **13**, 196–202 (1990a).

Sen Gupta, B. and T. K. Dutta, "Simulation of crystal size distribution in a continuous sodium chloride crystalliser under diffusion controlled conditions," *Chemie Ingenieur Technik*, **62**, 66–67 (1990b).

Sha, Z. and S. Palosaari, Size dependent classification function in imperfectly mixed suspension continuous crystallizer, In Sen

Gupta, B. and S. Ibrahim, editors, *Mixing and Crystallization*, pages 133–149. Kluwer Academic Publishers, Netherlands (2000).

Sha, Z., M. Louhi-Kultanen, P. Oinas, and S. Palosaari, CFD simulation of size dependent classification in an imperfectly mixed suspension crystallizer, In *Proc. of the 14th Int. Symp. on Industrial Crystallization*, Cambridge, UK (1999).

Shah, B. H., D. Ramkrishna, and J. D. Borwanker, "Simulation of particulate systems using the concept of the interval of quiescence," *AIChE J.*, **23**, 897–904 (1977).

Shangfeng, Y., S. Genbo, L. Zhengdong, and J. Rihong, "Rapid growth of KH_2PO_4 crystals in aqueous solutions with additives," *J. Crystal Growth*, **197**, 383–387 (1999).

Sherwin, M. B., R. Shinnar, and S. Katz, "Dynamic behavior of the well-mixed isothermal crystallizer," *AIChE J.*, **13**(6), 1141–1154 (1967).

Sikdar, S. K. and A. D. Randolph, "Secondary nucleation of two fast growth systems in a mixed suspension crystallizer: Magnesium sulfate and citric acid water systems," *AIChE J.*, **22**(1), 110–117 (1976).

Simmons, M. J. H., P. A. Langston, and A. S. Burbidge, "Particle and droplet size analysis from chord distributions," *Powder Technology*, **102**(1), 75–83 (1999).

Singh, P. N. and D. Ramkrishna, "Solution of population balance equations by MWR," *Comput. Chem. Eng.*, **1**, 23–31 (1977).

Skogestad, S. and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons, New York (1996).

Soderstrom, T. and P. Stoica, *System Identification*. Prentice Hall (1989).

Song, M. and X. J. Qiu, "Alternative to the concept of the interval of quiescence (IQ) in the Monte Carlo simulation of population balances," *Chem. Eng. Sci.*, **54**, 5711–5715 (1999).

Sotowa, K.-I., T. Hamamura, H. Taniguchi, S. Hasebe, and I. Hashimoto, Applicability of a Reduced Model to Study the Performance of Control System for a Continuous Crystallizer, In *AIChE Annual Meeting*, Dallas, TX (1999b).

Sotowa, K.-I., T. Hamamura, S. Hasebe, and I. Hashimoto, Optimal Operating Condition for a Continuous DTB Crystallizers—Qualitative Evaluation of Dynamic Behavior of Continuous Crystallizers, In *The 49th Canadian Chemical Engineering Conf.*, Saskatoon, Saskatchewan, Canada (1999a).

Sotowa, K.-I., K. Naito, M. Kano, S. Hasebe, and I. Hashimoto, A Study on the Effects of Fines Dissolution on the Stabilizing Control of Continuous DTB Crystallizers, In *Proc. of the 14th Int. Symp. on Industrial Crystallization*, Cambridge, UK (1999).

Sotowa, K., K. Naito, M. Kano, S. Hasebe, and I. Hashimoto, "Application of the method of characteristics to crystallizer simulation and comparison with finite difference for controller performance evaluation," *J. Proc. Cont.*, **10**, 203–208 (2000).

Sotowa, K.-I., H. Nagata, T. Hamamura, S. Hasebe, and I. Hashimoto, "Calculation of classification functions from CFD data," *AIChE J.* (2000). to be submitted.

Strom, C. S. and P. Bennema, "Combinatorial compatibility as habit-controlling factor in lysozyme crystallization II. Morphological evidence for tetrameric growth units," *J. Crystal Growth*, **173**, 159–166 (1997).

Tadayyon, A. and S. Rohani, "Determination of particle size distribution by Par-TecR 100: Modeling and experimental results," *Particle & Particle Systems Characterization*, **15**(3), 127–135 (1998).

Tahti, T., M. Louhi-Kultanen, and S. Palosaari, On-line measurement of crystal size distribution during batch crystallization, In *Industrial Crystallization 1999* (1999).

Tavare, N. S. and J. Garside, "Multiplicity in continuous MSMPR crystallizers, Part 1: Concentration multiplicity in an isothermal crystallizer," *AIChE J.*, **31**(7), 1121–1127 (1985).

Tavare, N. S., "Multiplicity in continuous crystallizers: Adiabatic reactive precipitation," *Chem. Eng. Commun.*, **80**, 135–152 (1989).

Tavare, N. S., *Industrial Crystallization*. Plenum Press (1995).

Togkalidou, T. and R. D. Braatz, A bilinear matrix inequality approach to the robust nonlinear control of chemical processes, In *Proc. of the American Control Conf.*, pages 2548–2552, Piscataway, New Jersey. IEEE Press (2000).

Togkalidou, T., M. Fujiwara, S. Patel, and R. D. Braatz, A robust chemometrics approach to inferential estimation of supersaturation, In *Proc. of the American Control Conf.*, pages 1732–1736, Piscataway, New Jersey. IEEE Press (2000).

Togkalidou, T., R. D. Braatz, B. Johnson, O. Davidson, and A. Andrews, "Experimental design and inferential modeling in pharmaceutical crystallization," *AIChE J.*, **47**, 160–168 (2001).

Tsuruoka, S. and A. D. Randolph, "State Space Representation of the Dynamic Crystallizer Population Balance: Application to CSD Controller Design," *AIChE Symp. Ser.*, **83**(253), 104–109 (1987).

Tullken, H. J. A. F., Application of the Grey-Box Approach to Parameter Estimation in Physicochemical Models, In *Proc. of the 30th Conf. on Decision and Control*, pages 1177–1183, Brighton, England (1991).

Van Antwerp, J. G., R. D. Braatz, and N. V. Sahinidis, "Globally optimal robust control for systems with nonlinear time-varying perturbations," *Comput. Chem. Eng.*, **21**, S125–S130 (1997).

Van Antwerp, J. G., R. D. Braatz, and N. V. Sahinidis, "Globally optimal robust process control," *J. Proc. Cont.*, **9**, 375–383 (1999).

van de Hulst, H. C., *Light Scattering by Small Particles*. Dover Publications, New York (1981).

van den Hof, P. M. J. and R. J. P. Schrama, "An Indirect Method for Transfer Function Estimation from Closed Loop Data," *Automatica*, **29**(6), 1523–1527 (1993).

Van Peborgh Gooch, J. R. and M. J. Hounslow, "Monte Carlo simulation of size-enlargement mechanisms in crystallization," *AIChE J.*, **42**, 1864–1874 (1996).

Van Peborgh Gooch, J. R., M. J. Hounslow, and J. Mydlarz, "Disciminating between size-enlargement mechanisms," *Trans. Inst. Chem. Eng.*, **74**, 803–811 (1996).

Wei, H. and J. Garside, "Application of CFD modelling to precipitation systems," *Chem. Eng. Res. Dev.*, **75**, 219–227 (1997).

Wey, J. S., Batch Crystallization, In *Handbook of Industrial Crystallization*, chapter 10, page 209. Butterworth-Heineman (1993).

Winn, D. and M. F. Doherty, "Modeling crystal shapes of organic materials grown from solution," *AIChE J.*, **46**(7), 1348–1367 (2000).

Witkowski, W. R. and J. B. Rawlings, Modelling and control of crystallizers, In *Proc. of the American Control Conf.*, pages 1400–1405, Piscataway, New Jersey. IEEE Press (1987).

Witkowski, W. R., S. M. Miller, and J. B. Rawlings, "Light-scattering measurements to estimate kinetic parameters for crystallization," *ACS Symposium Series*, **438**, 102–114 (1990).

Witkowski, W. R., *Model Identification and Parameter Estimation of Crystallization Processes*, PhD thesis, Univ. of Texas at Austin (1990).

Worlitschek, J. and M. Mazzotti, On-line control of batch cooling crystallization, Technical report, Institute of Process Engineering, Swiss Federal Institute (ETH), Zurich, Switzerland (2000). `http://www.ivuk.ethz.ch/Staff/worlitschek/page.htm`. also presented as Paper 11 of the *2000 FBRM Users Conference*.

Linking Control Strategy Design and Model Predictive Control

James J. Downs
Eastman Chemical Company
Kingsport, Tennessee

Abstract

The purpose of this paper is to describe the importance of the underlying relationship between control strategy design and model predictive control. Successes and problems encountered when implementing model predictive control (MPC) on chemical processes have revealed that understanding this relationship provides insight into the nature of the process control problem. Model predictive control (MPC) has been used as an effective tool to gain the process control benefits that come from its ability to handle constraints, process interactions, and multiple time frames. The use of the MPC algorithm on a variety of chemical processes has led to insight on how to effectively use MPC along with traditional control strategy notions to improve process control. The development of control strategies using MPC has resulted in the typically reported benefits of increased throughput and reduced process variability. Several issues remain to be addressed. These include controller tuning, complex performance criteria, depth of integration of MPC with the regulatory control layer, redundant process information, and controller robustness to measurement loss or deterioration.

Keywords

Model predictive control, Control strategy design, Process control vendor, Process control education, Cascade control

Introduction

Process control strategy design has been the cornerstone of successful application of process control technology for many years. As new process control methods and algorithms have been conceived and developed, their successful application in the process industries has relied upon the underlying insight into the nature of the process. Certainly, if one chooses the right problem the success of a particular technology is enhanced. One technology can be shown to be superior to another simply by judicious choice of application. The success of model predictive control lies in its ability to cast process control strategy choices into a manageable framework. The capability to dynamically decouple process control loops, to handle process constraints, and to minimize deviation from set point are important but are more tactical in nature compared to the control strategy changes that take place.

As a model predictive controller is exercised throughout its allowable range, any number of control strategies may be manifested. The ability to understand the ramifications and consequences of each strategy or group of strategies is key to the successful implementation of model predictive control. In the past, we had a fixed control strategy. The process control designer was charged with the design of a strategy that would perform the best for as wide a range of circumstances as possible. Often if it were known beforehand that a strategy could not handle a particular set of conditions, those conditions were avoided during operation or protected against on a case-by-case basis. Given that the control strategy was fixed, it was studied by subjecting it to the variety of disturbances and operating scenarios that were plausible. As unexpected operating conditions and disturbances were encountered, the plant operators acted as "test pilots" having to manage the new operating regime as best they could.

Model predictive controllers provide the capability to change the fundamental control strategy while the process is operating—in its simplest sense acting as a control system override. The opportunity to exercise and explore complex MPC control designs is limited by the factorial number of possible strategies that can be in effect at a given time. How each of these strategies will respond to an array of disturbances and operating conditions must be answered or addressed to avoid having "test pilots" testing systems that look like black boxes.

The evolution of process control technology has expanded the role for the process control engineer. The notion of designing multiple, complex control strategies that can change during the normal course of operation is becoming more prevalent. Certainly high and low select overrides have been around for many years. However, the extent to which even mildly complex MPC applications result in unexpected control strategies is a new realm. The purpose of this paper is to discuss how this new focus is unfolding.

This paper starts with a perspective on the current process control work environment in the chemical industry. This perspective highlights what a process control engineer is likely to face in today's world. With this perspective as a backdrop, a linkage is developed between the familiar territory of control strategy design and the newer, possibly unfamiliar tool, model predictive control. This linkage demonstrates the need and value of accumulated process knowledge and traditional process control notions when faced with reaping the widely acknowledged benefits of model predictive control. The paper concludes with examples that highlight the variety of problems benefiting from our application of model predictive control and that illustrate some of the implementation issues that we have encountered.

The Process Control Landscape in the Chemical Industry

New Plants

Normally for new plants the process control design is determined as a fixed strategy that provides regulatory control for the array of expected disturbances and operation regimes. The focus during this activity is maintaining the plant operation at a nominal operating condition from which operations can move to achieve product properties. New plants often contain new technology that involves uncertainty of operation and of performance. All that is needed of the control system is to maintain stability and to be understandable by personnel with a wide range of experience and education levels. Forays into the use of advanced control techniques on unfamiliar unit operations or processes employing new process technology have usually demonstrated that starting up with a simple, understandable control system is best. Once the operating characteristics are more known then the operation can be optimized employing more advanced control techniques.

The design of control strategies for new facilities warrants the need for control strategy design and analysis. The formation of rugged, well thought out regulatory control strategy designs that can withstand the variety of disturbances and operating abnormalities encountered during the first year of plant operation is a requirement for future process control enhancements. Undoubtedly, operating a process closer to optimum conditions and determining where that is requires some semblance of stable operation. There has been much written and presented to help integrate the process and the control strategy design. Recent examples include Barolo and Papini (2000), Groenendijk et al. (2000), and Tseng et al. (1999). These efforts are directed at the development of process designs and regulatory control strategies that achieve good regulatory control. This work is to be recognized for providing guidance where a few years ago there was none. In a similar vein, Luyben (1998c; 1998a; 1998b; 2000) has published a series of articles to guide control strategy development based on process situations. While the incorporation of advanced control technology into the process design phase is a notable goal, current practice is to get new processes up and running first and follow later with control system enhancements.

Existing Plants

Once a plant has been in operation long enough to find and fix problems that preclude stable operation then the initiation of process improvement activities is a natural consequence. It is during this time that enough is learned about the operation that advanced control techniques can be successfully applied. During this time the linkage among what is needed, what is feasible, and the appropriate technology to apply is most important. Among the many improvement opportunities and the flood of available technologies, it is necessary for the process control engineer to discriminate between process equipment problems and control strategy problems. Often process improvements that come from process control changes are of the control strategy variety rather than control algorithm changes. From our viewpoint, MPC is regarded as a "control strategy change agent" instead of an algorithm for improved high performance control. Indeed, it is capable of both. The effort needed to develop and maintain models accurate enough for high performance control, however, often outweighs the marginal benefits. The need to change control strategies for differing modes of operation has been more persistent.

The process knowledge available for existing plants provides insight into the true objective that needs to be achieved by the control system. While cursory overviews of plant operation may yield process control objectives that appear reasonable, often a deeper process insight is needed to arrive at the desired process objectives. This deeper process insight comes from understanding process chemistry, unit operation objectives, business objectives, and process flow structures.

Control Objectives

The definition of process control objectives often involves an evolutionary path. Often an initial statement of what the control system should do is oriented around what the current control system cannot do. "If only we could control the temperature, we would be happy with our operation" leads to "The temperature control we have is great, but what we really need is to control the composition of ...". This in turn may lead to other objectives that may change once the successive performance plateaus are reached. As process control systems hold key variables within narrower and narrower limits, the costs in terms of increased variation in other non-key variables becomes apparent and control objectives change. Furthermore, tighter control allows process engineers to see process improvement opportunities that are otherwise hidden.

When product requirements change, control objectives may need to be altered. These changes may involve a simple change such as altering controller weighting parameters or require an entire control system structure change. Labor, retraining, and opportunity costs to maintain and improve advanced control systems when process objectives change are compared with the economic benefits. This situation results in large, single product, unchanging plants to be obvious candidates for advanced control applications. This type of application is common and has been reported often in the literature. On the other hand, for plants that are smaller, multiple product, or undergo occasional change there is a need to be able to reap the benefits available from advanced control without prohibitive costs. An MPC structure that can represent various control strategies can be a very

effective tool.

In this more fluid application environment, we are not driving for a lower IAE, ITAE, etc. as much as we are addressing opportunities to conveniently drive processes to the optimum steady state when constraints are encountered. In many of our applications it is the steady state targeting feature of model predictive control that is the important piece. Control objectives that we encounter are much more focused on where the process will line out under various conditions rather than on how a process will dynamically respond.

The Academic/Vendor/Industry Relationship

The implementation of process control technology, and in particular model predictive control, requires process control skills that may not be taught in the normal undergraduate curriculum. Certainly skills arising from formal training in this area may not be recent or deep enough to warrant a personal embarkation into an advanced control project. The role of the corporate process control group in the chemical industry is pivotal in channeling the technology to the appropriate applications and ensuring their success. Without such a group, the linkage is weak between those with the problems and those possessing the solutions. A central group can provide the standardization and stewardship needed for company wide application.

The process control vendor has historically provided the control toolkit needed to apply process control technology along with training and personnel to use their products. The current climate of specialization of service providers as modeled in the communications industries is becoming popular in the chemical industry as well. Vendors are moving from providing a product to providing a service. Academic institutions on the other hand provide trained personnel and technology ideas but no industrially hardened products. The process control toolkits on the market today have a variety of technologies that hopefully weave together to make their use easy. The relationship between vendors and academicians is becoming stronger. This is driven by a viewpoint that few companies have the wherewithal to incorporate new theoretical advancements into their day-to-day business. Process control vendors are becoming a more important avenue through which theoretical advancements make their way to industrial practice.

Companies are in transition to meet relentless market pressures on shorter and shorter time horizons and the lure of marketing suggestions that promise short payback times while requiring little long term corporate investment are strong. The choice of appropriate control technology requires an unbiased viewpoint. Often times the solution of a control problem can be accomplished via many technology avenues. If a vendor is selling hammers then the vendor sees most problems looking like nails. The implementation of advanced control is only warranted where simpler control techniques are inadequate. Corporate process control groups should have the knowledge to make choices among competing technologies based on life cycle costs and other intangible factors. Vendors while economically driven should nevertheless provide a similar unbiased approach to problem solution.

The identification of appropriate candidates for advanced control usually requires proprietary knowledge of process economics, process weaknesses, process chemistry, and even corporate politics. Relying upon operation personnel to identify candidates in the midst of regulatory, labor, and production demands is difficult for reasonably steady operations and nearly impossible for constantly changing production environments. Indeed, a process control specialist with knowledge of corporate objectives and a process viewpoint of the larger picture has a much better chance of identifying the best projects. Once projects are selected, the process knowledge needed to reach solutions is normally located within operations. Often that knowledge is shared with the in-house process control specialist because of past experiences, built up trust relationships, or personal relationships. The criticality of process knowledge cannot be overstated (Downs and Doss, 1991). How much process knowledge can be shared with non-company personnel, secrecy agreements notwithstanding, is always a subject of debate. Process discoveries during implementation, accumulated process operational savvy, and application tricks are all subject to loss after the project is complete. In addition, process control revelations arising from implementation become leveragable knowledge for the control vendor. Undeniably, early customers are in the role of guinea pig until adequate enhancements harden advanced control products.

One of the most important factors in the success of process control projects is the long-term maintenance of the finished product. Valves change, transmitter ranges change, processes change. There is an inevitable march toward a process that sooner or later does not match the process control system. For large volume plants with only a few products the process control system may remain valid for several years. However, as the variety of projects increases, the applications require more support. Local personnel can change simple items, however, software upgrades, process changes, and even retuning will probably require specialized support. This support can be provided by service contracts or by in-house specialist.

In this environment, Eastman has thus far benefited from having a corporate group to manage this activity. The Eastman process control group has maintained the strength to objectively evaluate the cost/benefit tradeoff in the spectrum ranging from an entirely in-house process control program to one that is entirely contracted to a service provider. Our current approach is to purchase

those products and technology that provide value and are generally one-time in cost. Those products that entail on-going costs for each use or application have been used sparingly due to their continual drain on process control profitability. Each company trades off between the expense of maintaining in-house talent and purchasing that talent through vendors. However, recent informal survey data indicates that the need for in-house process control expertise remains strong (Downs, 2000).

We believe that there is a strong case for academia to continue to provide people knowledgeable in process control not only for service vendors but also for corporate in-house needs. The propagation of process control technology from the academic realm to the industrial shop floor requires both vendor and user comprehension. Movement toward a strictly "vendor/supplier sells control system hardware/solutions/knowledge to corporate consumer" may appeal to the compartmentally minded. However, once the corporate user becomes ignorant in the technology, the synergy between process design, control, and operation is lost. The lack of process control talent in any of academic research, process control vendor, or corporate consumer is a weakness and handicap for all concerned.

Instantiation of Supervisory Control Systems

Underlying the application of advanced technology is the computational platform and distributed control system in place. Decisions of complexity and distributive reliability are important factors in the definition of scope for advanced control applications. As advanced control algorithms become available on regulatory level distributed control systems, the process control engineer is faced not only with a technology decision but also with a choice of vendor instantiation of advanced control technology. While algorithm fundamentals are published and well known, it is often the subtle modifications of published technology that make the technology actually work in practice. Each vendor claiming that their implementation is superior to their competitors can create a confusing climate that clouds the more important issues of control system objectives.

While it may seem that the development of process control technology is a mature area, the application of the technology available is quite young. The field is much bigger than the $3M project to apply MPC on the next mega-sized olefins plant. It is much bigger than the application on the "off-the-shelf" polymers facility. Mining the industry for valuable applications that may not be of the high throughput/low margin genre is widespread for exploitation. However, to do this effectively the application costs must be low. Eastman has written its own MPC code to enable "free" replication of MPC technology in addition to the learning of the technology that comes with such an endeavor. We have applied the infinite horizon model predictive control algorithm as described by Muske and Rawlings (1993). The IHMPC algorithm is based on a state space description of the process. A Kalman filter is used as a plant observer to reconstruct plant states, a quadratic program formulation is used to determine steady state process targets, and an infinite horizon linear quadratic regulator problem is solved to determine process inputs. Additional implementation details are described in Downs and Vogel (2001). Certainly, as Qin and Badgwell (1997) point out, there are numerous implementations of model predictive control algorithms. With as much research and development effort that model predictive control has commanded it would be a shame if questions such as which implementation of MPC to use, implementation costs, etc. inhibited the harvesting of the fruit this technology offers.

Our experience suggests that there are numerous good applications that require MPC to be integrated with other process control technologies. This integration demands an understanding of our chemical processes, their regulatory control strategies, the array of process control technologies available, and how to apply various technologies effectively. The variety of process control needs, process control technology, and the underlying hardware available have led to an increased need for broad based, knowledgeable process control talent.

Motivating Example—A Low Selector

The linkage between control strategy design and model predictive control has provided insight into the problem formulation and design of advanced control systems. The examples in this paper are intended to highlight the variety of application needs, MPC control structure, and the control strategy viewpoint. A motivating example centered on a distillation column control problem demonstrates the relationship between control strategy design and MPC design. Each example illustrates the importance of control strategy concepts when developing good MPC problem statements.

The concept of controlling unit operations within process constraints has been around for many years. The use of high and low selectors to prevent an operation from violating constraints can be viewed as a control strategy change agent. Consider the column illustrated in Figure 1. One control objective is to manipulate steam flow rate to control the underflow composition. If the column feed rate becomes large enough then the steam rate may increase to the point of flooding the column. A low select can be used to choose the lower of two desired steam flow set points, that requested by the column delta pressure controller or that requested by the underflow composition controller.

The low selector changes the column control strategy from an underflow composition to steam strategy to a column delta pressure to steam strategy. In the former,

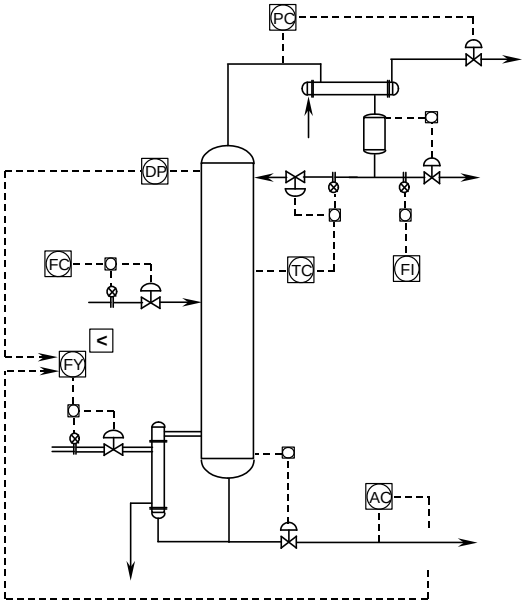


Figure 1: Distillation column with low selector on heat input.

the column loading will change in an attempt to maintain a relatively constant underflow composition. In the latter, the underflow composition will vary and yield to maintaining the column loading at a maximum value.

The steady state that this column will approach is apparent in this simple example, as are the two control loops that can be invoked. The two control loops, composition to steam and delta pressure to steam, will have different dynamics and may need to be tuned differently. Other control blocks can be added to make the transition from one strategy to another a smooth one. Current distributed control systems usually handle initialization of the non-selected controller. The influence of other control loops such as the temperature to reflux rate loop could also be incorporated. Other variables that depend on the steam rate and that need to be maintained within constraints could be added to the low selector. For example, the distillate flow rate may feed a downstream operation that has a maximum feed rate limit. If the column control demands a distillate rate that exceeds this limit, we may instead want to give up on the underflow composition to maintain temperature control. The list of possible constraint additions obviously could go on. As plant designs become more integrated this type of constraint escalation becomes more prevalent.

For the case of one manipulated variable, each constraint represents a different control strategy. However, if there is more than one manipulated variable then the number of possible control strategies is much larger. The understanding of what control strategy might be instantiated at any given time is an integral part of the design of the high and low selectors. Each possible pairing can be examined and verified for practical sense. If the number of possible strategies becomes too large to be reasonably evaluated, the high/low selectors are reconsidered and alternatives to achieve the control objective are developed. There is a self-regulating nature to the control design process—if the strategy becomes too complex to understand all that might happen, then simplifications are made.

Contrast the high/low selector design process to the model predictive control application mentality of today. It is so easy to add input or output constraints that a complete analysis of the resulting controller can become practically impossible. The unusual controller pairings that may result can be quite unexpected. Viewing model predictive control as a control strategy change agent can lead to insight into what the controller might end up controlling with what. This insight can provide guidance in what dynamic relationships are important in controller performance and robustness. Furthermore, do the operation regimes where the controller may end up make sense—even if they are stable? At Eastman we spend considerable time on determining why an undesirable outcome has occurred only to find out that the controller has done exactly what we programmed it to do. It has become evident to us that viewing MPC in light of control strategy design has made our MPC design job much easier and more intuitive.

Multivariable Control Applied to a Distillation Column

Problem Statement

Using high and low selectors for constraint control when more than one manipulated variable is involved quickly leads to application of MPC to more easily manage process constraints. When viewed as a control strategy change agent it is realized that the number of different control strategies that can be active at any one time is large. Each of these strategies can be evaluated based on numerous tools that have been developed over the years such as RGA, Niederlinski Index, SVD, etc. (Bristol, 1966; McAvoy, 1983). The issues of control loop interaction, degeneracy of degrees of freedom, and sensitivity to model error that control strategy analysis tools address can be applied to understand underlying problems in MPC applications.

Our approach in using MPC is one of understanding the control strategy that we want to invoke and how we want that strategy to change under different operating scenarios and then using MPC to accomplish this. As a result our MPC applications are studied more from the steady state viewpoint than a dynamic one. Certainly

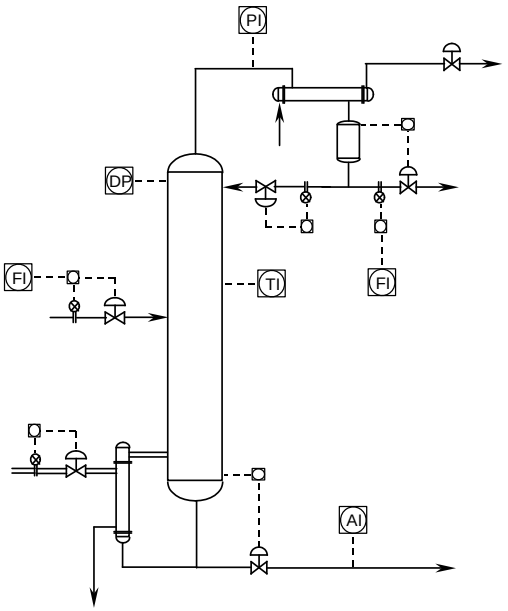


Figure 2: Distillation column.

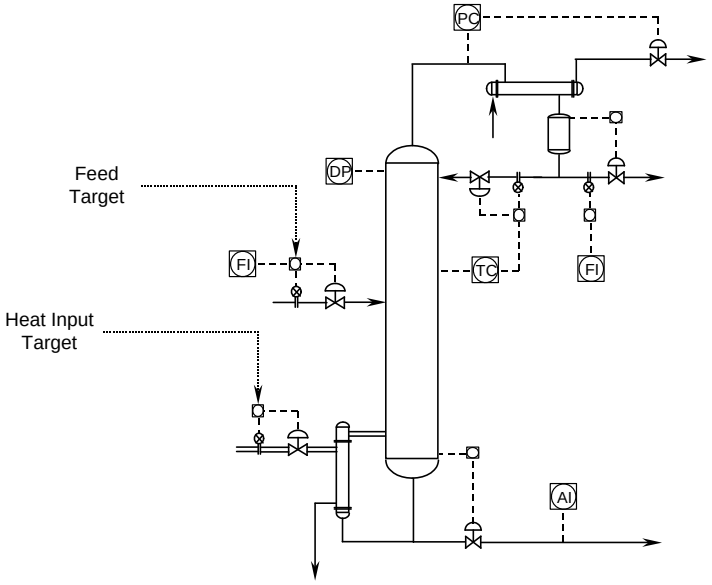


Figure 3: Distillation column hybrid control strategy.

we have cases where the dynamic advantages of MPC are exploited, but we have found that the steady state features are of most benefit.

Another important design consideration for supervisory control that employs model predictive controllers is the controller architecture. At one end of the spectrum is the flat architecture that has all the measurements and manipulated variables in one MPC controller. This structure takes on the appearance of a "black box" and it is sometimes difficult to diagnose underlying controller problems. At the other end of the spectrum is a vertical architecture that resembles a multi-layer cascade structure. This structure has the advantage of building control strategies using conventional process control notions and of segregating unrelated parts of the controller. Using the vertical structure, however, requires that the issues of controller initialization, constraint passing between layers, and controller speeds of response be managed.

To illustrate some of these issues consider the distillation column illustrated in Figure 2. We assume that process analysis has been completed to determine that the following process objectives are to be achieved:

1. Maintain tray temperature in the rectifying section at set point
2. Maintain underflow composition at set point
3. Maintain feed rate at set point
4. Maintain column pressure drop less than a given maximum
5. Maintain the column distillate rate less than a given maximum
6. Minimize energy usage

We have reflux rate, heat input rate, feed rate and cooling rate available to manipulate. Of course, even getting to this step required a decision to control reflux drum level with distillate rate and reboiler level with underflow rate. These controllers could also be added to the control strategy development problem but will be assumed as given here. Developing a control strategy to achieve these objectives is incomplete until we know what set points to give up on if the column becomes constrained—a ranking of importance is also required. For this example we will assume that the temperature in the rectifying section is the most important followed by column feed rate followed by underflow composition, which is the least important.

Case 1—A Hybrid Strategy

Consider first a hybrid strategy that employs an underlying SISO strategy that is illustrated in Figure 3. This may be a strategy that has been successfully used for many years and is effective in maintaining the rectifying temperature at its set point. Our job is simply to achieve the stated objectives by overlaying an "advanced control system" above the regulatory SISO strategy. This is common when the regulatory strategy is sound and provides stabilizing control during periods of process upsets, start-ups, etc.

A steady state process gain matrix for the resulting variables is

	Heat Input	Feed Rate
Underflow Composition	-1	0.5
Feed Rate	0	1
Distillate Rate	0.01	0.5
Pressure Drop	1	0

During normal operation with no constraint active an MPC controller will line out with the underflow composition and feed rate at set point and the distillate rate and pressure drop within limits. The control strategy active at this time has the steady state gain matrix,

	Heat Input	Feed Rate
Underflow Composition	-1	0.5
Feed Rate	0	1

Clearly we can see that the resulting MPC controller will look a lot like underflow composition to heat input and feed rate set constant with a feed forward term between feed rate and heat input. Interaction measures would say that this strategy should work fine—in fact, the relative gain for each loop is equal to one.

Consider how operations change if the pressure drop constraint becomes active. If we are to give up on underflow composition first and maintain feed rate then we end up with a gain matrix,

	Heat Input	Feed Rate
Feed Rate	0	1
Pressure Drop	1	0

which again indicates that MPC will work well. However, if it is desired to give up instead on feed rate and maintain underflow composition then the gain matrix is

	Heat Input	Feed Rate
Underflow Composition	-1	0.5
Pressure Drop	1	0

and we will have a more difficult control problem. In fact the model relating underflow composition to feed rate at constant heat input becomes more important because it is the only link that the feed rate has into the control strategy. During normal operation this model only influences the feedforward relationship between heat input and feed rate whereas in this constrained case it is the primary relationship for underflow composition control. This difference in control problem characteristics resulted from a change in the steady state weighting of the controlled variable importance. Certainly, this is an innocent change that has important ramifications on the resulting control problem.

Next consider the case where the distillate rate is constrained and again where we are to give up on underflow composition and maintain feed rate. This scenario yields the following gain matrix

	Heat Input	Feed Rate
Feed Rate	0	1
Distillate Rate	0.01	0.5

which is almost degenerate. Large steady state heat input changes are needed to have any effect on the distillate rate. Certainly for this simple example, process insight might key us into the fact that the distillate rate and the feed rate are so closely tied together that this requirement is unreasonable. However, this case was quite reasonable when it was the column pressure drop that was the constraint instead of the column distillate rate. This behavior points to the fact that steady state weighting preferences may easily lead to difficult dynamic control problems that have poor characteristics regarding interaction or robustness.

Finally consider the case where both constraints are active. The gain matrix again becomes docile and well behaved.

	Heat Input	Feed Rate
Distillate Rate	0.01	0.5
Pressure Drop	1	0

The heat input maximizes the column pressure drop to keep the underflow composition as close to target as possible and the feed rate is reduced to keep the column distillate rate within its limits. If we knew these constraints were always going to be active, this may even be a strategy we would design.

Each constraint scenario yields a different control strategy; some of these strategies are well behaved and some are clearly not strategies that we would want to deploy. The transition between different strategies is seamless and it may appear that if you can describe the control objective in terms of an MPC structure your problems are over—simply configure, tune, and start counting the savings. However, this is not the case and for MPC controllers that have these time bombs buried within them, it usually happens that the poor, unforeseen strategy gets invoked late at night or on holiday weekends.

Case 2—A Less Hybrid Strategy

If the temperature in the rectifying section of the column is in fact less important than the feed rate or underflow composition, then it is advantageous to move that control loop from the regulatory level to the supervisory level. This allows the importance of holding the temperature at set point to be given a lower weight in the steady state target calculation. As long as the temperature is controlled at the regulatory layer, the reflux will change in an attempt to get the temperature to target even though this may have a low priority. Consider a different hybrid strategy that employs an underlying SISO structure that is only used to handle level control loops. In this case the rectifying temperature and reflux rate are

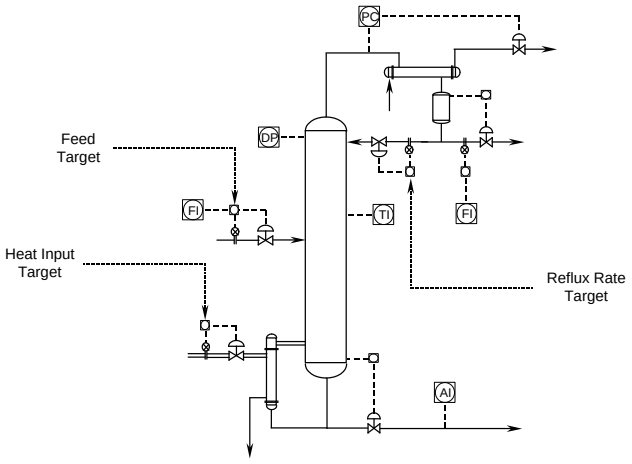


Figure 4: Distillation column hybrid control strategy with reflux rate in supervisory control layer.

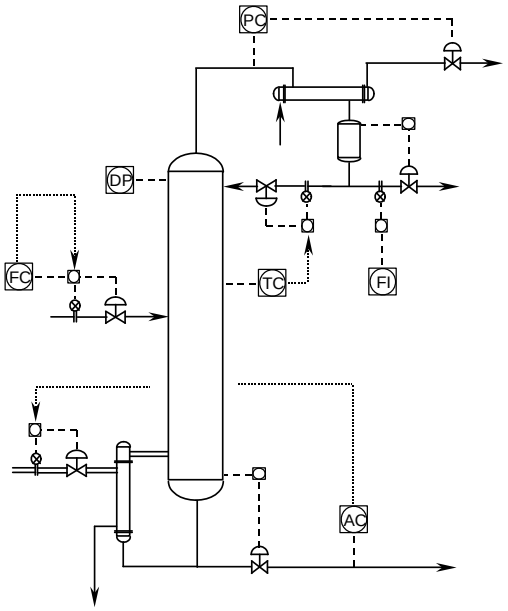


Figure 5: Distillation column control strategy during unconstrained operation.

in the supervisory layer. A candidate strategy is illustrated in Figure 4. There is advantage to leaving inventory loops at a PID level because it keeps the supervisory controller dealing only with self regulating loops and can help avoid reliability issues around advanced control systems. Our job is simply to achieve the stated objectives by overlaying an "advanced control system" above the regulatory control strategy. It is noted, however, that once the underlying inventory control strategy is chosen, many options for the overall control are eliminated.

A steady state process gain matrix for the resulting variables is given by

	Heat Input	Feed Rate	Reflux Rate
Underflow Composition	-2	0.3	1
Feed Rate	0	1	0
Temperature	1	0.1	-0.8
Distillate Rate	2	0	-1
Pressure Drop	1	0	0.1

When no constraints are active, controlling the underflow composition, feed rate, and temperature using the heat input set point, feed rate set point, and the reflux rate set point leads to a well-behaved process gain matrix. MPC in this case simply provides mild decoupling of the implied SISO loops. The RGA for this case is

	Heat Input	Feed Rate	Reflux Rate
Underflow Composition	2.66	0	-1.66
Feed Rate	0	1	0
Temperature	-1.66	0	2.66

and using it to pair loops one can envision an SISO strategy as shown in Figure 5.

When constraints become active, the implied control strategy changes and the resulting implied strategy is dependent on the importance placed on the variables having set points, that is, which controlled variables are allowed to deviate from their set point. Using a linear program (LP) for the determination of steady state operation results in answers that lie at a vertex and an implied control strategy that has controlled variables that do not line out at set point entering the problem in a sequential manner as constraints become active. Using a quadratic program (QP) for the determination of steady state operation results in answers that can look blended. For example, two controlled variables can be allowed to deviate from set point equally and can enter the problem in a parallel manner. There has been recent work to explore the steady state target problem formulation and calculation (Kassmann et al., 2000).

Consider the case where the column pressure drop constraint becomes active and the column rectifying temperature is to be allowed to deviate from set point. The resulting controller gain matrix is

	Heat Input	Feed Rate	Reflux Rate
Underflow Composition	-2	0.3	1
Feed Rate	0	1	0
Pressure Drop	1	0	0.1

An RGA calculation of this matrix is given by

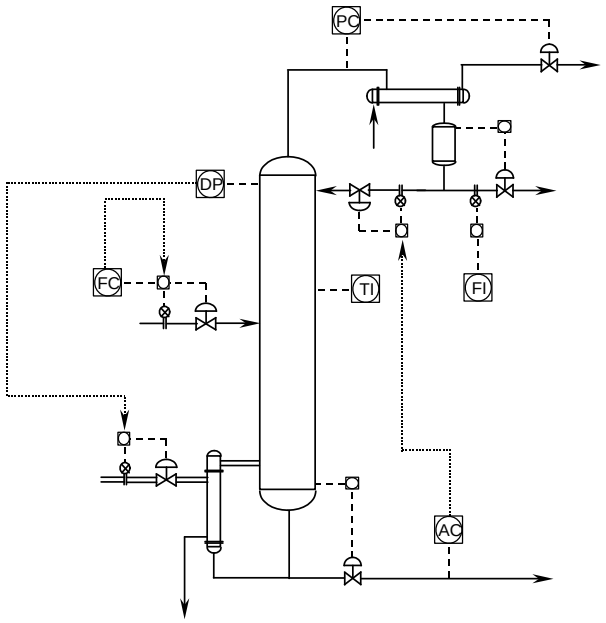


Figure 6: Distillation column control strategy when delta pressure becomes constrained

	Heat Input	Feed Rate	Reflux Rate
Underflow Composition	1/6	0	5/6
Feed Rate	0	1	0
Pressure Drop	5/6	0	1/6

and suggests a control strategy as shown in Figure 6. From the gain matrix we can see that this pairing is a pretty obvious one if the variables we are required to control are the ones shown. However, presented with this control design problem from an SISO point of view, we would probably be looking for alternatives—controlling the underflow composition with the reflux just doesn't look too promising. If we expect this case to occur then we would probably want to spend additional effort determining the relationship between reflux and underflow composition when the feed and pressure drop are constant. This is, of course, different than the relationship determined during open loop testing when the feed and heat input are constant. The understanding of what control strategies can look like under different constraint scenarios leads to insight into why an MPC controller might fail or perform poorly.

Finally, the question of including the reflux drum level, column reboiler level, and column pressure in the supervisory control layer must also be addressed. Including the level measurements and their control in MPC leads to handling a mix of self-regulating and integrating variables. It is not clear to us which predictive control technologies on the market are equipped to handle this case. Depending on the column reflux ratio, which may change during the course of operation, the level control strategy may be best left alone and on the regulatory layer or it may be paramount that it to be given over to the supervisory controller.

Incorporating the column pressure control into the supervisory layer may at first seem unwise. However, the ability to change the operating pressure of the column can lead to increased energy efficiency provided it can be done in a coordinated way with the other column controls. That, of course, is exactly what MPC does.

Process Applications

This paper contains examples of the variety of applications benefiting from our use of model predictive control. Our successful record of gaining benefit from this technology has relied upon several basic tenets. First, our ability to develop good regulatory control strategies has provided a solid foundation on which to build higher-level supervisory control systems. The benefit and results from this step sometimes indicate that this is all that needs to be done. Second, the identification of good advanced control candidates has required an understanding of the process economics to screen for high value applications. Third, the costs of solution development and implementation has been kept low and not hindered the "leap of faith" often required of operations. Fourth, a building block approach to reaching intermediate process control milestones has led to increasing complexity and value that could not be envisioned at the start of projects. Fifth, the integration of process improvement functions already in place (e.g. design of experiments, equipment design, process chemistry experiments) has led to control objectives that were unknown at project initiation. Finally, maintenance of our applications has led to new opportunities and relationships that have in turn grown this aspect of our work.

Reactor Product Crystallization Train

Illustrated in Figure 7 is a common situation where the control strategy needs to change during operation. Consider the problem of controlling the four crystallizer levels and a total throughput rate using the five manipulated variables shown. A common SISO strategy, Figure 8, would be to set the throughput rate using the reactor feed and then have a level to outflow pairing for the four crystallizers. This strategy has two problems. First the process variation will be propagated downstream and the fourth and perhaps the most important crystallizer will be the one getting the most variability in its feed flow rate. Second, if the process bottleneck is somewhere other than the reactor, then the throughput rate needs to be lowered enough to insure that the valve or crystallizer that is the bottleneck does not exceed its capability when throughput rates cycle through a maximum.

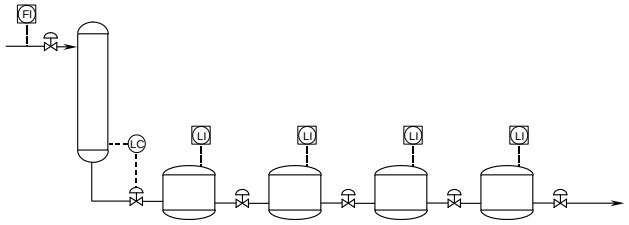


Figure 7: Reactor followed by crystallization train.

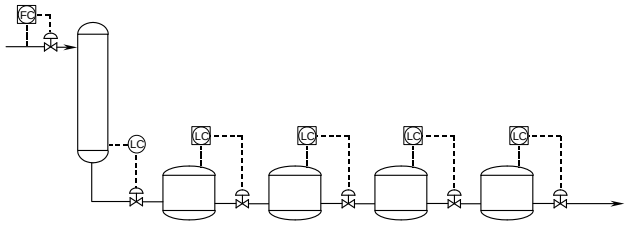


Figure 8: Original control strategy for reactor followed by crystallization train.

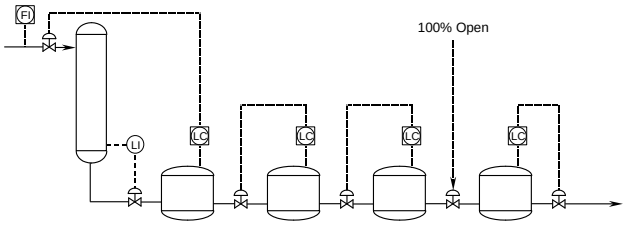


Figure 9: Level control strategy for reactor followed by crystallization train when downstream valve becomes constrained.

A model predictive controller has the ability to change the level control strategy as needed via the constraint handling. If the feed to the fourth crystallizer is the process constraint then the level control strategy becomes that shown in Figure 9. Of course this constraint, which may be more complicated than a simple valve limit, can move to different locations and a model predictive controller can accommodate this. Another advantage of a model predictive controller in this application is that it can be tuned to distribute the variability to the units that are least upset by flow variations. The level control variability in this example can be directed more toward the first and second crystallizers and away from the later ones.

There are implementation considerations needed to maintain operation when more than one manipulated

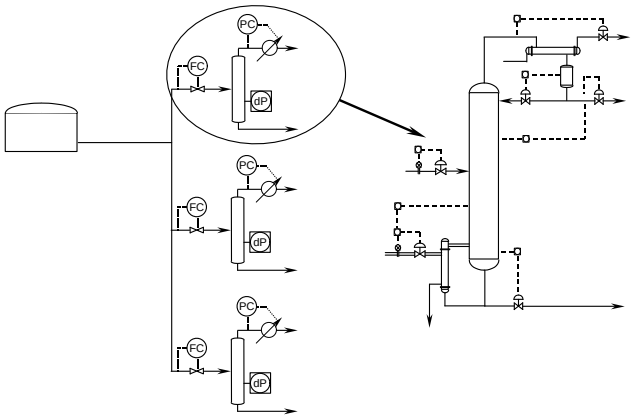


Figure 10: Parallel distillation column loading.

variable becomes constrained or gets put into manual. Procedures are needed to control the process to some extent when no steady state solution to the model equations exists. This example is one of many where management of process inventories is important. These inventory control problems can span a single process like this example or cover large networks of in-process tankage. The ability to handle integrating variables and to distribute level and flow variability is important in this category of problems. Benefits of reduced flow variability often translate into increased production rates. The ability to handle unit operation feed constraints that move from unit to unit based on processing conditions is also an important benefit arising form this type of problem.

Parallel Distillation Column Loader

Illustrated in Figure 10 is a common situation where several parallel unit operations need to be used in an efficient manner. In this example there are three isomer separation distillation columns that process a reactor effluent. The control objective is to maintain the total feed to the system at a specified target and to load the columns in an efficient manner. Each column has an effective SISO control strategy that controls end compositions by manipulating distillate rate and heat input. Manipulating cooling duty controls the operating pressure of each column. The feed rate capacity as measured by column differential pressure and the separation efficiency are a function of the column operating pressure.

The control objective can be met by manipulating the feed rate to each column and the operating pressure of each column. Certainly other choices can be made. In particular, the depth to which the individual column regulatory control strategy is included in the model predictive controller is an important decision. The column regulatory controls will have to respond when the column operating pressure is changed. If this is expected to be a slow change the supervisory system may simply layer

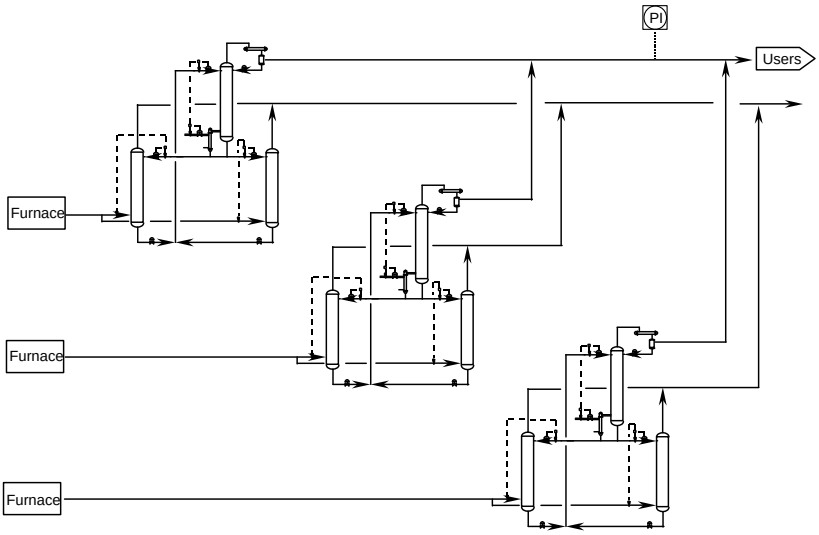


Figure 11: Process gas absorber/regenerator systems.

on top of the existing regulatory control. However, if the total feed rate and hence the pressure are expected to be changed more quickly, then the composition loops may need to be incorporated into the supervisory layer. If the economics strongly suggest that operation at minimum pressure is required, then the notion of running each column at its maximum pressure differential and manipulating the column pressures to control feed rates is not too far away. Unusual strategies like this one require process understanding to uncover pitfalls and unusual unit operation behavior that may make such a suggestion laughable.

This type of process loading to parallel unit operations is common. Usually the parallel operations have differing efficiencies that can be determined to minimize the processing costs. Often the operating efficiencies of the units are a function of how loaded the unit is. The efficiency often goes through a maximum indicating more efficient operation at higher loads up to a point after which efficiency drops off, usually very quickly.

Flue Gas CO_2 Absorber Control

A similar but different situation is illustrated in Figure 11 where three CO_2 absorber systems recover CO_2 from three different furnaces. There is a varying demand for recovered CO_2. The control objective is to recover the demand amount of CO_2 at the minimum costs. Each system has a different recovery efficiency and also has varying amounts of CO_2 that are available for recovery. Each recovery system has its own process constraints that must be honored.

A model predictive controller can be employed to manage the system. There are at least two major model predictive control strategies that are suggested. One is a flat, horizontal architecture and the other is a vertical architecture. The horizontal architecture has all the manipulated variables for each recovery system in the same MPC. This has the advantage of making all the information available in one MPC. As constraints become active in one system this information is part of the MPC calculations for the other systems. The downside is that changes in one system directly influence the other systems when local handling of disturbances might be a better alternative. In addition, there is always the possibility of a system being down or off-line requiring it to be removed from the controller.

Process control strategy notions suggest measurements that naturally reject some common process disturbances. For example, controlling percent CO_2 recovery, Figure 12, for a system may reject most of the feed rate disturbances that a local system may experience without propagating them to the other systems. Similarly, energy efficiency may also be normalized by feed rate. The characteristics and patterns of the CO_2 users can also be incorporated into the design of measurements that reject common disturbances. Process knowledge that absorber/regenerator systems of this type are suitable for ratio-oriented strategies can unload the supervisory control system to perform primarily the optimization work that it can do best.

Energy Recovery Pressure Controller

Illustrated in Figure 13 is a process environmental control/energy recovery situation where the control strategy needs to change during operation. In this example a process effluent gas stream contains components that need

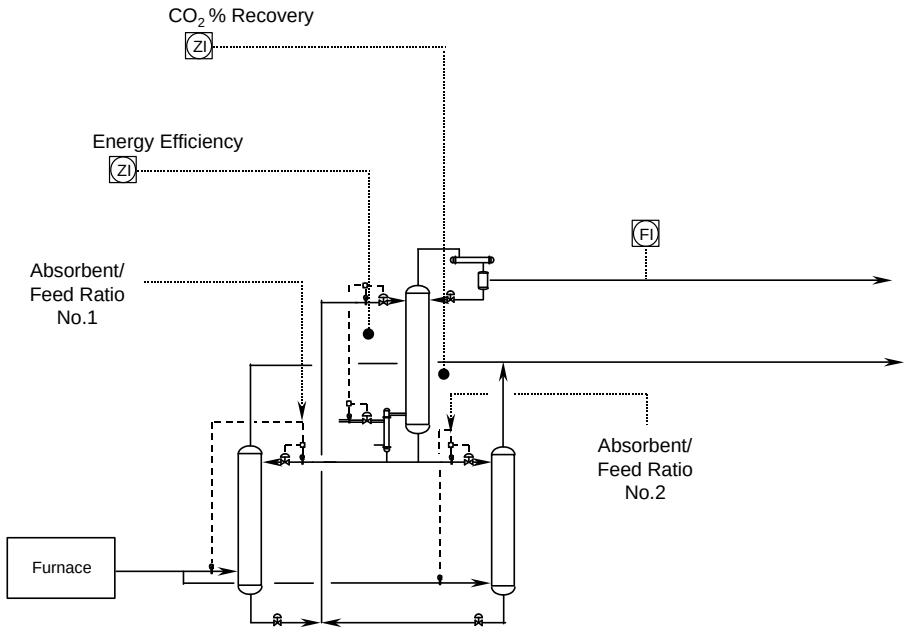


Figure 12: Individual process gas absorber/regenerator system.

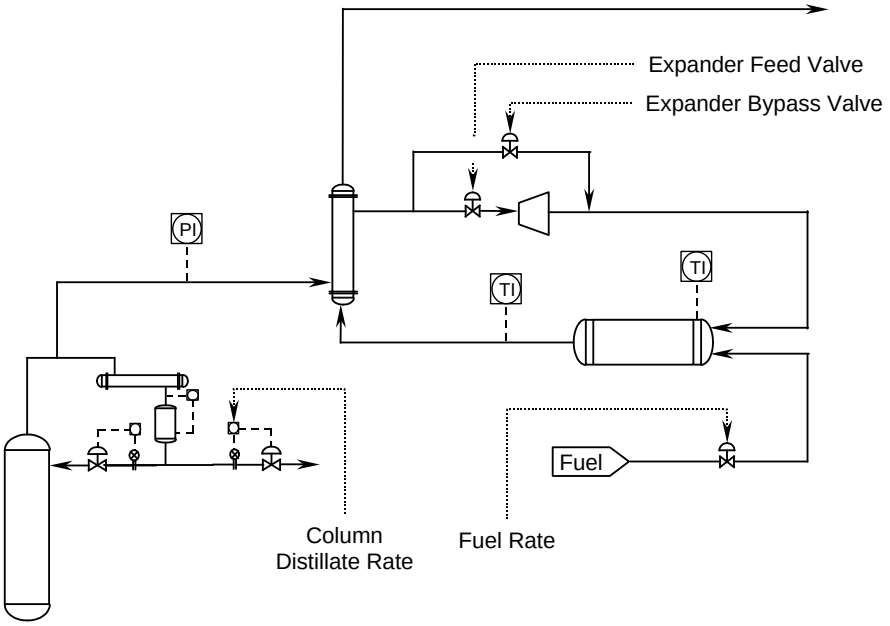


Figure 13: Process effluent catalytic oxidation process.

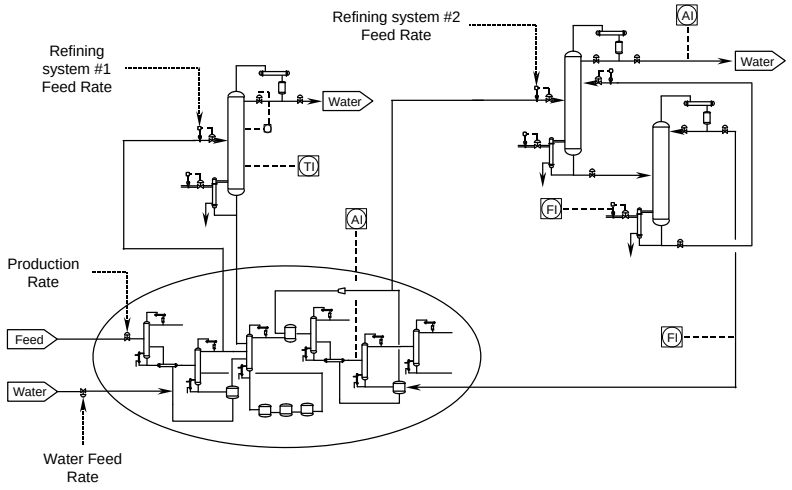


Figure 14: Plant wide water balance control.

	Refining System #1 Feed	Refining System #2 Feed	Process Production Rate	Process Water Feed (measured disturbance)
Composition	-1	-2	1	1
Temperature in Refining System #1	-2	0	-0.2	-0.1
Steam Flow in Refining System #2	0	1	0	0
Composition in Refining System #2	0	0.2	0	0
Recycle from Refining System #2	0	0.5	0	0

Table 1: Steady state process gain matrix for plant wide water balance control.

to be catalytically oxidized. The effluent gas comes from the top of a distillation column whose condenser can condense and remove some of the effluent as liquid products. The uncondensed gas is routed to a gas expander used to recover energy from this high pressure stream before it is sent to the catalytic oxidizer. The gas expander cools the gas. There is natural gas fuel that inexpensively preheats the gas to temperatures needed for catalytic oxidation to occur. There are maximum and minimum inlet and outlet temperatures that must be honored to insure proper component destruction.

The process gain matrix is given by

	Column Distillate Rate	Fuel Rate	Expander Feed Valve	Expander Bypass Valve
Pressure	-2	0	-1	-1
Inlet Temp.	0.5	2	-1	0
Exit Temp.	0.5	2	-1	0

This example incorporates the fast process dynamics associated with pressure control of gas systems. The incorporation of pressure control overrides for safety systems and the combustion control system used for the fuel can complicate the design of the advanced control strategy. The use of the advanced control system is to layer on top of the existing safety and burner control systems and not compromise their operation. In fact, these safety systems may operate on process control hardware that is distinct and loosely linked to the platform used for supervisory control. These issues may dictate the incorporation of information indicating how such systems are interacting with the supervisory control, if at all. Understanding the process control hardware, the process operational requirements, and the safety and environmental consequences are as much a part of the control system design as the control technology.

Plant Wide Water Balance Control

Illustrated in Figure 14 is a plant wide control situation where an overall control strategy needs to change during operation. In this example a plant contains a component that travels throughout the process. The component in this case is water and it is in a plant feed, is produced by reaction and is removed via two distillation systems. Refining system #1 is a simple distillation column and refining system #2 is a combination of columns. The cost of removal is different for each system and each system has process constraints that must be honored. An important control issue in this case is what measurement

or combination of measurements indicates the status of the water balance. The measurement used in this case is an on-line analyzer on a key stream within the process.

The objective of a model predictive controller is to maintain the water in balance by manipulating the feed rates to the two refining systems and, if needed, the overall plant production rate. The steady state process gain matrix is shown in Table 1.

The complications that arise in this example are: (1) the process time frames are widely diverse, (2) the process economics suggest a preferred use of the manipulated variables, and (3) the best approach to identify the status of the plant water composition may be unclear. These issues can be addressed using a layered, vertical hierarchy or a flat, "all-in-one" strategy. How one chooses to design the structure depends on not only understanding the process control ramifications of each design but also the appropriateness and timing of making changes in one part of the plant in response to upsets in another part of the process.

The ability to add additional process constraints as they are discovered integrates a longer term support role for applications of this type. While the initial installation may incorporate only the constraints listed in the process gain matrix, good process improvement work will probably eliminate them and identify new unforeseen constraints. This in turn will require the identification of new models and support for this application. For many of our applications process improvement work thrives when process constraints are clearly identified and operated against. This creates an environment of process development whose costs and benefits can clearly be identified and realized.

Conclusion

The application of process control and, in particular, model predictive control remains an active and profitable area for the chemical industry. Considerable progress has been made to provide a theoretical foundation for model predictive control and to move it into the mainstream of application. As this technology becomes more widespread the implementation issues encountered every day will begin to be addressed and this technology will mature into a powerful tool with routine application. In the meantime there remains much to develop to reach this destination. As demonstrated in the process examples, there are numerous application issues that complicate the use of MPC. These issues include controller tuning, characterizing complex performance criteria, using redundant process information, and controller robustness. The solution to these issues requires identification of the problem, short term and possibly ad-hoc fixes, and research to address the problem in more fundamental ways. The need for process control talent for academic research, process control vendors, and corporate consumers alike remains strong. Each group brings a unique and indispensable viewpoint to the effective application of process control.

Incorporating the process control strategy viewpoint into advanced control design has provided Eastman with a very high success rate when applying advanced control technology. The understanding and incorporation of process knowledge continues to be invaluable in the successful application of new control techniques. The advent and routine use of model predictive control has not supplanted the need nor the value of process understanding in the successful application of process control technology. On the contrary, model predictive control has amplified the need for process and control strategy analysis and understanding.

References

Barolo, M. and C. A. Papini, "Improving Dual Composition Control in Continuous Distillation by a Novel Column Design," *AIChE J.*, **46**(1), 146–159 (2000).

Bristol, E. H., "On a new measure of interaction for multivariable process control," *IEEE Trans. Auto. Cont.*, pages 133–134 (1966).

Downs, J. J. and J. E. Doss, Present status and future needs—a view from North American industry, In Arkun, Y. and W. H. Ray, editors, *Chemical Process Control CPC IV*, pages 53–77, Austin, TX. CACHE (1991).

Downs, J. J. and E. F. Vogel, Industrial Experience with State-Space Model Predictive Control, In Rawlings, J. B., B. A. Ogunnaike, and J. W. Eaton, editors, *Chemical Process Control CPC VI*, Austin, TX. CACHE (2001).

Downs, J. J., Private communication with PIP representatives in process control (2000).

Groenendijk, A. J., A. C. Dimian, and P. D. Iedema, "Systems Approach for Evaluating Dynamics and Plantwide Control of Complex Plants," *AIChE J.*, **46**(1), 133–145 (2000).

Kassmann, D. E., T. A. Badgwell, and R. B. Hawkins, "Robust steady-state target calculation for model predictive control," *AIChE J.*, **46**(5), 1007–1024 (2000).

Luyben, W. L., "Plantwide Design and Control of Processes with Inerts. 1. Heavy Inerts.," *Ind. Eng. Chem. Res.*, **37**, 528–534 (1998a).

Luyben, W. L., "Plantwide Design and Control of Processes with Inerts. 1. Intermediate Inerts.," *Ind. Eng. Chem. Res.*, **37**, 535–546 (1998b).

Luyben, W. L., "Plantwide Design and Control of Processes with Inerts. 1. Light Inerts.," *Ind. Eng. Chem. Res.*, **37**, 516–527 (1998c).

Luyben, W. L., "Design and Control of Gas-Phase Reactor/Recycle Processes with Reversible Exothermic Reactions," *Ind. Eng. Chem. Res.*, **39**, 1529–1538 (2000).

McAvoy, T. J., *Interaction Analysis—Principles and Applications.* Instrument Society of America, Research Triangle Park, NC (1983).

Muske, K. R. and J. B. Rawlings, "Model Predictive Control with Linear Models," *AIChE J.*, **39**(2), 262–287 (1993).

Qin, S. J. and T. A. Badgwell, An overview of industrial model predictive control technology, In Kantor, J. C., C. E. García, and B. Carnahan, editors, *Proceedings of Chemical Process Control—V*, pages 232–256. CACHE, AIChE (1997).

Tseng, J. L., W. R. Cluett, and W. L. Bialkowski, "New Analysis and Design Tool for Achieving Low Variability Process Designs," *AIChE J.*, **45**(10), 2188–2202 (1999).

Evolution of an Industrial Nonlinear Model Predictive Controller

Robert E. Young, R. Donald Bartusiak and Robert W. Fontaine
Engineering and Manufacturing Services Department
ExxonMobil Chemical Company
Baytown, Texas 77520

Abstract

Motivated by a specific manufacturing problem in 1990, Exxon Chemical Company embarked on the development of a nonlinear multivariable model-based predictive controller. The controller's evolution has included collaboration among academic researchers, engineers from industry, and process control software vendors. The resulting control algorithm was patented by Exxon Chemical Company and commercialized by Dynamic Optimization Technology Products, Inc. At the same time, several other academic interactions produced results supporting the implemention of these controllers in our manufacturing facilities. This paper chronicles the evolution of the controller development, and presents the details of the control algorithm. The control algorithm features are discussed, and where applicable, compared to other model predictive control (MPC) algorithms. Finally, two industrial examples are presented to illustrate the methodology.

Keywords

Model-based control, Predictive control, Nonlinear control, Industrial control

Introduction

For the past 10 years, Exxon Chemical Company has pursued development of a methodology to address industrial process control challenges characterized by nonlinear process responses. Motivated by problems in manufacturing plants, the evolution of this nonlinear control methodology has drawn on the expertise and experience of practitioners (both from manufacturing sites and technology organizations), of academic researchers, and of process control vendors.

Over these same 10 years, many changes have occurred within our company, the process control industry, and to the process control systems technology. Academic research has significantly increased our understanding of MPC, especially stability of linear MPC algorithms. Computer science and optimization technologies have improved vendor's packages, making industrial implementation easier and more effective. Through all of these changes, the motivation and development of this algorithm have persisted. While by no means complete, the evolution of this control algorithm is an interesting story about how diversely motivated groups of people can interact to produce a tool capable of solving commercially relevant and intellectually challenging process control problems.

Motivating Problem

In 1990, Exxon Chemical Company started up a new polymerization plant using a new catalyst system. While the details of the process are proprietary, the process involved a single reactor vessel with a simple monomer recovery/compression recycle. The product is a copolymer composed of two monomers. The control objective is to control particular polymer resin properties, specifically polymer melt viscosity and polymer density. In this case, these variables are controlled by manipulating reactor feed temperature and feed composition. The reactor pressure, feed flow rate and feed temperature are measured disturbance variables. None of the in-reactor compositions are measured. When compared to similar plants, the control of this unit was unable to achieve the expected prime or "right-first-time" production.

Polymerization processes have been considered challenging process control problems for many years (Ray, 1986). In these processes, the goal is to control polymer product properties, such as polymer melt viscosity and comonomer incorporation, as well as manufacturing targets such as production rate and slurry concentration. Reaction temperature has very significant effects on reaction rates, and hence, both the polymer properties and the process operability. Typically, these variables are controlled to targets by manipulating the feed rate and composition, catalyst feed rates, and reactor cooling. Often, regulatory control of polymerization reactor is achieved with a combination of PID feedback/feedforward and ratio controllers (Congalidis et al., 1989). These control schemes are often adequate for regulatory control because the process is *linear* enough near the operating point that more sophistication is not warranted. This observation continues to be true for *many* industrial polymer processes operating today.

However, the apparent gains and time constants between the control variables and manipulated variables often exhibit significant nonlinear behavior when a polymer plant makes different grades of polymer. Often, the simple regulatory control schemes must be tuned at each operating condition to achieve good control over the entire operating window of given plant. To maximize prime production, manufacturing planning attempts set schedules with every adjacent grade having overlapping specifications with the previous grade. Also, the process nonlinearity must not be severe enough to cause signif-

icant overshoot. Often, planning can not achieve either of these objectives, resulting in off-prime polymer production.

The transition control problem has been examined by both academic researchers and industrial practitioners (McAuley and MacGregor, 1992; Debling et al., 1994). Gain scheduling or multiple-model controller designs have been suggested approaches to solving this problem. However, these approaches require a model or data for each grade and, need an algorithm to decide when to switch. The downside to either approach is the significant added cost of controller maintenance. Industrially, a nonlinear predictive controller designed to execute polypropylene reactor transitions was successfully implemented (Hillestad and Andersen, 1994). This controller design is characteristic of linear MPC applications with the exceptions that it employs a nonlinear model and includes a state estimator.

Finally, polymerization processes are very susceptible to changes in unmeasured disturbances such as very small concentrations of polymerization poisons in the feed or catalyst activation changes for any variety of reasons. A recent study of a polymerization process demonstrates that linear MPC can not achieve acceptable controller performance when faced with typical industrial disturbance signals (Bindlish and Rawlings, 2000).

In the case of the new Exxon Chemical facility, the combination of the process/equipment design and new catalyst chemistry resulted in a highly-interactive nonlinear process. The nonlinear effects of both measured and unmeasured disturbances could not be rejected by the state-of-the art control technology used on similar reactor systems. In short, both the regulatory control and the transition performance was limited by the nonlinear behavior of the process to either servo or load changes.

Early Controller and Model Development

During this same time frame, multivariable model predictive control based on identified linear process models (Cutler and Ramaker, 1980; Richalet et al., 1978) was being used to solve significant industrial control problems. At the time, nonlinear control was already a strong area of academic research and significant effort had been made to develop nonlinear MPC algorithms (Bequette, 1991). The program at CPC IV (Arkun and Ray, 1991) contained several presentations on both topics indicating that both industry and academia had already recognized the importance of both technologies.

After careful examination of the motivating polymerization control problem, the Exxon Chemical process control technology organization determined that using linear MPC could not yield the desired process performance. Even given the significant academic research, there was no commercially available software to bring nonlinear MPC technology to bear on the polymerization problem.

Early Academic Collaboration

Specifically focusing on how to model and control the nonlinear behavior of the process, Exxon Chemical Company elected to collaborate with academic researchers. A request for competitive bids was issued and awarded for two specific projects. Both of these projects were contractual agreements with specific milestone dates, deliverables, and non-disclosure agreements.

First, to focus on understanding the process, they contracted the University of Maryland to develop nonlinear models of the polymerization reactor. As a result of this effort, Professor K. Y. Choi and co-workers developed a fundamental model of the reaction process and estimated kinetic parameters for this model from pilot-plant data. This process model is composed of the dynamic mass and energy balances that describe the polymer reaction system. The polymer population balances were condensed through the use of moments (Ray, 1972) after applying the quasi-steady-state assumption to the growing polymer chains. The model is very similar to others found the open literature (McAuley et al., 1990; Zacca and Ray, 1993; Ogunnaike, 1994) The fundamental model was combined with empirical correlations to relate the polymer moments to polymer resin properties.

Also, Georgia Institute of Technology was contracted to investigate and develop a nonlinear state estimator and multivariable predictive controller to be used to control the polymerization reactor. This work was conducted by Professor Yaman Arkun, Professor Joseph Schork and their co-workers. The state estimation work revolved around the implementation of different Kalman filter and Luenberger observer designs. The controller algorithm developed was a quasi-linearized QDMC algorithm (Peterson et al., 1992; Charos and Arkun, 1993; Sriniwas et al., 1995). The controller used the nonlinear model to predict process trajectories and to compute disturbance estimates. During the final stages of this contract work, the observer/NLMPC algorithms were tested using the fundamental model developed at the University of Maryland.

Both of these programs were two year contracts, successfully meeting all of the expectations set out at the beginning. However, as is often the case, these initial investigations were most successful at providing a more detailed specification about how Exxon Chemical wanted to address both the specific polymerization control problem and the general nonlinear control problem.

Internal Development Program

In the first quarter of 1993, the Exxon Chemical project team evaluated the results of the two academic contracts and defined an internal development project. This project was focused on both the specific polymerization reactor control problem and the development of a nonlin-

ear MPC (NLC) structure for use within Exxon Chemical Company.

The earliest milestones involved evaluation and on-line implementation of the modeling equations. The predictions made using the academic model did not match the plant data well. After some analysis, the model structure was modified and new parameters estimated from plant data. After these two changes, the model tracking errors were significantly reduced.

During the modeling phase of the development, work was underway formulating the controller. Characteristic of the NLMPC developed at Georgia Institute of Technology, any nonlinear controlled variable (CV) dynamics were retained in the closed-loop response of the controller. In other words, the closed loop CV response was still a function of the operating point. Illustrating this behavior is easiest with an example. Consider the following nonlinear single-input single-output(SISO) process:

$$400x^3\frac{dx}{dt} = -x + 0.7(1-x)e^{\frac{-1}{u}} \qquad (1)$$

This contrived example is motivated by a component mass balance, a single irreversible reaction, and a concentration dependent density. The parameters have no special meaning but were selected to provided a simple and illustrative example. Figure 1 is the response of this system of two setpoint changes—from 0.25 to 0.30 and from 0.30 to 0.35. The process is controlled using a linearized MPC controller as implemented in the Mathworks Model Predictive Control Toolbox (Morari and Ricker, 1995). The controller parameters specified for these simulations are the prediction horizon equal to the control horizon set to 25, the output weight (ywt) equal to one and the input weight (uwt) equal to 0.04. The manipulated variable (MV) or input is constrained to be greater than 0.05 and less than 10.0. The process model used by the controller is a linearized model of Equation 1 around the initial operating point for each setpoint change respectively.

Without dwelling on the tuning of the controller, the observation made by the development team is easily observed in this example (see Figure 1). The dynamics of the nonlinear process are not compensated for by the controller and appear in the closed-loop response. This observation is not surprising for MPC algorithms that use move suppression as their primary tuning mechanism. Move suppression allows the engineer to indicate how much input energy can be used. Excessive use of MV moves is penalized, sacrificing CV response. This specification amounts to stating that the amount of allowable change in the process inputs does not depend on the operating conditions. In the case of the polymerization problem, this is equivalent to stating that for some products, a slow transition response is acceptable even though a faster response is achievable. Re-tuning the controller for each operating point or, perhaps, parameterizing the move suppression factor as a function of the operating point, may provide a way to circumvent this issue.

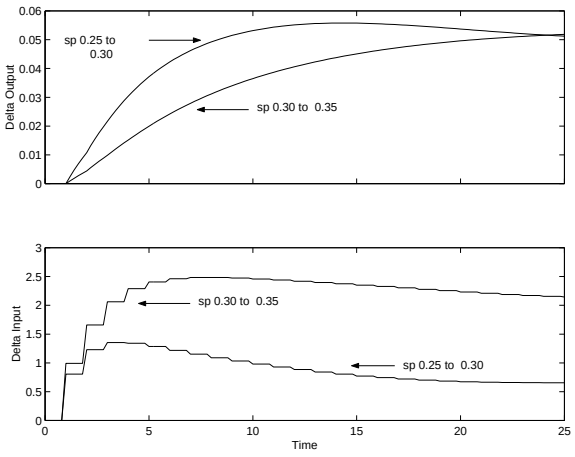


Figure 1: Quasi-linearized MPC simulation responses.

However, to the credit of the development team, the philosophy driving this effort was to avoid the inclusion of techniques that add to the life-cycle cost of the controller. Techniques like this include the use of:

- multiple linear models to approximate nonlinear models,
- linearization at each sample time to approximate nonlinear models,
- gain or tuning scheduling as a function of operating point.

Each of these techniques requires additional overhead in both the development and the maintenance of the controller. If the controller can be designed and implemented directly, this additional cost can be avoided and the controller is more likely to remain on-line.

To address the nonlinear CV dynamics, a reference system performance specification was added to the NLC design. Reference system synthesis (Bartusiak et al., 1989) is one of a class of differential geometric methods used for nonlinear control design (Lee and Sullivan, 1988; McLellan et al., 1988; Kravaris and Kantor, 1990a,b). Reference system synthesis employs a performance specification on the error trajectory to design a nonlinear control law. In a non-predictive form, this design methodology was studied as way to control product properties in gas-phase polyethylene reactors (McAuley and MacGregor, 1993).

In this development, a reference system specification is added as a soft constraint to a model predictive controller. The resulting control algorithm is tuned by specifying the desired process error trajectory for each CV and the relative weight for each of the CVs. Figure 2

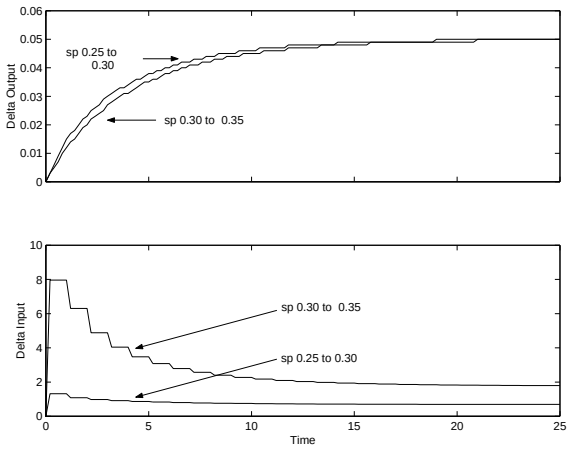


Figure 2: NLC simulation responses.

presents the result of applying this control algorithm to the nonlinear SISO example (Equation 1) for the same two setpoint changes.

The reference system specified describes a second-order over-damped response with a dominant time constant of 5. The prediction horizon and control horizon are set to 25, the same in the previous simulation. In the unconstrained case, both CV responses are identical and equal to the setpoint change from 0.25 to 0.30. However, the simulation shows the effect of the high MV constraint of 10 on the setpoint change from 0.30 to 0.35 (note: u(x=0.30)=2.048, so $\Delta u = 7.952$ at t = 0 is on the constraint). The quasi-linearized MPC move suppression factor, uwt = 0.04, was selected to generate a MV response peak value for the setpoint change from 0.25 to 0.30 nearly equal to the NOVA Nonlinear Controller (NLC) peak value. Without reducing this move suppression factor, the quasi-linearized MPC controller will not make use of the available input energy to achieve the desired performance in the second setpoint change.

Besides the addition of the reference system tuning, the project team chose a controller design based on a simultaneous optimization/solution algorithm. The implementation was augmented with the necessary code to initialize the controller. With these modification, the project team completed the controller in early 1994 and completed closed-loop testing by year end 1994. All of the original performance issues that motivated the work beginning in 1990 were addressed. Regulatory control achieved performance on par with similar processes. Transition times were reduced by at least a factor of 2 and were no longer limited by controller performance.

The Nonlinear Control (NLC) Algorithm

The inventors of the control methodology, Fontaine and Bartusiak, were granted a patent (Bartusiak and Fontaine, 1997). The methodology was later commercialized by Dynamic Optimization Technology (DOT) Products and is called NOVA Nonlinear Controller (NOVA NLC, 1997). The algorithm is a nonlinear program (NLP) optimization problem with a multi-objective cost function. The optimization problem is solved using the NOVA DAE solver over a finite time horizon. The solver uses orthogonal collocation to discretize the equations in time. The mathematical formulation of the controller is as follows:

$$\min_{u_{MV}} \Phi = \mu_1 J_1(e) + \mu_2 J_2(y,x,u) + \mu_3 J_3(\Delta u_{MV}) \quad (2)$$

subject to

$$0 = f(y, \dot{x}, x, u, \theta) \quad (3)$$

$$0 = g(y, \dot{x}, x, u, \theta) \quad (4)$$

$$x(0) = x_0 \;\; and \;\; y(0) = y_0$$

$$0 = \frac{\tau_i}{4\xi_i^2}\frac{d^2 y_{cv_i}}{dt^2} + \frac{dy_{cv_i}}{dt} + \frac{1}{\tau_i}\left(y_{cv_i} - y_i^{sphi}\right) - e_i^{sphi} + s_i^{sphi} \quad (5)$$

$$0 = \frac{\tau_i}{4\xi_i^2}\frac{d^2 y_{cv_i}}{dt^2} + \frac{dy_{cv_i}}{dt} + \frac{1}{\tau_i}\left(y_{cv_i} - y_i^{splo}\right) - e_i^{splo} + s_i^{splo} \quad (6)$$

$$e_i^{sphi}, e_i^{splo}, s_i^{sphi}, s_i^{splo} \geq 0$$

$$u_{mv_j}^{LB} \leq u_{mv_{j,k}} \leq u_{mv_j}^{UB} \quad (7)$$

$$|u_{j,k} - u_{j,k-1}| \leq \Delta u_{mv_j}^{B} \quad (8)$$

The objective function (2) is composed of three components. J_1 is the cost associated with the dynamic response of the closed-loop system. J_2 is the economic cost associated with each of the output and input variables. Finally, J_3 is the cost of moving each of the individual manipulated input variables. Each of these components is weighted by the μ weights in the equation.

Equations 3 to 4 define the nonlinear process model where y are the outputs, x are the states, u are the inputs and θ are parameters. The set of outputs, y, is composed of controlled variables, y_cv, measured outputs, and auxiliary outputs. The set of inputs u, is composed of manipulated variables, u_{mv}, feed forward variables, u_{ff}, and disturbance variables, u_d. The initial conditions on both the outputs and the states are specified.

Equations 5 and 6 describe the reference system performance equations for each of the controlled variables. Changing the tuning parameters τ_i and ξ_i, changes the reference system specifying the desired closed-loop performance for that variable. The errors, e_i^{sphi} and e_i^{splo}, are the absolute deviation between the reference system trajectories and the predicted response trajectories.

These equations are flexible enough to support a single target setpoint or a setpoint high/low window.

Equations 7 and 8 specify constraints on manipulated input variable values and changes.

In addition to the upper and lower bounds on u_{mv} and the bound on Δu_{mv}, the user has to specify the prediction horizon, the control horizon, the sampling period, the weights that rank the controlled variable errors (both negative and positive), the costs of any variables in J_2, the cost of each Δu_{mv}, and finally the relative weight of each of the three objective function components.

Also worthy of note, the implementation of this controller in the NOVA DAE system does not categorize variables this same way. This presentation of the variables focuses on variables from a control perspective. In the NOVA DAE implementation, variables are classified in a much more mathematical way. Specifically, variables are either integrated or non-integrated variables. Integrated variables are generally the dependent variables (outputs and states) while the non-integrated variables are typically the independent variables (inputs and parameters). The specification of both the initial state and output variables in the model is a result of this mathematical view of the problem.

Objective Function Details

The NLP objective function is a weighted composite of three cost functions. The first component is a measure of the cost to get the closed-loop system to target. This cost is computed by

$$J_1(e) = \frac{1}{n_p} \sum_{i=1}^{n_{cv}} \sum_{k=1}^{n_p} w_i^{sphi} e_{ik}^{sphi} + w_i^{splo} e_{ik}^{splo} \tag{9}$$

where n_p is the length of the prediction horizon, n_{cv} is the number of controlled output variables. The errors are computed at those times corresponding to knots in the collocation grid. This discrete sampling structure also applies to the other components of the objective function, J_2 and J_3.

The economic cost of the outputs, states, and inputs make the J_2 component of the objective function as follows:

$$J_2(y, x, u) = \frac{1}{n_p} (C_y + C_x + C_u) \tag{10}$$

where

$$C_y = \sum_{i=1}^{n_y} \sum_{k=1}^{n_p} c_{y_i} y_{i,k} \tag{11}$$

$$C_x = \sum_{m=1}^{n_x} \sum_{k=1}^{n_p} c_{x_m} x_{m,k} \tag{12}$$

$$C_u = \sum_{j=1}^{n_u} \sum_{k=1}^{n_p} c_{u_j} u_{j,k} \tag{13}$$

Note that any output, input or state can be included in the evaluation of this cost function.

Finally, the last term of the objective function penalizes manipulated variable moves as follows:

$$J_3(\Delta u_{mv}) = \sum_{j=1}^{n_u} \sum_{k=1}^{n_q} \lambda_j \Delta u_{mv_{j,k}} \tag{14}$$

This term in the objective function serves to ensure a unique solution of the NLP for nonzero weights. Also, this term is much like a move-suppression term in a DMC-type MPC controller. In an unconstrained DMC controller, move-suppression alters the performance of the controller by altering the singular values of the dynamic matrix. The dynamic matrix is the mapping of the control moves to future prediction errors. If this matrix is nearly singular, the inverse mapping will generate large control moves for small predicted errors. Increasing the move suppression factors serves to stabilize this inverse mapping, making it more robust to small errors. Increasing the move suppression factors even further serves to de-tune the controller. The penalty on $\Delta u_{mv_{j,k}}$ has the same effect on this controller.

Model Specification

The model described by equations 3 through 4 define a very general structure. The model equations must be compliant with the NOVA DAE format—a continuous-time open-equation residual form. Our experience has been to use fundamental models based on mass and energy balances coupled with empirical correlations and algebraic relationships. Many of the non-empirical nonlinear algebraic relationships are derived from the applicable kinetic, transport, and thermodynamic relationships. Besides supporting these types of models, this framework can easily support algebraic empirical models, neural network models and linear state-space models. Because the DAE system expects to use collocation, using neural network or discrete linear/nonlinear models will require some additional effort by the user. Typically, the empirical correlations used relate product properties to model states and/or other outputs.

Most of these models include parameters that must be specified. The software currently includes that capability to estimate these parameters from steady-state data. The parameter estimation case is defined as a least-squares fit of the parameters subject to the model and the steady-state data. In practice, getting sufficient data to estimate all parameters is difficult. There is still significant art in estimating parameters for these models.

Fundamental models have provided several advantages over models identified from process data. The fundamental models have extrapolated well to new operating conditions. When the process design changes, these models can be changed more easily than equivalent models identified from process data. This comparison is not neces-

sarily true if the process chemistry changes because of the parameter estimation issues discussed above.

Process Feedback and Initialization

Since the user is responsible for writing all of the model equations, it is possible to incorporate process feedback in several different ways. The base implementation has a default state correction built into the algorithm. The state estimates at the previous sampling time (from the previous controller execution) are assumed to be known. The state estimates at the current time are computed from this "known" starting point by integrating forward with the measured or known MVs. As implemented, this correction produces corrected state estimates when the actual inputs do not match the values computed at the previous execution.

Given this scheme of updating the state estimates, to start execution of the controller,"cold-start" estimates of the states must be computed to turn the control on. These "cold-start" estimates can be computed by solving the steady-state model equations before the first controller execution.

Output measurements can be used in a variety of ways to provide feedback. Probably the easiest and most common method is to compute a bias between the modeled value and the measured value. This bias can be filtered and used as a feedforward input into the model. This form of output feedback is similar to the approach used by linear MPC controllers.

As indicated in a review of commercial nonlinear MPC offerings (Qin and Badgwell, 2000), this controller provides state estimation through an extended Kalman filter(EKF). While technically correct, the model designer must incorporate the disturbance model into the model equations and augment these equations with the EKF equations. Alternatively, since the initial conditions can be imported into the problem, an external state estimation application can generate new initial state and disturbance estimates at each execution. There are many available methods to incorporate process feedback associated with nonlinear MPC (Muske and Edgar, 1997), however, the burden of implementing these options in the NLC methodology rests with the application/model developer.

Nonlinear Controller Stability

There are no known stability results for the controller as defined by equations 2 through 8. However, controller stability was one of the most significant areas of academic research during the decade. The formulation of the linear constrained MPC problem by Rawlings and Muske (1993) opened up many new academic studies in the area, see (Lee and Cooley, 1996; Meadows and Rawlings, 1997; Mayne et al., 2000) for reviews. However, even now, very few of these results have found their way into commercially available linear MPC products.

However, the academic stability research provides "comfort" to users of the NLC technology. The NLC objective function, equation 2, is very similar to a linear MPC objective function with soft constraints that has been shown to be nominally stable (Scokaert and Rawlings, 1999). Given these similarities, there is guarded optimism that careful selection of horizon lengths and objective function weights will result in stable closed-loop behavior.

There is also concern about optimization algorithmic stability, specifically, will the solution algorithm converge every sampling time. Again, academic work in this area (Wright, 1996; Biegler, 2000) would indicate that NLP codes are improving and can be tailored to the MPC problem to improve the convergence properties. Finally, our experience with industrially-used codes for real-time optimization and other on-line applications has been good and would indicate that reliability of this solver will not impact the success of the controller.

Controller Commercialization

After successfully implementing the first generation of the in-house version of the control methodology, Exxon Chemical Company had to choose how best to deploy this technology to the rest of the organization and how to keep the controllers on-line. Several of the following factors were studied:

- the cost of maintaining the controller software,
- the competitive advantage gained by keeping the technology proprietary,
- the competitive advantage in the corporate capability to develop process and disturbance models for our processes.

This analysis led Exxon Chemical to commercialize the nonlinear control methodology. After the award of the patent in 1997, a vendor evaluation was conducted. Subsequently, a contract was awarded to DOT Products to develop a commercial version of the software.

The on-line version of this commercial product has been used for subsequent implementations. A PC-based configuration and tuning tool has been recently released to assist with the development of new controller implementations. The configuration tool allows the engineer to specify the controller parameters, and given a nonlinear model, build an off-line controller. This model/controller combination can be used to estimate parameters from steady-state data and to perform steady-state, dynamic, and interactive simulations of the controller. While there is significant room for improvement with the graphics and user interface, the configuration tool represents a significant step forward on the path to new implementations.

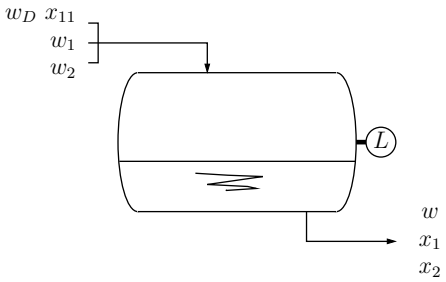


Figure 3: Blending drum schematic.

Blending Example

To illustrate the controller, a small blending example is presented. In this example, three flows are blended into a horizontal tank as show in Figure 3.

Two of the flows, w_1 and w_2 are pure components 1 and 2, respectively. The third flow, w_D is a mixture of the third component and component 1. The weight fraction of component 1 in w_D is x_{11}. The controlled variables are the tank level (L), and the weight fraction of components 1 and 2 (x_1 and x_2) in the effluent flow, w. The manipulated variables are the three inlet flow rates. The system often sees disturbances in the effluent flow, w, and the weight fraction, x_{11}, both measured.

The model for this blending process is the overall mass balance and the two component mass balance for components 1 and 2 given by

$$M = \rho V \tag{15}$$

$$V = aL^3 + bL^2 + cL + d \tag{16}$$

$$\frac{dV}{dL} = 3aL^2 + 2bL + c \tag{17}$$

$$0 = -\rho \frac{dV}{dL}\frac{dL}{dt} + w_D + w_1 + w_2 - w \tag{18}$$

$$0 = -\frac{d(x_1 M)}{dt} + x_{11} w_D + w_1 - x_1 w \tag{19}$$

$$0 = -\frac{d(x_2 M)}{dt} + w_2 - x_2 w \tag{20}$$

where M is the mass of the tank contents, V is the associated volume, L is the associated tank level, and the other variables as shown in Figure 3. The code to model this blending process for the NLC is less than one page. Except for the coefficients of the cubic polynomial relating level to volume, there are no parameters. The volume of the tank as a function of level is known from the vessel strapping chart and can be fit to a cubic polynomial. The error in this fit is less than the expected error in the level measurement. More importantly, there are no parameters that need to be fit to process data or that change with the operating point. Also, there is no special treatment of the model because it contains an integrating mode. The nonlinearities in this problem are mild, resulting from the cubic relationship between level and volume and the bilinear relationships between flows and compositions.

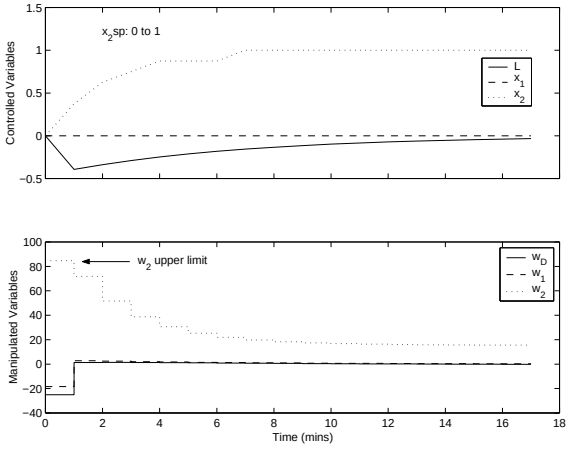


Figure 4: Blending concentration setpoint response.

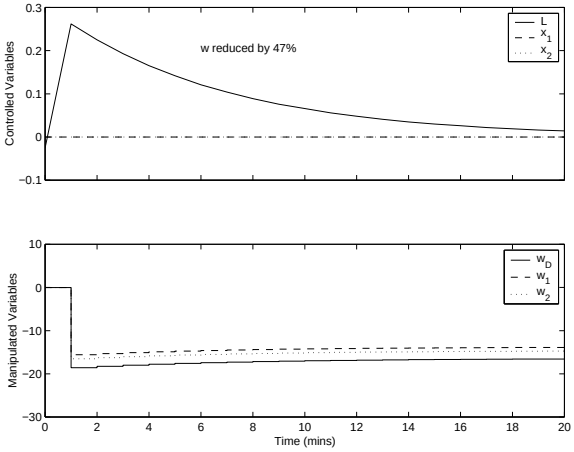


Figure 5: Effluent flow disturbance response.

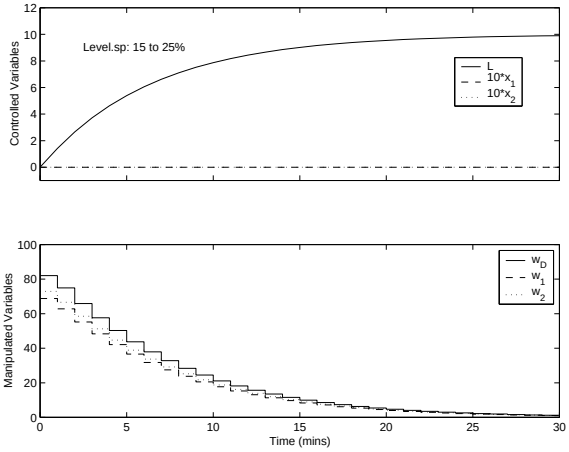


Figure 6: Level setpoint change response.

The disturbances in the effluent flow can be as great as 50% of the nominal value. Setpoint changes in the composition, x_2, are often made for process reasons. Level setpoint changes are made occasionally to reduce the residence time. Under normal operating conditions, this drum has about a one hour hold-up capacity. Level control is not a priority but it is important to keep the inventory to a manageable level without running the drum dry. Control of the effluent concentration is very important. Given these relative priorities, the controller tuning was specified so that the level response time constant would be approximately six minutes and the concentration response time constants would be approximately 1.8 minutes each. The prediction and control horizon are both five, with a sampling time of one minute. The relative weights on the CV response were all specified to be one and the weights on the economic and input move components of the objective function were set to zero.

Figure 4 shows the response to a setpoint change doubling the x_2 concentration, presented as a deviation from normal operating conditions, normalized to the full-scale value. This response is very much like the minimum-time optimal control that one should expect given the tuning. The w_2 flow is driven to a constraint and the other flows are set to achieve the desired concentration responses, temporarily sacrificing the level response. The level response returns to target along the specified error trajectory.

The response of the closed-loop system to change in the effluent flow to 53% of the nominal value is presented in Figure 5. This simulation shows that the expected level deviation is very small, the result of having a perfect model. Since the conventional controls on this unit operation are a combination of PID and ratio controls with no feed forward compensation for this disturbance, the NLC significantly reduces variation for these typical effluent flow changes.

Unlike the response in Figure 4, the response to a level setpoint change, shown in Figure 6, does not appear to have the same minimum-time optimal controller appearance. Instead, the controller responds smoothly according to the specified error trajectory. Note that neither of the concentration responses deviate from target to achieve this closed-loop performance in the level.

While this example is somewhat simplistic, it illustrates how the reference system tuning can be specified to achieve different desired responses in various controlled variables.

Polymerization Reactor Example

Finally, the control of a polymerization reactor is used to demonstrate the application of the NLC to a larger-scale industrial example. This polymerization reactor process is actually two reactors in series. Each reactor has independent feed and cooling systems. Catalyst is fed only to the first reactor. The model for each reactor includes mass balances for as many as seven species and multiple phases as well as energy balances around the reactor and cooling systems.

The controlled variables for this application are the polymer melt viscosity and the polymer comonomer incorporation in each reactor. The manipulated variables are setpoints in the distributed control system (DCS) that affect the addition of the comonomer and a transfer agent into the feed to each reactor. The current goal for this application is to control the transition to a desired trajectory.

The model for this 4 input×4 output problem has on the order of 50 state variables and is described by approximately 120 DAEs. In this application, a simple output feedback scheme is used. Lab measurements of the polymer properties are made on product collected at a sampling point well downstream of the second reactor. When new results are received, they are compared to the predicted value at the time the process was sampled. The difference is filtered and used to modify the output predictions until new feedback is received. Figures 7 through 10 show a typical transition response achieved with the NLC controller. Note that the laboratory data is shown on an as-measured basis. To compare the laboratory measurements with the measurement estimate, the laboratory data must be shifted back in time by the time period required to complete the analysis.

The controller is tuned so that the ratio of closed-loop settling time to open-loop settling time is close to one for the fastest transitions and significant improvement is achieved for slower transitions. The control and prediction horizons are equal and approximately two-thirds of the nominal process residence time. The controller is executed every six minutes on a DEC Alpha System 1000 processor running at 266 MHz (circa 1997). Normally, controller execution completes in a two to three minute range, less than half of the six minute control period.

The characteristics of the reference system tuning are most evident in the polymer melt viscosity responses. The computed transfer agent command signal for the second reactor becomes constrained at its maximum value. The polymer comonomer incorporation illustrates the effect of corrupted measurements on the controller performance. Overall, the performance achieved in this particular transition represents a significant improvement over past performance.

Perhaps most importantly, this example demonstrates that nonlinear MPC problems of industrial significance can be solved in real-time on modest computing hardware. This demonstration should not discourage the innovation of techniques designed to permit implementation of nonlinear predictive control to even larger problems. However, it provides a counter example for those who claim that modifications are required to solve industrial nonlinear MPC problem in general.

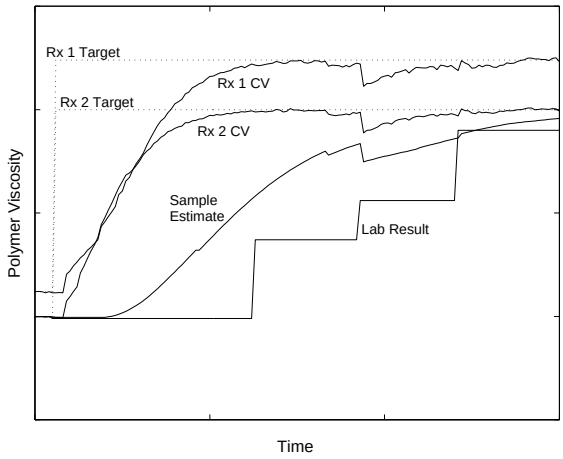


Figure 7: Polymer melt viscosity transition response.

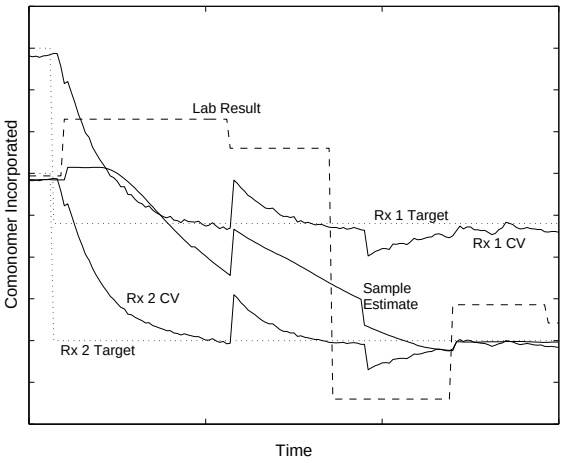


Figure 9: Polymer comonomer incorporation transition response.

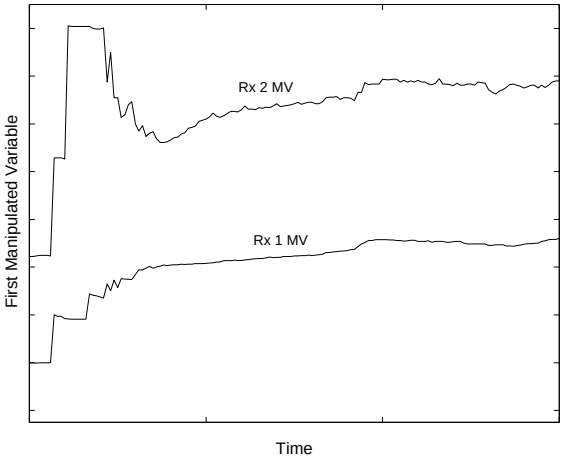


Figure 8: Transfer agent transition command signal.

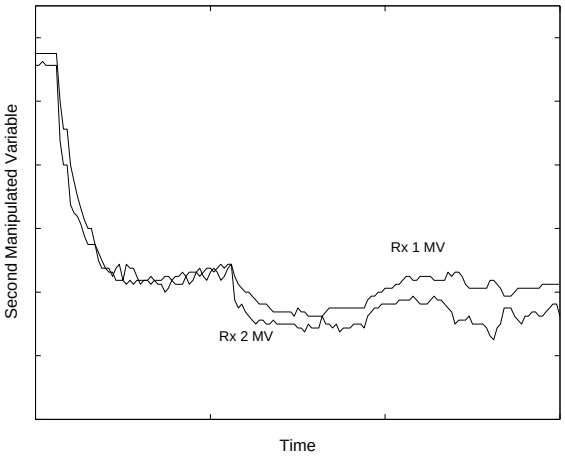


Figure 10: Comonomer transition command signal.

Concluding Remarks

There are no signs that the activity in the area of nonlinear model predictive control are slowing. There is continued academic study of reference system synthesis applied to model predictive control (Kalra, 1997). The presentations at the CPC V conference (Kantor et al., 1996) and the 1998 Workshop on Nonlinear Model Predictive Control at Ascona, Switzerland (Allgower and Zheng, 2000) continue to reinforce the industrial need and academic interests in these areas. A survey of nonlinear model predictive control products by vendors was presented at Ascona (Qin and Badgwell, 2000). Besides the controller described in this paper, four other vendor offerings were available. Of the five, no two were pursuing exactly the same approach to solve this problem.

As stated earlier in this paper, there are very active academic research programs examining nonlinear control, model-predictive control, parameter estimation, and optimization that can be used to further the development of the NLC. Several academic programs have been working on demonstrating nonlinear MPC techniques on industrial-like models (Doyle III and Wisnewski, 2000; Schley et al., 2000; Nagy et al., 2000; Tenny et al., 2001; Findeisen et al., 2001). ExxonMobil Chemical Company actively participates in several university industrial consortia and interacts with several individual faculty members, knowing this participation is important to our continued success in this endeavor.

The NLC controller technology described in this paper is by no means mature. Reminiscent of linear MPC packages of the late 1980's, significant insight and detailed knowledge is required to successfully implement a NLC. However, this control methodology provides a way to pursue opportunities that have been previously beyond the reach of industrial process control engineers.

Acknowledgments

The authors would like to thank ExxonMobil Chemical Company for it's support of this presentation and CPC VI. Additionally, the first author would like to thank Dr. Louis P. Russo for the insightful discussions associated with the development of this presentation.

References

Allgower, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, volume 26 of *Progress in Systems and Control Theory*. Birkhauser, Switzerland (2000).

Arkun, Y. and W. H. Ray, editors, *Fourth International Conference on Chemical Process Control*, Amsterdam. CAHCE Corporation, Elsevier (1991).

Bartusiak, R. D. and R. W. Fontaine, Feedback Method for Controlling Non-Linear Processes, United States Patent, number 5,682,309 (1997).

Bartusiak, R. D., C. Georgakis, and M. J. Reilly, "Nonlinear Feedforward/Feedback Control Structures Designed by Reference System Synthesis," *Chem. Eng. Sci.*, **44**(9), 1837–1851 (1989).

Bequette, B. W., "Nonlinear Control of Chemical Processes: A Review," *Ind. Eng. Chem. Res.*, **30**(7), 1391–1413 (1991).

Biegler, L. T., Efficient Solution of Dynamic Optimization and NMPC Problems, In (Allgower and Zheng, 2000), pages 219–243 (2000).

Bindlish, R. and J. B. Rawlings, "Model Predictive Control of a Prototypical Industrial Polymerization Process," *AIChE J.* (2000). Submitted.

Charos, G. N. and Y. Arkun, "A Decentralized Quadratic Dynamic Matrix Control Algorithm," *J. Proc. Cont.*, pages 75–83 (1993).

Congalidis, J. P., J. R. Richards, and W. H. Ray, "Feedforward and Feedback Control of a Solution Copolymerization Reactor," *AIChE J.*, **35**(6), 891–907 (1989).

Cutler, C. R. and B. L. Ramaker, Dynamic Matrix Control—A Computer Control Algorithm, In *Proc. of Joint Automatic Control Conference*, volume Paper WP5-B, San Francisco, CA (1980).

Debling, J. A., G. C.Han, F. Kuijpers, J. Verburg, J. Zacca, and W. H. Ray, "Dynamic Modeling of Product Grade Transitions for Olefin Polymerization Processes," *AIChE J.*, **40**(3), 506–520 (1994).

Doyle III, F. J. and P. A. Wisnewski, Nonlinear Multi-rate MPC with Large Scale Fundamental Models: Application to a Continuous Kamyr Digester, In (Allgower and Zheng, 2000), pages 419–432 (2000).

Findeisen, R., F. Allgower, M. Diehl, G. Bock, J. Schlooder, and Z. Nagy, Efficient Nonlinear Model Predictive Control, In *Chemical Process Control (CPC) VI* (2001).

Henson, M. A. and D. E. Seborg, *Nonlinear Process Control*. Prentice Hall PTR, Upper Saddle River, New Jersey (1997).

Hillestad, M. and K. S. Andersen, Model Predictive Control for Grade Transitions of a Polypropylene Reactor, In *ESCAPE-4*, Dublin (1994).

Kalra, L., *Reference System Based Model Predictive Control of Nonlinear Processes*, PhD thesis, Lehigh University (1997).

Kantor, J. C., C. E. Garcia, and B. Carnahan, editors, *Fifth International Conference on Chemcial Process Control*. CACHE Corporation, AIChE and CACHE Corporation (1996).

Kravaris, C. and J. C. Kantor, "Geometric Methods for Nonlinear Process Control. 1. Background," *Ind. Eng. Chem. Res.*, **29**(12), 2295–2310 (1990a).

Kravaris, C. and J. C. Kantor, "Geometric Methods for Nonlinear Process Control. 2. Controller Synthesis," *Ind. Eng. Chem. Res.*, **29**(12), 2310–2323 (1990b).

Lee, J. H. and B. Cooley, Recent Advances In Model Predictive Control and Other Related Areas, In (Kantor et al., 1996), pages 201–216 (1996).

Lee, P. L. and G. R. Sullivan, "Generic Model Control (GMC)," *Comput. Chem. Eng.*, **12**(6), 573–580 (1988).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and optimality," *Automatica*, **36**, 789–814 (2000).

McAuley, K. B. and J. F. MacGregor, "Optimal Grade Transitions in a Gas Phase Polyethylene Reactor," *AIChE J.*, **38**(10), 1564–1576 (1992).

McAuley, K. B. and J. F. MacGregor, "Nonlinear Product Property Control in Industrial Gas-Phase Polyethylene Reactors," *AIChE J.*, **39**(5), 855–866 (1993).

McAuley, K. B., J. F. MacGregor, and A. E. Hamielec, "A Kinetic Model for Industrial Gas-Phase Ethylene Copolymerization," *AIChE J.*, **36**(6), 837–850 (1990).

McLellan, P. J., T. J. Harris, and D. W. Bacon, "Error Trajectory Descriptions of Nonlinear Controller Designs," *Chem. Engineering*, **45**(10), 3017–3034 (1988).

Meadows, E. S. and J. B. Rawlings, Model Predictive Control, In (Henson and Seborg, 1997), chapter 5, pages 233–310 (1997).

Morari, M. and N. L. Ricker, *Model Predictive Control Toolbox* (1995).

Muske, K. R. and T. F. Edgar, Nonlinear State Estimation, In (Henson and Seborg, 1997), chapter 6, pages 311–370 (1997).

Nagy, Z., R. Findeisen, M. Diehl, F. Allgower, G. Bock, S. Agachi, J. Schlooder, and D. Leineweber, Real-time Feasiblity of Nonlinear Predictive Control for Large Scale Processes—a Case Study, In *Proc. of American Control Conference*, Chicago, Il (2000).

NOVA NLC, *NOVA Nonlinear Control System* (1997).

Ogunnaike, B. A., "On-line Modelling and Predictive Control of an Industrial Terpolymerization Reactor," *Int. J. Control*, **59**, 711–729 (1994).

Peterson, T., E. Hernandez, Y. Arkun, and F. J.Schork, "A Nonlinear DMC Algorithm and Its Application to a Semi-Batch Polymerization Reactor," *Chem. Eng. Sci.*, **47**(4), 737–753 (1992).

Qin, S. J. and T. A. Badgwell, An Overview of Nonlinear Model Predictive Control Applications, In (Allgower and Zheng, 2000), pages 369–392 (2000).

Rawlings, J. B. and K. R. Muske, "Stability of Contrained Receding Horizon Control," *IEEE Trans. Auto. Cont.*, **AC-38**(10), 1512–1516 (1993).

Ray, W. H., "On the Mathematical Modeling of Polymerization Reactors," *J. Macromol. Sci.-Revs.* (1972).

Ray, W. H., "Polymerization Reactor Control," *IEEE Cont. Sys. Mag.*, **6**(4), 3–8 (1986).

Richalet, J., A. Rault, J. L. Testud, and J. Papon, "Model Predictive Heuristic Control: Application to Industrial Processes," *Automatica*, **14**, 413–428 (1978).

Schley, M., V. Prasad, L. P. Russo, and B. W. Bequette, Nonlinear Model Predictive Control of a Styrene Polymerization Reactor, In (Allgower and Zheng, 2000), pages 403–417 (2000).

Scokaert, P. O. and J. B. Rawlings, "Feasiblity Issues in Model Predictive Control," *AIChE J.*, **45**(8), 1649–1659 (1999).

Sriniwas, G. R., Y. Arkun, and F. J. Schork, "Estimation and Control of an α-Olefin Polymerization Reactor," *J. Proc. Cont.*, **5**(5), 303–313 (1995).

Tenny, M. J., J. B. Rawlings, and R. Bindlish, Feasible real-time nonlinear model predictive control, In *Chemical Process Control (CPC) VI* (2001).

Wright, S. J., Applying New Optimization Algorithms to Model Predictive Control, In (Kantor et al., 1996), pages 147–155 (1996).

Zacca, J. J. and W. H. Ray, "Modelling of the Liquid Phase Polymerization of Olefins in Loop Reactors," *Chem. Eng. Sci.*, **48**(22), 3743–3765 (1993).

Emerging Technologies for Enterprise Optimization in the Process Industries

Rudolf Kulhavý*
Honeywell Laboratories
Pod vodárenskou věží
CZ-182 08 Prague, Czech Republic

Joseph Lu†
Honeywell Hi-Spec Solutions
16404 N. Black Canyon Highway
Phoenix, AZ 85023 U.S.A.

Tariq Samad‡
Honeywell Laboratories
3660 Technology Drive
Minneapolis, MN 55418 U.S.A.

Abstract

We discuss three emerging technologies for process enterprise optimization, using a different application domain as an example in each case. The first topic is cross-functional integration on a model predictive control (MPC) foundation, in which a coordination layer is added to dynamically integrate unit-level MPCs. Enterprise optimization for oil refining is used to illustrate the concept. We next discuss data-centric forecasting and optimization, providing some details for how high-dimensional problems can be addressed and outlining an application to a district heating network. The final topic is adaptive software agents and their use for developing bottom-up models of complex systems. With learning and adaptation algorithms, agents can generate optimized decision and control strategies. A tool for the deregulated electric power industry is described. We conclude by emphasizing the importance of seeking multiple approaches to complex problems, leveraging existing foundations and advances in information technology, and a multidisciplinary perspective.

Keywords

Model predictive control, Cross-functional integration, Oil refining, Data-centric modeling, District heating, Adaptive agents, Electric power

Introduction

The history of advances in control is the history of progress in the level of automation. From single-loop regulation to multivariable control to unit-level optimization, we have seen step changes in the efficiency, throughput, and autonomy of operation of process plants. The next step in this progression is the optimization of the enterprise—a process plant, a multifacility business, or even a cross-corporate industry sector.

Enterprise optimization is the object of significant research today. Given the diversity of potential applications and the relative newness of the topic, it is not surprising that no one approach has been accepted as a universal solution. Although it is conceivable that future research will identify a cross-sector solution, at this point it appears that multiple approaches are likely to be necessary to best cover the full spectrum of potential applications.

In this paper, we discuss three emerging technologies for process enterprise optimization: cross-functional integration on a model predictive control (MPC) foundation, data-centric modeling and optimization, and adaptive agents. The discussions are generic, but different process applications are used in each case for illustration: oil refining, district heating, and electric power.

Enterprise Optimization on an MPC Foundation[1]

MPC has proven to be the most viable multivariable control solution in the continuous process industries and it is becoming increasingly popular in the semibatch and batch industries as well. Most industrial MPC products also include a proprietary economic optimization algorithm that is essential for driving the process to deliver more profit. In terms of economic benefit, MPC is one of the most significant enabling technologies for the process industries in recent years, with reported payback times between 3 and 12 months (Hall and Verne, 1993, 1995; Smith, 1993; Sheehan and Reid, 1997; Verne and Escarcega, 1998).

Cross-Functional Integration as a New Trend

Traditionally, most MPC applications are used for stabilizing operations, reducing variability, improving product qualities, and optimizing unit production. In most cases, a divide-and-conquer approach to a complex plantwide production problem is adopted. In this approach, a large plant is divided into many process units, and MPC is then applied on appropriate units. The divide-and-conquer approach reduces the complexity of the plantwide problem, but each application can reach only its local optimum at best. In a complex plant, the composition of local optima can be significantly less than the potential global optimum. For example, the estimated latent benefit for a typical refinery is 2-10 times more than what the combination of MPC controllers can capture (Bodington, 1995).

One possible approach to plantwide control is employing a single controller that is responsible for the whole plant; however, this option is infeasible. To note just the most obvious issue, commissioning or maintenance would require the whole plant to be brought offline. An alternative and practical approach for delivering global optimization benefit is to add a coordination layer on top of all the MPC applications to achieve the global optimum. The coordination layer usually covers multiple functions of the plant, such as operations, production

*rudolf.kulhavy@honeywell.com
†joseph.lu@honeywell.com
‡tariq.samad@honeywell.com
[1]This section is adapted from Lu (1998).

scheduling and production planning.

Complexity of MPC Coordination. With the divide-and-conquer approach, transfer price is traditionally used for measuring the merit of an advanced control application from a plantwide view and over an appropriate time period. The transfer price of a product (or a feed) is an artificial price assigned to determine the benefit contribution of a process unit to the overall plant under an assumed level of global coordination. A common assumption in calculating transfer price is that the product produced in a unit will travel through the designed paths (or designed processes) with the designed fractions to reach the final designated markets. This assumption is not always valid due to a lack of dynamic coordination among the units.

This phenomenon is referred to as benefit erosion[2] (e.g., see Bodington, 1995) and is typically alleviated by manual coordination between different sections of the production processes. For example, after an advanced control application is implemented on a refinery's crude unit, the yield of the most valuable component often increases, whereas the yield of the least valuable component decreases. The scheduling group would detect the component inventory imbalances (which could cause tank level problems if not corrected in time) rippling through various parts of the refinery, and it would then coordinate the affected parts of the refinery to "digest" the imbalances. With feedback from scheduling and operations groups, the planning group would update its yield models to reflect the yield improvement and rerun the plantwide optimization to generate a new production plan and schedule.

The fundamental cause of benefit erosion, however, is a lack of global coordination or optimization. Generally, the more complex the production scheme, the greater the problem. Therefore, a complex plant, such as a refinery or a chemicals plant, presents a higher benefit potential for cross-functional integration. If a new dynamic or steady-state bottleneck is encountered, or if an infeasible production schedule results, the scheduling group and the planning group would have to work together with the operations group to devise a new solution. The final solution may take a few adjustments or iterations. For complex cases, this process can take a few weeks or even months.

Only when all operations and activities in the refinery are coordinated together will benefit erosion be minimized. Although the situation described in this refinery example may sound primitive, it is still one of the better cases. In reality, different parts of a refinery usually use different tools with different models on different platforms. Engineers and operators in different units look at the same problem with very different time horizons. All of these factors complicate plantwide coordination in practice.

Although the cross-functional integration approach requires existing infrastructure to be revamped, including the DCS hardware and supporting software systems, this is increasingly less of an issue. Hardware and infrastructure costs have dropped significantly in recent years, particularly when viewed as a percentage of the total advanced control project budget. Support and organizational "psychology" sometimes hinder progress, but the situation has improved as more and more refineries and chemical plants realize the benefits of advanced control.

Long- and Short-Term Goals. Cross-functional integration takes a holistic approach to the plantwide problem. The integration encompasses a large number of process units and operating activities. Practical considerations suggest that it proceed bottom up, integrating one layer at a time until the level of enterprise optimization is finally reached, as depicted in Figure 1.

In Figure 1, the various layers in the pyramid describe the plantwide automation solution structure and the decision-making hierarchy. Around the pyramid structure is the circle of supply chain, production planning and scheduling, process control, global optimization, and product distribution. As more layers are integrated into the cross-functional optimization, the integrated system will perform more tasks and make more decisions that are made heuristically and manually today. Long envisioned by many industrial researchers, such as Prett and Garcìa (1988), this concept is now being further developed with design details of hardware systems, software structures, network interfaces, and application structures.

As a long-term goal, enterprise optimization integrates all activities in the whole business process, from the supply chain to production, and further to the distribution channel. In addition, risk management can also be included as a key technical differentiator. Parallel to the structural development for such a general-purpose complex system, some proof-of-concept projects have been piloted, and experiments have been conducted on several different structures (Bain et al., 1993; del Toro, 1991; Watano et al., 1993). The multiple-layer structure described in Figure 1 is believed to provide the flexibility needed for implementing and operating such a system. Moreover, each layer can be built at an appropriate level of abstraction and over a suitable time horizon. The lower layers capture more detailed information of various local process units over a shorter time horizon, whereas the higher layers capture more of the business essence of the plant over a longer horizon.

As a short-term goal, cross-functional integration could include, in a refinery, for example, raw material allocation, inventory management, production management, unit process control, real-time optimization,

[2]The benefit loss when the assumed level of global coordination is not reached.

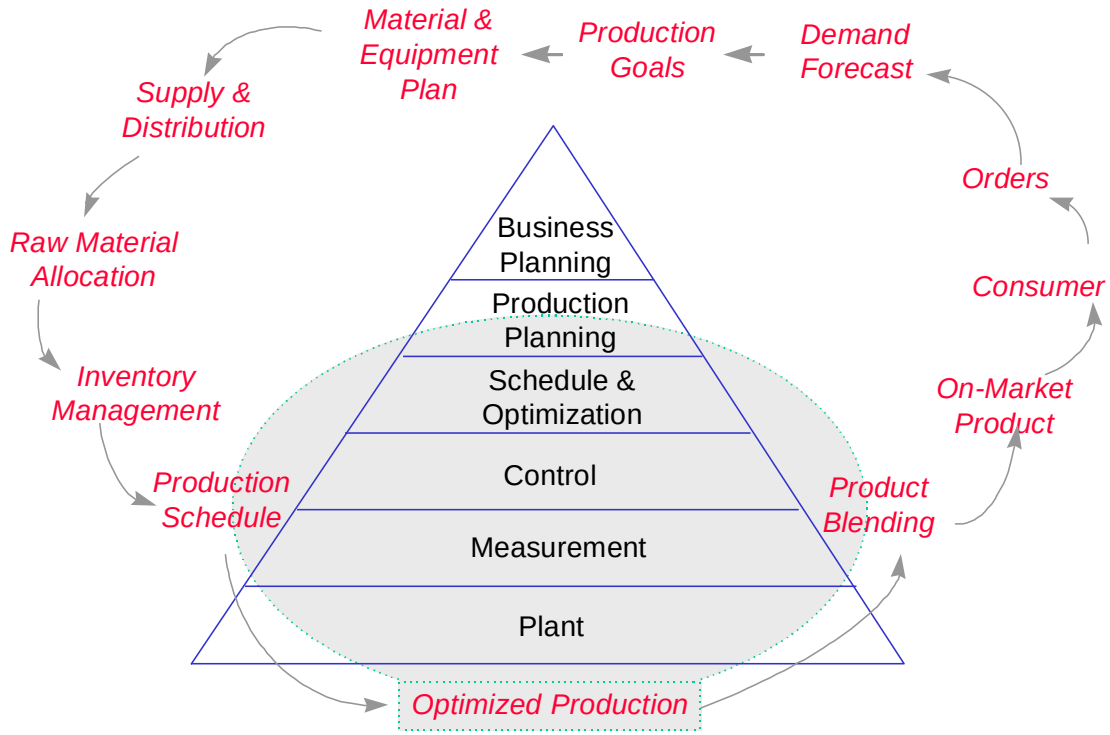


Figure 1: Enterprise optimization—refinery example.

product blending, production planning, and production scheduling. The instrumentation and real-time database are in the bottom layer. The regulatory control comes in the second layer. Following that are the MPC layer, the global/multiunit optimization layer, the production scheduling layer, and, last, the top production planning layer.

More relevant to the control community, this integration will have a large impact on almost all control, optimization, and scheduling technologies currently employed in the process industries. Advanced process control, primarily MPC and real-time optimization (RTO), will certainly be affected, particularly in terms of connectability, responsiveness, and compatibility.

MPC Considerations

This integration requires us to reassess the MPC and RTO designs in terms of their online connectability and their dynamic integration. There is a need for codesigning MPC and RTO, or at least designing one taking into account that it will be working with the other dynamically. Furthermore, from a control perspective, global coordination requires MPC applications to perform over a much wider operating region and more responsively. This poses new challenges to MPC technology, three of which are discussed below.

Nonlinear MPC. The majority of chemical process dynamics is, loosely speaking, slightly nonlinear and can be well modeled with a linearized model around a given operating point. With the application of CV/MV linearizing transformations, linear MPC can be extended to effectively handle a variety of simple nonlinear dynamics. Practical examples include pH control, valve position control, high-purity quality control, and differential pressure control. However, under cross-functional integration, a process will be required to promptly move its operating point over a significant span, or else to operate under very different conditions. Linear MPC may not be adequate in such cases.

Two examples can serve to illustrate. First, as the operating point of the process migrates, the dynamics of some processes may change dramatically, even to the extent that the gain will change sign. For example, some yields of the catalytic cracking unit in a refinery can change their signs when operated between undercracking and overcracking regions. Second, transition control presents a unique type of nonlinear control problem. During a transition, the process operating point typically moves much faster than usual, and the process typically responds more nonlinearly than usual. Examples include crude switch in a refinery and grade transition in many series production processes such as polymer and paper.

Both of these problems can be seen as instances of multizone control, where the process needs to be regulated

most of the time in each of the zones and to be occasionally maneuvered from one zone to another according to three performance criteria. The first criterion is for normal continuous operation in the first zone. The second criterion is for normal operation in the second zone. The third criterion is for special requirements during the transition or migration.

Although solving the multizone control problem alone has merit, solving it while coordinating with other parts of the plant will potentially provide much greater benefit. As enterprise optimization further advances, there is an increasing need for a nonlinear MPC tool that can solve multizone control problems. Likewise, there is an increasing need for MPC formulations, linear and nonlinear, that take into account the requirements of cross-functional optimization.

Model Structure. The MPC model structure is preferred to have a linear backbone or linear substructure. The linear model can be developed experimentally, and the nonlinear portion of the dynamics can be added as the need arises. This preference stems from the fact that, in most cases, one does not know a priori if a nonlinear model is necessary.

The model structure should also be scalable. Adding controlled variables (CVs) and manipulated variables (MVs) should not require the existing model to be reidentified or regenerated. The user should easily be able to eliminate any input-output pair of the model. The preferred solution should not require the system to be square or all the MVs to be available all the time.

The ability to share model information in the cross-functional integration scheme is important. The model obtained in an advanced control project is often the single most costly item in the implementation. Sharing model information with the other layers of cross-functional integration can show substantial monetary savings. This model sharing can be a two-way process: bottom up and top down.

The preferred model-sharing scheme is bottom up. It is much easier for engineers to find and correct modeling errors on a unit-by-unit basis in the control implementation than on a multiunit or plantwide basis under cross-functional integration. Bottom-up model sharing enhances usability, as engineers can verify their models as they commission the MPC applications unit by unit.

The solution technique needs to be algorithmically robust. If multiple solutions exist, it is preferable to stick to one branch of solutions throughout, or better yet, stick to the branch of solutions that requires only the minimum amount of control effort by some criterion.

Coordination Port. An essential requirement for cross-functional integration is an independent port for coordination. Simply sending a steady-state target as a set-point to an MPC controller is inadequate. The dynamic response required for coordination is often very different from that for set-point change, and the performance criterion can be completely different from set-point tracking or disturbance rejection. With a single port, it is difficult, if not impossible, to always satisfy both requirements.

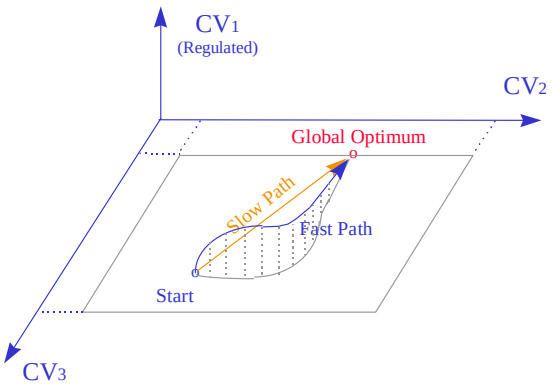


Figure 2: An example of different dynamic optimization paths.

A good coordination port implementation in MPC is one of the essential links in instituting cross-functional optimization. The nature of the problem is illustrated in Figure 2 as a simplified example of three CVs and two MVs. The first CV has a set-point, and the other two both have high and low bounds. (The MVs are not shown in the diagram.) The coordination or optimization target is solved by a global optimizer. The controller needs to drive the system to the target in an independent response for coordination. This is similar to designing a two-degree-of-freedom controller for the coordination port, except that the input direction and the response requirement may vary from one transition to another.

Furthermore, when the coordination port problem is not fully specified dynamically, as in many practical problems that we encounter, multiple solution paths to the destination exist, as depicted in Figure 2. Treating this as a traditional control problem by specifying a desired response trajectory is not always suitable for two reasons. First, the CV error violation is usually much less important in the transition than in normal operation. Second, except in trivial cases, one rarely knows a priori which transitional path would be financially optimal yet dynamically feasible.

An alternative solution is to solve for all equal solution paths in terms of the performance criteria (including financial terms) and choose the one that requires the minimum MV movement by some criterion. See Lu and Escarcega (1997) for one such solution to the coordination port implementation.

Summary

If MPC has been the main theme of the '80s and '90s in the process control industries, cross-functional integration and enterprise optimization will be the main theme of the next two decades. As we advance toward a higher level of computer-integrated manufacturing, the problem definition and the scope of classical model predictive control will be expanded. The cross-functional integration approach requires the dynamic coordination of multiple MPCs, not just in terms of steady-state operation, as in many steady-state RTO and composite LP approaches. One practical way of achieving dynamic coordination is by designing a coordination port in the MPC control strategies.

Although truly enterprise-scale applications along these lines are yet to be implemented, our initial projects in cross-functional integration are yielding exciting results. A cross-functional RMPCT[3] integration is discussed by Verne and Escarcega (1998), who report significant benefits. More recently, Nath et al. (1999) describe an application of Honeywell's Profit® Optimizer technology to an ethylene process at Petromont's Varennes olefins plant. In the Profit Optimizer approach, controllers are not operating in tandem, and the dynamic predictions of bridged disturbance variables are used by individual RMPCTs to ensure dynamic coordination/compensation among them. The controllers thereby work together to protect mutual CV constraints. The optimizer coordinates 10 controllers and other areas, covering the entire plant except for the debutanizer. During the acceptance test for the system, a sustained increase of over 10% in average production was achieved. This production level surpassed the previous all-time record for the plant by over 3.7% and was well above the expectation of a 2.7% increase.

Our experience to date in enterprisewide advanced control coordination has largely been limited to refining and petrochemical plants. For chemical plants, one of the key additional requirements is the integration of scheduling, product switchover, and other discrete-event aspects of plant operation within the formulation. The research of Morari and colleagues (Bemporad and Morari, 1999) on the optimization of hybrid dynamical systems is especially promising in this context.

Exploiting the Data-Centric Enterprise

From a technology that leverages the state-of-the-art in empirical-model-based control, we next turn to a more radical alternative to large-scale optimization. The key idea is that, where first principles or identified models cannot usefully be developed, we can consider historical data as a substitute. In other words, "the data is the model." This mantra was not especially useful even a few years ago, but now, with modern storage media available at affordable prices and computer performance increasing at a steady pace, entire process and business histories can be stored in single repositories and used online for enhanced forecasting, decision making, and optimization. This makes it possible to implement a data-centric version of intelligent behavior:

- Focus on what matters—by building a local model, on demand and for the immediate purpose.
- Learn from your errors—by consulting all relevant data in your repository.
- Improve best practices—by adapting proven strategies first.

The data-centric paradigm combines database queries for selecting data relevant to the case, fitting the data retrieved with a model of appropriate structure, and using the resulting model for forecasting or decision making.

When the size of the data repository becomes large enough, the architecture of the database (the data model) becomes crucial. To guarantee sufficiently fast retrieval of historical data, the queries need to be run against a specifically designed data warehouse rather than the operational database. Once the relevant data are retrieved, nonparametric statistical methods can be applied to build a local model fitting the data. Data-centric modeling can thus be seen as a synergistic merger of data warehousing and nonparametric statistics.

Data-centric models can form the basis for enterprise optimization. For example, the central problem of business optimization is matching demand with supply in situations when the supply or demand, or both, are uncertain. Suppose that a reward due to supply is defined as a response variable dependent on the decisions made and other conditions. Supply optimization then amounts to searching for maximum reward over the response surface.

In contrast to traditional response surface methods, data-centric optimization does not assume a global model of the response surface; rather it constructs a local model on demand—for each decision tested in optimization. Since the response is estimated through locally weighted regression (as discussed below), noise is automatically filtered. The uncertainty of the response estimate can also be respected in the optimization by replacing the estimated reward with the expected value of a utility function of the reward. By properly shaping the utility function, one can make decision making either risk-prone or risk-averse. As for the optimization algorithm itself, data-centric models lend themselves to any stochastic optimization method, including simulated annealing, genetic algorithms, and tabu search. (Response surface optimization is only one example of a data-centric

[3]RMPCT (Robust Multivariable Predictive Control Technology) is a Honeywell MPC product.

optimization scheme. Other schemes, such as optimization over multiunit systems and distributed optimization, are currently being investigated.)

We next contrast data-centric modeling with the more established global and local modeling approaches and provide some technical details. An overview of a reference application concludes this section.

Empirical Modeling

When exploring huge data sets, one must choose between trying to fit the complete behavior of the data and limiting model development to partial target-oriented descriptions.

The *global* approach generally calls for estimation of comprehensive models such as neural networks or other nonlinear parametric models (Sjöberg et al., 1995). The major advantage of global modeling is in splitting the model-building and model-exploitation phases. Once a model is fit to the data, model look-up is very fast. Global models also provide powerful data compression. After a model is built, the training data are not needed any further. The drawback is that the time necessary for estimation of unknown model parameters can be very long for huge data sets. Also, global models are increasingly sensitive to changes in the data behavior and may become obsolete unless they are periodically retuned.

The *local* approach makes use of the fact that often it is sufficient to limit the fit of the process behavior to a neighborhood of the current working point. Traditionally, local modeling has been identified with recent-data fitting. Linear regression (Box and Jenkins, 1970) and Kalman filtering (Kalman, 1960) with simple recursive formulae available for parameter/state estimation have become extremely popular tools. Their simplicity comes at some cost, however. Adaptation of local-in-time models is driven solely by the prediction error. When a previously encountered situation is encountered again, learning starts from scratch. Further, when a process is cycling through multiple operating modes, adaptation deteriorates model accuracy.

The data-centric approach is an alternative to both these prevailing paradigms. It extends *recent-data* fitting to *relevant-data* fitting (see Figure 3). This "shift of paradigm" combines the advantages of global and local modeling. Namely, it provides

- a global description of the data behavior,
- through a collection of simple local models built on demand,
- using all data relevant to the case.

The price we pay for this powerful mix is that all data need to be at our disposal at all times (i.e., no single compact model is returned as a result of modeling). The data-centric model can be built out of the original data stored in the database without any data compression.

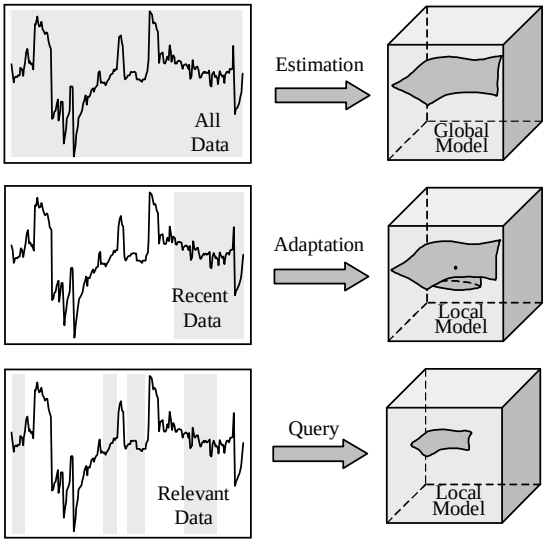


Figure 3: Global, local-in-time, and local-in-data modeling.

When applied for forecasting at the enterprise level, the data needs to be aggregated properly. The forecasting literature (see, e.g., West and Harrison, 1989) shows that forecasting from aggregated data yields more robust and more precise forecasts. The design of proper aggregation or more sophisticated preprocessing of data thus becomes a crucial part of data-centric modeling. Poor preprocessing strategies will need to be corrected before data-centric models can be effective.

Data-Centric Modeling with Locally Weighted Regression

To be more specific, we describe one possible implementation of data-centric modeling, namely, locally weighted regression with local variable bandwidth. The algorithm is by no means the only option and should be understood only as an illustration of the general idea.

It is difficult to trace the originator of local modeling. The concept appeared independently in various fields under names such as locally weighted smoothing (Cleveland, 1979), nonparametric regression (Härdle, 1990), local learning (Bottou and Vapnik, 1992), memory-based learning (Schaal and Atkeson, 1994), instance-based learning (Deng and Moore, 1994), and just-in-time estimation (Cybenko, 1996).

General Regression. Assume that a single dependent variable or response y depends on n independent or regressor variables $\varphi_1, \varphi_2, \ldots, \varphi_n$ through an unknown static nonlinear function $f(\cdot)$ with precision up to an unpredictable or stochastic component e

$$y = f(\cdot) + e.$$

The objective is to estimate response y_0 for any particular regressor φ_0.

Linearization in Parameters. Rather than trying to fit the above global model to all the data available, data-centric modeling suggests the application of a simpler model to data points $(y_1, \varphi_1), \ldots, (y_N, \varphi_N)$ selected so that $\varphi_1, \varphi_2, \ldots, \varphi_N$ are close to a specified regressor φ.

A typical example is a model linearized around a given column vector φ:

$$y = \theta'(\varphi)\varphi + e,$$

where θ denotes a column vector of regression coefficients and θ' its transposition.

Curse of Dimensionality. A word of caution applies here. As the dimension of φ increases, the data points become extremely sparsely distributed in the corresponding data space. To locate enough historical data points within a neighborhood of the query regressor φ_0, the neighborhood can become so large that the actual data behavior cannot be explained through a simplified model.

Consider, for instance, a 10-dimensional data cube built over 10-bit data; it contains $2^{100} \approx 10^{30}$ bins! Even a trillion (10^{12}) data points occupy an almost zero fraction of the data cube bins. Even with 5-bit (drastically aggregated) data, 99.9% of bins are still empty.

We have an obvious contradiction here. To justify a simple model, we must apply it in a relatively small neighborhood of the query point. To retrieve enough data points for a reliable statistical estimate, we must search within a large enough neighborhood. One must ask: Can data-centric modeling work at all?

Coping with Dimensionality. Luckily, real data behavior is rarely that extreme. First, the data is usually concentrated in several regions around typical operating conditions, which violates the assumption of uniform distribution assumed in mathematical paradoxes. Second, the regressor φ often lives in a subspace of lower dimension. That is, $\varphi = \varphi(x)$ where x is a vector of dimension smaller than the dimension of φ. Suppose, for instance, that y, x_1, and x_2 denote process yield, pressure, and temperature, respectively, and the model is a polynomial fit with

$$\varphi_i(x_1, x_2) = x_1^{m(i)} x_2^{n(i)}.$$

The dimension of regressor φ can easily be much larger than the dimension of x, but it is the dimension of x that matters here; it defines the dimension of a data space within which we search for "similar" data points.

Under the assumption $\varphi = \varphi(x)$, the model is linearized around the vector x:

$$y = \theta'(x)\varphi + e$$

and applied to data points $(y_1, x_1), \ldots, (y_N, x_N)$ selected so that $\|x_k - x_0\| \le d$ for $k = 1, 2, \ldots, N$, where $\|\Delta\|$ is an Euclidean norm of vector Δ, x_0 is a query vector, and d is a properly chosen upper bound on the distance.

Weighted Least Squares. The simplest statistical scheme for estimating the unknown parameters θ is based on minimizing the weighted sum of prediction errors squared:

$$\min_\theta \sum_{k=1}^{N} K(\|x_k - x_0\|)(y_k - \theta^T \varphi_k).$$

Each data point (y_k, x_k), $k = 1, 2, \ldots, N$ is assigned a weight inversely proportional to the Euclidean distance of x_k from x_0 through a kernel function $K(\cdot)$. Typical examples of kernel functions are the Gaussian kernel $K(\Delta) = \exp(-\Delta^2)$ or the Epanechnikov kernel $K(\Delta) = \max(1 - \Delta^2, 0)$.

The kernel function assigns zero or practically zero weight to data points (y, x) that appear too far from the query vector x_0. We can use this fact to accelerate database query by searching only for data points within the neighborhood of x_0 defined by $K(\|x_k - x_0\|) \le \varepsilon$, where ε is close to zero.

Performance Tuning. The performance of the above algorithm crucially depends on the definition of the Euclidean norm $\|\Delta\|$. In general, the norm is shaped by the "bandwidth" matrix S:

$$\|\Delta\|^2 = \Delta' S^{-1} \Delta.$$

Through S, it is possible to emphasize or suppress the importance of deviations from the query point x_0 in selected directions in the space of x-values. The matrix S depends on x_0 and can be determined, for example, by the nearest neighbor or cross-validation method (Hastie and Tibshirani, 1990).

Bayesian Prediction. The precision of prediction is an important issue with any statistical method, but in data-centric modeling it is even more pressing due to the relatively small number of data points used for local modeling. To quantify consistently the prediction uncertainty, one can adopt the Bayesian approach to estimation and prediction (Peterka, 1981). Compared with the least-squares method chosen above for simplicity, the Bayesian method calculates a complete probability distribution of the estimated parameters. The distribution is then used for calculating a probability distribution of the predicted variable(s). From the predictive distribution, any derived statistic such as variance or confidence interval can be computed. Due to the relative simplicity of local models, the Bayesian calculations, notorious for their computational complexity in many tasks, can be performed here analytically, without any approximation (for more details, see Kulhavý and Ivanova, 1999).

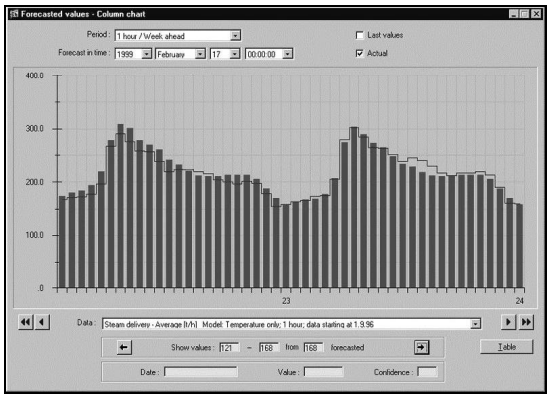


Figure 4: A comparison of predicted (bars) and actual (line) steam.

Reference Application

The first operating application of data-centric forecasting and decision making is an operator advisory system in a large combined heat and power company in the Czech Republic. The company supplies heat to a district heating network and electricity to a power grid. The whole system is composed of five generation plants, a steam pipeline network totaling 60 miles, and five primary, partially interconnected hot-water pipeline networks of a total length of 46 miles.

Data-centric forecasting is being applied to predict the total steam, heat, and electricity demand, heat demand in individual hot-water pipelines, and total gas consumption. The forecasts are performed in three horizons—15-minute average one day ahead, 1-hour average one week ahead, and 1-day average one month ahead (Figure 4). Altogether, this requires the computation of 1,632 forecasts every 15 minutes, 2,856 forecasts every hour, and 1,581 forecasts every midnight. The use of highly optimized data marts makes it possible to perform all the computations, including thousands of database queries, while still leaving enough time for other analytic applications.

Data-centric decision making is applied to optimization of set-points (supply temperature and pressure) on hot-water pipeline networks and to economic load allocation over 16 boilers. Insufficient instrumentation of hot-water pipeline networks is solved by complementing data-centric optimization with a simple deterministic model-based procedure for evaluating the objective function. The lack of data is thus compensated for with prior knowledge, combining the underlying physical laws and some simplifying assumptions. The solution can be adapted to a wide range of heat distribution networks with different levels of system knowledge and process monitoring. Data-centric modeling takes into account the outdoor temperature, current (or planned) electricity production, time of day, day of week, whether it is a working day/holiday, and the recency of data. Electricity production is considered to reflect the effect of operators' decisions in addition to external conditions.

Incomplete measurements on the boilers do not allow for online estimation of combustion efficiency for each boiler separately. Data-centric optimization is configured so as to exploit only available information. From the total fuel consumptions of generation plants and the total heat generated, the overall efficiency of the current boiler configuration is calculated. Data-centric optimization then searches in the process history and suggests possible improvements. For frequent situations and configurations, important for the company, enough points in the history are retrieved and reliable results are obtained. Different configurations are tested and evaluated while taking into account the costs of reconfiguration. One of the most valuable features of data-centric optimization is that it automatically adapts to changing operating conditions (e.g., variations in fuel calorific value, changes in boiler parameters, and aging of equipment).

Summary

A criticism frequently voiced about the data-centric approach is that it is incapable of modeling plant behavior in previously unvisited operational regimes and of handling changes in the plant, such as minor equipment failures, catalyst aging, and general wear and tear. In general, as with any statistical model, the quality of data-centric models depends crucially on the availability and quality of historical data. The lack of data can be compensated only by prior knowledge. One option is to populate the database with both actual and virtual data, the latter coming from a simulation model or domain expert (Kulhavý and Ivanova, 1999).

Thus, the data-centric approach itself suggests a solution to the problem of integrating data and knowledge. A historical database can be merged with a database populated with examples generated based on heuristics or conventional models. Different weightings can be associated with records that depend on their source. Thus historical data can be preferred for operational regions that are well represented in the recent history of plant operation, whereas "synthetic" data can be preferred for contexts for which process history provides few associated samples. In effect, a synthesis of multiple types of information sources can be achieved using the database as a common foundation.

To sum up, data-centric models can be applied to a range of problems in the process industries, subject only to an ability to satisfy the database-intensive search requirements and the absence of a single compact model that would permit closed-form analytic solutions. The former will cease to be a real constraint in a few years due to continuing progress in database and computer technol-

ogy. The latter issue is fundamental—with on-demand modeling, all decision making becomes iterative. For this reason, the data-centric model should not be considered a replacement for the classical results of decision and control theory, but as a tool that can extend the reach of automation and control to processes that have not been amenable to analytic methods.

Optimization with Adaptive Agents

The term *agent* is used in multiple senses in the computational intelligence and related communities. We use it here to mean software objects that represent problem domain elements. Agents can, for example, represent units and equipment in a chemical plant, parts of an electricity transmission and distribution network, or suppliers and consumers in supply chains (García-Flores et al., 2000). An agent must be capable of modeling the input/output behavior of a system at a level of fidelity appropriate for problems of interest. Couplings between systems, whether material or energy flows, sensing and control data, or financial transactions, can be captured through interagent communication mechanisms. In some sense, agent-based systems can be seen as extensions of object orientation, although the extensions are substantive enough that the analogy can be misleading.

One common use of agents is to develop bottom-up, componentwise models of complex systems. With subsystem behaviors captured with associated agents, the overall multiagent system can constitute a useful model of an enterprise such as a process plant or even an industry structure. Given some initial conditions and input streams, the evolution of the computational model can track the evolution of the physical system. Often a quantitative match will be neither expected nor obtained, but if the essential elements driving the dynamics of the domain are captured, even qualitative trends can provide insight into and guidance for system operation.

The increasing interest in agent systems can be attributed in part to advances in component software technology. Provided that common interface specifications are defined, agents can be developed in different programming languages and can be executing on different processes or computers. Agents can vary from very simple to extremely sophisticated, and heterogeneous agents may work together in a single application. "Plug-and-play" protocols for agent construction encourage code reuse and modularity while enabling agent developers to work in a variety of programming languages. Generic agent toolkits are now available (e.g., swarm `www.swarm.org`, Lost Wax `www.lostwax.com`, and Zeus `http://193.113.209.147/projects/agents/zeus/`) that can facilitate the design of agent applications for diverse problems. Languages for interagent communication for broad-based applications have also been developed. One such language, KQML, has been fairly widely adopted (Finin et al., 1994).

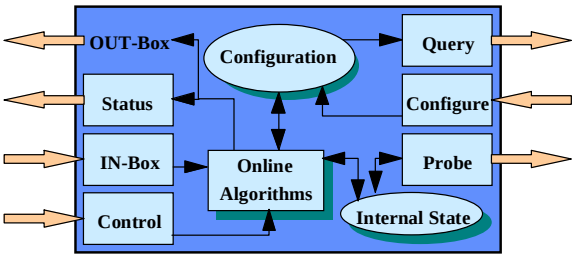


Figure 5: Abstract agent architecture.

Decision Making and Adaptation in Agents

Figure 5 shows an example abstract agent architecture, identifying some of the functions that are often incorporated. The In and Out boxes are the ports through which the agent exchanges messages with other agents; Status and Control are interfaces for the agent simulation framework; and the Query, Configure, and Probe features are used for initialization, monitoring, debugging, and so on. An agent can maintain considerable internal state information, which can be used by algorithms that implement its decision logic (its input/output behavior). Other online algorithms can be used to adapt the decision logic to improve some performance criteria (which may be agent-centric or global). The adaptation mechanism is often internal to an agent, but it need not necessarily be so.

The adaptation mechanism may be based on genetic algorithms (Mitchell, 1996), genetic programming (Koza, 1992), evolutionary computing (Fogel, 1995), statistical techniques, or artificial neural networks. These algorithms act on structures within an agent, but the adaptation is often based on information obtained from other agents. For example, a low-performing agent may change its program by incorporating elements of neighboring, better-performing agents.

The choice of learning algorithm interacts strongly with how the agent represents its decision-making knowledge and the kind of feedback available. For example, LISP programs may lend themselves to genetic programming. In a supervised learning situation, neural networks may employ an algorithm such as back-propagation or Levenburg-Marquardt minimization. However, in many agent applications, only weaker information is available (e.g., a final score, forcing agents to rely on reinforcement learning algorithms). The learner must solve temporal and spatial credit assignment problems—determining what aspect of its sequence of decisions led to the final (or intermediate) score and what part of its internal representation is responsible. Strategies such as temporal differencing and Q-learning have been proposed for this (Watkins and Dayan, 1992).

Multiple learning mechanisms can be incorporated

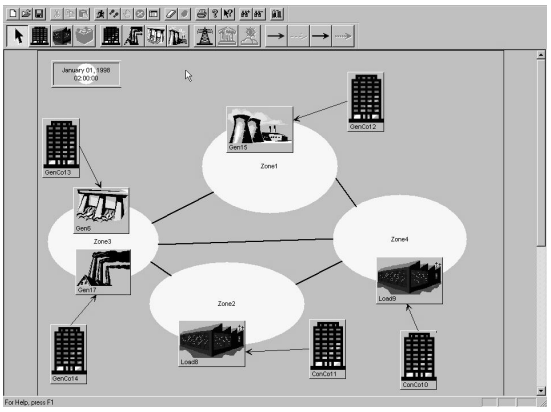


Figure 6: SEPIA interface showing a four-zone, three-generator, two-load scenario.

into the same agent environment, or even the same agent. "Classifier systems" frequently incorporate a genetic algorithm to evolve new classifiers over time. The use of multiple adaptation mechanisms can offer unsuspected advantages. One well-known example is the Baldwin effect, in which learning interacts with evolution even though the learned information is not inherited (Hinton and Nowlan, 1987).

Example: A Simulator for Electric Power Industry Agents

One industry where the integration of business and physical realms has recently taken on a new importance is electric power. Deregulation and competition in the power industries in several countries over the last few years have resulted in new business structures. At the same time, generation and transmission facilities impose hard physical constraints on the power system. For a utility to attempt to maximize its profit while ensuring that its power delivery commitments can be accommodated by the transmission system—which is simultaneously being used by many other utilities and power generators—economics and electricity must be jointly analyzed.

A prototype modeling and optimization tool recently developed under the sponsorship of the Electric Power Research Institute provides an illustration. The tool, named SEPIA (for Simulator for Electric Power Industry Agents), integrates classical power flow models, a bilateral power exchange market, and agents that represent both physical entities (power plants) and business entities (generating companies). Through a full-featured GUI, users can define, configure, and interconnect agents to set up specific industry-relevant scenarios (see Figure 6). A scenario can be simulated with adaptation capabilities enabled in selected agents as desired. For example, a generation company agent can search a space of pricing structures to optimize its profits subject to its generation constraints, the transmission capabilities of the system, and the simultaneous optimization of individualized criteria by other agents that may bear a cooperative or competitive relationship to it. Two reusable and configurable learning/ adaptation mechanisms have been incorporated within SEPIA: Q-learning and a genetic classifier system.

SEPIA models a power system as a set of zones interconnected by tie-lines. A transmission operator agent conducts security analysis and available transfer capacity calculation, including first contingency checks for all proposed transactions. A dc power flow algorithm and a simplified ac algorithm are included for this purpose.

Tools such as SEPIA have several potential uses:

- To identify optimized (although not necessarily optimal) operational parameters (such as pricing structures or consumption profiles in the power industry application).
- To determine whether an industry or business structure is stable or meets other criteria (such as fairness of access).
- To evaluate the potential benefits of new technology (e.g., superconducting cables and high-power transmission switching devices).
- To generally help decision makers gain insight into the operation and evolution of a complex system that may not be amenable to more conventional analysis techniques.

Further details on SEPIA are available in Harp et al. (2000). A self-running demonstration can be downloaded from the project Web site at `http://www.htc.honeywell.com/projects/sepia`.

Summary

Over the last several years, a technological convergence—increasing processing power and memory capacities, component software infrastructure developments, adaptive agent architectures—has made possible the development of a new class of simulation and optimization tools. With careful design, these tools can be used by nonexperts, and they can provide a level of decision support and insight to analysts and planners. It is, however, important to recognize that agency architectures—although the object of significant attention in both the software research community and the high-technology media—is no panacea (Wooldridge and Jennings, 1998). Problems that will benefit from an agent-oriented solution are those that have an appropriate degree of modularity and concurrency. Even then, considerable effort must be invested to endow agents with the necessary domain knowledge. For optimization applications, adaptation capabilities must be carefully enabled in agents to ensure that the flexibility of the agent-based approach

does not immediately imply drastic computational inefficiency.

Conclusions

We conclude with some "slogans" that reflect the motivations and philosophy underlying the research reported above.

Embrace pluralism. The more complex and encompassing the problems we attempt to solve, the less likely that any one solution approach will suffice. Key problem characteristics-such as the degree of knowledge and the amount of data available-will vary considerably across the range of enterprise optimization applications. Covering the space of problems of interest requires a multipronged research agenda and the development of approaches that are applicable across the diversity of problem instances.

Leverage existing foundations. New classes of problems need not mean new built-from-scratch solutions. Especially from a pragmatic perspective, an ability to create technology that can "piggy-back" on existing infrastructure can be a key differentiator. The extent of disruption of systems and processes is always a consideration in the adoption of new research results.

Exploit IT advances. Advances in hardware, software, and communication platforms do not just allow more complex algorithms to be run; they also suggest new ways of thinking about problems. The doing away with traditional technology constraints can be a liberating event for the research community, although it is important to remember that the inertia of a mature, established industry is a significant constraint in its own right.

Pursue multidisciplinary collaborations. Another corollary of complexity is that its management is a multidisciplinary undertaking (Samad and Weyrauch, 2000). In the current context, this is not only a matter of marrying control theory, chemical engineering, and computer science. As we attempt to automate larger-scale systems and pursue the autonomous operation of entire enterprises, any delimiting of multidisciplinary connections seems arbitrary.

It appears to be a law of automation that the larger the scale of the system to be automated, the more specialized and less generic the solution. At one extreme, the PID controller is ubiquitous across all industries (process, aerospace, automotive, buildings, etc.) for single-loop regulation. At the multivariable control level, MPC is the technology of choice for the process industries but has had little impact in others. For enterprise optimization, effective solutions will likely be even more domain-specific. The fact that we have discussed three very different technologies does not imply a fundamental uncertainty about which one of these will ultimately be the unique "winner"; rather, it reflects a fundamental belief that process enterprise optimization is too complex and diverse a problem area for any one solution approach to satisfactorily address.

References

Bain, M. L., K. W. Mansfield, J. G. Maphet, W. H. Bosler, and J. P. Kennedy, Gasoline blending with an integrated on-line optimization, scheduling, and control system, In *Proc. National Petroleum Refiners Association Computer Conference*, New Orleans, LA (1993).

Bemporad, A. and M. Morari, "Control of systems integrating dynamics, logic, and constraints," *Automatica*, **35**(3), 407–427 (1999).

Bodington, E., *Planning, Scheduling, and Control Integration in the Process Industries*. McGraw-Hill, New York (1995).

Bottou, L. and V. Vapnik, "Local learning algorithms," *Neural Computation*, **4**, 888–900 (1992).

Box, G. E. P. and G. M. Jenkins, *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco (1970).

Cleveland, W. S., "Robust locally-weighted regression and smoothing scatterplots," *J. Amer. Statist. Assoc.*, **74**, 829–836 (1979).

Cybenko, G., Just-in-time learning and estimation, In Bittanti, S. and G. Picci, editors, *Identification, Adaptation, Learning*, NATO ASI Series, pages 423–434. Springer-Verlag, New York (1996).

del Toro, J. L., Computer-integrated manufacturing at the Monsanto Pensacola plant, Presented at AIChE Spring Meeting, Houston, TX (1991).

Deng, K. and A. W. Moore, Multiresolution instance-based learning, In *Proc. 14th Int. Joint Conference on Artificial Intelligence*. Morgan Kaufmann (1994).

Finin, T., R. Fritzson, D. McKay, and R. McEntire, KQML as an agent communication language, In *Proc. Third International Conference on Information and Knowledge Management (CIKM'94)*. ACM Press (1994). Online at `http://www.cs.umbc.edu/kqml/papers/kqml_acl.ps`.

Fogel, D. B., *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, Piscataway, NJ (1995).

García-Flores, R., X. Z. Wang, and G. E. Goltz, Agent-based information flow for process industries' supply chain modeling, In *Proc. Process Control and Instrumentation 2000*, Glasgow (2000).

Hall, J. and T. Verne, "RCU optimization in a DCS gives fast payback," *Hydrocarbon Process.*, pages 85–92 (1993).

Hall, J. and T. Verne, "Advanced controls improve operation of Lubes Plant dewaxing unit," *Oil and Gas J.*, pages 25–28 (1995).

Härdle, W., *Applied Non-parametric Regression*. Cambridge University Press (1990).

Harp, S. A., S. Brignone, B. F. Wollenberg, and T. Samad, "SEPIA: A simulator for electric power industry agents," *IEEE Cont. Sys. Mag.*, pages 53–69 (2000).

Hastie, T. J. and R. J. Tibshirani, *Generalized Additive Models*. Chapman & Hall, London (1990).

Hinton, G. E. and S. J. Nowlan, "How learning can guide evolution," *Complex Systems*, **1**, 495–502 (1987).

Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME, J. Basic Engineering*, pages 35–45 (1960).

Koza, J., *Genetic Programming*. MIT Press, Cambridge, MA (1992).

Kulhavý, R. and P. Ivanova, Memory-based prediction in control and optimisation, In *14th World Congress of IFAC*, volume H, pages 289–294, Beijing, China (1999).

Lu, J. and J. Escarcega, RMPCT: Robust MPC technology simplifies APC, Presented at AIChE Spring Meeting, Houston, TX (1997).

Lu, J., Multi-zone control under enterprise optimization: Needs, challenges and requirements, In *Proc. International Symposium on Nonlinear MPC: Assessment and Future Directions*, Ascona, Switzerland (1998).

Mitchell, M., *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA (1996).

Nath, R., Z. Alzein, R. Pouwer, and M. Leseur, On-line dynamic optimization of an ethylene plant using Profit(r) Optimizer, Paper presented at NPRA Computer Conference, Kansas City, MO (1999).

Peterka, V., Bayesian approach to system identification, In Eykhoff, P., editor, *Trends and Progress in System Identification*, pages 239–304, Elsmford, NY. Pergamon Press (1981).

Prett, D. M. and C. E. Garcìa, *Fundamental Process Control*. Butterworths, Boston (1988).

Samad, T. and J. Weyrauch, editors, *Automation, Control and Complexity: An Integrated View*. John Wiley and Sons, Chichester, UK (2000).

Schaal, S. and C. G. Atkeson, "Robot juggling: An implementation of memory-based learning," *Control Systems Magazine*, **14**, 57–71 (1994).

Sheehan, B. and L. Reid, "Robust controls optimize productivity," *Chemical Engineering* (1997).

Sjöberg, J., Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: A unified overview," *Automatica*, **31**, 1691–1724 (1995).

Smith, F. B., An FCCU multivariable predictive control application in a DCS, The 48th Annual Symposium on Instrumentation for the Process Industries (1993).

Verne, T. and J. Escarcega, Multi-unit refinery optimization: Minimum investment, maximum return, In *The Second International Conference and Exhibition on Process Optimization*, pages 1–7, Houston, TX (1998).

Watano, T., K. Tamura, T. Sumiyoshi, and P. Nair, Integration of production, planning, operations and engineering at Idemistu Petrochemical Company, Presented at AIChE Spring Meeting, Houston, TX (1993).

Watkins, C. J. C. H. and P. Dayan, "Q-Learning," *Machine Learning*, **8**, 279–292 (1992).

West, M. and J. Harrison, *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York (1989).

Wooldridge, M. and N. R. Jennings, Pitfalls of agent-oriented development, In *Proc. of the 2nd Int. Conf. on Autonomous Agents (AA'98)*, pages 69–76, New York. ACM Press (1998).

A Definition for Plantwide Controllability

Surya Kiran Chodavarapu* and Alex Zheng†
Department of Chemical Engineering
University of Massachusetts
Amherst, MA 01003

Abstract

Chemical process synthesis typically accounts for model uncertainty by ensuring process flexibility. However, ensuring process flexibility does not guarantee steady-state robust feasibility (i.e., the existence of a plantwide control system to maintain the process at a desired steady state in the presence of uncertainty). We show, through examples, that the difference between process flexibility and steady-state robust feasibility can be observed at three levels of severity. A definition for plantwide controllability that guarantees steady-state robust feasibility is proposed.

Keywords

Flexibility, Controllability, Plantwide control, Steady-State robust feasibility

Introduction

A minimum requirement for process synthesis is to design an operable process. A process is operable if there exists a plantwide control system to maintain the process at a desired steady state in the presence of uncertainty. Considering uncertainty explicitly during process synthesis is important for many reasons. For example, models used for process synthesis are rarely perfect and assumptions made during process synthesis may not hold exactly. As a result, the process may not be operable if the uncertainty is not properly accounted for.

Researchers have proposed using various process flexibility conditions during process synthesis to ensure that there exist feasible steady-state operating conditions in the presence of uncertainty. In general, these conditions are independent of plantwide control systems. Grossmann and co-workers (Swaney and Grossmann, 1985a,b, etc.), for example, have developed several methods to deal with uncertainty optimally. More recent work in this area has been by Pistikopoulos and co-workers (Pistikopoulos, 1995; Georgiadis and Pistikopoulos, 1999, etc.) who focus on combining more than one operability characteristics during synthesis—such as flexibility, controllability and reliability. In this paper, we will show that ensuring process flexibility does not guarantee the existence of a plantwide control system. In fact, we identify three cases in which the process design satisfies the process flexibility conditions although the process is not plantwide controllable at the desired steady state.

Controllability is generally considered after a process has been synthesized. Many tools have been proposed to study process controllability issues (Lee et al., 1991; Braatz et al., 1991; Braatz and Morari, 1994; Skogestad and Wolff, 1992, etc.). However, most of the tools assume a fixed control structure (i.e., a fixed set of controlled variables and manipulated variables) and focus on the controller synthesis. The plantwide controllability definition proposed in this paper does not assume a fixed control structure.

*schodava@ecs.umass.edu
†To whom all correspondence should be addressed. zzheng@ecs.umass.edu

Process Flexibility

Process flexibility is defined by Grossmann and Swaney (Swaney and Grossmann, 1985a) as "the ability of a design to tolerate and adjust to variations in conditions which may be encountered during operation." Mathematically, this definition means that, for each allowable value of p and d, there exist allowable values of u and y such that the following is feasible:

$$\begin{cases} f(u,d,y,x,p) & = 0 \\ g(u,d,y,x,p) & \leq 0 \end{cases} \tag{1}$$

where x are the design variables, u are the manipulated variables, d are the disturbances, y are the outputs and p are the system parameters. The equation set, f, corresponds to all the energy and material balances and other algebraic equations (e.g., vapor-liquid relations), while g corresponds to the process constraints for the system (e.g., input and output constraints). Notice that for simplicity, we have only considered the parametric uncertainty and ignored the structural uncertainty.

This process flexibility definition assumes that both u and y can be adjusted based on the values of d and p to ensure (1) is feasible. However, in practice, both d and p are rarely known exactly. A common strategy is to adjust u to maintain some of the outputs at desired setpoints via a plantwide control system, an aspect not considered by the process flexibility definition. Therefore, we would expect that ensuring process flexibility is not sufficient to guarantee the steady-state robust feasibility, defined below.

Steady-State Robust Feasibility: A process is robustly feasible at steady state if there exists a plantwide control system so that (1) is feasible for all allowable values of p and d.

Plantwide Controllability

We propose a definition for plantwide controllability which explicitly considers the plantwide control system. In this definition, the steady-state relations for the plantwide control system are represented by $h(u, d, y, x, p, r) = 0$, where r represents the setpoints for a set of controlled variables.

Steady-State Plantwide Controllability: A process is plantwide controllable at steady state if there exist h and r such that the following is feasible for all allowable values of d and p:

$$\begin{cases} f(u, d, y, x, p) & = 0 \\ g(u, d, y, x, p) & \leq 0 \\ h(u, d, y, x, p, r) & = 0 \end{cases} \quad (2)$$

Notice that h and r are not unique. For fixed h and r, the plantwide controllability reduces to the conventional controllability.

Comparing the process flexibility condition (1) and the proposed plantwide controllability condition (2), it should be clear that ensuring process flexibility is a necessary but not a sufficient condition to guarantee steady-state plantwide controllability (i.e., a design which satisfies the process flexibility condition may not be steady-state plantwide controllable). In fact, there are three cases where (2) is not feasible even if (1) is feasible:

- Case 1: The steady-state plantwide controllability is not guaranteed for a fixed set of controlled variables at fixed setpoints (i.e., h and r are fixed). However, the steady-state plantwide controllability may be restored by simply choosing alternative setpoints.
- Case 2: The steady-state plantwide controllability is not guaranteed for a fixed set of controlled variables regardless of their setpoints (i.e., h is fixed but r is not). The steady-state plantwide controllability may be restored by choosing an alternative set of controlled variables.
- Case 3: The steady-state plantwide controllability is not guaranteed regardless of the controlled variable set and their setpoints (i.e., both h and r are not fixed). The steady-state plantwide controllability can only be restored through process retrofits or redesign.

We illustrate the first two cases using the Hydrodealkylation of Toluene (HDA) process and the third case using a reactor-separator-recycle (RSR) process.

HDA Process—Cases 1 and 2

We assume that the reactions involved in this process are

Rxn 1 $Toluene + Hydrogen \rightarrow Benzene + Methane$

Rxn 2 $2Benzene \leftrightarrow Diphenyl + Hydrogen$

Parameter	Reaction 1	Reaction 2
Rate constant	6×10^{14}	7.6×10^{14}
ΔH_{rxn} (kJ/kgmol)	-42,000	8,100
Activation Energy (kJ/kgmol)	220,000	130,000

Table 1: Kinetic parameters for the HDA process: units for reaction 1 rate constants are $\frac{(ft^3)^{1/2}}{lbmol^{1/2}-hr}$ while the units for reaction 2 rate constants are $\frac{ft^3}{lbmol-hr}$.

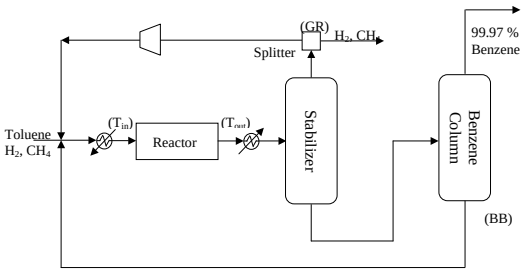


Figure 1: Simplified flowsheet for HDA process.

Both reactions are elementary and the nominal values of the kinetic parameters for these reactions are summarized in Table 1. To study the effect of model uncertainty, we assume a 50% uncertainty in the rate constant for Reaction 1.

The flowsheet of the HDA process (Figure 1) is simplified by assuming that diphenyl is recycled to extinction and that there is no impurity in the feed stream nor any by-product other than diphenyl. The stabilizer is also assumed to yield a perfect split. Finally, fixing the pressure in the benzene column (P = 2 atm) and the fresh feed flowrates, there are four degrees of freedom—one for the reactor, two for the benzene column, and one for the gas recycle splitter.

Constraints for this process include a benzene product purity constraint (99.97%), flooding and weeping constraints for the benzene column (2000 and 500 kgmol/hr respectively (Kister, 1989)), and a maximum reactor outlet temperature (700 deg C) to prevent coking.

In the following sections, we present designs for Cases 1 and 2. Both designs consist of a 54 stage benzene column with the feed entering at the 38th stage and a reactor diameter fixed at 3.05 meters. However, the reactor length in the two designs differs.

Case 1: Infeasible at Fixed Setpoints

Consider the process with a reactor length of 24.4 meters. We can verify that this design satisfies condition (1) by fixing the benzene bottoms composition and gas recycle ratio at 0.05% and 88% and decreasing the reac-

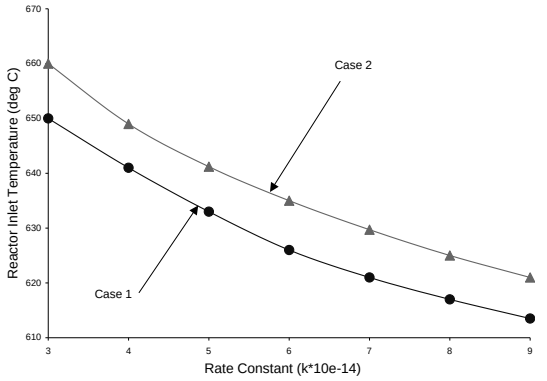


Figure 2: Reactor inlet temperature setpoint as function of the rate constant for the designs in Cases 1 and 2.

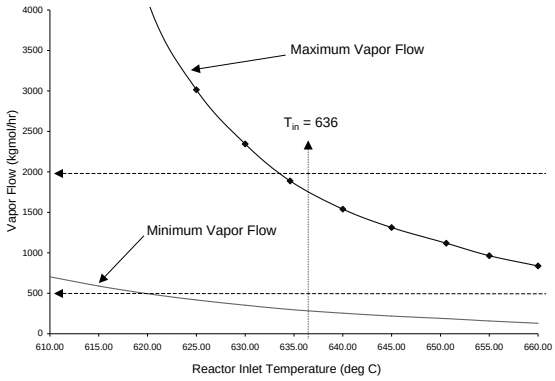


Figure 3: Minimum and maximum column vapor flows as functions of reactor inlet temperature setpoint with gas recycle setpoint of 88% and benzene bottoms composition setpoint of 0.05%.

tor inlet temperature with increasing values of the rate constant as shown in Figure 2. The maximum vapor flow, minimum vapor flow and maximum reactor outlet temperature obtained when following this trajectory are 1800 kgmol/hr, 1300 kgmol/hr and 689 deg C respectively. Since there exist steady-state operating conditions which satisfy all process constraints despite the uncertainty, this design satisfies condition (1).

However, to change the reactor inlet temperature setpoint according to this trajectory, exact values of the rate constant are needed. Since this is usually not the case in practice, we instead control the reactor inlet temperature, benzene product purity, benzene bottoms composition, and gas recycle ratio setpoints at 636 deg C, 99.97%, 0.05%, and 88%, respectively. Note that since we are only interested in the steady-state plantwide controllability, we merely need to select a set of controlled variables and their setpoints. Figure 3 shows

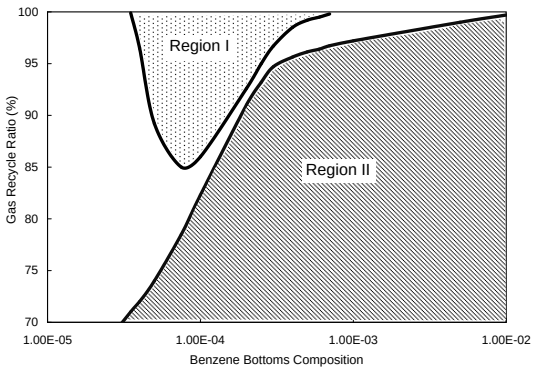


Figure 4: Feasibility regions for the design in Case 2.

that, for this reactor inlet temperature setpoint, the minimum vapor flow is less than the weeping constraint of 500 kgmol/hr. Therefore, the design with the proposed plantwide control system is not feasible at the specified setpoints although condition (1) is satisfied. In fact, no setpoint for the reactor inlet temperature would satisfy both flooding and weeping limits in this case.

Steady-state plantwide controllability can be restored to this process by choosing alternative setpoints. For example, changing the benzene bottoms composition setpoint to 0.01% and the reactor inlet temperature to 649 deg C would guarantee that all process constraints are satisfied.

Case 2: Infeasible for Fixed Controlled Variables

With a reactor of length 18 meters, the design still satisfies condition (1) (Figure 2 for Case 2). The maximum and minimum vapor flows in this case are 1900 and 1400 kgmol/hr while the maximum reactor outlet temperature is 695 deg C. However, steady-state plantwide controllability is still not guaranteed since the rate constant is not known exactly in practice.

Suppose we choose the benzene bottoms composition, gas recycle ratio, reactor inlet temperature and the benzene product purity as controlled variables. Figure 4 then maps the regions in which the flooding, weeping and reactor outlet temperature constraints are satisfied for different gas recycle ratios and benzene bottoms compositions setpoints. In this diagram, Region I corresponds to the setpoints which satisfy the flooding and weeping constraints, while Region II corresponds to the setpoints which satisfy the flooding and the reactor outlet temperature constraints. Since Region I and II do not overlap, there does not exist a gas recycle ratio and benzene bottoms composition setpoint satisfying all process constraints. This controlled variable set is therefore not feasible regardless of the setpoints.

Steady-state plantwide controllability can be restored to this system, by choosing an alternative set of con-

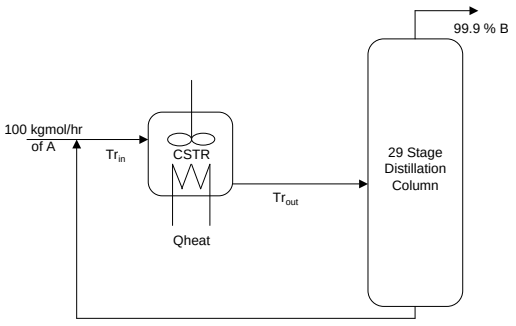


Figure 5: Reactor-Separator-Recycle process.

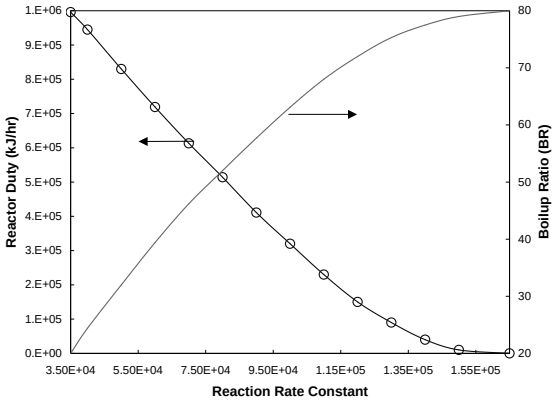


Figure 6: Reactor duty and reboil ratio setpoints as functions of the rate constant.

trolled variables. This can be accomplished, for example, by using the reboil duty (with a setpoint of 3×10^6 kJ/hr) instead of the benzene bottoms composition as a controlled variable.

Case 3: RSR Process

Determining a HDA design that is infeasible for all steady-state control structures requires the evaluation of all controlled variable alternatives and setpoints. To illustrate this third case without the associated complexity, we use a simpler process with one less degree of freedom—a reactor-separator-recycle (RSR) process (Figure 5).

The reactor in this process is a CSTR with a volume of 0.6 m^3, while the separator is a 29-stage distillation column with the feed entering at the 15th stage. The reaction is $A \rightarrow B$ with elementary kinetics. The activation energy and heat of reaction are 5×10^4 and -8.5×10^3 kJ/kgmol, respectively, while the nominal value of the rate constant, k_r, is 10^5. The effects of uncertainty are studied by assuming that the actual rate constant could be any value between $k_{rmin} = 3.5 \times 10^4$ and $k_{rmax} = 1.65 \times 10^5$.

Constraints for this process include flooding and weeping limits (12000 kgmol/hr and 3000 kgmol/hr, respectively) for the distillation column, a minimum reactor feed temperature ($Tr_{in} \geq 55$ deg C), a maximum reactor effluent temperature ($Tr_{out} \leq 117.5$ deg C), and a product purity constraint in the distillate of 99.9% B. We can verify that condition (1) is satisfied for this design by choosing setpoints according to the rate constant as shown in Figure 6: The maximum and minimum vapor flows in the column are 12000 and 7000 kgmol/hr, respectively, the minimum reactor feed temperature is 65 deg C and the maximum reactor effluent temperature is 114 deg C.

However, choosing setpoints in this fashion assumes that the rate constant is known exactly. Since this may not be the case in practice, the question is: Can we synthesize a plantwide control system such that all constraints can be satisfied for all possible values of the rate constant? For this system, we have to choose three controlled variables—two for the distillation column and one for the reactor. Choosing the product purity of B as one controlled variable, we are left with selecting two more controlled variables. Controlled variable alternatives for the reactor include the reactor duty and the reactor effluent temperature while controlled variable alternatives for the stripping section of the distillation column include the reboil ratio, the reboil duty and the bottoms composition of A.

With the reactor effluent temperature as the controlled variable for the reactor, there are two important limits. When the reaction rate constant is at the minimum possible value (i.e., $k_r = k_{rmin}$), a reactor effluent temperature setpoint greater than 114 deg C is required to ensure that the maximum vapor flow in the column would be less than 12000 kgmol/hr. On the other hand, when the rate constant is at the maximum possible value (i.e., $k_r = k_{rmax}$), a reactor effluent temperature setpoint less than 109 deg C is necessary to satisfy the weeping constraint. Since there does not exist a reactor effluent temperature setpoint satisfying both the flooding and weeping constraints for all possible values of the rate constant, any controlled variable set with the reactor effluent temperature would not be feasible and can be eliminated. The only controlled variable alternative remaining for the reactor is the reactor duty. Note that the reactor duty setpoint has to be less than 10^6 kJ/mol to satisfy the reactor effluent temperature constraint. With the product purity and reactor duty as controlled variables, we now discuss three alternatives with the bottoms composition of A, reboil ratio, or reboil duty as the other controlled variable.

For the controlled variable set with the bottoms composition of A, the setpoints for which the flooding and weeping constraints are satisfied are shown in Figure 7.

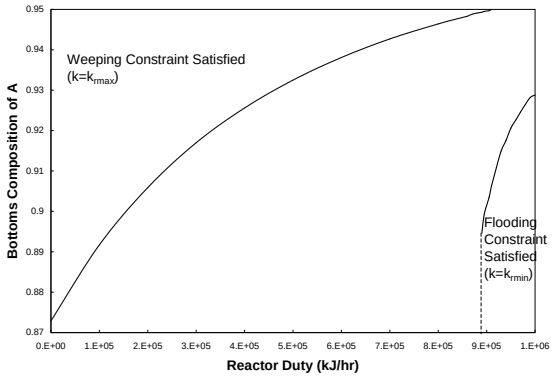


Figure 7: Feasibility regions with reactor duty and bottoms composition as controlled variables.

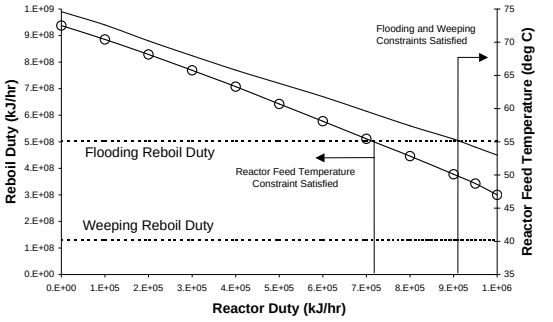


Figure 8: Feasibility regions with reactor duty and reboiler duty as controlled variables (minimum reactor feed temperature plotted by circles).

Since there do not exist setpoints which satisfy both the constraints, this controlled variable alternative is also not feasible. A graph similar to Figure 7 is obtained when using the reboil ratio instead of the bottoms composition of A and therefore, the controlled variable set with reboil ratio and reactor duty is also not feasible.

The final alternative to be studied utilizes the reactor duty, reboil duty and the product purity of B as controlled variables. With this controlled variable set, Figure 8 indicates that for reactor duties greater than 9×10^5 kJ/hr, there exist reboiler duty setpoints which satisfy both the flooding and weeping constraints. However, these setpoints are yet again infeasible since the minimum reactor feed temperature is below 55 deg C at these operating conditions.

From this discussion, it should be apparent that although this design satisfies condition (1), there may not exist a set of controlled variables and corresponding setpoints which would satisfy all process constraints for the given uncertainty. It should be emphasized that we only explored the obvious controlled variable alternatives and that we did not explore all possible alternatives (e.g., control of tray temperature, control of a variable that is a function of the column composition profile, etc.). While the steady-state plantwide controllability can be restored to designs corresponding to cases 1 and 2 with relatively small modifications, processes which fall in this third case have no recourse but for a major redesign or retrofit.

Conclusions

In this paper, we showed that ensuring process flexibility does not guarantee steady-state robust feasibility. A definition for steady-state plantwide controllability that ensures steady-state robust feasibility is proposed. The difference between this definition of plantwide controllability and the conventional controllability definition is that plantwide controllability does not assume a fixed controlled structure.

References

Braatz, R. D. and M. Morari, "Minimizing the Euclidean Condition Number," *SIAM J. Cont. Optim.*, **32**(6), 1763–1768 (1994).

Braatz, R. D., J. H. Lee, and M. Morari, Screening Plant Designs and Control Structures for Uncertain Systems, In *IFAC Workshop on Integration of Process Design and Control*, Baltimore, Maryland (1991).

Douglas, J. M., *Conceptual Design of Chemical Processes*. McGraw-Hill, New York (1988).

Georgiadis, M. C. and E. N. Pistikopoulos, "An Integrated Framework for Robust and Flexible Process Systems," *Ind. Eng. Chem. Res.*, **38**, 133–143 (1999).

Kister, H. Z., *Distillation Operation*. McGraw–Hill Book Company, New York (1989).

Lee, J. H., R. D. Braatz, M. Morari, and A. Packard, "Screening Tools for Robust Control Structure Selection," *Automatica*, **31**(2), 229–235 (1991).

Pistikopoulos, E. N., "Uncertainty in Process Design and Operations," *Comput. Chem. Eng.*, **19**, S553–S563 (1995).

Skogestad, S. and E. A. Wolff, Controllability Measures for Disturbance Rejection, In *IFAC Workshop on Interactions between Process Design and Process Control*, London, England (1992).

Swaney, R. E. and I. E. Grossmann, "An Index for Operational Flexibility in Chemical Process Design. Part I: Formulation and Theory," *AIChE J.*, **31**, 621 (1985a).

Swaney, R. E. and I. E. Grossmann, "An Index for Operational Flexibility in Chemical Process Design. Part II: Computational Algorithms," *AIChE J.*, **31**, 631 (1985b).

Nonlinear Process Control: Novel Controller Designs for Chemical Processes with Uncertainty and Constraints

Nael H. El-Farra* and Panagiotis D. Christofides
Department of Chemical Engineering
University of California
Los Angeles, CA 90095

Abstract

This paper highlights some of our recent results on control of single-input single-output constrained uncertain nonlinear processes. The main issues and challenges that arise in control of such processes are discussed and a novel Lyapunov-based framework for nonlinear controller synthesis is presented to address these issues. The proposed framework leads to the explicit synthesis of nonlinear robust optimal state feedback controllers, with well-characterized stability and performance properties, that cope effectively with the simultaneous problems of severe process nonlinearities, significant model uncertainty, and hard constraints on the manipulated input. The impact of the proposed framework is analyzed and its implications for nonlinear process control are identified.

Keywords

Uncertainty, Lyapunov's direct method, Input constraints, Bounded control, Inverse optimality

Introduction

In designing an effective process control system, there are several fundamental issues that transcend the boundaries of specific process applications. Although they may vary from one application to another and have different levels of significance, these issues remain generic in their relationship to the control design objectives. Central to these issues is the requirement that the control system provide satisfactory performance in the face of severe process nonlinearities, modeling errors, process variations, and actuator constraints. Nonlinear behavior, model uncertainty and input constraints represent some of the more salient features whose frequently-encountered co-presence in a multitude of chemical processes renders the task of controlling such processes a challenging one. For example, many important industrial processes including highly exothermic chemical reactions, high purity distillation columns, and batch systems exhibit strong nonlinear behavior and cannot be effectively controlled with controllers designed on the basis of approximate linear or linearized process models. The limitations of traditional linear control methods in dealing with nonlinear chemical processes have become increasingly apparent as chemical processes may be required to operate over a wide range of conditions due to large process upsets or set-point changes. This realization has consequently motivated intense research activity in the area of nonlinear process control over the last fifteen years. The literature on nonlinear process control is really extensive (see, e.g., (Allgöwer and Doyle III, 1997) for references).

In addition to nonlinear behavior, many industrial processes are characterized by the presence of model uncertainty such as unknown process parameters and external disturbances which, if not accounted for in the controller design, may cause performance deterioration and even closed-loop instability. Significant research work has consequently focused on this problem including the use of integral action in conjunction with feedback linearizing controllers to compensate for constant model uncertainty and the design of robust controllers through Lyapunov's direct method for processes with time-varying uncertain variables (e.g., (Christofides et al., 1996)).

The above approaches, however, do not lead in general to controllers that are optimal with respect to a meaningful cost and therefore do not guarantee achievement of the control objectives with the smallest possible control action. This is an important limitation of these methods, especially in view of the fact that the capacity of control actuators used to regulate chemical processes is almost always constrained. The problems caused by input constraints have consequently motivated many recent studies on the dynamics and control of chemical processes subject to input constraints. Notable contributions in this direction include controller design and stability analysis within the model predictive control framework (e.g., (Rawlings, 1999; Kurtz et al., 2000)), constrained quadratic-optimal control (Chmielewski and Manousiouthakis, 1996), the design of "anti-windup" schemes in order to prevent excessive performance deterioration of an already designed controller when the input saturates (Kothare et al., 1994; Valluri and Soroush, 1998; Kapoor and Daoutidis, 1999). However, these control methods do not explicitly account for robust uncertainty attenuation.

Summarizing, a close look at the available literature reveals the fact that, at this stage, existing nonlinear process control methods lead to the synthesis of controllers that can deal with either model uncertainty or input constraints, but not simultaneously or effectively with both. This clearly limits the achievable control quality and closed-loop performance, especially in view

*Financial support in part by UCLA through a Chancellor's Fellowship for N. H. El-Farra and the University of California Energy Institute is gratefully acknowledged.

of the commonly-encountered co-presence of uncertainty and constraints in chemical processes. Therefore, the development of a unified framework for control of nonlinear processes that explicitly accounts for the presence of model uncertainty and input constraints is expected to have a significant impact on chemical process control. Motivated by this, we outline in this paper some of our recent results on the development of this framework and discuss, in particular, its implications for enriching our existing arsenal of nonlinear control tools.

Nonlinear Processes with Uncertainty and Constraints

We consider single-input single-output nonlinear processes with uncertain variables and input constraints modeled by the class of continuous-time nonlinear ordinary differential equation systems with the following state-space description:

$$\begin{aligned} \dot{x} &= f(x) + g(x)sat(u) + \sum_{k=1}^{q} w_k(x)\theta_k(t) \\ y &= h(x) \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ denotes the vector of process state variables, $u \in \mathbb{R}$ denotes the manipulated input, $\theta_k(t) \in \mathcal{W} \subset \mathbb{R}$ denotes the k-th uncertain (possibly time-varying) but bounded variable taking values in a nonempty compact convex subset $\mathcal{W}$ of $\mathbb{R}$, $y \in \mathbb{R}$ denotes the output to be controlled, and *sat* refers to the standard saturation nonlinearity. The presence of the *sat* operator in Equation 1 signifies the presence of hard constraints on the manipulated input. The uncertain variable $\theta_k(t)$ may describe time-varying parametric uncertainty and/or exogenous disturbances. It is assumed that the origin is the equilibrium point of the nominal (i.e., $u(t) = \theta_k(t) \equiv 0$) system of Equation 1.

Main Issues on Control of Constrained Uncertain Nonlinear Processes

Towards our end goal of developing an effective control strategy that enforces the desired stability and performance properties in nonlinear processes subject to model uncertainty and actuator constraints, we begin in this section by identifying some of the outstanding issues that arise in this problem and must be addressed properly. While the individual presence of either model uncertainty or input constraints poses its own unique set of problems that must be dealt with at the stage of controller design, the combined presence of both uncertainty and constrains is far more problematic for process stability and performance. The difficulty here emanates not only from the cumulative effect of the co-presence of the two phenomena but is, more importantly, due to the additional issues that arise from the interaction of the two.

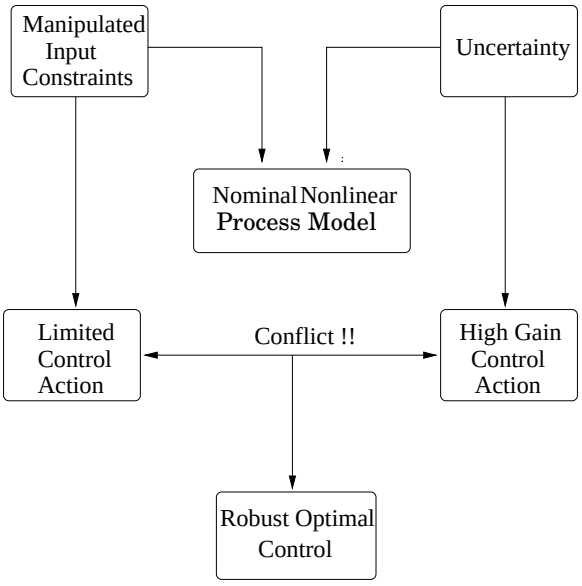


Figure 1: Paradigm for control of constrained uncertain nonlinear processes.

At the core of these issues are the following two problems:

1. The co-presence of model uncertainty and input constraints creates an inherent conflict in the controller design objectives and control policy to be implemented. On one hand, suppressing the influence of significant external disturbances requires typically large (high-gain) control action. On the other hand, the availability of such action is often limited by the presence of input constraints. Failure to resolve this conflict will render any potential control strategy essentially ineffective. A schematic representation of this conflict is depicted in Figure 1.
2. The set of feasible process operating conditions under which the process can be operated safely and reliably is significantly restricted by the co-presence of uncertainty and constraints. While, on their own, input constraints place fundamental limitations on the size of this set (and consequently on our ability to achieve certain control objectives), these limitations become even stronger when uncertainty is present. At an intuitive level, many of the feasible operating conditions permitted by constraints under nominal conditions (i.e. predicted using a nominal model of the process) cannot be expected to continue to be feasible in the presence of significant plant-model mismatch. The success of a control strategy in effectively addressing the problems of uncertainty and constraints hinges, therefore, not only on the design of effective controllers, but also on the ability to predict a priori the feasible conditions under which the designed controllers are guar-

anteed to work in the presence of both uncertainty and constraints.

Having discussed some of the main issues involved, we are now motivated to proceed in the next two sections with the presentation of some of our recent results on this problem, which directly address the issues outlined above and culminate in the development of a general Lyapunov-based framework for control of nonlinear processes with model uncertainty and input constraints. More specifically, the developed framework entails:

1. The synthesis of nonlinear robust optimal controllers that account explicitly for the problems of significant model uncertainty and input constraints and enforce the desired stability, robustness, optimality, and explicit constraint-handling properties in the closed-loop system.

2. The explicit and quantitative characterization of the limitations imposed by the co-presence of uncertainty and constraints on our ability to steer the nonlinear process in a desired direction.

To lay the appropriate foundation for the development of the proposed framework, we begin in the next section by presenting first some of the key tools necessary to address the first issue. Then, in the following section, we show how these tools can be built upon to address the second issue, leading finally to the desired framework.

Robust Optimal Control

The control paradigm presented in Figure 1 suggests a natural approach to resolve the inherent conflict between the need to compensate for the effect of model uncertainty and the limitations imposed by input constraints on the available control action. This is the robust optimal control approach which involves the synthesis of robust controllers that use minimal or reasonably small control action to compensate for the effect of significant model uncertainty and achieve robust stabilization. In (El-Farra and Christofides, 2001a), using a novel combination of Lyapunov's direct method and the inverse optimal control approach, we synthesized robust optimal controllers that enforce, in the presence of significant model uncertainty and absence of constraints: a) global stability, b) robust asymptotic output tracking with arbitrary degree of attenuation of the effect of uncertainty on the output, and c) optimal performance, in the closed-loop system. The controllers take the general form

$$u = -\frac{1}{2}R^{-1}(x, \theta_b, \phi)L_{\bar{g}}V \tag{2}$$

where

$$\frac{1}{2}R^{-1}(\cdot) = c_0 + \frac{L_{\bar{f}}V + \sqrt{(L_{\bar{f}}V)^2 + (L_{\bar{g}}V)^4}}{(L_{\bar{g}}V)^2} + \frac{\rho + \chi\sum_{k=1}^{q}\theta_{bk}|L_{\bar{w}k}V||2b^TPe|}{(L_{\bar{g}}V)^2(|2b^TPe| + \phi)} \tag{3}$$

is a strictly positive definite function, θ_{bk}'s represent the available bounds that capture the magnitude of uncertainty for all time, c_0, ρ, χ, and ϕ are tuning parameters that can be adjusted to achieve the desired degree of uncertainty attenuation, $V = e^TPe$ is a scalar quadratic Lyapunov function of the tracking error whose time-derivative is rendered negative definite by the feedback law of Equation 2 along the trajectories of the closed-loop system, and $L_{\bar{g}}V$ is the standard Lie derivative notation. Further details on the controller synthesis procedure and notation used can be found in (El-Farra and Christofides, 2001a). In what follows, we outline some of the key desirable features of the robust optimal controllers proposed and their implications for the problem of controlling constrained uncertain nonlinear processes.

1. The robust optimal controllers of Equation 2 possess two desirable properties not present in other nonlinear controllers designed on the basis of feedback linearization concepts. The first property is their ability to recognize the presence of beneficial (stabilizing) nonlinearities in the process and prevent their unnecessary cancellation. As a result, these controllers do not waste unnecessary control effort to accomplish the desired closed-loop objectives. An important implication of this property is the fact that these controllers (though not designed to handle constraints explicitly yet) are better equipped to cope with the limitations imposed by input constraints on the available control action than other controllers which may request unnecessarily large control effort. The second property is the use of domination, rather than cancellation, to eliminate the effects of harmful (destabilizing) nonlinearities. This property guards against the non-robustness of nonlinear cancellation designs which increases the risk of instability due to the presence of other uncertainty not taken into account in the controller design.

2. The use of Lyapunov's direct method in the controller design allows the synthesis of robust controllers that can effectively attenuate the effect of time-varying persistent uncertainty on the closed-loop output which cannot be achieved using classical uncertainty compensation techniques, including the incorporation of integral action and parameter adaptation in the controller. For constant disturbances, the Lyapunov approach offers an alternative method for disturbance rejection that avoids

the use of integral action which contributes to the problem of windup in the presence of constraints. Furthermore, robust stabilization is accomplished regardless of how large the disturbances are so long as bounds are available that capture their magnitude. For nonlinear controller design, in general, Lyapunov methods provide useful and systematic tools (see, e.g.,(Kazantzis and Kravaris, 1999)).

3. Using the inverse optimal control approach (Freeman and Kokotovic, 1996; El-Farra and Christofides, 2001a), one can rigorously prove that the robust Lyapunov-based controller of Equation 2 is optimal with respect to a meaningful performance index of the form:

$$J = \int_0^\infty (l(e) + uR(x)u)dt \tag{4}$$

which imposes physically sensible penalties on both the tracking error ($l(e)$ is a smooth positive definite nonlinear function bounded below by a quadratic function of the norm of the tracking error) and the control action. The inverse optimal approach provides a sound theoretical basis for explaining the origin of the controllers' optimality properties outlined in the first remark. Another major advantage of using the inverse optimal approach in controller design is the fact that it provides a convenient route for the synthesis of robust controllers that are optimal with respect to meaningful cost functionals without recourse to the unwieldy task of solving the Hamilton-Jacobi-Isaacs partial differential equation which is the optimality condition for the robust stabilization problem for nonlinear systems.

Integrating Robustness, Optimality, and Constraints

While optimality is certainly a desirable feature that every well designed controller must possess to cope with the limitations imposed by input constraints on the control action, it might not always be sufficient to address the problem in its entirety. For example, one can envision situations where the control objectives cannot be achieved in the presence of constraints, irrespective of the particular choice of the control law. Therefore, for an optimal control policy, such as the one presented in the previous section, to effectively address the problem of constraints, it is imperative that it also identifies, explicitly, the feasible operating conditions under which stability of the process can be guaranteed in the presence of constraints. In this regard, we note that although the robust optimal controllers of Equation 2 are equipped with the appropriate tools to resolve the conflict between uncertainty and constraints, they are not designed to address the second issue of explicitly characterizing the limitations imposed by input constraints on the feasible operating conditions and cannot therefore be expected to enforce the same closed-loop properties in the presence of arbitrary input constraints.

To address this issue, we developed in (El-Farra and Christofides, 2001a) a novel scaling procedure that bounds the robust optimal controllers in Equation 2:

$$|u| \leq u_{max} \tag{5}$$

where $|\cdot|$ is the Euclidean norm and u_{max} represents the available information on the actuator constraints. The result of this bounding procedure was the direct synthesis of nonlinear bounded robust optimal controllers that are conceptually aligned with the robust optimal controllers of Equation 2, but possess the additional capabilities of: a) incorporating both uncertainty and actuator constraints explicitly in the controller synthesis formula, and b) characterizing explicitly the set of feasible initial conditions starting from where the desired stability and performance properties are guaranteed in the closed-loop system under uncertainty and constraints. The resulting bounded robust optimal controllers have the general form:

$$u = -\frac{1}{2}\bar{R}^{-1}(x, u_{max}, \theta_b, \phi)L_{\bar{g}}V \tag{6}$$

where

$$\frac{1}{2}\bar{R}^{-1}(\cdot) = \frac{L^*_{\bar{f}}V + \sqrt{(L^*_{\bar{f}}V)^2 + (u_{max}L_{\bar{g}}V)^4}}{(L_{\bar{g}}V)^2[1 + \sqrt{1 + (u_{max}L_{\bar{g}}V)^2}]} \tag{7}$$

is a strictly positive definite function of its arguments and $L^*_{\bar{f}}V = L_{\bar{f}}V + \sum_{k=1}^{q}\theta_{bk}|L_{\bar{w}k}V| + \rho|e|$. For details on the notation used, see (El-Farra and Christofides, 2001a). We have shown in (El-Farra and Christofides, 2001a) that whenever the following inequality is satisfied:

$$L_{\bar{f}}V + \rho|e| + \chi\sum_{k=1}^{q}\theta_{bk}|L_{\bar{w}k}V| \leq u_{max}|L_{\bar{g}}V| \tag{8}$$

the bounded robust optimal controllers of Equation 6 enforce the following properties in the constrained closed-loop system, including: a) stability, b) robust asymptotic output tracking with arbitrary degree of uncertainty attenuation, and c) optimal performance with respect to a meaningful performance index of the general form of Equation 4 that imposes meaningful penalties on the tracking error and control action. In what follows, we outline some of the key features of the proposed bounded robust optimal controller design method and discuss its implications for the development of the desired unified framework for control of constrained uncertain nonlinear processes.

1. The inequality of Equation 8 describes explicitly the largest region in state space where the time-derivative of the Lyapunov function is guaranteed to be negative definite along the trajectories of the

closed-loop system, in the presence of uncertainty and constraints. Therefore, guided by this inequality, one can explicitly identify the set of admissible initial states, starting from where the aforementioned closed-loop properties are guaranteed, and ascertain a priori (before implementing the controller) whether stability and robust set-point tracking can be guaranteed in the constrained closed-loop system for a given initial condition. This aspect of the proposed method has important practical implications for efficient process operation since it provides the plant operators with a systematic guide to select feasible operating conditions. This is particularly significant in the case of unstable plants (e.g., exothermic chemical reactors) where lack of such a priori knowledge can lead to undesirable outcomes.

2. In addition to providing the desired characterization of the region of guaranteed closed-loop stability, the inequality of Equation 8 also depicts the region where the control action satisfies the input constraints. Within this region, no mismatch exists between the controller output and the actual process input.

3. The inequality of Equation 8 captures, in an intuitive way, the limitations imposed by uncertainty and constraints on the size of the closed-loop stability region. To this end, note that Equation 8 predicts that the tighter the input constraints (i.e., smaller u_{max}) and/or the larger the plant-model mismatch (i.e., larger θ_{bk}), the fewer the initial conditions that can be used for stabilization.

4. In light of the above remarks, it's important to compare the above control method with other methods of nonlinear control for nonlinear processes with input constraints. For example, in contrast to the traditional two-step approaches employed in analytical process control, which involves first the design of a controller for the unconstrained process and then accounts for input constraints through a suitable anti-windup modification, the bounded robust optimal control method offers a direct approach to the problem whereby the controllers of Equation 6 use directly the available information on actuator constraints (u_{max}) and uncertainty (θ_b) to compute the necessary control action; thus integrating robustness, optimality, and explicit-constraint handling capabilities in a single design. Other direct approaches for dealing with input constraints include optimization-based methods such as model predictive control. In these methods, however, the feedback law is not given explicitly, but implicitly through the optimization problem, which must be solved at each time step. Furthermore, issues of computational effort, robustness, and the a priori characterization of the region of closed-loop stability have yet to be addressed satisfactorily within these approaches.

The proposed Lyapunov-based control approaches reviewed here were applied successfully in (El-Farra and Christofides, 2001a) to benchmark examples including nonisothermal chemical reactors with unstable dynamics. Finally, we note that the problem of output feedback was recently addressed in (El-Farra and Christofides, 2001b, 2000) through combination of the robust optimal state feedback controllers with high gain observers. This approach was shown to practically preserve both the optimality properties as well as the region of guaranteed closed-loop stability obtained under state feedback.

References

Allgöwer, F. and F. J. Doyle III, Nonlinear Process Control—Which Way to the Promised Land?, In *Proceedings of 5th International Conference on Chemical Process Control*, pages 24–45, Tahoe City, CA (1997).

Chmielewski, D. and V. Manousiouthakis, "On Constrained Infinite-time Linear Quadratic Optimal Control," *Sys. Cont. Let.*, **29**, 121–129 (1996).

Christofides, P. D., A. R. Teel, and P. Daoutidis, "Robust Semiglobal Output Tracking for Nonlinear Singularly Perturbed Systems," *Int. J. Control*, **65**, 639–666 (1996).

El-Farra, N. H. and P. D. Christofides, "Robust Optimal Control and Estimation of Constrained Nonlinear Processes," *Comput. Chem. Eng.*, **24**, 801–805 (2000).

El-Farra, N. H. and P. D. Christofides, "Integrating Robustness Optimality and Constraints in Control of Nonlinear Processes," *Chem. Eng. Sci.*, **56**, 1841–1868 (2001a).

El-Farra, N. H. and P. D. Christofides, "Robust Near-Optimal Output Feedback Control of Nonlinear Systems," *Int. J. Control*, **74**, 133–157 (2001b).

Freeman, R. A. and P. V. Kokotovic, *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques*. Birkhauser, Boston (1996).

Kapoor, N. and P. Daoutidis, "An observer-based anti-windup scheme for non-linear systems with input constraints," *Int. J. Control*, **9**, 18–29 (1999).

Kazantzis, N. and C. Kravaris, "Energy-predictive control: a new synthesis approach for nonlinear process control.," *Chem. Eng. Sci.*, **54**, 1697–1709 (1999).

Kothare, M. V., P. J. Campo, M. Morari, and C. N. Nett, "A Unified Framework for the Study of Anti-Windup Designs," *Automatica*, **30**, 1869–1883 (1994).

Kurtz, M. J., G. Y. Zhu, and M. A. Henson, "Constrained Output Feedback Control of a Multivariable Polymerization Reactor," *IEEE Trans. Cont. Sys. Tech.*, **8**, 87–97 (2000).

Rawlings, J. B., Tutorial: Model Predictive Control Technology, In *Proceedings of American Control Conference*, pages 662–676, San Diego, CA (1999).

Valluri, S. and M. Soroush, "Analytical Control of SISO Nonlinear Processes with Input Constraints," *AIChE J.*, **44**, 116–130 (1998).

Efficient Nonlinear Model Predictive Control

Rolf Findeisen* and Frank Allgöwer†
Institute for Systems Theory in Engineering
University of Stuttgart, Germany

Moritz Diehl‡, H. Georg Bock§, and Johannes P. Schlöder¶
Interdisciplinary Center for Scientific Computing (IWR)
University of Heidelberg, Germany

Zoltan Nagy‖
Faculty of Chemistry and Chemical Engineering
"Babes-Bolyai" University of Cluj, Romania

Abstract

The growing interest in model predictive control for nonlinear systems, also called NMPC, is motivated by the fact that today's processes need to be operated under tighter performance specifications to guarantee profitable and environmentally safe production. One of the remaining essential problems for NMPC is the high on-line computational load. At each sampling instant, a nonlinear optimal control problem must be solved. In this paper, we summarize recent results showing the practical applicability of NMPC for process control. We show how recent advances in NMPC theory and dynamic optimization can be used to make the real-time application of NMPC feasible even for high dimensional problems. As an application example the real-time control of a high purity distillation column is considered.

Keywords

Nonlinear model predictive control, Real-time optimization, Quasi-infinite horizon, Multiple shooting, Computational effort, Large scale, Distillation control

Introduction

Over the last two decades model predictive control (MPC), also referred to as moving horizon control or receding horizon control, has become an attractive feedback strategy. Linear MPC approaches have found successful applications, especially in the process industry (Qin and Badgwell, 1996). Nowadays, tighter product quality specifications, increasing productivity demands and environmental regulations require systems to be operated closer to the boundary of the admissible operating region. To allow operation near the boundary, a linear model is often not adequate to describe the process dynamics. This motivates the use of nonlinear system models, non-quadratic cost functions and nonlinear constraints in the predictive framework, thus leading to nonlinear model predictive control (NMPC).

Recently NMPC schemes with favorable properties including guaranteed closed-loop stability or reduced computational demand have been developed, see for example De Nicolao et al. (2000); Allgöwer et al. (1999) for a review. Despite these advances concern has been raised that due to the high on-line computational load none of the available NMPC schemes can be used for real-time control in practice. This concern is based on the fact that at every sampling instant a high-dimensional nonlinear, finite horizon optimal control problem has to be solved.

In this paper we summarize results of an ongoing study (Nagy et al., 2000; Bock et al., 2000b; Allgöwer et al., 2000) showing the practical applicability of NMPC to medium/high dimensional processes. We consider the control of a high purity distillation column using NMPC. In contrast to (Nagy et al., 2000; Bock et al., 2000b) we consider the output feedback case in this paper.

Our goal is to outline the key components for real-time application of NMPC. The conclusion is that a successful application of NMPC is possible even nowadays, if a combination of special dynamic optimization strategies (Bock et al., 2000b; Biegler, 2000) and NMPC schemes with reduced online computational load (Chen and Allgöwer, 1998; De Nicolao et al., 1996) is used.

The paper is organized as follows: In the first section, we review NMPC strategies that require reduced computational load. In the second section, one specially tailored dynamic optimization strategy for the solution of the occurring optimal control problems is described. Finally, the control of a high-purity distillation column is considered.

Nonlinear Model Predictive Control

In Figure 1 the general principle of model predictive control is shown. For simplicity of exposition, we assume that the control and prediction horizon have the same length. Based on measurements obtained at time t, the controller predicts the future dynamic behavior of the system over a control horizon T_c and determines the manipulated input such that a predetermined open-loop performance objective functional is optimized. In order to incorporate some feedback mechanism, the open-loop manipulated input function obtained is implemented only until the next measurement becomes available. We assume that this is the case every δ seconds (sampling time). Using the new measurement, at time $t+\delta$, the whole procedure—prediction and optimization—is repeated to find a new input function.

*findeise@ist.uni-stuttgart.de
†allgower@ist.uni-stuttgart.de
‡moritz.diehl@iwr.uni-heidelberg.de
§bock@iwr.uni-heidelberg.de
¶schloeder@iwr.uni-heidelberg.de
‖znagy@chem.ubbcluj.ro

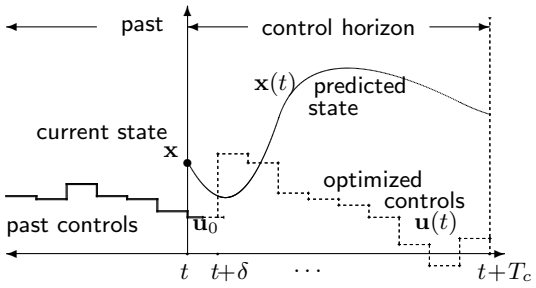


Figure 1: Principle of model predictive control.

Mathematical Formulation of NMPC

We assume that the system is given by nonlinear index-one differential algebraic equations (DAE) of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{z}(t), \mathbf{u}(t)), \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{1a}$$

$$\mathbf{0} = \mathbf{g}(\mathbf{x}(t), \mathbf{z}(t), \mathbf{u}(t)), \tag{1b}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ denotes the differential variables, $\mathbf{z}(t) \in \mathbb{R}^p$ the algebraic variables and $\mathbf{u}(t) \in \mathbb{R}^m$ the inputs. The control objective is to stabilize this system around a given setpoint, denoted by $(\mathbf{x}_s, \mathbf{z}_s, \mathbf{u}_s)$, while satisfying constraints on the input and states of the form:

$$\mathbf{h}(\mathbf{x}, \mathbf{z}, \mathbf{u}) \geq \mathbf{0}. \tag{2}$$

In the simplest case the constraints are box constraints, i.e. $\mathbf{u}_{min} \leq \mathbf{u} \leq \mathbf{u}_{max}$, $\mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max}$. The control is given by the following open-loop optimization problem that is solved at every sampling instant:

$$\min_{\bar{\mathbf{u}}(\cdot)} J(\bar{\mathbf{u}}(\cdot); \mathbf{x}(t)) \tag{3a}$$

with:

$$J(\bar{\mathbf{u}}(\cdot); \mathbf{x}(t)) = \int_t^{t+T_c} F(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau)) d\tau \tag{3b}$$

subject to:

$$\dot{\bar{\mathbf{x}}}(\tau) = \mathbf{f}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{z}}(\tau), \bar{\mathbf{u}}(\tau)), \quad \bar{\mathbf{x}}(t) = \mathbf{x}(t) \tag{3c}$$

$$\mathbf{0} = \mathbf{g}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{z}}(\tau), \bar{\mathbf{u}}(\tau)) \tag{3d}$$

$$\mathbf{h}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{z}}(\tau), \bar{\mathbf{u}}(\tau)) \geq \mathbf{0} \quad \tau \in [t, t + T_c]. \tag{3e}$$

Internal controller variables are denoted by a bar. The function F, often called stage cost function, specifies the desired control performance. Often, F is chosen as quadratic in x and u: $F(\mathbf{x}, \mathbf{u}) = (\mathbf{x} - \mathbf{x}_s)^T Q (\mathbf{x} - \mathbf{x}_s) + (\mathbf{u} - \mathbf{u}_s)^T R (\mathbf{u} - \mathbf{u}_s)$. The system input during the sampling time δ is given by the optimal input resulting from the solution of the open-loop control problem at time t: $\mathbf{u}(\tau) = \bar{\mathbf{u}}^\star(\tau)$, $\tau \in [t, t + \delta)$.

Efficient Formulation of the NMPC Problem

While the NMPC formulation described above can be applied straightforwardly in practice, in general nothing can be said about stability of the closed loop and performance. One way to achieve good closed-loop performance and stability is the use of an *infinite* horizon length (Keerthi and Gilbert, 1988), i.e. T_c is set to ∞. However, for an infinite horizon, the resulting nonlinear program (NLP) is in practice not solvable in finite time. If finite prediction and control horizons are used, the closed-loop system trajectories differ from the predicted open-loop ones. As a consequence it is not clear how the resulting performance is related to the "optimal" infinite horizon cost and whether the closed-loop is stable.

To allow an efficient solution of the resulting open-loop optimal control problem while guaranteeing stability and good performance several NMPC schemes have been proposed (Chen and Allgöwer, 1998; De Nicolao et al., 1996). These methods lead to a similar open-loop optimization problem as Equation 3a, however the cost function, Equation 3b, is augmented by a terminal penalty term $E_s(\cdot)$

$$J(\bar{\mathbf{u}}(\cdot); \mathbf{x}(t)) = \int_t^{t+T_c} F(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau)) d\tau + E_s(\bar{\mathbf{x}}(t + T_c)) \tag{4}$$

and the following final region constraint is added

$$r(\bar{\mathbf{x}}(t + T_c)) \geq 0. \tag{5}$$

Roughly speaking the terminal state penalty term E_s gives an (upper bound) estimate of the infinite horizon cost and thus approximately recovers the infinite horizon. For this approximation however, the final predicted state has to be restricted to a predetermined region given by the terminal region constraint r. Detailed descriptions of these approaches and how to obtain E_s and r can be found in Allgöwer et al. (2000); Chen and Allgöwer (1998); Findeisen and Allgöwer (2000); De Nicolao et al. (2000). The computational advantage of the described schemes lies in the fact, that shorter horizons can be used, while not jeopardizing performance and stability. We propose to use this kind of NMPC schemes in combination with specially tailored dynamic optimization strategies as outlined in the next section.

Efficient Solution of NMPC Problems

An numerically efficient solution of the NMPC optimization problem should: 1) take advantage of the special problem structure of the open loop optimization problem, 2) reuse as much information as possible from the previous sampling interval in the current sampling interval. One dynamic optimization scheme that can be adapted to provide all these properties is the so-called direct multiple shooting approach (Bock and Plitt, 1984).

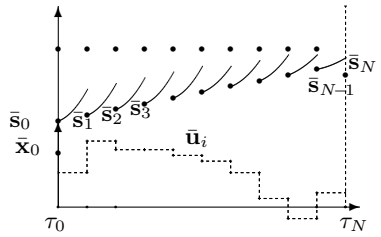


Figure 2: Principle of direct multiple shooting.

Direct Multiple Shooting for NMPC

In the following, the basic idea of direct multiple shooting is outlined. More can be found in Bock et al. (2000b). We assume, that the controls are parametrized as piecewise constant on each of the $N = \frac{T_c}{\delta}$ so-called *multiple shooting intervals*, i.e. $\bar{\mathbf{u}}(\tau) = \bar{\mathbf{u}}_i$ for $\tau \in [\tau_i, \tau_{i+1})$, $\tau_i = t + i\delta$. The DAE solution is decoupled on these intervals by considering the initial values $\bar{\mathbf{s}}_i^x$ and $\bar{\mathbf{s}}_i^z$ of differential and algebraic states at the times τ_i as additional optimization variables (compare Figure 2). The solution of the "decoupled" initial value problems obeys the following *relaxed* DAE formulation on the intervals $[\tau_i, \tau_{i+1})$:

$$\dot{\bar{\mathbf{x}}}_i(\tau) = \mathbf{f}(\bar{\mathbf{x}}_i(\tau), \bar{\mathbf{z}}_i(\tau), \bar{\mathbf{u}}_i) \tag{6a}$$

$$\mathbf{0} = \mathbf{g}(\bar{\mathbf{x}}_i(\tau), \bar{\mathbf{z}}_i(\tau), \bar{\mathbf{u}}_i) - \mathbf{g}(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i) \tag{6b}$$

$$\bar{\mathbf{x}}_i(\tau_i) = \bar{\mathbf{s}}_i^x, \qquad \bar{\mathbf{z}}_i(\tau_i) = \bar{\mathbf{s}}_i^z. \tag{6c}$$

The subtrahend in Equation 6b is introduced to allow an efficient DAE solution for initial values and controls $\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i$ that violate temporarily the consistency conditions $\mathbf{0} = \mathbf{g}(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i)$ (Bock et al., 2000a). The contribution of the integral cost term on $[\tau_i, \tau_{i+1})$ is determined by:

$$J_i(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i) = \int_{\tau_i}^{\tau_{i+1}} F(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}_i) d\tau.$$

The resulting large but structured NLP takes the form:

$$\min_{\bar{\mathbf{u}}_i, \bar{\mathbf{s}}_i} \sum_{i=0}^{N-1} J_i(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i) + E_s(\bar{\mathbf{s}}_N^x) \tag{7}$$

subject to:

$$\bar{\mathbf{s}}_0^x = x(t), \tag{8a}$$

$$\bar{\mathbf{s}}_{i+1}^x = \bar{\mathbf{x}}_i(\tau_{i+1}), \qquad i = 0, 1, \ldots, N-1, \tag{8b}$$

$$\mathbf{0} = \mathbf{g}(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i), \qquad i = 0, 1, \ldots, N-1, \tag{8c}$$

$$\mathbf{h}(\bar{\mathbf{s}}_i^x, \bar{\mathbf{s}}_i^z, \bar{\mathbf{u}}_i) \geq \mathbf{0}, \qquad i = 0, 1, \ldots, N-1, \tag{8d}$$

$$r_s(\bar{\mathbf{s}}_N^x) \geq 0. \tag{8e}$$

This large structured NLP problem is solved by a specially tailored *partially reduced* SQP algorithm (see Bock et al. (2000a) for a detailed description). To further improve the solution time, one should taken into account that the optimization problems at consecutive sampling instants are quite similar. We propose to consider the following strategy to decrease the computation time:

Initial Value Embedding Strategy: Optimization problems at subsequent sampling instants differ only by different initial values $\mathbf{x}(t)$, that are imposed via the initial value constraint, Equation 8a: $\bar{\mathbf{s}}_0^x = \mathbf{x}(t)$. Accepting an initial violation of this constraint, the complete solution trajectory of the previous optimization problem can be used as an initial guess for the current problem. Furthermore, all problem functions, derivatives as well as an approximation of the Hessian matrix have already been found for this trajectory and can be used in the new problem, so that the first QP solution can be performed *without any additional DAE solution.*
The application of this strategy does improve the robustness and speed of the optimization algorithm significantly. More details can be found in Diehl et al. (2001a).

Example Process

We outlined two key components for a computationally feasible application of NMPC: 1) the use of efficient, tailored dynamic optimization algorithms and 2) the use of efficient NMPC formulations. In this section we utilize these components to show that a successful application of NMPC to a nontrivial process control example is feasible already nowadays.

High Purity Distillation

As an application example the control of a high purity binary distillation column for the separation of Methanol and n-Propanol is considered (see Figure 3). The binary mixture is fed into the column with flow rate F and molar feed composition x_F. Products are removed at the top and bottom of the column with concentrations x_B and x_D. The liquid flow rate L and the vapor flow rate V are the control inputs (L/V configuration). The control problem is to maintain the specifications on the product concentrations x_B and x_D despite disturbances in the feed flow F and the feed concentration x_F. It is assumed that only the temperatures on the 14^{th} tray and 28^{th} tray can be measured and that the disturbance quantities x_F and F are not measured.

System Models and State Estimation

For comparison, two models of different complexity for the prediction are used in the controller. Modeling of the distillation column under the assumptions of constant relative volatility, constant molar overflow, no pressure losses, no energy balances and hydrodynamics leads to a 42^{nd} order ODE model. The second model considered is a 164^{th} order model with 122 algebraic states and 42 differential states. A more detailed description is given in Nagy et al. (2000). The controller needs estimates of all differential states, as well as of the disturbances F

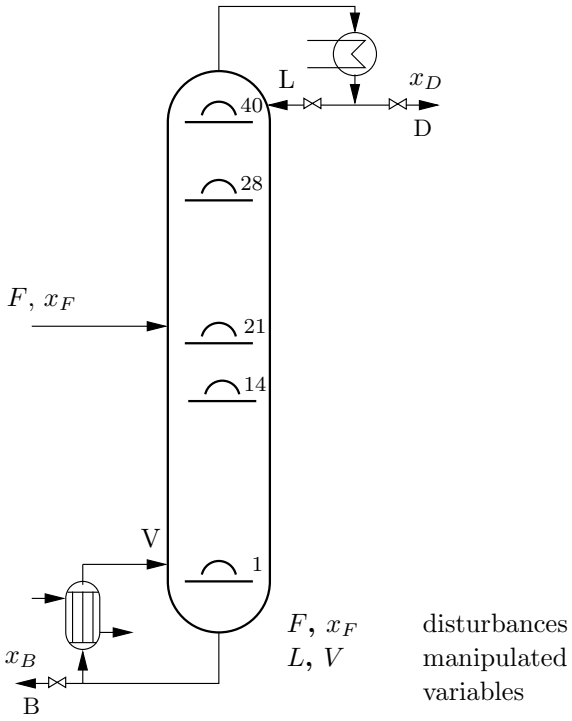


Figure 3: Considered distillation column.

model size	max	avrg
42	1.86s	0.89s
164	6.21s	2.48s

Table 1: Necessary CPU time for one sampling time.

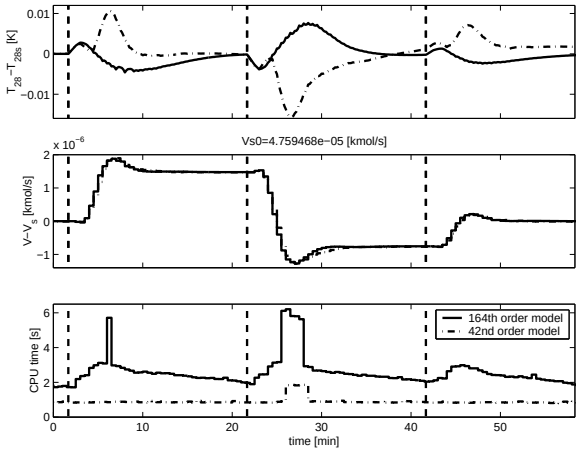


Figure 4: Behavior of the closed-loop.

and x_F. They are reconstructed from measurements of the temperatures T_{14}, T_{28} using an Extended Kalman Filter (EKF).

Controller Setup

As common in distillation control, the product concentrations x_B and x_D are not controlled directly, i.e. they do not appear directly in the cost function. Instead an inferential control scheme, which controls the deviation of the concentrations on tray 14 and 28 from the set-points, is used. Based on the estimates from the EKF in a first step, the system state at the next sampling instant is predicted. Using this state, the open-loop optimal control problem is solved. The resulting first input is implemented at the next control instant and the procedure is repeated. Note that this leads to a delay in the control scheme. This is necessary since the solution of the dynamic optimization problem cannot be obtained instantaneously. As NMPC scheme, the quasi-infinite horizon NMPC scheme for index–one DAE systems is applied (Findeisen and Allgöwer, 2000). A quadratic stage cost and a quadratic terminal penalty term are used. The choice of the weighting matrices and the derivation of the terminal region is described in (Nagy et al., 2000). For all simulations, the real plant is given by the 164^{th} order model. The 42^{nd} and 164^{th} order models are used for the controller predictions. The control input parameterization (controller sampling time) δ is $30s$, while the EKF is updated with the plant measurements every 10s. The control horizon T_c is fixed to 10 minutes ($N = 20$).

Performance and Computational Complexity

In Table 1, the necessary solution times for the input disturbance scenario as shown in Figure 4 are given[1]. One can see that the proposed strategy of combining NMPC schemes that require reduced computational time with a direct multiple shooting approach does lead to a rather low computational load. In our example the solution is easily feasible in the sampling time of 30s, even for the 164^{th} order model and a horizon length of $N = 20$. In Figure 4 the performance of the closed-loop for the different model sizes is compared. The temperatures T_{14} and T_{28} are kept in a narrow band, which is certainly more than satisfying. As shown, real-time application of NMPC is possible even for rather large models, if NMPC schemes with a low computational load and specialized optimization schemes are employed. Currently, the presented algorithms are experimentally applied to control a medium scale distillation column. Results will be presented in a forthcoming paper (Diehl et al., 2001b).

Conclusions

From an industrial/application point of view there is a strong demand to use NMPC schemes, since these methods allow to directly use (nonlinear) first principle models that are able to describe a wider range of operation than linear models can do. However, concern has been raised that NMPC cannot be applied in practice, since at every sampling instant a nonlinear optimization prob-

[1] All computations are carried out on a Compaq Alpha XP1000.

lem has to be solved. In this paper, we outlined the key components for a computationally feasible application of NMPC: The use of efficient NMPC schemes like quasi-infinite horizon NMPC in combination with specially tailored, efficient dynamic optimization techniques. Using these techniques, a successful application of NMPC even for high dimensional systems is feasible. This has been demonstrated considering the real-time control of a high purity distillation column.

References

Allgöwer, F., T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright, Nonlinear Predictive Control and Moving Horizon Estimation—An Introductory Overview, In Frank, P. M., editor, *Advances in Control, Highlights of ECC'99*, pages 391–449. Springer Verlag (1999).

Allgöwer, F., R. Findeisen, Z. Nagy, M. Diehl, H. G. Bock, and J. P. Schlöder, Efficient Nonlinear Model Predictive Control for Large Scale Constrained Processes, In *Proceedings of the Sixth International Conference on Methods and Models in Automation and Robotics*, pages 43–54. Miedzyzdroje, Poland (2000).

Biegler, L., Efficient Solution of Dynamic Optimization and NMPC Problems, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*. Birkäuser Verlag (2000).

Bock, H. G. and K. J. Plitt, A multiple shooting algorithm for direct solution of optimal control problems, In *Proc. 9th IFAC World Congress*, pages 1603–1608, Budapest (1984).

Bock, H. G., M. Diehl, D. Leineweber, and J. Schlöder, A Direct Multiple Shooting Method for Real-time Optimization of Nonlinear DAE Processes, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 245–268. Birkäuser Verlag (2000a).

Bock, H. G., M. Diehl, J. P. Schlöder, R. Findeisen, F. Allgöwer, and Z. Nagy, Real-time Optimization and Nonlinear Model Predictive Control of Processes Governed by Differential-algebraic Equations, In *Proc. Int. Symp. Adv. Control of Chemical Processes, ADCHEM*, pages 695–703, Pisa, Italy (2000b).

Chen, H. and F. Allgöwer, "A Quasi-infinite Horizon Nonlinear Model Predictive Control Scheme with Guaranteed Stability," *Automatica*, **34**(10), 1205–1218 (1998).

De Nicolao, G., L. Magni, and R. Scattolini, Stabilizing nonlinear receding horizon control via a nonquadratic terminal state penalty, In *Symposium on Control, Optimization and Supervision, CESA'96 IMACS Multiconference*, pages 185–187, Lille (1996).

De Nicolao, G., L. Magni, and R. Scattolini, Stability and Robustness of Nonlinear Receding Horizon Control, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 3–23. Birkäuser Verlag (2000).

Diehl, M., H. G. Bock, J. P. Schlöder, R. Findeisen, Z. Nagy, and F. Allgöwer, "Real-time Optimization and Nonlinear Model Predictive Control of Processes Governed by Differential-algebraic Equations," *J. Proc. Cont.* (2001a). Accepted for publication.

Diehl, M., R. Findeisen, S. Schwarzkopf, I. Uslu, F. Allgöwer, H. G. Bock, T. Bürner, E. D. Gilles, A. Kienle, J. P. Schlöder, and E. Stein, Real-Time Optimization of Large Scale Process Models: Nonlinear Model Predictive Control of a High Purity Distillation Column, In Groetschel, M., S. O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems: State of the Art*. Springer Verlag (2001b).

Findeisen, R. and F. Allgöwer, Nonlinear Model Predictive Control for Index-one DAE Systems, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*, pages 145–162. Birkäuser Verlag (2000).

Keerthi, S. S. and E. G. Gilbert, "Optimal Infinite-horizon Feedback Laws for a General Class of Constrained Discrete-time Systems: Stability and Moving-horizon Approximations," *J. Opt. Theory and Appl.*, **57**(2), 265–293 (1988).

Nagy, Z., R. Findeisen, M. Diehl, F. Allgöwer, H. G. Bock, S. Agachi, J. P. Schlöder, and D. Leineweber, Real-time Feasibility of Nonlinear Predictive Control for Large Scale Processes—a Case Study, In *Proc. of the American Control Conf.*, pages 4249–4254, Chicago. ACC (2000).

Qin, S. J. and T. A. Badgwell, An Overview of Industrial Model Predictive Control Technology, In Kantor, J. C., C. E. Garcia, and B. Carnahan, editors, *Fifth International Conference on Chemical Process Control—CPC V*, pages 232–256. American Institute of Chemical Engineers (1996).

Controller Design for Ventricular Assist Devices

Guruprasad A. Giridharan* and Mikhail Skliar†
Department of Chemical and Fuels Engineering
University of Utah
Salt Lake City, UT 84112

Abstract

In this paper, an integrated model of the human circulatory system with a brushless DC axial flow ventricular assist device (VAD) was developed. The resulting nonlinear hybrid model is then used to design the VAD feedback control system. The VAD controller is designed to maintain a physiologically motivated perfusion; it is tested using computer simulations of different scenarios ranging from the normal heart to the left heart asystole.

Keywords

Circulatory system model, Ventricular assist devices, Physiological control of rotary blood pumps

Introduction

Ventricular assist devices (VAD) have been in use for many years as a bridge to transplantation (Olsen, 1999) and hold a potential to become a long-term alternative to donor heart transplantation. VADs are mechanical support systems used in parallel with the failed heart to reduce the heart's workload.

Currently, a control system for continuous-flow VADs, which automatically responds to physiological cardiac demand, does not exist. In hospitals, the flow rate generated by the continuous flow VAD, such as the DeBakey pump, is selected manually by a physician or trained support personnel. Mobile patients can operate implanted continuous flow VADs in one of two ways: "automatic" or manual. During automatic control, the patient, following guidelines provided by the doctor, sets the desired pump rpm depending on the level of physical activity. The VAD controller automatically adjusts the current and voltage applied to the pump to achieve the desired rpm setpoint. A feedback system based on physiological measurements (such as pressures, flows, O_2 saturation, etc.) is currently not available. In manual mode, the patient directly adjusts the pump rpm by "twisting the knob" until the comfort level of perfusion is achieved.

A recent paper (Waters et al., 1999) is representative of the current state-of-the-art in developing an improved control of continuous flow VADs; in this paper, a PI controller was designed using a simple computer model of the circulatory system. The assumptions made by Waters et al. (1999) are unrealistic, including continuous flow throughout the circulatory system, no heart valves and linear correlation between pump generated pressure difference, ΔP, and pump voltage, current and rpm. Further research is clearly needed before a physiologically motivated continuous VAD control system can be developed for devices used in patients.

The selection of an adequate model, which avoids the overwhelming complexity of the full-scale CFD model, but retains all relevant characteristics of the circulatory system motivated our selection of the network-type circulation model. However, unlike the previous work by Waters et al. (1999), where the linear model with continuous flow throughout the system was assumed, we preserve such characteristics as nonlinearity, pulsativity and discontinuity due to the effects of the heart valves.

The selected control objective is to maintain the pressure difference between the left heart (LH) and the aorta close to the specified reference pressure. We show that this control objective leads to an adequate and physiologically motivated perfusion. At the same time, the simplicity of the objective allows for an implementation of simple control laws. In particular, a PI controller, developed to vary the VAD electrical current to minimize the difference between the reference and the actual differential pressure, results in a surprisingly good perfusion in vastly different clinical cases, ranging from the normal heart to a completely failed (asystolic) left heart.

During the development of the feedback system, we assumed that two implanted pressure sensors and an rpm sensor were available for the feedback. However, the need for implantable pressure sensors (the least reliable components) can potentially be eliminated by using readily available measurements of the pump rpm, voltage, and current to estimate the differential pressure between the left heart and the aorta.

Model Development

The model used in the controller design incorporates a model of the human circulatory system with a model of the continuous flow left ventricular assist device (LVAD).

Model of the Circulatory System

The Utah Circulation Model (UCM) is a network type model, which subdivides the human circulatory system into an arbitrary number of lumped parameter blocks, each characterized by its own resistance, compliance, pressure and volume of blood. In its simplest configuration, the UCM has eleven elements: 4 heart valves, and 7 blocks including the left heart, the right heart (RH), pulmonary and systemic circulation, the vena cava and

*gurug@eng.utah.edu
†mikhail.skliar@m.cc.utah.edu

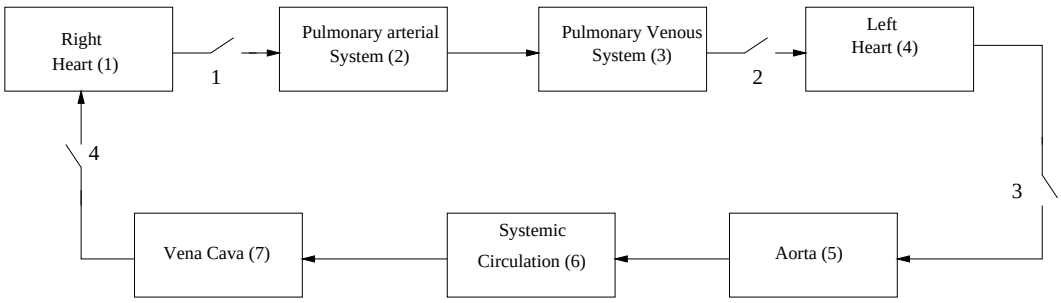


Figure 1: Schematic of the UCM.

the aorta, Figure 1. An arbitrary modeling detail can be achieved by increasing the number of blocks. The model was developed under the typical (McLeod, 1972) assumptions that blood is a Newtonian fluid, heart valves open and close instantaneously, and constant flow resistance in all blocks. Except for the heart blocks, the compliance of all other blocks remains constant.

Each block of the model is characterized by its resistance, R, to the flow, F, and its compliance, C, which characterizes the ability of a block to store a volume of blood, V. Two idealized elements, resistance and storage, are used to characterize each block. The storage element provides zero resistance to the flow, while the resistive element has zero volume. The resistance of the n-th block, R_n, is defined as a proportionality constant between pressure drop and blood flow across the block so that flow rates into and out of the block are given by

$$F_n^{in} = \frac{P_{n-1} - P_n}{R_{n-1}}, \quad F_n^{out} = \frac{P_n - P_{n+1}}{R_n}, \qquad (1)$$

where P_n is the pressure at the inlet of the n-th block, etc. The compliance, C_n, is defined as the ratio between the inlet pressure, P_n, to the stored volume of blood, V_n: $C_n = \frac{V_n}{P_n}$.

We further classify blocks as passive and active. Active blocks represent heart chambers; they are characterized by the varying compliance within each cardiac cycle. The rest of the blocks are referred to as passive. The varying compliance of the active blocks is responsible for the progression of a heartbeat. Figure 2 gives the typical value of the compliance of an active block used in the simulations.

The volume of blood in each block is described by a differential equation, which is an expression for the macroscopic material balance for a block.

The resistances and compliances will differ in different patients. In this work, typical values of Cs and Rs were assumed for all passive and active blocks. Parameters of the active blocks were adjusted to reflect different pathological conditions during the evaluation of the VAD control system performance under different scenarios.

The UCM includes four heart valves. Introducing valve conductance of the i-th valve, C_i^h, as an inverse of

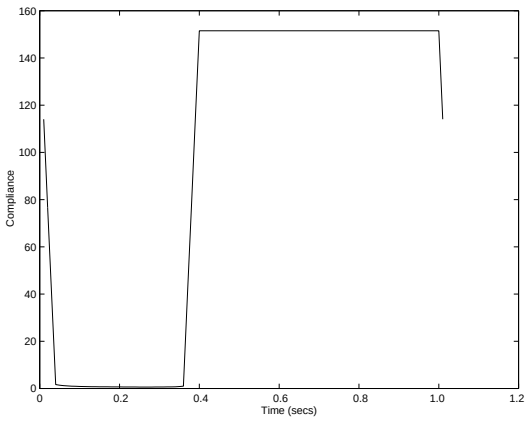


Figure 2: Typical compliance of an active block as a function of time.

the valve resistance, R_i^h, obtain: $C_i^h = \frac{1}{R_i^h} = c_i^h \delta_i$, $c_i^h = constant$, $i = \overline{1,4}$, where the Kronecker delta is a logical function of the differential pressure ΔP_i across the valve:

$$\delta_i = \begin{cases} 0 & \text{when } \Delta P_i \geq 0 \\ 1 & \text{when } \Delta P_i < 0 \end{cases}. \qquad (2)$$

The resulting model of the circulatory system now includes both dynamic and logical components, and is therefore, a hybrid system.

Model of the Axial Flow VAD

The integrated circulatory model includes a model of the axial flow LVAD, such as DeBakey LVAD, as an assist device. The LVAD is driven by a brushless DC motor. A typical brushless DC motor can be described by the following equations (Pillay and Krishnan, 1989):

$$J\frac{d\omega}{dt} = T_e - B\omega - T_p, \qquad (3)$$

where J is the inertia of the rotor, ω is the rotor speed in rads/sec, T_e is the motor torque, T_p is the load torque and B is the damping coefficient.

If the motor has a sinusoidal back EMF, the phase current also has a sinusoidal waveform. In this case

(Choi et al., 1997), the motor torque is related to the amplitude of the phase current, I, as

$$T_e = \frac{3}{2} K_B I, \tag{4}$$

where K_B is the constant of the back EMF. Following Choi et al. (1997), we adopt the following functional form for the correlation between the pump rotational speed, ω, generated flow rate, F_p, and the load torque:

$$T_p = a_0 \omega^3 + a_1 F_p \omega^2, \tag{5}$$

where a_0 and a_1 are correlation constants.

To obtain the LVAD model in the closed form, we need an additional correlation between the pump flow rate and the corresponding pressure rise across the pump, $\Delta P = P_5 - P_4$, and the rotational speed of the pump. Following Konishi et al. (1994), the following differential equation is used to describe the axial pump flow rate, and close the VAD model:

$$\frac{dF_p}{dt} = -\frac{b_0}{b_1} F_p - \frac{b_2}{b_1} \omega^2 + \frac{1}{b_1} \Delta P, \tag{6}$$

where b_0, b_1 and b_2 are experimental constants.

Model integration

In most cases, the assist device works in parallel with the natural heart taking blood from the LH and returning it to the aorta. Using the same block numbering as in Figure 1, we obtain the integrated model in the following form:

$$\begin{cases} \begin{bmatrix} \dot{\boldsymbol{v}} \\ \dot{F}_p \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ & \boldsymbol{A}_3 \end{bmatrix} \begin{bmatrix} \boldsymbol{v} \\ F_p \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ -\frac{b_2}{b_1} \end{bmatrix} \omega^2 \\ \dot{\omega} = -\frac{B}{J}\omega - \frac{a_1}{J} F_p \omega^2 - \frac{a_0}{J}\omega^3 + \frac{3K_B}{2J} I, \end{cases} \tag{7}$$

where $\boldsymbol{v} = \{V_i | i = \overline{1,7}\}$, $\boldsymbol{A}_2 = [0\ 0\ 0\ -1\ 1\ 0\ 0]^T$, $\boldsymbol{A}_3 = [0\ 0\ 0\ -\frac{1}{b_1 C_4}\ \frac{1}{b_1 C_5}\ 0\ 0\ -\frac{b_0}{b_1}]$, and $\boldsymbol{0}$ is a zero vector of an appropriate dimension. In diagonal form $\boldsymbol{A}_1 = \text{diags}\{\frac{1}{C_1(R_7+R_4^h)}\ \boldsymbol{0\ 0\ 0\ 0}\ \boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3\ \boldsymbol{0\ 0\ 0\ 0}\ \frac{1}{C_7(R_7+R_4^h)}\}$,

$$\boldsymbol{a}_1 = \begin{bmatrix} \frac{1}{C_1(R_1+R_1^h)} \\ \frac{1}{C_2 R_2} \\ \frac{1}{C_3(R_3+R_2^h)} \\ \frac{1}{C_4(R_4+R_3^h)} \\ \frac{1}{C_5 R_5} \\ \frac{1}{C_6 R_6} \end{bmatrix}, \boldsymbol{a}_3 = \begin{bmatrix} \frac{1}{C_2(R_1+R_1^h)} \\ \frac{1}{C_3 R_2} \\ \frac{1}{C_4(R_3+R_2^h)} \\ \frac{1}{C_5(R_4+R_3^h)} \\ \frac{1}{C_6 R_5} \\ \frac{1}{C_7 R_6} \end{bmatrix}, \tag{8}$$

where $R_i^h = \frac{1}{c_i^h \delta_i}$, $i = \overline{1,4}$ describes the resistance of the i-th heart valve, and the main diagonal

$$\boldsymbol{a}_2 = \begin{bmatrix} -\frac{1}{C_1}\left(\frac{1}{R_7+R_4^h} + \frac{1}{R_1+R_1^h}\right) \\ -\frac{1}{C_2}\left(\frac{1}{R_1+R_1^h} + \frac{1}{R_2}\right) \\ -\frac{1}{C_3}\left(\frac{1}{R_2} + \frac{1}{R_3+R_2^h}\right) \\ -\frac{1}{C_4}\left(\frac{1}{R_3+R_2^h} + \frac{1}{R_4+R_3^h}\right) \\ -\frac{1}{C_5}\left(\frac{1}{R_4+R_3^h} + \frac{1}{R_5}\right) \\ -\frac{1}{C_6}\left(\frac{1}{R_5} + \frac{1}{R_6}\right) \\ -\frac{1}{C_7}\left(\frac{1}{R_6} + \frac{1}{R_7+R_4^h}\right) \end{bmatrix}. \tag{9}$$

Equation 7 is the nonlinear, time-varying, hybrid model of the circulatory system with LVAD; its dimension depends on the number of blocks used to model the human circulatory system, and is equal to 9 in the present case. The pump current, I, is the manipulated variable.

In this work, we assume that the rotational speed of the pump, ω, and the pressure difference between the left heart and the aorta are directly measured. The rpm sensor can be integrated into the VAD design, as is the case with the DeBakey pump. However, measurements of differential pressure require an implantation of two pressure sensors, thus motivating an effort in developing "sensorless" VAD control systems, which relies on the estimation of ΔP from readily available measurements of pump current I, voltage V and rotational speed ω.

The compliances and resistances may differ from patient to patient. The variations with time within a patient can also be substantial. A limited number of identification schemes were previously proposed (Yu et al., 1996), which unfortunately require the implantation of additional pressure and flow sensors.

VAD Control

Control Objective and Design

The significance of the VAD control cannot be overstated. Though the design of a VAD itself is critical to the long-term success of the mechanical implant, the control of a VAD determines the confidence of doctors and patients in mechanically supported perfusion as a permanent solution and an alternative to the donor heart transplantation. The key requirement for the control system is the adaptation of the VAD generated flow to the changing physiological requirements of the patient.

Maintaining a reference differential pressure is known to be an effective way to correctly adapt the cardiac output to the changing requirements of the body. Such adaptation is possible because the vascular bed resistance can increase or decrease by a factor of 2 to 5 (Waters et al., 1999); since the flow is directly proportional to ΔP and is inversely proportional to the vascular bed resistance, maintaining a constant ΔP with changing bed

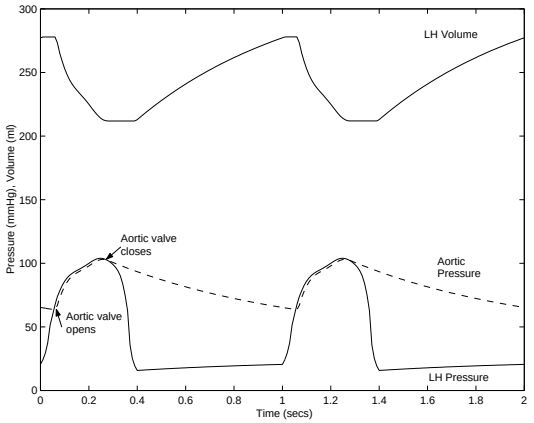


Figure 3: Volume and aortic pressure of a weakened heart.

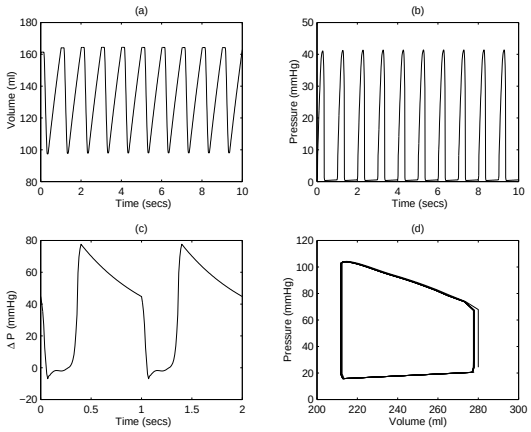


Figure 4: Characteristics of the weakened heart: (a) RH volume, (b) RH pressure, (c) ΔP between LH and aorta, (d) pressure-volume loop.

resistance can increase or decrease the flow rate by a factor of 2 to 5. A reference ΔP can be maintained by adjusting pump rpm within physiologically admissible limits despite changes in a patient's vascular resistance, stroke volume and pulse of the natural heart, all of which represent the response to natural regulatory mechanisms (Rao et al., 1999; Henson et al., 1995, 1994) to changing physiological cardiac output demands. By maintaining the prescribed ΔP we, in effect, synchronize the assist and natural perfusion, thus indirectly incorporating natural cardiovascular regulation into VAD control.

Controlling the ΔP also leads to relatively simple control algorithms and requires the implantation of only pressure sensors. An additional argument in favor of designing a ΔP controller is an observation that by controlling differential pressure we can ensure that pump rpm is maintained within limits dictated by physiological limitations related to possible collapse of the LH due to excessive suction, or back flow to the heart as a result of an inadequate pressure head developed by the VAD. The overall control problem can be formulated as the design of the feedback controller to regulate pump rpm within physiologically acceptable constraints, while minimizing the difference between the reference and the actual ΔP. Since pulsativity of the natural heart leads to the periodic changes in the ΔP, an additional objective is to keep oscillations of the pump rpm low, and, thus, increase pump life and the patient's comfort level. The formal expression of the control objective is to minimize the objective function J by selecting control input I subject to inequality constraints:

$$\min_I J = \int_0^t (\Delta P_r - \Delta P)^2 + r\dot{\omega}^2 dt, \qquad (10)$$

$$\omega_{min}(\boldsymbol{v}) \leq \omega(I) \leq \omega_{max}(\boldsymbol{v}), \qquad (11)$$

where $\boldsymbol{v}$ and ω must satisfy the system of nonlinear hybrid equations (7), and $r > 0$ is a user selectable weight.

The optimal solution to the formulated constrained quadratic optimal control problem for nonlinear hybrid systems is not known. Therefore, our approach is to select a fixed control structure followed by the optimization of the tuning parameters. In particular, for PI controller

$$I = K_P(\Delta P_r - \Delta P) + K_I \int_0^t (\Delta P_r - \Delta P)dt, \qquad (12)$$

the problem is to select the proportional and integral constants K_P and K_I, which minimize the objective function J. These constants were selected using an exhaustive, direct numerical search for the minimum of J for different weighting r until the desired trade-off between speed of response and rpm oscillations was achieved. The maximum value of r was limited to insure that the upper constraint in (11) is not violated.

Simulation Results

The controller performance was evaluated under different clinical conditions, ranging from the healthy heart to an asystolic LH. Figures 3 and 4 show characteristics of the weakened heart without assistance, indicating lower than normal stroke volume of approximately 65 ml and the aortic systolic and diastolic pressures are around 105/65 mmHg. The LH volume is considerably higher than normal. The RH pressure is also much higher at the normal 40 mmHg (Figure 4b), and is typical for RH pressure with the failing left heart. Though not shown in the figures for the weakened heart, the simulation predicts edema in the pulmonary circulation. Figure 4d shows the work done by the weakened heart, which is less than the work typical of the healthy heart.

The effect of the LVAD with the designed PI controller was tested with the weakened heart, assuming the pulse rate was 60 beats per minute, and the same LVAD pa-

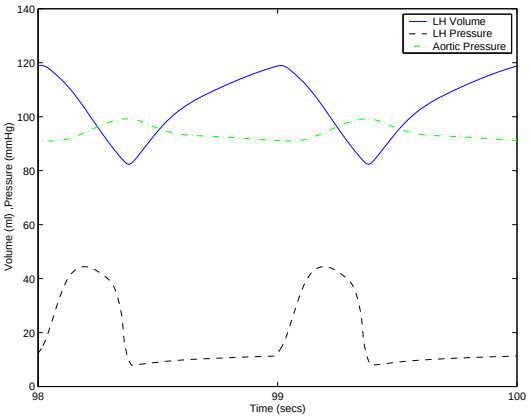


Figure 5: The LH characteristics and aortic pressure for a weakened heart with LVAD controlled by the PI controller.

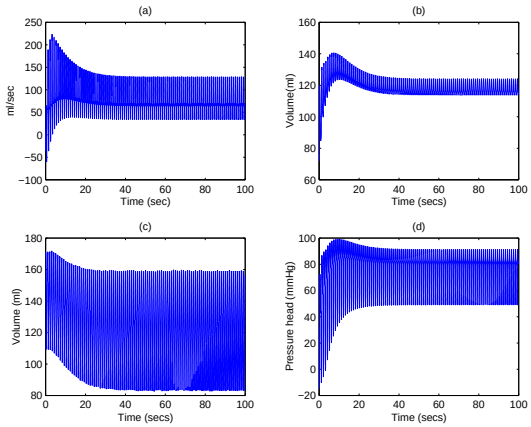


Figure 6: Weakened heart with LVAD assistance: (a) pump flow rate of the LVAD, (b) aorta volume, (c) RH volume and (d) ΔP between LH and aorta.

rameters as those used by Choi et al. (1997). At time $t = 0$, arbitrarily selected as the end of the diastole, the LVAD assistance was initiated with the reference differential pressure of 75 mmHg sent to the VAD controller. Initial VAD flow rate and rpm were set to zero, causing large initial back flow of blood to the left heart. Figures 5 and 6 show the results for the weakened heart assisted by the VAD with the designed controller. Figure 5 indicates a fairly constant aortic pressure 99/91 mmHg. The LH systolic and diastolic pressures are much closer to each other compared to a healthy heart with the LVAD. The volume of the LH with a VAD support reduces from 215/280 ml, observed without VAD assistance, to approximately 80/120 ml, which is in the normal range. The LH pressure is also reduced to about 45/12 mmHg. Though not shown in the figure, the RH pressure reduces to around 35/0 mmHg, which is within the normal range. The lung edema gradually reduces indicating an adequate perfusion. Figure 6 shows no pump back flow and the average pressure head of 75mmHg, which is the setpoint. After about 30 seconds the limit cycle is reached, at which time the rpm variations are reduced considerably, indicating inability of the weakened heart to produce high pressure variations.

Conclusions

The simulations show that maintaining an average pressure difference between the left heart and the aorta is an effective way to integrate the LVAD with the natural heart over a wide range of clinical conditions. The proposed control objective is effective in reflecting the physiological demands on perfusion, and simple enough to allow for simple control laws, which is a desirable feature because it simplifies FDA approval of a new device. However, the simplicity comes at a cost, since at least two implantable sensors are required. The feedback control without implantable sensors will necessitate the implementation of a more sophisticated control system, incorporating a ΔP estimator based on the measurements of the intrinsic pump characteristics. A further complication of adaptive control algorithms will be needed to account for inter- and intra- patient variability.

References

Choi, S., J. R. Boston, D. Thomas, and J. F. Antaki, Modeling and identification of an axial flow blood pump, In *Proceedings of the American Control Conference*, pages 3714–3715 (1997).

Henson, M. A., B. A. Ogunnaike, and J. S. Schwaber, "The Baroceptor Reflex: A Biological Control System with Applications in Chemical Process Control," *Ind. Eng. Chem. Res.*, **33**, 2453–2466 (1994).

Henson, M. A., B. A. Ogunnaike, and J. S. Schwaber, "Habituating Control Strategies for Process control," *AIChE J.*, **41**, 604–617 (1995).

Konishi, H., J. F. Antaki, J. R. Boston, J. P. Kerrigan, W. A. Mandarino, K. Yamaziki, M. Oe, P. Litwak, H. Borovetz, K. C. Butler, and R. L. Kormos, "Dynamic Systemic Vascular Resistance in Sheep Supported with Nimbus AxiPump," *ASAIO Journal*, **40**, M299–M302 (1994).

McLeod, J., "PHYSBE. A physiological simulation," *Simulation*, pages 152–156 (1972).

Olsen, D., "Sixth International Symposium for Rotary Blood Pumps," *Artificial Organs*, **26**, 475–476 (1999).

Pillay, P. and R. Krishnan, "Modeling, Simulation and Analysis of Permanent Magnet Motor Drives, Part 2: The Brushless DC Motor Drives," *IEEE Trans. Industrial Appl.*, **25** (1989).

Rao, R. R., B. Aufderheide, and B. W. Bequette, Multiple Model Predictive Control of hemodynamic variables, In *Proceedings of the American Control Conference*, pages 1253–1257 (1999).

Waters, T., P. Allaire, M. Adams, G. Bearnson, N. Wei, E. Hilton, M. Baloh, D. Olsen, and P. Khanwilkar, "Motor feedback physiological control for a continuous flow VAD," *Artificial Organs*, **23**, 480–486 (1999).

Yu, Y.-C., J. R. Boston, M. Simaan, and J. F. Antaki, Identification scheme for cardiovascular parameter estimation, In *13th IFAC World Congress*, pages 417–422 (1996).

Assessment of Performance for Single Loop Control Systems

Hsiao-Ping Huang* and Jyh-Cheng Jeng
Department of Chemical Engineering
National Taiwan University
Taipei 10617, Taiwan

Abstract

Assessment of performance in tracking set-point changes for single loop systems is presented. In contrast to the works that used stochastic performances, the current assessment uses a deterministic performance measure, i.e. the integration of absolute tracking errors (abbr. IAE). A benchmark system that has an open loop transfer function (abbr. OLTF) comprised of one LHP zero, one integrator, and dead time is established. This benchmark system is used to provide a goal of performance that the existing system can practically achieve. For assessing an existing control system, the performance of the system in terms of IAE is computed and an index that indicates the extent of achievement toward this benchmark system is computed. Evaluation of control based on this index can then be made. The model required in computing the aforementioned performance index can be obtained from an auto-tuning procedure.

Keywords

Single loop, Assessment, Benchmark system, IAE performance, Auto-tuning

Introduction

Assessment and monitoring of control systems with stochastic performance has been an active area of research for the last decade(Harris, 1989; Stanfelj et al., 1993; Harris et al., 1996; Qin, 1998; Harris et al., 1999; Leung and Romagnoli, 2000). The developments of research works have been focused on formulating the performance in terms of variance of the systems which are disturbed by stochastic inputs. As a result, with few exceptions, for example: Leung and Romagnoli (2000), controllers are implemented with discrete-time algorithms to pursuit minimum variance. On the other hand, the researches regarding assessment for deterministic performance have been reported, lately. For such assessments, technical developments have been focused on estimating maximum log modules ($L_{c,max}$) of closed-loops, (Chiang and Yu, 1993; Ju and Chiu, 1997), frequency responses (Kendra and Cinar, 1997), process characteristics(Piovoso et al., 1992), rise time (Åström et al., 1992), and settling time (Swanda and Seborg, 1999), etc. But, those works mentioned did not provided any implication regarding the best performance that an existing system can practically achieve.

In this paper, a benchmark system based on an IAE measure for an existing control system of single loop is presented. This benchmark system is to provide a practical goal of performance that the existing system can achieve. It has a loop transfer function comprised of one LHP zero, one integrator, and dead time. An index based on this benchmark system is thus presented to evaluate the existing system comparing with its control limit. This control limit is obtained based on what has been known about the process in terms of a model with specific dynamic order. Models to be required for this purpose can be obtained from an ATV experiment with relay feedback.

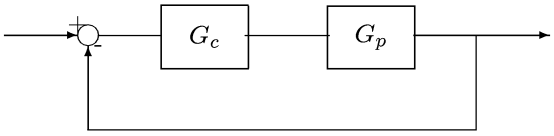


Figure 1: The conventional feedback control system.

Developing Benchmark System for Performance Assessment

According to the IMC design principle, an IMC equivalent controller in a single loop of Figure 1 is given by:

$$\bar{G}_c(s) = \frac{\bar{G}_{p,-}^{-1} F(s)}{1 - \bar{G}_{p,+} F(s)} \tag{1}$$

Where, $\bar{G}_p(s)$ designates the model for the process, and $\bar{G}_{p,-}$ and $\bar{G}_{p,+}$ designate the invertible and noninvertible parts of $\bar{G}_p$. The transfer function $F(s)$ is an IMC filter. For design purpose, the dynamics of a open loop process can in general be represented by a transfer function either of first-order-plus-dead-time (abbr. FOPDT) or of second-order-plus-dead-time (abbr. SOPDT) of the following:

FOPDT:

$$\bar{G}_p(s) = \frac{k_p e^{-\theta s}}{\tau s + 1} \tag{2}$$

SOPDT:

$$\bar{G}_p(s) = \frac{k_p (as + 1) e^{-\theta s}}{\tau^2 s^2 + 2\tau\zeta s + 1}; \qquad a \geq 0 \tag{3}$$

*Phone: 886-2-2363-8999 Fax: 886-2-2362-3935
Email: huanghpc@ccms.ntu.edu.tw

Notice that the SOPDT process with RHP zero can also be modeled with the one of Equation 3 by making use of the following approximation. Let the RHP zero is given in the form of $1-\beta s$. Then,

$$-\beta s+1 \approx e^{-2\beta s} \times (1+\beta s) \tag{4}$$

By using G_c in Equation 1 and G_p in Equations 2–3, the resulting open loop transfer function (abbr. OLTF) of the system becomes:

$$\begin{aligned} G_{OL}(s) &= G_c G_p(s) \\ &= \frac{G_{p,+}F(s)}{1-G_{p,+}F(s)} \\ &= \frac{e^{-\theta s}F(s)}{1-e^{-\theta s}F(s)} \end{aligned} \tag{5}$$

Thus, if G_c is implemented exactly with Equation 1, the performance of the system depends on the choice of $F(s)$. The limiting performance of such a system can be obtained from the result of Holt and Morari (1985) with a slight modification to take into account the additional dead time.

For conventional loops, G_c is used to be confined to have the following form:

$$G_c(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0} \tag{6}$$

To obtain $G_c(s)$ in the above form, $e^{-\theta s}$ in the denominator of Equation 1 should be replaced with a Pade' approximation of proper order. After introducing the same Pade' approximation into the denominator of Equation 5, it is easy to see that the OLTF becomes:

$$G_{OL} = H(s)\frac{e^{-\theta s}}{s} \tag{7}$$

where, $H(s)$ is considered a loop filter and its functional form varies with the choices of $F(s)$ and the Pade' approximation being used. But, in general, $H(s)$ consists of finite number of poles and zeros.

Due to the approximation that has been made for synthesizing $G_c(s)$, the performance limit set by the IMC will no longer be applicable to the system in Figure 1. Thus, the performance limit for a single loop system has to be obtained from a minimization process. The minimization process starts with an $H(s)$ of the following form:

$$H(s) = k_o\frac{\beta s+1}{\gamma s+1} \tag{8}$$

In other words, with an OLTF of the following:

$$G_{OL} = k_o\frac{\beta s+1}{\gamma s+1}\frac{e^{-\theta s}}{s} \tag{9}$$

If the performance in terms of IAE is used, then the procedure is to find the values of k_o and those of β and γ that minimize the following integral:

$$J^*_{IAE} = Min_{[k_o,\beta,\gamma]}\int_0^\infty |e(t)|dt \tag{10}$$

where $e(t)$ is given as the inverse transformation of $e(s)$, which is the tracking error of the system, i.e.:

$$e(t) = L^{-1}\left\{\frac{1}{1+G_{OL}}R(s)\right\} \tag{11}$$

This above optimization problem was solved numerically by simulations. The result turns out to be:

$$G^*_{OL} = \frac{k^*_o(1+\beta^* s)}{s}e^{-\theta s} \tag{12}$$

where,

$$k^*_o = \frac{0.76}{\theta}, \quad \beta^* = 0.47\theta, \quad \text{and} \gamma^* = 0 \tag{13}$$

Notice that G_{OL} should not have excess number of zeros than poles. Thus, for processes of Equations 2 and 3, the $H(s)$ in the form of Equation 8 is most appropriate for developing the benchmark system. Thus, with the optimization results given above, the benchmark system is selected as the one that has OLTF of the following:

$$G^*_{OL}(s) = \frac{0.76(1+0.47\theta s)}{\theta\ s}\ e^{-\theta s} \tag{14}$$

The IAE value of this benchmark system subjected to an unit step set-point change is found to be 1.38 θ. It is also found that this benchmark system has a gain margin of 2.11, and a phase margin of 64.4^o. The system with such margin values is considered to have acceptable stability robustness.

Based on the FOPDT model of Equation 2, and the the SOPDT processes of Equation 3, the controllers in the form of Equation 6 that yield the benchmark OLTF are given as follows:

Model or method	$\frac{IAE}{\theta}$	Remark
Benchmark system	1.38	$k^*_o = \frac{0.76}{\theta}$; $\beta^* = 0.47\theta$
Rovira(PI)*	1.93	FOPDT; $\tau = \theta$
Rovira(PID)*	1.52	FOPDT; $\tau = \theta$
Sung et al.(PID)†	2.06	$G_p(s) = \frac{1}{(s+1)^5}$
Sung et al.(PID)†	2.22	$G_p(s) = \frac{e^{-s}}{(9s^2+2.4s+1)(s+1)}$
Swanda and Seborg‡	2.0	PI control

*Smith and Corripio (1997, pg. 325)
†Sung et al. (1996)
‡Swanda and Seborg (1999)

Table 1: The minimum $\frac{IAE}{\theta}$ values for different control systems.

FOPDT:

$$G_{c,1}(s) = \left(\frac{0.76}{k_p\theta}\right)\frac{(\tau s+1)(0.47\theta s+1)}{s} \tag{15}$$

SOPDT:

$$G_{c,2}(s) = \left(\frac{0.76}{k_p\theta}\right)\frac{(\tau^2 s^2+2\tau\zeta s+1)(0.47\theta s+1)}{s(as+1)} \tag{16}$$

Obviously, these controllers are not physically realizable. For controllers to be realizable, one or two low pass filters with small time constants have to be introduced somewhere in G_c. The value of the resulting IAE will thus be degraded. But, the change is really too small to be considered. Thus, if $G_c(s)$ has not been confined to the conventional PID controllers, the minimum achievable IAE will be:

$$IAE^* = 1.38\theta \tag{17}$$

In Table 1, some IAE values of several systems are given. These systems include the benchmark one and some others, which have optimal controllers from different sources. It is to show none of these other systems from different sources has IAE less than the benchmark one.

This above equation can be adapted to apply to the case where $G_p(s)$ has $1-\beta s$ as a factor in the numerator (i.e. a RHP zero). In this case, the minimum IAE becomes:

$$IAE^* = 1.38\times(\theta+2\beta) \tag{18}$$

As an example, Consider a system that has open loop transfer function of the following:

$$G_cG_p(s) = \frac{k_o(1+as)(1-0.5s)e^{-s}}{(1+0.05s)s}$$

The minimum IAE of this closed loop system occurs when $k_o = 0.4$ and $a = 0.21$, and has a value of 2.8, which is about $1.38\times(1+2\times0.5)$ (i.e. 2.76).

Thus, assessment for a single loop control system that has $G_p(s)$ of Equation 2 or 3, will be targeting at the minimum IAE values of Equation 17 and 18.

Assessment for Control Based on IAE

In the previous section, it has been mentioned that the minimum achievable IAE value of a conventional feedback loop is 1.38 θ. In order to measure the achievement of an existing system toward this achievable target for set-point tracking, the following index is defined:

$$\Phi = \frac{1.38\theta}{\int_0^\infty |e(t)|dt} \tag{19}$$

Where, in the denominator, the IAE measure of the existing system for tracking step set-point change is used. This IAE measure can be obtained from experiment or from prediction based on a model for G_p. The index Φ, always less than one, is used to represent the ability of an existing system in eliminating the tracking errors. If Φ is close to one, it indicates that the system is near its performance limit.

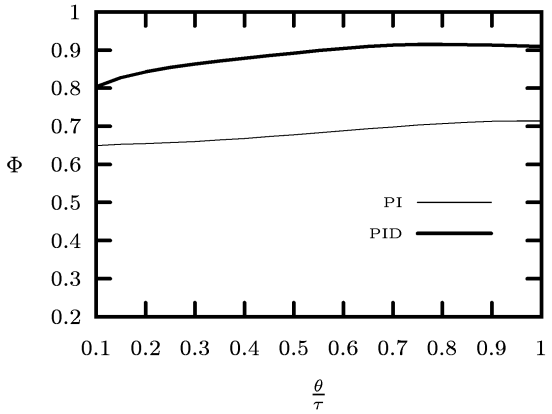


Figure 2: The Φ value for FOPDT process with Rovira's PI and PID controller.

To illustrate assessing with this presented index, first, the control of FOPDT process with PID controllers is considered. In order to be more inclusive for the results, the FOPDT model is normalized with dimensionless time unit. The normalized transfer function for the FOPDT model is:

$$\bar{G}_p(s) = \frac{k_pe^{-s}}{\bar{\tau}s+1} \tag{20}$$

where, $\bar{\tau} = \tau/\theta$. Then, for each $\bar{\tau}$, Rovira's tuning formula(Smith and Corripio, 1997, pg. 325) are used to tune the PI or PID controllers, which were claimed optimal for the IAE measure. As shown in Figure 2, the values of Φ resulting from such PI and PID control systems are given. It is thus found that, these optimal PID controllers give values of Φ higher than 0.8. On the other hand, those for optimal PI controllers, have values sitting between 0.65 and 0.75. Thus, as far as the control for FOPDT processes is concerned, the PID controller is a good choice.

Next, for dynamic systems of SOPDT, a normalized G_p with dimensionless time units are also considered for illustration:

$$\bar{\bar{G}}_p(s) = \frac{k_p(\bar{a}s+1)e^{-s}}{(\bar{\tau}^2s^2+2\bar{\tau}\zeta s+1)} \tag{21}$$

where, $\bar{a}$ designates $\frac{a}{\theta}$, and $\bar{\tau}$ designates $\frac{\tau}{\theta}$.

The normalized transfer function shows that $\bar{\tau}$ and ζ can be used to characterize the dynamic behaviors of such a process. Thus, PID control systems for G_p with different $\bar{\tau}$ and ζ are used for illustration. For PID control of SOPDT processes, tuning rules of Sung et al. (1996) are used to compute the values of Φ. The results

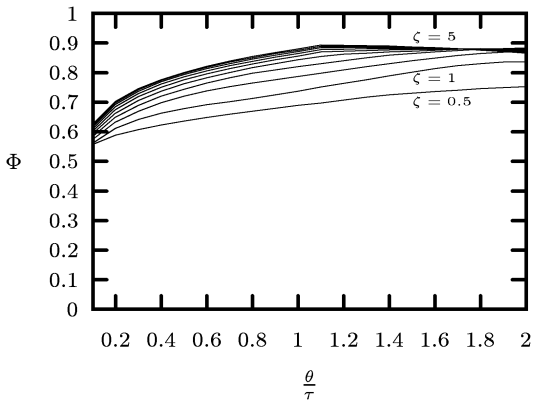


Figure 3: The Φ value for SOPDT process with PID controller given by Sung et al.

are given in Figure 3, where Φ is plotted along the values of θ/τ in the range between 0.1 and 2. The damping factor ζ has been used as a parameter. The computed Φ for such systems indicates that such PID controllers are more close to the limit for well overdamped G_p. In this figure, it is also observed that the Φ values computed based on the IAE^* from SOPDT models are lower than those from FOPDT models. This does not imply , for control purpose, an FOPDT model is superior to the SOPDT one. Instead, it indicates that with more knowledge about the dynamics, the control performance would have more stringent limit, and, if achievable, the performance would be superior to those with simpler (such as FOPDT) models.

In general, a $G_p(s)$ of high order may have an FOPDT and an SOPDT models at the same time as approximations of different accuracy. One may be questioned which type of models to be used. To justify, let θ_1 and θ_2 designate two apparent dead times in the FOPDT and SOPDT models, respectively. In general, $\theta_2 \leq \theta_1$. Two values of Φ will be resulted:

$$\Phi_1 = \frac{1.38\theta_1}{\int_0^\infty |e(t)|dt}$$
$$\Phi_2 = \frac{1.38\theta_2}{\int_0^\infty |e(t)|dt}$$

Then, we shall have:

$$\Phi_2 = (1-\eta) \times \Phi_1 \tag{22}$$

where,

$$\eta = \frac{\theta_1 - \theta_2}{\theta_1} \tag{23}$$

The value of η implies the portion of apparent dead time, from this FOPDT model, that can be reduced by the controllers based on SOPDT model. Thus, if the value of η is too high, that means FOPDT model is not sufficient for designing good control system.

Assessment with Relay Feedback Experiments

The information needed for assessing the control will be a model of FOPDT or of SOPDT that has an apparent dead time. To obtain this model, a relay feedback system can be used. The use of relay feedback has advantages on a few aspects. The most important one is that the control loop is still operated under closed loop and, hence, is still under control.

The relay feedback system is the same one as has been used in the ATV test of Åström and Hagglund (1984). The experiment consists of two stages. In the first stage, the system is perturbed with a bias to the output of the relay and wait until the system to appear constant cycling at the output. At this time, the cycling period (designated as P) and the amplitude (designated as a) of the cycles are measured to calculate the process gain:

$$k_p = \frac{\int_{t_0}^{t_0+P} y(t)dt}{\int_{t_0}^{t_0+P} u(t)dt} \tag{24}$$

Then, in the second stage, the bias to the relay is set to zero and wait again until the system appear constant cycling again. The period as well as the amplitude of the cycles are measured for estimating the other dynamic parameters. With these data obtained on-line, parameter estimations are carried out. The procedures and the algorithms for these estimations can be found elsewhere(Huang et al., 1996, 2000).

Thus, with the estimated model, we can calculate the predicted IAE for a set-point change by simulating the following with computer, i.e.:

$$e(s) = \frac{1}{1 + G_c(s)\bar{G}_p(s)} \frac{1}{s} \tag{25}$$

Predictions of IAE via this identification procedures and simulations have been carried out over several example processes of high order dynamics, and pretty close results are obtained. In other words, the computation of Φ for assessment can be performed with an ATV test mentioned above.

Conclusions

For a single loop control, a benchmark system that aims at minimuming IAE for set-point change has been established. This benchmark system system provides a performance limit for all feasible controllers in the form of Equation 6. It has an open loop transfer function (abbr. OLTF) comprised of one integrator, one simple lead, and dead time. For assessing an existing system, an index for indicating the extent of achievement toward this benchmark system is presented. This index is closely associated with the knowledge of dynamics being available, and the knowledge is usually contained in models of

different orders. To obtain these models an auto-tuning experiment with relay feedback can be used. Technically, this assessment reveals how close the existing system is to the benchmark one, and, if the knowledge (model) for design is sufficient.

References

Åström, K. J. and T. Hagglund, "Automatic Tuning of Simple Regulators with Specification on Phase Amplitude Margins," *Automatica*, **20**, 645–651 (1984).

Åström, K. J., C. C. Hang, P. Persson, and W. K. Ho, "Towards Intelligent PID Control," *Automatica*, **28**, 1–9 (1992).

Chiang, R. C. and C. C. Yu, "Monitoring Procedure for Intelligent Control: On-Line Identification of Maximum Closed-Loop Log Modulus," *Ind. Eng. Chem. Res.*, **32**, 90–99 (1993).

Harris, T. J., F. Boudreau, and J. F. MacGregor, "Performance Assessment of Multivariable Feedback Controllers," *Automatica*, **32**(11), 1505–1518 (1996).

Harris, T. J., C. T. Seppala, and L. D. Desborough, "A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems," *J. Proc. Cont.*, **9**, 1–17 (1999).

Harris, T. J., "Assessment of Control Loop Performance," *Can. J. Chem. Eng.*, **67**(10), 856–861 (1989).

Holt, B. R. and M. Morari, "Design of Resilient Processing Plants—VI. The Effect of Right-Half-Plane Zeros on Dynamic Resilience," *Chem. Eng. Sci.*, **40**(1), 59–74 (1985).

Huang, H. P., C. L. Chen, C. W. Lai, and G. B. Wang, "Autotuning for Model-Based PID Controllers," *AIChE J.*, **42**(9), 2687–2691 (1996).

Huang, H. P., M. W. Lee, and I. L. Chien, "Identification of Transfer Function Models from The Relay Feedback Test," *Chem. Eng. Commun.*, **180**, 231–253 (2000).

Ju, J. and M. S. Chiu, "Relay-Based On-Line Monitoring Procedures for 2×2 and 3×3 Multiloop Control Systems," *Ind. Eng. Chem. Res.*, **36**, 2225–2230 (1997).

Kendra, S. J. and A. Cinar, "Controller Performance Assessment by Frequency Domain Techniques," *J. Proc. Cont.*, **7**(3), 181–194 (1997).

Leung, D. and J. Romagnoli, "Real-Time MPC Supervisory System," *Comput. Chem. Eng.*, **24**, 285–290 (2000).

Piovoso, M. J., K. A. Kosanovich, and R. K. Pearson, Monitoring Process Performance in Real Time, In *Proceedings of the American Control Conference*, pages 24–26, Chicago, IL (1992).

Qin, S. J., "Control Performance Monitoring—A Review and Assessment," *Comput. Chem. Eng.*, **23**, 173–186 (1998).

Smith, C. A. and A. B. Corripio, *Principles and Practice of Automatic Process Control*. John Wiley & Sons, 2nd edition (1997).

Stanfelj, N., T. E. Marlin, and J. F. MacGregor, "Monitoring and Diagnosing Process Control Performance: The Single-Loop Case," *Ind. Eng. Chem. Res.*, **32**, 301–314 (1993).

Sung, S. W., O. Jungmin, I. B. Lee, and S. H. Yi, "Automatic Tuning of PID Controller Using Second-Order Plus Time Delay Model," *J. Chem. Eng. Japan*, **29**(6), 990–999 (1996).

Swanda, A. P. and D. E. Seborg, Controller Performance Assessment Based on Setpoint Response Data, In *Proceedings of the American Control Conference*, pages 3863–3867, San Diego, California (1999).

Feedback Control of Stable, Non-minimum-phase, Nonlinear Processes

Joshua M. Kanter* and Warren D. Seider†
Department of Chemical Engineering
University of Pennsylvania
Philadelphia, PA 19104

Masoud Soroush‡
Department of Chemical Engineering
Drexel University
Philadelphia, PA 19104

Abstract

A nonlinear control law is presented for stable, multiple-input, multiple-output processes, whether their delay-free part is minimum- or non-minimum-phase. It is derived by exploiting the connections between continuous-time model-predictive control and input-output linearization. The differential-geometric, control law induces a linear closed-loop response approximately. It has a few tunable parameters (one for each controlled output), and thus, is easily tuned.

Keywords

Nonlinear control, State feedback design, Input-output linearization, Non-minimum-phase systems, Model-based control

During the past decade, the problem of analytical (non-model-predictive) control of non-minimum-phase, nonlinear processes without deadtime has received considerable attention, leading to several solutions (Doyle III et al., 1996; Kravaris et al., 1998; Morari and Zafiriou, 1989; Isidori and Byrnes, 1990; Devasia et al., 1996; Isidori and Astolfi, 1992; van der Schaft, 1992; Isidori, 1995). These methods are applicable to a very small class of nonlinear processes, their application requires solving partial differential equations, or are not applicable to the class of general, nonlinear, stable, multivariable processes with time delays. More on advantages and disadvantages of these methods can be found in Kanter et al. (2000).

In the framework of model-predictive control, it is well known that large prediction horizons are needed for non-minimum-phase processes. For example, Hernández and Arkun (1992) developed a p-inverse (long prediction horizon) control law for nonlinear, single-input single-output (SISO), non-minimum-phase, discrete-time processes with arbitrary order and relative order.

An objective of this work is to derive a nonlinear control law for MIMO, stable, nonlinear, continuous-time processes, whether their delay-free part is non-minimum-phase or minimum-phase. This builds upon the single-input single-output controllers presented in Kanter et al. (2001), leading to the derivation of a continuous-time, differential-geometric control law that is approximately input-output linearizing (Allgöwer and Doyle III, 1998).

The paper is organized as follows. The scope of the study and some mathematical preliminaries are given in Section 2. Section 3 presents a method of nonlinear feedforward/state-feedback design. A nonlinear feedback control law with integral action is given in Section 4.

Scope and Mathematical Preliminaries

Consider the class of MIMO, nonlinear processes of the form:

$$\left.\begin{array}{rcll} \dfrac{d\bar{x}(t)}{dt} & = & f[\bar{x}(t), u(t)], & \bar{x}(0) = \bar{x}_0 \\ \bar{y}_i(t) & = & h_i[\bar{x}(t-\theta_i)] + d_i, & i = 1, \cdots, m \end{array}\right\} \quad (1)$$

where $\bar{x} = [\bar{x}_1 \cdots \bar{x}_n]^T \in X$ is the vector of the *process* state variables, $u = [u_1 \cdots u_m]^T \in U$ is the vector of manipulated inputs, $\bar{y} = [\bar{y}_1 \cdots \bar{y}_m]^T$ is the vector of *process* outputs, $\theta_1, \cdots, \theta_m$ are the measurement delays, $d = [d_1 \cdots d_m]^T \in D$ is the vector of constant unmeasured disturbances, $f(.,.)$ is a smooth vector field on $X \times U$, and $h_1(.), \cdots, h_m(.)$ are smooth functions on X. Here $X \subset \Re^n$ is a connected open set that includes $\bar{x}_{ss}$ and $\bar{x}_0$, $U \subset \Re^m$ is a connected open set that includes u_{ss}, and $D \subset \Re^m$ is a connected set, where $(\bar{x}_{ss}, u_{ss})$ denotes the nominal steady-state (equilibrium) pair of the process; that is, $f[\bar{x}_{ss}, u_{ss}] = 0$.

The system:

$$\left.\begin{array}{rcll} \dfrac{d\bar{x}(t)}{dt} & = & f[\bar{x}(t), u(t)], & \bar{x}(0) = \bar{x}_0 \\ \bar{y}_i^*(t) & = & h_i[\bar{x}(t)] + d_i, & i = 1, \cdots, m \end{array}\right\} \quad (2)$$

is referred to as the delay-free part of the process. The relative orders (degrees) of the controlled outputs $y_1, \cdots, y_m$ with respect to u are denoted by $r_1, \cdots, r_m$, respectively, where r_i is the smallest integer for which $\frac{d^{r_i}\bar{y}_i^*}{dt^{r_i}}$ explicitly depends on u for every $x \in X$ and every $u \in U$. The relative order (degree) of a controlled output y_i with respect to a manipulated input u_j is denoted by r_{ij} $(i = 1, \cdots, m,\ j = 1, \cdots, m)$, where r_{ij} is the smallest integer for which $\frac{d^{r_{ij}}\bar{y}_i^*}{dt^{r_{ij}}}$ explicitly depends on u_j for every $x \in X$ and every $u \in U$. The set-point and the set of acceptable set-point values are denoted by y_{sp} and Y, respectively, where $Y \subset \Re^m$ is a connected set.

The following assumptions are made:

(A1) For every $y_{sp} \in Y$ and every $d \in D$, there exists an equilibrium pair $(\bar{x}_{ss}, u_{ss}) \in X \times U$ that satisfies $y_{sp} - d = h(\bar{x}_{ss})$ and $f[\bar{x}_{ss}, u_{ss}] = 0$.

*kanterjm@seas.upenn.edu
†seider@seas.upenn.edu
‡masoud.soroush@coe.drexel.edu

(A2) The nominal steady-state (equilibrium) pair of the process, $(\bar{x}_{ss}, u_{ss})$, is hyperbolically stable; that is, all eigenvalues of the open-loop process evaluated at $(\bar{x}_{ss}, u_{ss})$ have negative real parts.

(A3) For a process in the form of (1), a model in the following form is available:

$$\left.\begin{array}{rcll} \dfrac{dx(t)}{dt} & = & f[x(t),u(t)], & x(0)=x_0 \\ y_i(t) & = & h_i[x(t-\theta_i)], & i=1,\cdots,m \end{array}\right\} \quad (3)$$

where $x=[x_1\cdots x_n]^T \in X$ is the vector of *model* state variables, and $y=[y_1,\cdots,y_m]^T$ is the vector of *model* outputs.

(A4) The relative orders, $r_1,\cdots,r_m$, are finite.

The following notation is used:

$$\begin{array}{rcl} h_i^1(x) & \doteq & \dfrac{dy_i^*}{dt} \\ & \vdots & \\ h_i^{r_i-1}(x) & \doteq & \dfrac{d^{r_i-1}y_i^*}{dt^{r_i-1}} \\ h_i^{r_i}(x,u) & \doteq & \dfrac{d^{r_i}y_i^*}{dt^{r_i}} \\ h_i^{r_i+1}(x,u^{(0)},u^{(1)}) & \doteq & \dfrac{d^{r_i+1}y_i^*}{dt^{r_i+1}} \\ & \vdots & \\ h_i^{p_i}(x,u^{(0)},u^{(1)},\ldots,u^{(p_i-r_i)}) & \doteq & \dfrac{d^{p_i}y_i^*}{dt^{p_i}} \end{array} \quad (4)$$

where $p_i \geq r_i$ and $u^{(\ell)}=d^\ell u/dt^\ell$.

Input-Output Linearization

For a process in the form of Equation 1, responses of the closed-loop process outputs are requested, having the linear form:

$$\begin{bmatrix} (\epsilon_1 D+1)^{r_1}\bar{y}_1(t+\theta_1) \\ \vdots \\ (\epsilon_m D+1)^{r_m}\bar{y}_m(t+\theta_m) \end{bmatrix} = y_{sp}, \quad (5)$$

where D is the differential operator (i.e., $D \doteq \frac{d}{dt}$), and $\epsilon_1,\cdots,\epsilon_m$ are positive, constant, adjustable parameters that set the speed of the response of the closed-loop process outputs $\bar{y}_1,\cdots,\bar{y}_m$ respectively. Substituting for the process output derivatives from the model in Equation 5, one obtains:

$$\begin{bmatrix} h_1(\bar{x})+\binom{r_1}{1}\epsilon_1 h_1^1(\bar{x})+\cdots+\binom{r_1}{r_1}\epsilon_1^{r_1}h_1^{r_1}(\bar{x},u) \\ \vdots \\ h_m(\bar{x})+\binom{r_m}{1}\epsilon_m h_m^1(\bar{x})+\cdots+\binom{r_m}{r_m}\epsilon_m^{r_m}h_m^{r_m}(\bar{x},u) \end{bmatrix} = y_{sp}-d \quad (6)$$

where

$$\binom{a}{b} \doteq \frac{a!}{b!(a-b)!}.$$

Under the assumption of the nonsingularity of the characteristic (decoupling) matrix:

$$\frac{\partial}{\partial u}\begin{bmatrix} h_1^{r_1}(\bar{x},u) \\ \vdots \\ h_m^{r_m}(\bar{x},u) \end{bmatrix}$$

on $X \times U$, Equation 6 represents a feedforward/state feedback. When the process delay-free part exhibits non-minimum-phase behavior, the input-output behavior of the closed-loop system under the feedforward/state feedback of Equation 6 is governed by the linear response of Equation 5, but the internal dynamics (unobservable modes) of the closed-loop system are unstable.

The dynamic feedforward/state feedback

$$\Phi_p(\bar{x},u,\mathcal{U}) = y_{sp}-d \quad (7)$$

where the ith component of $\Phi_p(\bar{x},u,\mathcal{U})$:

$$\begin{array}{c} [\Phi_p(\bar{x},u,\mathcal{U})]_i \doteq h_i(\bar{x})+\binom{p_i}{1}\epsilon_i h_i^1(\bar{x})+\cdots+ \\ \binom{p_i}{r_i-1}\epsilon_i^{r_i-1}h_i^{r_i-1}(\bar{x}) \\ +\binom{p_i}{r_i}\epsilon_i^{r_i}h_i^{r_i}(\bar{x},u)+\binom{p_i}{r_i+1}\epsilon_i^{r_i+1}h_i^{r_i+1}(\bar{x},u,u^{(1)}) \\ +\cdots+\binom{p_i}{p_i}\epsilon_i^{p_i}h_i^{p_i}(\bar{x},u,u^{(1)},\cdots,u^{(p_i-r_i)}) \end{array} \quad (8)$$

$$\mathcal{U}=\left[u^{(1)}\cdots u^{(\max[p_1-r_1,\cdots,p_m-r_m])}\right]^T, \quad p=[p_1\ \cdots\ p_m]$$

with

$$\partial\Phi_p/\partial\left[u_1^{\max(p_1-r_{11},\cdots,p_m-r_{m1})},\cdots,u_m^{\max(p_1-r_{1m},\cdots,p_m-r_{mm})}\right]^T$$

nonsingular, $\forall x \in X$, also induces a linear, closed-loop, output response of the form:

$$\begin{bmatrix} (\epsilon_1 D+1)^{p_1}\bar{y}_1(t+\theta_1) \\ \vdots \\ (\epsilon_m D+1)^{p_m}\bar{y}_m(t+\theta_m) \end{bmatrix} = y_{sp}, \quad (9)$$

Similarly, the dynamic feedforward/state feedback of Equation 7 cannot ensure asymptotic stability of the closed-loop system when the delay-free part of the process exhibits non-minimum-phase behavior. Consequently, an objective of this study is to design a feedback control law that ensures asymptotic stability of the closed-loop system, whether the delay-free part of the process is minimum- or non-minimum-phase.

Nonlinear Feedforward/State Feedback Design

Assume that for every $x \in X$, every $d \in D$, and every $y_{sp} \in Y$, the algebraic equation:

$$\phi_p(\bar{x}, u) = y_{sp} - d \tag{10}$$

where

$$\phi_p(\bar{x}, u) \doteq \Phi_p(\bar{x}, u, 0) \tag{11}$$

has a real root inside U for u, and that for every $\bar{x} \in X$ and every $u \in U$, $\frac{\partial \phi_p(\bar{x},u)}{\partial u}$ is nonsingular. The corresponding feedforward/state feedback that satisfies Equation 10 is denoted by

$$u = \Psi_p(\bar{x}, y_{sp} - d) \tag{12}$$

Note that the preceding feedforward/state feedback was obtained by setting all the time derivatives of u in Equation 7 to zero.

Theorem 1 *For a process in the form of Equation 1, the closed-loop system under the feedforward/state feedback of Equation 12 is asymptotically stable, if the following conditions hold:*

(a) The nominal equilibrium pair of the process, $(\bar{x}_{ss}, u_{ss})$, corresponding to y_{sp} and d, is hyperbolically stable.

(b) The tunable parameters $p_1, \cdots, p_m$ are chosen to be sufficiently large.

(c) The tunable parameters $\epsilon_1, \cdots, \epsilon_m$ are chosen such that for every $\ell = 1, \cdots, m$, all eigenvalues of $\epsilon_\ell J_{ol} \doteq \epsilon_\ell [\frac{\partial}{\partial x} f(x,u)]_{(\bar{x}_{ss}, u_{ss})}$ lie inside the unit circle centered at $(-1, 0j)$.

Furthermore, as $p_1, \cdots, p_m \longrightarrow \infty$, the state feedback places the n eigenvalues of the Jacobian of the closed-loop system evaluated at the nominal equilibrium pair at the n eigenvalues of the Jacobian of the open-loop process evaluated at the nominal equilibrium pair.

The proof can be found elsewhere (Kanter et al., 2000).

Condition (c) states that $\epsilon_1, \cdots, \epsilon_m$ should be chosen such that for every ϵ_ℓ ($\ell = 1, \cdots, m$) and for every λ_i ($i = 1, \cdots, n$), $|\epsilon_\ell \lambda_i + 1| < 1$, where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of J_{ol}. For an overdamped, stable process, $\epsilon_1, \cdots, \epsilon_m$ should be chosen such that for every ϵ_ℓ, $\ell = 1, \cdots, m$, and for every τ_i, $i = 1, \cdots, n$, $0 < \frac{\epsilon_\ell}{\tau_i} < 2$, where $\tau_1, \cdots, \tau_n$ are the open-loop time constants of the process. In other words, $\epsilon_1, \cdots, \epsilon_m$ should be chosen to be less than $2\tau_{min}$, where τ_{min} is the smallest time constant of the process [i.e., $\tau_{min} = min(\tau_1, \cdots, \tau_n)$].

Note that the feedforward/state feedback of Equation 12 does not induce the linear, closed-loop response of Equation 9, since in the derivation of the feedforward/state feedback the time derivatives of u were set to zero. The nonlinearity of the resulting delay-free output response is the price of ensuring closed-loop stability for processes with a non-minimum-phase delay-free part.

P	$\lambda_1(J_{cl})$	$\lambda_2(J_{cl})$	$\lambda_3(J_{cl})$
1	7.28	-20.00	-10.00
2	64.75	-13.19	-7.14
3	-4.34	-15.69+2.13i	-15.69-2.13i
4	-3.16	-11.18+3.25i	-11.18-3.25i
5	-2.60	-9.63+2.25i	-9.63-2.25i
6	-2.28	-8.76+0.96i	-8.76-0.96i
7	-2.10	-9.63	-6.74
10	-1.86	-9.97	-4.45
20	-1.78	-10.00	-2.38
30	-1.55	-10.00	-2.02
40	-1.31	-10.00	-2.00
50	-1.18	-9.70	-2.00

Table 1: Closed-loop eigenvalues of Example 1 for several values of $p_1 = p_2 = P$.

Example 1 *Consider a linear process without deadtime in state space form with*

$$A = \begin{bmatrix} -2 & 5 & -3 \\ 0 & -1 & 3 \\ 0 & 0 & -10 \end{bmatrix}, \; B = \begin{bmatrix} 2 & 5 \\ 5 & 0 \\ 2 & 22 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

It is non-minimum-phase (has a right-half plane [RHP] transmission zero at $s = 7.28$) and hyperbolically stable (has three left-half plane [LHP] eigenvalues at $s = -1$, $s = -2$ and $s = -10$). Its relative orders $r_1 = 1$ and $r_2 = 1$. With $\epsilon_1 = 0.1$ and $\epsilon_2 = 0.05$, the eigenvalues of $\epsilon_1 A$ are -0.2, -0.1 and -1, while the eigenvalues of $\epsilon_2 A$ are -0.5, -0.1 and -0.05. These eigenvalues are inside the unit circle centered at (-1,0i), so $\rho(\epsilon_1 A + I) = 0.9 < 1$ and $\rho(\epsilon_2 A + I) = 0.95 < 1$. The location of the closed-loop eigenvalues, $\lambda_1(J_{cl_p})$, $\lambda_2(J_{cl_p})$ and $\lambda_3(J_{cl_p})$, for several values of $p_1 = p_2 = P$ are given in Table 1; the closed-loop eigenvalues converge to the open-loop eigenvalues ($s = -1$, $s = -2$, $s = -10$), as $P \to \infty$.

Nonlinear Feedback Controller Design

The feedforward/state feedback of Equation 12 lacks integral action, and thus, it cannot induce offset-free response of the process output when process-model mismatch exists. To resolve this problem, the feedforward/state feedback of Equation 12 is implemented with a disturbance estimator, leading to the derivation of the feedback control law with integral action, given in Theorem 2.

Theorem 2 *For a process in the form of Equation 1 with incomplete state measurements, the closed-loop sys-*

tem under the error-feedback control law

$$\left.\begin{aligned} \frac{dx}{dt} &= f[x,u] \\ u &= \Psi_p\left[x, e + \left[\begin{array}{c} h_1[x(t-\theta_1)] \\ \vdots \\ h_m[x(t-\theta_m)] \end{array}\right]\right] \end{aligned}\right\} \quad (13)$$

where $e = y_{sp} - \bar{y}$, is asymptotically stable, if the following conditions hold:

(a) The nominal equilibrium pair of the process, $(\bar{x}_{ss}, u_{ss})$, corresponding to y_{sp} and d, is hyperbolically stable.

(b) The tunable parameters $p_1, \cdots, p_m$ are chosen to be sufficiently large.

(c) The tunable parameters $\epsilon_1, \cdots, \epsilon_m$ are chosen such that for every $\ell = 1, \cdots, m$, all eigenvalues of $\epsilon_\ell J_{ol} \doteq \epsilon_\ell [\frac{\partial}{\partial x} f(x,u)]_{(\bar{x}_{ss}, u_{ss})}$ lie inside the unit circle centered at $(-1, 0j)$.

Furthermore, the error-feedback control law of Equation 13 has integral action.

The proof can be found elsewhere (Kanter et al., 2000).

Acknowledgments

The first and third authors gratefully acknowledge partial support from the National Science Foundation through grant CTS-0101237. The second author would like to acknowledge financial support from the National Science Foundation through grant CTS-0101133.

References

Allgöwer, F. and F. J. Doyle III, Approximate I/O-Linearization of Nonlinear Systems, In Berber, R. and C. Kravaris, editors, *Nonlinear Model-Based Process Control*, NATO ASI Series, page 235. Kluwer, Dordrecht (1998).

Devasia, S., D. Chen, and B. Paden, "Nonlinear Inversion-based Output Tracking," *IEEE Trans. Auto. Cont.*, **41**, 930–942 (1996).

Doyle III, F. J., F. Allgöwer, and M. Morari, "A Normal Form Approach to Approximate Input-Output Linearization for Maximum Phase Nonlinear SISO Systems," *IEEE Trans. Auto. Cont.*, **41**, 305–309 (1996).

Hernández, E. and Y. Arkun, "Study of the Control-relevant Properties of Back-propagation Neural Network Models of Nonlinear Dynamical Systems," *Comput. Chem. Eng.*, **16**, 226–240 (1992).

Isidori, A. and A. Astolfi, "Disturbance Attenuation and H_∞ Control Via Measurement Feedback in Nonlinear Systems," *IEEE Trans. Auto. Cont.*, **37**, 1283–1293 (1992).

Isidori, A. and C. I. Byrnes, "Output Regulation of Nonlinear Systems," *IEEE Trans. Auto. Cont.*, **35**(2), 131–140 (1990).

Isidori, A., *Nonlinear Control Systems*. Springer-Verlag, New York, NY (1995).

Kanter, J., M. Soroush, and W. D. Seider, Nonlinear Feedback Control of Multivariable Non-minimum-phase Processes, Presented at AIChE Annual Meeting, Los Angeles, CA (2000).

Kanter, J., M. Soroush, and W. D. Seider, "Continuous-time, Nonlinear Feedback Control of Stable Processes," *I&EC Research*, **40**(9), 2069–2078 (2001).

Kravaris, C., M. Niemiec, R. Berber, and C. B. Brosilow, Nonlinear Model-based Control of Nonminimum-phase Processes, In Berber, R. and C. Kravaris, editors, *Nonlinear Model Based Process Control*. Kluwer, Dordrecht (1998).

Morari, M. and E. Zafiriou, *Robust Process Control*. Prentice-Hall, Englewood Cliffs, NJ (1989).

van der Schaft, A. J., "L_2-gain Analysis of Nonlinear Systems and Nonlinear State Feedback H_∞ Control," *IEEE Trans. Auto. Cont.*, **37**, 770–783 (1992).

Self-optimizing Control of a Large-scale Plant: The Tennessee Eastman Process

Truls Larsson*, Sigurd Skogestad† and Kristin Hestetun
Department of Chemical Engineering
Norwegian University of Science and Technology (NTNU)
N 7491 Trondheim, Norway

Abstract

The paper addresses the selection of controlled variables, that is, "what should we control". The concept of self-optimizing control provides a systematic tool for this, and in the paper we show how it may be applied to the Tennessee Eastman process which has a very large number of candidate variables. In the paper we present a systematic procedure for reducing the number of alternatives. One step is to eliminate variables which with constant setpoints result in large losses or infeasibility when there are disturbances (with the remaining degrees of freedom reoptimized).

Keywords

Process control, Plantwide control, Steady-state optimization, Control structure

Introduction

This paper addresses the selection of controlled variables for the Tenessee Eastman process. We base the selection on the concept of self-optimizing control using steady state models and steady state economics. "Self-optimizing control" is when an acceptable (economic) loss can be achieved using constant setpoints for the controlled variables, without the need to reoptimize when disturbances occur (Morari et al., 1980) (Skogestad, 2000). The constant setpoint policy is simple, but it will not be optimal (and thus have a positive loss) due to the following two factors

1. Disturbances, i.e. changes in (independent) variables and parameters compared to their nominal values, which cause the optimal setpoints to change.
2. Implementation errors, i.e. differences between the setpoints and the actual values of the controlled variables (e.g. due to measurement errors or poor control) (Skogestad, 2000).

The effect of these factors (the loss) depends on the choice of controlled variables, and the objective is to find a set of controlled variables for which the loss is acceptable.

(Downs and Vogel, 1993) introduced the Tennessee Eastman challenge problem at an AIChE meeting in 1990. The purpose was to supply the academics with a problem that contained many of the challenges that people in industry meet. There are eight components, including an inert (B) and a byproduct (F). The process has four feed streams (of A, D, E and A+C), one product stream (a mix of G and H) and one purge stream. The inert (B) enters in the A+C feedstream. The process has five major units; a reactor, a product condenser, a vapor-liquid separator, a recycle compressor and a product stripper, see Figure 1. There are 41 measurements and 12 manipulated variables. We here study the optimal operation of the base case (mode 1) with a given 50/50 product ratio between components G and H, and a given production rate. This plant has been studied by many authors, and it has been important for the development of plantwide control as a field. Many authors has used it to demonstrate their procedure for the design of a control system, e.g. (McAvoy and Ye, 1994), (Lyman and Georgakis, 1995), (Ricker, 1996), (Luyben et al., 1997), (Ng and Stephanopoulos, 1998), (Tyreus, 1999). To summarize, most authors do not control all the variables which are constrained at the optimum, thus they can not operate optimally in the nominal case. Most control reactor pressure, reactor level, reactor temperature and composition of B (inert) in reactor feed or in purge. It is common to control stripper temperature, separator temperature, and composition of C and/or A in reactor feed.

Stepwise Procedure for Self-optimizing Control

The main objective of operation, in addition to stabilization, is to optimize the economics of the operation subject, e.g in terms of minimizing the economic cost function J. To achieve truly optimal operation we would need a perfect model, we would need to measure all disturbances, and we would need to solve the resulting dynamic optimization problem on-line. This is unrealistic, and the question is if it is possible to find a simpler implementation which still operates satisfactorily (with an acceptable loss). More precisely, the *loss* L is defined as the difference between the actual value of the cost function and the truly optimal value, i.e. $L = J - J_{\text{opt}}$. *Self-optimizing control* is when we can achieve an acceptable economic loss with constant setpoint values for the controlled variables (without the need to reoptimize when disturbances occur). This sounds very simple, but it is not necessarily clear for a given problem what these controlled variables should be. The main objective of this

*Presently at Aker Offshore Partner, Stavanger, Norway
†skoge@chembio.ntnu.no

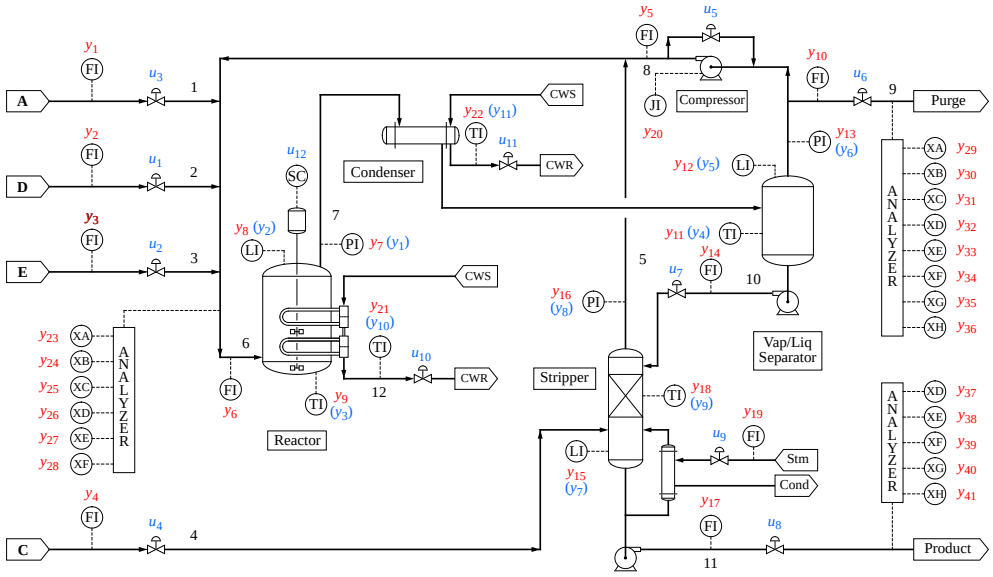


Figure 1: Tennessee Eastman process flowsheet.

paper is to search for a set of controlled variables which results in self-optimizing control for the Tennessee Eastman process. We will apply the stepwise procedure for self-optimizing control of (Skogestad, 2000). The main steps are

1. Degree of freedom analysis
2. Definition of optimal operation (cost and constraints)
3. Identification of important disturbances
4. Optimization
5. Identification of candidate controlled variables
6. Evaluation of the loss with constant setpoints for the alternative combinations of controlled variables (caused by disturbances or implementation errors)
7. Final evaluation and selection (including controllability analysis)

(Skogestad, 2000) applied this procedure to a reactor case and a distillation case, but in both cases there were only one unconstrained degree of freedom, so the evaluation in step 6 was managable. However, for the Tennessee Eastman process there are three unconstrained degrees of freedom, so it is necessary to do some more effort in step 5 to reduce the number of alternatives. We present below some general criteria that are useful for eliminating controlled variables.

Degrees of Freedom Analysis and Optimal Operation

The process has 12 manipulated variables and 41 measurements. All the manipulated variables have constraints and there are "output" constraints, including equality constraints on product quality and product rate. (Downs and Vogel, 1993) specify the economic cost J [\$/h] for the process, which is to be minimized. In words,

$$J = (unreacted\ feed) + (steam\ costs) + (compression\ costs)$$

The first term dominates the cost. An analysis, see Table 1, show that there are eight degrees of freedom at steady state which may be used for steady-state optimization. (Ricker, 1995) solved the optimization problem using the cost function of (Downs and Vogel, 1993) and gives a good explanation on what happens at the optimum. At the optimum there are five active constraints and these should be controlled to achieve optimal operation (at least nominally). This leaves three unconstrained degrees of freedom, which we want to select such that a constant setpoints policy results in an acceptable economic loss (self-optimizing control).

We consider the following three disturbanves:

- Disturbance 1: Change in A/C ratio in feed 4
- Disturbance 2: Change in %B (inert) in feed 4
- Throughput disturbances: Change in production rate by ±15 %.

	Manipulated variables	12
	D feed flow	
	E feed flow	
	A feed flow	
	A + C feed flow	
	Compressor recycle flow	
	Purge flow	
	Separator liquid flow	
	Stripper liquid product flow	
	Stripper steam flow	
	Reactor cooling water flow	
	Condenser cooling water flow	
	Agitator speed	
-	Levels without steady state effect	2
	Separator level	
	Stripper level	
-	Equality constraints	2
	Product quality	
	Production rate	
=	Degrees of freedom at steady state	8
-	Active constraints at the optimum	5
	Reactor pressure	
	Reactor level	
	Compressor recycle valve	
	Stripper steam valve	
	Agitator speed	
=	Unconstrained degrees of freedom	3

Table 1: Degrees of freedom and active constraints.

We use the same contraints (and safety margins) as given by (Ricker, 1995). The optimal (minimum) operation cost is 114.323 \$/h in the nominal case, 111.620 \$/h for disturbance 1, and 169.852 \$/h for disturbance 2. We define an "acceptable loss" to be 6 \$/h when summed over the disturbances.

Selection of Controlled Variables

What should we control? More precicely, we have 8 degrees of freedom at steady state, and we want to select 8 controlled variables which are to be controlled at constant setpoints. We can choose from 41 measurements and 12 manipulated variables, so there are 53 candidate variables. Even in the simplest case, where we do not consider variable combinations (such as differences, ratios, and so on), there are

$$\frac{53 \cdot 52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 \cdot 47 \cdot 46}{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 886 \cdot 10^6$$

possible combinations. It is clearly impossible to evaluate the loss with respect to disturbances and implementation errors for all these combinations. The following criteria are proposed to reduce the number of alternatives. Most of them are rather obvious, but nevertheless we find them useful.

Use Active Constraint Control

We choose to control the active constraints. This reduces the number of controlled variables to be selected from 8 to 3.

Eliminate Variables Related to Equality Constraints

The equality constraints must be satisfied, and if there are directly related variables then these must be eliminated from further consideration. The stripper liquid flow (product rate) is directly correlated with production rate which is specified (eliminates 1 manipulated variables and 1 directly related measurement).

Eliminate Variables with no Steady-state Effect

Two variables have no steady-state effect, namely stripper level and separator level (eliminates 2 measurements). (Of course, we need to measure and control these two variables for stabilization, but we are here concerned with the next control layer where the steady-state economics are the main concern).

Eliminate/Group Closely Related Variables

The controlled variables should be independent.

- Six of the remaining manipulated variables are measured (A feed, D feed, E feed, A+C feed, stripper liquid flow, purge flow) that is, there is a one to one correlation with a measurement (eliminates 5 mesurements).
- There is a only small differences between controlling the composition in the purge flow and in the reactor feed. We therefore eliminate reactor feed composition (eliminates 6 measurements)

Process Insight: Eliminate Further Candidates

Based on understanding of the process some further variables were excluded form the set of possible candidates for control (since they should *not* be kept constant): pressures in separator and stripper, condenser and reactor cooling water flowrates, reactor and separator cooling water outlet temperatures, and separator liquid flow. Finally the fractions of G in product and H in product should be equal (specified), so by keeping one of these fractions constant, we will idirectly specify their sum, which is optimally about 0.98. However, their sum cannot exceed 1.0, so taking into account the implementation error we should not keep G in product or H in product constant.

Eliminate Single Variables that Yield Infeasibility or Large Loss

The idea is to keep a single variable constant at its nominally optimal value, and evaluate the loss for (1) various disturbances (with the remaining degrees of freedom reoptimized), and (2) for the expected implementation

Variable	Nominal (constant)	Nearest feasible with disturb. 2
D feed [kg/h]	3657	3671
E feed [kg/h]	4440	4489
A+C feed [kscmh]	9.236	9.280
Purge rate [kscmh]	0.211	0.351

Table 2: Single variables with infeasibility

error. If operation is infeasible or the loss is large, then this variable is eliminated from further consideration.

Infeasibility. Keeping one of the following four manipulated variables constant results in infeasible operation for disturbance 2 (inert feed fraction): D feed flow, E feed flow , A+C feed flow (stream 4) and purge flow. This is independent on how the two remaining degrees of freedom are used, see Table 2.

Loss. We have now left 1 manipulated variable (A feed flow) and 17 measurements. Table 3 shows the loss (deviation above optimal value) for fixing one of these 18 variables at a time, and reoptimizing with respect to the remaning two degrees of freedom. The losses with constant A feed flow and constant reactor feedrate are totally unacceptable for disturbance 1 (eliminates 1 manipulated variable and 1 measurement), in fact, we could probably have eliminated these earlier based on process insight. The remaining 15 measurements yield reasonable losses. However, we have decided to eliminate variables with a loss larger than 6 \$/h when summed for the three disturbances. This eliminates the following 5 measurements: separator temperature, stripper temperature, B (inert) in purge, G in purge, and H in purge.

Eliminate Pairs of Constant Variables with Infeasibility or Large Loss

We are now left with 11 candidate measurements. that is, $(11 \cdot 10 \cdot 9)/(3 \cdot 2) = 165$ possible combinations of three variables. The next natural step is to proceed with keeping pairs of variables constant, and evaluate the loss with the remaining degree of freedom reoptimized. However, there are 55 combinations of pairs, so this does not result in a large reduction in the number of possibilities. We therefore choose to skip this step in the procedure.

Final Evaluation of Loss for Remaining Combinations

A quick screening indicates that one of the three controlled variables should be reactor temperature, which is the only remaining temperature among the candidate variables. A further evaluation shows that we should eliminate F (byproduct) in purge as a candidate variable, because the optimum is either very "sharp" in this variable, or optimal operation is achieved close to its maximum achievable value (see a typical plot in Figure 2). In either case, operation will be very sensitive to the implementation error for this variable.

Fixed variable	Dist.1	Dist.2	Throughput +15/-15%
A feed *	709.8	6.8	
Reactor feed*	53.5	0.5	
Recycle	0.0	0.8	0.5 / 0.3
Reactor T.	0.0	0.9	1.2 / 0.7
Sep T*	0.0	0.5	4.2 / 2.3
Stripper T*	0.1	0.3	4.3 / 2.3
Compr. Work	0.0	0.6	0.2 / 0.1
A in purge	0.0	0.7	0.4 / 0.2
B in purge*	0.0	7.4	3.1 / 1.6
C in purge	0.0	0.5	0.1 / 0.1
D in purge	0.0	0.0	0.2 / 0.1
E in purge	0.0	0.4	0.0 / 0.1
F in purge	0.0	0.5	0.0 / 0.0
G in purge*	0.0	0.4	4.1 / 2.2
H in purge*	0.0	0.4	4.2 / 2.2
D in product	0.0	0.1	0.2 / 0.1
E in product	0.0	0.0	1.2 / 0.7
F in product	0.0	1.5	1.4 / 0.8

Table 3: Loss [\$/h] with one variable fixed at its nominal optimal value and the remaining two degrees of freedom reoptimized. Variables marked with * have a loss larger than 6 \$/h.

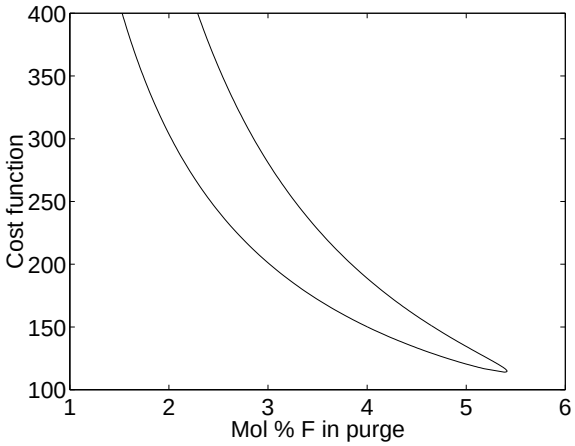


Figure 2: Unfavorable shape of cost function with F (byproduct) in purge as controlled variable. Shown for case with constant reactor temperature and C in purge.

The losses for the remaining $9 \cdot 8/2 = 36$ possible combinations of 2 variables were computed (not shown). Not surprisingly, keeping both recycle flow and compressor work constant results in infeasibility or large loss for disturbance 2 and for feed flow changes. This is as expected, because from process insight these two variables are closely correlated (and we could probably have elim-

inated one of them earlier). We note that constant F in product in all cases results in a large loss or infeasibility for disturbance 2. This, combined with the earlier finding that we should not control F in purge, leads to the conclusion that it is *not* favorable to control the composition of byproduct (F) for this process. The following four cases have a summed loss of less than 6 [$/h]:

I. Reactor temp., Recycle flow, C in purge (loss 3.8).

II. Reactor temp., Comp. work, C in purge (loss 3.9).

III. Reactor temp., C in purge, E in purge (loss 5.1).

IV. Reactor temp., C in purge, D in purge (loss 5.6).

Evaluation of Implementation Loss

In addition to disturbances, there will always be a implementation error related to each controlled variable, that is, a difference between its setpoint and its actual value, e.g. due to measurement error or poor control. By plotting plot for "best" case I the cost as a function of the three controlled variables (not shown) we find that the optimum is flat over a large range for all three variables, and we conclude that implementation error will not cause a problem. In comparison, for cases III and IV the cost is sensitive to implementation errors, and we get infeasibility if purge composition of D (case III) or E (case IV) becomes too small.

Should Inert be Controlled?

A common suggestion is that it is necessary to control the inventory of inert components, that is, in our case, to control the mole fraction of component B (Luyben et al., 1997) (McAvoy and Ye, 1994) (Lyman and Georgakis, 1995) (Ng and Stephanopoulos, 1998) (Tyreus, 1999). However, recall that we eliminated B in purge at an early stage because it gave a rather large loss for disturbance 2 (see Table 3). Moreover, and more seriously, we generally find that the shape of the economic objective function as a function of B in purge is very unfavorable, with either a sharp minimum or with the optimum value close to infeasibility. A typical example of the latter is shown in Figure 3. In conclusion, we do not recommend to control inert composition.

Summary

In conclusion, control of reactor temperature, C in purge, and recycle flow or compressor work (cases I or II) result in a small loss for disturbances and a flat optimum (and is thus insensitive to implementation error), and are therefore good candidates for self-optimizing control.

The analysis in the paper is based on steady-state economics, but we have also performed dynamic simulations that show that this proposal may be implemented in practice using a simple decentralized feedback control structure based on PI controllers.

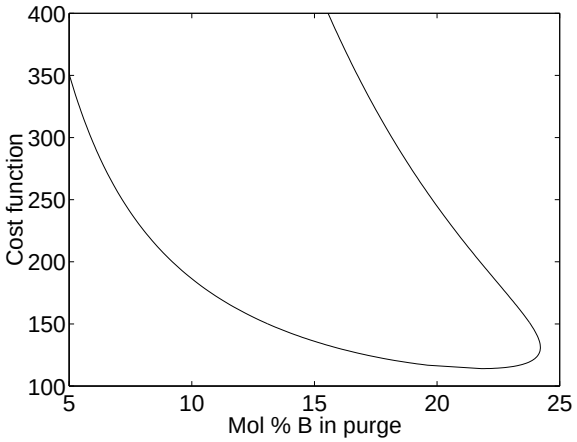


Figure 3: Typical unfavorable shape of cost function with B (inert) in purge as controlled variable (shown for case with constant reactor temperature and C in purge).

References

Downs, J. J. and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, **17**, 245–255 (1993).

Luyben, M. L., B. D. Tyreus, and W. L. Luyben, "Plantwide control design procedure," *AIChE J.*, **43**(12), 3161–3174 (1997).

Lyman, P. R. and C. Georgakis, "Plant-wide control of the Tennessee Eastman Problem," *Comput. Chem. Eng.*, **19**(3), 321–331 (1995).

McAvoy, T. J. and N. Ye, "Base control for the Tennessee Eastman problem," *Comput. Chem. Eng.*, **18**(5), 383–413 (1994).

Morari, M., G. Stephanopoulos, and Y. Arkun, "Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the classification of the control task. Analysis of the optimizing control structures.," *AIChE J.*, **26**(2), 220–232 (1980).

Ng, C. and G. Stephanopoulos, Plant-Wide control structures and strategies, In *Preprints Dycops-5*, pages 1–16. IFAC (1998).

Ricker, N. L., "Optimal steady-state operation of the Tennessee Eastman challenge process," *Comput. Chem. Eng.*, **19**(9), 949–959 (1995).

Ricker, N. L., "Decentralized control of the Tennessee Eastman Challenge Process," *J. Proc. Cont.*, **6**(4), 205–221 (1996).

Skogestad, S., "Plantwide control: The search for the self-optimizing control structure," *J. Proc. Cont.*, **10**, 487–507 (2000).

Tyreus, B. D., "Dominant Variables for Partial Control. 2. Application to the Tennessee Eastman Challenge Process.," *Ind. Eng. Chem. Res.*, pages 1444–1455 (1999).

Robust Passivity Analysis and Design for Chemical Processes

Huaizhong Li, Peter L. Lee* and Parisa A. Bahri†
School of Engineering
Murdoch University
Perth, WA 6150, Australia

Abstract

In this paper, a new approach is developed for robust control of nonlinear chemical processes. The methodology proposed is based on passivity theory. Uncertainty and perturbations are taken into account in the stability analysis and controller synthesis such that the resulting closed-loop system can achieve better robustness and performance. The suggested passivity framework can deal with a large class of uncertainties and perturbations in chemical processes.

Keywords

Process control, Passivity, Robust control

Introduction

The effective control of process plants using high performance control systems must consider both stability and performance in the face of uncertainty - the so called robust control problem. Many modern robust control methods are based on finite gain (H_∞) control theory. However, finite gain designs can be overly conservative for some uncertain systems because it ignores phase information of the feedback systems (Sakamoto and Suzuki, 1996).

Recent work on robust control design has employed the concept of passivity(Bao et al., 1999). Often robust stability of a system can be determined by evaluating passivity of a subsystem. Many uncertain systems can be converted into equivalent interconnected feedback systems which consist of a linear block and possibly a nonlinear and/or time-varying block. By studying the passivity of the interconnected systems, sufficient stability conditions can be derived for the original uncertain systems. Specifically, if the linear block is strictly passive and the nonlinear block is passive, then the original uncertain system is robustly stable.

In general, the nonlinear block of the interconnected system can be classified into four types: 1. Non-passive; 2. Near-passive; 3. Passive; and 4. Over-passive. It is very difficult to guarantee robust stability and robust performance if the nonlinear block is strictly non-passive. In order to apply passivity theory, it is necessary that the nonlinear is at least near passive.

However, it is not always straightforward or advantageous to combine all of the uncertainties into the nonlinear block. It is sometimes desired to leave some bounded uncertainties, for example, linearization errors, to the linear block such that the nonlinear block is passive or near-passive. It may be relatively simple to guarantee that the linear block with the bounded uncertainties is robust strictly passive.

This paper deals with robust stability and control design using the passivity approach. Near-passive and over-passive uncertain blocks are considered for the nonlinear block, while the the linear block may also contain uncertainty. A methodology for robust stability criteria and controller synthesis is proposed. The results presented in this work provide better robust stability and system performance for those systems that have a over passive nonlinear block or contain other uncertainties in the linear block. Two examples are presented to highlight the benefits of the proposed approach.

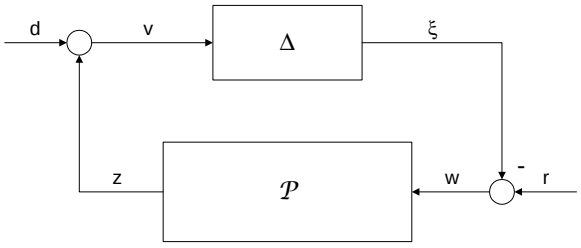


Figure 1: Interconnected feedback system.

Preliminaries

Consider the feedback system depicted in Figure 1, where $\mathcal{P}$ is a linear system and Δ represents a nonlinear/time-varying uncertainty.

The passivity concept is closely related to energy in an interconnected feedback system. For stability analysis, inputs r and d can be set to zero. For such a system, stability results can be established by observing energy consumption in the feedback loop. Let us assume that there is an abnormal energy burst occurring in the feedback loop. If the uncertain block Δ is passive, then at least this block does not inject energy into the feedback loop. In addition, if the linear block $\mathcal{P}$ is strictly passive, it means that this block absorbs energy. As there is no energy injection into the feedback loop, it is expected that energy in the feedback loop will be finally consumed by blocks Δ and $\mathcal{P}$, and hence stability is maintained. Technical details about the passivity approach can be found in Desoer and Vidyasagar (1975).

In the sequence, we call Δ block near passive if it is not

*plee@central.murdoch.edu.au. Corresponding author
†parisa@eng.murdoch.edu.au

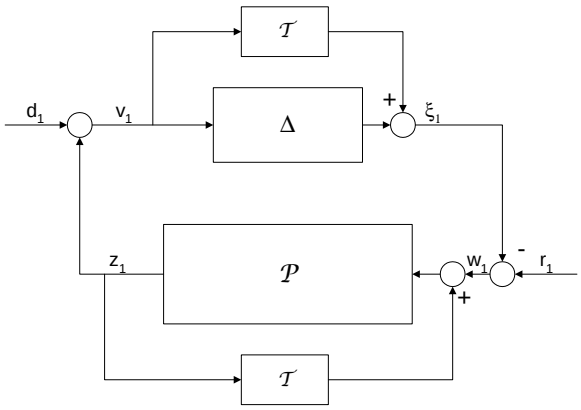


Figure 2: Transformed feedback system.

passive, but is close to being passive. Similarly, we call Δ over passive if it remains passive even under certain additive perturbation.

When Δ is near passive, transforms (Desoer and Vidyasagar, 1975; Xie et al., 1998) can be applied on the interconnected system to render the transformed Δ block passive, as depicted in Figure 2. Similarly, transforms can also be applied to reduce conservatism if Δ is over passive. In both cases, stability analysis are performed on the feedback system with two transformed blocks.

Methodology

In general, there is no unique way to find the transform $\mathcal{T}$. If Δ is linear time-invariant and near passive, then the easiest way to find $\mathcal{T}$ is to use the passivity index $v(\Delta)$ (Bao et al., 1999) such that $\mathcal{T} = v(\Delta)I$ and $\mathcal{T} + \Delta$ is passive. If Δ is over passive, then it is always possible to set $\mathcal{T} = \beta I$ where β can be any scalar such that $\mathcal{T} + \Delta$ is still passive. A general form for $\mathcal{T}$ when Δ is over passive can be found in Xie et al. (1998).

Once the transformed Δ block is rendered passive, a controller can be synthesized to render the other transformed block strictly passive.

A nonlinear chemical process can be represented by the diagram in Figure 3, where Δ_1 is generally used to represent the output sensor errors or neglected high frequency dynamics for the process, $\mathcal{P}_p$ is the linearized model for the process at a specific operating point, and K is a controller. However, due to nonlinearity and unknown perturbations, a better model for the process may be $\mathcal{P}_p(\Delta_2)$, where Δ_2 represents all of the uncertainties which are not absorbed by Δ_1. Note that Δ_2 can always be set to zero if the linearized model $\mathcal{P}_p$ is relatively accurate. In this case, further constraints can be manually imposed on Δ_1 to cope with the influence of omitting Δ_2 (Bao et al., 1999). It can be viewed as a special case under the framework of Figure 3.

For stability analysis, it is assumed that $r(t) = 0$. It

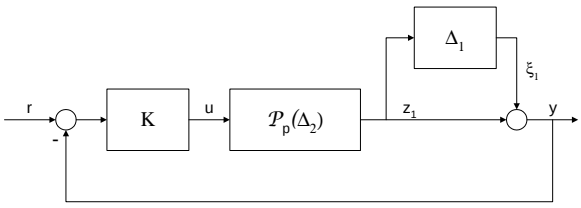


Figure 3: Diagram of a typical chemical process.

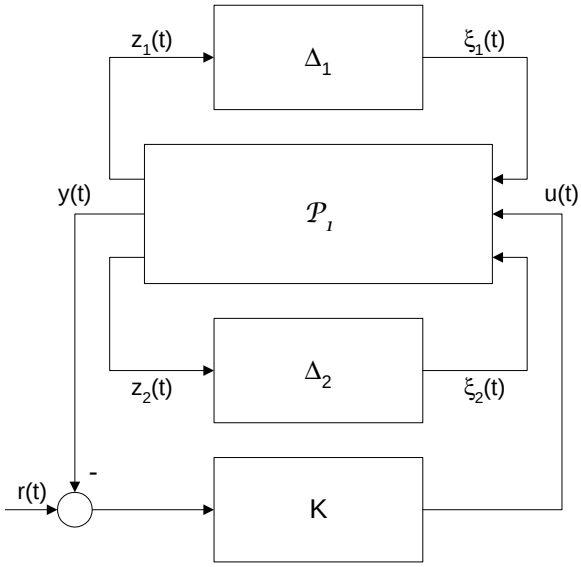


Figure 4: System diagram after 'pulled out' uncertainty.

is easy to check that for the special case $\Delta_2 = 0$, Figure 3 is equivalent to Figure 1 with $\Delta = \Delta_1$ and

$$\mathcal{P} = -\mathcal{P}_p K (I + \mathcal{P}_p K)^{-1}.$$

In general, the uncertainties Δ_1 and Δ_2 can be 'pulled out' to form an interconnected feedback system (see, for examples, Zhou et al., 1996), as depicted in Figure 4. We use operator matrix $\mathcal{P}_1$ to denote the remaining linear system after 'pulled out' uncertainties.

If Δ_1 is passive, a controller can be synthesized for the low half part with input $\xi_1(t)$ and output $z_1(t)$ to render it robustly strictly passive. The same controller also guarantees that the closed-loop system in Figure 3 is robustly stable.

If Δ_1 is near passive or over passive, a transform $\mathcal{T}$ can be applied to Δ_1 and the low half part in Figure 3, as shown in Figure 5. Note that $\mathcal{T}$ is set to $\mathcal{T} = \beta I$ for both near passive and over passive in this paper, and $\beta \geq v(\Delta_1)$ for the near passive case.

The transformed feedback system can be viewed as an interconnection of two transformed blocks, namely, 'Passive Uncertainty' block with input $z(t)$ and output $\xi(t)$ and 'System' block with input $\xi(t)$ and output $z(t)$, as

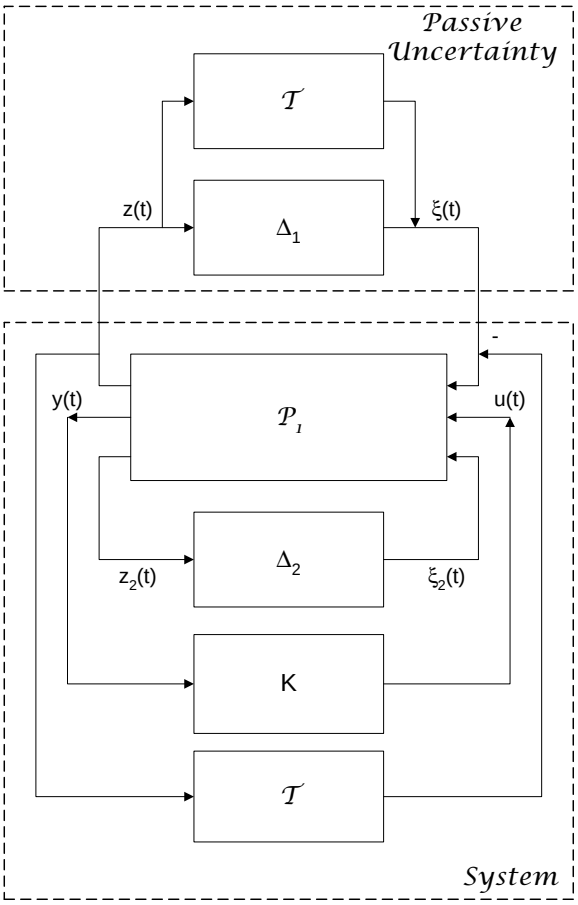


Figure 5: Transformed feedback system.

shown in Figure 5. K only needs to render the 'System' block robustly strictly passive.

The 'System' block can be described in the following state-space form:

$$\begin{aligned} \dot{x}(t) &= A(\beta)x(t) + B(\beta)u(t) + H_1\xi(t) + H_2\xi_2(t) \\ y(t) &= C(\beta)x(t) + D(\beta)\xi(t) \\ z(t) &= E_{11}x(t) + E_{12}u(t) + E_{13}\xi(t) \\ z_2(t) &= E_{21}x(t) + E_{22}u(t) + E_{23}\xi(t) \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state, $y(t) \in \mathbb{R}^m$ is the measured output of the process, $u(t) \in \mathbb{R}^r$ is the manipulated input to the process. Vectors $\xi(t)$ and $z(t)$ can be viewed as the input and output of the 'System' block. Mapping $z_2(t) \to \xi_2(t)$ describes the uncertainty Δ_2. This mapping can be very general, for example, it can be represented using integral quadratic constraints (IQCs) (see Boyd et al., 1994, for details). It has been shown that time delay uncertainties in systems can be represented by IQCs (see Xie et al., 1998, for details). However, to represent the linearization errors for a chemical process, it is often sufficient to set $E_{22} = 0$, $E_{23} = 0$ and $\xi_2(t) = \Delta_2(t)z_2(t)$ where $\|\Delta_2(t)\|_\infty \leq 1$.

β is a constant scalar. Although β can be pre-selected

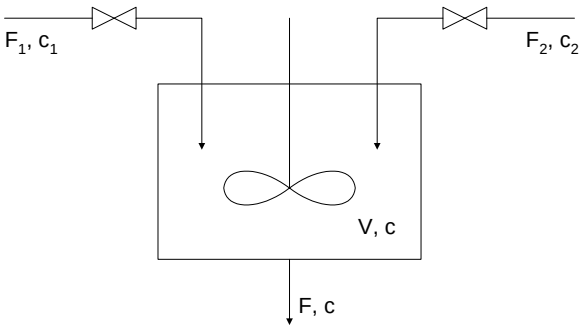


Figure 6: A mixing system.

based on the passivity index, we intentionally leave it intact as a design parameter. A β dependent controller can be synthesized, and then an iterative procedure or quasi-convex optimization can be applied to find the sub-optimal β which preserves the passive property of the 'Passive Uncertainty' block while achieving the best performance for the closed-loop system.

Based on the discussion above, we arrive at the following result:

Theorem 1 *If there exists a controller $u = Ky$ which renders system (1) robustly strictly passive, then the same controller guarantees that the closed-loop system shown in Figure 3 is robustly stable.*

Results are available in the literature to synthesize controller $u = Ky$ such that system (1) is robustly strictly passive. We omit the details in this paper due to page limitations. The interested readers may refer to Xie and Soh (1995) and the references therein.

Justification in Chemical Processes

In this section, we consider two examples. The main purpose of this section is to justify the framework proposed above for chemical processes. The details of the two examples are deleted for clarity.

The first example is a mixing system adopted from Bao et al. (1999). The well-stirred tank is fed with two inlet flows with flowrates $F_1(t)$ and $F_2(t)$. Both inlet flows contain the same dissolved material with concentrations c_1 and c_2 respectively. The outlet flowrate is $F(t)$. $F_1(t)$ and $F_2(t)$ are manipulated to control both $F(t)$ and the outlet concentration $c(t)$.

It was assumed in Bao et al. (1999) that inlet concentration disturbances cause the nominal plant transfer matrix $\hat{P}(s)$ to be shifted to $P(s)$. It was shown that the control problem can be described by Figure 3 with a near passive $\Delta_1(s) = (P(s) - \hat{P}(s))\hat{P}^{-1}(s)$ and $\mathcal{P}_p(\Delta_2) = \hat{P}(s)$. Following the settings in Bao et al. (1999), it is straightforward to verify that the feedback system in Bao et al. (1999) is equivalent to Figure 5 with $E_{13} = 0$, $E_{21} = 0$, $E_{22} = 0$ and $E_{23} = 0$ in (1). As (1)

is reduced to a linear system with no uncertainty in this case, standard passivity synthesis procedure can be applied to (1) to design the controller $u = Ky$.

It is also easy to show that $\Delta_1(s) = (P(s) - \hat{P}(s))\hat{P}^{-1}(s)$ can be over passive under different inlet concentration disturbances.

Next, we consider the stability problem of a CSTR. The component and energy balances for a CSTR with jacket cooling and first order irreversible exothermic reaction $A \rightarrow B$ are:

$$\frac{dC}{dt} = \frac{F}{V}(C_0 - C) - kC$$
$$\frac{dT}{dt} = \frac{F}{V}(T_0 - T) + \frac{\Delta H}{\rho c_p}kC - \frac{UA_r}{V\rho c_p}(T - T_C)$$

where V is tank volume, F is feed flowrate, C_0 is feed concentration, C is outlet concentration, V is heat transfer coefficient, ρ is density, c_p is specific heat, T is reactor temperature, T_0 is feed temperature, ΔH is exothermic heat of reaction, U is heat transfer coefficient, A_r is heat transfer area from jacket to reactor, and T_C is cooling water temperature. k is defined by Arrhenius relation $k = k_0 e^{-\frac{E}{RT}}$ where k_0 is a rate coefficient factor, E/R is a activation energy factor. The measured variables are outlet concentration C and reactor temperature T, and the manipulated variables are feed flowrate F and cooling water temperature T_C.

The component and energy balance equations are linearized at a specific operating point to obtain two linear differential equations. If there is no feed perturbation, then linearization errors for the two balance equations depend on the state variables C and T only. It was shown in Doyle III et al. (1989) that if the linearization errors are conic-sector bounded, then the process description can be given by

$$\dot{x}(t) = (A + \Delta_2 A_1)x(t) + Bu(t)$$
$$z_1(t) = Cx(t)$$

where Δ_2 is a bounded matrix.

However, the above plant description is obtained by assuming that the measurements are perfect. The measurement errors may be not significant when the process is operating at steady state. However, the instrument measurement response can often be important in the overall system response in the transition period (Mylroi and Calvert, 1986), or when the process is operating under the influence of uncertainty and perturbations. Therefore, the true measured output $y(t)$ is not identical to $z_1(t)$ in the above process description.

For simplicity, it is assumed that both temperature and concentration measurements can be represented by first-order lag models $1/(\tau_T s + 1)$ and $1/(\tau_C s + 1)$ where τ_T and τ_C are time constants which may change due to the variations in the unit inputs (Mylroi and Calvert, 1986). Therefore, the measured output is

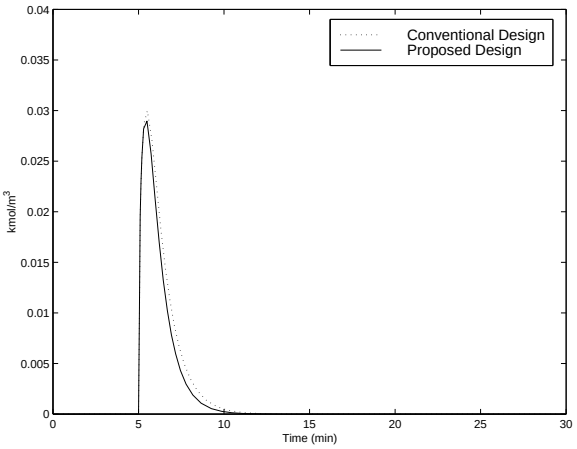


Figure 7: Control errors of outlet concentration for the mixing system.

$y = \text{diag}\{1/(\tau_T s + 1), 1/(\tau_C s + 1)\}z_1$. In the framework of Figure 3, this leads to an uncertain $\Delta_1 = \text{diag}\{-\tau_T s/(\tau_T s + 1), -\tau_C s/(\tau_C s + 1)\}$. It can be easily verified that Δ_1 is a near passive block and the passive index $v(\Delta_1) = 1$. Similarly, it is straightforward to see that the CSTR case fits into the framework of Figure 5 with $\beta = 1$, $E_{12} = 0$, $E_{13} = 0$, $E_{22} = 0$ and $E_{23} = 0$. These four matrices can be non-zero, for example, if there is a disturbance on the feed flowrate F.

Furthermore, it is possible to convert a multi-model control problem for multi-unit chemical processes in Lee et al. (2000) into a passivity framework shown in Figure 5. The passivity based multi-model approach will be addressed in a subsequent paper.

Illustrative Examples

We again consider the stability problem of the two examples mentioned above.

The first example verifies that over passive uncertainty needs to be properly addressed in order to reduce conservatism. This example takes the mixing system discussed in Bao et al. (1999). Differing slightly from (Bao et al., 1999), it is assumed that the $c_1(t)$ changes from $0.5kmol/m^3$ to $1.0kmol/m^3$ while $c_2(t)$ changes from $2.0kmol/m^3$ to $3.0kmol/m^3$. Rest of the values follow exactly as those in (Bao et al., 1999). Following the same procedure, it is easy to verify that Δ_1 block is over passive. Two passivity based designs were carried out. As Δ_1 is already passive, design based on conventional passivity approach only renders the nominal system $\hat{P}(s)$ strictly passive. Design based on the proposed framework first renders $\Delta_1(s)$ passive by transform βI where β is negative, then a controller is synthesized to render the 'System' block strictly passive.

It is also assumed that at the 5th minute both c_1 and c_2 change simultaneously, c_1 decreases from $1kmol/m^3$

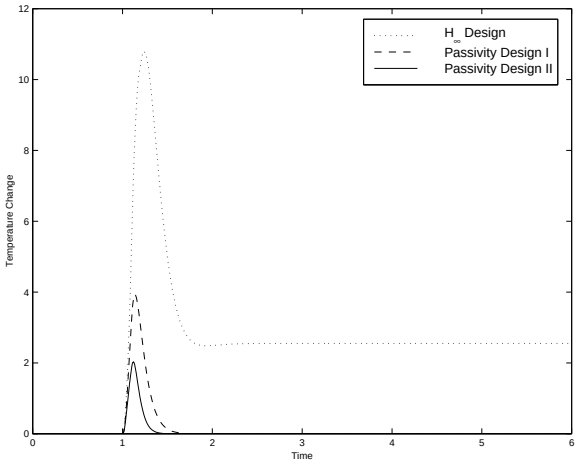


Figure 8: Closed-loop simulation results for the CSTR.

to $0.5kmol/m^3$ and c_2 increases from $2kmol/m^3$ to $3kmol/m^3$. Simulations shown in Figure 7 verifies that the proposed design is less conservative. The only purpose of this example is to show that conservatism can be further reduced in the over passive case by simply applying a loop transform, therefore, simulation results are less dramatic as uncertainty $\Delta_2 = 0$ in this case.

Our second example demonstrates the advantage of the framework proposed in this paper. This example uses the CSTR mentioned above. It is assumed that there are linearization errors and perturbations on the feed flowrate F, feed concentration C_0 and feed temperature T_0, as well as first-order measurement dynamics. The object is to design a controller such that the system is stable under feed perturbations, measurement dynamics and linearization errors. The details of the example have to be omitted due to space limitations.

Similar to Bao et al. (1999), a standard H_∞ design was carried out for comparison purpose. The H_∞ design assumes that there are no measurement dynamics.

Two passivity based design were studied, namely, Passivity Design I and Passivity Design II. Passivity Design I is similar to the one in Bao et al. (1999) by assuming that there are no measurement dynamics. The process was only perturbed by the feed perturbations which lead to a near passive Δ_1 in Figure 3. Similar to Bao et al. (1999), linearization error was partially taken into account by the Δ_1 block. However, unlike the procedure in Bao et al. (1999), frequency weighting was not used for simplicity. Passivity Design II uses the methodology proposed in this paper. There was no specific assumption on perturbation, linearization errors or measurement dynamics as those imposed in the H_∞ design and the Passive Design I. Similar to Passive Design I, frequency weighting was not used. However, it should be stressed that frequency weighting can be adopted for all three designs to further improve performance.

Three controllers were synthesized using the three design methods. They were then applied to the nonlinear component and energy balance model to simulate the behaviour of the closed-loop system. It was assumed that simultaneously feed perturbations occur at $t = 1$min after the closed loop reached steady state. Figure 8 shows the control errors of the reactor temperature corresponding to the three controllers. It can be observed that the H_∞ controller performs worse with a steady state offset and longer transition period, and the Passivity Design II controller is the best.

Conclusion

A new methodology is proposed in this paper to deal with robust stability analysis and controller synthesis of nonlinear chemical processes. The proposed approach is based on passivity theory which may be less conservative than the commonly used H_∞ approach in robust control. The framework used in this paper enables the consideration of near passive and over passive uncertainties in a unified way, and also allows a larger class of uncertainties existing in the chemical processes.

References

Bao, J., P. L. Lee, F. Wang, and W. Zhou, "New robust stability criterion and robust controller synthesis," *Int. J. Robust Nonlinear Contr.* (1999).

Boyd, S., L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia: SIAM (1994).

Desoer, C. A. and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. Academic Press, NY (1975).

Doyle III, F. J., A. K. Packard, and M. Morari, Robust controller design for a nonlinear CSTR, In *Proc. 1989 American Contr. Conf.*, pages 1087–1092 (1989).

Lee, P. L., H. Li, and I. T. Cameron, "Decentralized control design for nonlinear multi-unit plants: a gap metric approach," *Chem. Eng. Sci.*, **55**, 3743–3758 (2000).

Mylroi, M. G. and G. Calvert, *Measurement and Instrumentation for Control*. Peter Peregrinus Ltd., London (1986).

Sakamoto, N. and M. Suzuki, "γ-Passive System and Its Phase Property and Synthesis," *IEEE Trans. Auto. Cont.*, **41**, 859–865 (1996).

Xie, L. and Y. C. Soh, "Positive real control problem for uncertain linear time-invariant systems," *Sys. Cont. Let.*, **24**, 265–271 (1995).

Xie, L., M. Fu, and H. Li, "Passivity analysis and passification for uncertain signal processing systems," *IEEE Trans. Signal Proces.*, **46**(9), 2394–2403 (1998).

Zhou, K., J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall (1996).

On-line Optimization of a Crude Unit Heat Exchanger Network

Tore Lid*
Statoil Mongstad
N-5954 Mongstad, Norway

Stig Strand
Statoil Research Centre
Arkitekt Ebbellsvei 10
N-7005 Trondheim, Norway

Sigurd Skogestad
Department of Chemical Engineering, NTNU
N-7491 Trondheim, Norway

Abstract

This paper describes modeling and on-line optimization of a crude unit heat exchanger network at the Statoil Mongstad refinery. The objective is to minimize the energy input in the gas fired heater by optimally distributing the cold crude oil in the heat exchanger network. The steady state mass and energy balance of the 20 heat exchangers in the network yields the process model. This model is fitted to the measured values using data reconciliation and unmeasured values like heat exchanger duty and heat transfer coefficients are computed. The fitted model is used to compute the optimal split fractions of crude in the network. This system has been implemented at the refinery and has resulted in a 2% reduction in energy consumption. In operational modes where the unit is constrained on energy input this gives a increased throughput and a significant contribution to the refinery profit.

Keywords

Reconciliation, Optimization, Crude unit, Heat exchanger network

Introduction

This paper describes the development of a real time optimization system including model development, data reconciliation and on-line optimization. The case studied is a heat exchanger network for pre-heat of feed in a crude oil distillation unit at the Statoil Mongstad Refinery. The system is implemented and is now running in closed loop at the refinery. The optimal operation is computed using a steady state model which before each run is fitted to the current operation point. Process measurements contain uncertainties as random errors and possibly gross errors. This may be a result of miscalibration or failure in the measuring device. This uncertainty is reduced when the current operation point is estimated using a larger number of measurements, than the number of unknowns in the process model, to compute a set of reconciled data. Model parameters are estimated simultaneously or computed from the reconciled data. The optimal operation is computed as the maximum of the objective subject to the process model, current process operation and model parameters. The optimal operation is finally implemented as setpoints in the process control system. A large number of methods for data reconciliation have been suggested. These include robust objective functions (Chen et al., 1998), statistical tests, analysis of measurement redundancy and variable observability (Crowe et al., 1983). However, most examples and case studies presented in the literature are based on simulated processes, and most papers consider the data reconciliation decoupled from the optimization. One noteworthy exception is (Chen et al., 1998) who present an application of data reconciliation to a Monsanto sulfuric acid plant, but the paper is somewhat limited on details on the specific approach they have taken. The objective of this paper is therefore to present an actual industrial implementation, where we provide details about the data reconciliation approach, model and optimization.

*Tore.Lid@Statoil.com

Data Reconciliation

Data reconciliation is used to determine the current operation point. If measurement had no uncertainty the current operation point could be determined from $n-m$ independent measurements, where n is the number of variables and m the number of equations in the model. Since the measurements are uncertain and there are a surplus of measurements, compared to the number of unknown variables in the model, data reconciliation is used to reduce this uncertainty. The reconciled values minimizes some function of all measurement errors subject to the model equations. This is written as

$$\min_x \sum_{i=1}^{n_m} \psi(\epsilon_i/\sigma_i) \tag{1}$$

subject to

$$Ax = 0$$
$$g(x) = 0$$

All variables are collected in the vector x of dimension $n \times 1$. The measurement errors $\epsilon_i = x_i - y_i$ where y_i is a measurement of the variable contained in x_i. All measurement errors is scaled by its standard deviation σ_i. The process model is separated into a set of linear equations, $Ax = 0$, and nonlinear equations, $g(x) = 0$, since most NLP solvers take linear and nonlinear equations as separate arguments. If the uncertainty in the measurements are normal distributed with zero mean the summed squared measurement error is used as objective function, ψ, in equation 1. However, in the case of nonzero measurement error mean, gross errors, this method gives a biased estimate of the process variables. There are several methods for reducing of the effect of gross errors. In (Crowe et al., 1983) and (Crowe, 1986)

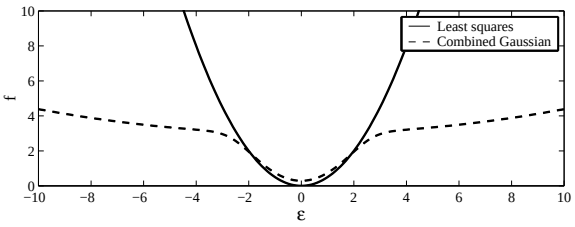


Figure 1: Combined Gaussian objective function. The standard deviation $\sigma = 1$, $p = 0.3$ and $b = 6$.

collective and individual statistical tests of the measurement errors are used to exclude measurements with gross errors. In (Chen et al., 1998),(Tjoa and Biegler, 1991) and (Johnston and Kramer, 1995) objective functions less sensitive to gross errors are used. In this work the *Combined Gaussian* objective function is selected due to its numerical robustness and promising "small example" results. In all robust objective functions the measurement error is scaled by its standard deviation. Normally this distribution is not known and the standard deviation has to be estimated from measured data or determined by a reasonable guess based on the actual measurement equipment installed and its measurement range. The Combined Gaussian function is based on a weighted sum of two Gaussian distributions, one distribution of the random errors and one of the gross errors. The combined Gaussian probability density function is written as

$$f_i = \frac{1}{\sigma_i\sqrt{2\pi}}\left[(1-p)\exp\left(-\frac{1}{2}\frac{\epsilon_i^2}{\sigma_i^2}\right) + \frac{p}{b}\exp\left(-\frac{1}{2}\frac{\epsilon_i^2}{\sigma_i^2 b^2}\right)\right] \quad (2)$$

with the probability of a gross error in the measurements p and the ratio of the standard deviations of the gross errors to that of the random errors b. The objective function to be minimized is the negative logarithm of the probability density function, $\sum_{i=1}^{n_m}\log(1/f_i)$. The Combined Gaussian objective function is graphed in Figure 1 with the least squares function for comparison. Compared to the least squares method the robust functions gives less penalty for measurement errors larger than 3σ. For the reconciled data this typically gives large measurement errors in few variables and small error in the other variables. At least intuitively this is what one would expect from the process measurements though it is difficult to verify. In equation 1 there is no limitation on the number of measurements and on which variable to measure. Before the reconciled variables are accepted some analysis has to be made to check if the unmeasured variables are observable. The measurements can also be classified as redundant or nonredundant measurements which can be used to evaluate the reconciled variables and decide if data reconciliation can be done if a specific measurement is out of service. Let x_r^* be a solution to the reconciliation problem in equation 1. The nonlinear constraints are linearized at the optimal solution x_r^* such that $g(x) \approx g(x_r^*)+G(x-x_r^*)$, where $G = \partial g(x)/\partial x|_{x=x_r^*}$. The linear and linearized constraints can now be written as

$$\hat{A}x - \hat{b} = 0 \quad (3)$$

where

$$\hat{A} = \begin{bmatrix} A \\ G \end{bmatrix} \qquad \hat{b} = \begin{bmatrix} 0 \\ g(x_r^*) - Gx_r^* \end{bmatrix}$$

The variables in x are separated into measured variables, y, and unmeasured variables, z. The matrix $\hat{A}$ is partitioned into $\hat{A}_1$ and $\hat{A}_2$ where $\hat{A}_1$ holds the columns of $\hat{A}$ corresponding to the measured variables and $\hat{A}_2$ the columns of $\hat{A}$ corresponding to the unmeasured variables. Equation 3 can now be written as

$$\hat{A}_1 y + \hat{A}_2 z = \hat{b} \quad (4)$$

To be able to compute the unmeasured variables, from the measured variables, the matrix $\hat{A}_2$ must have full column rank. If the number of measurements $n_y < n - m$, where n is the number of variables and m the number of equations, the size of $\hat{A}_2$ is $m \times n_z$ where $n_z > m$ and the matrix $\hat{A}_2$ has rank less than n_z. This implies that equation 4 has no unique solution for z when y is known. A requirement is that the number of measurements $n_y \geq n-m$, which is obvious, and that $\hat{A}_2$ has full column rank. The measurements can also be separated into redundant and nonredundant measurements. If a variables measurement is redundant it is possible to compute its value if its measurement is removed. This is not the case for a nonredundant measurement and removing this measurement causes $\hat{A}_2$ to be rank deficient. A simple test for redundancy is to check if $P^T\hat{A}_1$ has columns with only zero elements, where P is defined as a matrix that span the null space of $\hat{A}_2^T$. Any zero columns in corresponds to nonredundant measurements. Also note that for a nonredundant measurement i we always have that $y_i - y_{mi} = 0$ and that this measurement does not contribute directly in the calculations of the reconciled values.

Optimization

The typical process optimization problem has a linear objective function like product price times product flow which is to be maximized. For system simplicity the same process model and variable vector are used in both data reconciliation and process optimization. In the optimization problem some of the variable values are already known. These are typically disturbance variables and connects the data reconciliation with the optimization. The variable values are specified in the optimization problem as a set of linear constraints ($Rx = r$) where

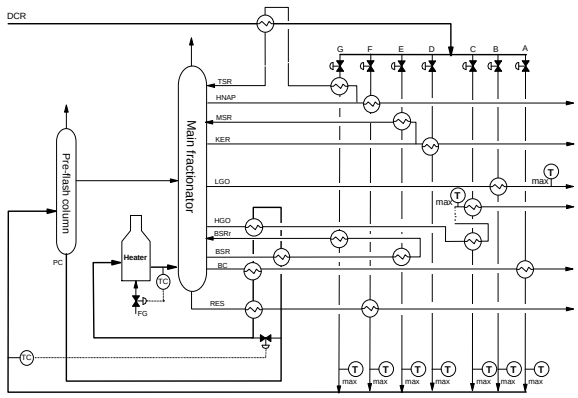


Figure 2: Simplified crude unit overview.

$r = Rx_r^*$. The matrix R has one nonzero element in each row, equal to one, corresponding to the element in x, which is set equal to its reconciled value. The optimization problem can now be written as

$$\min_x -p^T x \tag{5}$$

subject to

$$\begin{aligned} Ax &= 0 \\ g(x) &= 0 \\ Rx &= r \end{aligned}$$

$$x_{\min} \geq x \geq x_{\max}$$

Inequality constraints in process optimization are typically bounds on singe variables. Inequality constraints on combinations of variables may be added in this formulation by introducing slack variables.

A Case Study

In the crude unit the crude feed is separated into suitable components for production of propane, butane, gasoline, jet fuel, diesel and fuel oil. The crude is preheated in a heat exchanger network where heat is recovered from the hot products and circulating refluxes. As shown in Figure 2 the cold crude (DCR) is separated into seven parallel streams (A-G) and heated by the hot products. The flow in each pass and BSR heat exchanger bypasses provides the degrees of freedom necessary for optimization. The optimization objective is to save energy and to recover as much heat as possible. The heater is the main energy input in the process and heater outlet temperature is held constant. The minimum energy is then achieved by maximizing of the heater inlet temperature. Both distillation columns have feed conditions independent on the heat exchanger network operation. The inlet temperatures of both columns are assumed to have perfect temperature control. The feed flow and composition are then independent of operation of the heat exchanger network. With this simplification a model of the distillation columns is not needed and a mass and energy balance of the heat exchanger network is a sufficiently detailed model for optimization. The optimal solution must be within several process operating constraints. The total crude flow or throughput is to be unchanged. At each crude pass outlet there is a maximum temperature constraint to avoid flashing. On main column LGO and HGO products, exiting the heat exchangers, there is a maximum temperature limit as the products are fed to the LGO and HGO driers (the driers are not drawn in Figure 2). The preflash column inlet temperature is to be unchanged. Some of the heat exchangers are also included in the bottom circulating reflux (BSR) and the total duty in BSR is to be unchanged.

The Process Model

The heat exchanger network can be viewed as a set of nodes or unit operations connected by arcs or in this case pipes. A set of balance equations, mass and energy balance, describes the internals of each node. Variables for the arcs or pipes are fluid temperature and mass flow. The nodes in this network are stream mix nodes, stream split nodes and heat exchanger nodes. This selection of variables makes all nodes independent of other variables than those included in the input and output arcs. Heat exchanger nodes also have some internal variables like heat transfer coefficient and duty. This variable selection makes the model structure a simple and surveyable and it makes it practical possible to compute analytical derivatives of the nonlinear model equations. This reduces the numerical computational load in solving the model. The following describes the simplified balance equations for each type of node.

Mixing of Streams

In a node where n streams are mixed into one outlet stream the mass and energy balance equations can be written as

$$F_{\text{out}} - \sum_{i=1}^{n} F_{\text{in}_i} = 0 \tag{6}$$

$$F_{\text{out}} h(T_{\text{out}}) - \sum_{i=1}^{n} F_{\text{in}_i} h(T_{\text{in}_i}) = 0 \tag{7}$$

where $h(T)$ is the specific enthalpy of the fluid. The mass balance results in one linear equation and the energy balance in one nonlinear equation.

Splitting of Streams

In a node where one inlet stream are separated into n outlet streams the mass and energy balance equations

can be written as

$$F_{\text{in}} - \sum_{i=1}^{n} F_{\text{out}_i} = 0 \tag{8}$$

$$T_{\text{in}} - T_{\text{out}_i} = 0 \qquad \forall\, i = 1...n \tag{9}$$

The mass balance results in one linear equation and the energy balance results in n linear equations.

Heat Exchanger

For a heat exchanger node hot and cold side mass and energy balance and heat transfer is written as

$$F_{\text{c}_\text{in}} - F_{\text{c}_\text{out}} = 0 \tag{10}$$

$$F_{\text{h}_\text{in}} - F_{\text{h}_\text{out}} = 0 \tag{11}$$

$$Q + F_{\text{c}_\text{in}}(h(T_{\text{c}_\text{in}}) - h(T_{\text{c}_\text{out}})) = 0 \tag{12}$$

$$Q - F_{\text{h}_\text{in}}(h(T_{\text{h}_\text{in}}) - h(T_{\text{h}_\text{out}})) = 0 \tag{13}$$

$$Q - \varepsilon C_{\min}(T_{\text{h}_\text{in}} - T_{\text{c}_\text{in}}) = 0 \tag{14}$$

where the mass balance results in two linear equations (10 ,11) and the energy balance results in two nonlinear equations (12, 13). The heat transfer is described by equation 14. The heat exchangers in this unit is of multiple tube and multiple shell pass type and the ε-Ntu method (Mills, 1995) is used for calculation of the heat transfer. In equation 14 ε is the efficiency and $C_{\min}$ is the minimum capacity. $C_{\min}$ is calculated as

$$C_{\min} = \min(C_{\text{c}}, C_{\text{h}}) \tag{15}$$

$$C_{\text{c}} = F_{\text{c}_\text{in}} \text{C}_{\text{p}_\text{c}} \approx F_{\text{c}_\text{in}} \frac{h\left(T_{\text{c}_\text{out}}\right) - h\left(T_{\text{c}_\text{in}}\right)}{T_{\text{c}_\text{out}} - T_{\text{c}_\text{in}}} \tag{16}$$

$$C_{\text{h}} = F_{\text{h}_\text{in}} \text{C}_{\text{p}_\text{h}} \approx F_{\text{h}_\text{in}} \frac{h\left(T_{\text{h}_\text{out}}\right) - h\left(T_{\text{h}_\text{in}}\right)}{T_{\text{h}_\text{out}} - T_{\text{h}_\text{in}}} \tag{17}$$

The efficiency, ε, is a function of the number of transfer units, Ntu, and the capacity ratio, R_C. R_C and Ntu is calculated as

$$R_C = \frac{C_{\min}}{C_{\max}} \qquad Ntu = \frac{UA}{C_{\min}} \tag{18}$$

where $C_{\max} = \max(C_{\text{c}}, C_{\text{h}})$. The efficiency ε equals ε_1 for heat exchangers with single shell pass ($n = 1$) and a even number of tube passes. ε equals ε_2 for heat exchangers with even number of tube passes and n shell passes. ε_1 and ε_2 is calculated as

$$\varepsilon_1 = 2\left\{1 + R_C + \sqrt{1+R_C^2}\frac{1+\exp\left(-\frac{Ntu}{n}\left(\sqrt{1+R_C^2}\right)\right)}{1-\exp\left(-\frac{Ntu}{n}\left(\sqrt{1+R_C^2}\right)\right)}\right\}^{-1} \tag{19}$$

$$\varepsilon_2 = \left[\left(\frac{1-\varepsilon_1 R_C}{1-\varepsilon_1}\right)^n - 1\right]\left[\left(\frac{1-\varepsilon_1 R_C}{1-\varepsilon_1}\right)^n - R_C\right]^{-1} \tag{20}$$

When the equations for $C_{\min}$ and ϵ is substituted into equation 14 each heat exchanger is described by two linear and tree nonlinear equations.

Model Summary

There are totally 85 streams and 20 heat exchangers in the heat exchanger network. There are 9 stream mixes and 7 stream splits. The variables are 85 flows and 85 temperatures from the streams, 20 heat exchanger duties, 20 heat transfer coefficients and adds up to totally 210 variables. From the heat exchangers we have 40 linear and 60 nonlinear equations. From the stream mixing nodes we have 9 linear and 9 nonlinear equations and from the split nodes we have 29 linear equations. Coefficients for linear equations are collected in the matrix A where each equation occupy one row. The equation coefficients are placed in the column corresponding to its variable position in x. The nonlinear equation residues are collected in the residual vector $g(x)$. The model is now in the preferred form

$$Ax = 0 \tag{21}$$

$$g(x) = 0 \tag{22}$$

where A is a 78×210 matrix with the linear equation coefficients and $g(x)$ is a 1×69 vector of nonlinear equation residues.

On-line Data Reconciliation

Data is sampled from the process as one hour averages and reconciled using the Combined Gaussian objective function. Standard deviations for measurements are selected to be 1°C for temperature measurements and 2% of the maximum measuring range for flow measurements. The Combined Gaussian parameters p and b are set to 0.3 and 6. To avoid numerical difficulties in the model equations, like reversed flows, appropriate variable bounds are added to the data reconciliation problem in equation 1. There are 88 measurements in the process, which is a surplus of 25 compared to the number of unknown in the process model. The described analysis shows that all unmeasured variables are observable and that all measurements are redundant. As a example Figure 3 shows measured and reconciled values for 300 successive samples of one hour averages. The imbalance in the data is most likely caused by a gross error in the flow measurement of the hot stream F_{h}. The average error is 3.1 T/h and is fairly constant in all samples.

On-line Optimization

In the optimization problem the number of equality constraints are increased to 205 which leaves 5 degrees of freedom. These degrees of freedom corresponds to the flow trough each of the seven passes in the hot train minus two since the total flow and BSR duty is set equal to the reconciled value. As a example measured data is reconciled and optimum operation computed. Compared to current operation the pass flow (A-G) is changed by [0.0,-9.2,-0.1,+9.0,+0.1,+1.0,-0.8]%. In addition bypass

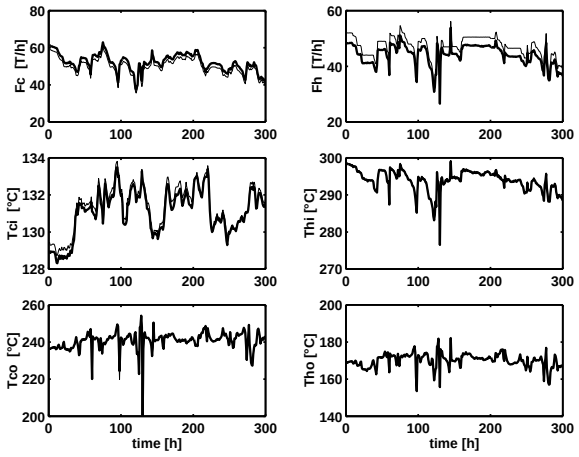


Figure 3: Measured (thin lines) and reconciled values (thick lines) for one of the heat exchangers.

flows of heat exchangers in the BSR is changed such that more heat is added in each pass while keeping the total duty constant. This increases the heater pre-heat duty by 2.3MW. Compared to the heater duty of $\approx$ 100MW this gives a 2% reduction of energy requirement. Constraints on pass G outlet temperature and LGO drier inlet temperature is active at optimal operation. Optimal operation is implemented as flow ratio setpoints in the MPC controller. Both the data reconciliation and optimization problems is solved using a software package for constrained optimization problems (NPSOL from Stanford University). This system runs on a DEC-Alpha computer and the average solution time is 3 minutes.

Conclusion

A process model describing the mass and energy balance is developed and used for data reconciliation and optimization. The model is fitted to the measured values and optimal feed split fractions are computed and implemented in the control system once an hour. The reconciled values provides valuable information about the current condition of the measurement equipment and of the condition of the heat exchangers. Comparison of reconciled values and measured values have detected several flow measurements with pour performance and also a temperature measurement that was found to be installed in the wrong pipe. The evolution of heat transfer coefficients during operation is also used to detect fouling and schedule cleaning of the heat exchangers. The model is sufficiently detailed for optimization purposes and the predicted optimal heater inlet temperature is achieved in the process.

Acknowledgments

The authors acknowledge the support of the process control staff at the Statoil Mongstad Refinery and the Statoil Research Centre for all help in implementing this system as a Septic MPC and RTO application.

References

Chen, X., R. W. Pike, T. A. Hertwig, and J. R. Hopper, "Optimal implementation of on-line optimization," *European Symposium on Computer Aided Process Engineering*, pages 435–442 (1998).

Crowe, C. M., Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of process flow rates by matrix projection, Part I: Linear case," *AIChE J.*, **29**(6), 881–888 (1983).

Crowe, C. M., "Reconciliation of process flow rates by matrix projection, Part II: Nonlinear case," *AIChE J.*, **32**(4), 881–888 (1986).

Johnston, L. P. M. and M. A. Kramer, "Maximum likelihood data rectification: Steady state systems," *AIChE J.*, **41**(11), 2415–2426 (1995).

Mills, A. F., *HEAT AND MASS TRANSFER*, The Richard D. Irwin series in heat transfer. RICARD D. IRWIN, INC, first edition (1995).

Tjoa, I. B. and L. T. Biegler, "Simultaneous strategies for data reconciliation and gross error detection of nonlinear systems," *Comput. Chem. Eng.*, **15**(10), 679–690 (1991).

Analysis of a Class of Statistical Techniques for Estimating the Economic Benefit from Improved Process Control

Kenneth R. Muske* and Conner S. Finegan
Department of Chemical Engineering
Villanova University, Villanova, PA 19085–1681

Abstract

This paper analyzes three statistical techniques commonly applied to estimate the potential benefit from improved process control. A brief overview of process control benefit estimation and a derivation of the statistical approaches used to estimate process control benefits is presented. The assumptions made in the derivation are outlined and the merits of each technique are discussed. The results from applying these techniques are then compared to the value actually obtained with improved process control using a semi-continuous distillation process.

Keywords

Process control improvement, Process control benefit estimation, Process control economic analysis

Introduction

Estimation of the economic benefit that can be obtained from implementing advanced process control and/or control system upgrades is essential for both justification and prioritization of these projects. Process control improvements result in both qualitative and quantitative benefits. The qualitative benefits can include more efficient evaluation of process and control system performance, improved access to real-time and historical process information, and better evaluation and management of abnormal conditions. Although improving the process operation, it can often be difficult to accurately and consistently assign a direct economic value to these benefits. The quantitative benefits can include improved energy efficiency, increased production rate, and decreased off-specification production. The direct economic benefits are more straightforward to determine for these cases. In this work, we consider three statistical techniques used for *a priori* estimation of the quantitative economic benefit from improved process control. These techniques have commonly been applied to justify the investment in process control projects for the petroleum and petrochemical industries (Tolfo, 1983), (Sivasubramanian and Penrod, 1990), (Martin et al., 1991), and (Latour, 1992).

These techniques are only applicable to controlled variables with an operating constraint or product specification limit and an economic incentive to operate as close as possible to this limit without excessive violation. Separation processes that require a minimum product purity are examples of a controlled variable with a product specification limit. A fired gas heater constrained by maximum tube skin or flue gas temperatures is an example of a controlled variable with an operating constraint.

Process Control Benefit Estimation

Quantitative estimation of the economic benefit from improved process control begins with determining the *base operation*. Process data, which includes the key economic controlled variables, are collected during a period of normal closed-loop operation. This collection period should be representative of the typical closed-loop operation of the process with the current control system. If the process is operated at a number of different conditions, a base operation is developed for each operating condition. The base operation mean value and variance for the controlled variables are determined from this data.

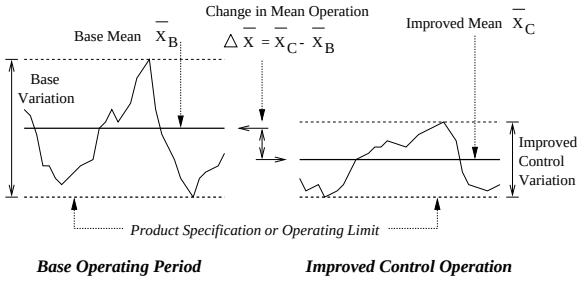


Figure 1: Improved process control operation.

Process control improvements are expected to reduce the variance of the controlled variables. Because of the reduction in the controlled variable variation, the mean operating value can be shifted closer to the product specification or operating constraint without increasing the frequency of violation. This operation is referred to as the *improved control operation* as shown in Figure 1.

The economic benefit is realized from operation at this new mean value. Quantification of the economic benefit is performed by using some form of a process model to determine the steady-state material and energy balance changes resulting from the improved control operation. Economic values are then used to estimate the monetary benefit. If there are a number of different operating conditions, this analysis is carried out for each and an average monetary benefit is determined by weighting the benefit realized from each operating mode by the fraction of time the process operates in that mode.

Predicting the change in the steady-state mean op-

*To whom all correspondence should be addressed. Phone: (610) 519-6195, Email: krmuske@kayak.che.villanova.edu

erating value obtainable from improved process control is performed in a variety of ways. Heuristic approaches include assuming the controller can consistently achieve the same performance as the best base operating data (Tolfo, 1983), assuming the controller can operate exactly at the specification or constraint (Martin et al., 1991), and assuming the process can be described by a low-order polynomial function (Stout and Cline, 1976). In this work, we discuss statistical techniques that determine the change in the mean operating value based on the reduction in the controlled variable variance.

Additional benefit may also be achieved from reducing the controlled variable variation. For many processes, such as polymerization, the product quality specifications are set by the desired end-use properties and cannot be changed due to a variance reduction. In these cases, the quantitative economic benefit comes from the reduction in variation alone. The statistical estimation techniques presented here are not appropriate for estimating this benefit and it will not be considered further.

Statistical Estimation Background

The probability that a normally distributed random variable X is less than a given value X_L is $P(X < X_L)$

$$P(X < X_L) = P(Z < Z_L) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_L} \exp\left(-\frac{Z^2}{2}\right) dZ$$

$$Z = \frac{X - \bar{X}}{S_X}, \qquad Z_L = \frac{X_L - \bar{X}}{S_X} \tag{1}$$

in which Z is the standard normal variable, $\bar{X}$ is the mean of X, and S_X is the standard deviation of X. The probability that a normally distributed random variable X is greater than a given value X_L is $P(X > X_L) = 1 - P(Z < Z_L) = P(Z < -Z_L)$ where the equalities follow from the symmetry of the normal distribution. The standard normal variable is used in these relationships to determine the probability since the integral is independent of the mean and variance of the process data.

Assuming that $P(Z < Z_L)$ also represents the fraction of samples below a given value X_L in a time series realized from a normally distributed stochastic process, $F_{\max}(Z_L) = P(Z < Z_L)$, in which $F_{\max}(Z_L)$ represents the fraction of samples in the base or improved control time series that is below the maximum limit. Under the same assumptions used in the maximum limit case, $F_{\min}(Z_L) = P(Z < -Z_L)$ where $F_{\min}(Z_L)$ represents the fraction of samples in the time series that is above the minimum limit. Note that Z_L is a negative number in the minimum limit case since the limit is less than the mean. Therefore, the fraction of samples in the time series below a maximum limit or above a minimum limit can be expressed as

$$F(Z_L) = P(Z < |Z_L|) \tag{2}$$

in which Z_L is the standard normal limit value determined from the actual process limit X_L using Equation 1.

The standard deviation of the base and improved control operation is required to determine the normalized limit values. The base operation variance, S_B^2, is determined from the base operating data. Assuming that the sensor noise is independent of the controlled variable, $S_B^2 = S_P^2 + S_M^2$, which is the sum of the contribution from the variance of the process, S_P^2, and the variance of the sensor, S_M^2. Under the same assumption, the improved control variance is the sum of the process variance after implementing the process control improvements and the sensor variance, $S_C^2 = S_I^2 + S_M^2$, in which the improved process variance is some function of the base process variance, $S_I^2 = \boldsymbol{f}(S_P^2)$. The measurement variance is typically neglected in this analysis since it is usually much smaller than the process variance. The exception is new and/or improved controlled variable sensors.

Estimating the reduction in the process variance obtainable from control system improvement is typically based on heuristics and prior experience. This reduction will depend on the process, the current control system, and the control system improvement under study. A typical assumption for advanced control implementation is a 50% reduction in the controlled variable variance (Sharpe and Latour, 1986) or standard deviation (Martin et al., 1991). Reductions in the standard deviation as large as 90% have been claimed (Tolfo, 1983). A lower bound on the achievable variance can also be estimated from the base operating data using controller performance assessment techniques (Qin, 1999) and a multiple of this value used. In this work, the improved control variance is determined from the experimental data and this value is used with each of the techniques.

Statistical Estimation Techniques

In this section, three published statistically-based techniques for determining an improved control mean operating value are presented. Each uses the base operation mean and variance and an improved control variance estimate. These techniques implicitly assume that the controlled variable time series for the base and improved control operation are realized from a strictly stationary, normally distributed, stochastic process. In addition, it is implicitly assumed that the controlled variable setpoints are not changed during the base operating period. If the setpoints change, there will be a contribution to the base variance due to the tracking control action. In this case, the setpoint deviation should be analyzed and these techniques implemented on a differential basis.

Method 1: Equal Operation at the Limit

The first method is referred to as *equal operation at the limit*. In this method, the improved control operation is required to respect the limit or specification the same

fraction of the time as the base operation. Since the variation in the improved control operation is reduced, the mean operation can be moved closer to the limit. This technique is claimed to be the most common method to estimate the change in the mean operation for product quality controlled variables (Sharpe and Latour, 1986) and is applicable if an acceptable fraction of the base operation violates the limit or specification (Sharpe and Latour, 1986), (Martin et al., 1991).

The fraction of time that the limit is respected for the base and improved control operation is determined using Equation 2. Assuming that these fractions are equal results in $F(Z_L^B) = F(Z_L^C)$ in which the superscript B refers to the base operation and the superscript C refers to the improved control operation. It then follows that

$$Z_L^B = Z_L^C \quad \Rightarrow \quad \frac{X_L - \bar{X}_B}{S_B} = \frac{X_L - \bar{X}_C}{S_C} \tag{3}$$

The change in mean operation is determined from Equation 3.

$$\Delta\bar{X} = \bar{X}_C - \bar{X}_B = \left(1 - \frac{S_C}{S_B}\right)(X_L - \bar{X}_B) \tag{4}$$

Method 2: Final Fractional Violation

The second method, referred to as *final fractional violation*, is recommended when the base operation does not violate the limit or specification (Sharpe and Latour, 1986). In this case, the improved control operation is allowed to violate the limit a specified fraction of the time f. The fraction of the time the improved control operation respects the limit is then $1 - f$ resulting in $F(Z_L^C) = 1 - f \quad \Rightarrow \quad Z_L^C = \alpha$ in which α is determined from f using the standard normal distribution $P(Z < |\alpha|) = 1 - f$. The change in the mean operation is determined from $Z_L^C = (X_L - \bar{X}_C)/S_C = \alpha$.

$$\Delta\bar{X} = \bar{X}_C - \bar{X}_B = X_L - \bar{X}_B - \alpha(S_C) \tag{5}$$

Note that *equal operation at the limit* is a special case of this method when the fractional violation and the distributions for the base and improved control operation are assumed to be the same. Substituting $(X_L - \bar{X}_B)/S_B$ for α in Equation 5 produces the expression in Equation 4.

Method 3: Equal Fractional Violation

The third method, referred to as *equal fractional violation*, is recommended when a significant fraction of the base operation data violates the specified limit (Sharpe and Latour, 1986), (Martin et al., 1991). In this method, the limit is replaced with one that results in a more reasonable fraction of violation by the base operation. The improved control operation is then allowed the same fractional violation of this new limit. The percent violation suggested to determine the new limit for this method is 5% (Sharpe and Latour, 1986), (Sivasubramanian and Penrod, 1990), (Martin et al., 1991), although no justification for the selection of this value is given.

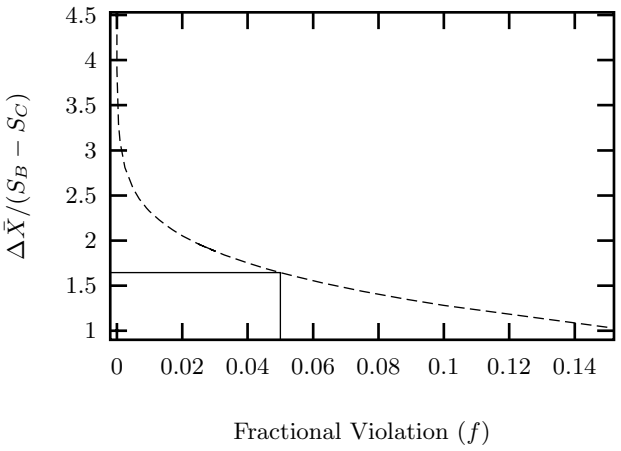


Figure 2: Equal fractional violation sensitivity to f.

The new limit, X_L^f, is that value violated by some fraction f of both the base and improved control operation. Therefore, the fraction of the time the base and improved control operation respect this limit is $1 - f$ resulting in $F(Z_f^B) = F(Z_f^C) = 1 - f$. It then follows that

$$Z_f^B = Z_f^C = \alpha \quad \Rightarrow \quad \frac{X_L^f - \bar{X}_B}{S_B} = \frac{X_L^f - \bar{X}_C}{S_C} = \alpha \tag{6}$$

The change in the mean operation is determined by eliminating the unknown limit X_L^f from Equation 6.

$$\Delta\bar{X} = \bar{X}_C - \bar{X}_B = \alpha(S_B - S_C) \tag{7}$$

The value of the new limit or specification, X_L^f, for this method is $\bar{X}_B + \alpha(S_B)$ which depends on the choice of f. The suggested value for f is 0.05 resulting in $|\alpha| = 1.645$. Note that the value of α is the ratio of the change in the mean to the change in the standard deviation, $\Delta\bar{X}/(S_B - S_C)$. As shown in Figure 2, this ratio is quite sensitive to f as the allowable violation is reduced.

Discussion of the Statistical Estimation Methods

We begin our discussion by suggesting that the *equal fractional violation* method is not an appropriate estimation technique. This method is recommended when a significant fraction of the base operation violates the specified process limit. If this limit is violated too often during the base operation, it is either not the true process limit or the base operation control system is functioning poorly. If it is not the true process limit, a more realistic limit should be determined based on process engineering, operation, and economics. It should not come from this *ad hoc* statistical procedure. If the specified limit is the true process limit, the improved control operation should be determined from the fractional violation of this

limit. There is no economic justification for the use of a different limit to estimate the benefit in this case.

The first method, *equal operation at the limit*, implicitly assumes that the base and improved control operation have the same distribution. Applying the same distribution to both the base and improved control operation may be a very poor assumption in many cases. The base operation fractional violation is also obtained from the base operating data mean and variance. The actual base operation fractional violation is not used. If this computed fractional violation deviates significantly from the actual fractional violation, this method can produce erroneous results. Finally, if equal violation of the limit is not acceptable, this method is not appropriate. These issues can limit the applicability of this technique.

The second method, *final fractional violation*, only requires an assumed distribution for the improved control operation. For advanced control applications, a normal distribution is often a reasonable assumption. The base operation fractional violation can be determined directly from the base operating data. If equal operation at the limit is desired, α in Equation 5 can be determined based on this value. If the base operation very seldom or never violates the limit, a value of α from a larger fractional violation can be used. If the limit is violated too often, the specified limit can be verified and either a more realistic value determined or a value of α based on more reasonable fractional violation selected.

Experimental Investigation

We present the results of an experimental investigation to compare the predicted mean operation from the three estimation techniques to the actual improved control operation of a semi-continuous distillation process.

Process Description

The process is a twenty tray ethanol/water distillation column used to produce concentrated ethanol from a dilute feed. The overhead ethanol product is recycled back to the process. The operating objective of the column is to maximize the recovery of ethanol subject to a minimum 74 wt% purity limit for the distillate product.

The improved control system for the column is shown in Figure 3. The column differential pressure is controlled by manipulating the reboiler steam flow rate. The differential pressure target is set slightly below the value in which jet flooding occurs in order to maximize separation. The concentration of ethanol in the distillate product is measured by an on-line density meter and used to reset the distillate to feed ratio target. Reflux is determined by the overhead liquid level. Since the total overhead liquid capacity in the system is very small, the composition responds quickly to changes in the distillate flow. The principal disturbances to the column are steam quality and reflux flow rate. The feed rate and composition were constant for this study. Base operation composition control was accomplished by manually adjusting the distillate to feed ratio target. The base operation differential pressure controller was poorly tuned.

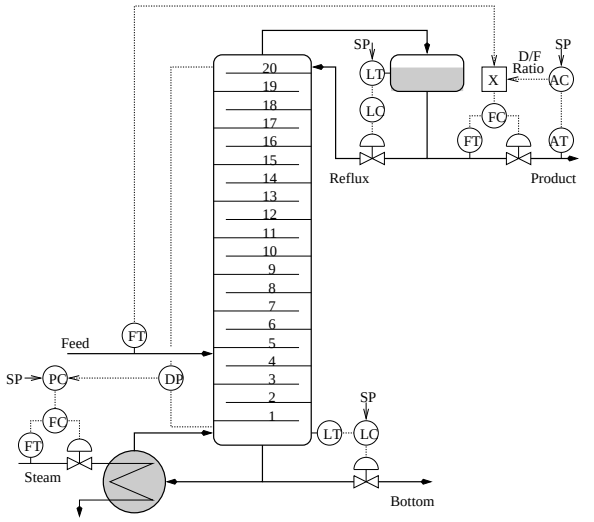


Figure 3: Distillation column control scheme.

Operation	*Mean*	*Std. Dev.*	*% Violation*
Base	79.82 wt%	2.495	4.2
Improved	74.54 wt%	0.263	4.3

Table 1: Base and improved control operation.

Process Data

Overhead composition data was sampled every two minutes for both the base and improved control operation. Base operating data was collected for 240 minutes which is approximately the normal operating cycle for the column. Improved control data was collected for 90 minutes due to adjustments made to the control system. However, we believe that this data is representative of an operating cycle with improved control. Table 1 presents the mean, standard deviation, and percent violation of the distillate product purity limit for both the base and improved control operating cycles. This study is based on the comparison between these two operating cycles.

The normalized base operating data distribution presented in Figure 4 is a bimodal distribution. Assuming a normal distribution is not appropriate in this case. The normalized improved control data distribution is presented in Figure 5. It more closely resembles a normal distribution although the tail below the mean is skewed.

Experimental Results

The improved control mean distillate composition predicted by each benefit estimation method is presented in Table 2. The value in parentheses for methods 2 and 3 is

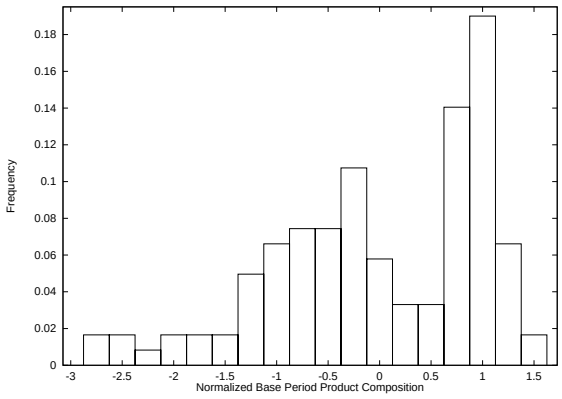


Figure 4: Distribution for base operation.

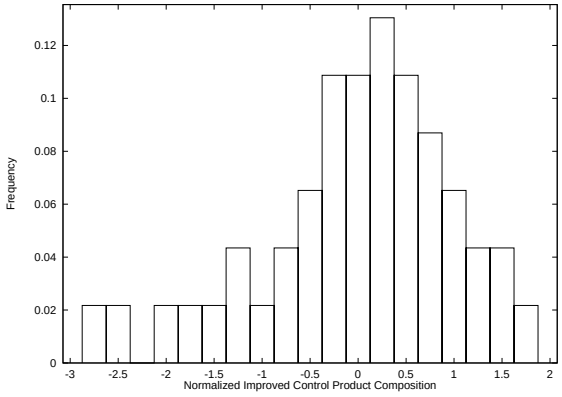


Figure 5: Distribution for improved control operation.

Benefit Estimation Method			
1	2 (5%)	2 (4.5%)	3 (5%)
74.62	74.43	74.44	76.15

Table 2: Estimated improved control mean comparison.

the percent violation specified for the method. The experimental base and improved control variance was used to determine these predictions. The actual improved control mean distillate composition was 74.54 wt%.

The first two methods, *equal operation at the limit* and *final fractional violation*, predict improved control mean distillate compositions quite close to the actual value. Since the first method assumes the same distribution for each operation, this result appears to be due to the interaction between the actual distributions in this case and is not believed to be a general result. The second method slightly under predicts the mean. Since the improved control distribution contains a larger fractional area that violates the limit than a normal distribution, this result is expected. The prediction from the third method, *equal fractional violation*, deviates significantly. The new limit assumed by this method is 75.72 wt% which helps explain the extent of the over prediction.

Conclusions

Three statistically-based techniques for *a priori* estimation of the mean operating controlled variable value after the application of process control improvements were presented. The *equal fractional violation* method is not recommended since the constraint limit or specification changes depending on the choice of the fractional violation. The *equal operation at the limit* method has limited applicability due to the restrictive assumptions made for the base operating data and equal fractional violation of the base and improved control operation. The *final fractional violation* method is the most general requiring an assumed distribution only for the improved control operation. The predictive capability of this method depends on how well the assumed distribution describes the improved control operation and how well the achievable improved control variance can be estimated.

References

Latour, P. R., "Quantify quality control's intangible benefits," *Hydrocarbon Process.*, **71**, 61–68 (1992).

Martin, G. D., L. E. Turpin, and R. P. Cline, "Estimating control function benefits," *Hydrocarbon Process.*, **70**, 68–73 (1991).

Qin, S. J., "Control performance monitoring—A review and assessment," *Comput. Chem. Eng.*, **23**(2), 178–186 (1999).

Sharpe, J. H. and P. R. Latour, Calculating real dollar savings from improved dynamic control, Technical report, SetPoint, Inc. (1986).

Sivasubramanian, R. and W. V. Penrod, "Computer control for third world contries," *Hydrocarbon Process.*, **69**, 84C–84L (1990).

Stout, T. M. and R. P. Cline, "Control system justification," *Intech*, **64**, 57–59 (1976).

Tolfo, F., A methodology to access the economic returns of advanced control projects, In *Proceedings of the 2nd American Control Conf.*, pages 315–319 (1983).

Connection between Model Predictive Control and Anti-Windup Control Schemes

Michael Nikolaou* and Mohan R Cherukuri
Dept. of Chemical Engineering
University of Houston
Houston, TX 77204-4792

Abstract

We show that the general Anti-Windup-Bumpless-Transfer (AWBT) controller structure naturally emerges from the structure of Model Predictive Control (MPC) with input constraints and plant model structure that is linear or nonlinear affine in the input variables. The key to establishing that relationship between AWBT control and MPC is a particular interpretation of the maximum principle.

Keywords

Model predictive control, Anti-windup-bumpless-transfer control, Maximum principle, Input saturation, Nonlinear control

Introduction

Controller design for linear or nonlinear processes with actuator saturation nonlinearities has long been studied within various contexts (Kothare, 1997). There are two distinct classes of control structures that handle input saturation nonlinearities: (a) On-line optimization based control structures, such as Model Predictive Control (MPC), and (b) anti-windup bumpless transfer (AWBT) controllers that have a closed form and do not perform on-line optimization. If properly designed, MPC can provide optimality, robustness, and other desirable properties. However, because of the time needed to perform the on-line optimization, MPC is usually implemented on relatively slow processes. On the other hand, AWBT controllers completely bypass on-line optimization; therefore they inherently have lower computational requirements and can be used on faster processes.

The AWBT controller design approach is based on the following two-step design paradigm: Firstly, a linear controller is designed ignoring input constraints. In the next step, an anti-windup scheme is added to compensate for the adverse effects of input constraints on closed-loop performance. Campo (1997) and Kothare et al. (1994) unified all heuristically developed AWBT control schemes into the structure shown in Figure 1, and developed a general framework for studying stability and robustness issues. The importance of that work lies in that model uncertainty can be taken into account systematically and theory exists to analyze, at least in principle, the closed-loop system for stability and robustness. However, that analysis is also based on the standard conic sector nonlinear stability theory. Therefore, the results could be potentially conservative. The design of AWBT controllers for SISO systems relies on a mix of intuitive and rigorous arguments, which become difficult to use in the MIMO case (Peng et al., 1998). As pointed by Doyle et al. (1987), for MIMO controllers, the saturation may cause a change in the plant input direction resulting in disastrous consequences. Through an example, Doyle et al. (1987) showed that all anti-windup schemes of the time failed to work on MIMO systems. Recently, Kothare and Morari (1997) described three performance requirements that should be incorporated in a multi-objective multivariable AWBT controller synthesis framework. Although promising lines for designing an AWBT controller using dynamic output feedback and one-step design were outlined, many of the details like "recovery of linear performance" need to be worked out.

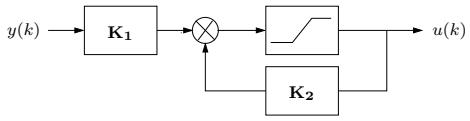


Figure 1: Classical AWBT controller structure.

The MPC design approach naturally and explicitly handles multivariable input and output constraints by directly incorporating them into the on-line optimization problem. The issues of stability and robustness of MPC are now a fairly well understood topic (Rawlings and Muske, 1993; Mayne et al., 2000; Nikolaou, 2000).

For linear plants the MPC problem can be reduced to a quadratic program (QP) which can be solved efficiently (Cutler and Ramaker, 1980; García and Morshedi, 1986). Alternatively, to reduce the computational load, MPC may use a cascaded on-line optimization approach in which a steady-state target is first calculated on-line via linear programming (cost minimization) and then an unconstrained least-squares problem steers the controlled system towards the steady-state optimum (Kassmann et al., 2000; Rao and Rawlings, 1999). It should be stressed that the solution of the least-squares problem (inputs to the controlled system) should satisfy constraints, even though the latter are not explicitly considered in the least-squares problem. Least-squares solution inputs that do not satisfy constraints are sim-

*Author to whom all correspondence should be addressed. Email address: nikolaou@uh.edu. Phone: (713) 743-4309. Fax (713) 743-4323.

ply clipped, clearly a non-optimal solution.

MPC for nonlinear plants naturally leads to nonlinear programs (NLP) which are in general non-convex and computationally demanding (Biegler and Rawlings, 1991; Mayne and Michalska, 1991, 1990; Kreshenbaum et al., 1994; Qin and Badgwell, 2000).

Both MPC and AWBT controllers each have their own advantages and disadvantages. However, to the best of our knowledge, no clear relationship between these two control schemes has been established.

In this work, we rigorously show that there is a direct relationship between MPC and AWBT control. In fact, we show that the heuristically proposed AWBT structure of Figure 1 naturally emerges from the structure of MPC with quadratic objective, input constraints, and (linear or nonlinear) plant model structure affine in the input variables. This realization is important for a number of reasons:

- It provides theoretical justification for the heuristically proposed AWBT structure of Figure 1.
- It allows direct substitution of MPC with input constraints by controllers with a closed-form structure, which allows computations to be performed significantly faster. Computational efficiency has been pursued by several other investigators using a variety of different approaches, such as approximation of the on-line optimization (Zheng, 1999), or a priori determination of active constraints (De Dona and Goodwin, 2000; Bemporad et al., 2000).
- It facilitates the design of both MPC and AWBT controllers, because the insight into a controller from either class can be augmented by using insight into an equivalent controller from the other class.
- It allows constrained least squares to be used with computational efficiency in MPC systems that follow a cascaded structure of linear programming followed by (unconstrained) least squares (Kassmann et al., 2000; Rao and Rawlings, 1999).

The proposed approach works for linear models as well as nonlinear models in which the input appears affinely. It also works equally well for SISO and MIMO systems.

From MPC to AWBT

MPC and On-line Optimization

Consider a discrete-time non-linear system in which the input $u(i)$ appears affinely in the right-hand side of the system difference equation, i.e.:

$$\begin{aligned} x(i+1) &= f[x(i)] + g[x(i)]u(i) + d(i), \quad x(0) = x_0 \\ y(i) &= h[x(i)] \end{aligned} \tag{1}$$

where $x(i) \in \Re^n$ is the state vector, $u(i) \in \Re^m$ is the control vector, $d(i) \in \Re^n$ is the disturbance vector, $f[x(i)] \in \Re^n$, and $g[x(i)] \in \Re^{n\times m}$. For the linear case the system is

$$\begin{aligned} x(i+1) &= \Phi x(i) + \Gamma u(i) + d(i), \quad x(0) = x_0 \\ y(i) &= Cx(i) \end{aligned} \tag{2}$$

where $\Phi \in \Re^{n\times n}$ and $\Gamma \in \Re^{n\times m}$.

The vector of manipulated variables is constrained as

$$u_{\min} \le u(i) \le u_{\max} \tag{3}$$

where $u_{\min}$ and $u_{\max}$ are real vectors.

To simplify the discussion, we assume that $x(i)$ is measured.

According to standard MPC practice, the optimization problem to be solved at time step k is

$$\min_{z,v} \sum_{i=0}^{N} \left[z_d^T(k+i) Q z_d(k+i) + \Delta v^T(k+i) R \Delta v(k+i) \right] \tag{4}$$

subject to

$$z_d(k+i+1) = z(k+i+1) - r(k+i+1) \qquad \text{(feedback error)}$$

$$z(k) = x(k) \qquad \text{(feedback measurement)}$$

$$\begin{cases} z(k+i+1) = f_m\left[z(k+i)\right] + g_m\left[z(k+i)\right] v(k+i) + d(k+i) \\ \text{or} \\ z(k+i+1) = \Phi_m z(k+i) + \Gamma_m v(k+i) + d(k+i) \end{cases} \qquad \text{(prediction)}$$

$$u_{min} \le v(k+i) \le u_{max} \qquad \text{(input constraints)}$$

where $i = 0, 1, \ldots, N-1$; Q and R are *diagonal* positive definite matrices; f_m and g_m are the nonlinear plant model vector functions; Φ_m and Γ_m are the linear plant model matrices; r is the desired state; z_d is the deviation from the desired state; x is the current state; $\Delta v(j) \equiv v(j) - v(j-1)$ is the change in control vector at time j; $d(k)$ is a load disturbance (bias) at time k estimated, for simplicity, as

$$d(k) = x(k) - \left(f_m[x(k-1)] + g_m[x(k-1)]u(k-1)\right)$$

for the nonlinear case, or

$$d(k) = x(k) - \left(\Phi_m x(k-1) + \Gamma_m u(k-1)\right)$$

for the linear case.

The controlled system is assumed to be controllable. Of the N control moves $\hat{v}(k), \ldots, \hat{v}(k+N-1)$ computed at time k, only the first one is implemented: $u(k) \equiv \hat{v}(k)$. At the next time instant $k+1$, when the new value for state $x(k+1)$ becomes available, the minimization of Equation 4 is performed with the new initial condition, to provide $u(k+1)$.

Necessary conditions satisfied by the solution of the optimization problem of Equation 4 can be obtained using the discrete maximum principle (Polak and Jordan, 1964; Halkin, 1966). Note that Boltyanskii (1978, pg. 54) has showed that not all formulations of the discrete

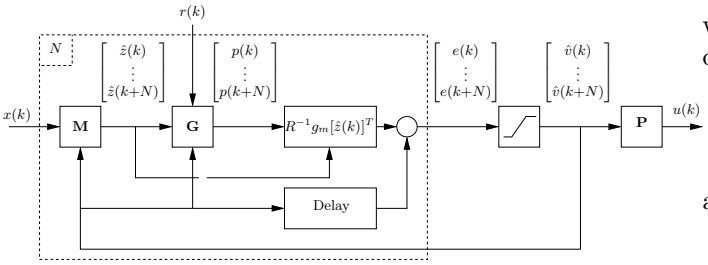


Figure 2: Block diagram for the system of Equations 5 through 10.

maximum principle that have appeared in literature are correct.

Next, we apply the discrete maximum principle to the MPC control problem with input bounds, in order to obtain an analytical solution. The maximum principle is a necessary condition. Under the additional assumption that the problem is convex, the maximum principle is also a sufficient condition for optimality (Theorem 47.7 in Boltyanskii, 1978).

The Discrete Maximum Principle and MPC

We apply the maximum principle to the optimization problem of Equation 4. There are two crucial points to stress:

- The matrix R in Equation 4 is *diagonal*, and
- The input constraints set is $U \equiv \{v | u_{\min} \leq v \leq u_{\max}\}$.

Under these conditions, we get the following necessary conditions for the optimal solution:

System Equations:

$$\begin{aligned}\hat{z}(k+i+1) - \hat{z}(k+i) &= f_m\left[\hat{z}(k+i)\right] \\ &\quad - \hat{z}(k+i) + g_m\left[\hat{z}(k+i)\right]\hat{v}(k+i) + \underbrace{d(k)}_{\substack{=x(k)-f_m(x(k-1))- \\ g_m(x(k-1))u(k-1)}} \qquad (5)\\ &\equiv F_i\left(\hat{z}(k+i), \hat{v}(k+i)\right), \qquad i = 0, \ldots, N-1 \\ \hat{z}(k) &= x(k)\end{aligned}$$

Adjoint equations:

$$\begin{aligned}p(k+i) = &\underbrace{\left(\frac{\partial f_m[\hat{z}(k+i)]}{\partial z} + \sum_{j=1}^{m} \frac{\partial g_{m,j}[\hat{z}(k+i)]}{\partial z}\hat{v}_j(k+i)\right)^T}_{\eta[\hat{z}(k+1),\hat{v}(k+i)]} p(k+i+1) \\ &- Qz_d(k+i), \quad i = 0, \ldots, N-1 \qquad (6)\end{aligned}$$

where $g_{m,j}$ denotes the j^{th} column of the matrix g_m.

Transversality equation:

$$p(k+N) = 0 \qquad (7)$$

Minimization of the Hamiltonian:

$$\hat{v}(k+i) = sat\left[\hat{v}(k+i-1) + R^{-1}g_m[\hat{z}(k+i)]^T p(k+i+1)\right] \qquad (8)$$

where $\hat{v}(k-1) = u(k-1)$ and the saturation function is defined in a standard way, i.e., for $u \in \Re^m$,

$$sat(u) \equiv \left[sat(u_1) \ \cdots \ sat(u_m)\right]. \qquad (9)$$

Based on the above, the input to the controlled process at time k is

$$u(k) = \hat{v}(k). \qquad (10)$$

A Revealing Block Diagram

Equations 5 through 10 correspond to a static (algebraic) system, namely knowledge of $x(k)$ is, in principle, sufficient for computation of everything else. Figure 2, shows a block diagram interpretation of Equations 5 through 10.

The block **M** is the following set of algebraic equations:

$$\begin{aligned}\hat{z}(k+1) &= f_m[x(k)] + g_m[x(k)]v(k) + d(k) \\ \hat{z}(k+2) &= f_m[\hat{z}(k+1)] + g_m[\hat{z}(k+1)]\hat{v}(k+1) + d(k) \\ &\vdots \\ \hat{z}(k+N) &= f_m[\hat{z}(k+N-1)] \\ &\quad + g_m[\hat{z}(k+N-1)]\hat{v}(k+N-1) + d(k)\end{aligned}$$

where

$$d(k) = x(k) = f_m(x(k-1)) - g_m(x(k-1))u(k-1). \qquad (11)$$

The block **G** is the following set of algebraic equations:

$$\begin{aligned}p(k+N) &= 0 \\ p(k+N-1) &= \eta\left[\hat{z}(k+N-1), \hat{v}(k+N-1)\right] p(k+N) \\ &\quad - Qz_d(k+N-1) \\ &\vdots \\ p(k) &= \eta\left[\hat{z}(k), \hat{v}(k)\right] p(k+1) - Qz_d(k)\end{aligned}$$

The block **P** is the projection matrix

$$\mathbf{P} = \begin{bmatrix} I_m & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{N \times N}.$$

Remarks.

- An essential trait of the structure in Figure 2 is that there is a feedback loop with the saturation function in the forward path and algebraic equations in both the forward and feedback paths. This structure bears strong resemblance to the general AWBT structure of Figure 1 (Kothare, 1997). However, unlike the classical AWBT structure, which is based on experience, the structure of Figure 2 emerges naturally as a result of formulating the controller design problem through MPC. In addition, Figure 2, provides guidelines on how to design an AWBT controller starting from MPC.

- Figure 2 implies that

$$e = N(Se, x, r), \quad \hat{v} = Se, \quad u = \mathbf{P}\hat{v} \tag{12}$$

where S denotes the saturation function (Equation 9). On the other hand, if the saturation block were placed right after the output of the feedback loop, before $\mathbf{P}$, in Figure 2 (i.e., if a saturation block were appended to a controller designed without taking input constraints into account), then we would have

$$e = N(e, x, r), \quad \hat{v} = Se, \quad u = \mathbf{P}\hat{v} \tag{13}$$

The above equations make it clear that the second alternative, Equation 13 is different from the first one, Equation 12 and show what is the missing element in controller design that does not take saturation explicitly into account during the design. The above comment will be more concrete in the linear case, discussed below.

On-line Implementation

The controller structure of Figure 2, albeit optimal in the MPC sense, is not suitable for direct on-line implementation, the reason being that the set of algebraic equations 12 must be solved at each time step. A time-recursive set of equations would be required, so that $\hat{v}(k)$ could be computed from data up to and including time k. To circumvent that difficulty, we use the following heuristic:

Let the optimal input sequence computed at time $k-1$ be $\{\hat{w}(k-1), \hat{w}(k), \ldots, \hat{w}(k+N-2)\}$. Then at time k we use

$$\{\hat{v}(k), \hat{v}(k+1), \ldots, \hat{v}(k+N-2), \hat{v}(k+N-1)\} = \{\hat{w}(k), \hat{w}(k+1), \ldots, \hat{w}(k+N-2), \hat{w}(k+N-2)\} \tag{14}$$

in Equations 5 and 8. This heuristic introduces a memory (delay) in the feedback path of Figure 2, thus making the structure suitable for on-line implementation. Note that the choice $\hat{v}(k+N-1) = \hat{w}(k+N-2)$ implicitly assumes that the optimal input sequence over the finite optimization horizon reaches a virtually flat profile towards the end of the horizon.

Analytical Solution for a Linear System with Quadratic Objective

Consider the system of Equation 1 and corresponding model used in Equation 4. In this case, the vector of predicted optimal states satisfies the following equation:

$$\begin{bmatrix} \hat{z}(k) \\ \vdots \\ \hat{z}(k+N) \end{bmatrix} = \underbrace{\begin{bmatrix} I \\ \Phi_m \\ \vdots \\ \Phi_m^N \end{bmatrix}}_{\mathbf{L}_2} \otimes x(k) + \underbrace{\begin{bmatrix} 0 & \cdots & \cdots 0 \\ I & \ddots & \vdots \\ \Phi_m & \ddots & \ddots \\ \vdots & \ddots & \ddots \vdots \\ \Phi_m^{N-1} & \cdots & \Phi_m^T \; I \; 0 \end{bmatrix}}_{\mathbf{L}_2} \otimes \Gamma_m \begin{bmatrix} \hat{v}(k) \\ \vdots \\ \hat{v}(k+N) \end{bmatrix} + \begin{bmatrix} d(k) \\ \vdots \\ d(k) \end{bmatrix}$$

where $d(k)$ is estimated as in Equation 11 and $\otimes$ denotes the Kronecker product. The adjoint equations, Equation 6, imply that the costate vectors , satisfy the following equations:

$$\begin{bmatrix} p(k+N) \\ \vdots \\ p(k) \end{bmatrix} = - \underbrace{\begin{bmatrix} 0 & \cdots\cdots 0 \\ I & \ddots \vdots \\ \Phi_m^T & \ddots \ddots \vdots \\ \vdots & \ddots \ddots 0 \\ \Phi_m^{T^{N-1}} & \cdots \Phi_m^T I \end{bmatrix}}_{\mathbf{L}_3} \otimes Q \begin{bmatrix} \hat{z}(k+N-1) \\ \vdots \\ \hat{z}(k) \end{bmatrix} \tag{15}$$

The minimization of the Hamiltonian, Equation 8, yields

$$\hat{v}(k+i) = sat[\hat{v}(k+i-1) + R^{-1}\Gamma_m^T p(k+i+1)], \quad i = 0, \ldots, N-1. \tag{16}$$

Again, the approximation of Equation 14 can be used.

Remarks.

- It is straight forward to modify the previous discussion for the objective function in Equation 4 containing a term quadratic in v instead of Δv. The well-known advantage of using the Δv is that step disturbance or setpoint changes result in zero offset.
- Figure 2 indicates that

$$e = L_x x + L_r r + L_v \hat{v} = L_x x + L_r r + L_v S e \tag{17}$$

$$\Rightarrow (I - L_v S)e = L_x x + L_r r \tag{18}$$

Because the quadratic minimization of Equation 4 is convex, there must exist a unique optimal solution, which implies that the above equation must have a unique solution for e, i.e., $e = (I - L_v S)^{-1}(L_x x + L_r r)$, from which we get

$$u = \mathbf{P}\hat{v} = \mathbf{P}Se = \mathbf{P}S(I - L_v S)^{-1}(L_x x + L_r r) \tag{19}$$

It is interesting to note again that if the controller was designed without taking input saturation into

account, and a saturation block were appended to it, then the corresponding mapping between (x, r) and u would be

$$u = \mathbf{P}\hat{v} = \mathbf{P}Se = \mathbf{P}S(I - L_v)^{-1}(L_x x + L_r r) \quad (20)$$

The above equation trivially shows that this design approach is not optimal.

Conclusions

In this work we established a direct relationship between multivariable AWBT control and MPC with quadratic objective, input constraints and plant model structure affine in the input variables. The key to establishing that relationship was application of the discrete maximum principle to the on-line optimization problem solved by MPC.

The results of this work are important for both theoretical and practical reasons.

From a theoretical viewpoint, these results provide fundamental justification for the empirical realization that virtually all heuristically developed AWBT control structures (Figure 1) follow a similar pattern involving a nonlinear (saturation) block and linear transfer functions (Figure 2). The structure of Figure 2 is actually valid for nonlinear systems as well.

From a practical viewpoint, the substitution of model predictive controllers with input constraints by controllers with a closed-form structure allows computations to be performed significantly faster. This is particularly important for MPC systems that follow a cascaded structure of linear programming followed by (unconstrained) least squares (Kassmann et al., 2000; Rao and Rawlings, 1999), because it allows *constrained* least squares to be used with *computational efficiency* in place of unconstrained least squares.

Of both theoretical and practical importance is the fact that insight into a controller from either the MPC or AWBT class can be augmented by using insight from an equivalent or related controller of the other class.

A number of simulation examples can be downloaded from `http://athens.chee.uh.edu`.

References

Bemporad, A., M. Morari, V. Dua, and E. N. Pistikopoulos, The Explicit Linear Quadratic Regulator for Constrained Systems, In *Proc. American Contr. Conf.* (2000).

Biegler, L. T. and J. B. Rawlings, Optimization Approaches to Nonlinear Model Predictive Control, In Arkun, Y. and W. H. Ray, editors, *Chemical Process Control—CPCIV*, pages 543–571. CACHE (1991).

Boltyanskii, V. G., *Optimal Control of Discrete Systems*. John Wiley & Sons, Inc., New York (1978).

Campo, P. J., *Studies In Robust Control of Systems Subject to Constraints*, PhD thesis, California Institute of Technology (1997).

Cutler, C. R. and B. L. Ramaker, Dynamic Matrix Control—A Computer Control Algorithm, In *Proceedings of the Joint Automatic Control Conference*, San Francisco, California (1980).

De Dona, J. A. and G. C. Goodwin, Elucidation of the state-space regions wherein model predictive control and anti-windup strategies achieve identical control policies, In *Proceedings of the 2000 ACC*, pages 1924–1928 (2000).

Doyle, J. C., R. S. Smith, and D. F. Enns, Control of plants with input saturation nonlinearities, In *Proceedings of the 1987 American Control Conference*, pages 1034–1039 (1987).

García, C. E. and A. M. Morshedi, "Quadratic Programming Solution of Dynamic Matrix Control (QDMC)," *Chem. Eng. Commun.*, **46**, 73–87 (1986).

Halkin, H., "A Maximum Principle of the Pontryagin Type for Systems Described by Nonlinear Difference Equations," *SIAM J. Cont.*, **4**(1), 90–111 (1966).

Kassmann, D. E., T. A. Badgwell, and R. B. Hawkins, "Robust steady-state target calculation for model predictive control," *AIChE J.*, **46**(5), 1007–1024 (2000).

Kothare, M. V. and M. Morari, Multivariable anti-windup controller synthesis using multi-objective optimization, In *Proceedings of the 1997 American Control Conference* (1997).

Kothare, M. V., P. J. Campo, M. Morari, and C. N. Nett, "A Unified Framework for the Study of Anti-Windup Designs," *Automatica*, **30**, 1869–1883 (1994).

Kothare, M. V., *Control of Systems Subject to Constraints*, PhD thesis, California Institute of Technology (1997).

Kreshenbaum, L., D. Mayne, R. Pytlak, and R. Vinter, Receding horizon control, In Clarke, D., editor, *Advances in Model-Based Predictive Control*, pages 523–535. Oxford University Press (1994).

Mayne, D. Q. and H. Michalska, An implementable receding horizon controller for stabilization of nonlinear systems, In *Proc. 29th Conference on Decision and Control*, pages 3396–3397, Honolulu, Hawaii (1990).

Mayne, D. and H. Michalska, Approximate global linearization of nonlinear systems via on-line optimization, In *Proceedings of the ECC*, pages 182–187, Grenoble, France (1991).

Mayne, D. Q., J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained Model Predictive Control: Stability and Optimality," *Automatica*, **36**(6), 789–814 (2000).

Nikolaou, M., *Model Predictive Controllers: A Critical Synthesis of Theory and Industrial Needs*, Advances in Chemical Engineering Series. Academic Press (2000).

Peng, Y., D. Vrancic, R. Hanus, and S. S. R. Weller, "Anti-Windup Designs for Multivariable Controllers," *Automatica*, **34**(12), 1559–1565 (1998).

Polak, E. and B. W. Jordan, "Theory of a Class of Discrete Optimal Control Systems," *J. Electron. and Control*, **17**(6) (1964).

Qin, S. J. and T. J. Badgwell, An overview of nonlinear model predictive control applications, In Allgöwer, F. and A. Zheng, editors, *Nonlinear Model Predictive Control*. Birkhäuser, Basel, Switzerland (2000).

Rao, C. V. and J. B. Rawlings, "Steady States and Constraints in Model Predictive Control," *AIChE J.*, **45**(6), 1266–1278 (1999).

Rawlings, J. B. and K. R. Muske, "Stability of Constrained Receding Horizon Control," *IEEE Trans. Auto. Cont.*, **38**(10), 1512–1516 (1993).

Zheng, A., "Reducing on-line computational demands in model predictive control by approximating QP constraints," *J. Proc. Cont.*, **9**(4), 279–290 (1999).

Efficient Nonlinear Model Predictive Control: Exploiting the Volterra-Laguerre Model Structure

Robert S. Parker*
Department of Chemical and Petroleum Engineering
University of Pittsburgh
Pittsburgh, PA 15261

Abstract

An analytical solution to the nonlinear model predictive control (NMPC) optimization problem is derived for single–input single–output (SISO) systems modeled by second–order Volterra–Laguerre models. All input moves except the current move ($m > 1$ in the NMPC framework) are approximated by solving an unconstrained linear MPC problem which utilizes a locally accurate linear model of the process. This linear MPC problem has an analytical solution; this is substituted into a nonlinear equation which is solved exactly for the current input move, $\Delta u(k|k)$. Results using this multi–m NMPC formulation are superior to a previously developed analytical NMPC controller that required $m = 1$ (Parker and Doyle III, 1998).

Keywords

Bioreactor control, Nonlinear model predictive control, Volterra-Laguerre models

Introduction

Model predictive control (MPC) is a control algorithm of industrial and academic interest (Allgöwer et al., 1999; Biegler, 1998) that solves an optimization problem on-line at each time step. For processes that display highly nonlinear behavior, whether due to the operating conditions or nonlinear dynamics (*e.g.* input multiplicative processes and high-purity distillation), performance degradation or instability can result when linear control algorithms are utilized. The use of nonlinear MPC (NMPC) can diminish this performance loss while retaining the multivariable and constraint handling capabilities of MPC.

A high-fidelity nonlinear process model and an optimization routine capable of solving the on-line optimization problem in real-time are required to reap the benefits of the NMPC algorithm. The use of a fundamental process model is conceptually appealing in that the process physics can be explicitly incorporated. Unfortunately, these models require significant time (often measured in man-months or more) and effort to construct, and the resulting NMPC optimization problems are nonconvex and computationally unattractive for most realistic systems (Zheng, 1997; Mayne, 1996).

In place of a fundamental model one can substitute a nonlinear empirical model identified from process data. These data-driven models capture only the input-output behavior of the process, thereby sacrificing physical understanding for rapid model development. One popular model structure, and the model form used in this work, is the second-order Volterra model (Doyle III et al., 1995; Zheng and Zafiriou, 1995) given by the equation:

$$\hat{y}(k) = \sum_{i=1}^{M} h_1(i)u(k-i) + \sum_{i=1}^{M}\sum_{j=1}^{M} h_2(i,j)u(k-i)u(k-j) \quad (1)$$

This model structure can capture the behavior of fading memory nonlinear processes, such as the bioreactor case study examined below. Highly parameterized Volterra models can be efficiently projected onto the Laguerre basis to produce a Volterra-Laguerre model (Schetzen, 1980; Dumont et al., 1994; Zheng and Zafiriou, 1995):

$$\ell(k+1) = A(\alpha)\ell(k) + B(\alpha)u(k) \quad (2)$$

$$\hat{y}(k) = C^T\ell(k) + \ell^T(k)D\ell(k) \quad (3)$$

Nonconvex NMPC optimization problems result from the use of this model form (the objective function is 4^{th}-order in u). If only a single input move is of interest, the optimization problem can be solved analytically (Dumont et al., 1994; Parker and Doyle III, 1998). Another approach to this NMPC problem is to solve for only the current input ($u(k|k)$) exactly, and solve for any future moves approximately using linear MPC and a locally accurate process model (Zheng, 1997) because these future moves are never actually implemented. This paper addresses the synthesis of an NMPC controller which analytically calculates an input profile by combining elements of Parker and Doyle III (1998) and Zheng (1997). Although a benefit of the current work is its computational efficiency, the concept of approximating future manipulated variable moves is employed here because it facilitates the analytical solution to the $m > 1$ problem.

*email: rparker@engrng.pitt.edu, phone: +1-412-624-7364

Input-Output Model Identification

Volterra model identification is accomplished using Algorithm 3 from (Parker et al., 2001). This involves a decomposition of (1) as follows:

$$y(k) = h_0 + \mathcal{L}(k) + \mathcal{D}(k) + \mathcal{O}(k) \tag{4}$$

$$\mathcal{L}(k) = \sum_{i=1}^{M} h_1(i)u(k-i)$$

$$\mathcal{D}(k) = \sum_{i=1}^{M} h_2(i,i)u^2(k-i)$$

$$\mathcal{O}(k) = 2\sum_{i=1}^{M}\sum_{j=1}^{i-1} h_2(i,j)u(k-i)u(k-j)$$

Here the linear, second-order diagonal, and off-diagonal coefficient contributions are given by $\mathcal{L}$, $\mathcal{D}$, and $\mathcal{O}$, respectively. Tailored input sequences, which excite specific contributions and minimize others, provide superior coefficient identification compared to cross-correlation techniques (Parker et al., 2001). The identified Volterra model is then projected onto the Laguerre basis to address the noise sensitivity of Volterra models and simultaneously reduce the model parameterization (9 vs. 860 unique parameters for the Volterra-Laguerre and Volterra models, respectively). This projection yields a Volterra-Laguerre model which optimally approximates the identified Volterra model in the mean-squared error sense. The C vector and D matrix are calculated via least-squares from the identified Volterra kernels, and the Laguerre time-scale, $0 < \alpha \leq 1$, is selected to minimize the error between the identified Volterra model and the expanded Volterra-Laguerre model (where expansion is the inverse of the projection operation). The resulting process model is given by (2) and (3).

Controller Synthesis

The NMPC controller utilizes the standard squared 2-norm objective function given by:

$$\min_{\Delta\mathcal{U}(k|k)} \|\Gamma_y \left[\mathcal{R}(k+1) - \mathcal{Y}(k+1|k)\right]\|_2^2 + \|\Gamma_u \Delta\mathcal{U}(k|k)\|_2^2 \tag{5}$$

The solution developed below employs the formalism of solving for $\Delta\mathcal{U}$ as opposed to absolute $\mathcal{U}$. Matrices Γ_y and Γ_u are used to trade off setpoint tracking versus manipulated variable movement, respectively. The minimization problem in (5) is solved at each sample time for a series of m manipulated variable moves which minimize the objective over a prediction horizon of length p. An analytical solution to the $m = 1$ problem has been developed for SISO problems modeled using the Volterra-Laguerre structure (2) and (3) (Parker and Doyle III, 1998). Limits on the input-output dimension and m resulted from an inability to solve a third-order vector or matrix polynomial.

Utilizing the linear approximation of future manipulated variable moves (Zheng, 1997), the setpoint tracking term of (5) can be decomposed according to the following equation:

$$\min_{\Delta\mathcal{U}(k|k)} \|\Gamma_y \left[\mathcal{R}(k+1) - \mathcal{Y}_N(k+1|k) - \mathcal{Y}_L(k+1|k)\right]\|_2^2 + \|\Gamma_u \Delta\mathcal{U}(k|k)\|_2^2 \tag{6}$$

Here the terms $\mathcal{Y}_N$ and $\mathcal{Y}_L$ represent the contributions of the first calculated input move (nonlinear, exact) and the remaining $m-1$ future input moves (linear, approximate), respectively. The linear component of the problem can be formulated as the solution to a modified reference signal, $\mathcal{R}_L(k+1|k) = \mathcal{R}(k+1) - \mathcal{Y}_N(k+1|k)$:

$$\min_{\Delta\mathcal{U}_L(k+1|k)} \|\Gamma_{yL} \left[\mathcal{R}_L(k+1|k) - \mathcal{Y}_L(k+1|k)\right]\|_2^2 + \|\Gamma_{uL} \Delta\mathcal{U}_L(k+1|k)\|_2^2 \tag{7}$$

An analytical solution to this problem exists (García et al., 1989). The controller model is developed by combining the linear process dynamics (2) with the linearization of the process output (3). The resulting linear controller model is given by (8) and (9).

$$x_L(k+i|k) = \begin{cases} 0 & i = 1 \\ \sum_{j=1}^{i} \bar{A}_{m-j-1} B \Delta u(k+j|k) & 2 \leq i \leq m-1 \\ A^{i-m+1} \sum_{j=1}^{m-1} \bar{A}_{m-j-1} B \Delta u(k+j|k) & i \geq m \end{cases} = G\Delta\mathcal{U}_L(k+1|k) \tag{8}$$

$$y_L(k+i) = \left[C^T + 2x_{lin}^T D\right] x_L(k+i|k) = H x_L(k+i|k) \tag{9}$$

The matrix $\bar{A}_i = (A^{i-1} + A^{i-2} + \ldots + I)$, and the $m-1$ future input moves are given by $\Delta\mathcal{U}_L$. For input multiplicative processes, the x_{lin} vector must change with operating point because no linear integrating controller can stabilize an input multiplicative process at the optimum (Morari, 1983). In this work the matrix H was recalculated at each time step using a local linearization about the current process state, $x_L(k)$ (García, 1984). The G matrix is static, and hence calculated off-line in this work.

The solution to the linear problem is given by the following equation:

$$\Delta\mathcal{U}_L(k+1|k) = \mathcal{K}(\mathcal{R}(k+2) - \mathcal{Y}_N(k+2|k)) \tag{10}$$

where

$$\mathcal{K} = \left(G^T H^T \Gamma_{yL}^T \Gamma_{yL} H G + \Gamma_{uL}^T \Gamma_{uL}\right)^{-1} G^T H^T \Gamma_{yL}^T \Gamma_{yL} \tag{11}$$

The weighting matrices $\Gamma_{yL} = \Gamma_y(2:p, 2:p)$ and $\Gamma_{uL} = \Gamma_u(2:m, 2:m)$ are of dimension $p-1$ and $m-1$, respectively. Equation (10) represents a solution for $\Delta\mathcal{U}_L(k+1|k)$ in terms of only the first input move, $\Delta u(k|k)$. The remainder of the derivation, given below, conceptually follows Parker and Doyle III (1998).

Substituting (10) into (6), yields the objective function

$$\min_{\Delta u(k|k)} \mathcal{E}^T \left[\mathcal{I}^T\Gamma_y^T\Gamma_y\mathcal{I} + \mathcal{K}^T\Gamma_{uL}^T\Gamma_{uL}\mathcal{K}\right]\mathcal{E} + \Delta u^2(k|k)\Gamma_u^2(1,1) \tag{12}$$

where

$$\mathcal{E} = (\mathcal{R}(k+1) - \mathcal{Y}_N(k+1|k)) \tag{13}$$

$$\mathcal{I} = I_{p-1\times p-1} - HG\mathcal{K} \tag{14}$$

$$\mathcal{Y}_N(k+i|k) = \epsilon_0 + \epsilon_1\Delta u(k|k) + \epsilon_2\Delta u^2(k|k) \tag{15}$$

$$\epsilon_2(k+i|k) = B^T\bar{A}_i^T D\bar{A}_i B \tag{16}$$

$$\epsilon_1(k+i|k) = C^T\bar{A}_i B + 2\ell^T(k)\left(A^i\right)^T D\bar{A}_i B + 2\epsilon_2(k+i|k)u(k-1) \tag{17}$$

$$\epsilon_0(k+i|k) = C^T\left(A^i\right)\ell(k) + \ell^T(k)\left(A^i\right)^T D\left(A^i\right)\ell(k) + \epsilon_1(k+i|k)u(k-1) \tag{18}$$

The matrices $\mathcal{K}$ and $\bar{A}_i$ are defined as above, and I is an appropriately sized identity matrix. This objective function is a function of the current Volterra-Laguerre state, $\ell(k)$, the immediate past input, $u(k-1)$, and $\Delta u(k|k)$. If E_i is defined as

$$E_i = \begin{bmatrix} \epsilon_i(k+1|k) \\ \epsilon_i(k+2|k) \\ \vdots \\ \epsilon_i(k+p|k) \end{bmatrix} \tag{19}$$

then the optimal $\Delta u(k|k)$ can be calculated by differentiating equation (12) with respect to $\Delta u(k|k)$ and setting the result equal to zero:

$$0 = \xi_3\Delta u^3(k|k) + \xi_2\Delta u^2(k|k) + \xi_1\Delta u(k|k) + \xi_0 \tag{20}$$

where

$$\xi_3 = 2E_2^T\Gamma_{yN}^T\Gamma_{yN}E_2 \tag{21}$$

$$\xi_2 = 3E_1^T\Gamma_{yN}^T\Gamma_{yN}E_2 \tag{22}$$

$$\xi_1 = 2E_2^T\Gamma_{yN}^T\Gamma_{yN}E_0 + E_1^T\Gamma_{yN}^T\Gamma_{yN}E_1 - 2E_2^T\Gamma_{yN}^T\Gamma_{yN}\mathcal{R}(k+1) + \Gamma_u^2(1,1) \tag{23}$$

$$\xi_0 = E_1^T\Gamma_{yN}^T\Gamma_{yN}E_0 - E_1^T\Gamma_{yN}^T\Gamma_{yN}\mathcal{R}(k+1) \tag{24}$$

$$\Gamma_{yN} = \begin{bmatrix} \Gamma_y(1,1) & 0 \\ 0 & \mathcal{I}^T\Gamma_{yL}^T\Gamma_{yL}\mathcal{I} + \mathcal{K}^T\Gamma_{uL}^T\Gamma_{uL}\mathcal{K} \end{bmatrix} \tag{25}$$

Provided that the ξ_i's are real, a solution to (20) exists (Tuma, 1987). The solution method involves a substitution for $\Delta u(k|k)$, and interested readers are referred to (Tuma, 1987) as the complete derivation is omitted due to space limitations. The cubic equation solution involves a term, $\mathcal{D}$, which lies beneath a radical similar to $\sqrt{b^2-4ac}$ in the quadratic equation. Roots of (20) are determined by the value of $\mathcal{D}$ as follows:

$$\begin{array}{ll} \mathcal{D} > 0 & \text{one real root, 2 complex roots} \\ \mathcal{D} = 0 & \text{3 real roots, at least 2 equal} \\ \mathcal{D} < 0 & \text{3 real unequal roots} \end{array} \tag{26}$$

All real roots are transformed back to the initial problem space and are analyzed for optimality by simulating the Volterra-Laguerre model over the prediction horizon to calculate an objective function value (OFV). The $\Delta u(k|k)$ which minimizes the OFV is implemented.

The analytical NMPC algorithm developed in this work retains the constraint handling capabilities of the algorithm developed previously (Parker and Doyle III, 1998). However, only the first (nonlinear) input move is constrained. In this formulation, solutions which cause the manipulated variable to violate imposed magnitude constraints are replaced by the constraint itself, and the OFV is recalculated. The unconstrained solution of the linear controller subproblem is used. This is done to maintain the existence of an analytical solution to the control problem, (5). Constraint handling on the future moves can be implemented in a "soft" framework, where Γ_{uL} can be tuned independently from $\Gamma_u(1,1)$, so that the linear moves are penalized for large magnitude changes. Although the algorithm developed in the current work has relaxed the limitation of $m = 1$ imposed on the analytical NMPC algorithm in Parker and Doyle III (1998), it is still limited to SISO problems. Formulation of the multivariable problem with $m > 1$ is feasible in the context of this algorithm. However, equation (20) would be a third-order vector polynomial, and an analytical solution for this problem requires further work. Rate constraints are not included in this formulation because they are not relevant to the case study below, but incorporation of rate constraints into the nonlinear problem would be straightforward.

Case Study: Continuous Bioreactor

A model for the growth of *Klebsiella pneumoniae* on glucose in a continuous-flow bioreactor was developed by Baloo and Ramkrishna (1991). Cell biomass exit concentration (g/L) was the output of interest, and dilution rate (hr^{-1}) was the manipulated variable. The nominal operating condition used in this study was 0.97 hr^{-1} yielding a biomass concentration of 0.2373 g/L.

From this system, a second-order Volterra-Laguerre model was identified using Algorithm 3 from Parker et al. (2001). The sample time was 15 minutes, and the Volterra model memory was $M = 40$. The inputs and outputs were scaled according to $u = \frac{u_{actual}-u_{nominal}}{0.08}$ and $y = \frac{y_{actual}-y_{nominal}}{0.01}$. The input sequence amplitude for identification was 2.375. After projection, the result-

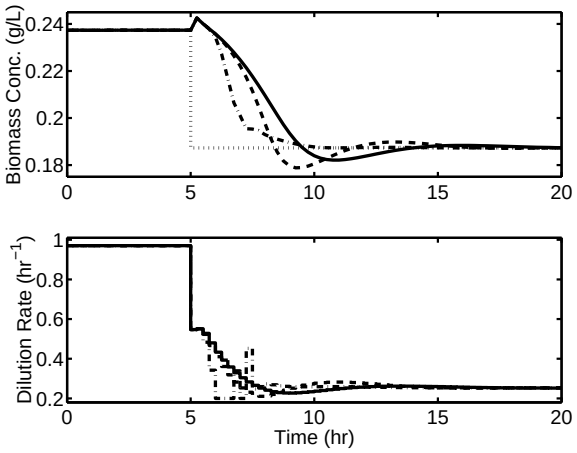


Figure 1: Response to a -0.05 g/L change in the reference (dotted) at $t = 5$ hr for various move horizon lengths: $m = 1$ (solid), $m = 14$, (dashed), and $m = 14$ w/ RLS (dash-dot). Other tuning parameters: $p = 16, \Gamma_y = I_{p\times p}, \Gamma_u(1,1) = 0, \Gamma_{uL} = I_{m-1\times m-1}$.

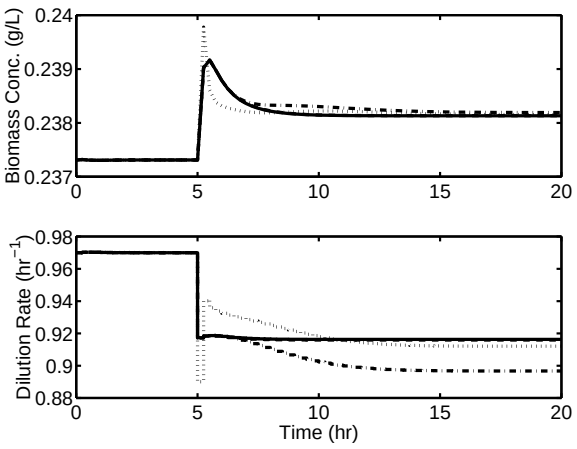


Figure 2: Response to a +0.025 g/L change in the reference at $t = 5$ hr for various move horizon lengths: $m = 1$ (solid), $m = 14$, (dashed), and $m = 14$ w/ RLS (dash-dot), and nonlinear programming MPC with $m = 14$ and RLS (dotted). Other tuning parameters: $p = 16, \Gamma_y = I_{p\times p}, \Gamma_u(1,1) = 0, \Gamma_{uL} = 0.1 I_{m-1\times m-1}$.

ing Volterra-Laguerre model ((2), (3)) with $\alpha = 0.59$ had the following matrices:

$$A = \begin{bmatrix} 0.59 & 0 & 0 \\ 0.652 & 0.59 & 0 \\ -0.385 & 0.652 & 0.59 \end{bmatrix} \tag{27}$$

$$B = [0.807 \quad -0.476 \quad 0.281]^T \tag{28}$$

$$C = [-0.249 \quad 0.128 \quad 0.002]^T \tag{29}$$

$$D = \begin{bmatrix} -0.050 & -0.006 & 0.010 \\ -0.006 & -0.023 & 0.008 \\ 0.010 & 0.008 & -0.013 \end{bmatrix} \tag{30}$$

Validation resulted in unbiased residuals of less than 0.5 mg/L, such that in the region of identification, the identified model and the actual differ by less than 2%.

Partial motivation for developing the analytical NMPC controller capable of handling $m > 1$ is the expectation that the use of larger move horizons would result in more aggressive controller response and therefore improved performance. Magnitude constraints of $0.2 \leq u(k) \leq 1.1$ hr^{-1} were imposed on the dilution rate, so that the cells did not starve or get washed out of the reactor, respectively. The response of the system under analytical NMPC control to a step change of -0.05 g/L in the reference signal is shown in Figure 1. The nonlinear programming MPC solution to this problem is not shown, because it is trapped in a local minimum at the high dilution rate constraint, and is therefore unable to track the setpoint change. The increase in move horizon improves tracking performance by 5%. A more aggressive response to the offset observed before $t = 10$ hr can be seen in the manipulated variable. Typical of more aggressive controllers, greater undershoot is also observed. The addition of a recursive least-squares (RLS) algorithm to update the Volterra-Laguerre model on-line is straightforward, and the C and D matrices are updated using a standard algorithm (Ljung, 1987). The improvement in the model accuracy is evidenced by the lack of undershoot and reduced oscillation around the reference. The more accurate model improves the controller prediction, thereby improving performance by 14% versus $m = 14$ without RLS and by 18% versus the $m = 1$ analytical NMPC controller. It should be noted that changes in the tuning weights, Γ_y and Γ_u, will change the degree of performance improvement.

In the case of an unreachable setpoint, the analytical NMPC algorithm performs as shown in Figure 2. All controllers remained stable. There was little observable difference between the $m = 1$ and $m = 14$ controllers. The full nonlinear solution implemented with RLS (nonlinear programming MPC, the dotted line) outperformed the controllers without RLS, but the analytical solution controller with $m = 14$ and RLS reached a steady state closest to the actual process optimum (dilution rate= 0.89hr^{-1}). It is possible that the full nonlinear solution could approach the analytical solution if the measurement signal were not noise-free, which limited the ability of the algorithm to update the model after $t = 10$ hr.

Summary

Given the nonlinear nature of many processes (*e.g.* bioreactors and CSTRs) and the nonlinear behaviors which result from operating in certain regimes (*e.g.* high purity

distillation), nonlinear MPC can offer significant performance improvements. The use of data-driven models in NMPC facilitates model development, and the choice of a particular structure can further simplify controller synthesis. By exploiting the structure of the Volterra-Laguerre (or equivalently, Volterra) model, an analytical solution to the NMPC problem was derived. Although this solution is not exact for $m > 1$, as it includes an approximation for the future input moves, significant performance improvement was observed in comparison with controllers synthesized using $m = 1$. A standard recursive least-squares algorithm, used in conjunction with the analytical NMPC controller, led to superior performance due to the ability of the algorithm to further update the model based on the current operating point.

Acknowledgments

This work was initiated while the author was at the University of Delaware and was partially funded by the University of Delaware Process Control and Monitoring Consortium. Additional funding has been provided by the Lai Family Foundation. The author would like to thank F. J. Doyle III for his insightful comments on this work.

References

Allgöwer, F., T. A. Badgwell, J. S. Qin, J. B. Rawlings, and S. J. Wright, *Advances in Control—Highlights of ECC '99*, chapter Nonlinear Predictive Control and Moving Horizon Estimation—An Introductory Overview, pages 391–449. Springer, London (1999).

Baloo, S. and D. Ramkrishna, "Metabolic Regulation in Bacterial Continuous Cultures: I," *Biotech. Bioeng.*, **38**, 1337–1352 (1991).

Biegler, L. T., Efficient Solution of Dynamic Optimization and NMPC Problems, In *International Symposium on Nonlinear Model Predictive Control: Assessment and Future Directions*, pages 46–63, Ascona, Switzerland (1998).

Doyle III, F. J., B. A. Ogunnaike, and R. K. Pearson, "Nonlinear Model-based Control Using Second-order Volterra Models," *Automatica*, **31**, 697–714 (1995).

Dumont, G. A., Y. Fu, and G. Lu, Nonlinear Adaptive Generalized Predictive Control and Applications, In Clarke, D., editor, *Advances in Model—based Predictive Control*, pages 498–515. Oxford University Press (1994).

García, C. E., D. M. Prett, and M. Morari, "Model Predictive Control: Theory and Practice—A Survey," *Automatica*, **25**(3), 335–348 (1989).

García, C. E., Quadratic Dynamic Matrix Control of Nonlinear Processes. An Application to a Batch Reaction Process, In *Proc. AIChE Annual Meeting*, San Francisco, CA (1984).

Ljung, L., *System Identification: Theory for the User*. P T R Prentice Hall, Englewood Cliffs, NJ (1987).

Mayne, D. Q., Nonlinear Model Predictive Control: An Assessment, In *Proceedings of the Fifth International Conference on Chemical Process Control*, Tahoe City, CA (1996).

Morari, M., Robust Stability of Systems with Integral Control, In *Proceedings of the IEEE Conference on Decision and Control*, pages 865–869, San Antonio, TX. IEEE Press (1983).

Parker, R. S. and F. J. Doyle III, Nonlinear Model Predictive Control of a Continuous Bioreactor at Near-Optimum Conditions, In *Proc. American Control Conf.*, pages 2549–2553, Philadelphia, PA (1998).

Parker, R. S., D. Heemstra, F. J. Doyle III, R. K. Pearson, and B. A. Ogunnaike, "The Identification of Nonlinear Models for Process Control Using Tailored "Plant-Friendly" Input Sequences," *J. Proc. Cont.*, **11**(Special Issue SI), 237–250 (2001).

Schetzen, M., *The Volterra and Wiener Theories of Nonlinear Systems*. John Wiley & Sons, New York, NY (1980).

Tuma, J. J., *Engineering Mathematics Handbook*. McGraw-Hill Inc., New York, NY (1987).

Zheng, Q. and E. Zafiriou, Nonlinear System Identification for Control Using Volterra-Laguerre Expansion, In *Proc. American Control Conf.*, pages 2195–2199, Seattle, WA (1995).

Zheng, A., A Computationally Efficient Nonlinear MPC Algorithm, In *Proc. American Control Conf.*, pages 1623–1627, Albuquerque, NM (1997).

Partial Differential Equation Model Based Control: Application to a Bleaching Reactor

Stéphane Renou* and Michel Perrier
Département de génie chimique
École Polytechnique de Montréal
C.P. 6079, succ. Centre-ville Montréal, H3C 3A7, Canada

Denis Dochain
CESAME, Université Catholique de Louvain
Av. G. Lemaître 4-6, 1348 Louvain-La-Neuve, Belgium

Sylvain Gendron
PAPRICAN, 570 Blvd St-Jean
Pointe-Claire, H9R 3J9, Canada

Abstract

Reactor engineering generally uses distributed parameter models for design purpose. These models are not often used for process control design. May the use of this kind of complex models improve control performance? This paper compares different control strategies based on a distributed parameter model to a time-scaled DMC that only uses a simple input-output model for the control of a bleaching reactor.

Keywords

Distributed parameter systems, Late-lumping control, Early-lumping control, DMC, Bleaching reactor

Introduction

The design of tubular reactors is usually performed by using mass and energy balances on a thin slice of the reactor. This modelling approach leads to partial differential equation models. However, process control practice for this type of reactor often uses lumped models such as first order plus delay transfer functions. Could there be some advantage to use the distributed parameter model for control purposes? On one hand, answering this question is easier when actuators and sensors are also distributed like in furnace heat control. Using a distributed parameter model then allows to use all the information in a structured manner. On the other hand, when sensors and actuators are only present at boundaries, performance enhancement using a distributed parameter model is not obvious. This question will be explore in this paper on a bleaching reactor application.

The bleaching process is the last step of pulp preparation. Its purpose is to improve the brightness of the pulp to a specified level which fulfills customers needs. The control objective for a bleaching reactor is then to obtain the desired brightness with a minimum output brightness variance at the lowest chemical cost. Traditional approaches to this control problem include variations around compensated brightness and scheduling, but the increase of computer power and the introduction of on-line analyzers offer new possibilities for model-based control such as directly using the PDE model.

Different models for the bleaching are presented in the literature for control purposes. Traditionally transfer function or other input-output models are used. But the need for more complex models is pointed out with the use of mixed model. Barrette and Perrier (1995) use multiple CSTR and Wang et al. (1995) use a combination of CSTR and PFR. Recently, a PDE model have been proposed by Renou et al. (2000b).

Various approaches have been considered to use a PDE phenomenological model directly. Ray (1981) proposed to divide control approaches on PDEs in two groups . The first group is composed of early lumping methods where a preliminary discretization of the PDE model is used to obtain a set of ODEs. This lumping is often realized by numerical techniques such as finite difference, orthogonal collocation or finite elements. Christofides (1996) has used the Galerkin method for the control of parabolic PDE. Early lumping techniques also includes the use of global differentiation proposed by Dochain (1994) as an approximation of partial derivative. This approach have been applied to hyperbolic PDE on a bioreactor by Bourrel (1996) and on a bleaching reactor model by the authors (Renou et al., 2000b). The second group of techniques is based on late lumping methods where the controller design problem is solved directly with the PDE model. When necessary, lumping may be applied for controller implementation. Christofides (1996) has used this approach in the case of hyperbolic system with a distributed control action. The control of parabolic PDE has been previously addressed by Hong and Bentsman (1994). They provide a design solution for systems in which the control action appears explicitly in the PDE system. For the boundary control problem, the authors have proposed a direct adaptive control strategy in Renou et al. (2000a) for the linear case.

The objective of this study is to present some results on the use of more complex models to enhance control performance. For this purpose, an early lumping and a late lumping strategy are compared to a simple time-scaled Dynamic Matrix Control (DMC) algorithm. The first section of the paper presents the PDE model development for a ClO_2 bleaching reactor. The second section briefly show the design ideas for each controller. The fol-

*Now with GE Corporate Research & Development, General Electric Company, Building KW Room D-211, P.O. Box 8 Schenectady NY 12301. Email: renou@crd.ge.com.

lowing section presents the main comparative results in terms of response to flow variations, step point changes and kinetic parameter mismatch.

Bleaching Reactor Model

The bleaching process for chemical pulp consists of extracting lignin from wood fibre. This brownish colored complex polymer is responsible for wood fiber coloration. It could be degraded by using a strong oxidant like ClO_2. A PDE model for this process reactor can be obtained by mass balances on lignin (L) and ClO_2 (C) on a thin section of the reactor. The following space axial dispersion model is then obtained:

$$\begin{aligned}\frac{\partial L(z,t)}{\partial t} &= -v\frac{\partial L(z,t)}{\partial z} + D\frac{\partial^2 L(z,t)}{\partial z^2} \\ &\quad - r_L(L(z,t), C(z,t)) \qquad (1)\\ \frac{\partial C(z,t)}{\partial t} &= -v\frac{\partial C(z,t)}{\partial z} + D\frac{\partial^2 C(z,t)}{\partial z^2} \\ &\quad - r_C(L(z,t), C(z,t)) \qquad (2)\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial L(z,t)}{\partial z}\right|_{z=0} &= \frac{v}{D}(L(0,t) - L_{in}(t))\\ \left.\frac{\partial C(z,t)}{\partial z}\right|_{z=0} &= \frac{v}{D}(C(0,t) - C_{in}(t))\\ \left.\frac{\partial C(z,t)}{\partial z}\right|_{z=\ell} &= \frac{\partial C_{out}}{\partial z} = 0\\ \left.\frac{\partial L(z,t)}{\partial z}\right|_{z=\ell} &= \frac{\partial L_{out}}{\partial z} = 0\end{aligned} \qquad (3)$$

In this model, reaction kinetics explicitly appears and can be identified by laboratory batch experiments. Hydrodynamical parameters v and D can be determined by tracer analysis. Here the kinetic data obtained by Savoie and Tessier (1997) have been considered and hydrodynamical parameters were deduced from Pudlas et al. (1999) as shown in Renou et al. (2000b). The following kinetics model and hydrodynamical parameters are used:

$$r_{L1}(L,C) = -k_L\ C^2L^2 = -0.0065\ C^2L^2 \qquad (4)$$

$$r_{C1}(L,C) = -k_C\ C^2L^2 = -0.0010\ C^2L^2 \qquad (5)$$

$$v = 1\,m/s,\ D = 0.001\,m^2/s \qquad (6)$$

for a 30 meter tower. Finally we consider the inlet concentration of ClO_2, C_{in}, and the lignine concentration at the outlet, L_{out}, as the manipulated variable and the controlled variable, respectively. Lignine and ClO_2 measurement are assumed to be available at the both ends of the reactor.

Time-Scaled DMC

The DMC controller is designed using two dynamic matrices: β_{CL} for ClO_2 input to lingin output response and

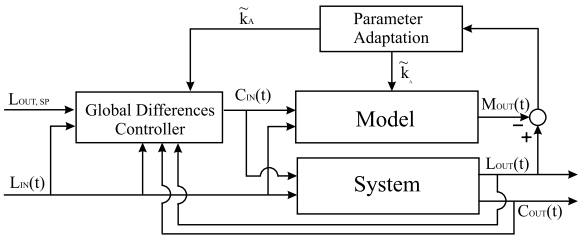


Figure 1: Global differences controller.

β_{LL} for lignin input to lignin output response. At each control step, the following criterion is applied:

$$\begin{aligned}\min_{\Delta u(k)} \phi = [e(k+1) - \beta_{CL} + K^2\Delta u(k)^T\Delta u(k)^T]^T\\ [e(k+1) - \beta_{CL} + K^2\Delta u(k)^T\Delta u(k)^T]\end{aligned} \qquad (7)$$

The prediction error takes the input lignin disturbances into account such as:

$$\begin{aligned}e(k+1) \equiv y^*(k+1)\\ - [\hat{y}^0(k) + w(k+1) + \beta_{LL}\Delta L(k)]\end{aligned} \qquad (8)$$

In the preceding equation, y^* is the set point, $\hat{y}^0$ is the prediction if no further control action is taken, w is the estimation of disturbance and ΔL is the variation of lignin at the inlet. Traditional DMC is sensitive to flow rate variations since they represent, in fact, a variation of dead-time from an input-output point of view. To overcome this problem efficiently, information about the flow rate has to be transmitted to the controller. To reach this goal, the prediction time span is scaled by the variation of flow rate. Thus, the Δt between each calculation of the control action is scaled by the ratio between the new flow rate and the old flow rate. This approach can be practically implemented by using oversampling or interpolating dynamic matrices and prediction.

Early Lumping Approach

One of the problems with the PDE model described by Equations 1–3 is that the control action does not appear explicitly in the PDE equations. Dochain (1994) have proposed to use global differences as an approximation for space partial derivatives. This early lumping approach introduces ClO_2 input and lignin output in an approximate model. An exact linearization approach of this model can then be considered. The following approximation are used for both species:

$$\frac{\partial L(1,t)}{\partial z} = \frac{L(1,t) - L(0,t)}{\Delta z} = L_{out}(t) - L_{in}(t) \qquad (9)$$

$$\begin{aligned}\frac{\partial^2 L(1,t)}{\partial z^2} &= \frac{L(2,t) - L(1,t) - L(0,t)}{\Delta z^2}\\ &= L_{in}(t) - L_{out}(t)\end{aligned} \qquad (10)$$

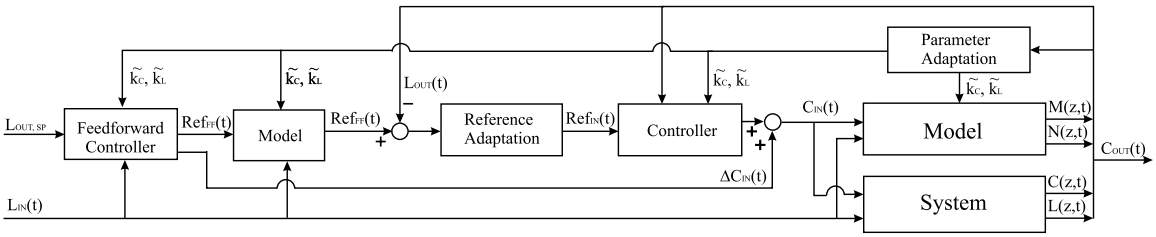


Figure 2: Late lumping internal model controller.

To obtain the approximate model, system mass balances are expressed at the reactor outlet, global differences are introduced and both equations are combined by their kinetic term. These operations give the following result:

$$\frac{dL_{out}(t)}{dt} = -v(L_{out}(t) - L_{in}(t)) + D(L_{in}(t) - L_{out}(t)) + \frac{k_{L1}}{k_{C1}}\Big(\frac{dC_{out}(t)}{dt} + v(C_{out}(t) - C_{in}(t)) - D(C_{in}(t) - C_{out}(t))\Big) \quad (11)$$

Using a backward finite difference to approximate the ClO_2 time derivative, an input-output relation between ClO_2 input and lignin output can be obtained. Exact linearization principle can be applied on this equation to obtain the following control law in which λ and γ are external loop tuning parameters:

$$C_{in}(t) = \frac{1}{v+D}\Big(vC_{out}(t) + DC_{out} + \frac{C_{out}(t) - C_{out}(t-1)}{\Delta t} + \frac{k_{C1}}{k_{L1}}\Big[u + v(L_{out}(t) - L_{in}(t)) + D(L_{in} - L_{out})\Big]\Big) \quad (12)$$

$$u(t) = \lambda\Big[(L_{sp} - L_{out}(t)) + \gamma\int_0^t (L_{sp} - L_{out}(t))dt\Big] \quad (13)$$

To insure more robustness to this algorithm, an adaptation mechanism is added for the reaction rate ratio as shown in Figure 1. A model is simulated in parallel with a variable k_{L1}, noted k_A, which is modified according to the error between the adaptation model and the system model on lignin using a linear first order filter.

Late Lumping Approach

To use the whole information of the PDEs model an internal model approach is considered. The error between the model and the system is then used in direct adaptive control scheme. To account for lignin inlet variation a feedforward compensation is added to this controller. The feedforward controller action is divided in two parts. The first part uses an internal model of the process to give an estimation of the reference output to the controller. The second part directly gives a correction of ClO_2 needed to compensate for the deviation of lignin from the nominal operating point. Those calculation are based on a relaxation algorithm. Figure 2 shows the proposed control structure.

The controller design is performed using the Lyapunov second method following the approach presented in Renou et al. (2000a). The controller and adaptation laws are defined by:

$$\dot{\tilde{C}}_{in} = \frac{(C_{in} - Ref_{in})^2}{w + (C_{in} - Ref_{in})^2}\tilde{C}_{in} + \frac{(C_{in} - Ref_{in})}{w + (C_{in} - Ref_{in})^2}f_c \quad (14)$$

$$\dot{C}_{in} = \tilde{C}_{in} \quad (15)$$

$$f_c = \langle e_C, -v\frac{\partial e_C}{\partial z}\rangle + \langle e_L, -v\frac{\partial e_L}{\partial z}\rangle + \langle e_C, \tilde{k}_C Q\rangle + \langle e_L, \tilde{k}_L Q\rangle + \frac{\theta}{\epsilon}(C_{in} - Ref_{in})\dot{Ref}_{in} \quad (16)$$

$$\dot{Ref}_{in} = -\theta(L_{out}(t) - Ref_{ff}(t)) \quad (17)$$

$$\dot{\tilde{k}}_L = -a\langle e_L, Q\rangle \quad (18)$$

$$\dot{\tilde{k}}_C = -b\langle e_C, Q\rangle \quad (19)$$

using the following error functions:

$$e_L(z,t) = L(z,t) - M(z,t) \quad (20)$$

$$e_C(z,t) = C(z,t) - N(z,t) \quad (21)$$

$$Q(z,t) = L^2C^2 - M^2N^2 \quad (22)$$

In those equations, $M(z,t)$ and $N(z,t)$ are the lignin and chlorine dioxide model profiles, respectively. The controller law uses, or implementation purposes, only information from sensors at both ends of the reactor. Overall this control structure will behave as a feedforward controller if the model match the system. Otherwise the feedback part will account for model mismatch.

Simulation Results

Numerical simulation of the control algorithm applied to the system has been performed using a sequencing algorithm with a 100 node mesh (Renou et al., 2000c). In this algorithm, convection, dispersion and reaction

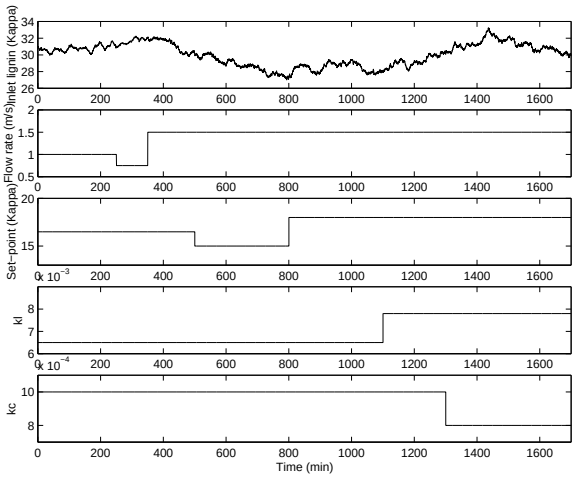


Figure 3: Operating conditions variations.

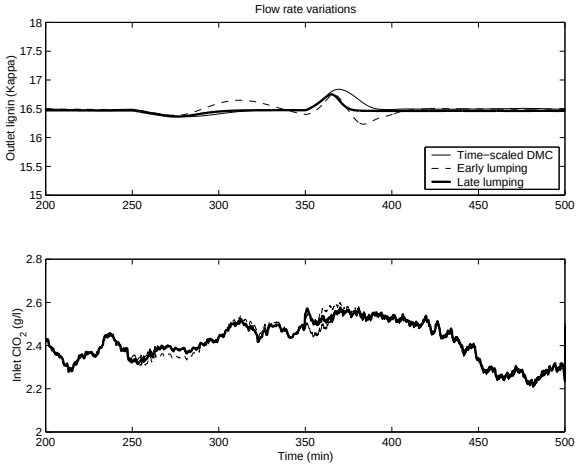


Figure 4: Flow rate variations.

phenomena are successively considered at each time step. The controller parameters have been chosen to minimize overshoot and oscillations. Simulations are started at steady state with $L_{in} = 31\ Kappa$ and $C_{in} = 2.35\ g/l$. The Kappa index is a measure of pulp whiteness. A sequence of events is applied to deviate the process from its nominal operating point as shown in in Figure 3.

Figure 4 shows the response of the system to flow rate variations. In each case the response of the controller to the variation of the delay is adequate. This result is guaranteed in the DMC case by the time-scaling of the model. In PDE based models, the time delay is implicit, and therefore, including flow rate variation directly in the control law accounts for time delay variation. The late lumping controller gives the less important deviation from set point in transient.

Figure 5 shows the response of the system to set-point variations. In this simulation, time-scaled DMC and late lumping controller give similar results that match open-loop dynamics of the reactor. The early lumping con-

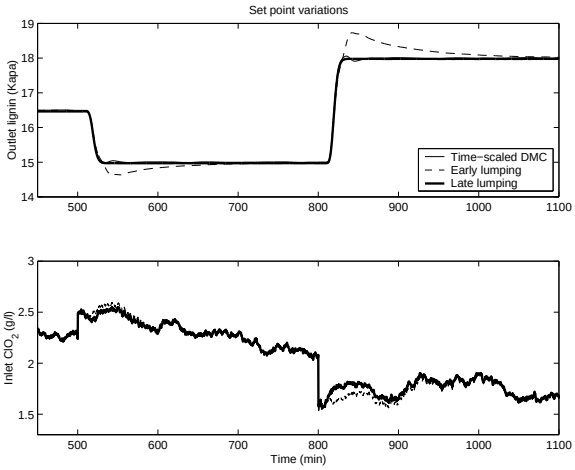


Figure 5: Set-point variations.

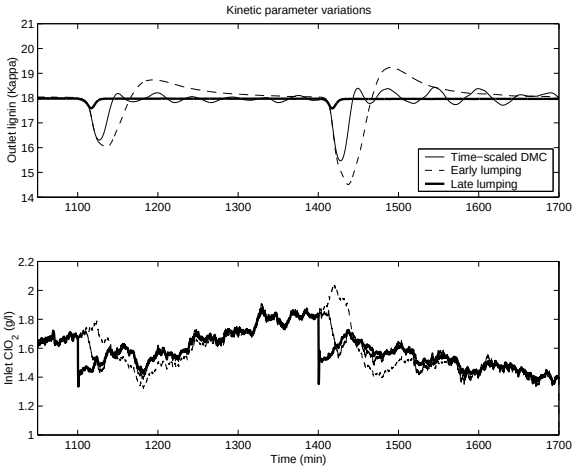


Figure 6: Model perturbations.

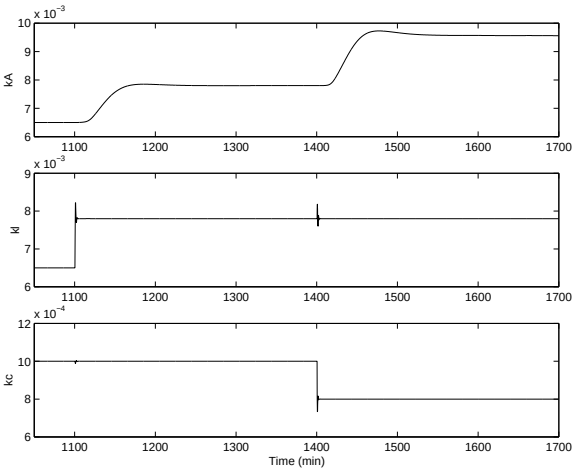


Figure 7: Kinetic parameter estimations.

troller exhibits an overshoot in case of set-point variation : this is due to the use of important simplifications in the PDE model. This overshoot can be reduced at the cost of a slower response time. Figure 6 shows the response of the system to kinetic parameter disturbances and parameter adaptations in PDE-based controller are shown in Figure 7. In this simulation time-scaled DMC exhibits oscillations. The linear model use in this controller is showing its limits to the successive deviations from the nominal operating point. The early lumping controller induces a large deviation from the set point as for the set-point variation, but the transient is still smooth. The late lumping controller provides a fast response to kinetic parameter variation.

Conclusion

A comparison between three levels of modeling for control have been presented: time-scaled DMC, a early lumping approach based on global differentiation of partial derivatives and a late lumping approach based on Lyapunov second method with feedforward action. The simulation results show the improvement by using a PDE model for tubular reactors. This improvement is particularly important when the process moves away from its nominal operating point where the nonlinearities in the kinetics cannot be followed adequately by a simple linear model.

References

Barrette, M. and M. Perrier, Modelling and Control of a Hypochlorite Bleaching Stage for Dissolving Pulp, In *Process Control 1995*, pages 11–22 (1995).

Bourrel, S., *Estimation et commande d'un procédé paramtres répartis utilisé pour le traitement biologique de l'eau à potabiliser*, Ph.D. Thesis, Université Paul Sabatier (1996).

Christofides, P. D., *Nonlinear Control of Two-Time-Scale and Distributed Parameter Systems*, Ph.D. Thesis, University of Minnesota (1996).

Dochain, D., *Contribution to the Analysis and Control of Distributed Parameter Systems with Application to (Bio)Chemical Processes and Robotics*, Agrégation Thesis, Université catholique de Louvain (1994).

Hong, K. S. and J. Bentsman, "Direct Adaptive Control of Parabolic Systems: Algorithm Synthesis and Convergence and Stability Analysis," *IEEE Trans. Auto. Cont.*, **39**(10), 2018–2033 (1994).

Pudlas, M., P. Tessier, and M. Savoie, Implementation of a Bleach Plant Supervisory Control System at Northwood Inc., In *Paptac Western Conference* (1999).

Ray, W. H., *Advanced Process Control*. McGraw-Hill, New York (1981).

Renou, S., M. Perrier, D. Dochain, and S. Gendron, Direct Adaptive Control of a Linear Parabolic System, In *IFAC SYSID 2000*, pages ThMD1–1, Santa-Barbara (2000a).

Renou, S., M. Perrier, D. Dochain, and S. Gendron, Nonlinear Control Design for Pulp Bleaching, In *Control System 2000*, pages 351–354, Victoria (2000b).

Renou, S., M. Perrier, D. Dochain, and S. Gendron, "Solution of the Convection-Dispersion-Reaction Equation by a Sequencing Method," *Submit to Comp. & Chem. Eng.* (2000c).

Savoie, M. and P. Tessier, "A Mathematical Model for Chlorine Dioxide Delignification," *Tappi J.*, **80**(6), 145–152 (1997).

Wang, R. X., J.-C. Tessier, and C. P. J. Bennington, "Modeling and Dynamic Simulation of a Bleach Plant," *AIChE J.*, **41**(12), 2603–2613 (1995).

Steady State Multiplicity and Stability in a Reactive Flash

Iván E. Rodríguez, Alex Zheng and Michael F. Malone*
Department of Chemical Engineering
University of Massachusetts
Amherst, MA 01003

Abstract

In this paper, we study the steady state solutions and their stability for an isobaric, adiabatic reactive flash. We show that the presence of vapor-liquid equilibrium can remove or create steady state multiplicity. For example, a system that does not have multiple steady states in a one phase CSTR can display multiple steady states as a two phase reactive flash unit. Both input multiplicity and output multiplicity are possible. Some of the multiplicities can be observed within the reactive flash operation. Others are observed when combining the operation of the reactive flash with the one phase CSTR. The existence of multiple steady states and their stability properties are related to a dimensionless quantity that depends on the heats of reaction and vaporization as well as the vapor-liquid equilibrium behavior. Three typical examples are shown.

The effect of reflux is also studied by adding a condenser for an exothermic reaction. This system can display multiple steady states with respect to the reflux ratio. At a sufficiently large reflux ratio, the system eventually ceases to display steady state multiplicity.

Keywords

Reactive distillation, Stability, Multiple steady states

Introduction

Reactive distillation (RD) systems combine chemical reaction and distillation, which are traditionally done separately. Hybrid combinations also include catalytic distillation, reactive extraction, phase-transfer catalysis, etc. By combining different unit operations into one consolidating system (whenever feasible), one could take advantage of the properties of one operation to enhance the others and sometimes improve the profitability of the overall process. As stated by Doherty and Malone (2001, page 427): "Reactive or catalytic distillation has captured the imagination of many recently because of the demonstrated potential for capital productivity improvements (from enhanced overall rates, by overcoming very low reaction equilibrium constants, and by avoiding or eliminating difficult separations), selectivity improvements (which reduce excess raw materials use and byproduct formation), reduced energy use, and the reduction or elimination of solvents."

Dynamic modeling of RD columns has received attention recently. Early studies were based on models that included the material balances and simultaneous phase and reaction equilibrium (e.g., Grosser et al. (1987)). More recent work has involved more complex models involving material and energy balances and kinetically controlled reactions (e.g., Alejski and Duprat (1996); Kumar and Daoutidis (1999)). Despite significant progress on design and dynamic modeling of reactive distillation systems, more general results are needed concerning the steady state and dynamic behavior of RD systems. Steady state multiplicity has been demonstrated for specific RD systems (e.g., Ciric and Miao (1994); Kumar and Daoutidis (1999); Hauan et al. (1997)). Some of the causes of steady state multiplicity have been established for the case of simultaneous reaction and phase equilibrium (Güttinger and Morari (1999)), but this is not the case for kinetically controlled RD systems. This paper establishes some of the conditions under which the steady state multiplicity is possible in a simple model system. The current work focuses on a simple RD system, the isobaric, adiabatic flash with chemical reaction for a binary mixture. We assume mass-transfer rates between the liquid and the vapor to be fast when compared to other characteristic times in the system; therefore the vapor composition is in equilibrium with the liquid composition at all times.

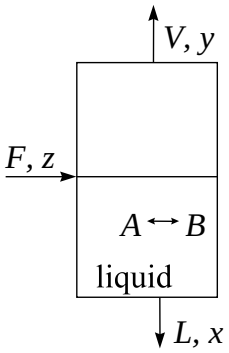


Figure 1: Diagram of a reactive flash.

The Reactive Flash

Figure 1 shows a graphical representation of the reactive flash system. An exothermic isomerization reaction

*To whom correspondence should be addressed. Address: 154A Goessmann Laboratory, 686 North Pleasant Street, Amherst, MA 01003-9303. Telephone: (413) 545-0838. Fax: (413) 545-1133. E-mail: mmalone@ecs.umass.edu

occurs in the liquid phase ($A \rightleftharpoons B$) only. The reactant is the heavier component. Both phases are assumed to be well mixed. Including the energy balance along with the overall mass balance and the species balance (for the reactant) yields:

$$\frac{dH}{dt} = F - V - L \tag{1}$$

$$\tau \frac{dx}{dt} = z - x - \phi(y - x) - \tau r \tag{2}$$

$$\tau C_p \frac{dT}{dt} = \tau r(-\Delta H_{rxn}) - C_p(T - T_i) - \phi(\Delta H_{vap}) \tag{3}$$

In these equations, H is the molar liquid holdup, F is the molar feed flowrate, L is the liquid molar flowrate out of the reactor, V is the vapor molar flowrate out of the reactor, t is the time, and x and y are the mole fractions of reactant A in the liquid and the vapor, respectively. T_i and z are the feed's temperature and mole fraction of A, respectively. Meanwhile, T is the temperature of the reactive flash and C_p is the liquid heat capacity (which has been assumed to be constant). The reaction rate is given by r which has units of $(time)^{-1}$. The fraction of the feed that leaves as vapor is defined as ϕ $(= V/F)$. The residence time is given by τ $(= H/F)$.

Steady State Behavior

If one considers the mass-transfer within the phases to be relatively fast, then one can relate the compositions of the phases by a vapor-liquid equilibrium relationship. Since the system is isobaric, knowing either T, x or y allows one to determine the other two quantities. One could use one of the balances, Equation 2 or Equation 3, to calculate T and the other one to determine the value of ϕ. For example,

$$0 = \frac{Da \cdot r}{k_{ref}}(-\Delta H_{rxn}) - C_p(T - T_i) - \phi(\Delta H_{vap}) \tag{4}$$

$$\phi = \frac{z - x + \frac{C_p(T_i - T)}{-\Delta H_{rxn}}}{\frac{\Delta H_{vap}}{-\Delta H_{rxn}} + y - x} \tag{5}$$

$$\Lambda \equiv \frac{\Delta H_{vap}}{-\Delta H_{rxn}} + y - x \tag{6}$$

In Equation 4, the Damköhler number, Da (Damköhler, 1939) is defined as the ratio of the residence time to a characteristic reaction time, $Da = \frac{\tau}{1/k_{ref}}$. The characteristic reaction time is given by the reciprocal of the forward reaction rate constant evaluated at a reference temperature, k_{ref}. From Equation 4, we have

$$\frac{dT}{dDa} = \frac{-\Delta H_{rxn} \cdot r/k_{ref}}{C_p + \frac{d(\phi \Delta H_{vap})}{dT} - Da(-\Delta H_{rxn}/k_{ref})\frac{dr}{dT}} \tag{7}$$

The steady state characteristics of this equation can be obtained by studying its denominator (since its numerator is always positive). After some manipulations, the denominator of this equation becomes

$$\Big[(y - x)C_p - \frac{dr}{dT}Da(-\Delta H_{rxn}\Lambda)/k_{ref} + \frac{d\Delta H_{vap}}{dT}(y - x)\phi - (\phi\frac{dy}{dT} + (1 - \phi)\frac{dx}{dT})\Delta H_{vap}\Big]\, 1/\Lambda \tag{8}$$

Some isomerizations are exothermic with large heats of reaction (see Frenkel et al. (1993) for examples). One example of such a reaction is the isomerization of quadricyclane to norbornadiene. Its heat of reaction has been measured to be around $-\Delta H_{rxn} = 89,000$ J/mol while the heat of vaporization of the two compounds lie in the range of $34,830$ to $37,850$ J/mol (An and Xie (1993)). In the examples that follow, the heat released by the reaction is used to heat the system and vaporize part of the reacting mixture. The heat of reaction $(-\Delta H_{rxn})$ is assumed to be greater than the heat of vaporization of the mixture (ΔH_{vap}). Also, since the mole fractions correspond to those of the heavier component the value of Λ can be either positive or negative depending on the particular properties of a given system.

Case 1: $\Lambda > 0$

The mole fractions in Expression 8 are those of the heavier component, therefore $y - x < 0$ (for systems without an azeotrope) and the first term in expression (8) is negative. Since $\frac{dx}{dT} > 0$, $\frac{dy}{dT} > 0$, and $0 \leq \phi \leq 1$, the last term is negative. Since the reaction rate generally increases with the temperature (i.e., $\frac{dr}{dT} > 0$) and $\Lambda > 0$, the second term is negative. The third term involves $\frac{d\Delta H_{vap}}{dT}$. If the value of the heat of vaporization is approximately constant for the temperature range of interest, then $\frac{d\Delta H_{vap}}{dT} \approx 0$. Therefore, if $\Lambda > 0$, then expression (8) is negative. Since the sign of Equation 7 depends on the sign of its denominator, then $\frac{dT}{dDa} < 0$ and there is only one steady state for each Da within the two phase region. It can also be shown that this steady state is stable (Rodríguez et al., 2001).

Figure 2 shows the temperature for a case with $\Lambda > 0$ (parameters are given in Table 1). Here, a case with a small relative volatility and a small ratio $\frac{\Delta H_{vap}}{-\Delta H_{rxn}}$ is depicted. The two phase region starts at $\phi = 0$ and ends at $\phi = 1$, after which steady state solutions are infeasible with the current setup. The small relative volatility ensures that $\Lambda > 0$ and a single steady state exists within the two phase region.

The one phase adiabatic CSTR steady states are also shown for comparison. All of those steady states (CSTR) are attainable at high pressures such that the system does not boil. If one decreases the pressure and allows the system to reach a boiling point, then one starts to move along the two phase region trajectory. Therefore, the CSTR states that lie above the two phase region are

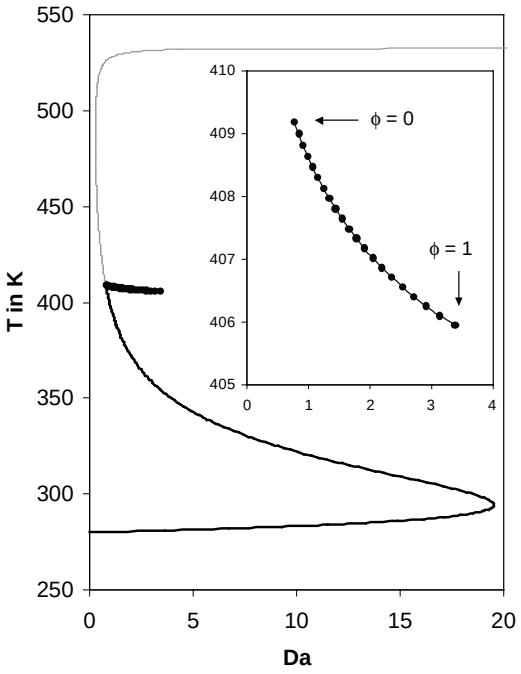


Figure 2: Temperature as a function of the Damköhler number for a CSTR and a reactive flash (also as inset). A system that displays multiple steady states as a CSTR only displays a single steady state as a reactive flash.

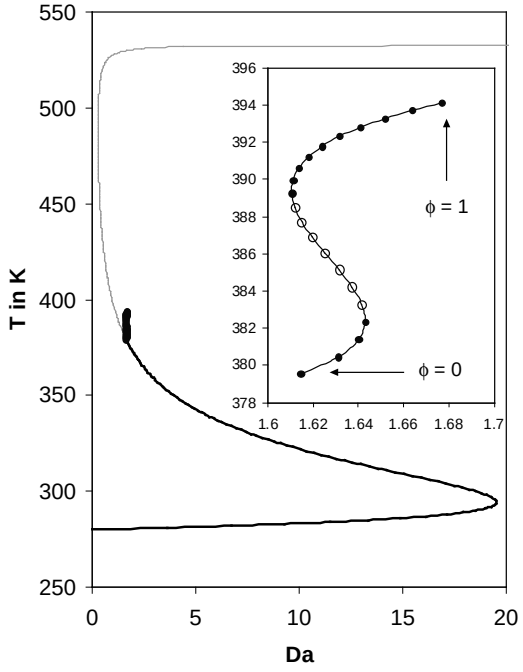


Figure 3: The reactive flash (also as inset) displays output multiplicity within the two phase region as well as with the CSTR operation.

Quantity	*Description*	*Value*
k_0	Pre-exponential factor	1000 (1/s)
Ea/R	Activation energy/ universal gas constant	6362 (K)
Keq	Equilibrium constant	$1.4843\exp[117.7(1/T - 1/298)]$
r	Reaction rate law	$k[x - \frac{1-x}{Keq}]$ (1/s)
C_p	Liquid heat capacity	252 (J/mol K)
T_i	Feed Temperature	280 (K)
z	Mole fraction of A in the feed	1
Figure 2		
α	Relative volatility	2
ΔH_{vap}	Heat of vaporization	36,600 (J/mol)
$-\Delta H_{rxn}$	Heat of reaction	115,000 (J/mol)
Figure 3		
α	Relative volatility	12
Figure 4		
ΔH_{vap}	Heat of vaporization	9,900 (J/mol)
$-\Delta H_{rxn}$	Heat of reaction	25,000 (J/mol)
Figures 5 and 6		
ΔH_{vap}	Heat of vaporization	36,622 (J/mol)
$-\Delta H_{rxn}$	Heat of reaction	155,000 (J/mol)

Table 1: Values used for the examples.

no longer accessible at this lower pressure (light colored line).

Comparing the one phase and two phase regions simultaneously one can see that there is the possibility of input multiplicity for this system. The behavior of the system in those two regions is significantly different, as is the required control policy for each case. If one is measuring temperature to control the system then it is important to be aware of this characteristic of the system as it affects the required control strategy.

Case 2: $\Lambda < 0$

If $\Lambda < 0$, the second term inside the brackets in expression (8) changes sign and becomes positive. This means that expression (8) and consequently $\frac{dT}{dDa}$ could change sign and result in multiple steady states. If the slope, $\frac{dT}{dDa}$, is positive the system is stable and unstable otherwise (Rodríguez et al., 2001).

Figure 3 shows the results of such a case. In this situation, a system with a higher relative volatility is depicted. The value of Λ is negative and one can observe the appearance of multiple steady states within the two phase region. The lower and upper branches of steady states are stable and the middle one is unstable.

When $\Lambda < 0$, the temperature increases as the system goes from $\phi = 0$ to $\phi = 1$. In doing so, there is a multi-directional enhancement in the reaction rate for

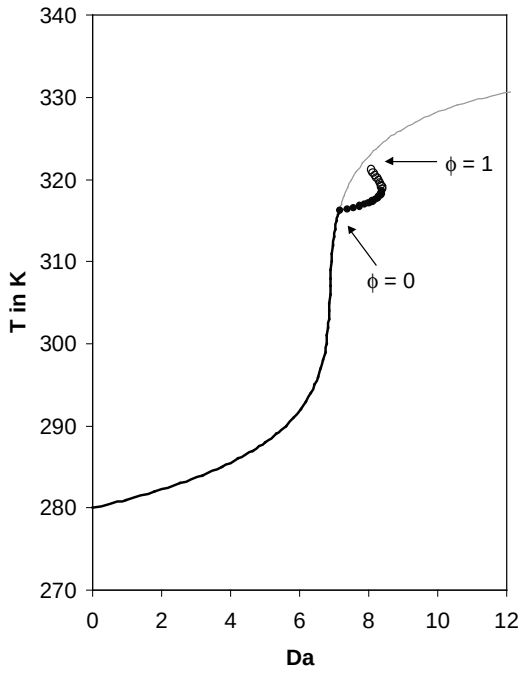


Figure 4: The presence of a separation operation allows the system to display output multiplicity as a reactive flash.

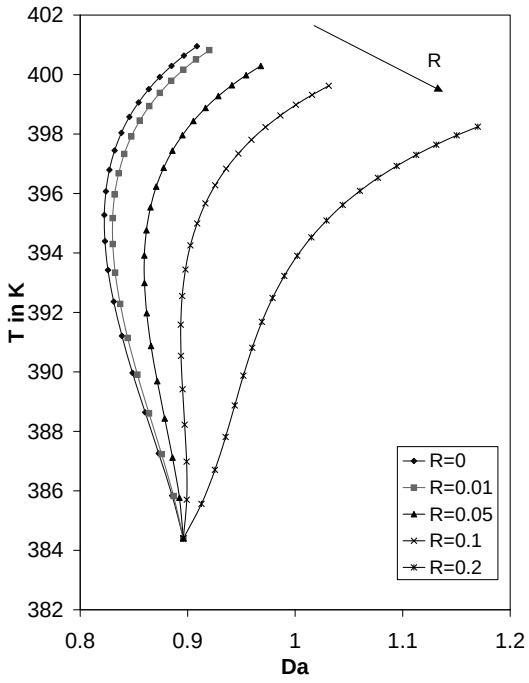


Figure 5: As the reflux ratio R increases, the region of multiplicity decreases and eventually only one steady state exists for the reactive flash with condenser.

the system, as both T and x increase simultaneously. This increase in reaction rate can promote the appearance of steady state multiplicity in the system. Therefore the fact that one is carrying out a separation and reaction simultaneously can lead to the appearance of multiple steady states even for some cases in which a single steady state is observed for every Da for the one phase CSTR.

Figure 4 presents the results for a system that does not display multiple steady states as a one phase CSTR but does as a reactive flash. One can observe the locus of steady states for the one phase CSTR in the figure and the fact that there is a single steady state for each value of Da. On the other hand, as the pressure is decreased and the system is allowed to reach a boiling point one can notice the appearance of two steady states for different values of Da within the two phase region. As before the two phase region starts at $\phi = 0$ and ends at $\phi = 1$.

Flash and Condenser

The influence of heat removal on the behavior of the system can be studied with the addition of a condenser to the vapor stream leaving the reactive flash. A reboiler or heating the flash directly would be necessary if the heat of reaction was not large enough to boil the reacting mixture (since enough heat is provided by the heat of reaction, external heating is not included in this analysis). The total condenser provides saturated liquid product and reflux but no chemical reaction. The exiting saturated liquid from this unit is partly removed as distillate and the rest returned as reflux to the reactive flash.

One could use this heat removal unit to possibly remove the steady state multiplicity observed in the cases when $\Lambda < 0$. The following figures show the results for a system in which the value of Λ is negative at zero reflux. When the reflux is equal to zero, this system corresponds to a reactive flash with $\Lambda < 0$, similar to the ones presented earlier.

In Figure 5, adding a condenser to the system can reduce the region of multiplicity until one is left with a single steady state for each value of Da. The addition of this condenser has more than one effect on the system. It removes energy released by reaction and lowers the temperature of the flash by introducing a cooler stream. For the reactive flash examples discussed previously, multiplicity exists as the separation in the system improves and the reaction is enhanced by the increase in temperature as well as reactant concentration. By returning this colder stream which has a larger concentration of product than the reacting mixture does, one reduces the enhancing effect that the separation has on the reaction (and the heat generated by it). Ultimately this causes the reduction and eventual elimination of the multiple

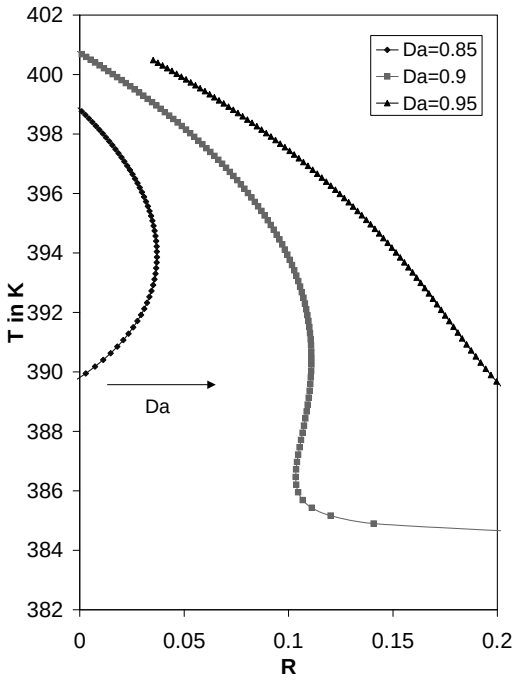


Figure 6: There is output multiplicity for some values of the reflux ratio R.

steady states for the reactive flash/condenser system.

Figure 6 depicts the situation for which we already have a fixed Da and would like to affect the purity or conversion by adjusting the reflux ratio, R. Once again, at high R there is a single steady state. At lower values of the Damköhler number there are two and sometimes three steady states possible for certain R values. At $R = 0$ the model reduces to that of the reactive flash where there were only two steady states for the particular Da chosen ($Da = 0.85$), for example. As Da increases the region of output multiplicity decreases and eventually only a single steady state is possible for each value of R.

Conclusion

The presence of vapor-liquid separation can remove or promote the appearance of multiple steady states in a reactive flash operated adiabatically and isobarically. Under some reasonable assumptions, the quantity Λ, relating the heat of vaporization, heat of reaction and the compositions in the system, is shown to be a key quantity to determine when steady state multiplicity is possible for the system. A negative value of Λ allows the system to operate in a region where the rate of reaction is enhanced by an increase in temperature as well as reactant composition in the liquid phase. In considering both the CSTR one phase steady states and the reactive flash, input and output multiplicities can be observed in some systems.

The use of a heat transfer unit for removing heat can reduce the region of instability and eventually eliminate steady state multiplicity. This could be accomplished by adding a total condenser to the vapor stream effluent from the reactive flash. As the reflux increases, the steady state multiplicities eventually disappear.

Acknowledgments

We gratefully acknowledge financial support from the sponsors of the Process Design and Control Center, University of Massachusetts-Amherst and from the donors of the Petroleum Research Fund, administered by the American Chemical Society (grants ACS-PRF #35612-AC9 and ACS-PRF #32221-G9).

References

Alejski, K. and F. Duprat, "Dynamic simulation of the multicomponent reactive distillation," *Chem. Eng. Sci.*, **51**, 4237 (1996).

An, X. W. and Y. D. Xie, "Enthalpy of isomerization of quadricyclane to norbornadiene," *Thermochimica Acta*, **220**, 17 (1993).

Ciric, A. R. and P. Miao, "Steady state multiplicities in an ethylene glycol reactive distillation column," *Ind. Eng. Chem. Res.*, **33**, 2738 (1994).

Damköhler, G., "Stromungs und warmeubergangsprobleme in chemischer technik und forschung," *Die Chemische Fabrik*, **12**, 469 (1939).

Doherty, M. F. and M. F. Malone, *Conceptual Design of Distillation Systems*. McGraw-Hill, New York (2001).

Frenkel, M. L., G. Y. Kabo, and G. N. Roganov, *Thermodynamic Properties of Isomerization Reactions*. Hemisphere Publishing Corporation, Washington, DC (1993).

Grosser, J. H., M. F. Doherty, and M. F. Malone, "Modeling of reactive distillation systems," *Ind. Eng. Chem. Res.*, **26**, 983 (1987).

Güttinger, T. E. and M. Morari, "Predicting multiple steady states in equilibrium reactive distillation. 1. Analysis of nonhybrid systems," *Ind. Eng. Chem. Res.*, **38**, 1633 (1999).

Hauan, S., T. Hertzberg, and K. M. Lien, "Multiplicity in reactive distillation of MTBE," *Comput. Chem. Eng.*, **21**, 1117 (1997).

Kumar, A. and P. Daoutidis, "Modeling, analysis and control of ethylene glycol reactive distillation column," *AIChE J.*, **45**, 51 (1999).

Rodríguez, I. E., A. Zheng, and M. F. Malone, "The stability of a reactive flash," *Chem. Eng. Sci.*, **56**, 4737 (2001).

Feasible Real-time Nonlinear Model Predictive Control

Matthew J. Tenny* and James B. Rawlings†
Department of Chemical Engineering
University of Wisconsin-Madison
Madison, WI 53706

Rahul Bindlish‡
Dow Chemical Company
Freeport, TX 77541

Abstract

This paper discusses an algorithm for efficiently calculating the control moves for constrained nonlinear model predictive control. The approach focuses on real-time optimization strategies that maintain feasibility with respect to the model and constraints at each iteration, yielding a stable technique suitable for suboptimal model predictive control of nonlinear process. We present a simulation to illustrate the performance of our method.

Keywords

Nonlinear processes, Model predictive control, Suboptimal control, Optimization strategies, Feasiblity

Introduction

Model predictive control (MPC) has become a well-established tool for advanced control applications, now one of the most prominent industrial control strategies. For many applications, linear MPC has proven to be a sufficient control tool. However, many processes exist that possess such a high degree of nonlinearity that linearized MPC produces inadequate results. For these processes, nonlinear model predictive control (NMPC) has been proposed. Some of the major challenges of NMPC are the solution of a global optimization problem in real-time and the appropriate handling of process constraints. In this paper, we present a method for real-time application of constrained NMPC with the possibility of running suboptimally, if required by a short sampling time.

Recent Advances in NMPC

In NMPC, stability is guaranteed for the finite horizon problem by applying a terminal constraint. Originally, the terminal state was required to be at the origin (Keerthi and Gilbert, 1988; Mayne and Michalska, 1990). The historical trend in the research that followed had the common theme of relaxing the constraints for an easier optimization formulation. Termination in a neighborhood around the origin with a terminal penalty is becoming a more popular formulation (Michalska and Mayne, 1993; Parisini and Zoppoli, 1995; Nicolao et al., 1998; Chen and Allgöwer, 1998a).

Less stringent criteria for stability were developed by Scokaert et al. (1999) and Chen and Allgöwer (1998b), who require only a decrease in the cost function at every time for stability, resulting in a suboptimal control law. The alternative formulation is that any controller that yields a closed-loop trajectory that ends in the terminal region is stable.

After the advances in NMPC theory, the next challenges came from the desire to calculate the control moves on-line in real-time. Significant effort has been made in reducing the computational burden of integrating the model over time; Bock et al. (1999) have investigated the use of simultaneous direct multiple shooting methods in which the solution to an ODE is determined at all points simultaneously. This method has been demonstrated to be faster than the traditional method of integration in series (Nagy et al., 2000).

Significant research also has been performed on the optimization approach utilized by the regulator. Some investigators have focused on global optimization, specifically genetic algorithms (Staus et al., 1996; Onnen et al., 1997; Rauch and Herremoës, 1999). However, this approach tends to be slow, and thus not implementable in real-time unless the time constants for the process are large.

Others have chosen to increase the speed of local optimization methods by tailoring them to take advantage of the specific structure of the MPC formulation. The approach uses an interior point method to solve a sequential quadratic programming problem (SQP) that, due to causal structure of the model, has a banded or almost block diagonal structure (Rao et al., 1998; Albuquerque et al., 1999). This method has been successful in easing the computational burden of optimizing large systems, and has recently come into favor with researchers.

This article describes the desired qualities of a real-time NMPC algorithm that can be run suboptimally, if required. We then discuss the method for achieving the desired objectives. Finally, we demonstrate our method on a nonlinear example.

Algorithm Requirements

We employ a typical ordinary differential equation model throughout this discussion. The process model has the form

$$\frac{dx}{dt} = f(x, u) \tag{1}$$

that, when integrated, becomes the discrete model

$$x_{i+1} = F(x_i, u_i) = x_i + \int_{t_i}^{t_{i+1}} f(x(\tau), u_i) d\tau \tag{2}$$

*tenny@bevo.che.wisc.edu
†jbraw@bevo.che.wisc.edu, author to whom correspondence should be addressed
‡RBindlish@dow.com

in which i represents the sampling time and a zero-order hold is assumed for the inputs.

One of the strengths of model predictive control is its ability to handle constraints in the regulator. The finite horizon regulator problem we consider is

$$\begin{aligned} \min_{u_i} \quad & \Phi(u_i, x_0) = \sum_{i=0}^{N} L(x_i, u_i) + h(x_N) \\ \text{subject to: } & x_N \in W_\alpha \\ & x_{i+1} = F(x_i, u_i) \\ & x \in X, u \in U \end{aligned} \tag{3}$$

Here, a terminal penalty is added to approximate the cost of the infinite horizon problem from the end of the trajectory in the terminal region. The approximation is valid because the terminal region is constructed in such a way that the nonlinear plant is not significantly different from the linear approximation in that neighborhood of the steady-state target. The problem is also constrained in the states and inputs. We assume the existence of a sequence of control moves that steers the state to setpoint asymptotically.

It is not necessary, however, to find the global optimum of Equation 3. Instead, asymptotic stability is guaranteed provided that the trajectory of states terminates in the region W_α (Scokaert et al., 1999). Therefore, the horizon length N need only be long enough to satisfy this restriction and the cost function does not need to be minimized for the control technique to reach the set-point. This concept eases the computation since a *suboptimal* set of control moves provides a stable controller. Therefore, the optimizer can stop early if the sampling time is small, and the process can still be controlled.

Algorithm Description

In the last section, we developed the requirements for regulating a nonlinear system with finite horizon MPC. Now, we must address the issue of solving the optimization problem according to our requirements. For this optimization, a sequential quadratic programming (SQP) technique is used. We exploit the structure of the problem to speed up the computation to run in real-time.

First, we present the fundamentals of the SQP method. Suppose we wish to solve the following problem:

$$\begin{aligned} \min_{w} \quad & h(w) \\ \text{subject to: } & c(w) = 0 \\ & d(w) \geq 0 \end{aligned} \tag{4}$$

The method reduces the problem in Equation 4 to a series of quadratic programs. Quadratic programs are well-studied and quickly solved with available methods, making SQP methods a suitable choice for nonlinear problems.

We present two methods for performing the quadratic programming approximation; the first does not have quadratic convergence properties, but may be easier to set up, depending on the structure of $h(w)$. The second has quadratic convergence properties for points near the solution, but the set-up may be computationally intensive.

Straightforward Formulation. First, define the superscript j to represent an iteration of the SQP method. In this formulation, we approximate $h(w)$ as a quadratic function around the current iterate w^j. We then compute a linear approximation to the constraints $c(w)$ and $d(w)$ around w^j. Defining $p = w - w^j$, we solve

$$\begin{aligned} \min_{p} \quad & \frac{1}{2} p^T \nabla^2 h(w^j) p + \nabla h(w^j)^T p \\ \text{subject to: } & \nabla c(w^j)^T p + c(w^j) = 0 \\ & \nabla d(w^j)^T p + d(w^j) \geq 0 \end{aligned} \tag{5}$$

No knowledge of the Lagrange multipliers by the user is required to form Equation 5. However, no special local convergence properties exist for this problem. It converges, but not quadratically near the solution. Quadratic convergence is important only when the initial guess is good enough that the minimizer is known to be close to the open loop prediction.

Quadratic Convergence Formulation. In order to obtain quadratic local convergence properties, the Lagrangian is first defined:

$$\mathcal{L}(w, \lambda) = h(w) - \lambda^T \begin{bmatrix} c(w) & d(w) \end{bmatrix} \tag{6}$$

The Hessian of the Lagrangian with respect to the variables is denoted by

$$\mathcal{H}(w, \lambda) = \nabla^2_{ww} \mathcal{L}(w, \lambda) \tag{7}$$

By the first-order KKT conditions, Equation 5 is equivalent to

$$\begin{aligned} \min_{w} \quad & \frac{1}{2} p^T \mathcal{H}(w^j, \lambda^j) p + \nabla h(w^j)^T p \\ \text{subject to: } & \nabla c(w^j)^T p + c(w^j) = 0 \\ & \nabla d(w^j)^T p + d(w^j) \geq 0 \end{aligned} \tag{8}$$

The solution to this problem is equal to a step of Newton's method (Nocedal and Wright, 1999), which is quadratically convergent in the region near the solution. This method requires an estimate of the Lagrange multipliers, which may not be estimated well until a few iterations have been performed. Also, a second order approximation of the constraints are required, which may be computationally intensive. We now relate the method of solution to the nonlinear MPC problem structure.

Dense Hessian SQP

In this naïve approach, we substitute Equation 2 into Equation 3 and solve it using an SQP method. We need first to define $w^j = u^j$. The equality constraints $c(w)$ are eliminated. The function that must be minimized, $h(w)$, is a complex nonlinear function. To form the quadratic approximation of $h(w)$, the value of $\nabla^2 h(w^j)$ is required. It can be seen that each x_i relies on all the $u_{i-k}, k = 1, 2, \ldots, i$. In fact, the last term relies on *all* of the u_i. The implications of this fact are that the Hessian of $h(w)$ is *dense*, making it difficult to take advantage of the specific architecture of the MPC problem in this formulation. To handle the MPC problem more efficiently, we propose a different approach.

Banded Hessian SQP

The banded Hessian SQP approach for MPC is described by Rao, Wright, and Rawlings (Rao et al., 1998) for the case of linear MPC. However, we can naturally extend it to the nonlinear case with a few modifications.

The key difference in the banded Hessian approach is not to plug the model equation into the Equation 3. Instead, it is left as an explicit equality constraint. Note that the constraint matrix is highly structured due to the causality of *only* the past state and input on the current state. In the quadratically convergent method, we require the Hessian of $c(w)$ as well. The differentiation could be performed by finite differences, which is slow and innaccurate, or an approximate Hessian could be calculated. However, popular Hessian update strategies, such as the BFGS update, destroy the banded structure of the Hessian, yielding instead a dense matrix. Until a more sophisticated update strategy is employed, it may be more suitable to solve Equation 8 for real-time applications.

The terminal region may be calculated offline (Tenny, 2000). However, we exclude the terminal region constraint in the quadratic program. This is done for a number of reasons:

1. If the horizon length N is chosen to be too short, the problem is infeasible.
2. If the horizon length is nominally long enough to reach the terminal region, the closed-loop solution may differ appreciably from the open-loop prediction.
3. Ellipsoidal constraints in quadratic programs are not exact; the constraint would be approximated.

A banded QP solver is then used for the structured optimization. The method of solution is an interior point method (Rao et al., 1998) that has been geared for the MPC structure. The cost of this approach is linear with respect to horizon length N, compared to cubic growth for the dense Hessian approach.

A solution to the approximate problem is then calculated. Instead of solving the nonlinear problem, the solution to the quadratic program points in a direction of objective function decrease. The next iterate is found using a trust region constraint.

Feasibility

One of the desired properties of the optimization algorithm is feasibility at all times with respect to the constraints. State inequalities are handled as soft constraints; they may be violated, but a term is added to the cost function to penalize the violation of such a constraint. The integration of the model can be accomplished using multiple shooting methods (Bock et al., 1999). We now describe how we maintain feasibility with respect to the input constraints and the equality constraints at all times so that we may terminate the optimization at any point to run the regulator suboptimally.

To guarantee feasibility of the initial guess w^0, we generate each u_i^0 as follows:

- In the case of small or no disturbances, the result from the previous open-loop prediction is feasible and becomes the initial guess for the current regulator problem.
- For startup or large disturbances, u_i^0 is determined based on the feedback law $u_i^0 = Kx_i^0$ in which K is the linear quadratic regulator feedback gain of the system linearized about the origin. If $Mu_i^0 > m$, we simply take those elements of u_i^0 that violate the constraint and clip them such that they meet the constraint. We maintain an initial guess of a feedback law in case large unmeasured disturbances are present and previous iterates are no longer valid. Alternatively, a local constrained linear MPC problem can be solved and one can use its solution as an initial guess for the regulator problem.
- Each x_i^0 is generated by substituting u_{i-1}^0, x_{i-1}^0 into Equation 2.

We now have a w^0 that satisfies the input and equality constraints.

The linearized system of constraints is formed and a quadratic program is solved. The state variables $\bar{x}$ from the result $\bar{w}$, are discarded and replaced with the states resulting from injecting the inputs $\bar{u}$ into the nonlinear model. If the cost function increases, we refine the trust region or line search method and regenerate the state predictions via the nonlinear model. However, if the cost function decreases, the method yields a new iterate w^1 that is feasible with respect to both the input and equality constraints *and* has a lower cost function. We now repeat the process using w^1 as the new initial guess. The algorithm repeats this process until the next sampling time is reached (suboptimal MPC) or until the iterates converge ($\|w^j - w^{j-1}\| \leq \delta$).

Grade	Polymer production rate (kg/min)	MMA Mole fraction in copolymer	Copolymer viscosity ($10^{-6}m^3/kg$)
A	0.3	0.60	35000
B	0.25	0.75	36725

Table 1: Product grades for MMA-VA copolymerization.

Once the iterates converge, the resulting vector w^{final} is the minimizer of the cost function without a terminal region constraint. Now, the final state in the w^{final} vector is x_N^{final}. We check to see if $x_N^{\text{final}} \in W_\alpha$. If it is, then the algorithm terminates. If not, then the horizon length N is not long enough. Therefore, N is increased (Rao et al., 1998) and the initial guess is the solution w^{final} with new inputs generated by Kx_N^{final}, etc. The initial guess is still feasible, and it takes advantage of the work of the previous optimization. We ignore the case of degenerate terminal regions for the purposes of this discussion.

An Example Process

We investigate the process of MMA-VA copolymerization as described by Bindlish (1999), which consists of a well-mixed reactor followed by a product separator. Several grades of polymer, characterized by the mole fraction of monomer A (MMA) in the copolymer product and the intrinsic viscosity of the copolymer product, are manufactured using this process. Depending on the demands of the market, the desired copolymer viscosity and composition are varied during operation. The feed to the reactor consists of the monomers (MMA and VA), initiator (AIBN), transfer agent (acetaldehyde) and inhibitor (m-dinitrobenzene) dissolved in a solvent (benzene). To remove heat released by the polymerization reaction, a coolant is employed. The polymer product is then separated from the unreacted hydrocarbons in a downstream separator. The reactor and separator are represented by a physical model based on first principles. The model has 15 states, 3 inputs, and 7 output variables, of which only three have set-points. The sampling time for this process is 5 minutes and the prediction horizon is 20 time steps (100 minutes). We wish to switch from product grade A to grade B, with output set-points as shown in Table 1. Figure 1 shows the output variables during the simulation for both the proposed nonlinear controller and nominal linear model predictive control based on a linearization of the nonlinear model at Grade B. In the case of the linear model predictive controller, the system becomes unstable. We note that the nonlinear model predictive controller returns a local optimal solution in this scenario because the computation time is smaller than the sampling time for this example.

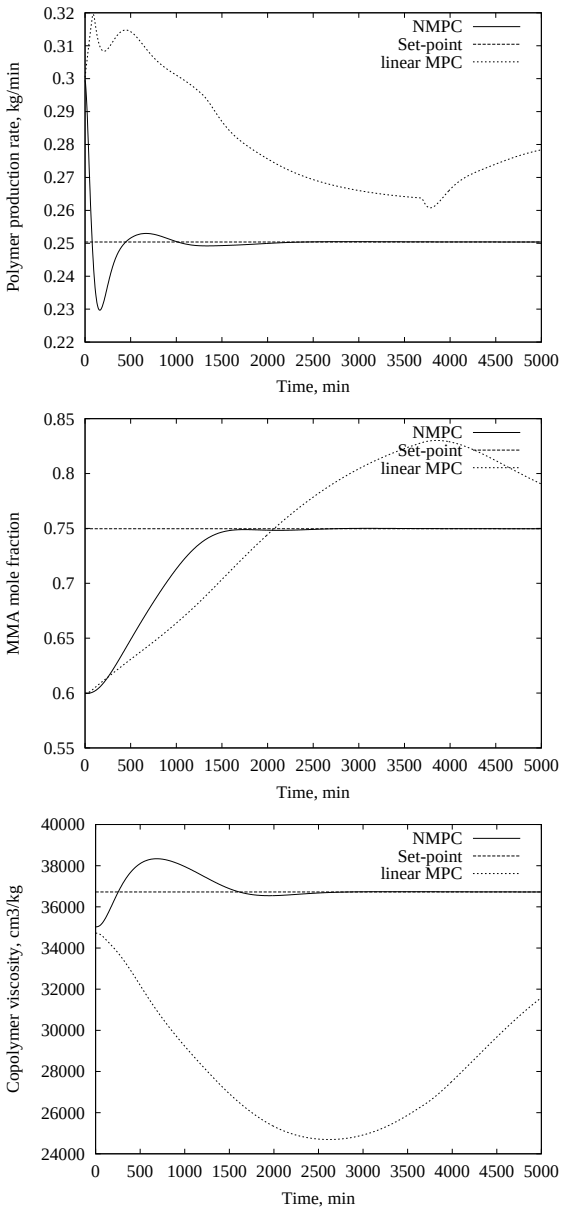


Figure 1: Output of transition from Grade A to B.

Conclusion

In this paper, we have devised and demonstrated a prototype algorithm for NMPC that can run suboptimally, if required. The method is capable of stabilizing nonlinear systems and obtaining optimal performance in real-time. Future directions for this research include real-time nonlinear state estimation.

References

Albuquerque, J., V. Gopal, G. Staus, L. T. Biegler, and E. B. Ydstie, "Interior Point SQP Strategies for Large-scale, Structured Process Optimization Problems," *Comput. Chem. Eng.*, **23**(4), 543–554 (1999).

Bindlish, R., *Modeling and Control of Polymerization Processes*,

PhD thesis, University of Wisconsin-Madison (1999).

Bock, H. G., M. Diehl, D. Leineweber, and J. Schlöder, Efficient Direct Multiple Shooting in Nonlinear Model Predictive Control, In Keil, F., W. Mackens, H. Voss, and J. Werther, editors, *Scientific Computing in Chemical Engineering II, Volume 2: Simulation, Image Processing, Optimization, and Control*. Springer, Berlin, Heidelberg (1999).

Chen, H. and F. Allgöwer, "A Computationally attractive nonlinear predictive control scheme with guaranteed stability for stable systems," *J. Proc. Cont.*, **8**(5–6), 475–485 (1998a).

Chen, H. and F. Allgöwer, Nonlinear Model Predictive Control Schemes with Guaranteed Stability, In Berber, R. and C. Kravaris, editors, *NATO ASI on Nonlinear Model Based Process Control*, pages 465–494. Kluwer (1998b).

Keerthi, S. S. and E. G. Gilbert, "Optimal Infinite-Horizon Feedback Laws for a General Class of Constrained Discrete-Time Systems: Stability and Moving-Horizon Approximations," *J. Optim. Theo. Appl.*, **57**(2), 265–293 (1988).

Mayne, D. Q. and H. Michalska, "Receding Horizon Control of Nonlinear Systems," *IEEE Trans. Auto. Cont.*, **35**(7), 814–824 (1990).

Michalska, H. and D. Q. Mayne, "Robust Receding Horizon Control of Constrained Nonlinear Systems," *IEEE Trans. Auto. Cont.*, **38**(11), 1623–1633 (1993).

Nagy, Z., R. Findeisen, M. Diehl, F. Allgöwer, H. G. Bock, S. Agachi, J. P. Schlöder, and D. Leineweber, Real-time Feasibility of Nonlinear Predictive Control for Large Scale Processes—a Case Study, *IST Report No. 2000-4*, Institut für Systemtheorie technischer Prozesse, Universität Stuttgart (2000).

Nicolao, G. D., L. Magni, and R. Scattolini, "Stabilizing Receding-Horizon Control of Nonlinear Time-Varying Systems," *IEEE Trans. Auto. Cont.*, **43**(7), 1030–1036 (1998).

Nocedal, J. and S. J. Wright, *Numerical Optimization*. Springer-Verlag, New York (1999).

Onnen, C., R. Babuska, U. Kaymak, J. M. Sousa, H. B. Verbruggen, and R. Isermann, "Genetic Algorithms for Optimization in Predictive Control," *Control Eng. Practice*, **5**(10), 1363–1372 (1997).

Parisini, T. and R. Zoppoli, "A receding-horizon regulator for nonlinear systems and a neural approximation," *Automatica*, **31**(10), 1443–1451 (1995).

Rao, C. V., S. J. Wright, and J. B. Rawlings, "On the application of interior point methods to model predictive control," *J. Optim. Theo. Appl.*, **99**, 723–757 (1998).

Rauch, W. and P. Herremoës, "Genetic Algorithms in Real Time Control Applied to Minimize Transient Pollution from Urban Wastewater Systems," *Water Research*, **33**(5), 1265–1277 (1999).

Scokaert, P. O. M., D. Q. Mayne, and J. B. Rawlings, "Suboptimal Model Predictive Control (Feasibility Implies Stability)," *IEEE Trans. Auto. Cont.*, **44**(3), 648–654 (1999).

Staus, G. H., L. T. Biegler, and B. E. Ydstie, Adaptive control via non-convex optimization, In Floudas, C. and P. Pardolas, editors, *Nonconvex optimization and its applications*, pages 1–19. Kluwer (1996).

Tenny, M. J., Practical Development in Nonlinear Model Predictive Control (2000). PhD Preliminary Report, University of Wisconsin-Madison.

Industrial Experience with State-Space Model Predictive Control

Ernest F. Vogel and James J. Downs
Eastman Chemical Company
Kingsport, Tennessee

Abstract

Experience with infinite-horizon state-space model predictive control confirms that the algorithm offers several advantages over the more conventional finite-horizon step-response based model predictive control algorithms, particularly in the specification of sample time and handling a wide range of process time constants. Examples illustrate our use of state space based model predictive control and its integration with conventional control techniques.

Keywords

Model predictive control, State space, Application, Constraint control, Multivariable control strategy

Introduction

The Advanced Controls Technology Group at Eastman Chemical Company has nine years of experience applying predictive control on industrial processes. The first five of those years, we applied our own variation of Dynamic Matrix Control (DMC). Our DMC is similar to the conventional finite-horizon, step-response-model based predictive control technology commonly applied in the chemical process industries (Cutler and Ramaker, 1980; Richalet et al., 1978). Four years ago, we began applying infinite-horizon state-space model predictive control, denoted here as MPC. We now have 35 installations of this technology. The state space formulation offers several advantages over the convolution model approach. As a result, all of our new applications use MPC. The DMC applications are still in service and, so far, we have not converted any of them to MPC. While the MPC algorithm offers several advantages, it still has some features which make it challenging to implement. In this paper, we briefly describe the strengths and weaknesses we have found in our experience with the MPC algorithm. We also discuss control strategy design with MPC and give two example problems.

MPC Implementation

The state-space predictive controller implemented at Eastman follows that documented by Muske and Rawlings (1993). Our implementation uses a fixed gain Kalman filter for the observer and a quadratic program formulation to determine the steady state targets. The regulator portion of the algorithm uses input parameterization as described by Muske (1995). This technique assumes that the inputs follow a $u = -Kx$ path from the end of the move horizon onward. The feasibility of output constraints is achieved by softening the constraints as presented by Ricker et al. (1988) and Zheng and Morari (1995). Although there can be significant performance limitations with this approach for non-minimum phase systems as discussed by Scokaert and Rawlings (1999), we have not observed the problems they discuss. The steady-state and dynamic optimization problems resulting from the state-space control algorithm are solved using a quadratic programming algorithm described by Powell (1985) and Schmid and Biegler (1994).

Strengths and Weaknesses of State-Space MPC

Process Dynamics

One of the strengths of state-space MPC is the relative independence of controller sample time and process time constants. Because the algorithm does not use a step-response model, the sample time can be very small relative to the process time constants. Also, the range of process time constants can be very wide. This capability is important when the MPC strategy includes variables that can respond quickly, such as distillation column differential pressure, and variables that can respond much more slowly, such as column product composition. The variables that respond quickly require a short sample time. The ability to have small sample times relative to process time constants makes state-space MPC applicable on a broader range of problems and provides more flexibility in its implementation, compared to step-response based algorithms. Additionally, the sample time is much easier to select with state-space MPC, because of the independence from the process time constants.

Dead time can still present a problem for state-space MPC. Because every sample time of dead time creates an additional state, small sample times relative to the dead time lead to a large number of states. While we have not encountered any implementation problems associated with the number of dead time states, our experience in this area is limited. So far, our largest number of dead time states for a single input/output pair has been 11.

Tuning

The state-space MPC algorithm has three major parts: the state observer, the steady-state target calculation, and the dynamic regulator. The calculations for each of these parts are performed each control interval. Each

of these parts must be tuned for the specific application. The final steady-state is affected only by the steady-state tuning. The dynamic performance of the controller is affected by the tuning of both the state observer and the regulator.

The separation of the steady-state performance and the dynamic performance within state-space MPC is a good feature. The steady-state weighting/tuning parameters have proven relatively easy to set to achieve the desired final steady-state. However, the dynamic weighting/tuning parameters in both the state observer and the regulator are more difficult to set to achieve the desired dynamic response. Further, the dynamic response can be adjusted by changing parameters in either the state observer or in the dynamic regulator. The best method for dynamic tuning is not fully clear.

State-space MPC offers the option of modeling disturbances on either the process input or output. Our algorithm includes both options. We have experimented with each option on some applications, but the best choice for that decision is seldom clear. We normally assume unmeasured disturbances enter at the process input unless process understanding dictates otherwise. Our applications have not required the identification of disturbance models other than the common step disturbance. In a proportional-integral controller sense, we view the regulator as providing the proportional part and the observer providing the integral mode. Unmeasured disturbances are picked up by the observer via the disturbance model and require reset action. Tuning of the disturbance model has become similar to tuning resets on conventional controllers.

State-space MPC has many tuning parameters. They require adjustment based on observed performance. For most problems, our experience has been that MPC tuning needs to be done off-line with a simulation.

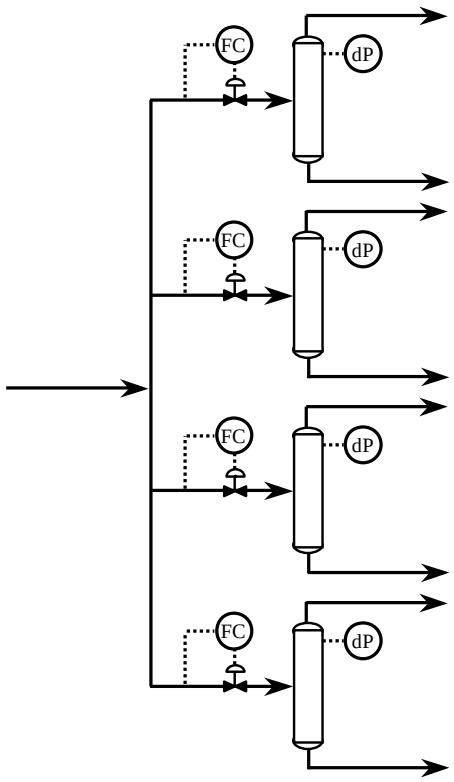


Figure 1: Distilation columns in parallel.

Control Strategy Design

While the previous section discussed aspects of the state-space MPC algorithm itself, we have found that the control strategy design with MPC has a far bigger impact on the success of a project than the performance of the MPC algorithm itself. By control strategy design, we mean definition of the control objectives, selection of control technology (MPC or traditional SISO structures), and the selection of controlled, manipulated, and constraint variables for MPC.

To illustrate our use of MPC, two examples are given below. A common control problem at Eastman is the distribution of load between parallel unit operations. The ability to handle non-square systems makes model predictive controllers an attractive control technology for this problem. In the first example, the capability to characterize the process control objective and to prioritize competing objectives is illustrated along with the capability to specify widely different closed-loop time constants within the same controller. In the second example, the capability to cascade multiple MPC layers and to integrate MPC with conventional control is shown to further enhance the strength of this technology.

Example 1—Distillation Columns in Parallel

Figure 1 shows four distillation columns in parallel service. The columns already have an excellent regulatory SISO strategy in place to control top and bottom compositions. The control problem is to distribute the total throughput subject to the hydraulic limits of the columns as indicated by their maximum differential pressures. Therefore, this problem has one controlled variable (total throughput), four manipulated variables (column feed rate set points), and four constraint variables (column differential pressures).

The control strategy objectives are summarized below in order of priority with the first item being the most important.

Objective	Speed
1. Satisfy manipulated variable constraints (hard constraints).	Instantaneous
2. Satisfy differential pressure constraints (soft constraints).	Fast
3. Satisfy total throughput target or maximize total throughput subject to constraints.	Slow
4. Distribute feed so all columns are equally loaded as measured by the distance from their differential pressure constraints.	Very slow

The specification of the control objectives is a major step in an advanced control application and typically goes through several iterations. Often, the best way to operate the equipment is discovered through experimentation with different objectives. In this case, the objectives are given above and the next step is to decide on the control technology and configuration required to meet the objectives. There are several possibilities including an exclusively SISO strategy, an exclusively MPC strategy, or a hybrid. The best choice depends not only on which technology will best meet the objectives, but also on several other factors: the size of the MPC problem (number of controlled and manipulated variables), maintenance of the algorithm, maintenance schedules for the physical units, reliability of the measurement and communication links, and understandablity.

For this particular application, we chose to use MPC, but with four additional controlled variables. In order to distribute the feed so all columns are an equal distance from their differential pressure constraints, we included for each column the difference between the high differential pressure limit and the current differential pressure. Those four controlled variables have a set point of zero. With a high steady-state weight in state-space MPC, we specify that the total throughput should be at set point. The four "constraint distance" controlled variables have a low steady-state weight. Since there are five controlled variables and four manipulated variables, all the set points cannot be reached. The total throughput goes to its set point and state-space MPC distributes the set point error equally for the four constraint distance variables.

The four column differential pressures are still needed as high constraint variables in MPC for speed of response reasons. The tuning is such that the distribution of the feed happens very slowly, but if something happens which drives a differential pressure to the constraint, the associated feed rate is moved quickly to compensate.

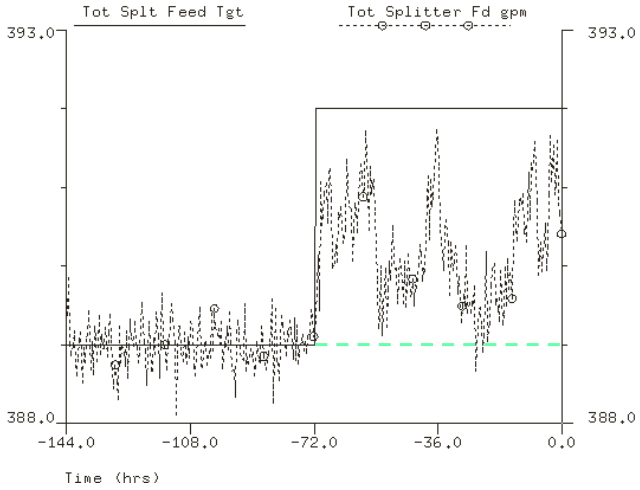


Figure 2: Total production rate for distillation columns in parallel.

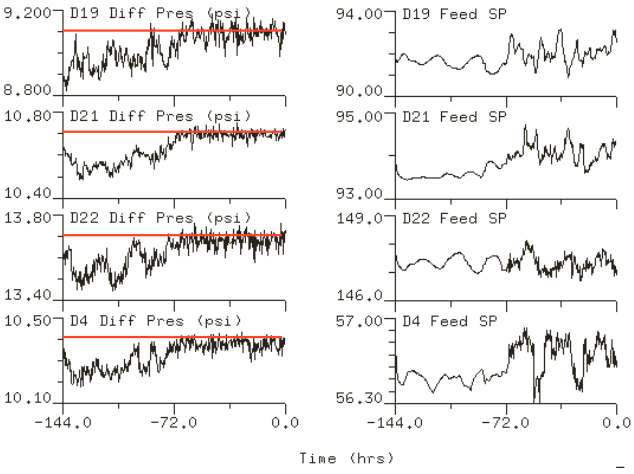


Figure 3: Differential pressures and feed rates for distiallation columns in parallel.

Figures 2 and 3 demonstrate the response for a production rate target increase beyond what is achievable. Figure 2 shows that prior to the change, the total rate is at target. After the target change, the total flow does not reach the target, but is maximized subject to the constraints. Figure 3 shows the column differential pressures all at approximately the same distance from the high limit prior to the target change. After the target change, the differential pressures are running at the high limits. Also, note that prior to the target change, the

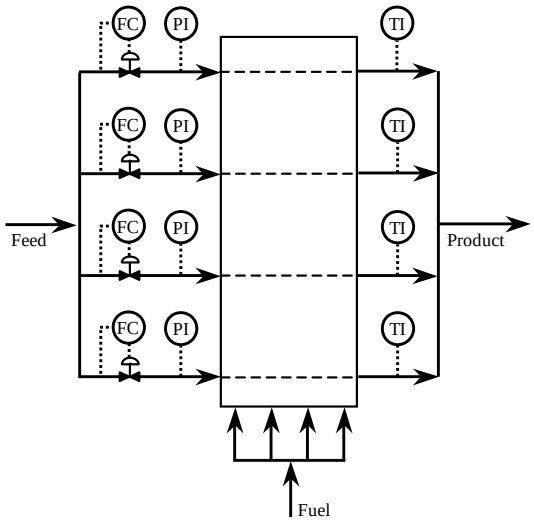


Figure 4: Cracking furnace.

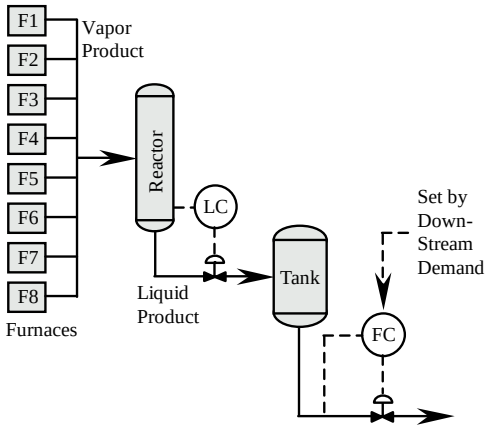


Figure 5: Cracking furnaces in parallel.

column feed rates move relatively slowly. However, the column feed rates move more aggressively after the target change when the columns are running on their differential pressure constraints.

Example 2—Cracking Furnaces in Parallel

A process gas is cracked in a natural gas fired furnace. Figure 4 shows a process diagram of a furnace. The process gas flows through four separate coils in the furnace. Each coil has its own inlet feed flow controller, inlet pressure measurement (downstream of the feed valve), and outlet temperature measurement. The combustion air flow is ratioed to the natural gas flow. There are 5 manipulated variables for controlling the furnace: each of the 4 inlet feed flow controller set points and the fuel flow controller set point.

The furnace production rate (sum of all 4 coil flows) is set by product demand (sometimes to be maximized). Coil outlet temperature is indicative of conversion and needs to be controlled on each coil. Coil inlet pressure can float but must not exceed a high limit because excessive coil inlet pressure causes a relief valve to open. The maximum fuel rate limit is determined by the furnace emission permit.

Changing a coil flow not only affects its own outlet temperature, but it also affects the outlet temperatures of the other 3 coils as well. As a result, most process changes or disturbances require adjustment to all 5 manipulated variables. Thus, it is a true multivariable control problem.

MPC is an excellent control technology for this problem because it is multivariable and because of the process constraints. For this application we used state-space MPC with 5 controlled variables (total rate and 4 outlet temperatures), 5 manipulated variables (fuel flow and 4 inlet feeds), and 4 constraint variables (4 coil inlet pressures).

In our application, there are 8 of these furnaces in parallel supplying a common user. Figure 5 shows a diagram of the system. The control objectives for this system are very similar to those for the distillation column system and are summarized below in order of priority with the first item being the most important.

Objective	Speed
1. Satisfy manipulated variable constraints (hard constraints).	Instantaneous
2. Satisfy coil inlet pressure constraints (soft constraints).	Fast
3. Control coil outlet temperatures at their target.	Medium
4. If a furnace becomes constrained or if a furnace is shut down or started up, redistribute load to maintain total rate.	Medium
5. Control the downstream inventory at target or maximize throughput subject to constraints.	Slow
6. Distribute feed so all furnaces are at equal total rate to maximize run length.	Very slow

Again, MPC is attractive for this load distribution application. However, it is a significantly bigger problem

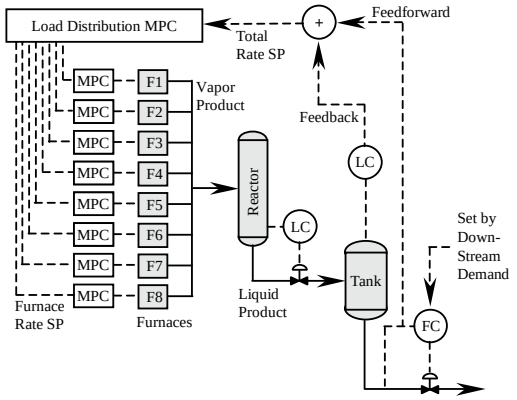


Figure 6: MPC control strategy for cracking furnaces in parallel.

than the distillation example above: 33 controlled variables (1 level, 32 temperatures), 40 manipulated variables (32 process feeds, 8 fuel rates), and 32 constraint variables (32 coil inlet pressures). Because of the size of the problem, several MPC control strategy options should be considered. One option is to do this problem with one big MPC. A second option is to control each furnace with an MPC and layer a master load distribution MPC on top of the 8 slave MPCs.

The solution we implemented is a variation of the latter choice and is shown in Figure 6. We chose this configuration for several reasons:

1. The top MPC layer can easily be turned off and furnace MPCs operated individually.
2. A PID controller was used for level controller because it was easier to tune and it met the need.
3. The smaller MPCs are less of a computational load.
4. The system could be built in steps. Successive furnace MPCs could be built and brought on-line without disturbing ones commissioned earlier.
5. The system is easier to maintain. Any code/configuration maintenance can be done while a furnace is down without disturbing running furnaces.

Except for very small problems, such as the distillation example, we normally choose to layer a master MPC on top of slave MPCs that manage parallel unit operations.

Conclusions

While state-space MPC offers several advantages over step response based algorithms, implementation still requires a great deal of expertise simply to use and tune the algorithm. The control strategy design when incorporating MPC has a far bigger impact on success of a project than the performance of the MPC algorithm itself. The examples illustrated the flexibility of state-space MPC to achieve a variety of control objectives, both in importance and speed of response. The state-space MPC technology has become an important tool for Eastman Chemical Company when solving control problems that have complex control objectives and widely different time frames.

References

Cutler, C. R. and B. L. Ramaker, Dynamic Matrix Control—A Computer Control Algorithm, In *Proceedings of the Joint Automatic Control Conference* (1980).

Muske, K. R. and J. B. Rawlings, "Model Predictive Control with Linear Models," *AIChE J.*, **39**(2), 262–287 (1993).

Muske, K. R., *Linear Model Predictive Control of Chemical Processes*, PhD thesis, The University of Texas at Austin (1995).

Powell, M. J. D., "On the quadratic programming algorithm of Goldfarb and Idnani," *Math. Prog. Study*, **25**, 46 (1985).

Richalet, J., A. Rault, J. L. Testud, and J. Papon, "Model Predictive Heuristic Control: Applications to Industrial Processes," *Automatica*, **14**, 413–428 (1978).

Ricker, N. L., T. Subrahmanian, and T. Sim, Case Studies of Model-Predictive Control in Pulp and Paper Production, In McAvoy, T. J., Y. Arkun, and E. Zafiriou, editors, *Proceedings of the 1988 IFAC Workshop on Model Based Process Control*, pages 13–22, Oxford. Pergamon Press (1988).

Schmid, C. and L. T. Biegler, "Quadratic programming methods for reduced hessian SQP," *Comput. Chem. Eng.*, **18**, 817 (1994).

Scokaert, P. O. M. and J. B. Rawlings, "Feasibility Issues in Linear Model Predictive Control," *AIChE J.*, **45**(8), 1649–1659 (1999).

Zheng, A. and M. Morari, "Stability of Model Predictive Control with Mixed Constraints," *IEEE Trans. Auto. Cont.*, **40**(10), 1818–1823 (1995).

A Computationally Efficient Formulation of Robust Model Predictive Control using Linear Matrix Inequalities

Zhaoyang Wan* and Mayuresh V. Kothare†
Department of Chemical Engineering
Lehigh University
Bethlehem, PA 18015, U.S.A.

Abstract

In this paper, we present an off-line approach for robust constrained MPC synthesis that gives an explicit control law using Linear Matrix Inequalities (LMIs). This off-line approach can address a broad class of model uncertainty descriptions with guaranteed robust stability of the closed-loop system and substantial reduction of the on-line MPC computation.

Keywords

Model predictive control, Linear matrix inequalities, Asymptotically stable invariant ellipsoid

Introduction

The practicality of Model Predictive Control (MPC) is partially limited by its ability to solve optimization problems in real time. Moreover, when MPC incorporates explicit plant-model uncertainty, the additional constraints imposed to guarantee robust stability result in significant growth of the on-line MPC computation. Researchers have begun to study methods for fast computation of an optimal or suboptimal solution to the quadratic programming problem associated with nominal MPC (Bemporad et al., 1999; Zheng, 1999; Van Antwerp and Braatz, 2000). For systems with polytopic model uncertainty and input constraints, receding horizon dual-mode paradigm can be used to reduce the computational complexity in MPC (Lee and Kouvaritakis, 2000). In this paper, we present an off-line approach for robust constrained MPC synthesis for both polytopic uncertain systems and norm bounded uncertain systems.

Background

Models for Uncertain Systems

Consider a linear time varying (LTV) system

$$\begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= Cx(k) \\ \begin{bmatrix} A(k) & B(k) \end{bmatrix} &\in \Omega \end{aligned} \tag{1}$$

where $u(k) \in \mathcal{R}^{n_u}$ is the control input, $x(k) \in \mathcal{R}^{n_x}$ is the state of the plant and $y(k) \in \mathcal{R}^{n_y}$ is the plant output.

For polytopic uncertainty, Ω is the polytope $Co\{[A_1 \quad B_1], ..., [A_L \quad B_L]\}$, where Co denotes the convex hull, $[A_i \quad B_i]$ are vertices of the convex hull. Any $[A \quad B]$ within the convex set Ω is a linear combination of the vertices $A = \sum_{j=1}^{L} \alpha_j A_j$, $B = \sum_{j=1}^{L} \alpha_j B_j$ with $\sum_{j=1}^{L} \alpha_j = 1,\ 0 \le \alpha_j \le 1$.

For norm bounded uncertainty, the LTV system is expressed as a LTI system with uncertainties or perturbations appearing in a feedback loop:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + B_p p(k) \\ y(k) &= Cx(k) \\ q(k) &= C_q x(k) + D_{qu} u(k) \\ p(k) &= (\Delta q)(k) \end{aligned} \tag{2}$$

where the operator $\Delta = diag(\Delta_1, ..., \Delta_l)$ with $\Delta_i : \mathcal{R}^{n_i} \longrightarrow \mathcal{R}^{n_i}$, $i = 1, ..., l$. Δ can represent either a memoryless time varying matrix with $\|\Delta_i(k)\|_2 \equiv \bar{\sigma}(\Delta_i(k)) \le 1$, $k \ge 0$, or a convolution operator (for, e.g., a stable LTI dynamical system), with the operator norm induced by the truncated l_2 -norm less than 1, i.e., $\sum_{j=0}^{k} p_i(j)^T p_i(j) \le \sum_{j=0}^{k} q_i(j)^T q_i(j),\ i = 1, ..., l,\ \forall k \ge 0$.

On-line Robust Constrained MPC Using LMIs

Consider the following problem, which minimizes the robust or worst case infinite horizon quadratic objective function:

$$\min_{u(k+i|k)=F(k)x(k+i|k)} \ \max_{[A(k+i)\ B(k+i)]\in\Omega,\ i\ge 0} J_\infty(k) \tag{3}$$

subject to

$$|u_r(k+i|k)| \le u_{r,\max}, \quad i \ge 0,\ r = 1, 2, ..., n_u \tag{4}$$

$$|y_r(k+i|k)| \le y_{r,\max}, \quad i \ge 1,\ r = 1, 2, ..., n_y \tag{5}$$

where $J_\infty(k) = \sum_{i=0}^{\infty} [x(k+i|k)^T Q_1 x(k+i|k) + u(k+i|k)^T R u(k+i|k)]$ with $Q_1 > 0, \ R > 0$. In (3), we assume that at each sampling time k, a constant state feedback law $u(k+i|k) = F(k)x(k+i|k)$ is used to minimize the worst case value of $J_\infty(k)$. Following an approach given in (Kothare et al., 1996), it is easy to derive an upper bound on $J_\infty(k)$. First, at sampling time k, define a quadratic function $V(x) = x^T P(k) x,\ P(k) > 0$. For any $[A(k+i)\ B(k+i)] \in \Omega,\ i \ge 0$, suppose $V(x)$ satisfies the

*email: zhw2@lehigh.edu
†Corresponding author: phone (610) 758-6654, fax (610) 758-5057, email: mayuresh.kothare@lehigh.edu

following robust stability constraint:

$$V(x(k+i+1|k)) - V(x(k+i|k)) \leq -\left[x(k+i|k)^T Q_1 x(k+i|k) + u(k+i|k)^T R u(k+i|k)\right] \quad (6)$$

Summing (6) from $i = 0$ to $i = \infty$ and requiring $x(\infty|k) = 0$ or $V(x(\infty|k)) = 0$, it follows that

$$\max_{[A(k+i)\ B(k+i)]\in\Omega,\ i\geq 0} J_\infty(k) \leq V(x(k|k)) \leq \gamma \quad (7)$$

(6) and (7) give an upper bound on $J_\infty(k)$. The condition $V(x(k|k)) \leq \gamma$ in (7) can be expressed equivalently as the LMI

$$\begin{bmatrix} 1 & x(k|k)^T \\ x(k|k) & Q \end{bmatrix} \geq 0, \quad (8)$$

where $Q = \gamma P(k)^{-1}$.

The robust stability constraint (6) for the system (1) is satisfied if for each vertex of Ω

$$\begin{bmatrix} Q & QA_j^T + Y^T B_j^T & QQ_1^{1/2} & Y^T R^{1/2} \\ A_j Q + B_j Y & Q & 0 & 0 \\ Q_1^{1/2} Q & 0 & \gamma I & 0 \\ R^{1/2} Y & 0 & 0 & \gamma I \end{bmatrix} \geq 0,\ j = 1, ..., L \quad (9)$$

where, $Q = \gamma P(k)^{-1}$ and $F(k)$ is parameterized by YQ^{-1}. This set of conditions is convex in Ω. So if (9) is satisfied, then for any $[A(k+i)\ B(k+i)] \in \Omega$, $i \geq 0$, (6) is satisfied. For the system (2), the stability constraint (6) is satisfied if

$$\begin{bmatrix} Q & QA^T + Y^T B^T & QC_q^T + Y^T D_{qu}^T & QQ_1^{1/2} & Y^T R^{1/2} \\ AQ + BY & Q - B_p \Lambda B_p^T & 0 & 0 & 0 \\ C_q Q + D_{qu} Y & 0 & \Lambda & 0 & 0 \\ Q_1^{1/2} Q & 0 & 0 & \gamma I & 0 \\ R^{1/2} Y & 0 & 0 & 0 & \gamma I \end{bmatrix} \geq 0 \quad (10)$$

where $\Lambda = diag(\lambda_1 I_{n_1}, ..., \lambda_l I_{n_l}) > 0$. The input constraints (4) are satisfied if there exists a symmetric matrix X such that

$$\begin{bmatrix} X & Y \\ Y^T & Q \end{bmatrix} \geq 0 \quad (11)$$

with $X_{rr} \leq u_{r,\max}^2$, $r = 1, 2, ..., n_u$. Similarly, the output constraints (5) for the system (1) are satisfied if there exists a symmetric matrix Z such that for each vertex of Ω,

$$\begin{bmatrix} Z & C(A_j Q + B_j Y) \\ (A_j Q + B_j Y)^T C^T & Q \end{bmatrix} \geq 0,\ j = 1, ..., L \quad (12)$$

with $Z_{rr} \leq y_{r,\max}^2$, $r = 1, 2, ..., n_y$. The output constraints (5) for the system (2) are satisfied if for each row of C

$$\begin{bmatrix} y_{r,\max}^2 Q & (C_q Q + D_{qu} Y)^T & (AQ + BY)^T C^{<r>T} \\ C_q Q + D_{qu} Y & T_r & 0 \\ C^{<r>}(AQ + BY) & 0 & I - C^{<r>} B_p T_r B_p^T C^{<r>T} \end{bmatrix} \geq 0 \quad (13)$$

where $C^{<r>}$ is the rth row of C, $T_r = diag(t_{r,1} I_{n_1}, ..., t_{r,l} I_{n_l}) > 0$, $r = 1, 2, ..., n_y$

Theorem 1 *(On-line robust constrained MPC) (Kothare et al., 1996). For the system (1) or (2), at sampling time k, let $x(k) = x(k|k)$ be the state. Then the state feedback matrix $F(k)$ in the control law $u(k+i|k) = F(k)x(k+i|k)$, $i \geq 0$, which minimizes the upper bound γ on the worst case robust MPC objective function $J_\infty(k)$, is given by $F(k) = YQ^{-1}$ where $Q > 0$ and Y are obtained from the solution (if it exists) of one of the following linear objective minimization problems:*
(a) for system (1), $\min_{\gamma,Q,X,Y,Z} \gamma$ subject to (8), (9), (11) and (12).
(b) for system (2), $\min_{\gamma,Q,X,Y,\Lambda,T_1,...,T_{n_y}} \gamma$ subject to (8), (10), (11) and (13).

Furthermore, the time varying state feedback matrix $F(k)$ robustly asymptotically stabilizes the closed-loop system.

Off-line Robust Constrained MPC

In this section, we present an off-line approach based on the concept of *the asymptotically stable invariant ellipsoid*. Without loss of generality, we use the algorithm for polytopic uncertain systems (Theorem 1 (a)) to illustrate the following lemmas, corollaries and theorem. Similar results can be obtained for Theorem 1 (b).

Asymptotically Stable Invariant Ellipsoid

Definition 1 *(Asymptotically stable invariant ellipsoid). Given a discrete dynamical system $x(k+1) = f(x(k))$, a subset $\mathcal{E} = \{x \in \mathcal{R}^{n_x} | x^T Q^{-1} x \leq 1\}$ of the state space $\mathcal{R}^{n_x}$ is said to be an asymptotically stable invariant ellipsoid, if it has the property that, whenever $x(k_1) \in \mathcal{E}$ at $k_1 \geq 0$, then $x(k) \in \mathcal{E}$ for all times $k \geq k_1$ and $x(k) \longrightarrow 0$ as $k \longrightarrow \infty$.*

Lemma 1 *Consider a closed-loop system composed of a plant (1) and a static state feedback controller $u = YQ^{-1}x$, where Y and Q^{-1} are obtained by applying the robust constrained MPC defined in Theorem 1 (a) to a system state x_0. Then, the subset $\mathcal{E} = \{x \in \mathcal{R}^{n_x} | x^T Q^{-1} x \leq 1\}$ of the state space $\mathcal{R}^{n_x}$ is an asymptotically stable invariant ellipsoid.*

Proof. When the robust constrained MPC defined in Theorem 1 (a) is applied to a state x_0 of a plant (1), the only LMI in Theorem 1 (a) that depends on the system state is (8) which is automatically satisfied for all states within the ellipsoid $\mathcal{E}$. So the minimizer γ, Q, X, Y and Z at the state x_0 is also feasible (though not necessarily optimal) for any other state in $\mathcal{E}$, which means we can apply the state feedback law $u = YQ^{-1}x$ to all the states in $\mathcal{E}$ with the satisfation of (9), (11) and (12).

Consider the closed-loop system composed of the plant (1) and the static state feedback controller $u = YQ^{-1}x$, where Y and Q^{-1} are obtained by applying the robust constrained MPC defined in Theorem 1 (a) to a system state x_0. Then, for any state $\tilde{x}(k) \in \mathcal{E}$, the satisfaction of (9) ensures that in real time $\tilde{x}(k+i+1)^T Q^{-1} \tilde{x}(k+i+1) < \tilde{x}(k+i)^T Q^{-1} \tilde{x}(k+i) \leq 1$, $i \geq 0$. Thus, $\tilde{x}(k+i) \in \mathcal{E}$, $i \geq 0$ and $\tilde{x}(k+i) \longrightarrow 0$ as $i \longrightarrow \infty$, which establish that $\mathcal{E}$ is an asymptotically stable invariant ellipsoid.

Remark 1 *From Lemma 1, we know that for system (1), an asymptotically stable invariant ellipsoid can be constructed by applying Theorem 1 (a) to an arbitrary feasible state x_0. The minimization at x_0 gives a state feedback law $u = YQ^{-1}x$. Once a state enters into the ellipsoid, the static state feedback controller $u = YQ^{-1}x$ can keep it within the ellipsoid and converge it to the origin. The following corollaries state the dependence of the ellipsoidal weighting matrix Q^{-1} and the state feedback matrix YQ^{-1} on the state x_0.*

Corollary 1 *For a nominal and unconstrained system, the ellipsoid weighting matrix Q^{-1} and the feedback matrix YQ^{-1} obtained by applying Theorem 1 (a) without (11) and (12) to an arbitrary state x_0 are $\frac{1}{x_0^T P_{are}\, x_0} P_{are}$ and $-(R+B^T P_{are} B)^{-1} B^T P_{are} A$ respectively, where P_{are} is the solution of the Algebraic Riccati Equation $P_{are} - A^T P_{are} A + A^T P_{are} B(R + B^T P B)^{-1} B^T P_{are} A - Q_1 = 0$.*

Corollary 2 *For an uncertain and unconstrained system (1), the ellipsoid weighting matrix Q^{-1} and the feedback matrix YQ^{-1} obtained by applying Theorem 1 (a) without (11) and (12) to an arbitrary state x_0 are $\frac{1}{x_0^T P\, x_0} P$ and F, where P and F are constant along an arbitrary one-dimensional subspace $\mathcal{S} = \{x \in R^{n_x} | \alpha x_0,\ \alpha \in R\}$ of the state space $\mathcal{R}^{n_x}$.*

Proof. Along the one-dimensional subspace $\mathcal{S}$, let γ_{opt}, Q_{opt} and Y_{opt} be the minimizers at x_0. It is easy to verify that $\alpha^2\gamma_{\text{opt}}$, $\alpha^2 Q_{\text{opt}}$ and $\alpha^2 Y_{\text{opt}}$ is the minimizers at αx_0. Therefore, at αx_0, $F = \alpha^2 Y_{\text{opt}} \frac{1}{\alpha^2} Q_{\text{opt}}^{-1} = F_{\text{opt}}$ and $P = \alpha^2 \gamma_{\text{opt}} \frac{1}{\alpha^2} Q_{\text{opt}}^{-1} = P_{\text{opt}}$.

Remark 2 *From Corollary (1) and (2), we can see that when we apply Theorem 1 (a) without (11) and (12) to an arbitrary state x_0, for a nominal system, the state feedback matrix is independent of the state, while for an uncertain system (1), the state feedback matrix is independent of the state along a one dimensional subspace, but may change according to different orientations of the one dimensional subspace.*

Off-line Robust Constrained MPC

For a constrained system, we know that the nearer the state is to the origin, the less restrictions on the choice of the feedback matrix. To provide the state space with a sense of distance, we define the norm of any vector within the asymptotically stable invariant ellipsoid $\mathcal{E} = \{x \in \mathcal{R}^{n_x} | x^T Q^{-1} x \leq 1\}$ as $\|x\|_{Q^{-1}} \triangleq \sqrt{x^T Q^{-1} x}$. The distance between the state and the origin is the norm of the state. On-line MPC in Theorem 1 has the advantage of determining the control law based on the norm of the state (see the derivation of (11) and (12) in (Kothare et al., 1996)). Off-line we can still achieve this. When we apply Theorem 1 to a state far from the origin, the resulting asymptotically stable invariant ellipsoid has a more restrictive feedback matrix. It is not necessary to keep this feedback matrix constant while the state is converging to the origin. We can construct inside the ellipsoid another asymptotically stable invariant ellipsoid based on a state nearer to the origin, which can have a less restrictive feedback matrix. By adding asymptotically stable invariant ellipsoids one inside another, we have more freedom to adopt suitable feedback matrices based on the distance between the state and the origin.

Lemma 2 *(Existence) Consider the minimization defined in Theorem 1 (a). If there exists a minimizer γ, Q, X, Y and Z at x, then, at an arbitrary $\tilde{x}$ satisfying $\|\tilde{x}\|^2_{Q^{-1}} < 1$ there exists a minimizer $\tilde{\gamma}, \tilde{Q}, \tilde{X}, \tilde{Y}$ and $\tilde{Z}$ for the minimization defined in Theorem 1 (a) with an additional constraint $Q > \tilde{Q}$.*

Proof. Consider the minimization defined in Theorem 1 (a) at x. The minimizer defines an ellipsoid $\mathcal{E} = \{x \in \mathcal{R}^{n_x} | x^T Q^{-1} x \leq 1\}$. An arbitrary $\tilde{x}$ satisfying $\|\tilde{x}\|^2_{Q^{-1}} < 1$ means that the state $\tilde{x}$ is inside $\mathcal{E}$. So $\exists \alpha > 1$, $\|\alpha\tilde{x}\|^2_{Q^{-1}} = 1$ and $\frac{1}{\alpha^2}\gamma$, $\frac{1}{\alpha^2}Q$, $\frac{1}{\alpha^2}X$, $\frac{1}{\alpha^2}Y$, $\frac{1}{\alpha^2}Z$ is a feasible solution for the minimization defined in Theorem 1 (a) with the additional constraint $Q > \frac{1}{\alpha^2}Q$ satisfied.

Algorithm 1 *(Off-line robust constrained MPC) Off-line, given an initial feasible state x_1, generate a sequence of minimizers γ_i, Q_i, X_i, Y_i and Z_i $(i = 1, ...N)$ as follows.*

1. compute the minimizer γ_i, Q_i, X_i, Y_i and Z_i at x_i by using Theorem 1 with an additional constraint $Q_{i-1} > Q_i$ (ignored at $i = 1$), store Q_i^{-1} and $F_i (= Y_i Q_i^{-1})$ in a look-up table;

2. if $i < N$, choose a state x_{i+1} satisfying $\|x_{i+1}\|^2_{Q_i^{-1}} < 1$. Go to 1.

On-line, given a dynamical system (1) and an initial state $x(0)$ satisfying $\|x(0)\|^2_{Q_1^{-1}} \leq 1$, let the state be $x(k)$

at time k. Perform a bisection search over Q_i^{-1} in the lookup table until a Q_i^{-1} is found satisfying $\|x(k)\|^2_{Q_i^{-1}} \leq 1$ and $\|x(k)\|^2_{Q_{i+1}^{-1}} > 1$ ($i = 1, ..., N-1$), or $\|x(k)\|^2_{Q_i^{-1}} \leq 1$ ($i = N$). Apply the control law $u(k) = F_i x(k)$.

Theorem 2 *Given a dynamical system (1) and an initial state $x(0)$ satisfying $\|x(0)\|^2_{Q_1^{-1}} \leq 1$, the off-line robust constrained MPC algorithm 1 robustly asymptotically stabilizes the closed-loop system.*

Proof. For the minimization at x_i, $i = 2, ..., N$, the additional constraint $Q_{i-1} > Q_i$ is equivalent to ${Q_{i-1}}^{-1} < Q_i^{-1}$. This implies that the constructed asymptotically stable invariant ellipsoid $\mathcal{E}_i = \{x \in \mathcal{R}^{n_x} | x^T Q_i^{-1} x \leq 1\}$ is inside $\mathcal{E}_{i-1}$, i.e., $\mathcal{E}_i \subset \mathcal{E}_{i-1}$. So for a fixed x, $\|x\|^2_{Q_i^{-1}}$ is monotonic with respect to the index i, which ensures the on-line bisection search over the lookup table finds a unique Q_i^{-1}.

Given a dynamical system (1) and an initial state $x(0)$ satisfying $\|x(0)\|^2_{Q_1^{-1}} \leq 1$, the closed–loop system becomes

$$x(k+1) = \begin{cases} (A(k) + B(k)F_i)\, x(k) & \|x(k)\|^2_{Q_i^{-1}} \leq 1 \\ & \cap\ \|x(k)\|^2_{Q_{i+1}^{-1}} > 1 \\ & (i = 1, ..., N-1) \\ (A(k) + B(k)F_N)\, x(k) & \|x(k)\|^2_{Q_N^{-1}} \leq 1 \end{cases}$$

When $x(k)$ satisfies $\|x(k)\|^2_{Q_i^{-1}} \leq 1$ and $\|x(k)\|^2_{Q_{i+1}^{-1}} > 1$, $i = 1, ..., N-1$, the control law $u(k) = F_i x(k)$ corresponding to the ellipsoid $\mathcal{E}_i$ is guaranteed to keep the state within $\mathcal{E}_i$ and converge it into the ellipsoid $\mathcal{E}_{i+1}$, and so on. Lastly, the smallest ellipsoid $\mathcal{E}_N$ is guaranteed to keep the state within $\mathcal{E}_N$ and converge it to the origin.

Remark 3 *Algorithm 1 is a general approach to construct a Lyapunov function for uncertain and constrained systems. The Lyapunov function is*

$$V(x) = \begin{cases} x^T Q_i^{-1}\, x & \|x(k)\|^2_{Q_i^{-1}} \leq 1\ \cap\ \|x(k)\|^2_{Q_{i+1}^{-1}} > 1 \\ & (i = 1, ..., N-1) \\ x^T Q_N^{-1} x & \|x(k)\|^2_{Q_N^{-1}} \leq 1 \end{cases}$$

Note that this Lyapunov function is not continuous on the boundary of ellipsoids. Due to the special characteristics of the robust asymptotically stable invariant ellipsoid, it is enough to have $V(x)$ be monotonically decreasing within the smallest ellipsoid and within each ring region between two adjacent ellipsoids to stabilize the closed-loop system.

From algorithm 1, we can see that the choice of the state x_{i+1} satisfying $\|x_{i+1}\|^2_{Q_i^{-1}} < 1$ is arbitrary. From the point of view of easy implementation, we provide the following suggestions. We can choose an arbitrary one dimensional subspace $\mathcal{S} = \{x \in R^{n_x} | \alpha x_0,\ \alpha \in R\}$, and discretize it and construct a set of discrete points, $\mathcal{S}^d = \{x \in R^{n_x} | \alpha_i x^{\max},\ 1 \geq \alpha_1 > ... > \alpha_N > 0\}$. Because the asymptotically stable invariant ellipsoid constructed for each discrete point actually passes through that point, $\|\alpha_{i+1} x^{\max}\|^2_{Q_i^{-1}} < \|\alpha_i x^{\max}\|^2_{Q_i^{-1}} = 1$ is satisfied. And in order to obtain a look-up table that can cover a very large portion of the state space with a limited number of discrete points, we suggest a discretization of the one dimensional subspace using a logarithmic scale.

Complexity Analysis

The on-line computation mainly comes from the bisection search in a lookup table. Because a sequence of N stored Q_i^{-1} (N generally less than 20) requires $\log_2 N$ searches and the matrix-vector multiplication in one search has quadratic growth $O(n_s^2)$ in the number of flops with n_s the number of state variables, the total number of flops required to calculate an input move is $O(n_s^2 \log_2 N)$. On the other hand, the fastest interior point algorithms are cubic growth in the number of flops as a function of problem size, which is $O((\frac{n_s^2}{2} + n_s n_c)^3)$ for the robust algorithm in Theorem 1 (a) with n_c the number of manipulated variables. So we can conclude that this off-line approach can substantially reduce the on-line computation.

Example

Consider the linearized model derived for a single, nonisothermal CSTR (Marlin, 1995), which is discretized using a sampling time of 0.15 min and given in terms of perturbation variables as follows:

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 0.85 - 0.1\alpha(k) & -0.001\alpha(k) \\ \alpha(k)\beta(k) & 0.05 + 0.01\alpha(k)\beta(k) \end{bmatrix} x(k) \\ &\quad + \begin{bmatrix} 0.15 & 0 \\ 0 & -0.9 \end{bmatrix} u(k) \\ y(k) &= x(k) \end{aligned}$$

where x is a vector of the reactor concentration and temperature, and u is a vector of the feed concentration and the coolant flow, $1 \leq \alpha(k) = k_0/10^9 \leq 10$ and $1 \leq \beta(k) = -\Delta H_{rxn}/10^7 \leq 10$. The polytopic uncertain set has four vertices. The robust performance objective function is defined as (3) subject to $|u_1(k+i|k)| \leq 0.5$ kmole/m^3 and $|u_2(k+i|k)| \leq 1$m^3/min, $i \geq 0$, with $Q_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $R = 0.2$. We choose the x_1 axis as the one dimensional subspace, and discretize it in ten points. Figure 1 shows the ellipsoids defined by Q_i^{-1} for all ten discrete points.

Given an initially perturbed state $x(0) = \left[\begin{smallmatrix} 0.1 \\ 2 \end{smallmatrix}\right]$, Figure 2 shows the closed-loop responses of the system cor-

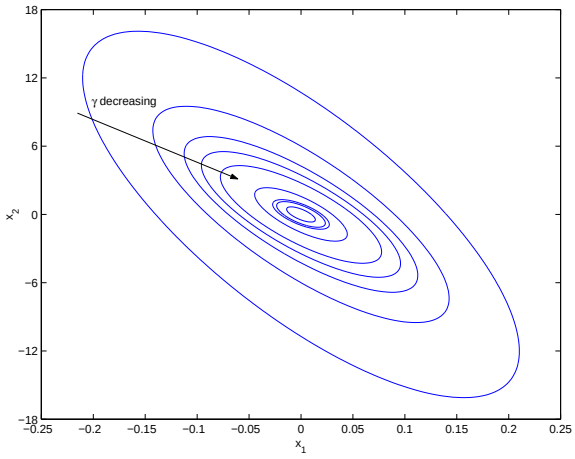


Figure 1: The ellipsoids defined by Q_i^{-1} for all ten discrete points.

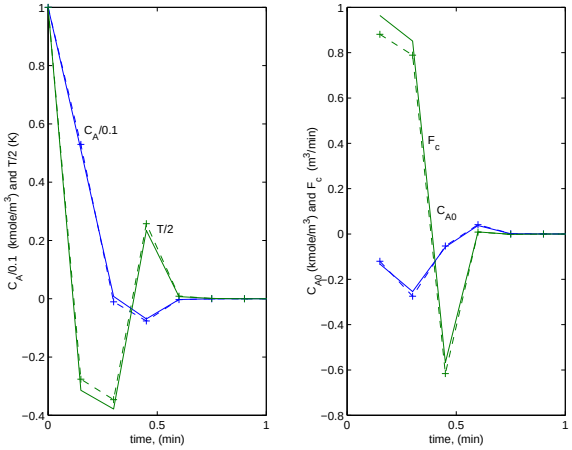


Figure 2: Closed-loop responses: solid lines, on-line MPC algorithm in Theorem 1; dashed lines with (+), off-line MPC in Algorithm 1.

responding to $\alpha(k) \equiv 1.1$ and $\beta(k) \equiv 1.1$. The off-line approach gives nearly the same performance as the on-line robust constrained MPC algorithm (Kothare et al., 1996). The average time for the off-line MPC to get a feedback gain is 6.6×10^{-4} sec, while the average time required for the on-line MPC is 0.6 sec. The calculation of this off-line approach is 900 times faster than that required for on-line MPC.

Conclusions

In this paper, based on the concept of *asymptotically stable invariant ellipsoid* and by using LMIs, we developed an off-line robust constrained MPC algorithm with guaranteed robust stability of the closed-loop system and substantial reduction of on-line MPC computation.

References

Bemporad, A., M. Morari, V. Dua, and E. N. Pistikopoulos, The explicit linear quadratic regulator for constrained systems, Technical Report AUT 99-16, Automatic Control Laboratory, Swiss Federal Institute of Technology (ETH), Zürich (1999).

Kothare, M. V., V. Balakrishnan, and M. Morari, "Robust constrained model predictive control using linear matrix inequalities," *Automatica*, **32**(10), 1361–1379 (1996).

Lee, Y. I. and B. Kouvaritakis, "A Linear Programming Approach to Constrained Robust Predictive Control," *IEEE Trans. Auto. Cont.*, **45**(9), 1765–1770 (2000).

Marlin, T. E., *Process Control: Designing Processes and Control Systems for Dynamic Performance*. McGraw Hill, New York (1995).

Van Antwerp, J. G. and R. D. Braatz, "Fast model predictive control of sheet and film processes," *IEEE Trans. Cont. Sys. Tech.*, **8**(3), 408–17 (2000).

Zheng, A., "Reducing on-line computational demands in model predictive control by approximating QP constraints," *J. Proc. Cont.*, **9**(4), 279–290 (1999).

Time Series Reconstruction from Quantized Measurements

M. Wang
Centre for Process Systems Engineering
Imperial College of Science Technology and Medicine
Prince Consort Road, London SW7 2BY

S. Saleem and N. F. Thornhill*
Department of Electronic and Electrical Engineering
UCL (University College London)
Torrington Place, London WC1E 7JE

Abstract

This paper describes a Quantization Regression (QR) algorithm which generates a nonlinear estimate of an autoregressive time series from quantized measurements. Its purpose is to retrieve the underlying information from quantized signals such as those from the analogue to digital converter of a plant instrument. The reconstructed signals have uses in data-centric applications such as controller performance assessment and system identification. The algorithm based upon Ziskand and Hertz is a combination of the 'Gaussian Fit' scheme of Curry with expectation-maximization (EM) algorithm of Dempster et al. The performance of Quantization Regression algorithm is compared with two other methods in fitting of an autoregressive time series for the reconstruction of a quantized signal.

Keywords

EM algorithm, Linear regression, Quantization, Performance analysis, Time series analysis

Introduction

A great deal of engineering data occurs in the form of time series where observations are dependent and where the nature of this dependence is of interest. For example, Desborough and Harris (1992) used routine closed loop process data to estimate the normalized control loop performance index. In that case, an AR model was implemented. Thornhill et al. (1999) observed that quantization of the measurements influenced the normalized performance index and thus motivated the need for an algorithm to accurately recover a signal from a quantized time series. Quantization is often observed in the outputs of process instruments even though 10 bit A/D conversion provides 1024 quantization levels because a measurement controlled to a steady value is likely to sample just a few of the available quantizer levels.

Ziskand and Hertz (1993) proposed an algorithm to estimate coefficients of a quantized autoregressive (qAR) process. Their implementation and examples illustrated the case of one-bit (two level) quantization. The algorithm was based on the following two developments: Dempster et al. (1977) presented a method for computing maximum likelihood estimates from "incomplete" data, i.e. data having a many-to-one mapping in the measurement function. Quantized process data are incomplete because many values of the underlying signal map to each quantizer level. The algorithm comprised two steps: the first expectation, the second maximization, and was called the EM algorithm. Shumway and Stoffer (1982) proposed an approach to smoothing and forecasting for time series with incomplete observations. The EM algorithm was used in conjunction with Kalman smoothed estimators to derive a recursive procedure for estimating the qAR parameters by maximum likelihood. The algorithm is a general technique for finding maximum likelihood estimates from incomplete data (Little and Rubin, 1986).

Expectation calculations for the maximum likelihood step are provided by the Gaussian Fit algorithm. The method was proposed by Curry (1970) for a discrete-time nonlinear filter that recursively fits a Gaussian distribution to the first two moments of the conditional distribution of a system state vector. The Gaussian Fit algorithm is easy to compute and can handle non-stationary data and its operation is independent of the quantization scheme used. However, it requires more computation than a linear filter and can be applied only to Gauss-Markov processes. It is applicable to quantized process data because such data can be expressed as a Gauss-Markov process in state-space form.

The algorithm presented by Ziskand and Hertz (1993) estimated the coefficients of several superimposed qAR signals of known model order using a two-level quantizer. The signals represented sinusoids at different frequencies. The contribution of this paper is the extension of quantized regression to multiple quantizer levels and to qAR signals of order 2 and higher. The Akaike Information Criterion was used to determine the order of the AR model. Further, the QR algorithm was compared with two other methods to illustrate why it can recover the underlying signal better. Simulations and experimental data support the studies. The conclusion is that the QR algorithm can optimally estimate the model parameters and recover the underlying signal at the same time.

Methods

Problem Description

Quantized autoregressive (AR) time series are the subject of the work. The process may be written as follows:

$$S(n) = \phi \cdot S(n-1) + w(n)$$
$$x(n) = h \cdot S(n) + v(n)$$
$$z(n) = g(x(n))$$

*Author to whom all correspondence should be addressed. email: n.thornhill@ee.ucl.ac.uk, Tel: +44 20 7679 3983, Fax: +44 20 7388 9325

where $S(n) = (s(n-m+1), \ldots, s(n))^T$ is a state vector, $\{x(n)\}_{n=1}^N$ the N samples from an autoregressive process of order m and $\{z(n)\}_{n=1}^N$ the quantized measurements of the process. $w(n)$ and $v(n)$ are white noise samples with mean zero and variances σ_w^2 and σ_0^2 respectively and independent from each other. The state transition matrix, ϕ, has the autoregressive coefficients $a_1, a_2, \ldots, a_m$ in the last row:

$$\phi = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ a_m & \cdots & \cdots & a_2 & a_1 \end{pmatrix}$$

The observation vector $h = (0, 0, \ldots, 0, 1)$ is of length m, and g is the non-linear quantizer function whose input is the AR signal plus noise. The algorithm is suitable for both uniform and non-uniform quantizers, although in this paper uniform quantization intervals were used. The quantization interval is the distance between quantizer levels. The problem is to determine the model order, m, to estimate the AR coefficients $a_1, a_2, \ldots, a_m$ and to recover the underlying signal $\{s(n)\}_{n=1}^N$.

Implementation of QR Algorithm

The Quantization Regression (QR) algorithm is iterative and includes two main conceptual steps:

Step 1: A modified Kalman smoothing algorithm finds smoothed estimates and their covariances. At step n the key calculations are:

$$S_n^n = S_n^{n-1} + K_n \left((E(x(n))|z_n) - hS_n^{n-1} \right)$$
$$P_n^n = P_n^{n-1} - K_n h P_n^{n-1} + K_n \mathrm{cov}\left(x(n)|z_n\right) K_n^T$$

where S_n^{n-1} and P_n^{n-1} are one step ahead predictions of the state vector and its variance, z_n the sequence of quantized measurements and K_n is the Kalman gain. A conventional Kalman filter would use the quantized measurement z_n directly, but the modified algorithm calculates the expected value of x_n given that the quantized measurement falls within the observed quantization interval. The expectation is computed using the Gaussian Fit approximation (Curry, 1970). The last term in the calculation for P_n^n captures the inflation in the variance of the estimate caused by quantization.

Step 2: The likelihood function is maximised by iterative adjustment of the AR coefficients $a_1, a_2, \ldots, a_m$, the estimated signal and σ_w^2 and σ_0^2.

Typing errors found in Ziskand and Hertz were corrected as follows where the underline indicates the altered terms,

$$K_n = P_n^{n-1} h^T \underline{(\sigma_x^2)^{-1}}$$
$$P_{n-1}^N = \underline{P_{n-1}^{n-1}} + J_{n-1}\left(P_n^N - P_n^{n-1}\right)\underline{J_{n-1}^T}$$
$$P_{n-1,n-2}^N = P_{n-1}^{n-1} J_{n-2}^T + \underline{J_{n-1}}\left(P_{n,n-1}^N - \phi(r) P_{n-1}^{n-1}\right)\underline{J_{n-2}^T}$$

and the loop in the lag-one covariance calculation was for $n = N, N-1, \ldots, \underline{2}$.

The following improvements were made: (1) σ_0^2 was updated during the iteration rather than being taken as known; (2) the innovative form of the log liklihood function (Shumway and Stoffer, 1982) was used as a stopping criterion.

$$\Delta(\log L) = -\frac{1}{2}\sum_{n=1}^{N} \log(\sigma_x^2) - \frac{1}{2\sigma_x^2}\sum_{n=1}^{N}\left(z(n) - hS_n^{n-1}\right)^T\left(z(n) - hS_n^{n-1}\right)$$

Iterations stopped when the innovation became small; (3) the initial variance in the iteration was fixed and the state vector was initialised thus: $S_0^0(r+1) = S_0^N(r)$.

Determination of Model Order and Quality of Fit

Ziskand and Hertz assumed the model order m was known. Where the model order is unknown one approach to determining m is to fit AR process of progressively high order, to calculate the sum of squared errors (SSE) for each value of m and to plot SSE against m (Chatfield, 1989). One chooses the value of m where the addition of extra parameters gives little improvement in fit. The method does not work here since the aim is to recover the underlying signal not the quantized observation. An approach suited to maximum likelihood estimation is to use Akaike Information criterion (AIC) to determine model order. The required complete-data log likelihood function ($\log L$) is presented in Ziskand and Hertz (1993), and Shumway and Stoffer (1982). AIC can be calculated from $\log L$ and the model order m as $AIC = -2\log L + 2m$. The model order is the value of m where the largest decrease occurs.

The multiple coefficient of determination R^2 (Scheaffer and McClave, 1995) was used to determine the quality of the reconstruction:

$$R^2 = 1 - \frac{\sum_{n=1}^{N}\left(y(n) - \hat{y}(n)\right)^2}{\sum_{n=1}^{N}\left(y(n) - \bar{y}\right)^2}$$

where $\hat{y}(n)$ is the reconstruction of a true underlying signal $y(n)$ and $\bar{y}$ is its mean. $R^2 = 0$ implies the variance of the reconstruction error is as large as the variance of the measurements, and thus that there is a complete lack of fit of the model, while $R^2 = 1$ implies a perfect reconstruction.

Other Methods for Comparison

Two other methods for estimation of the AR coefficients were implemented for comparison.

One method was a one-step linear least squares estimate (LLS) of the AR coefficients. The quantized data $\{z(n)\}_{n=1}^{N}$ themselves were modelled as an AR series:

$$\begin{pmatrix} z(n) \\ z(n-1) \\ \vdots \\ z(m+1) \end{pmatrix} = \begin{pmatrix} z(n-1) & \cdots & z(n-m) \\ z(n-2) & \cdots & z(n-m-1) \\ \vdots & & \vdots \\ z(m) & \cdots & z(1) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} + \begin{pmatrix} e(n) \\ e(n-1) \\ \vdots \\ e(m+1) \end{pmatrix} \quad (1)$$

The above matrix equation can be written as:

$$Y = X \cdot a + e$$

and least squares estimates of the AR parameters determined from:

$$\hat{a} = \left(X^T X\right)^{-1} X^T Y$$

The above calculation coincides with the linear regression approach in Desborough and Harris (1992) using a one-step prediction. It was also used in Thornhill et al. (1999) where the adverse influence of quantization was noted.

The other method, termed Kalman smoothing, used the same formulation as the QR algorithm but without the Gaussian Fit step. The Kalman smoother equations became:

$$S_n^n = S_n^{n-1} + K_n \left(z(n) - hS_n^{n-1}\right)$$
$$P_n^n = P_n^{n-1} - K_n h P_n^{n-1}$$

Results

Simulation Examples

A unit amplitude sinusoid wave was used as the quantizer input with quantization interval 0.5. Therefore the signal used just five quantizer levels and was coarsely quantized. The sine wave had a period of 51.2 samples, the value being selected so that 10 cycles used 512 samples (a power of 2). Using the method of backward differencing the sine wave $\sin(2\pi t/51.2)$ can be expressed as a two term AR series:

$$y(n) = 1.990y(n-1) - 0.998y(n-2)$$

Therefore we require the QR algorithm to give a model with $m = 2$ and coefficients $a_1 = 1.990$ and $a_2 = -0.998$. The AIC indicated that the optimum model order for the qAR time series was indeed $m = 2$. The QR algorithm was used to reconstruct an AR model with two terms and the results were compared with the reconstruction of $m = 2$ AR models using Kalman smoothing alone and the one-step linear least squares estimate. The recovering ability is compared in the left hand panel of Figure 1, where it can be seen visually that the QR algorithm provided the best reconstruction of the underlying sine wave. The results with the LLS method were the least satisfactory.

The estimated AR coefficients and R^2 values were:

QR:	$a_1 = 1.977$,	$a_2 = -0.993$,	$R^2 = 0.998$
Kalman:	$a_1 = 1.276$,	$a_2 = -0.304$,	$R^2 = 0.996$
LLS:	$a_1 = 0.982$,	$a_2 = -0.018$,	$R^2 = 0.968$

The QR algorithm recovered both coefficients with less than 1% error from the true values, while the LLS coefficients had large errors and could only achieve a model that said the next sample would be almost the same as the previous sample. Kalman smoothing without the Gaussian Fit step gave an intermediate result with the model coefficients in error by 36% and 70%. It is concluded that the major benefit of the QR algorithm with coarsely quantized data is the expectation step using the Gaussian Fit approximation.

A noisy sine wave with a signal to noise ratio of 1:1 was used as a second test signal into the quantizer input. Application of AIC indicated a model order of 5 or 6 for the QR algorithm and 6 for Kalman smoothing. The recovering ability of the three methods for $m = 6$ AR models is compared in the right hand panel of Figure 1. The QR algorithm and Kalman smoothing provided good reconstruction of the underlying sine wave and both were superior to the LLS reconstruction which provided less filtering of the noise. The R^2 values were:

QR:	$R^2 = 0.981$
Kalman:	$R^2 = 0.978$
LLS:	$R^2 = 0.764$

Influence of Quantization Interval

Figure 2 compares the R^2 measure across a range of quantization intervals. The left panel shows the results for the quantized sine wave signal using AR models with model order $m = 2$. The uppermost trend is for the QR algorithm, the lowest is the LLS algorithm and the Kalman smoothing algorithm is in-between. As was suggested by Figure 1, the Gaussian Fit element of the QR algorithm gave benefits over and above its Kalman smoothing component. The benefit increased as the quantization interval increased.

The right hand panel of Figure 2 shows the reconstruction performance of the quantized noisy sine wave using AR models with $m = 6$. The R^2 values for Kalman smoothing and QR were almost identical in this case, and both gave improvements over the LLS method. Thus when the unquantized signal included white noise, no matter what the quantization interval was, the recovery was contributed by Kalman smoothing. It is concluded that the benefit of the Gaussian Fit to the QR algorithm reduces as the influence of quantization reduces relative to the noise.

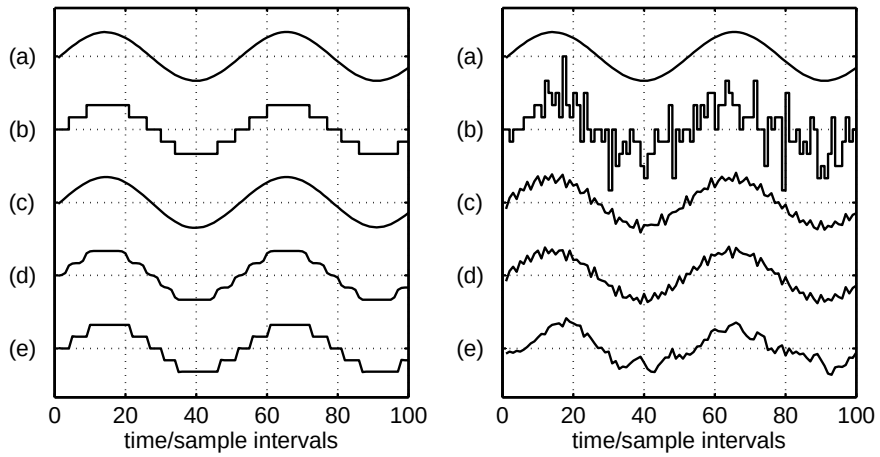


Figure 1: Reconstruction of a sine wave from quantized samples (left) and noisy quantized samples (right). (a): the sine wave signal; (b) quantized signal; (c) reconstructed with QR; (d) reconstructed with Kalman smoothing; (e) reconstructed with LLS.

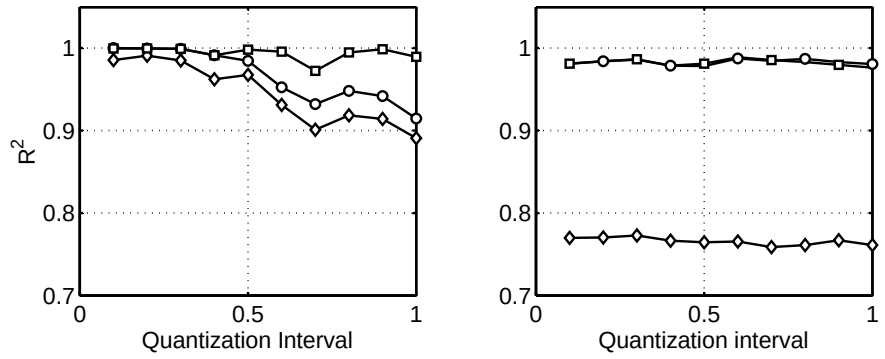


Figure 2: Reconstruction performance for a sine wave from quantized samples (left) and noisy quantized samples (right). QR: squares; Kalman smoothing: circles; LLS: diamonds.

Ensemble Length	a_1	a_2	R^2
N = 1440	1.977	-0.993	0.9981
N = 1280	1.979	-0.994	0.9981
N = 1024	1.977	-0.992	0.9980
N = 768	1.974	-0.989	0.9979
N = 512	1.968	-0.984	0.9977
N = 256	1.952	-0.968	0.9976

Table 1: The influence of ensemble length.

Influence of Ensemble Length

Table 1 shows the influence of data ensemble length, which was examined using the sine wave signal with quantization interval of 0.5.

The table shows that the values of the AR coefficients recovered by the QR algorithm diverged from the correct values as the data ensemble length became smaller. In this example a data ensemble length of 800 gave errors compared to the true values of about 1% in each of the AR coefficients, therefore it is recommended that the data ensemble length be at least 800 samples.

Performance with Plant Data

Figure 3 shows the behaviour of the QR, Kalman smoothing and LLS algorithms with pilot plant measurements. The measurements were from the transient response of the pH control in a buffered fed-batch yeast fermentation process. They were quantized by the A/D converter of the pH probe. The aim of the analysis was

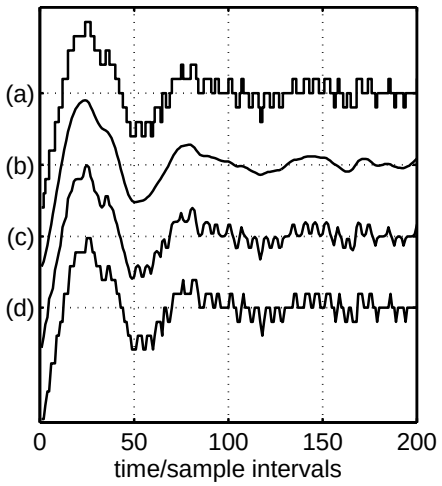


Figure 3: Reconstruction of quantized plant data. (a) quantized trend from plant; (b) reconstructed using QR; (c) using Kalman smoothing; (d) using LLS.

to recover the underlying smooth transient response of the fermenter pH. The closed loop under proportional-plus-integral control of the process has dominant under-damped complex conjugate poles and therefore the AR time series should have $m = 2$. The s-plane poles are $s = -0.025 \pm 0.108j$ corresponding to damped oscillations with a period of 58 samples and an exponential time constant of 40 samples. These s-plane poles map to z-plane poles at $z = 0.969 \pm 0.105j$.

Application of AIC to the QR algorithm indicated that the model order should indeed be $m = 2$. The reconstructions of $m = 2$ AR models using QR, Kalman smoothing and the LLS methods are shown in Figure 3. It was not possible to determine the R^2 values because the true underlying signal was not known. A quantitative comparison of the z-plane poles shows the QR algorithm has achieved the correct reconstruction because the AR model was:

$$y(n) = 1.914y(n-1) - 0.934y(n-2)$$

which has z-plane poles at $z = 0.957 \pm 0.090j$. These poles are near those of the closed loop process and it is concluded that the AR coefficients determined by the QR algorithm captured the under-damped oscillatory behavior of the pH response. The other models were:

$$\begin{array}{ll} \text{Kalman:} & y(n) = 1.210y(n-1) - 0.231y(n-2) \\ \text{LLS:} & y(n) = 0.972y(n-1) + 8 \cdot 10^{-3}y(n-2) \end{array}$$

both of whose z-plane poles are real. These recovered AR time series have no oscillatory behavior of their own and the reconstructions appear oscillatory only because they are driven by the experimental data.

Conclusion

For coarsely quantized signals, the QR (quantized regression) algorithm can recover the underlying signal from quantized observations better than the LLS (linear least squares) algorithm, as measured by the R^2 values. The fundamental reason behind the phenomenon is that QR algorithm based on the Gaussian Fit algorithm and Kalman smoothing can recover some nonlinear parts from the quantized observation, while the LLS algorithm assumes the quantizer has a fixed input-output relationship.

Kalman smoothing was also used without the Gaussian Fit scheme. Its performance was also superior to that of the LLS method. When the effects of quantization were dominant the performance of the QR algorithm was significantly better than that of Kalman smoothing alone. When the underlying signal was noisy, however, the performances of the QR and Kalman smoothing algorithms were similar. Therefore it is concluded that the Gaussian Fit scheme offers the most improvement when quantization is severe, but that when noise is predominant the majority of the benefit is due to Kalman smoothing.

QR reconstruction of a transient response signal from experimental plant data gave an AR series of the correct order and with z-plane poles close to the true values. It can be concluded that QR reconstruction has the capacity to be useful in the accurate reconstruction of autoregressive time series from quantized process data.

Acknowledgments

Meihong Wang gratefully acknowledges the financial support of the Centre for Process Systems Engineering, Imperial College of Science, Technology and Medicine.

References

Chatfield, C., *Time Series Analysis: an introduction*. Chapman and Hall Ltd., London (1989).

Curry, R. E., *Estimation and control with quantized measurements*. MIT Press, Cambridge, MA (1970).

Dempster, A. P., N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc.*, **B-39**, 1–37 (1977).

Desborough, L. and T. Harris, "Performance Assessment Measures for Univariate Feedback Control," *Can. J. Chem. Eng.*, **70**, 1186–1197 (1992).

Little, R. and D. B. Rubin, *Statistical Analysis with Missing Data*. Wiley, New York, NY (1986).

Scheaffer, R. L. and J. T. McClave, *Probability and Statistics for Engineers*. Duxbury Press, Belmont (1995).

Shumway, R. H. and D. S. Stoffer, "An approach to time series smoothing and forecasting using the EM algorithm," *J. Time Series Analysis*, **3**, 253–263 (1982).

Thornhill, N. F., M. Oettinger, and P. Fedenczuk, "Refinery-wide control loop performance assessment," *J. Proc. Cont.*, **9**, 109–124 (1999).

Ziskand, I. and D. Hertz, "Multiple frequencies and AR parameters estimation from one bit quantized signal via the EM algorithm," *IEEE Trans. Signal Proces.*, **41**, 3202–3206 (1993).

Author Index

Subject Index