STUDIES ON PLANTWIDE CONTROL

by

Truls Larsson

A Thesis Submitted for the Degree of Dr. Ing.

Department of Chemical Engineering
Norwegian University of Science and Technology

Submitted July 2000

Abstract

A chemical plant may have thousands of measurements and control loops. By the term *plantwide control* it is *not* meant the tuning and behavior of each of these loops, but rather the *control philosophy* of the overall plant with emphasis on the *structural decisions*. The structural decision includes the selection/placement of manipulators and measurements as well as the *decomposition* of the overall problem into smaller subproblems (the control configuration).

Based on a review of the existing methods, a plantwide control design procedure is proposed. The procedure starts with a top-down analysis of the plant. Where the emphasis is on selecting controlled variables, which will give an easy and more robust optimization. This is achieved by controlling the active constraints, and for the unconstrained degrees of freedom variables with a flat optimum is preferred. A flat optimum indicates that an implementation error or a disturbance will have a small effect on the economic performance. The next step is to choose the throughput manipulator.

The top-down analysis is followed by a bottom-up design of the control system. The bottom-up design is guided by controllability analysis. The goal is first to stabilize the plant (including nearly unstable poles), such that it is possible to operate the plant manually. This is the regulatory control layer. Finally, the supervisory control layer is designed.

One issue that needs to be resolved is if such a control hierarchy can impose new and fundamental limitations, which is not present in the original plant. It is shown, that if the setpoints and measurements of the lower layer are available to the next layer and that the lower layer controller is stable and minimum phase, it is not possible to introduce new fundamental limitations. When the lower layer measurements and/or the lower layer setpoints are unavailable it is possible to introduce new limitations.

The procedure is applied on several applications:

1. A liquid phase reactor with a distillation column and recycle.

2. A gas phase reactor with separator, compressors and recycle.

3. The methanol synthesis loop (a special case of application 2).

4. The Tennessee Eastman challenge problem.

5. An industrial heat integrated distillation columns (from the methanol plant).

All these plants have in common that their behavior is changed by recycle or heat integration.

In the case of a liquid phase reactor, there is no economic penalty for increasing the holdup in the reactor. In fact, the holdup should be as large as possible in order to increase the conversion per pass, which will make separation cheaper. Luyben has proposed a control strategy in which he uses the reactor holdup as a throughput manipulator. This will give an economic penalty which most other authors so far have neglected. For the remaining degree of freedom there is a flat optimum for all but a few variables. One of the conclusion is that the "Luyben-rule", i.e. "fix one flow in recycle" has bad self-optimizing properties and should not be applied to this plant.

For the gas phase plant, the situation is different. Due to compression costs there are a cost associated with the hold-up (pressure). In fact, the optimum is unconstrained in this variable. Control of recycle-rate, purge fraction, or reactor pressure gives a system with good self-optimizing properties. This is linked to the behavior of these variables as conversion increases. As expected purge flow is a bad alternative as a controlled variable. More unexpectedly, inert composition in the recycle turned out to have bad self-optimizing properties. This is also explainable by the behavior of this variable when conversion is increased. The results for the simple gas phase reactor carries well over to the methanol case study.

The Tennessee Eastman problem is a well-studied test problem, but few have studied the selection of controlled variables based on the economics of the plant. In addition to the constrained variables, reactor temperature, C in purge and recycle flow or compressor work, should be controlled. A very common claim is that it is necessary to control the inventory of inert components, this is not true. The shape of the objective function is very unfavorable, and a small implementation error leads to infeasibilities.

The heat-integrated distillation columns are similar too simple distillation in many aspects. But there are differences, e.g. the number of degrees of freedom are different. We argue that the heat transfer area between the two columns, top compositions (of valuable products), and pressure in the low-pressure column should be controlled at their constraints. There is one unconstrained degree of freedom, and for this particular case control of a temperature in the lower part of the column shows good self-optimizing properties.

It is shown that it is not given that poles at the origin will not show up in the relative gain array. It may happen if it is possible to stabilize the pole with two different control loops. This should be seen as an argument for using the frequency dependent relative gain array.

The emphasis in this thesis has been on case studies. By the use of systematic tools for analysis, some “rules” that have been presented in the process control community are shown to have had a weak theoretical basis. The thesis has improved the understanding of the control of a large scale processing plants.

Acknowledgment

When I started to work on my doktor ingeniør degree I had no idea how much work that was waiting for me. A large part of the work has been to get on top of the field of process control, which I feel that I have managed to do. Another large part of the work has been struggling with matlab, and at times I have felt that I was doing a dr.ing. degree in matlab. But still these years has been rewarding, I have gained what I wanted from my dr.ing. degree: A strong theoretical background.

For this I am in debt to professor Ph.D. Sigurd Skogestad for his guidance through these years. Sigurd had always time for discussions, and he has given good advises and valuable inputs. I would also like to thank Sigurd for “jule-grøtene” at Stokkanhaugen, and for the conferences that I have been allowed to visit.

I would also thank all of the members of the process control group here in Trondheim. It has been nice to work with you.

Finally I would like to thank my wife Åshild for supported. She was probably the person who most eagerly awaited the completion of my thesis. Together we have become parents of the most loveliest child ever: Johan Emil. To him I dedicate this thesis.

The Norwegian Research Council and the department of Chemical Engineering are acknowledged for financing.

Contents

Chapter 1

Introduction

1.1 Motivation

The behavior of a complete chemical processing plant is not only given by its individual units, the connections between the units are equally important. The behavior of a plant with the units connected in series, is easy to predict form the behavior of the individual units. This does not imply that the units can be operated like individual units: The output of one unit will act as a disturbance on the next unit, and at steady state they must have the same throughput. Even for a system with simple connection, certain considerations needs a perspective above the unit operation. A simple example is the placement of level controllers for a plant with units in series. It is exactly such a type of structural question that the field of plantwide control seeks to answer. Chapter 2 gives a more precise definition of plantwide control.

The presence of heat integration and mass recycle changes the dynamic and steady state behavior of the plant in ways which are difficult to predict from the behavior of the individual units. Therefor heat integration and mass recycle makes the need for a plantwide perspective much more pronounced when the control structure is designed.

The field of plantwide control is divided in two different approaches:

- A mathematically oriented approach.
- A process oriented approach.

The process oriented approach has proposed heuristics for plantwide design based upon case studies and their experience. This approach has two main drawbacks: The insight gained from a specific case study may be too narrow to make the conclusions general. Secondly, since the control objectives often are unclear the rules that are proposed has a weak basis.

In this thesis a number of cases are studied in a systematic manner. In this way we hope to better understand the issues that are involved in plantwide control. In particular we will question some of the heuristic rules, which we feel has a weak theoretical basis.

A better understanding of plantwide control will lead to a better design of control system. Better control systems will give plants with lower energy consumption and better utilization of raw material. This is important for both the society and the company.

1.2 Main contributions and thesis overview

This thesis contains seven chapters, they may be read independently. This is particular valid for Chapter 3 and 8. It is however recommended to read chapter 2 first. Chapter 4, 5 and 6 are strongly related and should be read together.

In Chapter 2 we present a large and comprehensive literature review of plantwide control. Based on this review we have presented a control structure design procedure. This chapter is based on an article submitted to Journal of Process Control.

Chapter 3 shows that the lower control layer may impose fundamental limitation if some information from the lower layer is unavailable to the higher layer. Chapter 3 was presented at the AIChE annual meeting 1998, Miami Beach.

In Chapter 4, 5 and 6 we deal with the control of processes with recycle. In the case studies we have been using systematic method for selection of control structure. We have shown that Luybens basis for his rule “fix one flow in recycle” is wrong. For the liquid phase plant it has bad self-optimizing properties. The heuristic “maximize recycle flow” by Fisher, is not correctly formulated. It was not economically optimal to maximize the recycle flow. The correct interpretation should be to avoid unnecessary reductions in recycle flow (open valves etc.).

For a chemical plant it is important to avoid the accumulation of chemical components, and inert may be particularly tricky. This has led many to believe that inert composition should always be controlled. This is not true, with a reasonable control structure the level of inert is normally self-regulating, and in our cases we have shown that it is a bad candidate for self-optimizing control.

Chapter 4 and 5 has been presented, in several versions, on NPCW-1998 Stockholm, CAPE Forum 99 Liége and AIChE annual meeting 1999, Dallas. A version of Chapter 6 is to be submitted to Ind. Eng. Chem. Res.

In Chapter 7 has looked on the control of an industrial heat integrated distillation column. For this case there where only one unconstrained degree of freedom at the optimum. The active constraints are, pressure in the low pressure column, heat transfer area between the column and both top compositions. Temperature on tray six has good self-optimizing properties. This chapter was presented at the AIChE annual meeting, Dallas November 1999.

In Chapter 8 we have shown how poles at the origin may be present in the relative gain array.

Chapter 2

Plantwide control -
A review and a new design procedure

Truls Larsson and Sigurd Skogestad

Based on a paper submitted to Journal of Process control.

Abstract

Most (if not all) available control theories assume that a control structure is given at the outset. They therefore fail to answer some basic questions that a control engineer regularly meets in practice (Foss, 1973): "Which variables should be controlled, which variables should be measured, which inputs should be manipulated, and which links should be made between them?" These are the questions that plantwide control tries to answer.

There are two main approaches to the problem, a mathematically oriented approach (control structure design) and a process oriented approach. Both approaches are reviewed in the paper. Emphasis is put on the selection of controlled variables, and it is shown that the idea of "self-optimizing control" provides a link between steady-state optimization and control.

We also provide some definitions of terms used within the area of plantwide control.

2.1 Introduction

A chemical plant may have thousands of measurements and control loops. By the term *plantwide control* it is *not* meant the tuning and behavior of each of these loops, but rather the *control philosophy* of the overall plant with emphasis on the *structural decisions*. The structural decision include the selection/placement of manipulators and measurements as well as the *decomposition* of the overall problem into smaller subproblems (the control configuration).

In practice, the control system is usually divided into several layers. Typically, layers include scheduling (weeks), site-wide optimization (day), local optimization (hour), supervisory/predictive control (minutes) and regulatory control (seconds), see Figure 2.1. The

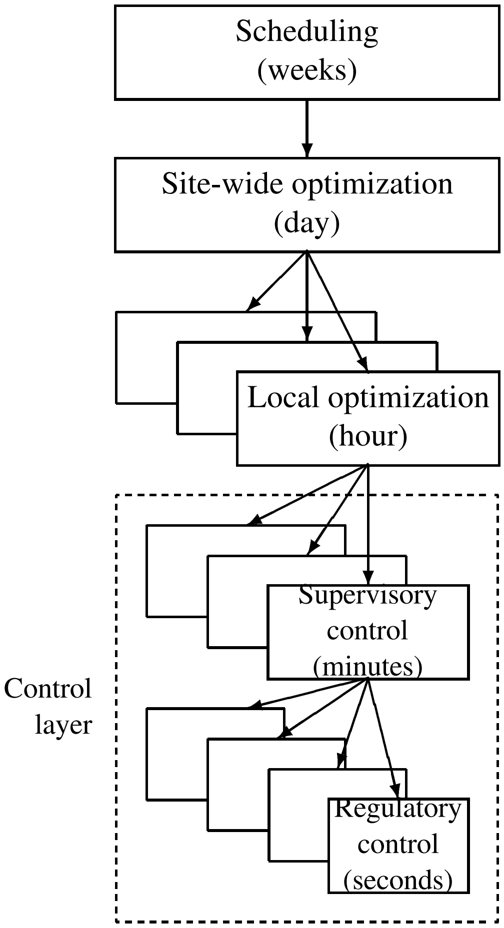


Figure 2.1: Typical control hierarchy in a chemical plant.

optimization layer typically computes new setpoints only once a hour or so, whereas the feedback layer operates continuously. The layers are linked by the controlled variables, whereby the setpoints are computed by the upper layer and implemented by the lower layer. An important issue is the selection of these variables.

Of course, we could imagine using a single optimizing controller who stabilizes the process while at the same time perfectly coordinates all the manipulated inputs based on dynamic on-line optimization. There are fundamental reasons why such a solution is not the best, even with today's and tomorrows computing power. One fundamental reason is the cost of modeling, and the fact that feedback control, without much need for models, is very effective when performed locally. In fact, by cascading feedback loops, it is possible to control large plants with thousands of variables without the actual need to develop any models. However, the traditional single-loop control systems can sometimes be rather complicated, especially if the cascades are heavily nested or if the presence of constraints during operation make it necessary to use logic switches. Thus, model based control should be used when the modeling effort gives enough pay-back in terms of simplicity and/or improved performance, and this will usually be at the higher layers in the control hierarchy.

A very important (if not the most important) problem in plantwide control is the issue of determining the *control structure*:

- Which "boxes" should we have and what information should be send between them?

Note that that we are here not interested in what should be inside the boxes, which is the controller design or tuning problem. More precisely, **control structure design** is defined as the *structural decisions* involved in control system design, including the following tasks ((Foss, 1973); (Morari, 1982); (Skogestad and Postlethwaite, 1996))

1. *Selection of controlled outputs* c (variables with setpoints)
2. *Selection of manipulated inputs* m
3. *Selection of measurements* v (for control purposes including stabilization)
4. *Selection of control* **configuration** (a structure interconnecting measurements/setpoints and manipulated variables, i.e. the structure of the controller K which interconnects the variables c_s and v (controller inputs) with the variables m)
5. *Selection of controller type* (control law specification, e.g., PID, decoupler, LQG, etc.).

In most cases the control structure design is solved by a mixture of a top-down consideration of control objectives and which degrees of freedom are available to meet these (tasks 1 and 2), and a with a bottom-up design of the control system, starting with the stabilization of the process (tasks 3,4 and 5).

In most cases the problem is solved without the use of any theoretical tools. In fact, the industrial approach to plantwide control is still very much along the lines described by Page Buckley in his book from 1964. Of course, the control field has made many advances over these years, for example, in methods for and applications of on-line optimization and predictive control. Advances have also been made in control theory and in the formulation of tools for analyzing the controllability of a plant. These latter tools can be most helpful in screening alternative control structures. However, a systematic method for generating promising alternative structures has been lacking. This is related to the fact that plantwide control problem itself has not been well understood or even acknowledged as important.

The control structure design problem is difficult to define mathematically, both because of the size of the problem, and the large cost involved in making a precise problem definition, which would include, for example, a detailed dynamic and steady state model. An alternative to this is to develop heuristic rules based on experience and process understanding. This is what will be referred to as the **process oriented approach**.

The realization that the field of control structure design is underdeveloped is not new. In the 1970's several "critique" articles where written on the gap between theory and practice in the area of process control. The most famous is the one of (Foss, 1973) who made the observation that in many areas application was *ahead* of theory, and he stated that

> The central issues to be resolved by the new theories are the determination of the control system structure. Which variables should be measured, which inputs should be manipulated and which links should be made between the two sets? ... The gap is present indeed, but contrary to the views of many, it is the theoretician who must close it.

A similar observation that applications seem to be ahead of formal theory was made by Findeisen *et al.* (1980) in their book on hierarchical systems (p. 10).

Many authors point out that the need for a plantwide perspective on control is mainly due to changes in the way plants are designed – with more heat integration and recycle and less inventory. Indeed, these factors lead to more interactions and therefore the need for a perspective beyond individual units. However, we would like to point out that even without any integration there is still a need for a plantwide perspective as a chemical plant consists of a string of units connected in series, and one unit will act as a disturbance to the next, for example, all units must have the same throughput at steady-state.

Outline

We will first discuss in more detail some of the terms used above and provide some definitions. We then present a review of some of the work on plantwide control. In section 2.4 we discuss the mathematically oriented approach (control structure design). Then, in section 2.5 we look at the process oriented approach. In section 2.6 we consider a fairly simple plant consisting of reactor, separator and recycle. In section 2.7 we consider the most studied plantwide control problem, namely the Tennessee Eastman problem introduced by Downs and Vogel (1993), and we discuss how various authors have attempted to solve the problem. Finally, in section 2.8 we propose a new plantwide control design procedure.

2.2 Terms and definitions

We here make some comments on the terms introduced above, and also attempt to provide some more precise definitions, of these terms and some additional ones.

Let us first consider the terms *plant* and *process*, which are almost synonymous terms. In the control community as a whole, the term plant is somewhat more general than process: A *process* usually refers to the "process itself" (without any control system) whereas a *plant* may be *any system to be controlled* (including a partially controlled process). However, note that in the chemical engineering community the term plant has a somewhat different meaning, namely as the whole factory, which consists of many process units; the term *plantwide* control is derived from this meaning of the word plant.

Let us then discuss the two closely related terms *layer* and *level* which are used in hierarchical control. Following the literature e.g. Findeisen et al. (1980) the correct term in our context is *layer*. In a layer the parts acts at different time scales and each layer has some feedback or information from the process and follows setpoints given from layers above. A lower layer may not know the criterion of optimality by which the setpoint has been set. A *multi-layer* system cannot be strictly optimal because the actions of the higher layers are discrete and thus unable to follow strictly the optimal continuous time pattern. (On the other hand, in a *multilevel* system there is no time scale separation and the parts are coordinated such that there are no performance loss. Multilevel decomposition may be used in the optimization algorithm but otherwise is of no interest here.)

Control is the adjustment of available degrees of freedom (manipulated variables) to assist in achieving acceptable operation of the plant. Control system design may be divided into three main activities

1. Control structure design (structural decisions)
2. Controller design (parametric decisions)
3. Implementation

The term **control structure design**, which is commonly used in the control community, refers to the structural decisions in the design of the control system. It is defined by the five tasks (given in the introduction):

1. *Selection of controlled outputs* (c with setpoints c_s).
2. *Selection of manipulated inputs* (m).
3. *Selection of measurements* (v)
4. *Selection of control* **configuration**
5. *Selection of controller type*

The result from the *control structure design* is the *control structure* (alternatively denoted the *control strategy* or *control philosophy* of the plant).

The term *plantwide control* is used only in the process control community. One could regard *plantwide control* as the “process control” version of control structure design, but this is probably a bit too limiting. In fact, Rinard and Downs (1992) refer to the control structure design problem as defined above as the “strict definition of plantwide control”, and they point out that plantwide control also include important issues such as the operator interaction, startup, grade-change, shut-down, fault detection, performance monitoring and design of safety and interlock systems. This is more in line with the discussion by Stephanopoulos, (1982).

Maybe a better distinction is the following: *Plantwide control* refers to the structural and strategic decisions involved in the control system design of a complete chemical plant (factory), and *control structure design* is the systematic (mathematical) approach for solving this problem.

The control configuration, is defined as the restrictions imposed by the overall controller K by decomposing it into a set of local controllers (sub-controllers), units, elements, blocks) with predetermined links and possibly with a predetermined design sequence where sub-controllers are designed locally.

Operation involves the behavior of the system once it has been build, and this includes a lot more than control. More precisely, the control system is designed to aid the operation of the plant. *Operability* is the ability of the plant (together with its control system) to achieve acceptable operation (both statically and dynamically). Operability includes switchability and controllability as well as many other issues.

Flexibility refers to the ability to obtain feasible steady-state operation at a given set of operating points. This is a steady-state issue, and we will assume it to be satisfied at the operating points we consider. It is not considered any further in this paper.

Switchability refers to the ability to go from one operating point to another in an acceptable manner usually with emphasis on feasibility. It is related to other terms such as optimal operation and controllability for large changes, and is not considered explicitly in this paper.

We will assume that the "quality (goodness) of operation" can be quantified in terms of a scalar performance index (objective function) J, which should be minimized. For example, J can be the operating costs.

Optimal operation usually refers to the *nominally* optimal way of operating a plant as it would result by applying steady-state and/or dynamic optimization to a model of the plant (with no uncertainty), attempting to minimize the cost J by adjusting the degrees of freedom.

In practice, we cannot obtain optimal operation due to uncertainty. The difference between the actual value of the objective function J and its nominally optimal value is the *loss*.

The two main sources of *uncertainty* are (1) *signal uncertainty* (includes disturbances (d) and measurement noise (n)) and (2) *model uncertainty*.

Robust means insensitive to uncertainty. *Robust optimal operation* is the optimal way of operating a plant (with uncertainty considerations included).

Integrated optimization and control (or optimizing control) refers to a system where optimization and its control implementation are integrated. In theory, it should be possible to obtain robust optimal operation with such a system. In practice, one often uses a *hierarchical decomposition* with separate layers for optimization and control. In making this split we assume that for the control system the goal of "acceptable operation" has been translated into "keeping the controlled variables (c) within specified bounds from their setpoints (c_s)". The optimization layer sends setpoint values (c_s) for selected controlled outputs (c) to the control layer. The setpoints are updated only periodically. (The tasks, or parts of the tasks, in either of these layers may be performed by humans.) The control layer may be further divided, e.g. into supervisory control and regulatory control. In general, in a hierarchical system, the lower layers work on a shorter time scale.

In addition to keeping the controlled variables at their setpoints, the control system must *"stabilize"* the plant. We have here put stabilize in quotes because we use the word in an extended meaning, and include both modes which are mathematically unstable as well as slow modes ("drift") that need to be "stabilized" from an operator point of view. Usually, stabilization is done within a separate (lower) layer of the control system, often called the regulatory control layer. The controlled outputs for stabilization are measured output variables, and their setpoints may be used as degrees of freedom for the layers above.

For each layer in a control system we use the terms *controlled output* (y with setpoint y_s) and *manipulated input* (u). Correspondingly, the term "plant" refers to the system to be controlled (with manipulated inputs u and controlled outputs y). The layers are often structured hierarchically, such that the manipulated input for a higher layer (u_1) is the setpoint for a lower layer (y_{2s}), i.e. $y_{2s} = u_1$. (These controlled outputs need in general not be measured variables, and they may include some of the manipulated inputs (u).)

From this we see that the terms "plant", "controlled output" (y) and "manipulated input" (u) takes on different meaning depending on where we are in the hierarchy. To avoid confusion, we reserve special symbols for the variables at the top and bottom of the hierarchy. Thus, as already mentioned, the term *process* is often used to denote the uncontrolled plant as seen from the bottom of the hierarchy. Here the manipulated inputs are the physical manipulators (e.g. valve positions), and are denoted m. Correspondingly, at the top of the hierarchy, we use the symbol c to denote the controlled variables for which the setpoint values (c_s) are determined by the optimization layer.

Input-Output controllability of a plant is the ability to achieve acceptable control performance, that is, to keep the controlled outputs (y) within specified bounds from their setpoints (r), in spite of signal uncertainty (disturbances d, noise n) and model uncertainty, using available inputs (u) and available measurements. In other words, the plant is controllable if there is a controller that satisfies the control objectives.

This definition of controllability may be applied to the control system as a whole, or to parts of it (in the case the control layer is structured). The term controllability generally assumes that we use the best possible multivariable controller, but we may impose restrictions on the class of allowed controllers (e.g. consider "controllability with decentralized PI control").

A plant is *self-regulating* if we with constant inputs can keep the controlled outputs within acceptable bounds. (Note that this definition may be applied to any layer in the control system, so the plant may be a partially controlled process). "True" self-regulation is defined as the case where no control is ever needed at the lowest layer (i.e. m is constant). It relies on the process to dampen the disturbances itself, e.g. by having large buffer tanks. We rarely have "true" self-regulation because it may be very costly.

Self-optimizing control is when an acceptable loss can be achieved using constant setpoints for the controlled variables (without the need to reoptimize when disturbances occur). "True" self-optimization is defined as the case where no re-optimization is ever needed (so c_s can be kept constant always), but this objective is usually not satisfied. On the other hand, we must require that the process is self-optimizing within the time period between each re-optimization, or else we cannot use separate control and optimization layers.

A process is self-optimizing if there exists a set of controlled outputs (c) such that if we with keep constant setpoints for the optimized variables (c_s), then we can keep the loss within an acceptable bound within a specified time period. A steady-state analysis is usually sufficient to analyze if we have self-optimality. This is based on the assumption that the closed-loop time constant of the control system is smaller than the time period between each re-optimization (so that it settles to a new steady-state) and that the value of the objective function (J) is mostly determined by the steady-state behavior (i.e. there is no "costly" dynamic behavior e.g. imposed by poor control).

Most of the terms given above are in standard use and the definitions mostly follow those of Skogestad and Postlethwaite (1996).

2.3 General reviews and books on plantwide control

We here present a brief review of some of the previous reviews and books on plantwide control.

Morari (1982) presents a well-written review on plantwide control, where he discusses why modern control techniques were not (at that time) in widespread use in the process industry. The four main reasons were believed to be

1. Large scale system aspects.
2. Sensitivity (robustness).
3. Fundamental limitations to control quality.
4. Education.

He then proceeds to look at how two ways of decompose the problem:

1. Multi-layer (vertical), where the difference between the layers are in the frequency of adjustment of the input.
2. Horizontal decomposition, where the system is divided into noninteracting parts.

Stephanopoulos (1982) states that the synthesis of a control system for a chemical plant is still to a large extent an art. He asks: "Which variables should be measured in order to monitor completely the operation of a plant? Which input should be manipulated for effective control? How should measurements be paired with the manipulations to form the control structure, and finally, what the control laws are?" He notes that the problem of plantwide control is "multi-objective" and "there is a need for a systematic and organized approach which will identify all necessary control objectives". The article is comprehensive, and discusses many of the problems in the synthesis of control systems for chemical plants.

Rinard and Downs (1992) review much of the relevant work in the area of plantwide control, and they also refer to important papers that we have not referred. They conclude the review by stating that "the problem probably never will be solved in the sense that a set of algorithms will lead to the complete design of a plantwide control system". They suggests that more work should be done on the following items: (1) A way of answering whether or not the control system will meet all the objectives, (2) Sensor selection and location (where they indicate that theory on partial control may be useful), (3) Processes with recycle. They also welcome computer-aided tools, better education and good new test problems.

The book by Balchen and Mummé (1988) attempts to combine process and control knowledge, and to use this to design control systems for some common unit operations and also consider plantwide control. The book provides many practical examples, but there is little in terms of analysis tools or a systematic framework for plantwide control.

The book "Integrated process control and automation" by Rijnsdorp (1991), contains several subjects that are relevant here. Part II in the book is on optimal operation. He distinguishes between two situations, sellers marked (maximize production) and buyers marked (produce a given amount at lowest possible cost). He also has a procedure for design of an optimizing control system.

Loe (1994) presents a systematic way of looking at plants with the focus is on functions. The author covers "qualitative" dynamics and control of important unit operations.

van de Wal and de Jager (1995) list several criteria for evaluation of control structure design methods: generality, applicable to nonlinear control systems, controller-independent, direct, quantitative, efficient, effective, simple and theoretically well developed. After reviewing they conclude that such a method does not exist.

The book by Skogestad and Postlethwaite (1996) has two chapters on controllability analysis and one chapter on control structure design. Particularly in chapter 10 there is some topics, which are relevant for plantwide control, among them are partial control and self-optimizing control (a term introduced later).

The coming monograph by Ng and Stephanopoulos (1998*a*) deals almost exclusively with plantwide control.

The book by Luyben *et al.* (1998) has collected much of Luyben's practical ideas and summarized them in a clear manner. The emphasis is on case studies.

There also exists a large body of system-theoretic literature within the field of large-scale systems, but most of it has little relevance to plantwide control. One important exception is the book by Findeisen et al. (1980) on "Control and coordination in hierarchical systems" which probably deserves to be studied more carefully by the process control community.

2.4 Control Structure Design (The mathematically oriented approach)

In this section we look at the mathematically oriented approach to plantwide control.

Structural methods

There are some methods that use structural information about the plant as a basis for control structure design. For a recent review of these methods we refer to the coming monograph of Ng and Stephanopoulos (1998*a*). Central concepts are structural state controllability, observability and accessibility. Based on this, sets of inputs and measurements are classified as viable or non-viable. Although the structural methods are interesting, they are not quantitative and usually provide little information other than confirming insights about the structure of the process that most engineers already have.

In the reminder of this section we discuss the five tasks of the control structure design problem, listed in the introduction.

2.4.1 Selection of controlled outputs (c)

By "controlled outputs" we here refer to the controlled variables c for which the setpoints c_s are determined by the optimization layer. There will also be other (internally) controlled outputs which result from the decomposition of the controller into blocks or layers (including controlled measurements used for stabilization), but these are related to the control configuration selection, which is discussed as part of task 4.

The issue of selection of controlled outputs, is probably the least studied of the tasks in the control structure design problem. In fact, it seems from our experience that most people do not consider it as being an issue at all. The most important reason for this is probably that it is a structural decision for which there has not been much theory. Therefore the decision has mostly been based on engineering insight and experience, and the validity of the selection of controlled outputs has seldom been questioned by the control theoretician.

To see that the selection of output is an issue, ask the question:

> *Why are we controlling hundreds of temperatures, pressures and compositions in a chemical plant, when there is no specification on most of these variables?*

After some thought, one realizes that the main reason for controlling all these variables is that one needs to specify the available degrees of freedom in order to keep the plant close to its optimal operating point. But there is a follow-up question:

> *Why do we select a particular set c of controlled variables? (e.g., why specify (control) the top composition in a distillation column, which does not produce final products, rather than just specifying its reflux?)*

The answer to this second question is less obvious, because at first it seems like it does not really matter which variables we specify (as long as all degrees of freedom are consumed, because the remaining variables are then uniquely determined). However, this is true only when there is no uncertainty caused by disturbances and noise (signal uncertainty) or model uncertainty. When there is uncertainty then it does make a difference how the solution is implemented, that is, which variables we select to control at their setpoints.

Self-optimizing control

The basic idea of what we have called self-optimizing control was formulated about twenty years ago by Morari *et al.* (1980):

> "in attempting to synthesize a feedback optimizing control structure, our main objective is to translate the economic objectives into process control objectives. In other words, *we want to find a function c of the process variables which when*

> *held constant, leads automatically to the optimal adjustments of the manipulated variables, and with it, the optimal operating conditions.* [...] This means that by keeping the function $c(u, d)$ at the setpoint c_s, through the use of the manipulated variables u, for various disturbances d, it follows uniquely that the process is operating at the optimal steady-state."

If we replace the term "optimal adjustments" by "acceptable adjustments (in terms of the loss)" then the above is a precise description of what Skogestad (2000) denote a self-optimizing control structure. The only factor Morari *et al.* (1980) fail to consider is the effect of the implementation error $c - c_s$. Morari *et al.* (1980) propose to select the best set of controlled variables based on minimizing the loss ("feedback optimizing control criterion 1").

Somewhat surprisingly, the ideas of Morari *et al.* (1980) received very little attention. One reason is probably that the paper also dealt with the issue of finding the optimal operation (and not only on how to implement it), and another reason is that the only example in the paper happened to result in an implementation with the controlled variables at their constraints. The constrained case is "easy" from an implementation point of view, because the simplest and optimal implementation is to simply maintain the constrained variables at their constraints. No example was given for the more difficult unconstrained case, where the choice of controlled (feedback) variables is a critical issue. The follow-up paper by Arkun and Stephanopoulos (1980) concentrated further on the constrained case and tracking of active constraints.

Skogestad and Postlethwaite (1996) (Chapter 10.3) present an approach for selecting controlled output similar to those of Morari *et al.* (1980) and the ideas where further developed in (Skogestad, 2000) where the term self-optimizing control is introduced. Skogestad (2000) stresses the need to consider the implementation error when evaluating the loss. Skogestad (2000) gives four requirements that a controlled variable should meet: 1) Its optimal value should be insensitive to disturbances. 2) It should be easy to measure and control accurately. 3) Its value should be sensitive to changes in the manipulated variables. 4) For cases with two or more controlled variables, the selected variables should not be closely correlated. By scaling of the variables properly, Skogestad and Postlethwaite (1996) shows that the self-optimizing control structure is related to maximizing the minimum singular value of the gain matrix G, where $\Delta c = \Delta u$. Zheng *et al.* (1999) also use the ideas of Morari *et al.* (1980) as a basis for selecting controlled variables. The relationship to the work of Shinnar is discussed separately later.

Other work

In his book Rijnsdorp (1991) gives on page 99 a stepwise design procedure for designing optimizing control systems for process units. One step is to "transfer the result into on-line algorithms for adjusting the degrees of freedom for optimization". He states that this "requires good process insight and control structure know-how. It is worthwhile basing the algorithm as far as possible on process measurements. In any case, it is impossible to give a clear-cut recipe here."

Fisher *et al.* (1988*a*) discuss plant economics in relation to control. They provide some interesting heuristic ideas. In particular, hidden in their HDA example in part 3 (p. 614) one finds an interesting discussion on the selection of controlled variables, which is quite closely related to the ideas of Morari *et al.* (1980).

Luyben (1988) introduced the term "eigenstructure" to describe the inherently best control structure (with the best self-regulating and self-optimizing property). However, he did not really define the term, and also the name is unfortunate since "eigenstructure" has a another unrelated mathematical meaning in terms of eigenvalues. Apart from this, Luyben and coworkers (e.g. Luyben (1975), Yi and Luyben (1995)) have studied unconstrained problems, and some of the examples presented point in the direction of the selection methods presented in this paper. However, Luyben proposes to select controlled outputs which minimizes the steady-state sensitive of the manipulated variable (u) to disturbances, i.e. to select controlled outputs (c) such that $(\partial u/\partial d)_c$ is small, whereas we really want to minimize the steady-state sensitivity of the economic loss (L) to disturbances, i.e. to select controlled outputs (c) such that $(\partial L/\partial d)_c$ is small.

Narraway *et al.* (1991), Narraway and Perkins (1993) and Narraway and Perkins (1994)) strongly stress the need to base the selection of the control structure on economics, and they discuss the effect of disturbances on the economics. However, they do not formulate any rules or procedures for selecting controlled variables.

Finally Mizoguchi *et al.* (1995) and Marlin and Hrymak (1997) stress the need to find a good way of implementing the optimal solution in terms how the control system should respond to disturbances, "i.e. the key constraints to remain active, variables to be maximized or minimized, priority for adjusting manipulated variables, and so forth." They suggest that an issue for improvement in today's real-time optimization systems is to select the control system that yields the highest profit for a range of disturbances that occur between each execution of the optimization.

There has also been done some work on non-square plants, i.e. with more outputs than inputs, e.g. (Cao, 1995) and (Chang and Yu, 1990). These works assumes that the control goal is the keep all these variables as close to "zero" as possible, and often the effect of disturbances is not considered. It may be more suitable to reformulate these problems into the framework of self-optimizing control.

2.4.2 Selection of manipulated inputs (m)

By manipulated inputs we refer to the physical degrees of freedom, typically the valve positions or electric power inputs. Actually, selection of these variables is usually not much of an issue at the stage of control structure design, since these variables usually follow as direct consequence of the design of the process itself.

However, there may be some possibility of adding valves or moving them. For example, if we install a bypass pipeline and a valve, then we may use the bypass flow as an extra degree of freedom for control purposes.

Finally, let us make it clear that the possibility of not actively using some manipulated inputs (or only changing them rarely), is a decision that is included above in "selection of controlled outputs".

2.4.3 Selection of measurements (v)

Controllability considerations, including dynamic behavior, are important when selecting which variables to measure. There are often many possible measurements we can make, and the number, location and accuracy of the measurement is a tradeoff between cost of measurements and benefits of improved control. A controllability analysis may be very useful. In most cases the selection of measurements must be considered simultaneously with the selection of the control configuration. For example, this applies to the issue of stabilization and the use of secondary measurements.

2.4.4 Selection of control configuration

The issue of control configuration selection, including decentralized control, is discussed in Hovd and Skogestad (1993) and in sections 10.6, 10.7 and 10.8 of Skogestad and Postlethwaite (1996), and we will here discuss mainly issues which are not covered there.

The control configuration is the structure of the controller K that interconnects the measurements, setpoints c_s and manipulated variables m. The controller can be structured (decomposed) into blocks both in an vertical (hierarchical) and horizontal (decentralized control) manner.

Why is the controller decomposed? (1) The first reason is that it may require less computation. This reason may be relevant in some decision-making systems where there is limited capacity for transmitting and handling information (like in most systems where humans are involved), but it does not hold in today's chemical plant where information is centralized and computing power is abundant. Two other reasons often given are (2) failure tolerance and (3) the ability of local units to act quickly to reject disturbances (e.g. Findeisen et al., 1980). These reasons may be more relevant, but, as pointed out by Skogestad and Hovd (1995) there are probably other even more fundamental reasons. The most important one is probably (4) to reduce the cost involved in defining the control problem and setting up the detailed dynamic model which is required in a centralized system with no predetermined links. Also, (5) decomposed control systems are much less sensitive to model uncertainty (since they often use no explicit model). In other words, by imposing a certain control configuration, we are implicitly providing process information, which we with a centralized controller would need to supply explicitly through the model.

Stabilizing control

Instability requires the active use of manipulated inputs (m) using feedback control. There exist relatively few systematic tools to assist in selecting a control structure for stabilizing control. Usually, single-loop controllers are used for stabilization, and issues are which variables to measure and which manipulated inputs to use. One problem in stabilization is that measurement noise may cause large variations in the input such that it saturates. Havre and Skogestad (1996, 1998) have shown that the *pole vectors* may be used to select measurements and manipulated inputs such that this problem is minimized.

Secondary measurements

Extra (secondary) measurements are often added to improve the control. Three alternatives for use of extra measurements are:

1. Centralized controller: All the measurements are used to compute the optimal input. This controller has implicitly an estimator (model) hidden inside it.

2. Inferential control: Based on the measurements a model is used to provide an estimate of the primary output (e.g. a controlled output c). This estimate is send to a separate controller.

3. Cascade control: The secondary measurements are controlled locally and their setpoints are used as degrees of freedom at some higher layer in the hierarchy.

Note that both centralized and inferential control uses the extra measurements to estimate parameters in a model, whereas cascaded control they are used for additional feedback. The subject of estimation and measurements selection for estimation is beyond the scope of this review article; we refer to Ljung (1987) for a control view and to Martens (1989) for a chemometrics approach to this issue. However, we would like point out that the control system should be designed for best possible control of the primary variables (c), and not the best possible estimate. A drawback of the inferential scheme is that estimate is used in feed-forward manner.

For cascade control Havre (1998) has shown how to select secondary measurements such that the need for updating the setpoints is small. The issues here are similar to that of selecting controlled variables (c) discussed above. One approach is to minimize some norm of the transfer function from the disturbance and control error in the secondary variable to the control error in the primary variable. A simpler, but less accurate, alternative is to maximize the minimum singular value in the transfer function from secondary measurements to the input used to control the secondary measurements. Lee and Morari ((Lee and Morari, 1991), (Lee *et al.*, 1995) and (Lee *et al.*, 1997)) considered a similar problem. They used a more rigorous approach where model uncertainty is explicitly considered and the structured singular value is used as a tool.

Partial control

Most control configurations are structured in a hierarchical manner with fast inner loops, and slower outer loops that adjust the setpoints for the inner loops. Control system design generally starts by designing the inner (fast) loops, and then outer loops are closed in a sequential manner. Thus, the design of an "outer loop" is done on a *partially controlled system.* We here provide some simple but yet very useful relationships for partially controlled systems. We divide the outputs into two classes:

- y_1 – (temporarily) uncontrolled output

- y_2 – (locally) measured and controlled output (in the inner loop)

We have inserted the word *temporarily* above, since y_1 is normally a controlled output at some higher layer in the hierarchy. We also subdivide the available manipulated inputs in a similar manner:

- u_2 – inputs used for controlling y_2 (in the inner loop)
- u_1 – remaining inputs (which *may* be used for controlling y_1)

A block diagram of the partially controlled system resulting from closing the loop involving u_2 and y_2 with the local controller K_2 is shown in Figure 2.2.

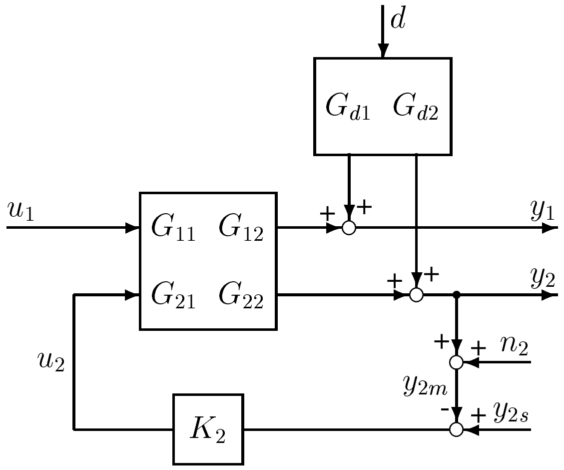


Figure 2.2: Block diagram of a partially controlled plant

Skogestad and Postlethwaite (1996) distinguish between the following four cases of partial control:

		Meas./Control of y_1 ?	Control objective for y_2 ?
I	Indirect control	No	No
II	Sequential cascade control	Yes	No
III	"True" partial control	No	Yes
IV	Sequential decentralized control	Yes	Yes

In all cases there is a control objective associated with y_1 and a measurement of y_2. For example, for indirect control there is no separate control objective on y_2, the reasons we control y_2 is to indirectly achieve good control of y_1 which are not controlled. The first two cases are probably the most important as they are related to vertical (hierarchical) structuring. The latter two cases (where y_2 has its own control objective so that the setpoints y_{2s} are not adjustable) gives a horizontal structuring.

In any case, the linear model for the plant can be written

$$y_1 = G_{11}(s)u_1 + G_{12}(s)u_2 + G_{d1}(s)d \tag{2.1}$$

$$y_2 = G_{21}(s)u_1 + G_{22}(s)u_2 + G_{d2}(s)d \tag{2.2}$$

To derive transfer functions for the partially controlled system we simply solve (2.2) with respect to u_2 (assuming that $G_{22}(s)$ is square an invertible at a given value of s)[1]

$$u_2 = G_{22}^{-1}(s)\,(y_2 - G_{21}(s)u_1 - G_{d2}(s)d) \tag{2.3}$$

Substituting (2.3) into (2.1) then yields

$$\boxed{y_1 = P_u(s)u_1 + P_d(s)d + P_y(s)y_2} \tag{2.4}$$

where

$$P_u(s) \stackrel{\text{def}}{=} G_{11}(s) - G_{12}G_{22}^{-1}G_{21}(s) \tag{2.5}$$
$$P_d(s) \stackrel{\text{def}}{=} G_{d1}(s) - G_{12}G_{22}^{-1}G_{d2}(s) \tag{2.6}$$
$$P_y(s) \stackrel{\text{def}}{=} G_{12}G_{22}^{-1}(s) \tag{2.7}$$

Here P_d is the *partial disturbance gain*, P_y is the gain from y_2 to y_1, and P_u is the partial input gain from the unused inputs u_1. If we look more carefully at (2.4) then we see that the matrix P_d gives the effect of disturbances on the primary outputs y_1, when the manipulated inputs u_2 are adjusted to keep y_2 constant, which is consistent of the original definition of the partial disturbance gain given by Skogestad and Wolff (1992). Note that no approximation about perfect control has been made when deriving (2.4). Equation (2.4) applies for any fixed value of s (on a frequency-by-frequency basis).

The above equations are simple yet very useful. Relationships containing parts of these expressions have been derived by many authors, e.g. see the work of Manousiouthakis *et al.* (1986) on block relative gains and the work of Häggblom and Waller (1988) on distillation control configurations.

Note that this kind of analysis can be performed at each layer in the control system. At the top layer we may assume that the cost J is a function of the variables y_1, and we can then interpret y_2 as the set of controlled outputs c. If c is never adjusted then this is a special case of indirect control, and if c is adjusted at regular intervals (as is usually done) then this may be viewed as a special case of sequential cascade control.

2.5 The Process Oriented Approach

We here review procedures for plantwide control that are based on using process insight, that is, methods that are unique to process control.

The first comprehensive discussion on plantwide control was given by Page Buckley in his book "Techniques of process control" in a chapter on *Overall process control* (Buckley, 1964). The chapter introduces the main issues, and presents what is still in many ways the industrial approach to plantwide control. In fact, when reading this chapter, 35 years later one is struck with the feeling that there has been relatively little development in this area. Some of the terms which are introduced and discussed in the chapter are material balance control

[1] The assumption that G_{22}^{-1} exists for all values of s can be relaxed by replacing the inverse with the pseudo-inverse.

(in direction of flow, and in direction opposite of flow), production rate control, buffer tanks as low-pass filters, indirect control, and predictive optimization. He also discusses recycle and the need to purge impurities, and he points out that you cannot at a given point in a plant control inventory (level, pressure) and flow independently since they are related through the material balance. In summary, he presents a number of useful engineering insights, but there is really no overall procedure. As pointed out by Ogunnaike (1995) the basic principles applied by the industry does not deviate far from Buckley (1964).

Wolff and Skogestad (1994) review previous work on plantwide control with emphasis on the process-oriented decomposition approaches. They suggest that plantwide control system design should start with a "top-down" selection of controlled and manipulated variables, and proceed with a "bottom-up" design of the control system. At the end of the paper ten heuristic guidelines for plantwide control are listed.

There exists other more or less heuristics rules for process control; e.g. see Hougen and Brockmeier (1969) and Seborg *et al.* (1995).

2.5.1 Degrees of freedom for control and optimization

A starting point for plantwide control is to establish the number of degrees of freedom for operation; both dynamically (for control, N_m) and at steady-state (for optimization, N_{ss}). These are defined as

N_m Degrees of freedom for control: The number of variables (temperatures, pressures, levels etc.) that may be set by the control system.

N_{ss} Degrees of freedom for steady state optimization optimization: The number of independent variables with a steady state effect.

Many authors suggest to use a process model to find the degrees of freedom. However this approach will be error prone, it is easy to write too many or too few equations. Fortunately, it is in most cases relatively straightforward to establish these numbers from process insight.

Ponton (1994) propose a method for finding N_{ss} by counting the number of streams and subtracting the number of "extra" phases (i.e. if there are more than one phase present in that unit). It is easy to construct really easy examples where the method fails. Consider a simple liquid storage tank with one inflow and one outflow. According to the above, we would have $N_{ss} = 2$, which is clearly wrong. Maybe we should have subtracted the vapor phase which probably exists above the liquid. This gives $N_{ss} = 2-1 = 1$, which gives the correct answer. However, if we add a reaction in the tank, then conversion depends on the holdup in the tank and N_{ss} should be equal to 2. A better approach is needed.

It is well known that N_m equals the number of number of adjustable valves plus the number of other adjustable electrical and mechanical variables (electric power, etc.). According to (Skogestad, 2000) the number of degrees at freedom at steady-state (N_{ss}) can be found by subtracting the number of variables with no steady state effects. These variables are

N_{m0} is the number of manipulated inputs (u's), or combinations thereof, with no steady-state effect.

N_{y0} is the number of manipulated inputs that are used to control variables with no steady-state effect.

The latter usually equals the number of liquid levels with no steady-state effect, including most buffer tank levels. However, note that some liquid levels *do* have a steady-state effect, such as the level in a non-equilibrium liquid phase reactor, and levels associated with adjustable heat transfer areas. Also, we should *not* include in N_{y0} any liquid holdups that are left uncontrolled, such as internal stage holdups in distillation columns.

We find that N_{y0} is nonzero for most chemical processes, whereas we often have $N_{m0} = 0$. A simple example where N_{m0} is non-zero is a heat exchanger with bypass on both sides, (i.e. $N_m = 2$). However, at steady-state $N_{ss} = 1$ since there is really only one operational degree of freedom, namely the heat transfer rate Q (which at steady-state may be achieved by many combinations of the two bypasses), so we have $N_{m0} = 1$.

The optimization is generally subject to several constraints. First, there are generally upper and lower limits on all manipulated variables (e.g. fully open or closed valve). In addition, there are constraints on many dependent variables; due to safety (e.g. maximum pressure or temperature), equipment limitations (maximum throughput) or product specifications. Some of these constraints will be active at the optimum. The number of "free" unconstrained variables "for steady-state optimization", $N_{ss,free}$, is then equal to

$$N_{ss,free} = N_{ss} - N_{active}$$

where N_{active} is the number of active constraints. Note that the term "left for optimization" may be somewhat misleading, since the decision to keep some constraints active, really follows as part of the optimization; thus all N_{ss} variables are really used for optimization.

Remark on design degrees of freedom. Above we have discussed operational degrees of freedom. The design degree of freedom (which is not really a concern of this paper) includes all the N_{ss} operational degrees of freedom plus all parameters related to the size of the equipment, such as the number of stages in column sections, area of heat exchangers, etc.

Luyben (1996) claims that "design degrees of freedom is equal to the number of control degrees for an important class of processes." This is clearly not true, as there is no general relationship between the two numbers. For example, consider a heat exchanger between two streams. Then there may be zero, one or two control degrees of freedom (depending on the number of bypasses), but there is always one design degree of freedom (heat exchanger area).

2.5.2 Production rate

Identifying the major disturbances is very important in any control problem, and for process control the production rate (throughput) is often the main disturbance. In addition, the location of where the production rate is actually set ("throughput manipulator"), usually determines the control structure for the inventory control of the various units. For a plant running at maximum capacity, the location where the production rate is set is usually somewhere inside the plant, (e.g. caused by maximum capacity of a heat exchanger or a compressor). Then, downstream of this location the plant has to process whatever comes in (given feed rate), and upstream of this location the plant has to produce the desired quantity (given product rate). To avoid any "long loops", it is preferably to use the input flow for inventory control upstream the location where the production rate is set, and to use the output flow for inventory control downstream this location.

From this it follows that it is critical to know where in the plant the production rate is set. In practice, the location may vary depending on operating conditions. This may require reconfiguring of many control loops, but often supervisory control systems, such as model predictive control, provide a simpler and better solution.

2.5.3 The framework of partial control and dominating variables

Shinnar (1981) introduced the following sets of variables

- Y_p (the "primary" or "performance" or "economic" variables) is "the set of process variables that define the product and process specifications as well as process constraints"
- Y_d is the set of dynamically measured process variables
- Y_{cd} (a subset of Y_d) is the "set of process variables on which we base our dynamic control strategy"
- U_d is the dynamic input variables

The goal is to maintain Y_p within prescribed limits and to achieve this goal "we choose in most cases a small set Y_{cd} and try to keep these at a fixed set of values by manipulating U_d" (later, in Arbel *et al.* (1996), he introduced the term "partial control" to describe this idea).

He writes that the overall control algorithm can normally be decomposed into a dynamic control system (which adjust U_d) and a steady-state control which determines the set points of Y_{cd} as well as the values of U_s [the latter are the manipulations which only can be changed slowly], and that we "look for a set Y_{cd}, U_d that contains variables that have a maximum compensating effect on Y_p". If one translates the words and notation, then one realizes that Shinnar's idea of "partial control" is very close to the idea of "self-optimizing control" presented in Morari *et al.* (1980), Skogestad and Postlethwaite (1996), and Skogestad (2000). The difference is that Shinnar assumes that there exist at the outset a set of "primary" variables Y_p that need to be controlled, whereas in self-optimizing control the starting point is an economic cost function that should be minimized. The authors provide some intuitive ideas and examples for selecting dominant variables which may be useful in some cases, especially when no model information is available.

However, it is not clear how helpful the idea of "dominant" variable is, since they are not really defined and no explicit procedure is given for identifying them. Indeed, Arbel *et al.* (1996) write that "the problems of partial control have been discussed in a heuristic way" and that "considerably further research is needed to fully understand the problems is steady-state control of chemical plants".

Tyreus (1999*b*) provides some additional interesting ideas on how to select dominant variables, partly based on the extensive variable idea of Georgakis (1986) and the thermodynamic ideas of Ydstie, (Alonso and Ydstie, 1996), but again no procedure for selecting such variables are presented.

2.5.4 Decomposition of the problem

The task of designing a control system for complete plants is a large and difficult task. Therefore most methods will try to decompose the problem into manageable parts. Four common ways of decomposing the problem are

1. Decomposition based on process units
2. Decomposition based on process structure
3. Decomposition based on control objectives (material balance, energy balance, quality, etc.)
4. Decomposition based on time scale

The first is a horizontal (decentralized) decomposition whereas the latter three provide hierarchical decompositions. Most practical approaches contain elements from several categories.

Many of the methods described below perform the optimization at the end of the procedure after checking if there are degrees of freedom left. However, as discussed above, it should be possible to identify the steady-state degrees of freedom initially, and make a preliminary choice on controlled outputs (c's) before getting into the detailed design.

It is also interesting to see how the methods differ in terms of how important inventory (level) control is considered. Some regard inventory control as the most important (as is probably correct when viewed purely from an operational point of view) whereas Ponton (1994) states that "inventory should normally be regarded as the least important of all variables to be regulated" (which is correct when viewed from a design point of view). We feel that there is a need to integrate the viewpoints of the control and design people.

The unit based approach

The unit-based approach, suggested by Umeda *et al.* (1978), proposes to

1. Decompose the plant into individual unit of operations
2. Generate the best control structure for each unit
3. Combine all these structures to form a complete one for the entire plant.
4. Eliminate conflicts among the individual control structures through mutual adjustments.

This approach has always been widely used in industry, and has its main advantage that many effective control schemes have been established over the years for individual units (e.g. Shinskey (1988)). However, with an increasing use of material recycle, heat integration and the desire to reduce buffer volumes between units, this approach may result in too many conflicts and become impractical.

As a result, one has to shift to plantwide methods, where a hierarchical decomposition is used. The first such approach was Buckley's (1964) division of the control system into material balance control and product quality control, and three plantwide approaches partly based on his ideas are described in the following.

Hierarchical decomposition based on process structure

The hierarchy given in Douglas (1988) for process design starts at a crude representation and gets more detailed:

Level 1 Batch vs continuous

Level 2 Input-output structure

Level 3 Recycle structure

Level 4 General structure of separation system

Level 5 Energy interaction

Fisher *et al.* (1988*a*) propose to use this hierarchy when performing controllability analysis, and Ponton and Laing (1993) point out that this hierarchy, (e.g. level 2 to level 5) could also be used for control system design. This framework enables parallel development for the process and the control system. Within each of the levels above any design method might be applied.

Ng and Stephanopoulos (1998*b*) propose to use a similar hierarchy for control structure design. The difference between Douglas (1988) and Ng and Stephanopoulos (1998*b*)'s hierarchy is that level 1 is replaced by a preliminary analysis and level 4 and on is replaced by more and more detailed structures. At each step the objectives identified at an earlier step is translated to this level and new objectives are identified. The focus is on construction of mass and energy balance control. The method is applied to the Tennessee Eastman case.

All these methods have in common that at each level a key point is to check if there are enough manipulated variables to meet the constraints and to optimize operation. The methods are easy to follow and give a good process understanding, and the concept of a hierarchical view is possible to combine with almost any design method.

Hierarchical decomposition based on control objectives

The hierarchy based on control objectives is sometimes called the tiered procedure. This bottom-up procedure focuses on the tasks that the controller has to perform. Normally one starts by stabilizing the plant, which mainly involves placing inventory (mass and energy) controllers.

Price *et al.* (1993) build on the ideas that was introduced by Buckley (1964) and they introduce a tiered framework. The framework is divided into four different tasks:

I Inventory and production rate control.

II Product specification control

III Equipment & operating constraints

IV Economic performance enhancement.

Their paper does not discuss points III or IV. They perform a large number (318) of simulations with different control structures, controllers (P or PI), and tunings on a simple process consisting of a reactor, separator and recycle of unreacted reactant. The configurations are ranked based on integrated absolute error of the product composition for steps in the disturbance. From these simulation they propose some guidelines for selecting the throughput manipulator and inventory controls. (1) Prefer internal flows as throughput manipulator. (2) the throughput manipulator and inventory controls should be self-consistent (self-consistency is fulfilled when a change in the throughput propagates through the process by “itself” and does not depend on composition controllers). They apply their ideas on the Tennessee Eastman problem (Price *et al.*, 1994).

Ricker (1996) comments upon the work of Price *et al.* (1994). Ricker points out that plants are often run at full capacity, corresponding to constraints in one or several variables. If a manipulated variable that is used for level control saturates, one looses a degree of freedom for maximum production. This should be considered when choosing a throughput manipulator.

Luyben *et al.* (1997) point out three limitations of the approach of Buckley. First, he did not explicitly discuss energy management. Second, he did not look at recycles. Third, he placed emphasis on inventory control before quality control. Their plantwide control design procedure is listed below:

1. Establish control objectives.

2. Determine the control degrees of freedom by counting the number of independent valves.

3. Establish energy inventory control, for removing the heats of reactions and to prevent propagation of thermal disturbances.

4. Set production rate. The production rate can only be increased by increasing the reaction rate in the reactor. One recommendation is to use the input to the separation section.

5. Product quality and safety control. Here they recommend the usual “pair close”-rule.

6. Inventory control. Fix a flow in all liquid recycle loops. They state that all liquid levels and gas pressures should be controlled.

7. Check component balances. (After this point it might bee necessary to go back to item 4.)

8. Unit operations control.

9. Optimize economics or improve dynamic controllability.

Step 3 comes before determining the throughput manipulator, since the reactor is typically the heart of the process and the methods for heat removal are intrinsically part of the reactor design. In order to avoid recycling of disturbances they suggest to set a flow-rate in all recycles loops; this is discussed more in section 2.6. They suggest in step 6 to control all inventories, but this may not be necessary in all cases; e.g. it may be optimal to let the pressure float (Shinskey, 1988). They apply their procedure on several test problems; the vinyl acetate monomer process, the Tennessee Eastman process, and the HDA process.

Larsson and Skogestad, in Chapter 4, question the rule proposed by Luyben. By using the simple recycle plant they are able to show that application of his rule may lead to large problem.

McAvoy (1999) present a method in which the control objective is divided into two categories: variables that must be controlled and product flow and quality. His approach is to find the set of input that will minimize valve movements, where only as many valves as controlled variables are allowed to be used. This is first solved for the “must” variables, then for product rate and quality. The optimization problem is simplified by using a linear stable steady state model. He gives no guidance into how to find which variables that must be controlled.

Hierarchical decomposition based on time scales

Buckley (1964) proposed to design the quality control system as high-pass filters for disturbances and to design the mass balance control system as low pass filters. If the resonance frequency of the quality control system is designed to be an order of magnitude higher than the break frequency of the mass balance system then the two loops will be non-interacting.

McAvoy and Ye (1994) divide their method into four stages:

1. Design of inner cascade loops.

2. Design of basic decentralized loops, except those associated with quality and production rate.

3. Production rate and quality controls.

4. Higher layer controls.

The decomposition in stages 1-3 is based on the speed of the loops. In stage 1 the idea is to locally reduce the effect of disturbances. In stage 2 there generally are a large number of alternative configurations. These may be screened using simple controllability tools, such as the RGA. One problem of selecting outputs based on a controllability analysis is that one may end up with the outputs that are easy to control, rather than the ones that are important to control. The method is applied to the Tennessee Eastman test problem.

Douglas (1988), at page 414, presents a hierarchy for control system design, based on the dynamics that are involved. In this hierarchy the viewpoint is not on the flowsheet but on steady-state, normal dynamic response and abnormal dynamic operation. Zheng *et al.* (1999) continue this work. They places a greater attention to feasibility of the constraints and robust optimality (self-optimizing control). In addition they proposes to use minimum surge capacity as a dynamic cost, (Zheng and Mahajannam, 1999). They have not documented that this will capture the true dynamic cost.

2.6 The reactor, separator and recycle plant

A common feature of most plants is the presence of recycle. A simple example is distillation, with recycle ("reflux") of liquid from the top of the column and of vapor from the bottom of the column.

In this section, we consider the reactor and separator process with recycle of unreacted feed from a reactor. This kind of problem has lately been studied by many authors, (Papadourakis *et al.*, 1987), (Wolff *et al.*, 1992), (Price *et al.*, 1993), (Luyben, 1994), (Luyben and Floudas, 1994), (Mizsey and Kalmar, 1996), (Wu and Yu, 1996), (Hansen, 1998), (Ng and Stephanopoulos, 1998*a*), Larsson and Skogestad (in Chapter 4) and many more. It may be difficult to follow all the details in the case studies presented, so instead we aim in this section to gain some basic insight into the problem.

In the simplest case, let the reactor be a CSTR where component A is converted to a product and the amount converted is given by

$$P = kz_A M \ [moleA/s]$$

The amount of unreacted A is separated from the product and recycled back to the reactor (for simplicity we will here assume perfect separation). To increase the conversion P one then has three options:

1. Increase the temperature which increases the reaction rate constant k.

2. Increase the amount of recycle, which indirectly increases the fraction of A in the reactor, z_A.

3. Increase the reactor holdup M. (In a liquid phase system the reactor holdup is determined by the reactor level, and in a gas phase system by the reactor pressure.)

Here we will assume that the temperature is constant, so there are two options left.

Since at steady-state with given product specifications the conversion of A in the reactor is given by the feed rate, it follows that only one of the two remaining options can be controlled independently (or more generally, one variable that influences these options), and we must let the second variable "float" and adjust itself.

Two common control strategies are then

(A) Control the reactor holdup (and let the recycle flow float)

(B) Control the recycle flow (and let the reactor holdup float).

In case (A) one may encounter the so-called "snowball effect" where the recycle goes to infinity. This occurs because at infinite recycle flow we have $z_A = 1$ which gives the highest possible production. In effect, the snowball effect occurs because the reactor is too small to handle the given feed rate, so it is really a steady-state design problem.

Luyben (1992, 1994) has studied liquid phase systems and has concluded that control strategy (B) (or a variant of it) with one flow fixed in the recycle loop should be used to avoid the "snowball effect".

Wu and Yu (1996) also study the snowball effect for the reactor/separator and recycle plant. They propose as a remedy to the snowball effect to distribute the "work" evenly between the different units. To achieve this they suggest to

(C) Control the reactor composition constant.

Also in this case the reactor volume varies depending on the disturbance.

However, from an economic point of view one should usually for liquid phase systems keep the reactor level at its maximum value, Chapter 4. This maximizes the conversion per pass and results in the smallest possible recycle, which generally will reduce the operational cost. Thus, the recommendation of Luyben (1992, 1994) and (Wu and Yu, 1996), has an economic penalty which it seems that most researchers so far have neglected.

On the other hand, for gas phase systems, there is usually an economic penalty from compression costs involved in increasing the reactor holdup (i.e. the reactor pressure), and strategy (B) where we let the holdup (pressure) float may in fact be optimal. Indeed, such schemes are used in industry, e.g. in ammonia plants. For example, for processes with gas recycle and purge, Fisher *et al.* (1988*a*) recommend to keep the gas recycle constant at the maximum value. For a simple gas phase plant and the recycle plant we have shown that the economic optimum does not coincide with the maximum recycle flow, Chapter 5.

Wolff *et al.* (1992) studied a similar plant. They included an inert component and looked on the effects of recycle on the controllability of the process. Their conclusion is that the purge stream flow should be used to control the composition of inert. They did not consider the reactor holdup as a possible controlled variable.

All the above works have in common that the authors are searching for the right controlled variables to keep constant (recycle flow, reactor volume, composition, etc.). However, a common basis for comparing the alternatives seems to be lacking. In terms of future work, we propose that one first needs to define clearly the objective function (cost) J for the operation of the reactor system. Only when this is given, may one decide in a rigorous manner on the best selection of controlled outputs, for example by using the idea of "self-optimizing" control and evaluating the loss. This is done in a systematic manner Chapter 4 and 5.

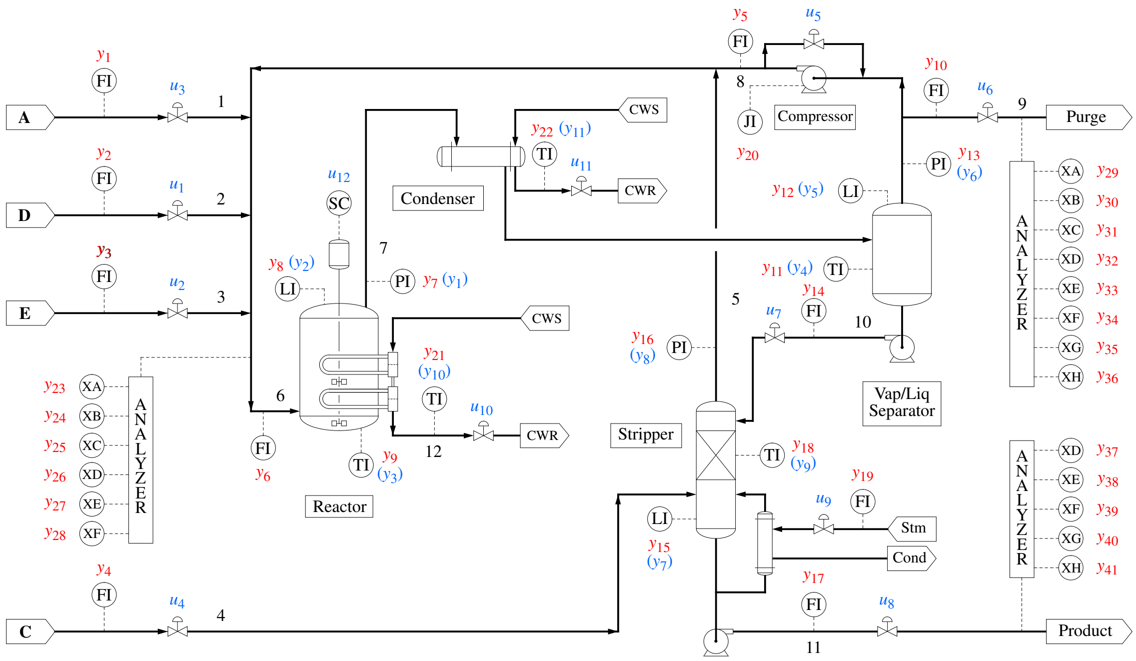


Figure 2.3: Tennessee Eastman process flowsheet

2.7 Tennessee Eastman Problem

2.7.1 Introduction to the test problem

The problem of Downs and Vogel (1993) was first proposed at an AIChE meeting in 1990 and has since been studied by many authors. The process has four feed streams, one product streams, and one purge stream to remove inert. The reactions are

$$\begin{aligned}
A(g) + C(g) + D(g) &\rightarrow G(liq), \quad \text{Product 1},\\
A(g) + C(g) + E(g) &\rightarrow H(liq), \quad \text{Product 2},\\
A(g) + E(g) &\rightarrow F(liq), \quad \text{Byproduct},\\
3D(g) &\rightarrow 2F(liq), \quad \text{Byproduct},
\end{aligned}$$

All reactions are irreversible, exothermic and temperature dependent via the Arrhenius expression. The process has five major units; a reactor, a product condenser, a vapor-liquid separator, a recycle compressor and a product stripper; see Figure 2.3. There are 41 measurements and 12 manipulated variables. For a more detailed description see (Downs and Vogel, 1993) We will here mainly look at the approaches used to solve the problem, not at the solutions themselves.

2.7.2 McAvoy and Ye solution

At stage 1 they close inner cascade loops involving eight flows and two temperature. This reduces the effect of the disturbances associated with these loops. At stage 3 they use a simple mass balance of the plant. This gives some constraints for stage 2, for example, that either the C-feed or the product flow must be left for the third stage.

At stage 2 decentralized loops are closed. They start with the level loops since they are the most important loops. There are three level loops; reactor, separator and stripper, and they consider four possible level configurations. Three of the configurations were ruled out based on controllability analysis. The alternative where the E-feed is used for reactor level control is analyzed in greater detail. They look at three 6x6, eighteen 5x5, and fifteen 4x4 systems, where the controlled outputs seem to be rather randomly chosen. After an analysis involving RGA, Niederlinski index and linear valve saturation, only four alternatives are left. These are further screened on their steady-state behavior for a range of disturbances.

2.7.3 Lyman, Georgakis and Price's solution

Georgakis and coworkers have studied the problem in several papers (Lyman and Georgakis, 1995), (Price *et al.*, 1994). They start by identifying the primary path, which is from the raw materials, through the reactor, condenser, the stripper, and to the product flow. They do not consider the C-feed since it is in excess in the recycle. In (Price *et al.*, 1994) they list all candidates for throughput manipulations along the primary path: The feed streams, flow of coolant to reactor condenser, the separator drum bottoms flows and final product flow. Of the feeds only D is considered. The C-feed was not a candidate as throughput manipulator since it was assumed not to be on the primary path. Next, they list the inventories that need to be controlled: pressure, reactor level, separator level and stripper level. Inventory controls are chosen so to construct a self-consistent path (which does not depend on quality controllers). At this point they have four different structures. After this reactor temperature controller and quality controllers are added.

Their procedure is simple and clear to follow. The result is a control system that is fairly simple to understand.

2.7.4 Ricker's solution

Ricker (1996) starts by listing the variables that must be controlled: production rate, mole % G in product, reactor pressure, reactor liquid level, separator level and stripper level. The production rate is chosen as the input that most likely is going to saturate; namely a combination of D and E. The remaining control system follows by applying quality control and inventory control. After that overrides are installed.

2.7.5 Luyben's solution

Luyben *et al.* (1997) look on two cases, control of throughput with the product flow or control with the A-feed. Here we only look at the case where the product flow is the throughput manipulator. At step 3 they look at energy inventory control, which in this case is to control the reactor temperature with the reactor cooling water. In step 5 they assign the stripper steam stream to control stripper temperature, and therefore also the product compositions. Since the pressure of the reactor must be kept within bounds, they use the largest gas feed (the feed of C) to control the reactor pressure. Step 7 is the check of component balances, which gives a composition controller for inert using the purge flow and a composition controller for A using the A-feed. After doing some simulations they add a controller for control of the condenser, using the reactor temperature.

The resulting control system is simple, but there could have been a better justification on what outputs to control.

2.7.6 Ng and Stephanopulos's solution

Ng and Stephanopoulos (1997, 1998) start by stabilizing the reactor. Then they proceed to look at the input/output level of the plant, where the central point is to fulfill material and energy balances. At this level it should have been possible to say something about how the feeds should be adjusted in order to achieve the right mix of G and H, but they do not. Rather they look at which feed or exit flows that should be used to maintain material balance control.

At the final level they translate the control objectives to measurements. Here material balance control is translated into inventory controllers, like product flow to control stripper level and bottom flow to control separator level. The next objective is then reactor pressure where purge is assigned. Finally feed E is assigned to control of product ratio, and E is assigned to throughput control. The A and C feed is used for controlling composition of A and C.

The method is somewhat difficult to follow and they seem to repeat many of the arguments in each phase.

2.7.7 Larsson, Hestetun and Skogestad's solution

Larsson, Hestetun and Skogestad (in Chapter 6) uses the concept of self-optimizing control for selecting the controlled variables. This problem is well suited since the economic objective was defined. However this problem is larger in size, and they needs to use process insight to reduce the number of alternative structures. For example, they choose to control reactor temperature since everybody else does it. Still they manage to come up with a solution with good self-optimizing properties.

2.7.8 Other work

The above review is by no means complete, and there are many authors who have worked on this problem, some are (Wu and Yu, 1997), (Banerjee and Arkun, 1995) and (Scali and Cortonesi, 1995). In addition there are several others that has looked at other aspects of the Tennessee Eastman plant.

2.7.9 Other test problems

There are several other test problem, that are suitable for studying issues related to plantwide control. These include the HDA-plant (Douglas, 1988), the vinyl acetate monomer process (Luyben and Tyreus, 1998), the recycle plant (Wu and Yu, 1996) and the Luyben and Luyben plant (Luyben and Luyben, 1995).

2.8 A new plantwide control design procedure

Based the above review and as a conclusion to this paper, we propose a plantwide design procedure. The procedure mainly follows the mathematically oriented approach, but with some elements from the process oriented approach. We propose to first perform a top-down analysis to select controlled variables, based on the ideas of self-optimizing control. For this we need a steady state model and operational objectives (steady state economics).The result is one or more alternative sets for (primary) controlled variables ($y_1 = c$).

This top-down analysis is followed by a bottom-up assignment and possibly design of the control loops. This is done in a sequential manner starting with stabilization. Next we consider the fast loops needed for local disturbance rejection. Here we may make use of (extra) secondary measurements (y_2). This is the "regulatory" control layer (system). The objective for the regulatory layer is that manual operation of the plant is possible after these loops are closed.

We now have as degrees of freedom the setpoints of the regulatory layer (r_2) plus any unused manipulators (u_1), these should be used to control the primary outputs (y_1). This control layer is often called the supervisory control layer. This name may be misleading or too limited. Often a supervisory controller is only "activated" when something has gone wrong (e.g. logical switches), but we will also use it for "active" controllers. An alternative name could have been advanced control layer, but advanced is a relative word and it is not given that the controllers in this layer are advanced.

There are two main approaches here: Use of single loop (decentralized) controllers with possible feed-forward links, or use of multivariable control, e.g. decoupling or model predictive control (MPC). Properly designed multivariable controllers give better performance, but this must be traded against the cost of obtaining and maintaining the models used in the controller.

The main result of this will be the control structure, but controller tuning may also be obtained. Finally, nonlinear dynamic simulations should be performed to validate the proposed control structure.

2.9 Conclusion

In this paper we have given a review on plantwide control with emphasis on the following tasks that make up the control structure design problem:

1. *Selection of controlled outputs* (c with setpoints c_s).

2. *Selection of manipulated inputs* (m).

3. *Selection of measurements* (v)

4. *Selection of control configuration*

5. *Selection of controller type*

The main emphasis has been on the selection of controlled outputs, where we have seen that the use of a steady-state economic criterion is very useful. It appears that the solution to this task provides the much needed link between steady-state optimization and process control, and that the idea of "self-optimizing control" to reduce the effect of uncertainty provides a very useful framework for making the right decision. We thus propose that the design of the control system should start with the optimization (or at least identifying what the control objectives really are) and thus providing candidate sets for the controlled outputs. The control problem is then defined, and one may proceed to analyze (e.g. using an input-output controllability analysis, whether the control objectives can be met).

The actual design of the control system is done at the end, after the control problem has been defined, including the classification of all variables (into inputs, disturbances, controlled variables, etc.). Control system design usually starts with stabilizing control where it is usually important to avoid input saturation. The control system is then build up in a hierarchical manner such that each controller is of limited size (usually with as few inputs and outputs as possible). Emphasis should be on avoiding "long" loops, that is, one should pair inputs and outputs with are "close" to another.

Most of the proposed process oriented procedures have elements from this way of thinking, although some procedures focus mostly on control and operation and seem to skip lightly over the phase where the overall control problem is defined.

Several case studies have been proposed, which is good. However, some of the work on these case studies seems to provide little general insight, and their value may therefore be questioned. A more systematic approach and a common ground of comparison are suggested for future work.

In summary, the field of plantwide control is still at a relatively early stage of its development. However, the progress over the last few years, both in terms of case studies and theoretical work, bears promise for the future. There is still a need for a clearer definition of the issues, and it is hoped that this paper may be useful in this respect. In the longer term, where automatic generation and analysis of alternative structures may become more routine, the main problem will probably be to be able to generate models in an efficient way, and to provide efficient means for their analysis (e.g. using input-output controllability analysis).

A plantwide control design procedure

	Step	Tools (in addition to insight)
	Top-down analysis:	
1.	CONTROLLED VARIABLES: What is the control objective and which variables should be controlled? Goal: Obtain primary controlled variables ($y_1 = c$)	*Steady-state model and operational objectives* Degree of freedom analysis. Determine the major disturbances. Evaluate the (economic) loss, with constant controlled outputs and look for "self-optimizing" control structure.
2.	PRODUCTION RATE: Where should the throughput be set? The throughput manipulator cannot be used for other control tasks, thus this choice will have a large effect on the structure of the control system.	The optimal choice may follow from steady-state optimization, but may require continuous reconfiguration (may use MPC to avoid this) The throughput manipulator should have a strong and direct effect on the production rate.

A plantwide control design procedure

	Bottom up design: (With given controlled and manipulated variables)	**Controllability analysis** Compute zeros, poles, relative gain array, minimum singular value, etc.
1.	REGULATORY CONTROL LAYER. The purpose of this control layer is to enable manual operation of the plant	
	Stabilization. Selection of measurements and inputs for stabilization (including slowly drifting modes).	*Pole vectors* Give insight about which measurements and inputs can be used for each unstable mode. Select large elements: Small input energy needed and large noise tolerated.
	Local disturbance rejection. Often based on secondary measurements.	*Partially controlled plant* Select secondary measurements (y_2) so that the effect of disturbances on the primary output (y_1) can be handled by the operators.
2.	SUPERVISORY CONTROL LAYER. The purpose of this control layer is to control the primary control variables (c).	
	Decentralized control. We use a decentralized control structure if the process is noninteracting and the constraints are not changing. Feed-forward control and ratio control may be useful here.	*Controllability analysis for decentralized control* Pair on relative gain array close to identity matrix at crossover frequency, provided not negative at stead-state. Closed loop disturbance gain (CLDG) and performance gain array (PRGA) may be used to analyze the interactions.
	Multivariable control Use for interacting process (coordination including feed-forward control). MPC is useful for tracking the active constraints (if the steady-state optimization shows that the active constraints are changing with disturbances).	
3.	REAL TIME OPTIMIZATION	
	Validation.	Nonlinear dynamic simulation

Chapter 3

Limitations imposed by lower layer partial control

T. Larsson and S. Skogestad

Extended version of paper presented at AIChE Annual meeting 1998, Miami.

Abstract

In this chapter we study how fundamental limitations can be introduced by improper design of the lower control layer. One useful tool for analyzing this problem is the concept of partial control, which shows how the plant looks as seen from the higher layer controller.

The conclusion is that we cannot introduce new limitations, provided that we have access to the measurements of the already closed loops, that their setpoints are available as degrees of freedom and that the lower layer controller is stable and minimum phase.

However, in some cases the requirements are not met, and we may introduce limitations. For example if the setpoints of the lower layer are unavailable, pairing on a negative RGA may introduce a negative RHP-zero (inverse response) in the remaining system. We may also introduce sensitivity to disturbances and ill-conditioning into the remaining system.

3.1 Introduction

A hierarchical control system arises when design is done in a sequential manner. Typically, a chemical plant has at least three layers; a regulatory (base) control layer, a supervisory control layer with local optimization (e.g. model predictive control), and an optimizing layer (usually based on steady-state optimization). The functions in the two upper layers are often performed by humans. The lower regulatory control layer usually consists of PID controllers. The task of this layer is to "stabilize" the plant, so that it can be operated manually, without the high level controllers in place.

Hovd and Skogestad (1993) proposed some rules for designing the lower layer. One of the points they mention is that one should not introduce new fundamental limitations, like RHP-zeros and ill-conditioning, into the remaining control problem. Is this really a problem? That is, can an improper selection of the base control layer lead to unnecessary limitations, which cannot be overcome by the higher layers? By "unnecessary" we mean limitations that were not present as fundamental limitations in the original plant.

The aim of this paper is to study this problem more generally: Can we by improper design of a lower layer control system, introduce (new) fundamental limitations not present in the original system?

We will use the terms "lower layer" and "inner loop" as synonyms, and also the terms "upper layer" and "outer loop" are used as synonyms. The term fundamental limitation will here imply a restriction on the attainable control performance.

Outline

We first introduce the idea of partial control, and then discuss how to cancel the effect of the lower control layer. We then discuss the possibility for RHP-zeros in subsystems under partial control. Next we look at the effect of disturbance under partial control, and then at ill-conditioning, as indicated by large RGA elements. Finally one example is devoted to a high-purity distillation column, where it seems that we can remove a fundamental limitation by closing the lower layer levels.

3.2 Partial Control

One useful tool for analyzing the effect of closing lower layer loops, is the concept of partial control, e.g. (Häggblom and Waller, 1988), (Skogestad and Wolff, 1992), (Havre and Skogestad, 1996*a*) and (Skogestad and Postlethwaite, 1996). We partition the inputs u and outputs y into two sets

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

The linear model $y = Gu + G_d d$ may then be written

$$y_1 = G_{11}u_1 + G_{12}u_2 + G_{d1}d \tag{3.1}$$

$$y_2 = G_{21}u_1 + G_{22}u_2 + G_{d2}d \tag{3.2}$$

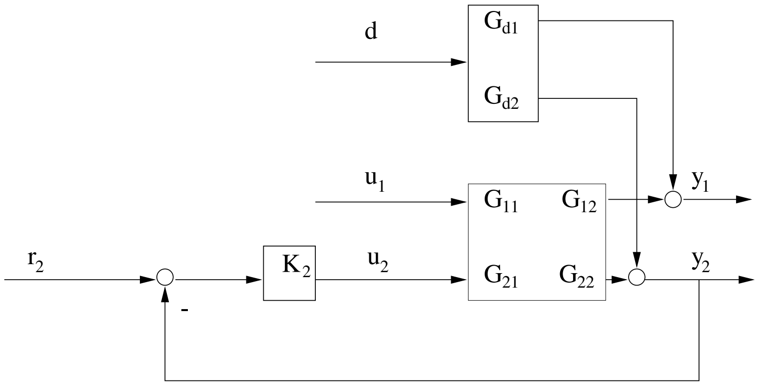


Figure 3.1: Partial control.

It is assumed that the plant is partially controlled, see Figure 3.1, by closing the loops involving the variables u_2 and y_2 (measurement noise is ignored),

$$u_2 = K_2(r_2 - y_2) \tag{3.3}$$

where r_2 is the reference for y_2 and K_2 the lower layer controller. The partially controlled system becomes (Havre and Skogestad, 1996*a*) and (Skogestad and Postlethwaite, 1996)

$$\begin{aligned} y_1 &= (G_{11} - G_{12}K_2(I + G_{22}K_2)^{-1}G_{21})u_1 + \\ &\quad (G_{d1} - G_{12}K_2(I + G_{22}K_2)^{-1}G_{d2})d + \\ &\quad G_{12}K_2(I + G_{22}K_2)^{-1}r_2 \\ y_2 &= (I + G_{22}K_2)^{-1}G_{21}u_1 + (I + G_{22}K_2)^{-1}G_{d2}d + \\ &\quad (I + G_{22}K_2)^{-1}G_{22}K_2r_2 \end{aligned} \tag{3.4}$$

In compact form the “new” partially controlled plant is written

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \underbrace{\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}}_{P} \begin{bmatrix} u_1 \\ r_2 \end{bmatrix} + \underbrace{\begin{bmatrix} P_{d1} \\ P_{d2} \end{bmatrix}}_{P_d} d \tag{3.5}$$

where from Equation 3.4

$$P = \begin{bmatrix} G_{11} - G_{12}K_2S_2G_{21} & G_{12}K_2S_2 \\ S_2G_{21} & G_{22}K_2S_2 \end{bmatrix} \tag{3.6}$$

$$P_d = \begin{bmatrix} (G_{d1} - G_{12}K_2S_2G_{d2}) \\ S_2G_{d2} \end{bmatrix} \tag{3.7}$$

$$S_2 = (I + G_{22}K_2)^{-1} \tag{3.8}$$

By closing the lower layer loops the number of degrees of freedom is not changed, at least provided $\dim y_2 \geq \dim u_2$ and that the setpoints r_2 may be used by the upper control layer.

Comment: It is worth noting that we can rewrite P as

$$P = \left(I + G \begin{bmatrix} 0 & 0 \\ 0 & K_2 \end{bmatrix} \right)^{-1} G \begin{bmatrix} I & 0 \\ 0 & K_2 \end{bmatrix} \tag{3.9}$$

where we see that $\begin{bmatrix} 0 & 0 \\ 0 & K_2 \end{bmatrix}$ inside the paranthesis is different from $\begin{bmatrix} I & 0 \\ 0 & K_2 \end{bmatrix}$ outside the paranthesis. This means that partial control is not just a special case of (a one degree of freedom) multivariable feedback control.

3.2.1 Perfect control

In the following we assume that $\dim y_2 = \dim u_2$ so that K_2 and G_{22} are square matrices and that G_{22}^{-1} exists (at least on a frequency by frequency basis). Often the lower layers, involving y_2, are much faster than the outer layers, involving y_1. This implies that the assumption of perfect control ($y_2 = r_2$) is useful for gaining insight into the behavior of the "new" plant as seen from the layer above. By letting $y_2 = r_2$ in equation 3.1 and 3.2, we get

$$y_1 = (G_{11} - G_{12}G_{22}^{-1}G_{21})u_1 + G_{12}G_{22}^{-1}r_2 + (G_{d1} - G_{12}G_{22}^{-1}G_{d1})d \tag{3.10}$$

$$y_2 = r_2 \tag{3.11}$$

In compact form:

$$P^* = \begin{bmatrix} G_{11} - G_{12}G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1} \\ 0 & I \end{bmatrix} \tag{3.12}$$

$$P_d^* = \begin{bmatrix} (G_{d1} - G_{12}G_{22}^{-1}G_{d2}) \\ 0 \end{bmatrix} \tag{3.13}$$

where * indicates perfect control.

These equations provide a good approximation at all frequencies where control of y_2 is thigh i.e. $y_2 \approx r_2$, $S_2 \approx 0$ and $K_2S_2 \approx G_{22}^{-1}$.

3.3 Cancellation of lower control layer

Is it always possible for the higher layer to cancel the effects introduced by the lower control layer? If this is the case then, in principle, we cannot introduce any new fundamental limitation by closing the lower layer loops.

To answer this question considers Figure 3.2, where the original partially controlled plant is shown within the dotted box. Outside this we have added a positive unity feedback that cancels the negative feedback. Note that this requires access to both y_2 and r_2. In addition, we have added a block K_2^{-1} at the input, which cancels the effect of K_2, assuming that K_2 is stable and minimum phase so that $K_2K_2^{-1}$ contains no unstable hidden modes.

Note that the scheme in Figure 3.2 makes use of feedback from y_2. Is it possible to cancel the effect of the lower layer by just using feed-forward from r_2? If this was the case then we

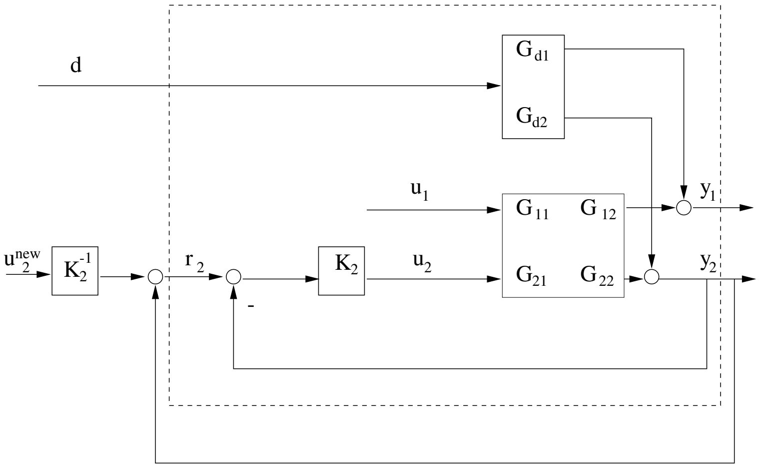


Figure 3.2: Cancellation of the lower layer controller by positive feedback.

would be able to cancel the controller without using y_2. For simplicity let us assume that all outputs are controlled, then our new plant is

$$y = Tr + SG_d d \tag{3.14}$$

By applying the feed-forward controller $r = T^{-1}Gu^{new}$ we would get

$$y = Tr + SG_d d = Gu^{new} + SG_d d \tag{3.15}$$

Although we have restored G, the effect of the disturbances are changed from G_d to SG_d. Thus, we cannot cancel the effect of the lower layer feedback loops if we only have access to r_2

We summarize the above findings in the following theorem (see Appendix 3.A for a complete proof):

Theorem 1 *By applying partial control to a plant, Figure 3.1, the achievable control performance for the remaining control problem, will be the same as it where for the original plant if (i.e. the lower control layer can not introduce any new fundamental limitations if)*

1. *r_2 is available as a manipulated variable at the layer above.*
2. *y_2 is available as a measurement at the layer above.*
3. *K_2 is stable and minimum phase so that $K_2K_2^{-1}$ contains no unstable hidden modes.*

These results are rather trivial but nevertheless important. At this point we would like to stress that the goal here is not to show how the lower layer controllers could be canceled, but to show that the lower layer control does not introduce any new fundamental limitations, provided that the conditions in the above theorem are met. (In practice one would set the controllers in manual.)

In practice the conditions in Theorem 1 (and 2), may not always be met and partial control may introduce some fundamental limitations. For example if the setpoints r_2 are not available as manipulated variables at the layer above, then limitations in P_{11} will limit the achievable control performance of y_1. Conditions for introducing RHP-zeros in P_{11} is discussed in Section 3.4.1 and in Section 3.6.1 the possibility for introducing ill-conditioning in P_{11} is discussed.

If the measurement y_2 is not available for the higher layer, then as indicated by Equation 3.15, we may impose limitations in the ability to reject disturbances. This is further discussed in Section 3.5.

It should also be noted that requirement 3 is not always met as some plants require unstable or non-minimum phase controller for stabilization, and in other cases, it may be optimal as seen from the lower layer to use a controller K_2 which is unstable or non-minimum phase. Nevertheless, these cases are relatively rare. An important exception to this is integral action in K_2, which means that K_2 is at the limit to instability. However, as shown in appendix 3.A, we may cancel the lower layer but not the intergral effect. Then the "new plant" would be equal to the original plant except that the "new plant" would contain controllable poles at the origin. But these poles would not represent a fundamental limitation.

3.4 RHP-zeros and partial control

Theorem 1 indicates that we cannot introduce new RHP-zeros in P, provided the controller is stable and minimum phase. The following theorem is valid only for RHP-zeros and is slightly stronger than Theorem 1, the condition on a stable controller is removed (see Appendix 3.B for proof):

Theorem 2 *A plant under partial control, Figure 3.1, will only have multivariable RHP-zeros from the plant G in the transfer-function from the manipulated variables and the measurements of the higher layer if*

1. *y_2 is available as a measurement at the layer above.*
2. *r_2 is available as a manipulated variable at the layer above.*
3. *K_2 is minimum phase.*

Any RHP-zeros in the controller K_2 will be RHP-zeros of the higher layer plant, i.e. from (u_1, r_2) to (y_1, y_2).

Thus we cannot introduce new multivariable RHP-zeros in P which where not present in either G or K_2. However, in many cases we may want to use only u_1 to control y_1, or only r_2 to control y_1. We now want to study if new RHP-zeros can occur in the resulting subsystem, P_{11} or P_{12}.

3.4.1 RHP-zeros in P_{11}

Let us first consider the case where the upper layer does not make use of the setpoints r_2. The relevant transfer function from u_1 to y_1 is then $P_{11} = G_{11} - G_{12}K_2S_2G_{21}$. It is well known that pairing on a negative steady-state RGA-element in G and use of integral control implies the introduction of new a RHP-zero in P_{11}, (Shinskey, 1979, Bristol, 1966 and Grosdidier and Morari, 1985). This was generalized to non-integral control by Jacobsen (1997). Jacobsen and Hong (1998) gives a general condition for P_{11} and G_{11} to have a different number of RHP-zeros.

The presence of a RHP-zero in P_{11} implies that there is a fundamental limitation in the use of u_1 to control y_1. However, the significance of the limitation depends on the location of the RHP-zero and a closer analysis shows that these new RHP-zeros will often be located at frequencies at or beyond the bandwidth of the control system involving the other outputs y_2. This means that if we require the bandwidth in the outer loops to be much less than in the inner loop, then the RHP-zeros will not be a limitation. This is illustrated by the following theorem, where we assume perfect (very fast) control in the inner loop.

Theorem 3 *Consider a plant G for which it is possible to achieve perfect control for a subset of inputs and outputs corresponding to the subsystem G_{22}. If perfect control is applied to this subsystem, the remaining part of the plant involving the unused input and outputs, $P_{11}^* = G_{11} - G_{12}G_{22}^{-1}G_{21}$ will contain the RHP-zeros of the original plant G.*

The proof is included in Appendix 3.C. Rosenbrock (1970) presented a similar result when all outputs but one are under perfect control. By continuity arguments, it then follows that any new RHP-zeros in P_{11} will move to higher frequencies as we "tighten" the control of the inner loop involving y_2.

Example 1 *To illustrate the points above, we consider the following plant*

$$G(s) = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \tag{3.16}$$

The $(2,2)$-element of the RGA for this plant is $\lambda_{22} = -1$. Pairing on this negative RGA element, e.g. by closing an "inner" loop $u_2 = \frac{K}{s}(r_2 - y_2)$, gives ($K$ is a constant positive gain.)

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = P \begin{bmatrix} u_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} \frac{s/K-1}{s/K+1} & \frac{1}{s/K+1}r_2 \\ \frac{2s/K}{s/K+1} & \frac{1}{s/K+1} \end{bmatrix} \begin{bmatrix} u_1 \\ r_2 \end{bmatrix} \tag{3.17}$$

Note that the overall plant P does not have any RHP-zeros, that is there are no multivariable RHP-zeros in the transfer function from (u_1, r_2) to (y_1, y_2). However, if we do not make use of r_2, then the remaining plant as seen from the upper layer is

$$y_1 = P_{11}u_1 = \frac{s/K - 1}{s/K + 1}u_1 \tag{3.18}$$

and there is a RHP-zero is at K, which is the same as the bandwidth frequency of the inner loop ($L_2 = G_{22}K_2 = K/s$). Thus if the outer loop, involving only u_1 and y_1, is significantly slower than the inner loop, the RHP-zero will not pose much of a limitation. Furthermore if we increase K, i.e. tighten the control of y_2 the RHP-zero will move further out in the RHP-plan.

It is, however, possible to construct examples where the RHP-zero appears at a frequency well below the bandwidth of the lower layer, see examples in (Jacobsen and Hong, 1998). But, as expected from Theorem 3, in all the cases the RHP-zero will move to higher frequencies (becoming a less severe limitation) if the bandwidth of the lower layer is increased.

3.4.2 RHP-zeros in P_{12} due to RHP-poles in G_{22}

It is well known that the presence of RHP-poles implies a fundamental limitation, e.g. (Skogestad and Postlethwaite, 1996), as it requires feedback control with active use of plant inputs. Here we will provide a consequence of this for partial control. Consider the case where the primary objective is to use the input u_2 to control the output y_1 (the set u_1 is empty). However, the plant is unstable, and we also need to use u_2 to stabilize the plant, for which we have available the measurement y_2. Thus we have (not considering disturbances)

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix} u_2$$

Will the instability limit the performance of the control of y_1? If the instability is not detectable in y_1 then the instability will limit the performance of the control of y_1. More specifically we have the following theorem:

Theorem 4 *Consider an unstable plant where the instability is not detectable in the output set y_1. Assume that the outputs (measurements) y_2 with setpoints r_2 are used for stabilizing control. Then the RHP-poles of the original plant G as RHP-zeros in the transfer function P_{12} from r_2 to y_1.*

Proof: *The transferfunction from r_2 to y_1 is $P_{12} = G_{12}K_2S_2$. By assumption G_{12} is stable and G_{22} is unstable. For internal stability all RHP-poles in G_{22} must appear as RHP-zeros in K_2S_2 (Youla* et al.*, 1974), and since G_{12} is stable as RHP-zeros in P_{12}.*

Thus we introduce a fundamental limitation in P_{12} whenever we attempt to control variables not containing this instability. A special case of the above is when $y_1 = u_2$ ($P_{12} = I$). This occurs when we want to control the input used for stabilization ("input-resetting").

Example 2 *Attempting of input resetting for liquid level control. Assume that a flow u_2 used for controlling the liquid level y_2, which is an integrating process. We would like to freely change the flow, i.e. $y_1 = u_2$. Thus the model is*

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = Gu_2 = \begin{bmatrix} 1 \\ \frac{1}{s} \end{bmatrix} u_2 \tag{3.19}$$

A P-controller, with unit gain, $u_2 = r_2 - y_2$, is used for stabilization. The resulting transfer function from r_2 to y_1 is

$$P_{12} = \frac{s}{s+1} \tag{3.20}$$

and we find as expected that the pole at the origin in G_{22} appears as a zero at the origin in P_{12}. Thus, it will not be possible to control the flow $y_1 = u_2$ at steady state. This is well known in practice, and follows physically from the fact that inflow must equal the outflow at steady state.

3.5 Disturbances and partial control

Without control, the effect of the disturbance d on the outputs is (not considering the inputs)

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} G_{d1} \\ G_{d2} \end{bmatrix} d \tag{3.21}$$

and under partial control (closing the loops involving y_2 and r_2) we have

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} P_{d1} \\ P_{d2} \end{bmatrix} d \begin{bmatrix} (G_{d1} - G_{12}K_2(I + G_{22}K_2)^{-1}G_{d2}) \\ (I + G_{22}K_2)^{-1}G_{d2} \end{bmatrix} d \approx \begin{bmatrix} (G_{d1} - G_{12}G_{22}^{-1}G_{d2}) \\ 0 \end{bmatrix} d \tag{3.22}$$

where the approximation holds at frequencies where we have tight control of y_2 i.e. $y_2 \approx r_2$. Although the objective of feedback control is to reduce the sensitivity for disturbances on y_1 (i.e. to have $\|P_{d1}\| < \|G_{d1}\|$), we see from the above equations that we may in some cases get a larger sensitivity, (i.e. $\|P_{d1}\| > \|G_{d1}\|$). For example, we see that if the disturbance originally had no effect on y_1, i.e. $G_{d1} = 0$, but do effect y_2, i.e. $G_{d2} \neq 0$, then closing the inner loop will yield a nonzero transfer function $P_{d1} = G_{12}K_2(I + G_{22}K_2)^{-1}G_{d2}$.

In summary, we find that if we close a lower loop involving the output y_2, then we may in some cases introduce a sensitivity to disturbance not present in the original transfer function from u_1 to y_1. If the lower layer measurement is not available in the next control layer then this may impose a fundamental limitation.

To study the other desired situation, where local feedback through control of y_2 reduces the disturbance sensitivity, we assume for simplicity perfect control of y_2. Then

$$y_1 = (I - G_{12}G_{22}^{-1}G_{d2}G_{d1}^{-1})G_{d1}d \tag{3.23}$$

and we see that if $G_{12} = G_{22}$ and $G_{d1} = G_{d2}$ then $y_1 = 0\ d$. To state this in a clear manner

- If the disturbance d and the secondary input u_2 have similar effects (on y_1 and y_2), i.e. $G_{d1} \approx G_{d2}$ and $G_{12} \approx G_{22}$ then the primary measurement y_1 will become less sensitive to the disturbance when u_2 is used to control y_2.

- If the disturbance d or the secondary input u_2 have different effects (on y_1 and y_2), then the primary measurement y_1 may become more sensitive to the disturbance when u_2 is used to control y_2.

This insight is useful to evaluate if cascaded control should be applied. There is more on this subject on page 429 in (Skogestad and Postlethwaite, 1996).

Comment: Much of this discussion is valid for model uncertainty and for noise.

3.6 Ill-conditioning and partial control

3.6.1 Introducing ill-conditioning

The relative gain array (RGA) is defined by $RGA(G) = G \times (G^{-1})^T$, where $\times$ denotes element-by-element multiplication. Large RGA-elements imply fundamental control problems, (Skogestad and Morari, 1987). More specifically, decoupling control can not be used due to sensitivity to input gain uncertainty, whereas decentralized (diagonal) control yields poor control, even nominally, since large RGA-elements imply strong two-way couplings.

The following example shows that partial control invovling u_2 and y_2, may introduce large RGA-elements into the remaining control problem involving u_1 and y_1 (P_{11}). (However, note that these large RGA elements do not appear if we look at the "full" partially controlled plant P including y_2 and r_2).

Example 3 *Consider the following plant*

$$G = \left[\begin{array}{c|c} G_{11} & G_{12} \\ \hline G_{21} & G_{22} \end{array}\right] = \left[\begin{array}{cc|c} nn1 & 0.51 & 1.02 \\ 1 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}\right] \quad RGA(G) = \begin{bmatrix} -1.89 & 0.96 & 1.91 \\ .96 & 0.04 & 0 \\ 1.91 & 0 & -0.92 \end{bmatrix} \tag{3.24}$$

All the RGA elements of G are small, indicating no particular control problem. However, applying partial control gives the partially controlled plant

$$P_{11} = \begin{bmatrix} 0.49 & 0.51 \\ 1 & 1 \end{bmatrix} \tag{3.25}$$

for which the relative gain is

$$RGA(P_{11}) = \begin{bmatrix} -24.5 & 25.5 \\ 25.5 & -24-5 \end{bmatrix} \tag{3.26}$$

These large values in the RGA indicates sever control problems for using u_1 to control y_1, when u_2 is used for controlling y_2.

3.6.2 Apparent removing ill-conditioning (distillation example)

Let us first make clear that ill-conditioning (e.g. as expressed by large RGA-element) is a fundamental problem, so it cannot be removed by applying partial control.

Nevertheless, this is exactly what seems to happen for some high-purity distillation columns for which the original plant has large RGA-elements. But somehow, by closing the level loops in a particular way to obtain the DV-configuration, we are able to eliminate the large RGA-elements. Thus a fundamental limitation seems to have disappeared, which is obviously not possible. This is puzzling and below we will study how this happens.

Example 4 *We consider "column A" of Skogestad and Morari (1988), with constant pressure. The column has four inputs, reflux (L), boil-up (V), distillate (D) and bottom product flow (B). The four outputs are, top composition of light component (y_D), bottom composition of light component (x_B), level in condenser (M_D) and level in the re-boiler (M_D). A simplified linear model is, (Skogestad and Morari, 1988)*

$$\begin{bmatrix} y_D \\ x_B \\ M_D \\ M_B \end{bmatrix} = \underbrace{\begin{bmatrix} g_{yL}(s) & g_{yV}(s) & 0 & 0 \\ g_{xL}(s) & g_{xV}(s) & 0 & 0 \\ -\frac{1}{s} & \frac{1}{s} & -\frac{1}{s} & 0 \\ \frac{1}{s} & -\frac{1}{s} & 0 & -\frac{1}{s} \end{bmatrix}}_{G} \begin{bmatrix} L \\ V \\ D \\ B \end{bmatrix} \tag{3.27}$$

At steady state

$$G_{11} = \begin{bmatrix} g_{yL} & g_{yV} \\ g_{xL} & g_{xV} \end{bmatrix} = \begin{bmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{bmatrix}; \quad RGA = \begin{bmatrix} 35.1 & -34.1 & 0 & 0 \\ -34.1 & 35.1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.28}$$

These large RGA values indicate a problem with decoupling control.

The tasks of composition control and stabilization of the levels usually require different bandwidths. We therefore partition the plant so that $y_1 = \begin{bmatrix} y_D \\ x_B \end{bmatrix}$ *(compositions) and* $y_2 = \begin{bmatrix} M_D \\ M_B \end{bmatrix}$ *(levels). The inputs may be partitioned in many ways, and each choice of* u_1 *is denoted a "configuration", e.g.* $u_1 = \begin{bmatrix} L \\ V \end{bmatrix}$ *is named the LV-configuration and* $u_1 = \begin{bmatrix} D \\ V \end{bmatrix}$ *is named the DV-configuration.*

With perfect level control the LV-configuration becomes

$$\underbrace{\begin{bmatrix} y_D \\ x_B \end{bmatrix}}_{y_1} = \underbrace{\begin{bmatrix} g_{yL} & g_{yV} \\ g_{xL} & g_{xV} \end{bmatrix}}_{P_{11}^{*LV}} \underbrace{\begin{bmatrix} L \\ V \end{bmatrix}}_{u_1} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}}_{P_{12}^{*LV}} \underbrace{\begin{bmatrix} M_{Ds} \\ M_{Bs} \end{bmatrix}}_{r_2} \tag{3.29}$$

The RGA of P_{11}^{*LV} *is thus the same as for the original plant, and at steady-state*

$$\Lambda\left(P_{11}^{*LV}\right) = \begin{bmatrix} 35.1 & -34.1 \\ -34.1 & 35.1 \end{bmatrix} \tag{3.30}$$

indicating fundamental problems with decoupled control of the compositions $y_1 = \begin{bmatrix} x_D \\ x_B \end{bmatrix}$.

However, with the DV-configuration, assuming perfect level control, the remaining plant is

$$\underbrace{\begin{bmatrix} y_D \\ x_B \end{bmatrix}}_{y_1} = \underbrace{\begin{bmatrix} -g_{yL} & g_{yV} + g_{yL} \\ -g_{xL} & g_{xV} + g_{xL} \end{bmatrix}}_{P_{11}^{*DV}} \underbrace{\begin{bmatrix} D \\ V \end{bmatrix}}_{u_1} + \underbrace{\begin{bmatrix} g_{yL}s & 0 \\ g_{xL}s & 0 \end{bmatrix}}_{P_{12}^{*DV}} \underbrace{\begin{bmatrix} M_{Ds} \\ M_{Bs} \end{bmatrix}}_{r_2} \tag{3.31}$$

and the steady state RGA of P_{11} *for this control configuration is*

$$\Lambda\left(P_{11}^{*DV}\right) = \begin{bmatrix} 0.46 & 0.54 \\ 0.54 & 0.46 \end{bmatrix} \tag{3.32}$$

Thus, the large RGA values have disappeared(!), and it should be possible to have decoupled control of the compositions. This seems inconsistent with the large values of the RGA.

A more careful study reveals that in the DV-configuration there are interactions from the level loops to the compositions, $P_{12}^{*DV} \neq 0$, *whereas there are no interactions in the LV-configuration,* $P_{12}^{*LV} = 0$. *Apparently these interactions from the level loops mean that in the DV-configuration we can have decoupling between the compositions.*

In practice, the coupling from the condenser level, imply that we must maintain constant condenser level when using the DV-configuration. This is clearly a disadvantage, which must be traded against the possibility for achieving decoupled control of the compositions.

3.7 Conclusion

This paper has followed up a point made by Hovd and Skogestad (1993); that the lower control layer should in such not impose any new fundamental limitation that was not present in the original plant.

We find that if both the measurements y_2 and setpoints r_2 of the lower layer are available for the next layer and the lower layer controller is stable and minimum phase, then no new fundamental limitation can be introduced. This follows because we can then cancel the lower layer controller with a positive feedback and inversion of the lower layer controller at the input. However, if either y_2 or r_2 are unavailable then it is possible to introduce new limitations for the higher layer.

3.A Proof of Theorem 1

We study the following system (i.e. the partially controlled system in Figure 3.1, where the loop involving y_2 and u_2 is closed)

$$\begin{aligned} y_1 &= (G_{11} - G_{12}K_2S_2G_{21})u_1 + \\ &\quad (G_{d1} - G_{12}K_2S_2G_{d2})d + \\ &\quad G_{12}K_2S_2r_2 \\ y_2 &= S_2G_{21}u_1 + S_2G_{d2}d + \\ &\quad S_2G_{22}K_2r_2 \end{aligned} \tag{3.33}$$

Here r_2 is a degree of freedom and we apply a "two degree of freedom" controller

$$r_2 = K_3u_2^{new} - K_4y_2 \tag{3.34}$$

The resulting closed-loop system is

$$\begin{aligned} y_2 &= (I + S_2G_{22}K_2K_4)^{-1}S_2(G_{21}u_1 + G_{d2}d) + \\ &\quad (I + S_2G_{22}K_2K_4)^{-1}S_2G_{22}K_2K_3u_2^{new} \end{aligned}$$

If $(I + S_2G_{22}K_2K_4)^{-1}S_2 = I$ then the effect of u_1 and d will become that of the original plant. This is achieved if $(I + \underbrace{S_2G_{22}K_2}_{T_2}K_4) = S_2$, which is satisfied if $K_4 = -I$ (i.e. positive feedback of unity gain). Then

$$\begin{aligned} y_1 &= G_{11}u_1 + G_{d1}d + G_{12}K_2K_3u_2^{new} \\ y_2 &= G_{21}u_1 + G_{d2}d + G_{22}K_2K_3u_2^{new} \end{aligned} \tag{3.35}$$

To also restore the original transfer function from u_2 to y, we need $K_3 = K_2^{-1}$. Thus, K_2 has to be minimum phase and stable (otherwise $K_2K_2^{-1}$ would contain hidden unstable modes). Thus minimum phase behavior and instabilities introduced by the controller can not be canceled by the higher layer.

If the controller K_2 uses integral action then by letting K_3 be equal $\tilde{K}_2^{-1}$, where $K_2 = \tilde{K}_2/s$ and $\tilde{K}_2$ is the stable, Equation 3.35 becomes

$$\begin{aligned} y_1 &= G_{11}u_1 + G_{d1}d + \frac{G_{12}}{s}u_2^{new} \\ y_2 &= G_{21}u_1 + G_{d2}d + \frac{G_{22}}{s}u_2^{new} \end{aligned} \tag{3.36}$$

Provided that either G_{12} or G_{22} does not have a RHP-zero at the origin then the pole at the origin will be controllable from u_2^{new}. Thus the only difference between the above plant and the original plant G is that it contains a pole at the origin which is not really a fundamental limitation.

Comment: It is well-known that zeros of a plant are not moved by feedback control, i.e. the zeros of GK and $(I + GK)^{-1}GK$ are the same. However, as noted from Equation 3.9, partial control is not a special case of feedback control, and this is why we have (re-)derived Theorem 2.

3.B Proof of Theorem 2

For the sake of argument we let K_3 in Equation 3.35 be equal to I, then

$$y = G \begin{bmatrix} I & 0 \\ 0 & K_2 \end{bmatrix} u_2^{new} + G_d d \tag{3.37}$$

From the above equation we can clearly see that any RHP-zeros of the partially controlled plant has to come from either G or K_2 even if K_2 has RHP-poles.

3.C Proof of Theorem 3

The partially controlled plant is

$$P_{11} = G_{11} - G_{12} G_{22}^{-1} G_{21} \tag{3.38}$$

The determinant of the partially controlled plant is

$$\det P_{11} = \det(G_{11} - G_{12} G_{22}^{-1} G_{21}) \tag{3.39}$$

Using Schurs formula gives (and assuming that $\det G_{22} \neq 0$)

$$\det P_{11} = \frac{\det G}{\det G_{22}} \tag{3.40}$$

The determinant of a system can be written in terms of the zero and pole polynomial, (MacFarlane and Karcanias, 1976):

$$\det G = c\frac{z(s)}{p(s)} \tag{3.41}$$

$$\det G_{22} = c_{22}\frac{z_{22}(s)}{p_{22}(s)} \tag{3.42}$$

where z and p are zero and pole polynomial (may not necessarily be relatively prime) and c is a constant. Thus

$$\det P_{11} = c/c_{22} \frac{z}{z_{22}} \frac{p_{22}}{p} \tag{3.43}$$

Since the roots of p_{22} has to be roots in p, they will not contribute to any zeros in $\det P_{11}$. Secondly since we assume that G_{22} is used for perfect control it can not have any RHP-zeros. Which means that any RHP-zeros in $\det P_{11}$ will have to come from z, e.g. from G.

Chapter 4

Control of Reactor, Separator and Recycle
Part I: Liquid phase systems

T. Larsson, S. Skogestad and C. C. Yu

Parts of this work where presented at NPCW-1998 Stockholm, CAPE Forum 99 Liége and AIChE annual meeting 1999, Dallas.

We consider the control structure selection for a simple plant with a liquid phase reactor, a distillation column and recycle of unreacted reactants. This class of plants has been studied extensively in the literature. Luyben has used it as a basis for his generic rule, "fix a flow in a liquid recycle loop". In this paper we shows that the basis for his rule is questionable. This is based on a systematic selection of the control structure, by using the concept of self-optimizing control.

4.1 Introduction

A common feature of many chemical processing plants is the presence of recycle of material. The most common material recycle is to recover unreacted reactant in from the reactor effluent flow and to recycle it back to the reactor. Figure 4.1 is a simple representation of such systems. It is a plant with a reactor, a distillation column and recycle of unconverted reactant. This system is simple but complex enough to capture some of effects encountered in real plants. This is explains why variations of Figure 4.1 has been extensively studied in the literature and why it is studied here.

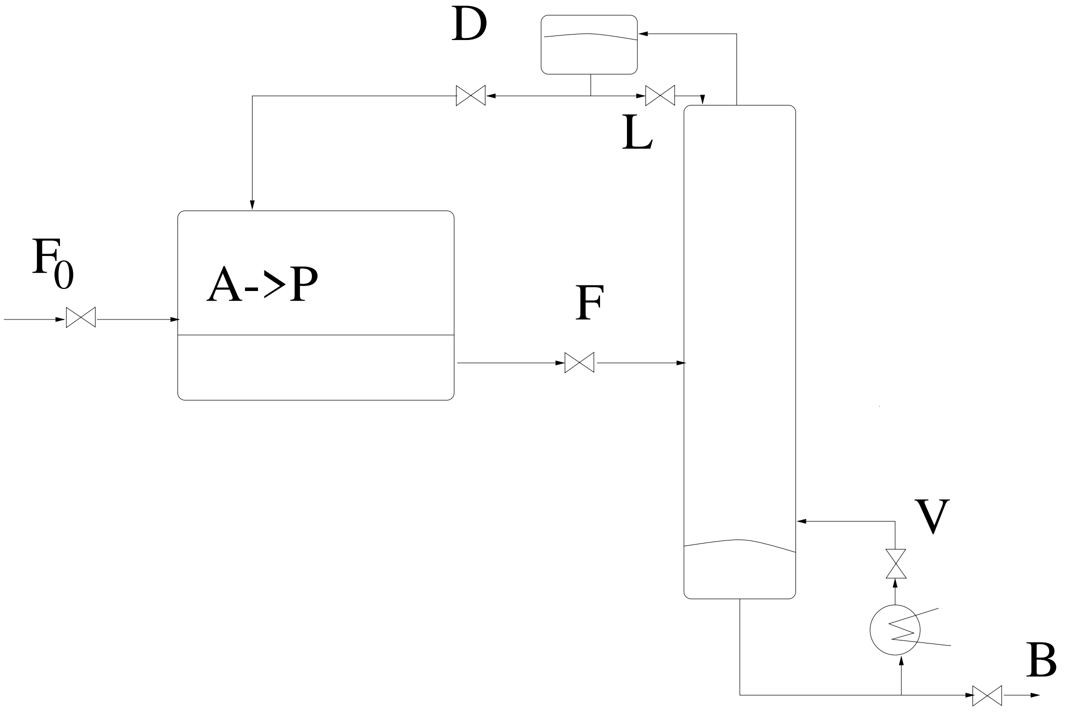


Figure 4.1: A liquid phase reactor with separator and recycle.

Gilliland *et al.* (1964) used it to illustrated how the dynamics and steady state behavior where changed by recycle. Luyben (1993a, 1993b and 1994) followed up Gilliand's points and described the high sensitivity in the recycle flowrate to the feed-flowrate. An effect for which he used the term "snowball effect". He deemed the high gain to be unfavorable. As a remedy he proposed a generic rule: "fix a flow in a liquid recycle loop".

Wu and Yu (1996) followed up the work of Luyben (1994) and showed that for a fixed reactor effluent flow, the reactor hold-up has a high gain for feed-rate changes. Their solution was to control reactor composition.

Price and Georgakis (1993) used the integrated absolute error of the product composition to rank different control structures. They found that control of internal compositions, either distillate or reactor composition, helps the control of bottom composition.

The above works rises some issues that need to be studied further. First, the "Luyben" rule has not been properly substantiated. Second, in most of the above works, the overall control objective for the plant was not clearly defined. Third, a liquid phase reactor should normally be operated at maximum volume (liquid level) in order to optimize economics, Wu and Yu (1996) and Luyben propose to let reactor level float. This has an impact on the steady state economics, an issue that has been overlooked by most researchers so far.

The recycle plant in Figure 4.1 has four degrees of freedom at steady-state (one for the throughput (feed-rate), one for the reactor and two for the distillation column), see Table 4.2. In the literature several alternative sets of controlled variables has been proposed for the case with given feed rate F_0:

1. "Conventional": Control of M_r, x_D and x_B.

2. "Luyben": Control of F, x_D and x_B, (Luyben, 1994). (In the scheme proposed by Luyben the reactor level is controlled by using the feed rate F_0, and the setpoint of this loop is used to achieve the correct feed rate.)

3. "Balanced scheme": Control of x_r, x_D and x_B, (Wu and Yu, 1996).

4. "Self-optimizing": Control of M_r, L/F and x_B, (this paper).

where M_r is reactor holdup, x_D is distillate composition, x_B is bottom composition, x_r is reactor composition, F is reactor effluent and L is reflux in the distillation column.

Note that Luyben's scheme and the balanced scheme the reactor level is uncontrolled (floating), which at first may seem strange. However, the reactor level can be stabilized in two indirect ways:

1. The composition in the reactor depends on the conversion (kx_rM_r), thus by controlling the reactor composition x_r the conversion (per pass) is fixed and reactor level is stabilized.

2. The (total) conversion may be fixed by recycle the unreacted A. This may be achieved by controlling the compositions in the distillation column (x_D and x_B) then any unreacted A will be recycle back to the reactor and since the total conversion is given the reactor level is stabilized.

The main objective of this paper is to study schemes such as the ones listed above. The difference between them is which variables that are to be controlled. This issue is studied in a systematic manner by using the concept of self-optimizing control, (Skogestad, 2000). A brief introduction to self-optimizing control is given in the next section.

4.2 Procedure for selecting controlled variables

As mentioned above (see Table 4.2), there are four degrees of freedom at steady state, and their value should be selected to optimize the plant economics. At the economic optimum, some of the degrees of freedom will be used to satisfy constraints, but normally some of them are unconstrained. The main issue is what to do with these unconstrained degrees of freedom. That is which variables should be controlled? If there were no uncertainty or disturbances, then every choice would give the same economics. However there will always be uncertainty and disturbances, and each set of controlled variables will give a different shape of the objective function around the optimum. An ideal set will yield a "flat" objective function, so that an implementation error will give a small loss from the economic optimum, (Skogestad, 2000). More precisely, we will consider the economic loss imposed by keeping a given set of variables constant.

From (Skogestad, 2000) we adopt the following procedure for selection of the controlled variables.

Step 1: Degrees of freedom analysis. Determine the degrees of freedom available for steady-state optimization. The easiest way is to count the number of inputs and subtract the number of variables without steady state effect, which has to be, controlled (e.g. non reacting levels).

Step 2: Cost function and constraints. Define the optimal operation problem by formulating a scalar cost function J to be minimized, and specify the constraints.

Step 3: Identify the disturbances (uncertainty). These may be caused by errors (uncertainty) in the assumed (nominal) model, disturbances or implementation errors in the controlled variables.

Step 4: Optimization. The optimization problem is solved both for the nominal case and for the identified range of disturbances.

Step 5: Identify candidate controlled variables c. Active constraints should be controlled, especially if there are a measurement associated with them. To select between the remaining candidates we proceed to step 6.

Step 6: Evaluation of loss. Compute the mean value of the loss for alternative sets of controlled variables. This is done by evaluating the loss, which is the difference between the cost when fixing the optimal cost and

$$L_u(u(c_s), d) = J(u(c_s), d) - J_{\text{opt}}(d) \tag{4.1}$$

with fixed setpoints c_s as a function of the disturbances d. Controlled variables c, which results in a small loss are preferred.

Step 7: Further analysis. Normally several candidates gives an acceptable loss, and further analysis may be based on a controllability analysis.

4.3 Selection of controlled variables

In this section use the concept of self-optimizing control, introduced above, to select the controlled variables for the recycle plant. We will look at two different objectives: minimization of operation cost and maximization of the production rate. Table 4.1 summarizes the results.

	Minimization of operation cost		Maximization of through-put	
Step 1: Degrees of freedom analysis. (see Table 4.2)				
	Degrees of freedom at steady state	4	Degrees of freedom at steady state	4
Step 2: Cost function.				
	Objective function		Objective function	
	minimize V		maximize F_0	
	Constraints		Constraints	
	Reactor level $M_r \leq 2800$		Reactor level $M_r \leq 2800$	
	Product quality $x_B \leq 0.015$		Product quality $x_B \leq 0.015$	
	Feedrate $F_0 = 460$		Vapor boilup $V \leq 1400$	
Step 3: Identify the most important disturbances.				
	Disturbances		Disturbances	
	Feed rate $F_0 \pm 20\%$		Maximum vapor boilup $V \pm 20\%$	
	Control error $\pm 20\%$		Control error $\pm 20\%$	
Step 4: Optimization.				
-	Active constraints at the optimum	3	Active constraints at the optimum	3
	M_r, x_B, F_0		M_r, x_B, V	
=	Unconstrained degree of freedom	1	Unconstrained degree of freedom	1
Step 5: Identify candidate controlled variables.				
	F, L, D, L/D, L/F, L/V, x_D		F, L, D, L/D, L/F, L/V, x_D	
Step 6: Evaluation of loss.				
	Good candidates		Good candidates	
	x_D, L/D, L/F		x_D, L/D, L/V, L/F	
	Bad candidates		Bad candidates	
	D, L, F, L/V		D, L, F	
Step 7: Further analysis.				
	Ratio control L/F is easier than composition control x_D.			
Conclusion.				
	Control: M_r, x_B, L/F		Control: M_r, V, x_B, L/F	

Table 4.1: Summary of self optimizing control.

The plant and the design data are taken from (Wu and Yu, 1996). The model is simple and it is assumed isothermal reactor, first order kinetics ($k = 0.3408$ hr^{-1}) where $A \rightarrow P$, A is the light component, constant relative volatility ($\alpha = 2$) and linearized flow dynamics ($\tau = 0.0011$ hr) in a distillation column with 20 stages and liquid feed at stage 12.

4.3.1 Given feed, minimize operation cost

Step 1: Degrees of freedom analysis There are 4 degrees of freedom at steady-state, see Table 4.2. (Feed rate F_0 is included for consistency.)

	Manipulable variables	6
	Product flow B	
	Vapor boilup V	
	Reflux L	
	Distillate D	
	Reactor effluent F	
	Feed F_0	
-	Levels without steady state effect	
	Condenser level M_D	
	Boiler level M_B	
=	Degrees of freedom at steady state	4

Table 4.2: Degrees of freedom analysis.

Step 2: Cost function and constraints The goal of the operation is to minimize the operating cost, we assume that the cost is given by the vapor boilup in the distillation column, i.e. $J = V$. There are constraints, on the reactor level and on the product quality x_B. In addition the feed rate F_0 is given.

Step 3: Disturbances The main disturbance is throughput changes, with $\pm 20\%$ of the nominal flowrate of F_0. The control error is assumed to be $\pm 20\%$ in all variables, except for the top composition x_D where it is assumed to be $\pm 20\%$ of the impurity $1 - x_D$.

Step 4: Optimization Table 4.3 shows the results of the nominal optimization.

We find that two constraints are active. Clearly bottom composition has to be at its constraint, if not we could reduce the vapor boilup and still satisfy inequality constraint. Since there are no penalty involved in increasing the reactor holdup, M_r, it is clearly optimal to keep it at its upper bound to maximize the conversion "per pass". Maximum conversion will reduce the need for recycle, and therefore reduce the load to the distillation column.

Since the feed rate F_0 is given, we are left with one unconstrained degree of freedom.

	Case 1: Min. V	Case 2: Max. F_0	
Feed rate F_0	460	493	[kmole/hr]
Reactor effluent	963	1097	[kmole/hr]
Vapor boilup V	1216	1400	[kmole/hr]
Reflux L	713	796	[kmole/hr]
Distillate D	503	604	[kmole/hr]
Top composition x_D	0.80	0.82	[mole A/(mole A+P)]
Bottom composition x_D	0.015	0.015	[mole A/(mole A+P)]
Reactor level M_r	2800	2800	[kmole]

Table 4.3: Nominal optimization results for the two cases.

Step 5: Candidates for control Active constraints should be controlled to optimize operation, i.e. M_r and x_B should be controlled. Note that this rules out the Luyben scheme from being candidates for self-optimizing control.

Some of the possible candidates for the remaining degree of freedom are listed in Table 4.1. It could be tempting to set V directly. But if it is set to at a too low value, we will have infeasible operation, and if it is set to a too high value it does not give optimal operation.

The reactor composition x_r is not a candidate for control. It is given by the total conversion of A. This can be shown from a total mass balance of component A (and $F_0 = B$):

$$x_r = \frac{F_0(x_0 - x_B)}{kM_r} \tag{4.2}$$

Since the feed (F_0 and x_0) is given, and x_B and M_r are controlled at their constraints, x_r has to float. This removes the balanced scheme as a candidate for self-optimizing control.

Step 6: Evaluation of the loss Figure 4.2 shows the losses due to control error and the disturbance. Control of x_D, L/F or L/D gives small losses, while D, F, L and L/V gives larger losses. If we should follow Luyben's rule then we would have to select either D or F. But as can be seen from Figure 4.2 fixing one of these variables would lead to an overload of the distillation column.

Step 7: Other considerations Since L/F gives a much easier control problem for the distillation column it will be preferred before control of x_D.

4.3.2 Maximize the feedrate

We now consider the case where the feedrate is free to choose, but the column has a capacity limitation. This may be more realistic in practice, and certainly of more interest from an economic point of view.

Step 1: Degrees of freedom analysis As before, there are four degrees of freedom at steady state, see Table 4.2.

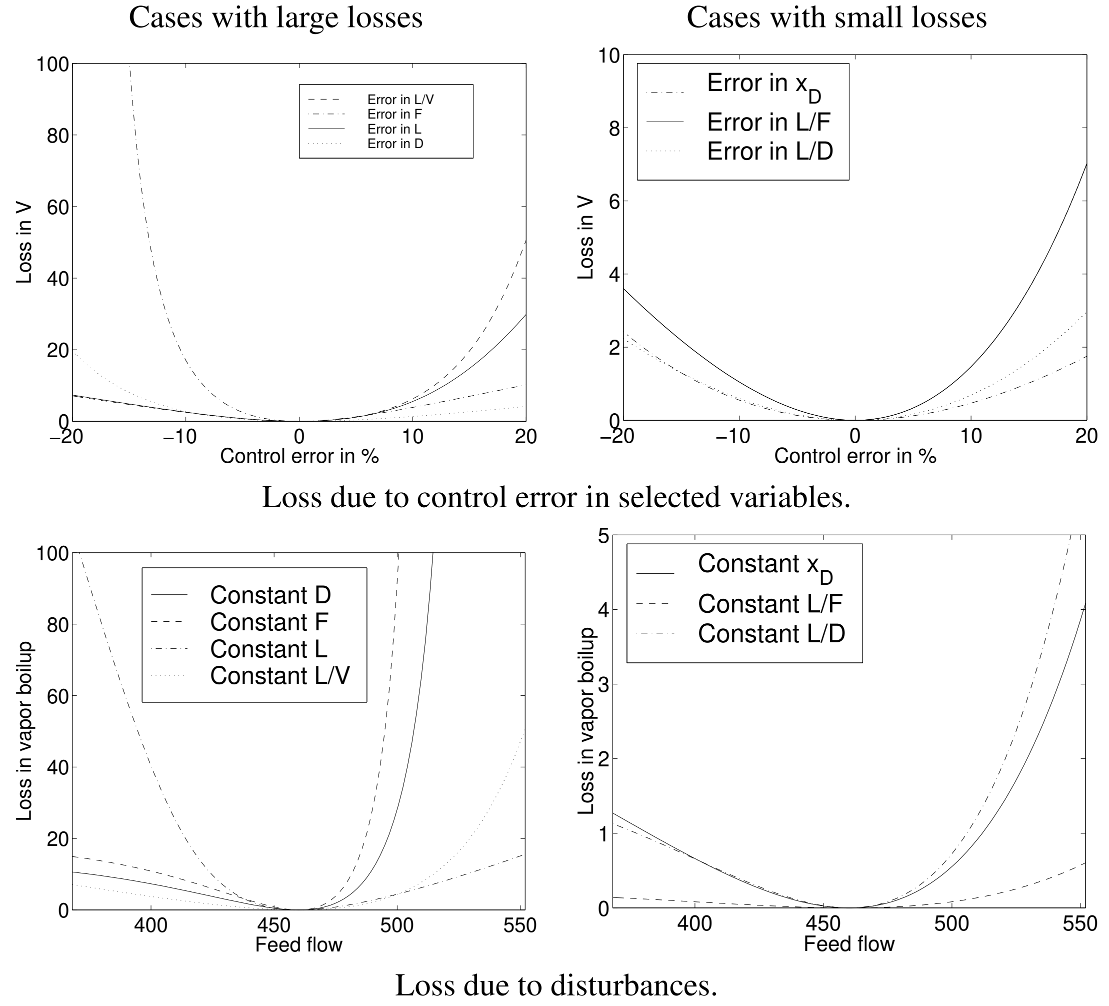


Figure 4.2: Losses in V for the case with given feed F_0.

Step 2: Cost function and constraints The goal here is to maximize the production rate, i.e. $J = -F_0$. In addition to the other constraints the capacity of the distillation column is limited (e.g. maximum vapor boilup).

Step 3: Disturbances The boilup V is in practice set indirectly, and the available V will not be constant.

Step 4: Optimization Table 4.3 shows the results from the nominal optimization. Vapor boilup, reactor level and product composition are constrained at the optimum.

This leaves one unconstrained degree of freedom. To understand this let us assume that we start with a feed flow of $F_0 = 350$ kmole/hr, which gives a certain recycle flow D. At low production rates there is almost a linear relationship. But as production is increased, the load to the distillation column increases (Fx_r increases) and recycle purity x_D decreases. As

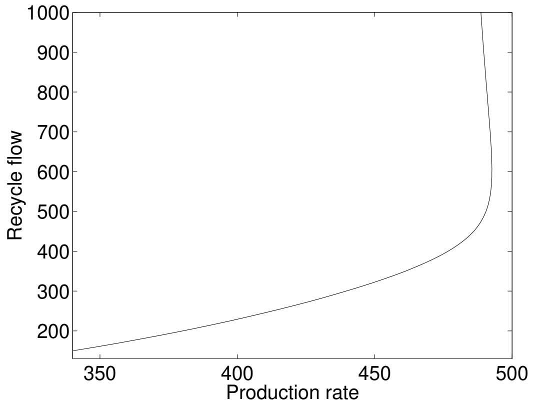


Figure 4.3: Distillate flow as a function of production rate.

x_D decreases, more D is needed to keep up the production, and at $F_0 = 493$ kmole/hr the production rate can not be increased any further. This is shown in Figure 4.3.

Step 5: Candidates for control Some of the possible candidates for control are listed in Table 4.1. Reactor composition should be as large as possible, and it is therefore not a candidate for control.

Step 6: Evaluation of the loss Figure 4.4 shows the losses in production rate due to disturbance and control error. (Note that if V is constant, then constant L, L/D etc. are the same as constant D.) Control of x_D, L/F, L/V and L/D give small losses, while D, F and L give larger losses.

Comment: It is somewhat surprising that the same variable L/F turned out to be the best variable to control for both cases.

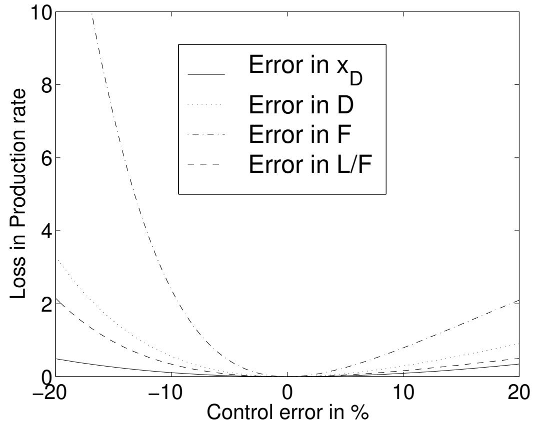
4.4 Comparisons to previous literature

4.4.1 The conventional approach

The conventional control configuration has very good self-optimizing properties. However the analysis shows that control of the internal composition x_D is not really needed. From a self-optimizing point of view x_D and L/F performs equally well. The latter is preferred because online measurements are costly and two-point control is often difficult.

4.4.2 The snowball effect and the Luyben rule

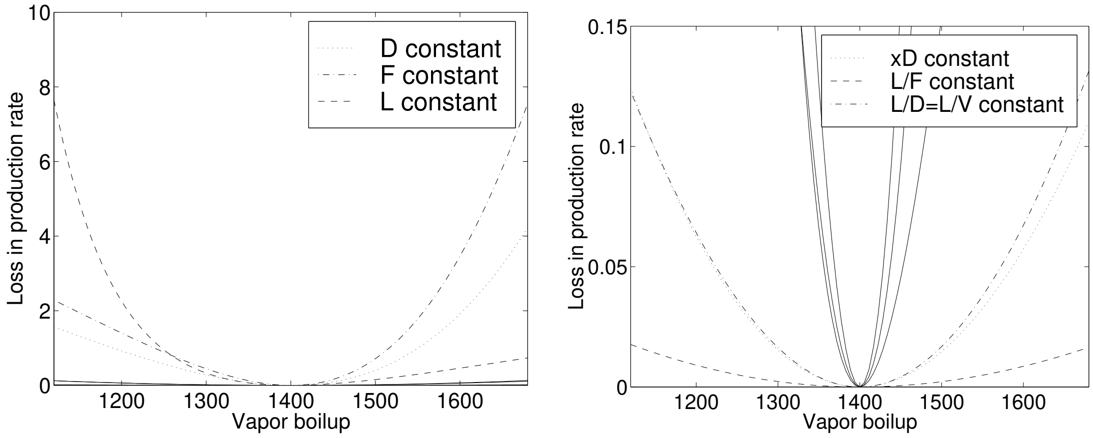
Luyben (1994) states "the use of a conventional control structure resulted in a 100% increase in the recycle flow rate for a 10% increase in the fresh feed flow rate. Such large changes are very undesirable because columns can only tolerate a limited turn-down ratio." (In the



The loss due to control errors.

Cases with large losses

Cases with small losses



The loss due to disturbances.

Figure 4.4: Loss in production due to control errors and disturbances.

conventional structure reactor level, bottom and top composition is controlled.) His remedy was to fix the reactor effluent flow F and to let reactor holdup be controlled with F_0. However his solution has some serious limitations which will be discussed here.

From a material balance we find (Luyben, 1994):

$$D = \frac{F_0(F_0(x_0 - x_B) - kM_r x_B)}{kM_r x_D - F_0(x_0 - x_B)} \tag{4.3}$$

and we see that as if the reactor holdup is too small compared to F_0, more precisely if

$$M_r \leq \frac{F_0(x_0 - x_B)}{kx_D} \tag{4.4}$$

then the recycle rate D will approach infinity. This is what Luyben (1994) refers to as the snowball effect.

We found above that both for the case where we minimize the boilup (given F_0) and the case of maximization the feed rate, it is optimal to have the reactor level M_r at its maximum value. In the Luyben scheme the level will "float".

Later, Luyben *et al.* (1997) proposed a more generic rule:

> *Fix a flow in every recycle loop.* [...] This is a simple and effective way to prevent potentially large changes in recycle flows if all flows in the recycle loop are controlled by levels. [...] the plants separation section is not subjected to large load disturbances.

We have already discussed that this rule is not economical optimal if it results in a scheme where the reactor level floats. However, there is also the possibility that we fix a flow in recycle and control the reactor level M_r at its maximum value. This is the scheme was studied above, and we found that it had poor self-optimizing (Figure 4.2), some of them where not even able to reject the disturbances. We thus conclude that the Luyben rule should not be recommended on a generic basis. But does it have advantages from a controllability point of view? This will be the topic Section 4.5.

Remark on the snowball effect: Thus the "snowball" effect is really not an issue in the operation of the plant. Rather, it is a design issue, since the reactor size is an unconstrained optimization variable during the design phase. It may be an option to over-design the reactor to avoid large fluctuation in recycles for throughput changes.

4.4.3 The balanced scheme

Wu and Yu (1996) indicate that with their control structure, i.e. control of reactor composition, it will be possible to operate the plant over a wider range of feedrates. However, we have shown that maximizing the reactor holdup is necessary to achieve maximum production. These two views are contradictory and the argument of (Wu and Yu, 1996) is correct only if we do not include any constraints. However, in practice there will always be a hard constraint on the reactor level.

4.5 Controllability analysis of the liquid phase system

Consider the case where the feed rate is given. In this section we will look on the controllability analysis of some schemes. (We will in this section assume a level is perfectly controlled if it is controlled.)

We will look at the following cases (the pairing is based on the relative gain array):

Conventional $M_r - F$, $x_D - L$, $x_B - V$, $M_D - D$ and $M_B - B$.

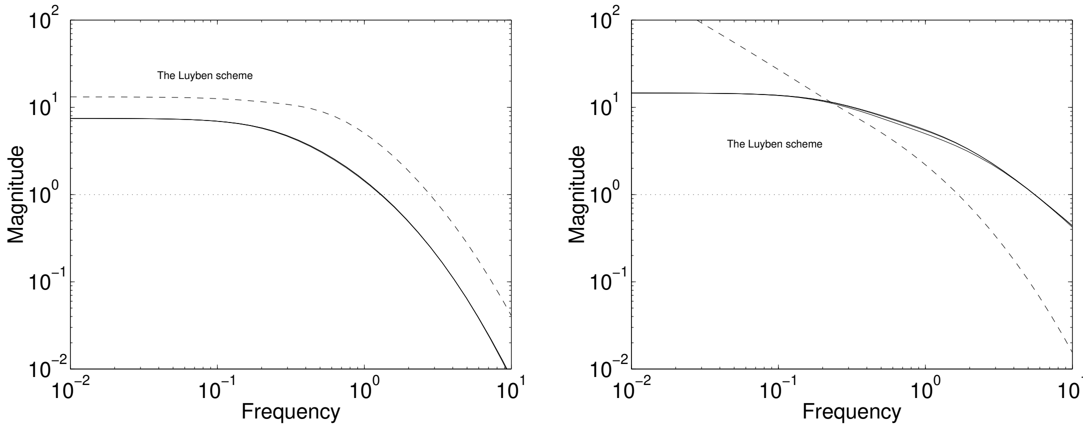
Luyben scheme F, $x_D - L$, $x_B - V$,$M_D - D$ and $M_B - B$.

Luyben rule F, $M_r - D$, x_B, $M_D - L$ and $M_B - B$.

Self-optimizing $M_r - F$, L/F, $x_B - V$, $M_D - D$ and $M_B - B$.

Note the difference between the Luyben scheme and the "Luyben rule". In the former the reactor level is floating and in the former it is fixed. In both cases we have applied the Luyben rule by fixing a flow in the liquid recycle.

From Figure 4.5 we see the closed loop disturbance gain, (Hovd and Skogestad, 1992). We note only the Luyben scheme is significantly different, and has better ability to reject disturbance in feed rate.



(a) Disturbance in feed composition x_0

(b) Feed rate changes

Figure 4.5: Closed loop disturbance gain.

This is confirmed in the nonlinear simulations, see Figure 4.6. The Luyben scheme rejects the disturbance in F_0 easily, this is related to the use of the reactor holdup as a buffer for feed rate changes, and not due to the Luyben rule. If we control both the reactor level and the reactor effluent flow (denoted by Luyben rule on the figure) the system is not even able to reject the disturbance. This is explained by the steady state analysis in the previous section: As the feedrate is increased, less D is recycled to maintain the reactor level, and therefor light component A will accumulate in the distillation column (e.g. V and L will approach infinity and the column will "saturate").

4.6 Conclusion

This paper has looked at the selection of controlled variables for a recycle plant. The basis for the selection has been the concept of self-optimizing control. A very central point is to base the selection on the steady state economics of the plant. We studied two different objectives:

1. Minimization of the operating cost, which was the boil-up.

2. Maximization of production rate.

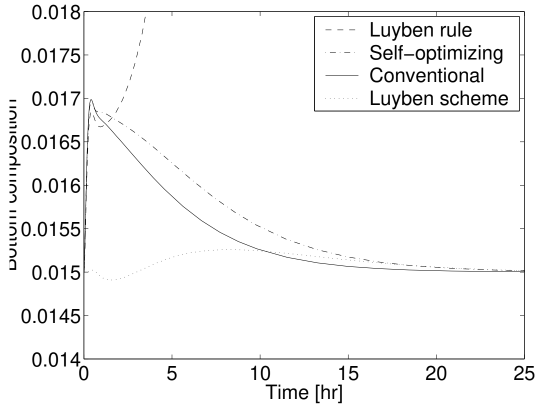


Figure 4.6: A 20% step in F_0.

In both cases there was no cost associated with increasing hold-up. Therefor it should be controlled at its constraint, to maximize the conversion "per-pass". In both cases control of x_D or L/F gives small economic losses.

We have shown that the basis given by Luyben (1993*c*) for the generic rule: "one flow rate somewhere in the recycle loop should be flow controlled", can lead to excess load in the distillation column (i.e. the opposite effect of what the rule seeks to avoid).

However, dynamic simulations showed that the scheme proposed by Luyben (1994) performs better than the conventional scheme and the self-optimizing scheme for disturbances in the feed rate. This is not due to fixing a flow in the liquid recycle, it is due to the buffer effect by the floating reactor level. The scheme with both fixed reactor effluent and level where not even able to reject a 20 % increase in feed rate. If one allows for a non-optimal control structure, like the Luyben scheme, it may be better to control the reactor level at its constraint and to over-purify the product.

In a gas phase system the situation would be different, there are a cost associated with reactor hold-up (pressure) so it is not given that optimal reactor holdup will be at its constrained. These systems are more common in the industry, and will be the subject of Part II of this work (Chapter 5).

Chapter 5

Control of Reactor, Separator and Recycle
Part II: Gas phase systems

T. Larsson and S. Skogestad

Presented at the AIChE annual meeting 1999, Dallas.

Abstract

This paper has looked at the selection of controlled variables for recycle processes. We looked at two main cases, a simple gas phase system and a more realistic case, the methanol synthesis loop. The paper is a continuation of work done on a simple liquid phase reactor. Our focus has been on the selection of controlled variables.

Control of recycle-rate, purge fraction, or reactor pressure gives a system with good self-optimizing properties. This is linked to the behavior of these variables as conversion increases. As expected purge flow is a bad alternative as a controlled variable. More unexpectedly, inert composition in the recycle turned out to have bad self-optimizing properties. This is also explainable by the behavior of this variable when conversion is increased.

The results for the simple gas phase reactor carries well over to the methanol case study.

5.1 Introduction

A common feature of many chemical processes is the presence of recycle. One example of mass recycle is a reactor, separator and recycle of unreacted reactant, see Figure 5.1. This configuration is very common in industry, some common examples are ammonia and methanol plants.

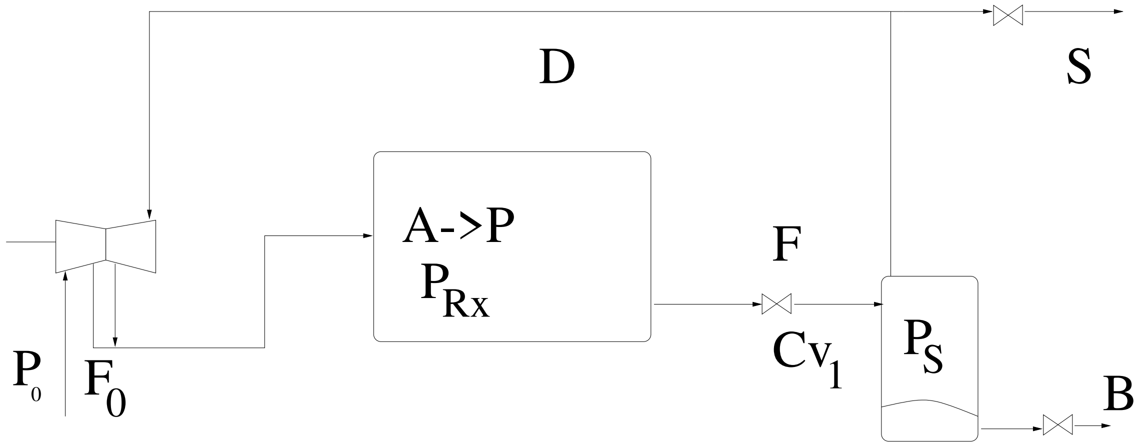


Figure 5.1: A gas phase with separator and recycle.

In part one of this paper we studied a similar plant. That plant had a liquid phase reactor, a distillation column as a separator and no inert. Even thought gas phase reactors are more common in the industry, the liquid phase plant has been the most studied. In many aspects the difference between the two systems is small, but we will show that there is a significant difference for the selection of control structure.

The gas phase plant has not been studied to the same extent by academic researchers. Hansen (1998) studied the methanol synthesis loop. He used a complex model for an interesting controllability analysis. Unfortunately, he did not discuss the selection of outputs. Loe (1994) discusses the industrial practice for control of ammonia plants, and concludes that it may not be necessary to control the pressure.

Fisher *et al.* (1988*b*) give some heuristics for plant control. The most commonly applied seems to be: “keep the gas recycle flow constant at its maximum value”. The reason is to maximize the yield. They used the HDA-plant as an example.

In this work we study first a simple system where component A is converted to P, and then an industrial example, the methanol synthesis loop. The focus is on the selection of controlled variables. To address the issue in a systematic manner we will use the concept of self-optimizing control (Skogestad *et al.*, 1999). This involves searching for the variables, which when kept constant give the minimum operating cost. The steps that are involved in this procedure is given in Chapter 4.

5.2 The simple gas phase system

The main modeling assumptions are:

- Isothermal reactor (no energy balance).
- Sharp split in the separator, reactant A and inert in the gas phase are separated from the liquid product P.
- Ideal gas law.
- The reaction rate of $A \rightarrow P$ is given by kx_rM_r, where M_r is the molar holdup in the reactor, x_r is the composition of A in the reactor and k is the reaction rate.
- The reactor effluent flow to the separator is driven by the pressure difference, i.e. $F = C_v\sqrt{|P_{rx}^2 - P_s^2|}$. Where P_{rx} is reactor pressure and P_s is separator pressure. This is assumed to be the only pressure drop in the loop.
- The compressor work is modeled as isothermal ideal gas compression.
- The compressors are mounted on a common shaft. This is implies that $w_{recycle}/w_{main}$ is constant, where w is the molar work of either the main or recycle compressor.

We have assumed that both the feed flow rate F_0 and up-stream pressure P_0 are independent variables (disturbances). This is not really possible, since a changes in F_0 must be followed by a change in pressure P_0 (the pressure P_0 represents some local up-steam gas hold-up). So implicitly we have assumed that there is a control system in place (i.e. the main compressor is adjusted to keep P_0 constant).

Step 1: Degrees of freedom We have assumed that feed rate and up-stream pressure P_0 is given, this implies that the main compressor work is given and therefore also the recycle compressor work (common shaft). Since any valves in the loop should be fully open, we only have two manipulated variables. One of these, B, is used to control the liquid level in the separator. There is one degree of freedom at steady state.

Step 2: The objective function The goal of operation is to maximize the values of the product stream diminished by the value of feed streams and utility costs. For the liquid phase plant there was only one feed and only one exit stream, hence during the optimization we need only to minimize the cost of utility. However for the gas phase system we have two product flows with different price. The objective function to be maximized is then

$$J = \$_B B + \$_S S - \$_{F_0} F_0 - Utility \tag{5.1}$$

Where B is the product flow, S is the purge flow, F_0 is the feed rate. Using the total mass balance, $S = F_0 - B$ gives

$$J = (\$_B - \$_S)B - (\$_{F_0} - \$_S)F_0 - Utility \tag{5.2}$$

F_0 is given and may be removed form the objective. The utility is compressor work, introduction of a relative cost gives:

$$J = B - w_r\left(W_1 + W_2\right) \tag{5.3}$$

where w_r is the relative cost of compressor works compared to $\$_B - \$_S$. W_1 and W_2 compressor work is

$$W_1 = F_0 RT \ln \frac{P_r}{P_0} \tag{5.4}$$

Unlike the liquid phase system, there is a penalty for increasing the hold-up (pressure) in the reactor. If the upstream pressure is higher than P_r then there would be no need for the main compressor and this term would vanish. The recycle compressor work is given by

$$\begin{aligned} W_2 &= DRT \ln \frac{P_r}{P_S} \\ &= DRT \ln \frac{P_r}{P_r - \Delta P} \end{aligned} \tag{5.5}$$

From this equation we see that the pressure drop should be as small as possible, thus any unnecessary pressure drop should be avoided (fully open valve). Furthermore as the reactor pressure is increased, this term will decrease. Very often it will be negligible compared to the main-compressor work. There will only be a small incentive for an independent shaft for the compressors. They are often mounted on a common shaft, which will be the case in this study (see section 5.3.3).

The last term in 5.3 represents the valuable product. Hence it should be as large as possible. Production rate can be increased by increasing the reactor pressure and there by decreasing the purge. This gives an increase in compressor work, the optimum may be unconstrained.

Step 3: Disturbances We assume that throughput is the main disturbance here: $F_0 = 1000$ mole/s $\pm 20\%$.

Step 4: Optimization Figure 5.2 shows the contributions to the objective function J. As purge flow is reduced, we produce more valuable product. But the pressure and compressor work increases. At full conversion, the reactor pressure and compressor work would become infinite.

In Figure 5.2(c), the objective has been plotted for several relative costs w_r (i.e. ranging from 0 to 4 mole/MJ. It shows how the optimum would move towards lower purge and higher conversion if the relative weight were reduced so that compressor costs are less. If w_r is low enough the optimum will be constrained, either by maximum pressure or by compressor load. We will study an unconstrained case, with a relative weight $w_r = 1$ mole/MJ, marked with a x in Figure 5.2(c).

Step 5: Candidates for control Our candidates for the remaining degree of freedom are: purge flow S, recycle flow R, purge flow fraction S/R, inert composition in recycle z_I and reactor pressure P_r.

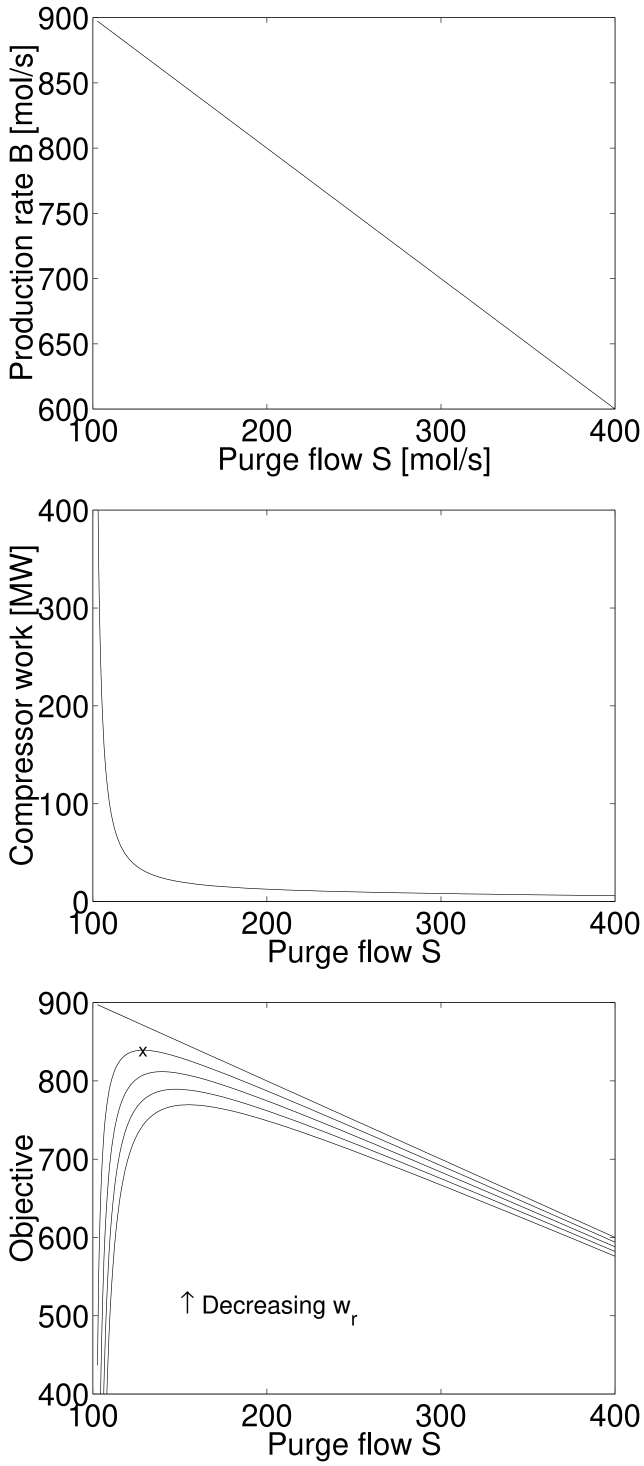


Figure 5.2: Production rate, compressor work and the objective for several relative weights.

Variable	Range		Max. loss
S	103 - 155	mole/s	330
R	2230 - 3345	mole/s	1.7
S/R	0.037 - 0.055		1.2
z_I	0.676 - 0.876		18.2
P_r	189 - 231	bar	1.0

Table 5.1: The worst loss due to control error.

Step 6: Evaluation of the losses From Table 5.1 we see that purge flow S and inert composition in purge z_I are very sensitive to the control error. Let us explain why, Figure 5.2 shows that a small error in purge flow leads to a large drop in the objective. When we approach the lower limit on purge flow, the inert composition z_I approaches one. There will only a small gain between from S to z_I, which implies that an error in z_I will be present in S. Thus control of z_I does not change the shape of the objective function much. However, the reactor pressure P_r will approach infinity as S is reduced. By control of P_r it will prevent us from implementing a value of S which is close to the minimum purge flow. This effect gives a small loss for control error in reactor pressure. Similar arguments applies for R and S/R.

The losses for controlling P_r, R and R/S for feed rate changes are shown in Figure 5.3. There are only small differences between the candidates, and the final choice should be based on other considerations.

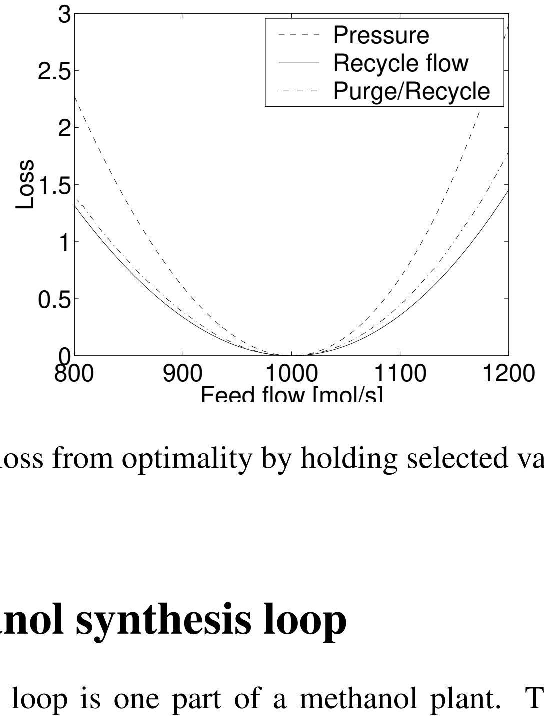


Figure 5.3: The loss from optimality by holding selected variables constant.

5.3 The methanol synthesis loop

The methanol synthesis loop is one part of a methanol plant. The feed comes from the reformer section, where synthesis gas is produced from natural gas. The product flow is sent to the distillation columns. There exist several good references that describes the methanol synthesis, (Olsvik *et al.*, 1997), (Lee, 1990) and (Skrzypek *et al.*, 1994). A brief overview of the process is given below.

5.3.1 The process and the model

The feed (which consists of CO, CO_2, H_2O and H_2) is compressed to the reaction pressure. Next it is preheated in the feed/effluent heat exchanger. In the reactor the catalyst is packed in vertical tubes, which are surrounded by boiling water. The reactor effluent is cooled, and product and reactant is separated in the flash drums. The overhead gas is divided into purge and recycle.

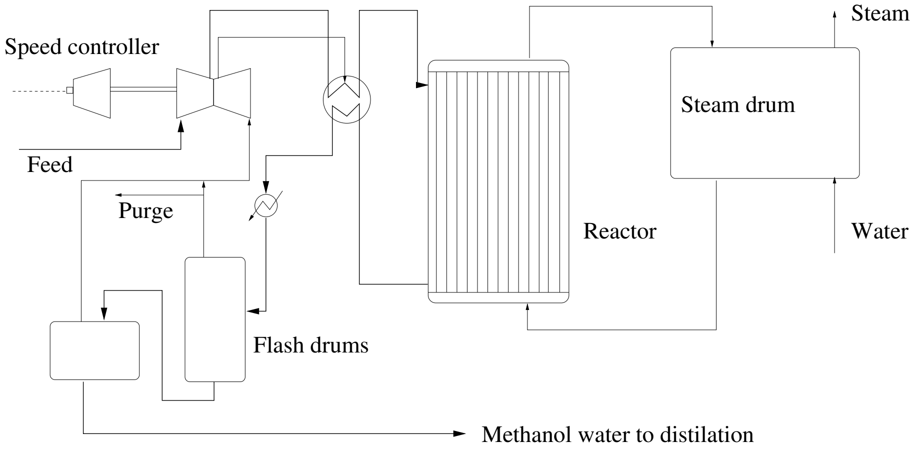


Figure 5.4: The methanol synthesis loop.

Two reactions are considered in the reactor, these are:

Methanol:	$CO_2 + 3H_2$	$=$	$CH_3OH + H_2O$	$-\Delta H_R =$	$6.09\ 10^4$ kJ/kmole
Water shift:	$CO_2 + H_2$	$=$	$CO + H_2O$	$-\Delta H_R =$	$-3.89\ 10^4$ kJ/kmole

Commercial methanol catalysts are highly selective, so side reactions are negligible.

The reactor is modeled as a series of isothermal CSTR with kinetic model from Vanden Bussche and Froment (1996).

For optimal operation it is important to avoid recycle of the product, i.e. an efficient separations between reactant (gas) and products (liquids) in the flash drum. This is achieved by cooling the reactor effluent as much as possible, i.e. the cooling in Figure 5.4 should be at its constraints. This and the fact that the solubility of CO_2 in the liquid phase is in the order of 1%, justifies the assumption of perfect separation in the flash drums. Therefor, we can ignore the heat exchangers (we have already assumed isothermal reactor).

The compressor work is modeled as polytropic compression of ideal gas. Since they are mounted on a common shaft, the work of the recycle and main compressor is related by $w_{recycle} = kw_{main}$ where w is the molar work and k is a constant. The reactor effluent flow is given by $C_V \Delta P$, where C_V is chosen to give reasonable pressure drop in the loop.

5.3.2 Selection of controlled variables

Step 1: Degrees of freedom analysis There is one degree of freedom at steady state.

Step 2: The objective function The key element for selection of controlled variables is the economic objective. In this case there is an important effect which will not be captured by steady state economic, namely the deactivation rate of the catalyst. This subject has been studied by Løvik *et al.* (1998). They showed that there is an optimal temperature profile as a function of time. Since the deactivation rate mainly depends on reactor temperature, we will assume that the reactor temperature is given. Therefor a steady state analysis will be used on the remaining degrees of freedom.

We also make one more assumption: that the synthesis loop is decomposed from the rest of the plant. There is a small recycle flow from the synthesis loop to the reformer section, but we do believe that the two-way interaction is small enough to allow for a vertical decomposition of the problem.

The objective function is

$$J_1 = \$_B B + \$_S S + \$_P P - \$_{F_0} F_0 - Utility \tag{5.6}$$

Where B is product flow, S is purge flow, P is steam, F_0 is the feed, and $Utility$ is the cost of utility (compressor work).

How are the different terms in the objective function influenced by our degree of freedom? Figure 5.5 shows the production of methanol, steam, and consumed work as a function of the purge flow-rate. As the purge flow approaches the minimum purge flow (full conversion) $S_{min} = 607$ mole/s, the reactor pressure approaches infinity. This is illustrated in Figure 5.5(c), where compression work approaches infinity.

As shown in appendix 5.A, there is almost a linear relationship between purge and product flow, and Figure 5.5(b) shows that the produced amount of steam is almost a linear function of purge. Therefor we choose to use a simplified objective

$$J = B - w_r W \tag{5.7}$$

Where B is the production of methanol, W is the compressor work and w_r is the relative cost.

Step 3: Disturbances Our disturbances are feedrate variations, inert composition and stoichiometric number in, as defined in appendix 5.A.

Step 4: Optimization Figure 5.6 shows the dependency of the objective function for different values of the relative weight w_r. For all cases except $w_r = 0$, there is a steep fall only on one side of the optimum. The smaller the value of w_r the closer the optimum is to the steep fall, which is easily explained by Figure 5.5. We will study an unconstrained optimum with $w_r = 1e^{-5}$.

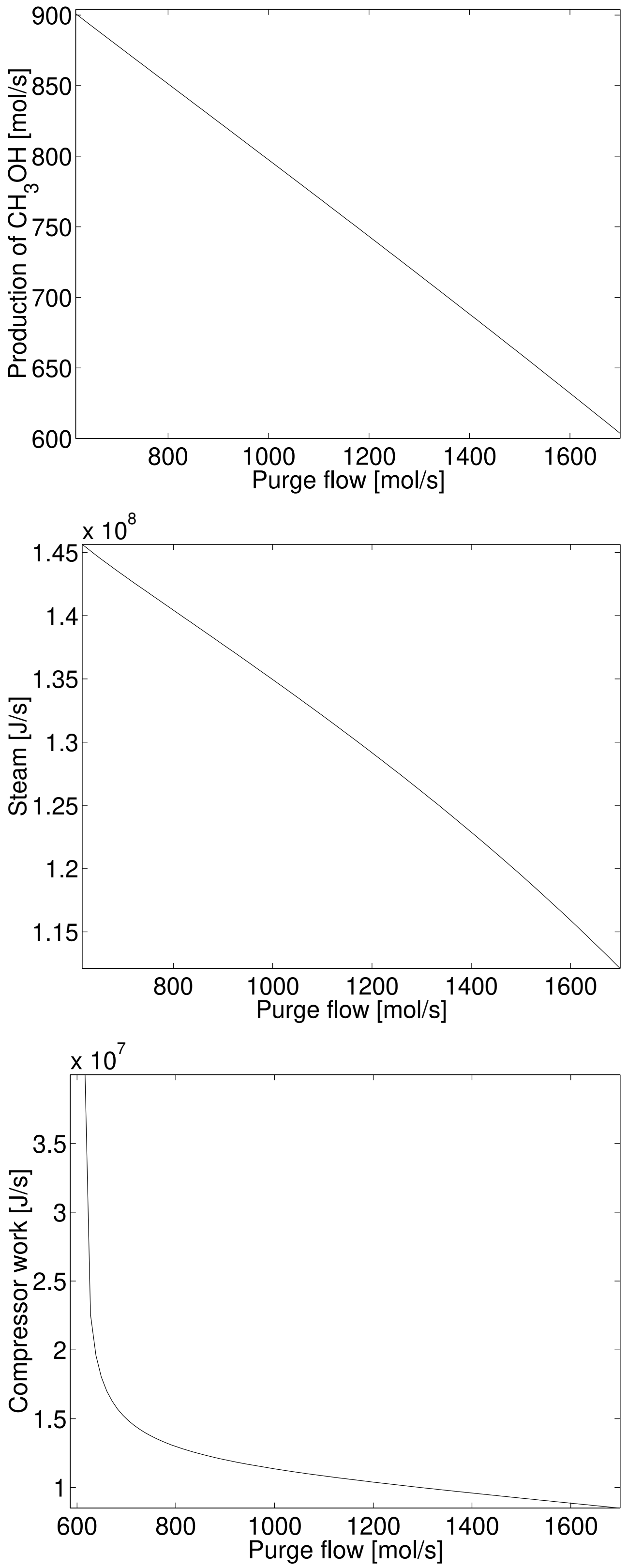


Figure 5.5: Production rate, steam production and the objective.

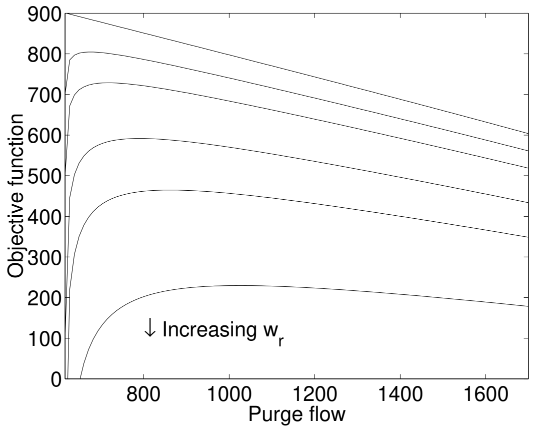


Figure 5.6: The objective for different values of the relative cost from 0 to $3e^{-5}$.

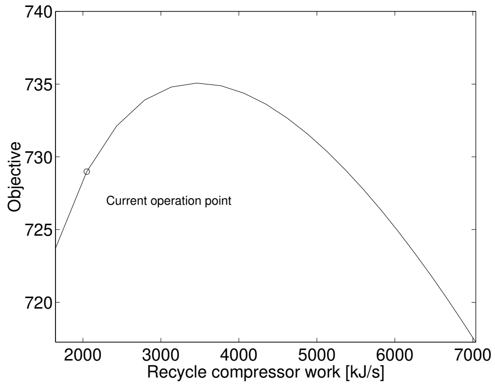


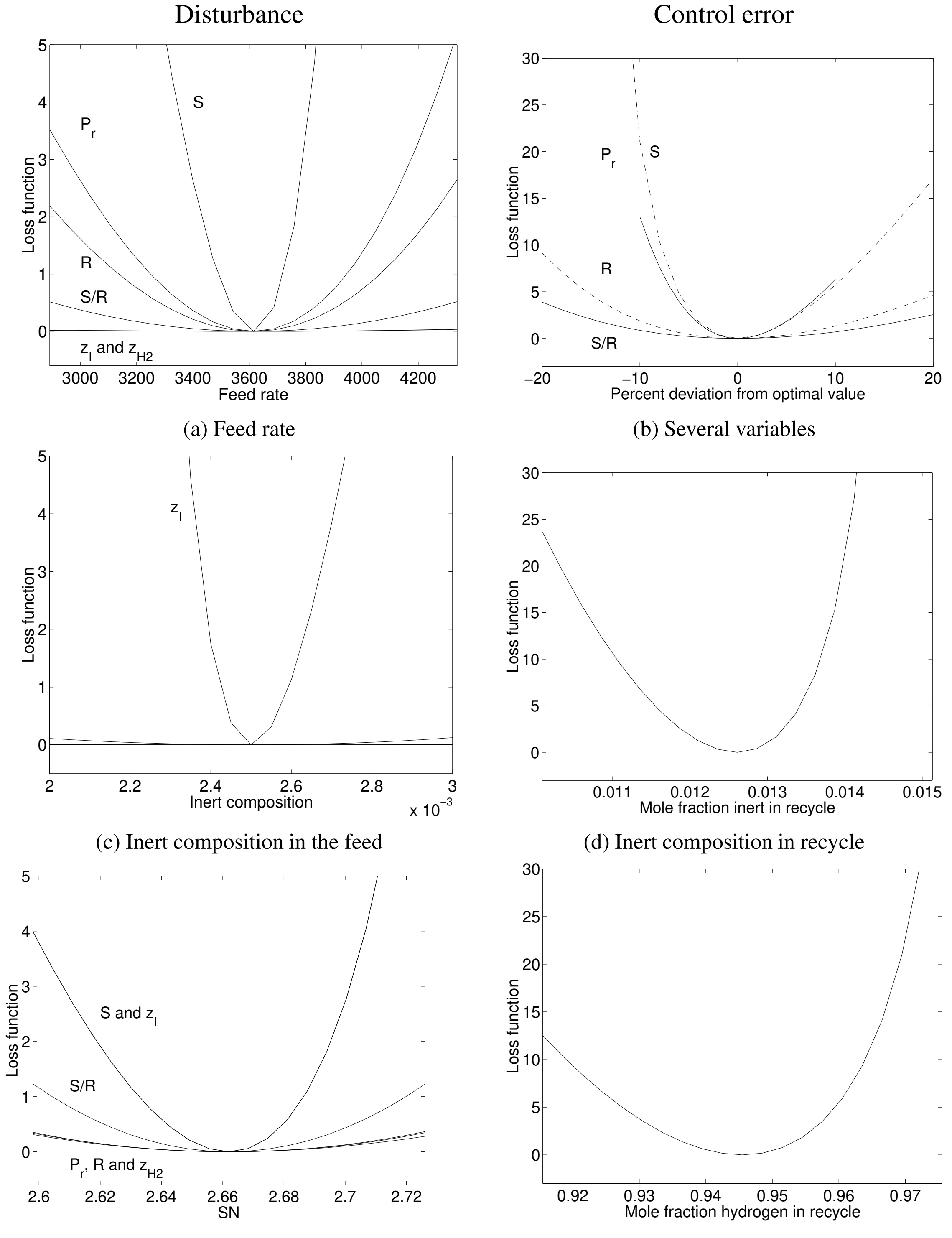
Figure 5.7: The objective as a function of recycle compressor work.

Step 5: Candidates for control Our alternatives for the remaining degree of freedom are: Reactor pressure, composition of hydrogen and inert in the recycle loop, the recycle flow, purge flow, recycle flow and purge fraction.

Step 6 Evaluation of the loss Figure 5.8 shows the loss function for the disturbances and control error. To summarize, purge flow, inert and hydrogen composition have large losses, while reactor pressure, recycle and purge fraction are good candidates for control.

5.3.3 The common shaft

To save investment costs the compressors are often mounted on a common shaft and we loose one degree of freedom for operation. This was taken into account in the above. Figure 5.7 shows how to objective could be improved if recycle work was added as an independent variable. The difference between the current operation point, and the operational optimum is so small that it will probably not be justifiable to have a separate shaft for the recycle



(a) Feed rate

(b) Several variables

(c) Inert composition in the feed

(d) Inert composition in recycle

(e) Stoichiometric number

(f) Hydrogen composition in recycle

Figure 5.8: The loss due to disturbances (left) and control error (right).

compressor. However if there where two different shafts, then we could increase the recycle compressor work. But as indicated by Figure 5.7 the optimum would most likely be at a constraint for the recycle compressor. The result obtained above would not have been changed.

5.4 Conclusion

We have considered the selection of controlled variables for recycle processes. We looked at two main cases, a simple gas phase system and the methanol synthesis process.

The gas phase and liquid phase systems are fundamentally different in economic operation. In the liquid phase system there is no cost associated with the reactor hold-up, therefore it is optimally to control it at its maximum value. In a gas phase system the situation is different, there are a cost associated with reactor hold-up (pressure). The optimum is therefore unconstrained in this variable.

Control of recycle rate, purge fraction or reactor pressure gives a system with good self-optimizing properties. This is linked to the behavior of these variables as conversion increases. As expected purge flow is a poor alternative as a controlled variable. More unexpectedly, inert composition in the recycle turned out to have poor self-optimizing properties. This is also explainable by the behavior of this variable when conversion is increased. This result was unexpected and diverts from the common view in the process control community.

The argument made by Fisher *et al.* (1988*b*) "keep the gas recycle flow at its maximum value", should not be misunderstood. We have shown that the optimal recycle flow is not at a constraint. However often there are inputs that can be used to increase recycle for a small cost, i.e. valves and recycle compressors, and these will often be at their constraint.

The results for the simple gas phase reactor carries well over to the methanol case study.

5.A Some simple relations

In order to understand some of the most basics facts of the methanol synthesis, we can derive some simple equations based on atomic balances.

$$\begin{array}{lrcl}
\text{Inert:} & Sy_{CH_4} & = & F_0 z_{CH_4} \\
\text{Carbon:} & S(y_{CO_2}+y_{CO})+Bx_{CH_3OH} & = & F_0(z_{CO_2}+z_{CO}) \\
\text{Oxygen:} & S(2y_{CO_2}+y_{CO})+B(x_{CH_3OH}+x_{H_2O}) & = & F_0(2z_{CO_2}+z_{CO}+z_{H_2O}) \\
\text{Hydrogen:} & Sy_{H_2}+B(2x_{CH_3OH}+x_{H_2O}) & = & F_0(z_{H_2}+z_{H_2O})
\end{array}$$

Where we have assumed that there is no methanol in the feed, and that the are no gases in the product and no liquid in the purge flow. For a given feed we then have six equations (the four above and two equations for summation of mole fractions) in eight variables, thus there are 2 degrees of freedom. (These equations do not say anything of the internal flows and pressures.)

For a given feed, with surplus of H_2 will have maximum production if all CO and CO_2 is converted to methanol. That is in the limiting case $y_{CO} = y_{CO2} = 0$. Then the production of methanol would be

$$x_{CH_3OH}B = F_0(z_{CO_2}+z_{CO}) \tag{5.8}$$

The minimum amount of hydrogen in the purge flow

$$Sy_{H_2} = F_0(z_{CO_2} + z_{CO})(SN - 2) \tag{5.9}$$

Where the stoichiometric number is defined as $SN = \frac{[H_2]-[CO_2]}{[CO]+[CO_2]}$. Which together with the amount of inert gives the minimum purge flow

$$S = F_0(z_{CH_4} + (z_{CO_2} + z_{CO})(SN - 2)) \tag{5.10}$$

If $SN = 2$ then we will lose no hydrogen in the purge for full conversion.

What these equation tells us is obvious, for a given feed there is an upper limit for produced methanol which is given by the amount of carbon in the feed. And there is a lower limit for purge gas, which is given by the amount of inert and surplus of hydrogen in the feed.

Purge flow is related to feed flow and product flow by (from a total mole balance)

$$S = F_0 - B(3x_{CH_3OH} + x_{H_2O}) \tag{5.11}$$

Chapter 6

Selection of controlled variables for the Tennessee Eastman problem

T. Larsson, K. Hestetun and S. Skogestad

Abstract

This chapter 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).

The following controlled variables are recommended for this problem:

- Equality constrained variables: Product composition and product-rate. (We have not assigned measurements or manipulated variables to these.)
- Optimally constrained variables: Reactor level, reactor pressure, compressor recycle valve, stripper steam valve and agitator speed.
- Unconstrained variables with good self-optimizing properties: Reactor temperature, composition of C in purge and recycle flow or compressor work.

A common suggestion is to control the inventory of inert components. However, this is a poor choice for this problem, since an unfavorable shape of the economic objective function implies that a small error can lead to infeasibility. Thus, we recommend leeting inert composition float.

6.1 Introduction

This paper addresses the selection of controlled variables for the Tennessee 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 reactions are

$$
\begin{array}{rcll}
A(g) + C(g) + D(g) & \rightarrow & G(liq) & \text{(Product)} \\
A(g) + C(g) + E(g) & \rightarrow & H(liq) & \text{(Product)} \\
A(g) + E(g) & \rightarrow & F(liq) & \text{(Byproduct)} \\
3D(g) & \rightarrow & 2F(liq) & \text{(Byproduct)}
\end{array}
$$

The process has four feed streams (of A, D, E and A+C), one product stream 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 6.1. 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 have used it to demonstrate their procedure for the design of a control system, e.g. see Chapter 2 for a review of the various approaches. We here only consider the selection of controlled variables.

McAvoy and Ye (1994) select the controlled variables in a somewhat *ad hoc* fashion. In addition to the liquid levels, they control reactor temperature, reactor pressure, recycle flow rate, compressor work, concentration of B (inert) in purge, concentration of E in product flow.

Lyman and Georgakis (1995) recommend a control structure where the following variables are controlled: Reactor temperature, reactor level, recycle flow rate, agitation rate, composition of A, D and E in reactor feed, composition of B (inert) in purge and composition of E in product. Even though they consider the operation cost for the control structure,

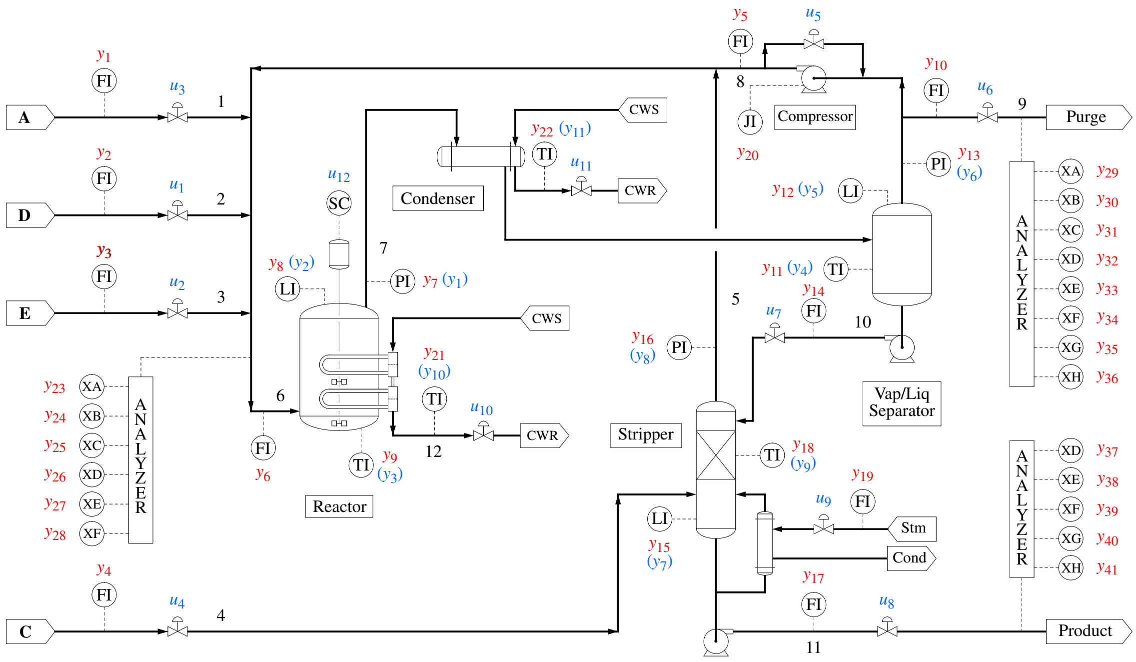


Figure 6.1: Tennessee Eastman process flowsheet

it can never become optimal since variables that should be kept at their constraints are used in control loops (like the recycle valve).

The approach of Ricker (1996) is similar to the one in this paper. First, he chooses to control the variables that optimally should be at their constraints ("active constraint control"). Second, he excludes variables for which the economic optimal value varies a lot. This is in agreement with the concept of self-optimizing control. He ends up controlling recycle valve position (at minimum), steam valve position (at minimum), reactor level (at minimum), reactor temperature, composition of C in reactor feed, and composition of A in reactor feed. He notes that it is important to determine appropriate setpoint values for the latter three controlled variables.

Luyben *et al.* (1997) (correctly) sets agitation rate and the recycle valve at their constraints. They choose to control the reactor pressure, reactor level, separator temperature, stripper temperature, ratio between E and D feedrates, A in purge, and B (inert) in purge.

Ng and Stephanopoulos (1998*b*) proposes to use a multivariable modular controller to control reactor temperature, reactor level, reactor pressure, G in product flow, stripper temperature, C in reactor feed, A in reactor feed and B (inert) in purge flow.

Tyreus (1999*a*) used a thermodynamic approach to solve the problem. He (correctly) sets the agitation on full speed, closes the steam valve and the recycle valve. In addition he controls reactor temperature, reactor pressure, reactor level and A in reactor feed and B (inert) in purge flow.

To summarize, most authors do not control all the variables that 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.

6.2 Stepwise procedure for self-optimizing control

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)

In the earlier Chapters, there was only one unconstrained degree of freedom, so the evaluation in step 6 was manageable. 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.

6.3 Degrees of freedom analysis and optimal operation

The process has 12 manipulated variables, 41 measurements and 20 disturbances. In addition, 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,

$$J = (loss\ of\ raw\ materials\ in\ purge\ and\ products) + (steam\ costs) + (compression\ costs) \tag{6.1}$$

The first term dominates the cost.

An analysis, see Table 6.1, shows 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).

	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 6.1: Degrees of freedom and active constraints.

6.4 Disturbances

A closer analysis reveals that disturbances 3, 4, 5 and 7 have no steady-state effect on the economics provided we make appropriate use of the available manipulated variables. For example, disturbance 4 (a step in the reactor cooling water inlet temperature), is easily counteracted by increasing the reactor cooling water flowrate; thus this disturbance will have no impact on the economics provided we adjust the cooling rate. Similar arguments can be made for disturbance 3, 5 and 7, provided we manipulate the reactor coolant flow, separator cooling water flow and the A+C feedrate. Disturbance 6 (loss of feed A) is considered to be so serious that it should be handled by overrides, therefore it is not included in this study.

This leaves only the following three disturbances:

- Disturbance 1: Change in A/C ratio in feedstream 4.
- Disturbance 2: Change in fraction of B (inert) in feedstream 4.

Variable	Disturbances						
	Nominal	IDV(1) A/C in stream 4	IDV(2) B in stream 4	IDV(3) D feed T	IDV(4) Rx coolant T	IDV(5) Cond coolant T	IDV(7) C header P
Cost	114	111	170	114	114	114	114
Rx. T	123	123	124	123	123	123	123
Rx. coolant flow	36.0	36.0	35.5	36.0	38.8	36.0	36.0
Sep T	92	92	93	92	92	92	92
Cond coolant flow	12	12	12	12	12	14	12
C feed flow	9.2	8.7	9.3	9.2	9.2	9.2	9.2
C feed valve	61	57	61	61	61	61	76

Table 6.2: Optimal operation for the nominal case and for disturbances.

- Throughput disturbances: Change in production rate by ± 15 %.

We use the same constraints (and safety margins) as given by Ricker (1995). Optimizing the operation with respect to the three unconstrained degrees of freedom, resulted in the same optimal values as found 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.

6.5 Selection of controlled variables

What should we control? More precisely, we have 8 degrees of freedom at steady state, and we want to select 8 controlled variables that 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.

To proceed, one approach is to select a smaller subset of candidates, for example, based on physical insight. Alternatively, one may consider the four requirements for a "good" controlled variable given by (Skogestad, 2000):

Requirement 1. Its optimal value should be insensitive to disturbances

Requirement 2. It should be easy to measure and control

Requirement 3. Its value should be sensitive to changes in the manipulated variables (alternatively, the optimum should be "flat" with respect to this variable)

Requirement 4. For cases with two or more controlled variables, the selected variables should not be closely correlated

However, these requirements require quite a lot of effort with respect to optimization, and are at the same time rather qualitative. We therefore want to find some more quantitative criteria for eliminating variables, until we are left with a manageable number.

The following criteria are proposed to reduce the number of alternatives. Most of them are rather obvious, but nevertheless we find them useful.

1. Active constraint control: We choose to control the active constraints. This reduces the number of controlled variables to be selected (in our case from 8 to 3). Of course, we must also eliminate the corresponding variables from further consideration.

2. Eliminate variables related to equality constraints

3. Eliminate variables with no effect on the economics (i.e. with no steady-state effect)

4. Eliminate/group closely related variables

5. Process insight: Eliminate further candidates

6. Eliminate single variables which with constant setpoints yield infeasibility or large loss when there are (1) disturbances (with the remaining degrees of freedom reoptimized) or (2) implementation errors.

7. Eliminate combinations (pairs, triplets, etc.) of variables that yield infeasibility or large loss

After this we enter into the final evaluation:

8. Evaluation of disturbance loss for remaining combinations

9. Evaluation of implementation loss

6.5.1 Active constraint control

As mentioned, there are 5 active constraints. 3 of the constraints are related to the manipulated variables (compressor recycle, stripper steam and agitator speed); this eliminates 3 manipulated variables and also 1 directly related measurement (stripper steam). 2 of the constraints are related to outputs (reactor level and pressure); this eliminates another 2 measurements.

We are now left with 38 measurements and 9 manipulated variables, from which we want to select 3 unconstrained controlled variables. This gives 16215 possible combinations, which is still much too large.

6.5.2 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) and should not be kept constant (eliminates 1 manipulated variables and 1 directly related measurement).
- The ratio of components G and H in the product is specified; this eliminates at least the combined use of the measurements of G and H in product.

Comment: We have not said that we shall not control throughput or product composition. The ratio of G and H has to be controlled. The throughput needs also to be set somewhere, but we will not address it here.

6.5.3 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 will measure and control these two variables for stabilization, but we have actually already accounted for that in our degrees of freedom analysis.

6.5.4 Eliminate/group closely related variables

The controlled variables should be independent (requirement 4).

- 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 measurements).
- Hestetun (1999) considered several pairs of variables and found that there are 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)

Note that the choice of which variables to keep and which to eliminate was more or less arbitrary, but since the variables are closely related it does not matter very much in the further analysis. The main idea is to keep one variable in each group of related variables.

6.5.5 Process insight: Eliminate further candidates

Based on understanding of the process some further variables can be excluded form the set of possible candidates for control:

- The pressures in separator and stripper are probably a measure of the same "quantity" namely gas holdup. We expect them to be correlated with reactor pressure and will not control them (eliminates 2 measurements).
- The condenser and reactor cooling water flowrates should not be held constant, since that would imply a loss for disturbances 4 and 5 (eliminates 2 manipulated variables). For the same reason we should not keep the reactor and separator cooling water outlet temperatures constant (eliminates 2 measurements).

- The separator liquid flow is strongly correlated with the production rate (which is specified) and should not be kept constant (eliminates 1 manipulated variable)
- The fractions of G in product and H in product should be equal (specified), so by keeping one of these fractions constant, we will indirectly specify their sum, which is optimally about 0.98. However, their sum cannot exceed 1.0, so taking into account the implementation error it is clear that we can not keep G in product or H in product constant (eliminates 2 measurements).

6.5.6 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 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 and one measurement constant results in infeasible operation for disturbance 2 (inert feed fraction): D feed flow, E feed flow , A+C feed flow (stream 4), purge valve purge rate (measurement). This is independent on how the two remaining degrees of freedom are used, see Table 6.3.

Variable	Nominal value (constant)	Nearest feasible value with disturbance 2
D feedrate [kg/h]	3657	3671
E feedrate [kg/h]	4440	4489
A+C feedrate [kscmh=k Sm^3/h]	9.236	9.280
Purge rate [kscmh]	0.211	0.351
Purge valve	26	46

Table 6.3: Single variables with infeasibility for disturbance 2

Loss. We have now 1 manipulated variable (A feed flow) and 17 measurements left. Table 6.4 shows the loss (deviation above optimal value) for fixing one of these 18 variables at a time, and reoptimizing with respect to the remaining 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.

6.5.7 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.

Fixed variable	Disturbance 1	Disturbance 2	Throughput +15/-15%
A feed flow *	709.8	6.8	
Reactor feed flow*	53.5	0.5	
Recycle flow	0.0	0.8	0.5 / 0.3
Reactor Temp.	0.0	0.9	1.2 / 0.7
Sep Temp.*	0.0	0.5	4.2 / 2.3
Stripper Temp.*	0.1	0.3	4.3 / 2.3
Compressor 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 6.4: 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.

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.

6.5.8 Final evaluation of loss for remaining combinations

As mentioned, there are 165 possible combinations of three variables. 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. Furthermore, reactor temperature is proposed by most authors, and it is normally easy to control, so we will now only consider combinations that include reactor temperature.

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 6.2). In either case, operation will be very sensitive to the implementation error for this variable.

The losses for the remaining $9 \cdot 8/2 = 36$ possible combinations of 2 variables are shown in Table 6.5. Not surprising, 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 eliminated one of them earlier).

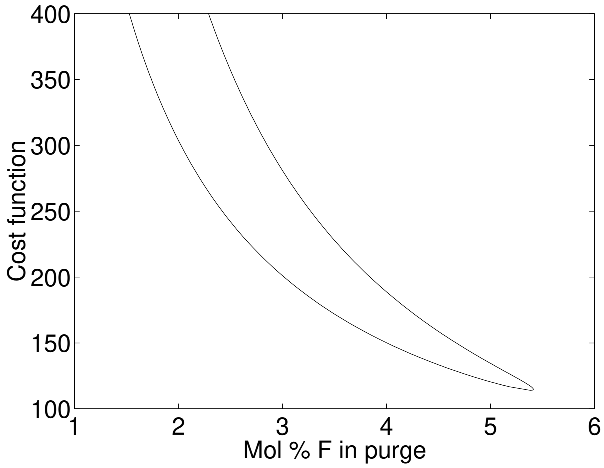


Figure 6.2: Unfavorable shape of cost function with F (byproduct) in purge as controlled variable. Shown for case with constant reactor temperature and C in purge.

We note that constant F in product 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]:

Case I. Reactor temperature, Recycle flow, and C in purge (loss 3.8).

Case II. Reactor temperature, Compressor work, and C in purge (loss 3.9).

Case III. Reactor temperature, C in purge, and E in purge (loss 5.1).

Case IV. Reactor temperature, C in purge, and D in purge (loss 5.6).

The choice of Ricker (1996), with reactor temperature, A in purge and C in purge, is somewhat less favorable with a summed loss of 9.8 $/h.

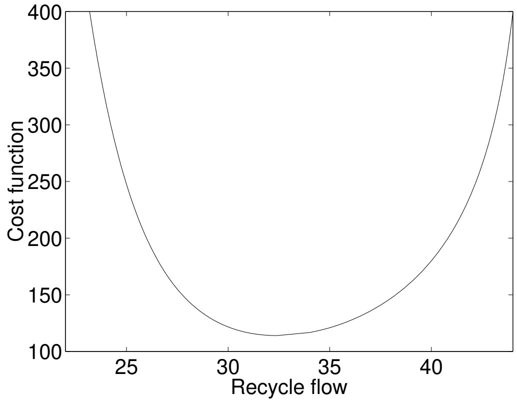
6.5.9 Evaluation of implementation loss

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. In Figure 6.3 we plot for "best" case I the cost as a function of the three controlled variables (the plots for case II are nearly identical and are not shown). We see that the optimum is flat over a large range for all three variables, and we conclude that implementation error will not cause a problem.

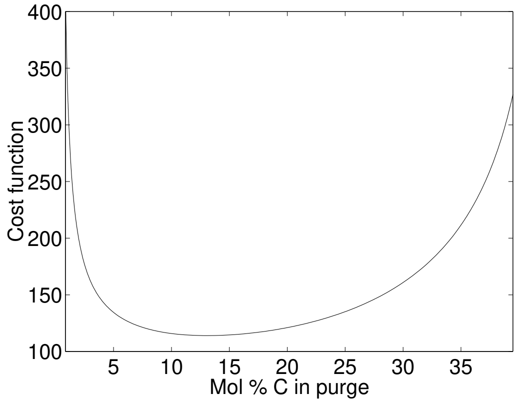
To compare, we see from Figures 6.4 and 6.5 that in cases III and IV the cost is sensitive to implementation errors, and we get infeasibility if purge composition of D (case III) or E (case IV) becomes too small.

Case	Fixed variables		Disturbance 1	Disturbance 2	Throughput +15%	Throughput -15%
	Recycle Flow	Comp. Work	0.1	Infeasible	Infeasible	40.4
	Recycle Flow	A in purge	0.0	1.2	Infeasible	9.1
I	Recycle Flow	C in purge	0.0	1.9	1.3	0.6
	Recycle Flow	D in purge	0.0	3.7	4.8	3.0
	Recycle Flow	E in purge	0.0	3.7	3.1	2.2
	Recycle Flow	D in prod.	0.2	2.6	38.0	11.9
	Recycle Flow	E in prod.	0.2	1.5	42.1	12.9
	Recycle Flow	F in prod.	0.2	37.7	1.8	0.8
	Comp. Work	A in purge	0.0	1.3	126.0	8.0
II	Comp. Work	C in purge	0.0	1.8	1.4	0.7
	Comp. Work	D in purge	0.0	4.0	5.5	3.6
	Comp. Work	E in purge	0.0	4.0	3.5	2.8
	Comp. Work	D in prod.	0.2	2.0	40.8	12.8
	Comp. Work	E in prod.	0.2	1.6	45.3	13.8
	Comp. Work	F in prod.	0.2	32.8	1.9	0.9
Ricker	A in purge	C in purge	0.0	2.4	5.3	2.1
	A in purge	D in purge	0.0	2.3	13.4	5.2
	A in purge	E in purge	0.0	2.3	10.2	4.6
	A in purge	D in prod.	0.0	1.6	50.5	10.6
	A in purge	E in prod.	0.1	1.3	54.6	11.1
	A in purge	F in prod.	0.1	17.0	4.5	2.1
IV	C in purge	D in purge	0.0	2.4	2.1	1.1
III	C in purge	E in purge	0.0	2.4	1.7	1.0
	C in purge	D in prod.	0.0	1.7	5.1	2.5
	C in purge	E in prod.	0.0	1.7	5.4	2.7
	C in purge	F in prod.	0.2	35.6	1.9	1.2
	D in purge	E in purge	0.0	2.6	77.3	Infeasible
	D in purge	D in prod.	6.2	5.4	52.6	Infeasible
	D in purge	E in prod.	5.5	Infeasible	52.2	Infeasible
	D in purge	F in prod.	0.5	Infeasible	2.4	1.0
	E in purge	D in prod.	4.5	5.3	54.9	Infeasible
	E in purge	E in prod.	3.8	Infeasible	54.3	Infeasible
	E in purge	F in prod.	0.5	Infeasible	1.6	0.9
	D in prod.	E in prod.	0.2	3.2	42.4	Infeasible
	D in prod.	F in prod.	0.2	Infeasible	Infeasible	3.3
	E in prod.	F in prod.	0.2	Infeasible	Infeasible	3.5

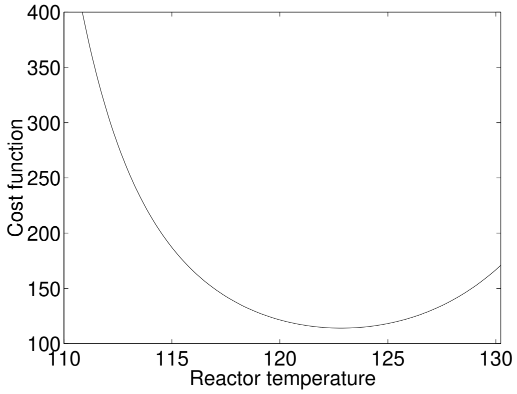
Table 6.5: Loss [$/h] when fixing all three degrees of freedom. Reactor temperature is fixed in all cases.



(a) Constant reac.T and C in purge



(b) Constant reac.T and recycle flow



(c) Constant C in purge and recycle flow

Figure 6.3: Shape of cost function for case I

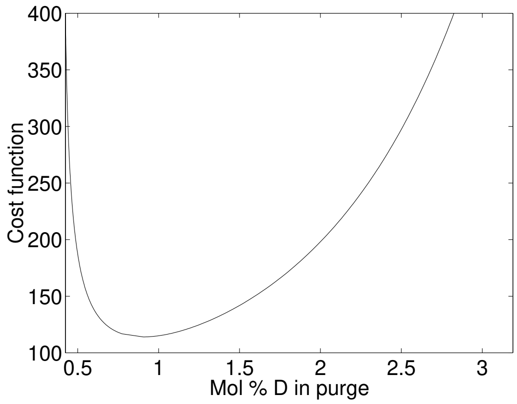


Figure 6.4: Shape of cost function for case III (with constant reactor temperature and C in purge)

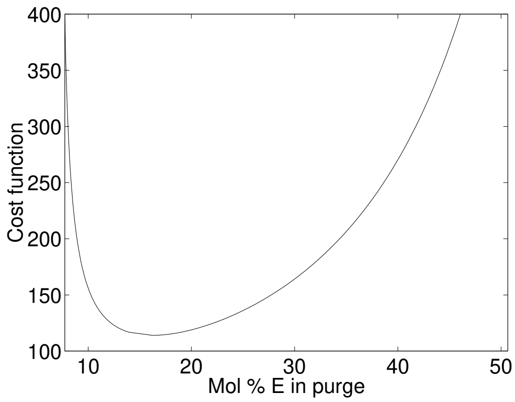


Figure 6.5: Shape of cost function for case IV (with constant reactor temperature and C in purge)

6.5.10 Summary

In conclusion, control of reactor temperature, C in purge, and recycle flow or compressor work (cases I or II) results in a small loss for disturbances, has a flat optimum (and is thus insensitive to implementation error), and are therefore good candidates for self-optimizing control.

6.5.11 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*b*) (Tyreus, 1999*a*). However, recall that we eliminated B in purge at an early stage because it gave a rather large loss for disturbance 2 (see Table reftab:2DF). 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 6.6. In conclusion, we do not recommend controlling inert composition.

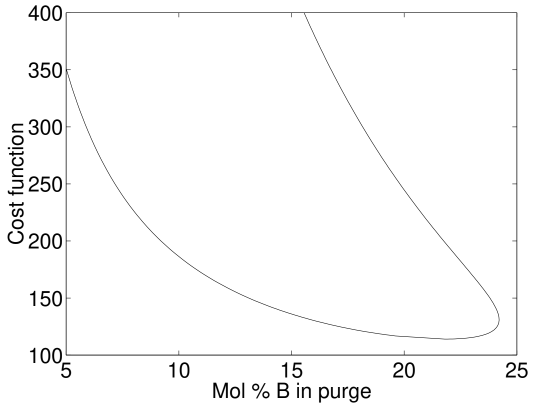


Figure 6.6: Typical unfavorable shape of cost function with B (inert) in purge as controlled variable (shown for case with constant reactor temperature and C in purge).

6.6 Conclusion

In this study of the Tennessee Eastman process, we have focused on the selection of the controlled variables using the concept of self-optimizing control (acceptable loss with constant setpoints and with implementation errors). The conclusion is that in addition to the constrained variables, reactor temperature, C in purge and recycle flow or compressor work, should be controlled.

Somewhat arbitrarily, we selected to control reactor temperature. However, since our final candidates have good self-optimizing properties, it is justifiable. This does not mean that they are the best alternative, but "acceptable" is good enough here.

A very common suggestion, is that it is necessary to control the inventory of inert components. However, this choice may lead to serious operational problems as demonstrated by Figure 6.6, and in a more careful evaluation we did not find any favorable combination that included the inert composition.

We have presented a number of criteria for reducing the number of alternatives. However, note that the number of alternatives would have been much larger if we also had considered combinations of variables, such as sums, differences, ratios and so on. In some applications, such as distillation, the use of variable combinations has proved to be very useful.

All the analysis in this paper is based on steady-state economics. We have not said anything on how the proposed control structure should be implemented. This could be the subject of future work, and should preferably be based on a controllability analysis.

Chapter 7

Control of an Industrial Heat Integrated Distillation Column

Truls Larsson and Sigurd Skogestad

Presented at the AIChE annual meeting, Dallas November 1999.

Abstract

Heat integration of distillation columns, where condenser and reboiler duties are coupled, is used to reduce the energy consumption used for separation. It is well known which variables to control in normal distillation columns, but for heat integrated distillation columns it is more an open issue. This work will focus on the selection of controlled variables.

We use the concept of self-optimizing control, (Skogestad *et al.*, 1999), in order to select the controlled variables. This involves a search for the variables which when kept constant give operation with acceptable economic loss. This provides us with a systematic framework for selection of controlled variables based on steady state economics.

7.1 Introduction

Much work has been done on the control of distillation columns, see (Skogestad, 1992) and references therein. This short review is limited to control studies on double effect distillation columns, and it is far from complete. There are good papers that have not been included.

Tyreus and Luyben (1976) published one of the first papers addressing control of heat integrated distillation columns. Their main conclusion was to decouple the two columns by introducing auxiliary boilers and condensers. Their conclusions where solely based on simulations. Lenhoff and Morari (1982) questioned their conclusion since they did not find such an effect.

The work by Roffel and Fontein (1979) is most similar to our work. They discuss some aspects related to constrained control. Much of their discussion is based on steady state economics and active constraints. Lenhoff and Morari (1982) point out that it is not always optimal that the overhead compositions of both distillation columns are at their constraints.

Frey *et al.* (1984) recommended using ratios of material flows as manipulated variables. They used the relative gain array as a controllability measure.

Much of the above work used simple models that did not include important effects, (like flow dynamics or heat transfer area). Gross *et al.* (1998) present results for a rigorous model. They used controllability analysis and non-linear simulation, and concluded that a detailed model is needed in order to capture essential details.

We will use the concept of self-optimizing control, (Skogestad *et al.*, 1999), for selection of the controlled variables. This involves a search for the variables that when kept constant will give minimum operation cost. This provides us with a systematic framework for selection of controlled variables based on steady state economics. Finally nonlinear simulation is used to confirm that the self-optimizing control structure will work.

7.2 The process and modeling

The plant is shown in Figure 7.1. Methanol and water are feed to the first column that operates at a high pressure. The bottom flow is feed to the second column that operates at a lower pressure. Since there is ethanol in the feed, there is a small side stream in the low-pressure column to avoid accumulation of ethanol. Heat is transferred from the condenser in the high-pressure column to the reboiler in the low-pressure column.

This paper is based on an industrial case but the design parameters have been somewhat modified. The main assumptions are

- Ideal equilibrium stages (with fewer stages than the actually number of trays).
- Thermodynamics is based on NRTL-activity coefficients and ideal gas, with parameters from (Gmehling and Onken, 1977) and (Reid *et al.*, 1987).
- Both columns operate below 10 bar, so the hold-up of vapor is neglected.
- The liquid flows are modeled by a simplified Francis weir formula.

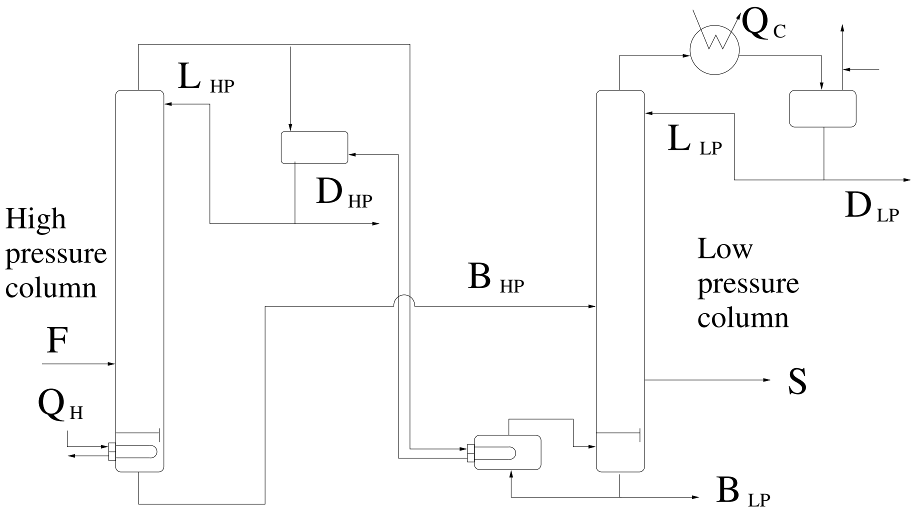


Figure 7.1: High and low-pressure columns.

- The gas flows between the stages are given by $F = c\sqrt{P_{j-1}^2 - P_j^2}$.
- Cooling is adjusted to achieve constant pressure in the top of the low-pressure column.
- Dynamics in the heat exchangers are ignored.

The columns are operated such that flooding and weeping is avoided. Olsen *et al.* (1997) present simulations of the same plant using a different model.

7.3 Selection of controlled variables

We will follow the procedure outlined in Chapter 4.2.

Step 1: Degrees of freedom and constraints

The eleven available manipulated variables are: feed rate, heat to the high-pressure column, reflux in high-and low-pressure column, distillate flow in high- and low-pressure column, the heat transfer area for the condenser/boiler, the bottom flow in high and low-pressure column, the side stream and cooling in the low-pressure column. We loose four degrees of freedom to the four levels, which must be controlled, and one degree of freedom since the feed is given.

If we do not include the side stream, see discussion below, we are left with 5 degrees of freedom at steady state, which is one less then two distillation columns without heat integration.

Step 2: Objective function

There are two conflicting elements in the operation of the plant: we would like as much valuable product as possible, and we would like to use as little energy as possible. This gives the objective (profit function) to be maximized:

$$J = D_{HP} + D_{LP} - Q_{HP} w_r \tag{7.1}$$

Where $D_{HP} + D_{LP}$ $[mole/s]$ is the top product (methanol), and Q_{HP} $[MW]$ is the heat load to the high- pressure column, and w_r is the relative cost of energy.

The constraints are:

- The low-pressure column pressure must be above 1 bar.
- The high-pressure column pressure must be below 10 bar.
- Maximum heat transfer area in the combined reboiler/condenser.
- Purity constraint on top product $D_{HP} + D_{LP}$ (mole fraction 0.99 methanol).
- The constraint on bottom purity B_{LP} (mole fraction 0.99 water) is only considered in relation to the side stream.

Step 3: Optimization

The five steady state degrees of freedom will be used for optimization and satisfying the constraints. We find that the following constraints are at active at the optimum:

- The low-pressure column pressure must be above 1 bar.
- Maximum heat transfer area in the combined reboiler/condenser.
- Purity constraints on top product $D_{HP} + D_{LP}$ (mole fraction 0.99 methanol).

It was noted in Roffel and Fontein (1979) and Lenhoff and Morari (1982) that it was optimal for the streams of the two distillation columns to have a different top composition. We did not find such an effect in this work; it was optimal with both compositions at the same values (their constraint).

This results in four constraints, and leaves one degree of freedom as unconstrained. Figure 7.2(a) shows the combined methanol production rate from the two columns as a function of the heat load. For low heat loads the production increases sharply. But as the bottom product gets purer the increase in production rate reduces, and approaches the upper limit slowly. This figure shows use that if energy is free (i.e. very low relative cost) then the process will be constrained, either by maximum heat load, flooding or pressure constraints.

Since Figure 7.2(a) has a sharp transition between a steep increase and no increase, it indicates that the optimum is nearly insensitive to the relative cost of energy. This is confirmed by Figure 7.2(b), which shows the objective as a function of heat load for several relative costs. With the relative cost of 0.6488 mole/MJ, the optimum heat load is 66.1 MW.

Step 4: Disturbances

The main disturbance is feed rate variations, with an assumed range of 1200 mole/s $\pm 20\%$ Feed composition has only small variations, and will not be considered.

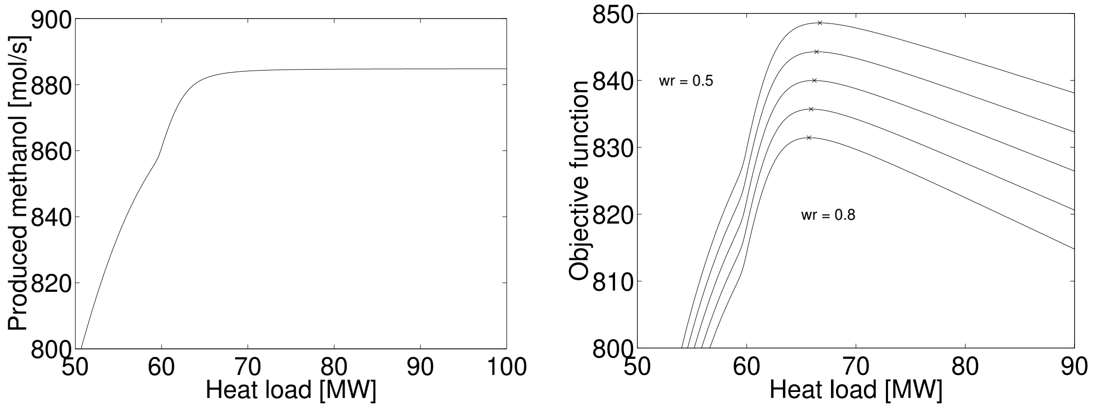


Figure 7.2: Production rate and objective function as a function of heat load.

Step 5: Candidates for controlled variables

Our candidates for the remaining degree of freedom are:

Heat load to high-pressure column (Q_{HP}).
Pressure in the high-pressure column (P_{HP}).
Pressure drop in the high-pressure column (ΔP).
Bottom composition in the high-pressure column (x_{BHP}).
Bottom composition in the low-pressure column (x_{BLP}).
Temperature in the low-pressure column (Reboiler T_{BLP}, on tray i $T_{i,LP}$).
Bottom flow from high-pressure column (B_{HP}).
Reflux flow in high-pressure and low-pressure column (L_{LP}, L_{HP}).
Ratio between heat load and feed rate Q_{HP}/F.
Ratio between heat load and reflux in high-pressure column (Q_{HP}/L_{HP}).
Ratio between heat load and reflux in low-pressure column (Q_{HP}/L_{LP}).
Ratio between bottom flow from high-pressure column to feed rate B_{HP}/F.

We are looking for candidates with a “flat” optimum. As a first screening, we eliminate candidates, which gives a large loss, for their expected range (based on changes in the optimal value due to disturbance and the implementation error). Table 7.1 shows us the six variables that gives large losses, and the three variables that gives infeasibilities, (see step 7 below). These nine variables are not considered any further. Note that all open-loop alternatives are ruled out.

To give a feel for the size of these numbers, if we had a loss of 1 unit in Table 1 and 2 during a whole year, we would lose approximately 100.000 US $. This means that there is a significant difference between the various alternatives.

Step 6: Evaluation of the loss

Table 7.1 is only an approximation, and Table 7.2 the maximum loss by controlling a variable to a constant setpoint when there are disturbances and implementation errors. The best

Variable	Range	Maximum loss
Q_{HP}	51 - 86 MW	68
P_{HP}	6.7 - 10.5 bar	26
ΔP	42 - 75 mbar	infeasible
$1 - x_{BLP}$	1e-05 - 0.001	19
x_{BHP}	0.36 - 0.38	24
T_{BLP}	379 - 387 K	23
$T_{2,LP}$	379 - 386 K	20
$T_{4,LP}$	378 - 384 K	8
$T_{6,LP}$	359 - 367 K	4
T_{HP}	402 - 419 K	25
B_{HP}	635 - 1018 mol/s	infeasible
L_{LP}	876 - 1470 mol/s	43
L_{HP}	915 - 1600 mol/s	47
Q/F	4.4e-02 - 6.6e-02 MW/mol/s	54
Q/L_{LP}	4.7e-02 - 7.0e-02 MW/mol/s	79
Q/L_{HP}	4.4e-02 - 6.6e-02 MW/mol/s	79

Table 7.1: The worst loss within the range.

alternative as is temperature on tray six in the low-pressure column, which gives an average loss of 6 units.

Variable	Max loss, disturbance	Max loss, control error	Average
P	21	23	22
T_{BHP}	21	22	21
T_{BLP}	18	26	22
T_{2LP}	15	21	18
T_{4LP}	7	12	9
T_{6LP}	2	10	6
x_{BLP}	2	20	11

Table 7.2: The worst loss within the range.

Step 7: Further analysis

Control of the pressure drop ΔP_{HP}, bottom composition x_{BHP}, or the bottom flow B_{HP}, (all in the high-pressure column) may lead to serious operational problems as indicated in Table 7.2. Figure 7.3 shows the objective as a function of x_{BHP} and B_{HP} . Due to the multiplicities in the objective, an implementation error could move the plant into a region with a very large loss or infeasibility problem.

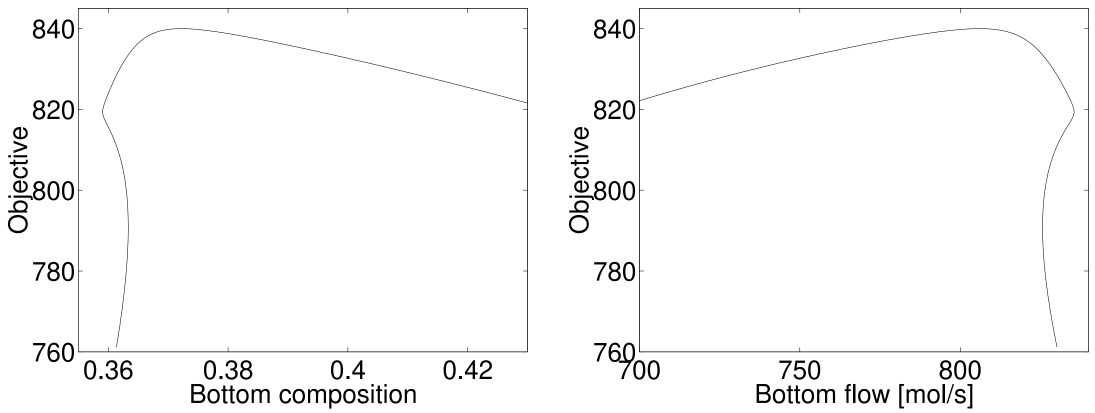
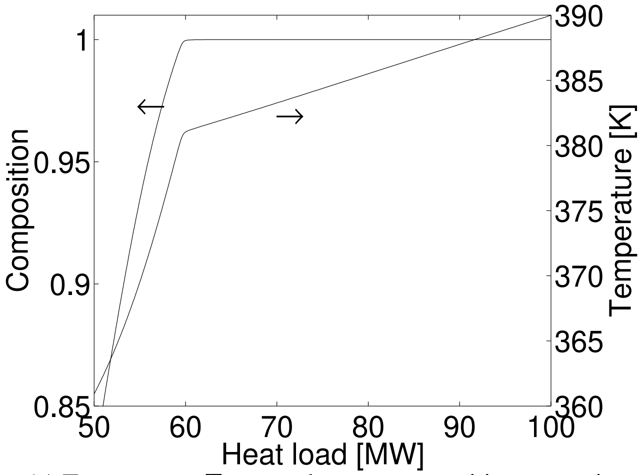


Figure 7.3: Multiplicity in the objective for some of the variables in the high-pressure column.

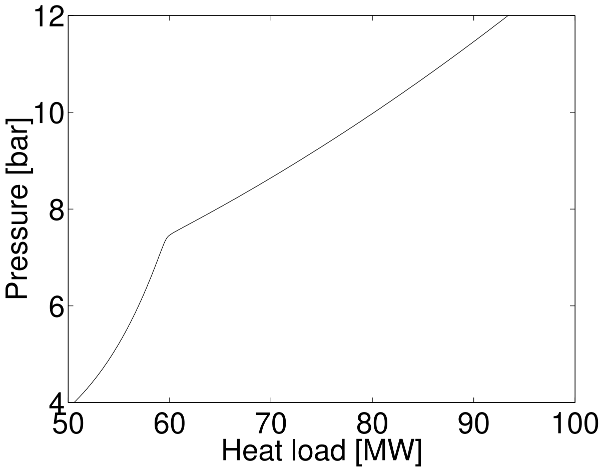
Why do these multiplicities occur? Let us start with high heat input and reduce it, see Figure 7.4. This results in a lower pressure in the high-pressure column, and less heat is supplied to the low-pressure column. Then the amount of water in the bottom flow will decrease in both columns. At a certain point we will get breakthrough of methanol in the bottom of the low-pressure column. This is accompanied with a steep decrease in temperature in the low-pressure column. This temperature decrease improves the heat transfer between the two columns, so that we get more cooling in the high-pressure column. This further reduces the pressure in the high-pressure column. The effect of lower pressure is that the separation improves and the amount of water in the bottom flow temporarily increases. But as the heat load is further decreased the bottom composition of water will decrease again. This also explains what happens to the bottom flow, since top composition is controlled, it has to change in the same manner as the bottom composition.

The side stream

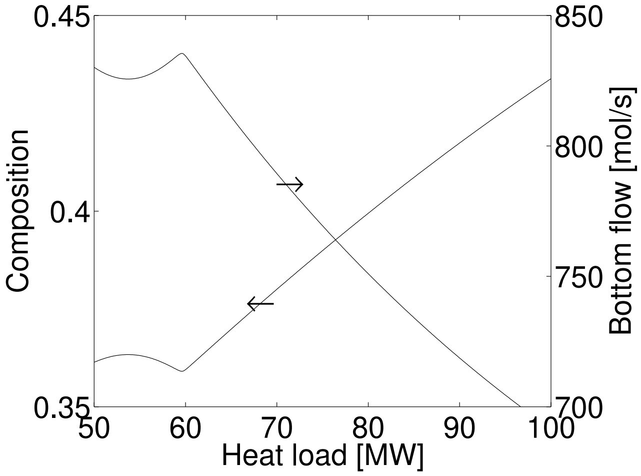
We did not take the side stream flow into account, since it was expected to have only a small effect on the economic objective. It is possible to operate the low-pressure distillation column without side stream. This would mean that all the ethanol would have to leave the column in the bottom product. This is undesirable since that would make reuse of water more difficult. However, since ethanol will accumulate in the lower parts of the low-pressure column, a small side stream will make it possible to produce purer water. In Figure 7.5 we have shown how methanol production rate depends on the side stream (with constant heat load for the high-pressure column). A small side stream gives a small increase in methanol production rate. If the side stream is below 1 kg/s then there is a steep increase in impurities in the bottom product, this is unacceptable. We will operate at with a side stream of 2 kg/s, where water is so pure that it will allow for reuse.



(a) Temperature T_{BLP} and water composition x_{BLP} in low-pressure column

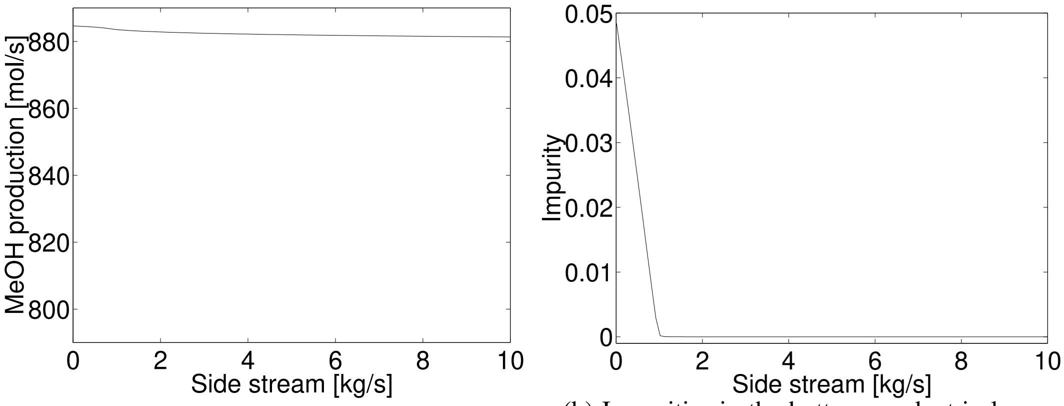


(b) Pressure in high-pressure column P_{HP}



(c) Bottom flow B_{HP} and water composition x_{BHP}

Figure 7.4: Selected variables as a function of heat load.



(a) Methanol production rate

(b) Impurities in the bottom product in low-pressure column

Figure 7.5: Production rate and bottom composition as a function of the side stream.

7.4 Selection of the throughput manipulator

We have assumed that the feed rate (throughput) was given (by upstream production). However there are a large buffer tank in between the upstream production and our distillation columns, which gives us some degrees of freedom for short-term adjustment of the feed rate.

Further more there is a maximum throughput in the columns, and a closer look shows that this is given by maximizing the pressure in the high-pressure column, Figure 7.6. To be able to maximize the throughput with out changing the control structure, we therefor select to control this pressure and use it setpoint as a handle for the throughput.

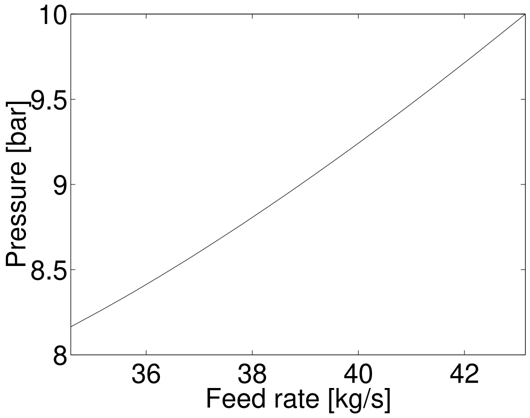


Figure 7.6: Pressure as a function of feed rate.

7.5 The control structure

The proposed control structure is shown in Figure 7.7.

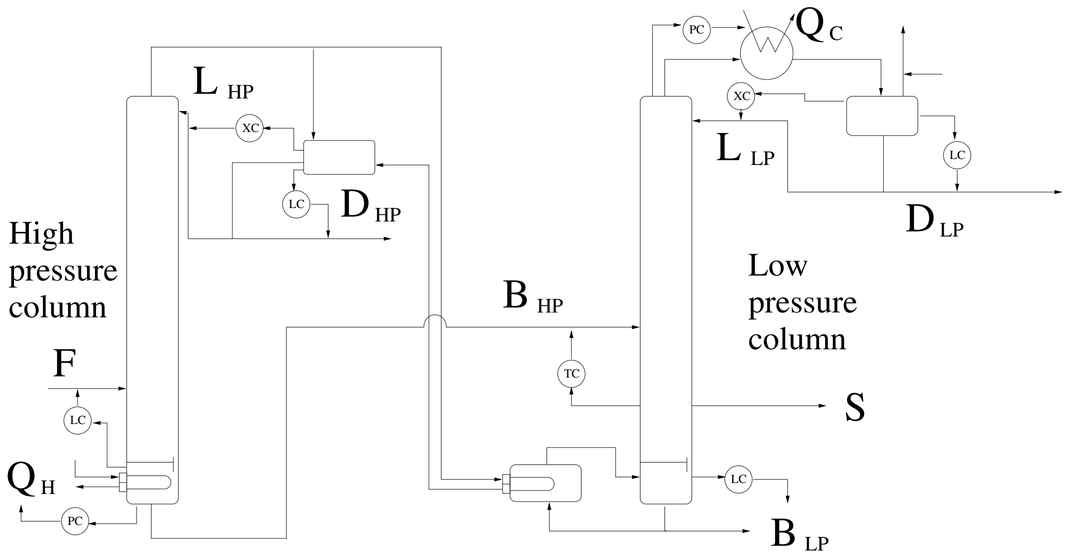


Figure 7.7: The proposed control system.

Some comments on the control system:

- The distillate flows are used for level control and reflux flows are used for composition control.
- The heat transfer between the columns is maximized and is therefor not available as a manipulated variable. This explains why Q_{HP} is used for pressure control in the reboiler in the high-pressure column (the highest pressure). B_{LP} is used for level control.
- The flow B_{HP} is then closed for temperature control in the low-pressure column.
- The reboiler level in the high-pressure column must then be controlled by the feed flow. (Actually, this is another reason for choosing the pressure as a throughput manipulator.)

7.6 Simulations

Closed loop simulations of the proposed control system are shown in Figure 7.8. We have shown the response in the product composition of methanol in both columns for a step in the setpoint of the pressure in the high-pressure column. (The plot is scaled such that acceptable performance is between ± 1, the scaling factor is 100.)

7.7 Conclusion

In this paper we have studied a heat integrated distillation column. The selection of controlled variables has been the main part of the paper. The selection of controlled variable has

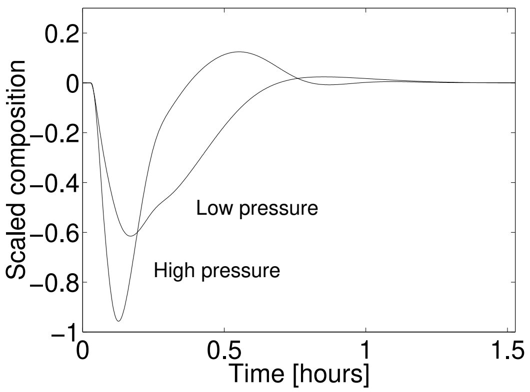


Figure 7.8: Responses to a step from 9 to 10 bar in the setpoint for pressure in the high-pressure column.

much in common with normal distillation, but there are some important differences.

The heat integration implies that we have fewer degrees of freedom than normal distillation. However since the bottom composition of the high-pressure column feeds into the low-pressure column, there are also one less product stream.

We have found that the heat transfer area between the two columns, top compositions (of valuable products), and pressure in the low-pressure column should be controlled at their constraints. There is one unconstrained degree of freedom and selecting a temperature in the lower part of the column as a controlled variable, has good self-optimizing properties.

Some of the other candidates for control, like bottom composition, pressure drop or bottom flow (all in high-pressure column) showed multiplicities in the objective function. These variables are then poor candidates for self-optimizing control since a small error could easily yield very poor operation (economically) and infeasibility problems. Two competing effects can explain these multiplicities as the heat load is decreased.

Acknowledgments

We acknowledge dr.ing. David Di Rucscio for supplying his identification code and for valuable assistance in this respect. I. Halvorsen is also acknowledged for valuable comments.

Chapter 8

Poles at the origin in the Relative Gain Array

Truls Larsson and Sigurd Skogestad

Abstract

This paper shows how it is possible to get poles at the origin in the relative gain array. The source of these poles can be zeros or poles of the original plant. It is possible that integrators will not cancel out in the RGA, therefor care must be taken if one wish to follow the approach outlined by Woolverton (1980) and McAvoy (1998). The work presented here is an argument for using the frequency dependent RGA.

8.1 Introduction

Bristol (1966) introduced the relative gain array (RGA). It was originally defined as a steady state interaction measure. Later it has been given a dynamically interpretation, (McAvoy, 1983), which will be used here. The RGA has gained a lot of attention; this work will only look into how poles at the origin may appear in the RGA.

Woolverton (1980) proposed to use the derivatives of the integrating variables instead of the gain for computation of the RGA. Later McAvoy (1983) showed that for a specific 3x3 example using the integrating variables or their derivatives are the same. Arkun and Downs (1990) showed that in some cases, it is possible to factor out the integrators. Since RGA is independent of scaling the RGA of the derivatives will be the same as for integration output. Or written with equations

$$G(s) = \begin{bmatrix} G_{NI} \\ G_I \frac{1}{s} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \frac{1}{s} I \end{bmatrix} \begin{bmatrix} G_{NI} \\ G_I \end{bmatrix} \tag{8.1}$$

where G_{NI} and G_I does not contain any integrators. Which gives for the RGA

$$RGA(G) = RGA\left(\begin{bmatrix} I & 0 \\ 0 & \frac{1}{s} I \end{bmatrix} \begin{bmatrix} G_{NI} \\ G_I \end{bmatrix} \right) = RGA\left(\begin{bmatrix} G_{NI} \\ G_I \end{bmatrix} \right) \tag{8.2}$$

Even though such a factorization is always possible, it does not imply that the relative gain of the stable part $\begin{bmatrix} G_{NI} \\ G_I \end{bmatrix}$ is non-singular at steady state.

In this article we will look into this problem, and we will show that the relative gain array may contain poles at the origin even for systems without zeros at the origin. That is, there exists cases where integrators in G will be present in $RGA(G)$. This paper gives the conditions for this to happen. Briefly the implications of these results are discussed. (However, the implication of integrators in the relative gain array for controllability or control structure selection is not discussed.)

8.2 Results

From the definition of the RGA one can derive

$$\lambda_{ij} = (-1)^{i+j} \frac{g_{ij} \det G^{ij}}{\det G} \tag{8.3}$$

Where λ_{ij} is the elements of the RGA, G^{ij} denotes the matrix with row i and column j removed. The next theorem gives conditions for the RGA to have poles at the origin.

Theorem 1 If one of the following conditions are true:

1. $\det G$ has more zeros at the origin than $g_{ij} \det G^{ij}$.

2. $g_{ij} \det G^{ij}$ has more poles in the origin than $\det G$.

then λ_{ij} will have poles in the origin.

Proof: The proof is straight forward by looking at equation 8.3.

For this theorem to be meaningful we will have to show that the conditions given in fact can be satisfied. The first condition is satisfied if there exist a zero of $\det G$ which is not a zero of g_{ij} or $\det G^{ij}$. Since zeros is a property of the whole plant this is possible, as shown in the example below.

Example 5 *Consider*

$$G = \begin{bmatrix} \frac{1}{1+s} & 1 \\ 1 & \frac{1}{1+10s} \end{bmatrix} \tag{8.4}$$

This plant has a zero at the origin. The $(1,1)$ *element in the RGA is*

$$\lambda_{11} = \frac{-1}{s(11+10s)} \tag{8.5}$$

which contains a pole at the origin.

The second condition can be satisfied in two ways. A pole which is present in g_{ij} or $\det G^{ij}$ and not present in $\det G$ due to pole zero cancellation, or if there exist poles which are present in both g_{ij} and $\det G^{ij}$. Below we give one example for of the latter.

Example 6 *Consider*

$$G = \left[\begin{array}{c|c} 0 & \begin{bmatrix} 1 & 1 \end{bmatrix} \\ \hline \begin{bmatrix} 1 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 1+1/s & 1/s \\ 1/s & 1+1/s \end{bmatrix} \tag{8.6}$$

This plant has a zero at $z = -2$, *and one pole at the origin. The* $(1,1)$ *element in the relative gain for this plant is*

$$\lambda_{11} = \frac{(s+1)^2}{s(s+1)} \tag{8.7}$$

The pole is observable in both outputs and stabilizable with both inputs. So when one output is perfectly controlled, the pole is also stabilized. λ_{11} *is the ratio between the open loop transfer function (between* u_1 *and* y_1*) and the transfer function (between* u_1 *and* y_1*) when* y_2 *is perfectly controlled. Only the former will contain the pole, hence* λ_{11} *will have a pole at origin.*

What if we factor out the integrator, as implied in equation 8.2

$$G = \begin{bmatrix} 1+1/s & 1/s \\ 1/s & 1+1/s \end{bmatrix} = 1/s \begin{bmatrix} s+1 & 1 \\ 1 & s+1 \end{bmatrix} \tag{8.8}$$

At first one may believe that since the factorization is possible, the relative gain will be well behaved at steady state. This is not the case; $\begin{bmatrix} s+1 & 1 \\ 1 & s+1 \end{bmatrix}$ *has a zero at the origin, which gives a pole in the RGA.*

The last example gives valuable insight. It shows how and explains why the pole appears in the relative gain. It leads us to the next theorem

Theorem 2 The relative gain will contain poles at the origin if the following conditions are met:

1. It is possible to stabilize a pole at the origin, with two different loops ($y_i - u_j$ or $y_k - u_l$ where $i \neq k$ and $j \neq l$). In mathematical terms this means that the pole is observable in more than one output and that it is controllable with more than one input.

2. There are equal or less numbers of zeros in the origin for the subsystem G^{ij} then there is in G.

Proof: If a pole at the origin is stabilizable with $y_i - u_j$ then element g_{ij} contains that pole and no zero at the origin. If the same pole is stabilizable with the pair $y_k - u_l$ where $i \neq k$ and $j \neq l$ then element g_{kl} will contain the same pole. Element g_{kl} is part of G^{ij}. If G^{ij} does not have a zero at the origin then $\det G^{ij}$ *will have a pole at the origin. Therefor g_{ij}* $\det G^{ij}$ *will have more poles at the origin than* $\det G$. *If G^{ij} contains zeros at the origin. For the poles to appear in the RGA, then these have to be canceled against zeros in G.*

At this point we would like to point out that it is not unlikely that this situation can occur in real system. In Chapter 4 we have shown an example of such a plant. In that plant the conversion in the reactor depends on the reactor level, which makes it observable in some compositions, and controllable with other inputs. The resulting RGA has a pole at the origin.

8.3 Conclusion

In this paper we have shown how the relative gain array can have poles at the origin. There are two sources for these poles which are either zeros or poles of the original plant. The main point that we would like to stress is that it is not necessary to have zeros at the origin to get a pole at the origin. We showed that when a pole at the origin is possible to stabilize with two different loops then the relative gain array has poles at the origin.

The implications of this work is that one should be carefully when one follows the approach outlined in (Woolverton, 1980), (Arkun and Downs, 1990) and (McAvoy, 1998). A better approach is to use a frequency dependent RGA.

Chapter 9

Conclusion

9.1 Discussion

The goal of this thesis has been to improve the understanding of plantwide control. By the term *plantwide control* it is *not* meant the tuning and behavior of each of these loops, but rather the *control philosophy* of the overall plant with emphasis on the *structural decisions*. The starting point for the studies has been a literature review of plantwide control.

Based on the review, we have proposed a plantwide control design procedure. It starts with a top-down analysis of the control objective, which we assume is possible to formulate as a constrained optimization problem, and a degree of freedom analysis. After the throughput manipulator has been chosen, the control system is designed bottom-up, starting with the fastest loops.

The procedure will result in a hierarchical control system with several control layers. An issue that has been resolved is if a control layer can introduce new fundamental limitations for the next control layer that was not present in the original plant. We have shown that if the lower layer controller is minimum phase, stable and the measurements and setpoints are available for the next layer, new fundamental limitations can not be introduced. If any of these criteria are not met then we can introduce new fundamental limitations for the next control layer.

In the process control community there has been developed some heuristic rules for the design of control systems. Two of these are related to the recycle flow, (Luyben *et al.*, 1997) and (Fisher *et al.*, 1988*b*).

Luyben *et al.* (1997) has proposed to fix a flow in every liquid recycle loop, as a generic rule. The basis for this rule is that he wants to avoid to overload the distillation column or to avoid moving it below its minimum turndown. However as shown in Chapter 4 if it is likely that the separation section will be overloaded then it will be much more efficient to maximizing the conversion in the reactor (by maximum holdup) then fixing a flow in the recycle loop. Further more we have shown that application of his rule can give a control system with (very) bad self-optimizing properties (large economic losses due to uncertainties and disturbances). However, letting the reactor level float additional buffer capacity is obtained which make rejection of some disturbances easier.

Fisher *et al.* (1988*b*) were more concerned with gas phase systems. They stated: “keep the gas recycle flow constant at its maximum value", which is not entirely correct. In Chapter 5 and 6 we showed that the recycle flow is an "unconstrained degree of freedom" (i.e. the optimal value of recycle flow is an unconstrained variable and it maximum value does not coincide with the economic optimum). What they probably meant was to keep equipment, like recycle compressor, at their constraints. Which is what they have done in the example in (Fisher *et al.*, 1988*b*). This is confirmed in our case studies, inputs like control-valves in the recycle loop, compressor-recycle valve and recycle compressor (if mounted on separate shaft) where at their constraints at the optimum. The recycle flow-rate can also be affected "indirectly" (for example by purge flow), and it actually has good self-optimizing properties.

In a chemical processing plant it is very important to avoid build-up of inert. Inert will accumulate when its mass balance is not satisfied. This can happen if a purge flow is too small. This does not meant that the inert composition should be controlled, with a proper sets of controlled variables the inert composition will be self-regulating (e.g. pressure control using the purge flow). Still it is often assumed that control of inert composition is needed, but as we have shown in our case studies that it is not true. Inert composition is likely to have bad self-optimizing properties. Often the optimal inert composition is close to the highest attainable inert composition, see figure 6.6. Thus a small implementation error can lead to infeasibilities. Based on (Downs, 1992), Luyben *et al.* (1998) recommends to make a table over all the components in order to check how the control system will correct an imbalance in a chemical component. If one applies the concepts of self-optimizing control, such a table will not be necessary.

9.2 Directions for future work

In this work we have successfully used the concept of self-optimizing control on several case studies. With a clear definition of the control objective and a systematic framework for analysis of the control structure we have been able to question some of the heuristics proposed in process control community. However, the case studies have also revealed some of the shortcomings of the procedure.

In our case studies the number of measurement and the degrees of freedom to assign them to where modest. But when the degrees of freedom that needs to be assigned grows, it quickly becomes impossible to analyze all the alternatives. Good process understanding can help at this point, but it may also prevent the engineer from seeing the best solution. A way to screen among the many alternatives is needed.

One alternative is to use the maximization of the minimum singular value (with proper scaling), unfortunately the local information is probably not enough information. As can be seen from some of the figures in this thesis, it is when we are at a distance from the optimum that we get the large losses. Still this approach should be further studied.

A different alternative could be a mixed-integer-nonlinear programming (MINLP) approach. The problem formulation and a good solver have to be found. The problem formulation is probably easy, but the solver has to exploit the structure of this problem. Firstly, the lower bound of a branch is easy to calculate from optimization. Secondly, the optimization and steady-state solver must be based on a continuation method. The last point may be important since infeasibilities are likely to occur. At least toolbox for automating many of the calculations is needed.

Very often the objective function turned out to have a flat surface for several of the possible controlled variables. This does imply that the final choice should be based on other considerations. There are a least two such very important issues, which has not been properly addressed: Constraints handling and abnormal situations.

In this thesis all constraints has been treated equally. However, there are differences. Some constraints are hard if they are violated it would lead to shutdown of the plant and therefor a large loss (like the constraints on reactor pressure in the Tennessee Eastman plant). Some constraints are "soft"; they may be violated dynamically (like composition in a distillation column). It is not even given that all the active constraints are measured. For the unmeasured constraints, the concept of partial control, (Skogestad and Postlethwaite, 1996), may be suitable.

For the case of hard constraints with a measurement, a back-off in the setpoint for this controlled variable is needed to avoid violation of the constraint. The size of the back-off should be as small as possible (i.e. as close to the optimum as possible). However the required size of the back-off will depend on the control systems ability to minimize variations in that measurement. In future works, more emphasis should be given to this problem, it may be wise to use some of the unconstrained degrees of freedom to achieve tighter control of the constraints. A dynamic analysis will be needed to address this issue properly. A starting point for future work could be along the lines of Narraway and Perkins (1993), combined with partial control ideas of Skogestad and Postlethwaite (1996).

A soft constraint may not be as critical, but there may be cases where there is a larger average cost for fluctuations then there is for steady behavior. Distillation is a good example, over purification requires more energy than what is saved by under purification. Fewer fluctuations in the composition will save energy. Such behavior is likely to appear frequently, due to the "law" of decreasing outcome. This dynamic cost is probably much more important than the dynamic cost that is given in (Zheng and Mahajannam, 1999). This will be challenging since a dynamic non-linear analysis is needed.

The most important task of the control system is to avoid a shutdown of the plant. An increase in the regularity of the plant will probably have a much larger effect on the economics of the plant than optimization. This implies that greater emphasis should be placed on the control systems ability to tackle abnormal situations. In practice this means that we will have to accept a larger loss for "the usual" disturbance, to be able to handle "rare" disturbance. It may even be needed to back-off from the optimum to handle large and "rare" disturbances.

As stated many places in this thesis, the control system is often designed in a hierarchical manner, i.e. with a base control layer, a supervisory control layer etc. This was touched upon in Chapter 3 where we looked at the possibility to introduce new fundamental limitations by improper design of the lower layer. More work should be done on how this structuring should be and which benefits it is possible to obtain from using a several layers. One such benefit of a multilayered control system is believed to robustness of the higher layers.

In practice the different control layers have a different sampling time, this implies that the higher layer reacts slower then the lower layer. Thus the higher layer may not be able to cancel the effects of the lower layer; hence if the sampling time is locked the lower layer may introduce limitations for the higher layer. This effect needs to be studied further.

Bibliography

Alonso, A.A. and B.E. Ydstie (1996). Process systems, passivity and the second law of thermodynamics. *Computers & Chemical Engineering* **20**, S1119–S1124.

Arbel, A., I.H. Rinard and R. Shinnar (1996). Dynamics and Control of Fluidized Catalytic Crackers. 3. Designing the Control System: Choice of Manipulated and Measured Variables for Partial Control. *Ind. Eng. Chem. Res.* **35**, 2215–2233.

Arkun, Y. and G. Stephanopoulos (1980). Studies in the synthesis of control structures for chemical processes: Part iv. design of steady-state optimizing control structures for chemical process units. *AIChE Journal* **26**(6), 975–991.

Arkun, Y. and J. Downs (1990). A general method to calculate input-output gains and the relative gain array for integrating processes. *Computers chem. Engng* **26**, 1101–1110.

Balchen, J.G. and K.I. Mummé (1988). *Process Control, Structures and applications*. Van Nostrand Reinhold.

Banerjee, A. and Y. Arkun (1995). Control configuration design applied to the tennessee eastman plant-wide control problem. *Computers. chem. Engng.* **19**(4), 453–480.

Bristol, E.H. (1966). On a new measure of interaction for multivariable process control. *IEEE Trans. on automatic control* pp. 133–134.

Buckley, P.S. (1964). *Techniques of Process control*. Chap. 13. John Wiley & Sons.

Cao, Y. (1995). Control Structure Selection for Chemical Processes Using Input-Output Controllability Analysis. PhD thesis. University of Exeter.

Chang, J.W. and C.C. Yu (1990). The relative gain for non-square multivariable systems. *Chem. Eng. Sci.* **45**, 1309–1323.

Douglas, J.M. (1988). *Conceptual Design of Chemical Processes*. McGraw-Hill.

Downs, J.J. (1992). Distillation control in a plantwide control environment. In: *Practical Distillation Control* (W.L. Luyben, Ed.). Chap. 20, pp. 413–439. Van Nostrand Reinhold.

Downs, J.J. and E.F. Vogel (1993). A plant-wide industrial process control problem. *Computers chem. Engng.* **17**, 245–255.

Findeisen, W, F.N. Bailey, M. Brdys, K. Malinowski, P. Tatjewski and A. Wozniak (1980). *Control and coordination in Hierarchical Systems*. John Wiley & sons.

Fisher, W.R., M.F. Doherty and J.M. Douglas (1988*a*). The interface between design and control. 1, 2 and 3.; 1: Process controllability, 2: Process operability 3: Selecting a set of controlled variables.. *Ind. Eng. Chem. Res.* **27**(4), 597–615.

Fisher, W.R., M.F. Doherty and J.M. Douglas (1988*b*). The interface between design and control. 3: Selecting a set of controlled variables.. *Ind. Eng. Chem. Res.* **27**, 611–615.

Foss, C.S. (1973). Critique of chemical process control theory. *AIChE Journal* **19**(2), 209–214.

Frey, R.M., M.F. Doherty, J.M. Douglas and M.F. Malone (1984). Controlling thermally linked distillation columns. *Ind. Eng. Chem. Res.* **23**, 483–490.

Georgakis, C. (1986). On the use of extensive variables in process dynamics and control. *Chemical Engineering Science* **47**(6), 1471–1484.

Gilliland, E.R., L.A. Gould and T.J. Boyle (1964). Dynamic effects of material recycle. *Preprints JACC, Stanford, Calif.* pp. 140–146.

Gmehling, J. and U. Onken (1977). *Vapor-Liquid equilibrium data collection*. Vol. 1. DECHEMA.

Grosdidier, P. and M. Morari (1985). Closed-loop properties from steady-state gain information. *Ind. Eng. Chem. Fundam.* **24**(2), 221–235.

Gross, F., E. Baumann, A. Geser, D.W.T. Rippin and L. Lang (1998). Modelling, simulation and controllability analysis of an industrial heat-integrated distillation process. *Computers. chem. Engng.* **22**(1), 223–237.

Häggblom, K.E. and K.V. Waller (1988). Transformations and consistency relations of distillation control structures. *AIChE J.* **34**(10), 1634–1648.

Hansen, J.E. (1998). Plant wide dynamic simulation and control of chemical processes. PhD thesis. Danmarks Tekniske Universitet.

Havre, K. (1998). Studies on controllability analysis and control structure design. PhD thesis. NTNU Trondheim. Aviable from http://www.chembio.ntnu.no/users/skoge/.

Havre, K. and S. Skogestad (1996*a*). Input/output selection and partial control. *Preprints IFAC '96, 13th World Congress of IFAC, San Francisco, CA, June 30-July 5, 1996, Volume M* pp. 181–186.

Havre, K. and S. Skogestad (1996*b*). Selection of variables for regulatory control using pole directions. *1996 AIChE Annual Meeting, Chicago, Illinois; Paper 45f.*

Havre, K. and S. Skogestad (1998). Selection of variables for regulatory control using pole vectors. *DYCOPS-5, 5th IFAC Symposium on Dynamics and Control of Process Systems, Corfu, Greece, June 8-10, 1998* pp. 614–619.

Hestetun, K. (1999). Regulering av Tennessee Eastman prosessen. Master's thesis. NTNU. In Norwegian.

Hougen, J. O. and N. F. Brockmeier (1969). Developing Process Control Strategies - I: Eleven Basic Principles. *Instrumentation Technology* pp. 45–49.

Hovd, M. and S. Skogestad (1992). Simple frequency-dependent ant tools for control system analysis, structure selection and design. *Automatica*.

Hovd, M. and S. Skogestad (1993). Procedure for Regulatory Control Structure Selection with Application to the FCC Process. *AIChE Journal* **39**(12), 1938–1953.

Jacobsen, E. W. (1997). On the dynamics of integrated plants - non-minimum phase behavior. To be published in journal of process control.

Jacobsen, E. W. and C. Hong (1998). Zero crossing due to loop closure in decentralized control systems. Presented at 1998 AIChE Annual meeting, Miami. Paper 233e.

Lee, J.H. and M. Morari (1991). Robust measurements selection. *Automatica* **27**(3), 519–527.

Lee, J.H., P. Kesavan and M. Morari (1997). Control structure selection and robust control system design for a high-purity distillation column. *IEEE Transactions on control systems technology* pp. 402–416.

Lee, J.H., R.D. Braatz, M. Morari and A. Packard (1995). Screening tools for robust control structure selection. *Automatica* **31**(2), 229–235.

Lee, S. (1990). *Methanol Synthesis technology*. CRC Press.

Lenhoff, A.M. and M. Morari (1982). Design of resilient processing plants - I. Process Design Under Consideration of Dynamic Aspects. *Chem. Eng. Sci.* **37**(2), 245–258.

Ljung, L. (1987). *System Identification - Theory for the User*. Prentice-Hall.

Loe, I. (1994). System-strukturer i prosessindustrien. Lecture notes, for a course given at Telemark College. In Norwegian.

Løvik, I, M Hillestad and T Hertzberg (1998). Long term dynamic optimization of a catalytic reactor system.. *Computers & Chemical Engineering* **22**(Supplement), S707–S710.

Luyben, M.L. and B.D. Tyreus (1998). An industrial design/control study for the vinyl acetate monomer process. *Computers. chem. Engng.* **22**(7-8), 876–877.

Luyben, M.L. and C.A. Floudas (1994). Analyzing the interaction of design and control. 2. reactor separator recycle system. *Computers & Chemical Engineering* **18**(10), 971–994.

Luyben, M.L. and W.L. Luyben (1995). Design and control of a complex process involving two reaction steps, three distillation columns, and two recycle streams. *Ind. Eng. Chem. Res.* **34**(11), 3885–3898.

Luyben, M.L., B.D. Tyreus and W.L. Luyben (1997). Plantwide control design procedure. *AICHE journal* **43**(12), 3161–3174.

Luyben, W.L. (1975). Steady-state energy conservation aspects of distillation column control system design. *Ind. Eng. Chem. Fundam.* **14**(4), 321–325.

Luyben, W.L. (1988). The concept of eigenstructure in process control. *Ind. Eng. Chem. Res.* **27**(4), 206–208.

Luyben, W.L. (1992). Design and control of recycle processes in ternary systems with consecutive reactions. In: *Interactions between process design and process control*. IFAC Workshop. Pergamon Press. pp. 65–74.

Luyben, W.L. (1993*a*). Dynamics and control of recycle systems. 1. simple open-loop and closed loop systems. *Ind. Eng. Chem. Res.* **32**(3), 466–475.

Luyben, W.L. (1993*b*). Dynamics and control of recycle systems. 2. comparison of alternative process designs. *Ind. Eng. Chem. Res.* **32**(3), 476–486.

Luyben, W.L. (1993*c*). Dynamics and Control of Recycle Systems. 3. Alternative Process Designs in a Ternary System . *Ind. Eng. Chem. Res.* **32**(6), 1142–1153.

Luyben, W.L. (1994). Snowball effect in reactor/separator processes with recycle. *Ind. Eng. Chem. Res.* **33**(2), 299–305.

Luyben, W.L. (1996). Design and control degrees of freedom. *Ind. Eng. Chem. Res.* **35**(7), 2204–2214.

Luyben, W.L., B.D. Tyreus and M. Luyben (1998). *Plantwide Process Control*. McGraw-Hill.

Lyman, P.R. and C. Georgakis (1995). Plant-wide control of the Tennessee Eastman Problem. *Computers chem. Engng.* **19**(3), 321–331.

MacFarlane, A.G.J. and N. Karcanias (1976). Poles and zeros of linear multivariable systems: A survey of the algebraic, geometric and complex-variable theory. *Int. J. Control* **24**(1), 33–74.

Manousiouthakis, V., R. Savage and Y. Arkun (1986). Synthesis of decentralized process control structures. *AIChE Journal* **32**(6), 991–1003.

Marlin, T.E. and A.N. Hrymak (1997). Real-time operations optimization of continuous processes. In: *Fifth international conference on chemical process control (CPC-5, Lake Tahoe, Jan. 1996)*. Vol. 93 of *AIChE Symposium Series*. pp. 156–164.

Martens, H. (1989). *Multivariate calibration*. Wiley.

McAvoy, T. J. (1983). *Interaction Analysis - Principles and Applications*. ISA.

McAvoy, T. J. (1998). A methodology for screening level control structures in plantwide control systems. *Computers. Chem. Engng* **22**(11), 1543–1552.

McAvoy, T.J. (1999). Synthesis of plantwide control systems using optimization. *Ind. Eng. Chem. Res.* **38**(8), 2984–2994.

McAvoy, T.J. and N. Ye (1994). Base control for the Tennessee Eastman problem. *Computers chem. Engng.* **18**(5), 383–413.

Mizoguchi, A., T.E. Marlin and A.N. Hrymak (1995). Operations opimization and control design for a petroleum distillation process. *Canadian J. of Chem. Engng.* **73**, 896–907.

Mizsey, P. and I. Kalmar (1996). Effects of recycle on control of chemical processes. *ESCAPE-6, 26-29 May 1996, Rhodes, Greece; Supplement to Computers & Chemical Engineering* pp. S883–S888.

Morari, M. (1982). Integrated plant control: A solution at hand or a research topic for the next decade?. In: *CPC-II*. pp. 467–495.

Morari, M., G. Stephanopoulos and Y. Arkun (1980). 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 Journal* **26**(2), 220–232.

Narraway, L. and J. Perkins (1994). Selection of process control structures based in economics. *Computers chem. Engng* **18**(supplement), S511–S515.

Narraway, L.T. and J.D. Perkins (1993). Selection of process control structure based on linear dynamic economics. *Ind. Eng. Chem. Res.* **32**(11), 2681–2692.

Narraway, L.T., J.D. Perkins and G.W. Barton (1991). Interaction between process design and process control: economic analysis of process dynamics. *J. Proc. Cont.* **1**, 243–250.

Ng, C. and G. Stephanopoulos (1998*a*). Plant-wide control structures and strategies. To be published in Process System Engineering Series of Academic press.

Ng, C. and G. Stephanopoulos (1998*b*). Plant-Wide control structures and strategies. In: *Preprints Dycops-5*. IFAC. pp. 1–16.

Ogunnaike, B.A. (1995). A contemporary industrial perspective on process control theory and practice. *Dycord+ '95, 4th IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes, Preprints, 7-9 June 1995* pp. 1–8.

Olsen, I., G.O. Endrestøl and T. Sira (1997). A rigorous and efficient distillation column model for engineering and training simulators. *Computers. chem. Engng.* **21**, 193–198.

Olsvik, O., S. Vada and R. Hansen (1997). Statoils metanolanlegg på Tjeldbergodden. *Kjemi* pp. 42–46.

Papadourakis, A., M.F. Doherty and J.M. Douglas (1987). Relative gain array for units in plants with recycle. *Ind. Eng. Chem. Res.* **26**(6), 1259–1262.

Ponton, J.W. (1994). Degrees of freedom analysis in process control. *Chemical Engineering Science* **49**(13), 2089–2095.

Ponton, J.W. and D.M. Laing (1993). A hierarchical approach to the design of process control systems. *Trans IChemE* **71**(Part A), 181–188.

Price, R.M. and C. Georgakis (1993). A plant wide regulatory control design procedure using a tiered framework. *Ind. Eng. Chem. Res.* **32**(11), 2693–2705.

Price, R.M., P.R. Lyman and C. Georgakis (1993). Selection of throughput manipulators for plant-wide control structures. *ECC '93* pp. 1060–1066.

Price, R.M., P.R. Lyman and C. Georgakis (1994). Throughput manipulation in plantwide control structures. *Ind. Eng. Chem. Res.* **33**(5), 1197–1207.

Reid, R.C., J.M Prausnitz and B.E Polling (1987). *The properties of gases and liquids*. McGraw-Hill.

Ricker, N.L. (1995). Optimal steady-state operation of the tennessee eastman challenge process. *Computers chem. Engng* **19**(9), 949–959.

Ricker, N.L. (1996). Decentralized control of the Tennessee Eastman Challenge Process. *J. Proc. Cont.* **6**(4), 205–221.

Rijnsdorp, J.E. (1991). *Integrated Process Control and automation*. Elsevier.

Rinard, I.H. and J.J. Downs (1992). Plant wide control: A review and critique. *AIChE Spring Meeting 1992, New Orleans, paper 67f.*

Roffel, B. and H.J. Fontein (1979). Constraint control of distillation processes.. *Chem. Eng. Sci.* **34**, 1007–1018.

Rosenbrock, H.H. (1970). *State Space and Multivariable Theory*. Nelson.

Scali, C. and C. Cortonesi (1995). Control of the tennessee eastman benchmark: Performance versus integrity tradeoff. *Proc. of 3rd European Control Conference, Rome, Italy, 1995* pp. 3913–3918.

Seborg, D.E., T.F. Edgar and D.A. Mellichamp (1995). *Process dynamics and control*. Wiley.

Shinnar, R. (1981). Chemical reactor modelling for purposes of controller design. *Cheng. Eng. Commun.* **9**, 73–99.

Shinskey, F.G. (1979). *Process Control Systems*. McGraw-Hill.

Shinskey, F.G. (1988). *Process Control Systems*. McGraw-Hill.

Skogestad, S. (1992). Dynamics and control of distillation columns - a critical survey. In: *Preprints IFAC-symposium DYCORD+'92*. IFAC. pp. 1–25.

Skogestad, S. (2000). Plantwide control: The search for the self-optimizing control structure.. *Accepted for publications in J. Proc. Control*.

Skogestad, S. and E.A. Wolff (1992). Controllability measures for disturbance rejection. *IFAC Workshop, London, Sept. 7-8 1992* pp. 23–30. Later printed in *Modeling, Indetification and Control, 1996, pp 167-182* .

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

Skogestad, S. and M. Hovd (1995). Letter to the editor on the decentralized versus multivariable control. *J. Proc. Cont.* **5**(6), 499–400.

Skogestad, S. and M. Morari (1987). Implications of large RGA-elements on control performance. *Ind. Eng. Chem. Res.* **26**(11), 2323–2330.

Skogestad, S. and M. Morari (1988). Understanding the dynamic behavior of distillation columns. *Ind. Eng. Chem. Res.* **27**(10), 1848–1862.

Skogestad, S., I.J. Halvorsen, T. Larsson and M.S Govatsmark (1999). Plantwide control: The search for the self-optimizing control structure. In: *Precedings of the 13th IFAC World Congress*.

Skrzypek, J., J. Slooczynski and S. Ledakowicz (1994). *Methanol Synthesis*. Polish Scientific Publishers.

Stephanopoulos, G. (1982). Synthesis of control systems for chemical plants - a challenge for creativity. *Computers & Chemical Engineering* **7**(4), 331–365.

Tyreus, B.D. (1999*a*). Dominant Variables for Partial Control. 2. Application to the Tennessee Eastman Challenge Process.. *Ind. Eng. Chem. Res.* pp. 1444–1455.

Tyreus, B.D. (1999*b*). Dominant variables for partial control; Part 1 and 2. *Ind. Eng. Chem. Res.* **38**(4), 1432–1455.

Tyreus, B.D. and W.L. Luyben (1976). Controlling heat integrated distillation columns. *Chemical Engineering Progress* pp. 59–66.

Umeda, T., T. Kuriyama and A. Ichikawa (1978). A logical structure for process control system synthesis. *Proc. IFAC Congress (Helsinki) 1978*.

van de Wal, M. and B. de Jager (1995). Control structure design: A survey. In: *Proceedings of the American control conference*. pp. 225–229.

Vanden Bussche, K. M. and G. F. Froment (1996). A steady-state kinetic model for methanol synthesis and the water gas shift reaction on a commercial $Cu/ZnO/Al_2O_3$ Catalyst. *Journal of Catalysis* pp. 1–10.

Wolff, E. and S. Skogestad (1994). Operability of integrated plants. *PSE '94, Korea 30 May-3 June 1994* pp. 63–69.

Wolff, E.A., S. Skogestad and M. Hovd (1992). Controllability of integrated plants. AIChE Spring National Meeting Paper 67a.

Woolverton, P.F. (1980). How to use relative gain analysis in systems with integrating variables. *InTech* pp. 63–65.

Wu, K.L. and C.-C. Yu (1996). Reactor/separator process with recycle-1. candidate control structure for operability. *Computers. chem. Engng.* **20**(11), 1291–1316.

Wu, K.L. and C.-C. Yu (1997). Operability for processes with recycles: Interaction between design and operation with application to the tennessee eastman challenge process. *Ind. Eng. Chem. Res.* **36**(6), 2239–2251.

Yi, C.K. and W.L. Luyben (1995). Evaluation of plant-wide control structures by steady-state disturbance sensitivity analysis. *Ind. Eng. Chem. Res.* **34**, 2393–2405.

Youla, D.C., J.J. jr. Bongiorno and C.N. Lu (1974). Single-loop feedback-stabilization of linear multivariable dynamical plants. *Automatica* **10**, 159–173.

Zheng, A. and R.V. Mahajannam (1999). A qualitative controllability index and its applications. *Ind. Eng. Chem. Res.* pp. 999–1006.

Zheng, A., R.V. Mahajannam and J.M. Douglas (1999). Hierarchical procedure for plantwide control system synthesis. *AIChE Journal* **45**, 1255–1265.