Bernt of Wittgens

STUDIES ON CONTROL STRUCTURE SELECTION AND DESIGN OF ROBUST DECENTRALIZED AND SVD CONTROLLERS

by

Morten Hovd

A Thesis Submitted for the Degree of Dr. Ing.

University of Trondheim

The Norwegian Institute of Technology

Submitted September 3, 1992

NTH-TRYKK
1992

ACKNOWLEDGEMENTS

This thesis would never have been written without the help and support I have received from a number of people. Most importantly, I am indebted to my wife Ellen, for her love, encouragement and patience.

I also would like to thank my supervisor, Sigurd Skogestad, for the insight and the inspiration he has given me during my Dr. Ing. studies. No graduate student can hope for a better supervisor.

Throughout my Dr. Ing. studies, Sigurd has had the largest group of Dr. Ing. students in the Chemical Engineering Department. I have therefore always been able to discuss the ideas and problems I have had with other people, also when Sigurd has not been available. My discussions with Petter Lundström have been especially helpful, because of his understanding of robust control and the related computer software.

The autumn of 1990 I spent at the California Institute of Technology, working in the group of Professor Manfred Morari. I am grateful to Professor Morari for letting me stay in his group and for the guidance he gave me during those months. I also would like to thank two of Professor Morari's students, Jay H. Lee and Richard D. Braatz, Jay for his contributions to Appendix A in this thesis and Richard for his help with writing Chapter 7.

I am also grateful to Dag Ljungquist of the Engineering Cybernetics Department for helping me getting started with FCC modelling and control, and for many interesting discussions on these topics.

Financial support from the Royal Norwegian Council for Scientific and Industrial Research (NTNF) is gratefully acknowledged.

ABSTRACT

This thesis is concerned with different aspects of structure in the control system for chemical plants.

- **Control structure selection for the regulatory control level.** A regulatory control level that performs well is a prerequisite for the higher levels in the control system (supervisory control and plant-wide optimization) to realize their full potential benefits. The regulatory control level is typically highly decentralized. Three decisions therefore have to be made in the control structure selection for the regulatory control level:

  1. Selection of controlled variables.
  2. Selection of manipulated variables.
  3. Pairing of controlled and manipulated variables.

  The control structure can have a strong influence on the achievable control quality for the regulatory control level. In this thesis tools are developed that can aid the engineer in selecting a good control structure, both for stable and unstable plants.

  The tools for control structure selection are applied to the riser-regenerator section of the fluid catalytic cracking process. The resulting findings resolve contradictory claims in the literature on control structure selection for the fluid catalytic cracking process.

- **Design of decentralized controllers.** Standard controller synthesis algorithms, e.g. H_2- or H_∞-synthesis, cannot handle a requirement for a specific structure for the controller. Instead three practical approaches to the design of decentralized (i.e., highly structured) controllers have evolved: parameter optimization, sequential design and independent design. The advantages and disadvantages of these three design methods are discussed, and improvements are made to the independent design and sequential design methods.

- **SVD controllers.** In some cases the requirement for a specific structure for the controller does not come from an à priori decision, but arises from analysis of the plant itself. In this thesis three classes of design problems are found for which the optimal controller has the structure of an SVD controller for H_2-, H_∞ or μ-optimal control. This knowledge about the structure of the optimal controller can be used to simplify controller design. The simplification of controller design is especially dramatic for plants consisting of nominally identical units working in parallel.

Model Predictive Control (MPC) has found wide application for supervisory control of chemical plants. Step response models are very common in MPC algorithms. If

fast control is desired for plants with a slow open loop response, it has been necessary to use a large number of step response coefficients. A large number of step response coefficients will put a large computational load on the control system. In Appendix A a method is developed for reducing the number of step response coefficients needed.

Contents

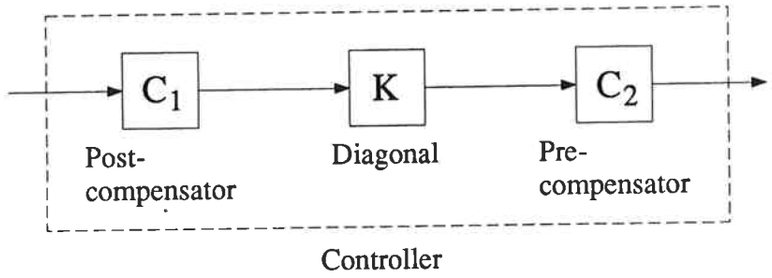


Figure 1.1: Structured multivariable controller decomposed into precompensator, diagonal matrix, and postcompensator.

- **"Control structure selection"**: Step 1 and (for a structured controller) Step 2a.
- **"Control configuration selection"**: Step 2a.

The reason we include Step 2a in the control structure selection is that for decentralized controllers the issue of actuator and measurement selection is intimately linked with the control configuration selection (pairing of actuators and measurements).

The difference between the words "structure" and "configuration" may seem minor, but note that they are significantly different within the context of this thesis. There is also a linguistic basis for making this distinction; Webster's dictionary [38] defines "structure" as *"the arrangement or interrelation of all parts of a whole"* and "configuration" as simply *"the arrangement of parts"*.

The term "structure of the controller" needs to be clarified (it is very different from "controller structure"). Many standard controller synthesis algorithms for multivariable systems (e.g., H_2- or H_∞ synthesis) generally give controllers with no apparent structure. However, there are many design methods which lead to highly structured controllers. These design methods typically involve controllers with the structure shown in Fig. 1.1. The compensators C_1 and C_2 usually contain only simple dynamics or no dynamics at all, and primarily deal with multivariable effects. The main dynamics of the controller is located in the diagonal matrix K[1], which handles requirements for integral action, speed of response, etc. Examples of controllers with the structure shown in Fig. 1.1 are:

1. Decentralized controllers: $C_1 = C_2 = I$.
2. SVD controllers: C_1 and C_2 come from the singular value decomposition of the plant.
3. Inverse-based controllers: $C_1 = I$, $C_2 = G^{-1}$, where G is the plant.

Other examples of controllers with the structure shown in Fig. 1.1 are one-way decouplers and the compensator designs of MacFarlane and of Rosenbrock, see [19].

[1]Chapter 6 also considers controllers for which K in Fig. 1.1 is a block diagonal matrix

One reason for using controllers with the structure of Fig. 1.1 is that the compensators C_1 and C_2 may not have to be changed much if the operating conditions change, and this makes retuning the controller relatively simple, as one just has to retune the diagonal elements of K in Fig. 1.1.

1.1 Hierarchical Structuring of the Control System

1.1.1 Objectives of the Control System

A control system is required to fulfill a multitude of objectives. These include:

- Safe operation of the plant. The control system must ensure that injuries to humans and damage to process equipment are avoided.
- Operation in accordance with environmental regulations. The release of pollutants should be within legal limits.
- The control system must facilitate startup and shutdown of the plant and of individual units within the plant.
- Maintain consistent operating conditions and quality specifications.
- Economically optimal operation. Within the limits imposed by the objectives mentioned above, the control system should identify the economically optimal operating conditions, and enable the operators to keep the operating conditions close to these conditions. The economically optimal operating conditions will depend on a number of external factors, such as the price and quality of the available feedstock, the market prices for the range of product qualities the plant can produce, and possibly ambient weather conditions. In addition scheduled or unscheduled shutdown of individual processes within a plant can affect the optimal operating conditions. The control system must therefore be able to identify and track changes in the optimal operating conditions.
- The control system should be easy to understand for the operating personnel, and the effect of changes should be easy to anticipate.

During the history of process control, the emphasis placed on each of these objectives have varied. Of course, safety and the ability to start up, operate and shut down the plant have always been emphasized. The importance of environmental regulations have increased as regulations have become more strict. The importance of operator understanding and acceptance has become clear as the development in hardware has enabled the implementation of more complex control systems. In the early years of the chemical process industry, the computational power necessary to calculate the economically optimal operating conditions (accurately and with frequent updates) for a large industrial plant was simply not available.

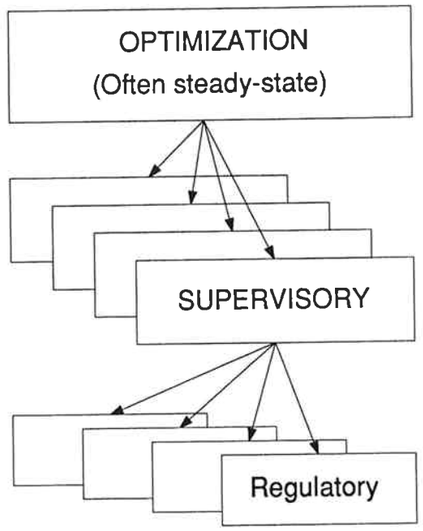


Figure 1.2: Typical decomposition of control tasks

1.1.2 Typical Control System Organization

The organization of a typical control system for a large modern chemical plant reflects how current practice in process control has evolved from earlier practice. The organization of a control system for a typical large chemical plant is illustrated in Fig. 1.2.

At the lowest level one finds the regulatory control system, which keeps a set of measurements at setpoints specified by the operators or the higher levels in the control system. The regulatory control system is usually highly decentralized, consisting mainly of single input single output (SISO) feedback loops. Feedforward may be used in some cases, and cascades are often used for flow control. True multivariable controllers are relatively rare. Thus, although the regulatory control system is implemented using electronics and computers, it has changed little conceptually from the days of SISO pneumatic controllers.

The middle level in the control system hierarchy may be termed supervisory control. This level in the control system usually coordinates several alternative manipulated variables in order to avoid saturation in critical loops in the regulatory control system, and maintain a set of measurements at setpoints or within maximum and minimum limits. The supervisory control system commonly uses a model of the process to predict the future values of the controlled variables. If no constraints are active or are predicted to become active, some manipulated variables are commonly reset to predefined ideal resting values. The concept of using a model to predict future outputs of the plant has become known as Model Predictive Control (MPC), and many control algorithms apply this concept, e.g., DMC [3] and GPC [2]. Supervisory control first appeared in

the 1970s, and has become very popular during the 1980s.

The next level in the control system hierarchy is steady state optimization. At this level the optimal plant conditions are calculated simultaneously for part of the plant or for the entire plant. Plant wide steady state optimization became routine in the 1960s and 1970s, and is now performed regularly in many modern plants. Usually, the lower level uses linear models, the middle level uses linear models but allows for constraints, while the top level uses a nonlinear model (or a linear model which is updated based on linearization).

In addition there is of course higher levels of decision-making, like purchase of raw materials, scheduled shutdown for maintenance, construction of new facilities and so on. Likewise, there is an emergency shutdown system which may be considered to lie below the regulatory control level. However, these additional levels are usually not considered to be a part of the control system.

Some work can be found in the literature about *how* to decompose the control system in a manner similar to what is described above (e.g. Morari et al. [22], Umeda et al. [37]). However, there is little justification in the literature for *why* this decomposition is needed, the most comprehensive discussion about the need for decomposition appears to be given by Mesarovic [21].

The tasks of control structure selection and control configuration selection can be applied to the entire control problem as discussed above. However, in this thesis these tasks are mainly addressed within the regulatory control system.

1.1.3 Centralized Control System

As explained above, the control system in most large plants can be decomposed into three different layers, and the middle layer can be further decomposed into separate units with limited information interchange between each unit. It is clear that a control system which can be decomposed in this way cannot be truly optimal in the sense that it will achieve a lower rate of profit (within the constraints imposed by safety considerations and environmental regulations) over any prolonged period of time compared to a hypothetical control system that uses a dynamic model of the entire plant to optimize plant operation continuously. All control actions in such a hypothetical, ideal control system would be perfectly coordinated, and the control system would use dynamic optimization in real time instead of steady state optimization performed comparatively infrequently[2]. Most chemical processes contain nonlinearities, and there are also nonlinearities in the relationship between quality and price for feedstocks and products. A hypothetical, ideal control system would take full advantage of these nonlinearities (although there are still many unresolved issues in nonlinear control theory), in contrast to a traditional control system which is based mainly on linear control theory.

With the increasing availability of computing power it is therefore pertinent to consider whether the conventional decomposition of the control system should be discarded and attempts should be made to design and implement control systems which

[2]Vendors of supervisory control systems may claim that their package performs online economic optimization, but the optimization performed in supervisory control systems is usually much too simple to be considered a realistic economic optimization.

more closely approximate the hypothetically optimal. Such a control system will in the following be termed an "optimal centralized control system", because it will consist of only one "layer", in which all the tasks of the three different layers of a traditional control system are centralized.

locally remove uncertainty in decentralized loops

1.1.4 Drawbacks of an Optimal Centralized Control System

The advantage of the optimal centralized control system has already been pointed out: When an optimal centralized control system is operating successfully, plant profitability will be higher than for the same plant controlled by a traditional control system. The drawbacks of an optimal centralized control system are not so clear, and need to be considered more carefully.

- **Economics.** It is clear that the view of economics taken above is much too simplistic. The costs of designing and maintaining a control system must also be considered. These costs will be much higher for an optimal centralized control system than for a traditional control system, because one cannot reduce problem complexity by decomposing the problem when designing an optimal centralized control system. The need for control systems with a minimum degree of complexity is emphasized by Reeves et al. (1991), for the reasons of system cost, maintainability and reliability. The top (optimization) layer in a traditional control system only uses a steady state model, which is much easier to build than the plant wide dynamic model needed for an optimal centralized control system. The middle, supervisory control level requires a dynamic model of its section of the plant, but can take advantage of the simplification in plant behavior caused by the control action of the lower (regulatory) control level.

  It should also be understood that regularity is very important for plant economics, as it often takes a long time to start up a chemical plant. Thus, if plant shutdown occurs more frequently with an optimal centralized control system than with a traditional control system, it is unlikely that overall profit will increase when installing an optimal centralized control system even if the profit per hour of operation increases significantly.

- **Redesign and retuning.** Because of changes in operating conditions, raw materials and equipment, it may be necessary to make adjustments to the control system. This is much easier in a traditional, structured control system, where it may be obvious where the changes should be made. In a centralized model-based system, a costly and time-consuming change in the plant model and in the mathematical objective function may be needed.

- **Startup and shutdown.** Common operating practice during startup is that virtually all controls are initially put on manual. The loops of the regulatory control system are then put in service when the equipment that they control approaches normal operating conditions. When the regulatory control level is in service, the supervisory control can be turned on. Shutdown is performed in the reverse sequence. Thus, there may be significant scope for improvement in the startup and shutdown procedures of a plant, as quicker startup and shutdown can reduce plant downtime. However, a model of a plant which in addition to normal operating conditions also is able to describe startup and shutdown is necessarily much more complex than a model which only covers the range of conditions encountered in normal operation. Building such a model will be difficult and costly. Startup and shutdown of a plant with an optimal centralized control system that does not cover startup and shutdown, may be more difficult than with a traditional control system, because it may not be possible to put an optimal centralized control system gradually into service.

- **Operator acceptance and understanding.** Control systems that are not accepted by the operators are likely to be taken out of service. An optimal centralized control system will often be complex and difficult to understand. Operator understanding obviously makes acceptance easier, and a traditional control system, being easier to understand, often has an advantage in this respect. Plant shutdowns may be caused by an operator with insufficient understanding of the control system. Such shutdowns should actually be blamed on the control system (or those who designed and installed the control system), since operators are an integral part of the plant operation, and their understanding of the control system must therefore be ensured.

  On the other hand, it is clear that operator acceptance of a control system will be influenced by the available alternatives. Thus optimal centralized control systems have a better chance of being accepted for processes where traditional control systems have failed. One such example appears to be the Light Metal Electrolysis studied by Balchen and coworkers (e.g. [36]).

- **Hardware and software failure.** In a traditional control system the operators retain the help of the regulatory control system in keeping the plant in operation if a hardware or software failure occurs in the higher levels of the control system. A hardware backup system for the regulatory control level is much cheaper than for the higher levels in the control hierarchy, as the regulatory control system can be decomposed into simple control tasks (mainly single loops). In contrast, an optimal centralized control system requires powerful computers, and it is therefore more costly to provide a backup system. However, with continued increase in availability and decrease in price for computing power this argument may weaken.

- **Robustness.** The complexity of an optimal centralized control system will make it difficult to analyze whether the system is robust with respect to model uncertainty and numerical inaccuracies. Analyzing robustness need not be trivial even with a traditional control system. The ultimate test of robustness will be in operation of the plant. A traditional control system can be applied gradually, first the regulatory control, then section by section of the supervisory control level. When problems arise, it will therefore be easier to analyze the cause of the problem with a traditional control system than for an integrated control system.

- **Existing traditional control systems.** Where existing control systems perform reasonably well, it makes more sense to put the effort into improving the existing control system rather than to make the risky decision to design a new control system. This will apply also to many new plants, as many chemical processes are not well known. For such processes it will therefore be necessary to carry out model identification and validation on the actual process. During this period some minimum amount of control will be needed. The regulatory level in a traditional control system requires much less information about the process, and can operate during this period.

1.1.5 Conclusions on Hierarchical Structuring of Control Systems

Control systems for large chemical plants will probably continue to evolve from their traditional structure. Major breaks with tradition are most likely to occur where traditional methods for process control have been clearly unsuccessful.

Current research into improving the supervisory control algorithms (e.g. [17]) appears to be effort well spent. However, improved supervisory control does not imply that the other levels of the control hierarchy become less important. A well designed and structured regulatory control system is a prerequisite for applying supervisory control, as a poorly designed regulatory control system may impose inherent control limitations which cannot be removed by the higher levels in the control system. Also, a well designed regulatory control system will simplify the design of a supervisory control system, and will be of help to the process operators during startup and shutdown, and when the higher levels of the control hierarchy are out of service.

1.2 The Regulatory Control Level

In the preceding section it was mentioned that the regulatory control level is usually highly decentralized, consisting mainly of SISO loops. Some reasons for choosing the regulatory control level to be decentralized are:

- **Operator understanding and acceptance.** The need for operator understanding and acceptance is especially relevant for the regulatory control level, as this is the control level that operators will most frequently interact with. The operators may also occasionally have to keep the plant in operation without the help of the higher levels in the control hierarchy, when the higher levels are out of service.

- **Ease of tuning.** A result of the fact that a decentralized control system is easy to understand, is that it is also relatively easy to retune. Operators may therefore be allowed to retune the loops of the regulatory control level in order to counter the effect of process changes. Retuning of the loops at the regulatory control level will often reduce the need for changing the higher levels of the control system.

- **Startup and shutdown.** A well-designed decentralized control system can be taken into or out of service loop by loop, as noted in the preceding section.

- **Implementation and maintenance.** A decentralized control system is relatively easy to implement and less costly to maintain, due to its low degree of complexity.

A more complete list of objectives for the regulatory control system is given in Chapter 4. Prior to the application of any controller design methodology, the manipulated and controlled variables of the control system must be identified.

Choosing all available manipulators and measurements will often make the control system unnecessarily complex [24]. The regulatory control level typically provides faster control than the higher levels of the control system. Thus, if fast control is required for a given measurement (e.g. for reasons of stability or safety), it will normally be controlled at the regulatory control level. It is also natural to control a measurement at the regulatory control level if the control of that measurement helps the operators keep the plant in operation when the higher levels of the control system are out of service.

Obviously, the manipulated variables in the regulatory control system are chosen to enable good control of the controlled variables. Using more manipulated variables than controlled variables (e.g. split range control, using one fast and one slower manipulated variable) can sometimes make the control easier. However, the number of manipulated variables should be the minimum necessary for effective control of the controlled variables, as one otherwise restricts the freedom for design of the higher levels of the control system.

A very important issue in the design of the regulatory control system is therefore the *control structure selection*. That is, one has to select controlled and manipulated variables, and decide what manipulated variable should be used for controlling a given controlled variable (also known as the *pairing* of controlled and manipulated variables). Several books contain recommendations and rules for control structure selection, e.g. Stephanopoulos [35], Rijnsdorp [25] and Balchen and Mummé [1]. However, the issue of control structure selection is still far from being resolved, and there is a need for a more quantitative treatment of the issue than what is given in these books.

1.3 Design of Controllers with a Specific Structure

1.3.1 Decentralized Controllers

After the control structure of a decentralized control system has been determined, one is faced with the problem of designing a decentralized controller. Unfortunately, standard controller synthesis algorithms, such as H_2 or H_∞ synthesis cannot handle a requirement for a specific structure for the controller. Instead, some practical approaches to the design of decentralized controllers have evolved:

- **Parameter optimization.** This involves an *à priori* parametrization of the controller, and a computable measure of control quality (e.g. the H_2 or H_∞

norm of the closed loop system). The tuning parameters of the controller are then optimized with respect to the chosen measure of control quality.

- **Sequential design.** Sequential design involves closing one control loop at the time, and treating the design of each controller element as a SISO problem. The controllers that have already been designed are assumed to be in service when designing controllers for subsequent loops. Sequential design was introduced into the control literature by Mayne [20].

- **Independent design.** In independent design one attempts to find bounds on the design of the individual loops. If these bounds are fulfilled for *all* loops, the overall closed loop system is guaranteed to perform satisfactorily. Independent design was introduced by Skogestad and Morari [30].

1.3.2 Other Controllers with a Specific Structure

In some cases a specific structure for the controller does not result from an à priori decision, but arises from analysis of the controller synthesis problem itself. This occurs when the controller synthesis problem has some structure which makes it possible to transform the problem such that it decomposes into non-interacting subproblems. After decomposition, each of the synthesis subproblems will be of lower dimension and contain fewer states than the original problem. Controller synthesis can therefore be significantly simplified by such decomposition. The resulting overall controller will then have a structure which is determined by the matrices used for transforming the synthesis problem.

1.4 The use of the term "Controllability"

In state-space control theory the term "controllability" has a rather limited definition in terms of Kalman's *state* controllability. If all the states of a system are controllable, any value for the state vector can be achieved in finite time by the appropriate use of the manipulated inputs.

However, in engineering practice, a system is called controllable if it is possible to achieve the specified aims of the control, whatever these may be (Rosenbrock, [26], p. 171). At the risk of upsetting some readers with a background from state-space control theory, the term controllability will throughout this thesis be used in a manner which is more consistent with its use in engineering practice than in state-space control theory. The definition of controllability used in this thesis is:

Controllability (of a plant) is the best quality of the response which can be obtained for the plant by use of feedback control.

Admittedly, this definition is not very precise, since, for example, "best" is not defined. The definition of controllability used here is similar to that used by Perkins [27]. A key idea is that controllability is an inherent property of the plant, and that it is independent of controller tuning (it is assumed that the optimal tunings are used).

On the other hand, one may restrict the class of allowable controllers, for example by considering "controllability using linear controllers" (which is done throughout this thesis) or "controllability using decentralized controllers" (which we do in most parts of this thesis).

1.5 Robust control

No model of a plant is perfect, any model will contain uncertainties and inaccuracies. It is therefore not sufficient that a controller gives a stable system and performs well on the model for which it was designed, when applied to the actual plant it must also give a stable system and acceptable performance. One therefore needs a method for ascertaining à priori whether the controller will perform acceptably when applied to the *actual plant*. The problem here is obvious: It is impossible to know exactly how the actual plant behaves. This problem can be circumvented by finding estimates of the *locations*, *structures* and *magnitudes* of the uncertainties in the model, and use a controller design method which guarantees stability and performance for the whole class of possible plants described by the model and the uncertainty estimates. That is, the controller should be *robust* with respect to the uncertainties of the plant. If the closed loop system can be shown to be stable for the whole class of possible plants, the system has Robust Stability (RS). Likewise, if the closed loop system can be shown to fulfill the performance specification for the whole class of possible plants, the system has Robust Performance (RP).

The H_∞-norm of a transfer function matrix $M(j\omega)$ is defined as $\sup_\omega \bar{\sigma}(M(j\omega))$ (where $\bar{\sigma}$ denotes the maximum singular value). A common frequency domain specification for control system is that the weighted H_∞-norm of some matrix should not exceed some specific value. Doyle [4] introduced the *structured singular value*, μ. The structured singular value is used to analyze whether a system fulfills H_∞-norm specifications in a robust sense, and for designing controllers which result in closed loop systems that robustly fulfill their H_∞-norm specifications.

1.6 Thesis Overview

Chapter 2. This chapter gives frequency-dependent tools to assist evaluating controllability and to use for control configuration selection and design of decentralized control systems. In a broader sense these tools may also be use for control structure selection. The relationship between the Relative Gain Array (RGA) and Right Half Plane (RHP) zeros is clarified. The use of the RGA as a measure of the sensitivity of the closed loop system with respect to input uncertainty and individual element uncertainty is discussed.

The Performance Relative Gain Array (PRGA) and Closed Loop Disturbance Gain (CLDG) are introduced. The PRGA and CLDG can be used to find bounds on the designs of the individual loops for decentralized control. These bounds are valid at frequencies below the closed loop bandwidth, and fulfilling the bounds ensures that the

performance requirements with respect to setpoint following and disturbance rejection are met. The bounds may be used to evaluate the inherent sensitivity to disturbances for a particular decentralized control configuration.

Chapter 3. This chapter deals with controllability analysis for open loop unstable systems. The implications of the presence of RHP transmission zeros on control structure selection are discussed.

The well known pairing criteria based on the Niederlinski index and the steady state RGA are generalized to unstable systems. The use of several frequency-dependent tool for control structure selection for unstable plants are discussed and compared.

Chapter 4. The tools developed in Chapter 2 are applied to the problem of control structure selection for the regulatory control system for the riser-regenerator section of a Fluid Catalytic Cracking process. Both the partial combustion mode and the complete combustion mode of operation are studied. However, the main focus is on the partial combustion mode, mainly in order to enable comparison with the results of previous authors.

The conventional control structure [28] and the Kurihara control structure [16] are both shown to have inherent bandwidth limitations due to RHP transmission zeros. It is demonstrated how RHP transmission zeros can be avoided by proper choice of controlled outputs. The RGA and PRGA are used to find pairings of controlled and manipulated variables, and the effects of disturbances are investigated using the CLDG.

The sensitivity of the results with respect to changes in the operating point and errors in the parameter values of the model are investigated. It is also shown how different assumptions about the model structure affect the results of the controllability analysis.

Chapter 5. This chapter is concerned with the design of decentralized controllers. It is here assumed that the control configuration (pairing) has been determined prior to the design, e.g. by using the tools of Chapters 2-3. Chapter 5 gives new results on independent and sequential design of decentralized controllers.

Independent design involves treating some controller-related quantity as uncertainty, and thereafter finding bounds on the allowable magnitude of this uncertainty. For independent design to be powerful, the controller-related quantity for which the bounds are found must be chosen such that the associated uncertainty is small. Expressing the set of possible designs in terms of bounds on the sensitivity function and complementary sensitivity function of the individual loops, as done by Skogestad and Morari [30], is shown to be conservative in most cases. Instead, finding bounds on the IMC filter time constant, which corresponds to a much smaller class of possible designs, is shown to be much less conservative. It is demonstrated that by finding bounds on the IMC filter time constant, it is possible to perform independent design on problems that have previously been claimed to be impossible to solve with independent design.

For sequential design, it is shown how simple estimates of the interaction can be included into the design problem at each stage of the design.

Chapter 6. This chapter is concerned with the control of plants consisting of identical, interacting units operating in parallel. Such plants are common in industry, as in many cases a single unit would be too large to be practical. This type of plants has structural

properties which allows decomposition of the controller design problem into smaller subproblems. For plants consisting of n units in parallel, each unit having n_o outputs and n_i inputs, the controller design problem can be performed on two "plants" with n_o outputs and n_i inputs. For most cases these two "plants" of dimension $n_o \times n_i$ can be considered independently. Dramatic reductions in the size of the controller design problems can therefore be achieved for problems where n is a large number.

The *structure* of the resulting controller will only depend on the numbers n, n_o and n_i. For plants with SISO units in parallel ($n_o = n_i = 1$), the resulting controller will have the structure of an SVD controller.

Chapter 7. SVD controllers are studied in a more general perspective in this chapter. The motivation for studying SVD controllers is the results in Chapter 6, and the fact that Engstad [5] found numerically that the μ-optimal controller found for the ill-conditioned distillation example introduced by Skogestad et al. [29] is on this form. Indeed, it is found by using simple unitary transforms that H_2-, H_∞-, and μ-optimal controllers are on the form of an SVD controller for quite a large class of problems, including the example in [29]. This is interesting from a theoretical point of view, but also suggests that multivariable controllers based on SVD-compensators plus simple single-loop controllers may be close to optimal in many cases. That is, one may introduce a controller of the form $C = VKU^T$ (see Fig. 1.1), where V and U are constant matrices, and use the methods outlined in Chapter 5 to design the decentralized controller K.

Chapter 8. Conclusions and suggestions for future research are given in this chapter.

Appendix A. The appendix contains material which is somewhat on the side of the main topics of this thesis. For Model Predictive Control (MPC) based on step response models, fast sampling is required for fast control. However, if the plant contains outputs in which the responses are slow, fast sampling means that a large number of step response coefficients are needed. A large number of step response coefficients leads to a large requirement for storage and computing power in the control system. A robust method for reducing the number of step response coefficients needed is given in Appendix A. This work was carried out under the guidance of Professor Manfred Morari during a visit to the California Institute of Technology in 1990.

Most of the material in this thesis has been presented previously or is submitted for presentation:

- Section 1.1 of Chapter 1 roughly corresponds to this authors contribution to [39].
- Chapter 2 consists of material from [32] and [10], and related material can be found in [33] and [34].
- Chapter 3 is presented in [12].
- Most of Chapter 4 is presented in [9], and some preliminary results are presented in [6] and [8].
- Chapter 5 will be presented in [13] and [15].

- Parts of Chapter 6 is published in [11] and [31].
- Chapter 7 is submitted for presentation [14].
- Appendix A is presented in [7].

Overall this thesis is expected to result in 13 conference presentations and 9 journal articles. Other work in which the author has participated, but which is not covered in this thesis, can be found in references [18, 40].

References

[1] Balchen J. G. and Mummé, K. I. (1988). *Process Control. Structures and Applications*, Van Nostrand Reinhold, New York, USA.

[2] Clarke, D. W., Mothadi, C. and Tuffs, P. S. (1987). Generalized Predictive Control — Part I. The Basic Algorithm. *Automatica*, **23**, 2, pp. 137-148.

[3] Cutler, C. R. and Ramaker, B. L. (1980). Dynamic Matrix Control — A computer control Algorithm. *Proc. Joint Automatic Control Conf.*, San Fransisco, California, USA.

[4] Doyle, J. C., Wall, J. E. and Stein, G. (1982). Performance and Robustness Analysis for Structured Uncertainty. *Proc. IEEE Conf. Decision Contr.*, Orlando, FL., pp. 629-636.

[5] Engstad, P. K. (1991). Comparison of Robustness Between Different Control Algorithms. *Diploma Thesis*, Dept. of Engineering Cybernetics, University of Trondheim-NTH.

[6] Hovd, M., Lundström, P. and Skogestad, S. (1990). Controllability Analysis Using Frequency-dependent Measures for Interactions and Disturbances. *AIChE Annual Meeting*, paper 321j, Chicage, Illinois, USA.

[7] Hovd, M., Lee, J. H. and Morari, M. (1991). Model Requirements for Model Predictive Control. *Proc. European Control Conference*, Grenoble, France, pp. 2428-2433.

[8] Hovd, M. and Skogestad, S. (1991). Impact of Model Uncertainty on Control Structure Selection for the Fluid Catalytic Cracking Process. *Preprints IFAC Symposium ADCHEM'91*, Toulouse, France, pp. 107-112.

[9] Hovd, M. and Skogestad, S. (1991). Controllability Analysis of the Fluid Catalytic Cracking Process. *AIChE Annual Meeting*, Los Angeles, CA, paper 145g.

[10] Hovd, M. and Skogestad, S. (1992). Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. To appear in *Automatica*.

[11] Hovd, M. and Skogestad, S. (1992). Robust Control of Systems Consisting of Symmetrically Interconnected Subsystems. *Proc. ACC*, Chicago, IL, pp. 3021-3025.

[12] Hovd, M. and Skogestad, S. (1992). Controllability Analysis for Unstable Processes. *Preprints IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

[13] Hovd, M. and Skogestad, S. (1992). Design of Robust Decentralized Controllers. *AIChE Annual Meeting*, Miami Beach, FL, paper 127d.

[14] Hovd, M., Braatz, R. D. and Skogestad, S. (1993). On the Structure of the Robust Optimal Controller. Submitted to *12th IFAC World Congress, Mini-Symposium on Robust Control Design*, Sydney, Australia.

[15] Hovd, M. and Skogestad, S. (1993). Improved Independent Design of Decentralized Controllers. Submitted to *12th IFAC World Congress, Mini-Symposium on Chemical Process Control*, Sydney, Australia.

[16] Kurihara, H. (1967). Optimal Control of Fluid Catalytic Cracking Processes. *Ph.D. Thesis* MIT.

[17] Lee, J. H., Morari, M. and Garcia, C. E., (1992). State Space Interpretation of Model Predictive Control. Accepted for presentation in *Automatica*, preprint.

[18] Lundström, P., Skogestad, S., Hovd, M. and Wang, Z.-Q. (1991). Non-uniqueness of Robust H_∞ Decentralized PI Control. *Proc. ACC*, Boston, MA, pp. 1830-1035.

[19] Maciejowski, J. M. (1989). *Multivariable Feedback Design*, Addison Wesley, Wokingham, England.

[20] Mayne, D. Q. (1973). The Design of Linear Multivariable Systems. *Automatica*, **9**, pp. 201-207.

[21] Mesarovic, M. D. (1970). Multilevel Systems and Concepts in Process Control. *Proceedings of the IEEE*, **58**, 1, pp. 111-125.

[22] Morari, M., Arkun, Y. and Stephanopoulos, G. (1980). Studies in the Synthesis of Control Structures for Chemical Processes. Part I. *AIChE Journal*, **26**, 2, pp. 220-232.

[23] Niederlinski, A. (1971). A Heuristic Approach to the Design of Linear Multivariable Interacting Control Systems. *Automatica*, **7**, pp. 691-701.

[24] Reeves, D. E., Nett, C. N. and Arkun, Y. (1991). Control Configuration Design for Complex Systems: A Practical Theory. Submitted to *IEEE Transactions on Automatic Control.*

[25] Rijnsdorp, J. E. (1991). *Integrated Process Control and Automation*, Elsevier, New York, USA.

[26] Rosenbrock, H. H. (1970). *State-Space and Multivariable Theory.* Nelson, London, England.

[27] Perkins, J. D. (1989). Interactions Between Process Design and Process Control. *Preprints IFAC Symposium DYCORD+89*, Maastricht, The Netherlands.

[28] Pohlenz, J. B., (1963). How operational variables affect fluid catalytic cracking. *Oil and Gas Journal*, April 1, pp. 124-153.

[29] Skogestad, S., Morari, M. and Doyle, J. C. (1988). Robust Control of Ill-conditioned Plants: High Purity Distillation. *IEEE Trans. Autom. Control*, **33**, 12, pp. 1092-1105.

[30] Skogestad, S. and Morari, M. (1989). Robust Performance of Decentralized Control Systems by Independent Design. *Automatica*, **25**, 1, pp. 119-125.

[31] Skogestad, S., Lundström, P. and Hovd, M. (1989). Control of Identical Parallel Processes. Presented at AIChE Annual Meeting, San Francisco, CA, Paper no. 167 Ba.

[32] Skogestad, S. and Hovd, M. (1990). Use of Frequency-dependent RGA for Control Structure Selection. *Proc. American Control Conference*, San Diego, CA.

[33] Skogestad, S., Hovd, M. and Lundström, P. (1991). Simple Frequency-dependent tools for Analysis of Inherent Control Limitations. *Modeling, Identification and Control*, **12**, pp. 159-177.

[34] Skogestad, S., Hovd, M. and Lundström, P. (1991). Towards Integrating Design and Control: Use of Frequency-dependent Tools for Controllability Analysis. *Proc. Process Systems Engineering (PSE)*, Canada, pp. III.3.1-III.3.15.

[35] Stephanopoulos, G. (1984). *Chemical Process Control. An Introduction to Theory and Practice.* Prentice-Hall, Englewood Cliffs, New Jersey, USA.

[36] Strand, S. (1991). Dynamic Optimisation in State-space Predictive Control Schemes. *Dr. Ing. Thesis*, University of Trondheim-NTH, Norway.

[37] Umeda, T., Kuriyama, T. and Ichikawa, A., 1978. A Logical Structure for Process Control System Synthesis. *Proc. IFAC Congress*, Helsinki, pp. 271-278.

[38] *Webster's New World Dictionary of the American Language, College Edition*, (1959). The New World Publishing Company, New York, USA.

[39] Wolff, E. A., Skogestad, S. and Hovd, M. (1992). Controllability of Integrated Plants. Presented at *AIChE Spring National Meeting*, New Orleans, paper 67a.

[40] Wolff, E. A., Skogestad, S., Hovd, M. and Mathisen, K. W. (1992). A Procedure for Controllability Analysis. *Preprints IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

Chapter 2

Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design

Morten Hovd and Sigurd Skogestad *
Chemical Engineering
University of Trondheim, NTH
N-7034 Trondheim, Norway

Extended version of paper accepted for publication in *Automatica*

Abstract

The paper presents results on frequency-dependent tools for analysis, structure selection and design of control systems. This includes relationships between the relative gain array (RGA) and right half plane zeros, and the use of the RGA as a sensitivity measure with respect to individual element uncertainty and diagonal input uncertainty. It is also shown how frequency-dependent plots of the closely related performance relative gains (PRGA) and a new proposed disturbance measure, the closed-loop disturbance gains (CLDG), can be used to evaluate the achievable performance (controllability) of a plant under decentralized control. These controller-independent measures give constraints on the design of the individual loops, which when satisfied e that the overall system satisfies performance objectives with respect to setpoint tracking and disturbance rejection.

*Author to whom correspondence should be addressed. E-mail: SKOGE@KJEMI.UNIT.NO, phone: 47-7-594154, fax: 47-7-591410

2.1 Introduction

The relative gain array (RGA) has found widespread use as a measure of interaction and as a tool for control structure selection for single-loop controllers. It was first introduced by Bristol [5]. It was originally defined at steady-state, but it may easily be extended to higher frequencies [6]. Shinskey [19, 20] and McAvoy [12] have demonstrated practical applications of the RGA. Important advantages with the RGA is that it depends on the plant model only and that it is scaling independent. It is straightforward to generalize the RGA from single-loop controllers to block-diagonal controllers by introducing the block relative gain (BRG) [10], and most of the results presented in this paper may be generalized in such a manner. However, to simplify the presentation, and because single-loop controllers are most common in practice, we shall consider only the RGA in this paper.

Our interest in the RGA as a frequency-dependent measure was initially focused on its use as a sensitivity measure with respect to model uncertainty [22] (see Theorem 2 and Eq. (2.12) below). However, based on its original definition as a steady-state interaction measure for single-loop control, it seemed reasonable that the frequency-dependent RGA should have some use as a performance or stability measure for decentralized control. Some interesting relationships and reports of encouraging applications presented by Nett [14] led us to investigate this in more detail.

Most authors have confined themselves to use the RGA at steady state, and a thorough review of the use and interpretation of the steady-state RGA is given by Grosdidier [8]. A frequency-dependent interaction measure Y, which is equivalent to the RGA for 2×2 systems, was introduced by Balchen [3] and Rijnsdorp [17] and is discussed for $n \times n$ systems in [4]. Balchen also gives some performance interpretation to his measure. Applications of the frequency-dependent RGA are given by McAvoy [11, 12].

We use the dynamic RGA as defined by Bristol [6]. Other definitions have also been proposed. Witcher and McAvoy [28] proposed a time domain definition of the RGA, as did Tung and Edgar [27]. Arkun [1, 2] has proposed measures (DBRG and Relative Sensitivity) which include the controller. Balchen and Mumme [4] generalize the measure Y to include the controller. However, one then looses one of the main advantages of the RGA which is that it depends on the plant model only. These alternative definitions are not considered in this paper.

In the paper we first consider the RGA as a general analysis tool and refer to some of its properties, which we believe are significant for engineering applications. However, the main part of the paper is devoted to decentralized control. We study stability and achievable performance (controllability) using simple frequency-dependent measures for interactions (PRGA) and disturbances (CLDG).

One of the main criticisms against the use of the RGA has been its "failure" to predict the poor performance one often has when using decentralized control for one-way interactive systems because the RGA matrix is identity in such cases. We propose a new measure, the Performance RGA (PRGA) which may be used to address also this performance issue.

2.2 Definitions

2.2.1 Relative Gain Array (RGA)

Consider a $n \times n$ plant $G(s)$.

$$y(s) = G(s)u(s) \tag{2.1}$$

The open loop gain from input u_j to output y_i is $g_{ij}(s)$ when all other outputs y are uncontrolled. Writing equation (2.1) as

$$u(s) = G^{-1}(s)y(s) \tag{2.2}$$

it can be seen that the gain from u_j to y_i is $1/[G^{-1}(s)]_{ji}$ when all other y's are perfectly controlled (e.g. [8]). The relative gain is the ratio of these "open-loop" and "closed-loop" gains. Thus a matrix of relative gains, the RGA matrix, can be computed using the formula

$$\Lambda(s) = G(s) \times (G^{-1}(s))^T \tag{2.3}$$

where the $\times$ symbol denotes element by element multiplication (Hadamard or Schur product). The inverse $G^{-1}(s)$ may be non-proper or non-causal, and a physical interpretation in terms of perfect control is of course not meaningful except at steady-state. This has caused many authors to discard use of a dynamic RGA, or to restrict its use to plants with no RHP-zeros [10]. This is unfortunate as the dynamic RGA as defined above proves to have a number of useful properties. Furthermore, we shall mainly consider the $\Lambda(s)$ as a function of frequency, $s = j\omega$, and in this case $\Lambda(j\omega)$ may be computed for any plant G except for frequencies corresponding to $j\omega$-axis zeros.

The RGA matrix as defined above has some interesting algebraic properties (eg., [8]):

a) It is scaling independent (eg., independent of units chosen for u and y). Mathematically, $\Lambda(D_1 G D_2) = \Lambda(G)$ where D_1 and D_2 are diagonal matrices.

b) All row and column sums equal one.

c) Any permutation of rows or columns in G results in the same permutation in the RGA.

d) If $G(s)$ is triangular (and hence also if it is diagonal), $\Lambda(G) = I$.

e) Relative perturbations in elements of G and in its inverse are related by $d[G^{-1}]_{ji}/[G^{-1}]_{ji} = -\lambda_{ij} dg_{ij}/g_{ij}$.

These properties are easily proven from the following expression for the individual elements of Λ

$$\lambda_{ij}(s) = (-1)^{i+j} \frac{g_{ij}(s) det(G^{ij}(s))}{det(G(s))} \tag{2.4}$$

Here G^{ij} denotes the matrix G with subsystem ij removed, that is, row i and column j is deleted.

2.2.2 Performance Relative Gain Array (PRGA)

One inadequacy of the RGA (eg., [12], p. 166) is that it, because of property d, may indicate that interactions is no problem, but significant one-way coupling may exist.

To overcome this problem we introduce the performance relative gain array (PRGA). The PRGA-matrix is defined as

$$\Gamma(s) = \tilde{G}(s)G(s)^{-1} \tag{2.5}$$

where $\tilde{G}(s)$ is the matrix consisting of only the diagonal elements of $G(s)$, i.e., $\tilde{G} = diag\{g_{ii}\}$. The matrix Γ was originally introduced at steady-state by Grosdidier [9] in order to understand the effect of directions under decentralized control. The elements of Γ are given by

$$\gamma_{ij}(s) = g_{ii}(s)[G^{-1}(s)]_{ij} = \frac{g_{ii}(s)}{g_{ji}(s)}\lambda_{ji}(s) \tag{2.6}$$

Note that the diagonal elements of RGA and PRGA are identical, but otherwise PRGA does not have all the nice algebraic properties of the RGA. PRGA must be recomputed whenever G is rearranged, whereas RGA only needs to be rearranged in the same way as G. PRGA is independent of *input* scaling, that is, $\Gamma(GD_2) = \Gamma(G)$, but it depends on output scaling. This is reasonable since performance is defined in terms of the magnitude of the outputs.

The measures above may be extended to non-square systems by introducing the pseudoinverse. However, the usefulness of the measures, at least for analyzing decentralized control, then seems limited. In the following $G(s)$ is assumed square.

2.3 The RGA as a General Analysis Tool

In this section we present some relationships involving the RGA of $G(s)$ which do not assume a decentralized control system. The results are based on the general definition of the RGA given by Eq. (2.3), and the physical interpretation preceding Eq. (2.3) is of limited interest in this case.

2.3.1 The RGA and Right Half Plane Zeros

Consider a transfer matrix $G(s)$. Bristol [5] claimed in his original paper and later [6, 7] that there is a relationship between RHP-zeros and negative values of $\lambda_{ii}(0)$, but Grosdidier et al. [8] showed with a counterexample that this is not true. However, as we shall see there proves to be a relationship if we assume that the loops have been paired such that $\lambda_{ii}(\infty)$ is positive.

Theorem 1 *Assume* $\lim_{s\to\infty}\lambda_{ij}(s)$ *is finite and different from zero. Let* $g_{ij}(s)det(G^{ij}(s))$ *have* z_{0n} *zeros at the origin,* z_{Rn} *zeros in the right half plane,* p_{0n} *poles at the origin and* p_{Rn} *poles in the right half plane. Similarly, let* $det(G(s))$ *have* z_{0d} *zeros at the origin,* z_{Rd} *zeros in the right half plane,* p_{0d} *poles at the origin and* p_{Rd} *poles in the right half plane. Define* $z_0 = z_{0n} - z_{0d}$, $p_0 = p_{0n} - p_{0d}$, $z_R = z_{Rn} - z_{Rd}$, *and* $p_R = p_{Rn} - p_{Rd}$. *Then the net change in the phase of* $\lambda_{ij}(j\omega)$ *as the frequency goes from 0 to* ∞ *is* $\frac{\pi}{2}(p_0 - z_0) + \pi(p_R - z_R)$.

Proof: See Appendix 1. If all elements of $G(s)$ are stable (all poles in the closed left half plane), then any net change in phase must have been caused by a different number of RHP zeros in $g_{ij}(s)det(G^{ij}(s))$ and $det(G(s))$. The direction of the phase change will then tell whether the numerator or denominator of (2.4) has the most RHP zeros. The theorem is useful, for example, if only the frequency-response of the plant is known. The following Corollary is even more useful since it only requires knowledge about the diagonal RGA-elements at $\omega = 0$ and $\omega = \infty$ which may be available even when the detailed dynamics are unknown.

Corollary 1 *Assume* $\lim_{s\to\infty}\lambda_{ij}(s)$ *is finite and different from zero. Consider a transfer matrix with stable elements and no zeros or poles at* $s = 0$. *If* $\lambda_{ij}(j\infty)$ *and* $\lambda_{ij}(0)$ *have different signs then at least one of the following must be true:*
a) $g_{ij}(s)$ *has a RHP zero.*
b) $G(s)$ *has a RHP transmission zero.*
c) $G^{ij}(s)$ *(ie., the subsystem with input* j *and output* i *removed) has a RHP transmission zero.*

That is, different signs of $\lambda_{ij}(j\infty)$ and $\lambda_{ij}(0)$ is a sufficient condition for the existence of RHP zeros or RHP transmission zeros. Any such zeros may be detrimental for control of the system. However, it is not a necessary condition, and there may be RHP-zeros present even if the RGA elements do not change sign. For example, adding a time delay or RHP-zero to an individual input or output channel will not affect the RGA as it may simply be viewed as a kind of scaling. [1] In most cases the pairings are chosen such that $\lambda_{ii}(\infty)$ is positive (usually close to 1, see pairing rule 2 below) and this confirms Bristols claim that negative RGA-elements imply presence of RHP-zeros. In the case when the process does contain zeros or poles at $s = 0$ the Corollary still applies if $\lambda_{ij}(0)$ is corrected by adding $\frac{\pi}{2}(p_0 - z_0)$ to the phase angle.

Example 1. Consider a plant

$$G(s) = \frac{1}{\tau s+1}\begin{pmatrix} s+1 & s+4 \\ 1 & 2 \end{pmatrix} \tag{2.7}$$

We have $\lambda_{11}(\infty) = 2$ and $\lambda_{11}(0) = -1$. Since none of the diagonal elements have RHP-zeros we conclude from Corollary 1 that $G(s)$ must have a RHP-zero. This is indeed confirmed as $G(s)$ has a transmission zero at $s = 2$.

2.3.2 The RGA and the Optimally Scaled Condition Number

Consider any complex matrix G. Bristol [5] pointed out the formal resemblance between the RGA and the condition number $\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G) = \bar{\sigma}(G)\bar{\sigma}(G^{-1})$ However, the condition number depends on scaling, whereas the RGA does not. Minimizing the

[1] Adding a time delay θ_i to each output i yields the plant D_1G where $D_1 = diag\{e^{-\theta_i s}\}$, but the RGA-matrix is unchanged since $\Lambda(D_1G) = \Lambda(G)$.

condition number with respect to all input and output scalings yields the optimal condition number

$$\gamma^*(G) = \min_{D_1,D_2} \gamma(D_1 G D_2) \tag{2.8}$$

It is commonly conjectured that there is a close relationship between $\gamma^*(G)$ and the magnitude of the elements in the RGA as is illustrated by the following lower and conjectured upper bounds on $\gamma^*(G)$:

$$\|\Lambda\|_m - \frac{1}{\gamma^*(G)} \leq \gamma^*(G) \leq \|\Lambda\|_1 + k(n) \tag{2.9}$$

where $\|\Lambda\|_m = 2\max\{\|\Lambda\|_{i1}, \|\Lambda\|_{i\infty}\}$ and $\|\Lambda\|_1 = \Sigma_{ij}|\lambda_{ij}|$,and $k(n)$ is a constant. The lower bound is proven by Nett and Manousiouthakis [15]. The upper bound is proven for 2×2 matrices with $k(2) = 0$ [8], but it is only conjectured for the general case with $k(3) = 1$ and $k(4) = 2$ [22].

2.3.3 RGA and Individual Element Uncertainty

Theorem 2 *The (complex) matrix G becomes singular if we make a relative change $-1/\lambda_{ij}$ in its ij-th element, that is, if a single element in G is perturbed from g_{ij} to $g_{pij} = g_{ij}(1 - \frac{1}{\lambda_{ij}})$.*

Proof. Let $G_p(s)$ denote $G(s)$ with g_{pij} substituted for g_{ij}. Using (2.4), we find by expanding the determinant of $G_p(s)$ by row i or column j that

$$det(G_p) = det(G) - \frac{det(G)}{(-1)^{i+j} det(G^{ij})}(-1)^{i+j} det(G^{ij}) = 0 \tag{2.10}$$

Theorem 2 provides necessary and sufficient condition for singularity of a matrix with element uncertainty. It is actually a quite amazing algebraic property of the RGA which seems to be little known. The theorem was originally presented by Yu and Luyben [30], but the proof above is much simpler.

Theorem 2 has some important control implications:

1. Element uncertainty. Consider a plant with transfer matrix $G(s)$. If the relative uncertainty in an element at a given frequency is larger than $|1/\lambda_{ij}(j\omega)|$ then the plant may have $j\omega$-axis zeros and RHP-zeros at this frequency. However, the assumption of element-by-element uncertainty is often poor from a physical point of view because the elements are usually always coupled in some way.

2. Process identification. Models of multivariable plants, $G(s)$, are often obtained by identifying one element at the time, for example, by using step or impulse responses. From Theorem 2 it is clear this method will most likely give meaningless results (eg., wrong sign of $det(G(0))$ or non-existing RHP- zeros) if there are large RGA-elements within the bandwidth where the model is intended to be used. Consequently, identification must be combined with physical knowledge if a good multivariable model is desired in such cases.

3. Uncertainty in state matrix. Consider a stable linear system written on state variable form; $dx/dt = Ax + \cdots$. Then changing the ij'th element in A from a_{ij} to $a_{ij}(1 - 1/\lambda_{ij}(A))$ yields one eigenvalue of A equal to zero. Thus, we may conclude that systems with large RGA-elements of A, will become unstable for small relative changes in the elements of A.

2.3.4 RGA and Diagonal Input Uncertainty

One kind of uncertainty that is *always* present is input uncertainty. Let the nominal plant model be $G(s)$, and the true (perturbed) plant be $G_p = G(I+\Delta)$. $\Delta = diag\{\Delta_i\}$ is a diagonal matrix consisting of the relative uncertainty (error) in the gain of each input channel. If an "inverse-based" controller (decoupler) is used, $C(s) = G^{-1}(s)K(s)$, where $K(s)$ is a diagonal matrix, then the true open loop gain G_pC is

$$G_pC = (I + G\Delta G^{-1})K \tag{2.11}$$

Result: The diagonal elements of $G\Delta G^{-1}$ are directly given by the RGA [22]

$$(G\Delta G^{-1})_{ii} = \Sigma_{j=1}^{n} \lambda_{ij}(G)\Delta_j \tag{2.12}$$

Thus, if the plant has large RGA elements and an inverse-based controller is used, the overall system will be extremely sensitive to input uncertainty.

Control implications. Consider a plant with large RGA-elements in the frequency-range of importance for feedback control. A diagonal controller is robust (insensitive) with respect to input uncertainty, but will be unable to compensate for the strong couplings (as expressed by the large RGA- elements) and will yield poor performance (even nominally). On the other hand, An inverse-based controller which corrects for the interactions may yield excellent nominal performance, but will be very sensitive to input uncertainty and will not yield robust performance. The physical reason for the problems with the inverse-based controller is that the controller tries to apply large input signals in certain directions to match weak directions in the plant. The input uncertainty changes these directions and ruins the desired match. In addition, stability problems are also expected for the inverse-based controller. In summary, plants with large RGA-elements around the crossover-frequency are fundamentally difficult to control, and inverse-based controllers should never be used for such plants.

2.3.5 RGA and Decentralized Integral Controllability (DIC)

Definition of DIC. A plant $G(s)$ (corresponding to a given pairing) is DIC if there exists a stabilizing decentralized controller with integral action such that each individual loop may be detuned independently by a factor ϵ_i $(0 \leq \epsilon_i \leq 1)$ without introducing instability. DIC is a property of the plant and the chosen pairings. Unstable plants are not DIC.

Theorem 3 *Assume $C(s)$ is diagonal and that $G(s)C(s)$ is stable and proper. Then $\lambda_{ii}(0) < 0$ for any $i \Rightarrow$ not DIC*

Proof: Follows from Theorem 6 [8].
This condition is tight for 2×2 systems since in this case [23] $DIC \Leftrightarrow \lambda_{11}(0) = \lambda_{22}(0) > 0$. For 3×3 systems with $\lambda_{ii} > 0$ the necessary and sufficient condition is $DIC \Leftrightarrow \sqrt{\lambda_{11}(0)} + \sqrt{\lambda_{22}(0)} + \sqrt{\lambda_{33}(0)} > 1$ [29].

2.3.6 RGA and Stability of Decentralized Control Systems

Apart from the relationships between RGA and DIC presented above, we have not found any strong relationships between the RGA and overall nominal stability (NS) of a decentralized control system. However the following theorem holds:

Theorem 4 *If $\Lambda(G) = I \quad \forall\omega$ then stability of the individual loops imply stability of the entire system.*

The proof is straight forward: We find from Eq. (2.3) that $\Lambda(G) = I$ can only arise from triangular G(s) (with diagonal G(s) as a special case) or from $G(s)$-matrices that can be made triangular by interchanging rows and columns in such a way that the diagonal elements remain the same but in a different order (the pairings remain the same). A plant with a triangularizable transfer matrix (as described above) controlled by a diagonal controller has only "one-way coupling" and will always be stable provided the individual loops are stable.

For plants that can not be made triangular by row and column interchanges Theorem 4 is of little use as it does not tell what deviations from $\Lambda(G) = I$ can be tolerated without impairing stability. Care should be taken to distinguish Theorem 4 from what may be termed the conventional pairing rule.

Conventional pairing "rule": *Prefer pairings ij with $\lambda_{ij}(j\omega)$ close to 1* (e.g. [5], [18], p. 457)

We emphasize that the conventional pairing rule is an engineering rule of thumb, and is not based on any proof. Indeed, pairing in accordance with the conventional pairing rule may result in unstable systems even if the individual loops are tuned to be stable (for systems of dimension larger than 2×2).

Example 2. Counterexample to the conventional pairing rule. Consider the plant

$$G(s) = \frac{(1-s)}{(1+5s)^2} \begin{pmatrix} 1 & -4.19 & -25.96 \\ 6.19 & 1 & -25.96 \\ 1 & 1 & 1 \end{pmatrix} \tag{2.13}$$

The corresponding RGA matrix is at all frequencies

$$\Lambda(G) = \begin{pmatrix} 1 & 5 & -5 \\ -5 & 1 & 5 \\ 5 & -5 & 1 \end{pmatrix} \tag{2.14}$$

If we use the pairing indicated by Eq. (2.13) and tune individual PI controllers according to the Ziegler-Nichols tuning rules we obtain controllers $c_i(s) = 4.46\frac{7.58s+1}{7.58s}$. However,

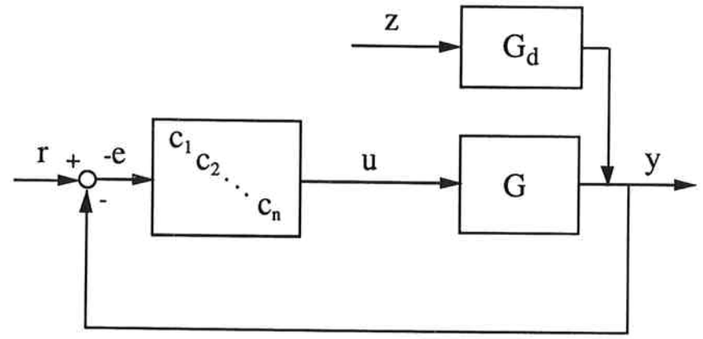


Figure 2.1: Block diagram of a decentralized control system.

the overall system becomes unstable even though the individual loops are stable. In order to obtain overall stability we have to detune the controller gains by a factor of 125.

2.4 Performance Relationships for Decentralized Control

In this section we consider the implications of overall performance requirements (nominal performance - NP) on the single-loop designs. We derive bounds on the designs of the individual loops

$$|g_{ii}c_i(j\omega)| > b_i(\omega); \quad \omega < \omega_B \tag{2.15}$$

which when satisfied yield performance (NP) of the overall system (with all loops closed). Note that the relationships for performance derived below require stability of the overall system (NS) as a prerequisite, that is, NS must be tested separately.

2.4.1 Notation

The controller $C(s)$ is diagonal with entries $c_i(s)$ (see Fig. 2.1). This implies that after the variable pairing has been determined, the order of the elements in y and u has been arranged so that the plant transfer matrix $G(s)$ has the elements corresponding to the paired variables on the main diagonal. Let $y(s)$ denote the output response for the overall system when all loops are closed and let $e(s) = y(s) - r(s)$ denote the output error. Then

$$e(s) = -S(s)r(s) + S(s)G_d(s)z(s); \quad S = (I + GC)^{-1} \tag{2.16}$$

Here z denotes the disturbances. G is assumed to be a $n \times n$ square matrix, but G_d may be nonsquare. The bandwidth of the system, ω_B, is defined as the frequency range

where the asymptotes of $\bar{\sigma}(S(j\omega))$ and $\underline{\sigma}(S(j\omega))$ cross one. This frequency range is also called called the "crossover region".

The matrix consisting of only the diagonal elements of $G(s)$ is denoted $\tilde{G}(s)$. $\tilde{y}(s)$ denotes the response of the individual subsystems, that is, $\tilde{y}_i(s)$ is the response when loop i is closed and the other loops are open. The closed-loop sensitivity functions for the individual loops may be collected in the diagonal matrix $\tilde{S}$:

$$\tilde{e}(s) = -\tilde{S}(s)r(s) + \tilde{S}(s)G_d(s)z(s) \tag{2.17}$$

$$\tilde{S} = (I + \tilde{G}C)^{-1} = diag\{\tilde{s}_{ii}\}; \quad \tilde{s}_{ii} = (1 + g_{ii}c_i)^{-1} \tag{2.18}$$

Note that the elements in $\tilde{S}$ are not equal to the diagonal elements of S.

2.4.2 Performance Requirements (Definition of NP)

Assume that G and G_d have been scaled such that at each frequency 1) the expected disturbances, $|z_k(j\omega)|$, are less than one, and 2) the outputs, y_i, are such that the expected setpoint changes, $|r_j(j\omega)|$ are less than one. As a NP performance specification we shall require for any setpoint change, r_j, that the offset e_i is bounded:

$$|e_i(j\omega)/r_j(j\omega)| = |S_{ij}(j\omega)| < 1/|w_{ri}(j\omega)|; \quad \forall\omega, \forall i, \forall j \tag{2.19}$$

Here $w_{ri}(s)$ is a scalar performance weight. For any disturbance z_k we require that

$$|e_i(j\omega)/z_k(j\omega)| = |[SG_d]_{ik}(j\omega)| < 1/|w_{di}(j\omega)|; \quad \forall\omega, \forall i, \forall k \tag{2.20}$$

Typically, both weights $|w_{di}(j\omega)|$ and $|w_{ri}(j\omega)|$ are large at low frequencies where small offset is desired. $|w_{ri}|$ is often about 0.5 at high frequencies to guarantee an amplification of high-frequency noise of 2 or less. Thus we have a number of performance specifications we want satisfied simultaneously.

2.4.3 Bounds on Single-loop Designs

In this section we shall use the above definition of performance to obtain bounds on the individual transfer functions $g_{ii}c_i$ at low frequencies. The Laplace variable s is omitted to simplify notation. For $\omega < \omega_B$ we may usually assume

$$S = (I + GC)^{-1} \approx (GC)^{-1} \tag{2.21}$$

We thus have:

$$e = -Sr + SG_dz \approx -C^{-1}G^{-1}r + C^{-1}G^{-1}G_dz \tag{2.22}$$

$$= -(\tilde{G}C)^{-1}\tilde{G}G^{-1}r + (\tilde{G}C)^{-1}\tilde{G}G^{-1}G_dz; \quad \omega < \omega_B \tag{2.23}$$

where $\Gamma = \tilde{G}G^{-1}$ is the PRGA matrix, and ΓG_d is known as the closed loop disturbance gain (CLDG) matrix [26]. The elements of Γ are denoted by γ_{ij} and those of ΓG_d are denoted by δ_{ij}. The step from (2.22) to (2.23) requires that the diagonal elements of G are nonzero. We have proven the following theorem:

Theorem 5 *For plants with nonzero diagonal elements in $G(s)$, and at frequencies $\omega < \omega_B$ where (2.21) holds, the NP specifications (2.19) and (2.20) are satisfied iff*

$$|g_{ii}c_i(j\omega)| > |\gamma_{ij}w_{ri}(j\omega)|; \quad \forall\omega < \omega_B, \forall i, \forall j \tag{2.24}$$

$$|g_{ii}c_i(j\omega)| > |\delta_{ik}w_{di}(j\omega)|; \quad \forall\omega < \omega_B, \forall i, \forall k \tag{2.25}$$

For a given choice of pairings Theorem 5 provides lower bounds on the individual loop gains to achieve NP. We get one bound on the loop gain $g_{ii}c_i$ for each setpoint j and each disturbance k. The bounds may be difficult to satisfy if γ_{ij} or δ_{ik} are large. A plot of $|\gamma_{ij}(j\omega)|$ as a function of frequency will give useful information about for which input-output pairs ij we may expect interactions. A plot of $|\delta_{ik}(j\omega)|$ will give useful information about which disturbances k are difficult to reject.

Comparison with all loops open. To get a better physical interpretation of the PRGA and CLDG consider the response e_i to a setpoint change r_i and a disturbance z_k when all the other loops are *open*. We get

$$e_i = -(1 + g_{ii}c_i)^{-1}r_i + (1 + g_{ii}c_i)^{-1}g_{dik}z_k \tag{2.26}$$

When all loops are *closed* simultaneously and we assume $\tilde{S} \approx (\tilde{G}C)^{-1}$ we get

$$e \approx -\tilde{S}\Gamma r + \tilde{S}\Gamma G_d z; \quad \omega < \omega_B \tag{2.27}$$

or

$$e_i \approx -(1 + g_{ii}c_i)^{-1}\gamma_{ij}r_j + (1 + g_{ii}c_i)^{-1}\delta_{ik}z_k; \quad \omega < \omega_B \tag{2.28}$$

Comparing (2.26) and (2.28) we see for a setpoint change r_i in loop i that the performance relative gain, γ_{ii}, gives the approximate change in offset caused by closing all the loops. In addition, γ_{ij} gives the effect of setpoint change r_j on output e_i when the other loops are closed. That is, for $\omega < \omega_B$ we have $s_{ij}/\tilde{s}_{ii} \approx \gamma_{ij}$, and we see that γ_{ii} is a measure of performance degradation at low and intermediate frequencies. Similarly, for loop i and disturbance z_k we see that the closed loop disturbance gain, δ_{ik}, gives the approximate gain from disturbance z_k to offset e_i when all the loops are closed, which explains why the name closed loop disturbance gain is chosen for ΓG_d.

2.4.4 Limitations of Theorem 5

1. The main limitation with the bounds in Theorem 5 is that they apply only to lower and intermediate frequencies.

2. Furthermore, they only address performance, and stability must be considered separately. For example, for input disturbances, i.e. $G_d = G$, we get the closed-loop disturbance gains $\delta_{ik}(s) = g_{ii}(s)$. Thus, it seems from performance considerations with respect to input disturbances that large diagonal elements in G (when appropriately scaled for disturbances) should be avoided. This is opposite of the conventional pairing rule of selecting inputs that have large effects on the controlled variables (i.e. $|g_{ii}(j\omega)|$ should be large, eg., [4], p. 48, [18], p. 683). The reason for the apparent discrepancy is stability issues (NS) and even more importantly input constraints which generally favor pairing on large elements.

3. Theorem 5 requires that the approximation $S \approx (GC)^{-1}$ holds for individual elements in S. It may appear that this approximation is poor for elements in S corresponding to elements in G^{-1} equal to zero. However, we show in Appendix 2 that if I^{ji} and G^{ji} have zero gain in the same direction, the approximation in (2.21) holds also for this element. Thus, in most cases Theorem 5 will hold (structurally) also for the zero elements in G^{-1} provided $g_{ii} \neq 0$. For example, it holds for all elements when G is triangular.

4. Another limitation with Theorem 5 is the assumption that $g_{ii} \neq 0, \forall i$. However, we may derive alternative bounds when $g_{ii} = 0$ as shown below for the 2×2 case.

2×2 *plant with diagonal element zero.* Without loss of generality assume $g_{11} = 0$. In this case the previously derived performance bounds (2.24) apply neither to loop 1 nor to loop 2. To derive appropriate bounds consider the elements s_{ij} of $S = (I + GC)^{-1}$ directly, and assume

$$|g_{22}c_2| \gg 1; \quad |c_1| \gg |\frac{g_{22}}{g_{12}g_{21}}| \tag{2.29}$$

such that $\det(I+GC) \approx \det(GC)$. The performance requirements for setpoint tracking to replace (2.24) then become

$$|c_1| > |\frac{g_{22}}{g_{12}g_{21}}w_{r1}|; \quad |c_1| > |\frac{1}{g_{21}}w_{r1}|; \quad |g_{22}c_2| > |\frac{g_{22}}{g_{12}}w_{r2}|; \quad |c_1c_2| > |\frac{1}{g_{12}g_{21}}w_{r2}| \tag{2.30}$$

The last bound puts a requirement on the product of the controller gains. This is reasonable since with $g_{11} = 0$ input u_1 can only affect output y_1 by the indirect action of control loop 2. For disturbance rejection, closed loop disturbance gains can be calculated for loop 2 as if g_{11} is non-zero. For loop 1, bounds on the controller gain c_1 for disturbance rejection are found from the elements in the first row of $[G^{-1}G_d]$.

2.4.5 Comparison with Previous Work

Mathematically, the performance specification (2.19) and (2.20) used above may be written

$$\|[W_r S \quad W_d S G_d](j\omega)\|_e < 1, \quad \forall\omega \tag{2.31}$$

where the e-norm used spatially (channels) is the largest modulus of the elements in the matrix. $W_r = diag\{w_{ri}\}$ and $W_d = diag\{w_{di}\}$ are diagonal matrices specifying the desired performance in in each output. This performance specification is very similar to the H_∞-norm, but in the latter case the induced 2-norm is used spatially. Consider the special case where $W_r = W_d = W_P$ and we have the H_∞-performance specification

$$\bar{\sigma}(W_P[S \quad SG_d](j\omega)) < 1, \forall\omega \Leftrightarrow \| W_p[S \quad SG_d] \|_\infty < 1 \tag{2.32}$$

Skogestad and Morari [25] have shown how one from the NP-condition (2.32) may derive the tightest possible bounds on the individual loops, for example, in terms of bounds on $|h_i|$, $|s_i|$ or $|g_{ii}c_i|$. These results are very powerful, but unfortunately the same bound is used for all loops, and this may be conservative. It is possible to derive

less conservative bounds by introducing additional adjustable parameters ("weights"), but it is not all obvious how this should be done *a priori* (see [16] for an example on how difficult it is even for a very simple case). However, using the spatial ∞-norm for the matrix as in (2.19), (2.20) and (2.31) makes it is much simpler to derive tight bounds on the individual loops.

In the paper we have shown that $\gamma_{ii} = \lambda_{ii}$ is a measure of performance degradation in terms of the diagonal elements in the sensitivity function, S. These results apply at small frequencies below crossover, but for control purposes the most important frequency region is close to crossover. Nett (eg. [13]) has presented results which relate $(\Lambda - I)$ and performance degradation in terms of $H = I - S$. These bounds are most useful at frequencies *beyond* crossover, but this frequency region by itself is not too interesting. However, our results complement each other and indicate that we should have $\Lambda \approx I$ at crossover in order to avoid degradation in performance when other loops are opened or closed. This provides a performance justification for the conventional pairing "rule". However, this justification only applies if we have as a design objective to maintain the same performance for the overall system as for the individual loops. However, we may want to sacrifice the latter in order to meet some other design objective, as demonstrated in example 4 below.

2.5 Examples

2.5.1 Example 2 (continued)

We return to example 1 in section 2.3.6 to illustrate that pairing according to $\lambda_{ii} = 1 \forall i$ may be undesirable from the point of view of performance. We consider the two alternative pairings corresponding to positive RGA values. For the pairing corresponding to $\lambda_{ii} = 1$ the transfer function matrix $G(s)$ is given by (2.13), whereas for the pairing corresponding to $\lambda_{ii} = 5$ the transfer function matrix is rearranged to give

$$G'(s) = \frac{(1-s)}{(1+5s)^2}\begin{pmatrix} -4.19 & -25.96 & 1 \\ 1 & -25.96 & 6.19 \\ 1 & 1 & 1 \end{pmatrix} \tag{2.33}$$

The control problem is formulated as follows: With a diagonal performance weight $Wp(s)$ with all diagonal elements equal

$$Wp(s)_{i,i} = 0.5\frac{\tau_{CL}s + 1}{\tau_{CL}s} \tag{2.34}$$

minimize τ_{CL} subject to an upper bound on the weighted sensitivity

$$\| WpS \|_\infty \leq 1 \tag{2.35}$$

This means that we allow a maximum peak in the sensitivity function at high frequencies of 2, and seek the controller which minimizes the closed loop time constant in the slowest direction. Only PI controllers were considered, and the tunings were obtained

	$\lambda_{ii}=5$	$\lambda_{ii}=1$
k_1	−0.6840	0.1230
τ_1	24.15	32.40
k_2	−0.02425	0.1443
τ_2	7.270	34.54
k_3	0.007685	0.002940
τ_3	0.3688	3.988
τ_{CL}	220	1160

Table 2.1: PI controller parameters for Example 1 ($c_i(s) = k_i\frac{\tau_i s+1}{\tau_i s}$).

by a numerical search. The results demonstrate that it is advantageous to choose the pairing corresponding to $\lambda_{ii}=5$ rather than $\lambda_{ii}=1$. For the pairing corresponding to $\lambda_{ii}=5$ we were able to fulfill (2.35) with $\tau_{CL}=220$ whereas for the pairing corresponding to $\lambda_{ii}=1$ we had to increase τ_{CL} to 1160 in order to be able to fulfill (2.35). Although the resulting closed loop systems are quite slow for both pairings (relative to the RHP zero at $s=1$) the pairing corresponding to $\lambda_{ii}=5$ is significantly better. The controller parameters giving the above values for τ are given in Table 2.1.

2.5.2 Example 3: Distillation Column Control

In order to demonstrate the use of the frequency dependent PRGA and CLDG for evaluation of expected control performance and control structure selection, a binary distillation column with 40 theoretical trays plus a total condenser is considered. This is the same example as studied in [24], but we use a more rigorous model which includes liquid dynamics in addition to the composition dynamics. Using model reduction, the number of states in the model was reduced from 82 to 5. Disturbances in feed flowrate F (z_1) and feed composition z_F (z_2), are included in the model. The LV configuration is used, that is, the manipulated inputs are reflux L (u_1) and boilup V (u_2). Outputs are the product compositions y_D (y_1) and x_B (y_2). The model then becomes

$$\begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} = G(s)\begin{pmatrix} du_1 \\ du_2 \end{pmatrix} + G_d(s)\begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix} \tag{2.36}$$

A state space description is given in Appendix 3. The disturbances and outputs have been scaled such that a magnitude of 1 corresponds to a change in F of 30%, a change in z_F of 20%, and a change in x_B and y_D of 0.01 molefraction units.

Pairings. We choose u_1 to control y_1 and u_2 to control y_2, as indicated by (2.36), in order to have positive steady state relative gains. This is in agreement with industrial practice.

Analysis of the model. Fig. 2.2 shows the open-loop disturbance gains, g_{dik}, as a function of frequency. These gains are quite similar in magnitude and rejecting disturbances z_1 and z_2 seems to be equally difficult. However, this conclusion is incorrect.

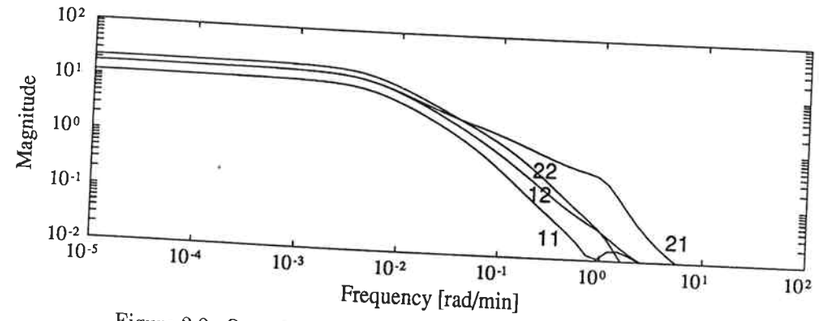
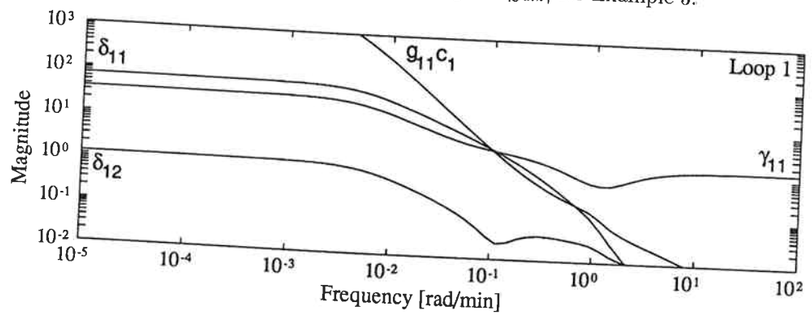


Figure 2.2: Open loop disturbance gains $|g_{dik}|$ for Example 3.



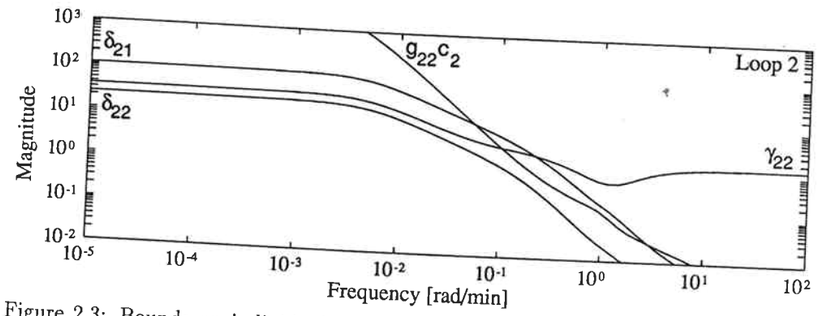


Figure 2.3: Bounds on individual loop gains for Example 3. NP at low frequencies ($\omega < \omega_B$) requires $|g_{ii}c_i|/|\gamma_{ij}| > |w_{ri}|$ and $|g_{ii}c_i|/|\delta_{ik}| > |w_{di}|$. Performance weights w_{ri} and w_{di} are not shown, but these are generally large at low frequencies and approach 1 at $\omega \approx \omega_B$.

The reason is that the *direction* of these two disturbances is quite different, that is, disturbance 2 is well aligned with G and is easy to reject, while disturbance 1 is not [21]. This is seen from Fig. 2.3 where the closed-loop disturbance gains, δ_{i2}, for z_2 are

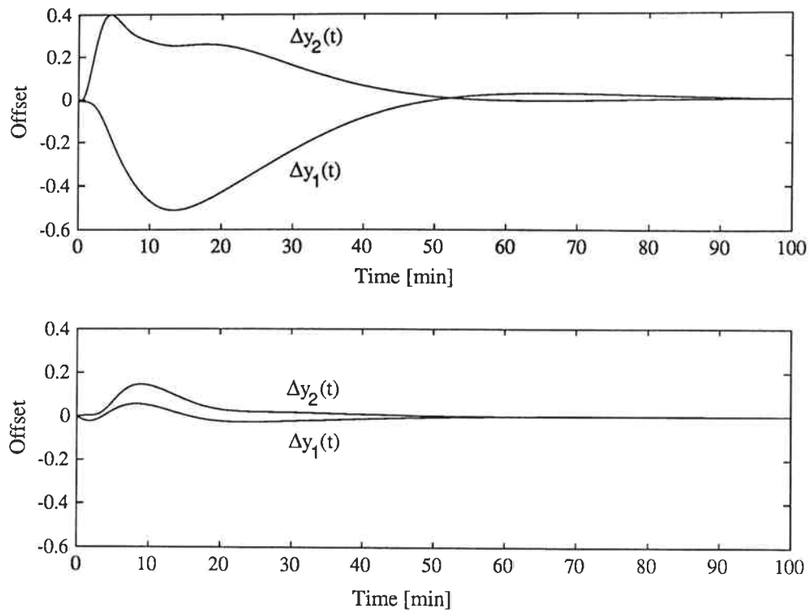


Figure 2.4: Responses for Example 3 to unit steps in z_1 (top) and z_2 (bottom).

seen to be much smaller than δ_{i1} for z_1. The performance relative gains for the loops are also included in Fig. 2.3 (note that $\gamma_{11} = \gamma_{22}$ for 2×2 plants). We see that rejection of disturbance 1 (as indicated by $|\delta_{i1}|$) and setpoint following (as indicated by $|\gamma_{ii}|$) put similar bounds on the loop gain $|g_{ii}c_i|$. Assuming that the performance requirement around crossover corresponds to performance weights $|w_d(j\omega_B)| \approx |w_r(j\omega_B)| \approx 1$ we find that the minimum bandwidth requirement for both loops is about 0.5 rad/min. That is, interactions become severe and performance will deteriorate drastically if the loops are detuned.

Observed control performance. To check the validity of the above results we designed single-loop PI controllers by optimizing robust performance using (2.32) as the performance specification, assuming up to one minute unmodelled deadtime in the inputs. The controllers obtained are:

$$c_1(s) = 0.261\frac{1+3.76s}{3.76s}; \quad c_2(s) = -0.375\frac{1+3.31s}{3.31s} \tag{2.37}$$

The loop gains, $|g_{ii}c_i|$, with these controllers are also shown in Fig. 2.3. The loop gains are seen to be larger than the closed-loop disturbance gains, $|\delta_{ik}|$, at all frequencies up to crossover. Closed-loop simulations with these controllers are shown in Fig. 2.4. The simulations confirm that disturbance 2 is much easier rejected than disturbance

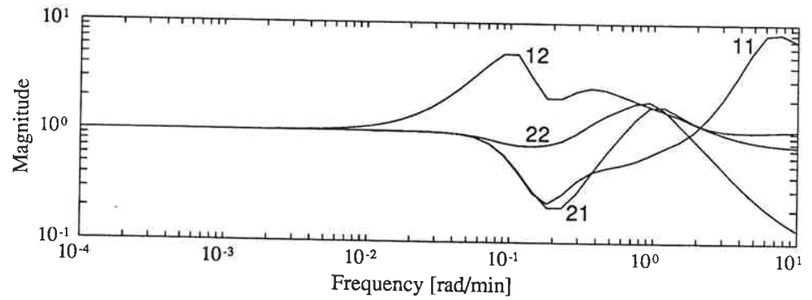


Figure 2.5: Check of approximation $SG_d \approx \tilde{S}\Gamma G_d$ for Example 3. The figure shows the magnitude of $[SG_d]_{ik}/[\tilde{S}\Gamma G_d]_{ik}$.

1. In summary, there is an excellent correlation between the analysis based on $|\delta_{ik}|$ in Fig. 2.3 and the simulations. This is not surprising when one considers Fig. 2.5 which shows the accuracy of the approximation $[S(s)G_d(s)]_{ik} \approx [\tilde{S}\Gamma G_d]_{ik}$ which formed the basis for the analysis in Fig. 2.3. The approximation is very good at low frequencies, but as expected poorer at frequencies around the closed loop bandwidth. The most significant deviation occurs for $i = 1, k = 2$ at frequencies around 0.1 rad/min, where we see that the actual disturbance rejection is poorer than the approximation. This explains why the effect of z_2 on y_1 is somewhat poorer than might be expected from Fig. 2.3.

2.5.3 Example 4: Pairing Corresponding to $\lambda_{ii} = 0$

In order to demonstrate that acceptable performance may be achieved even with pairings corresponding to $\lambda_{ii} = 0$, consider control of the top part of a distillation column. It is desired to control the top product composition (y_1) and the level in the condenser (y_2). The manipulated inputs are the distillate flowrate (u_1) and the reflux flowrate (u_2). The vapor flowrate entering the top of the column is considered to be the only disturbance (z_1). The achievable bandwidth is limited by unmodeled measurement delay in y_1 of one minute and valve dynamics in u_2 equivalent to a time delay of 0.1 minute. After scaling, the resulting transfer functions are

$$\begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} = \begin{pmatrix} 0 & \frac{100}{1+100s} \\ \frac{-1}{s} & \frac{-1}{s} \end{pmatrix} \begin{pmatrix} du_1 \\ du_2 \end{pmatrix} + \begin{pmatrix} \frac{-100}{1+100s} \\ \frac{1}{s} \end{pmatrix} dz_1 \tag{2.38}$$

This pairing corresponds to $\lambda_{11} = \lambda_{22} = 0$ at all frequencies. This pairing may be preferred in some cases, for example, if the reflux is large such that *constraints* on the distillate flowrate make level control with this input difficult. The chosen controllers are

$$c_1(s) = -0.5\frac{1+10s}{10s}; \quad c_2(s) = -5 \tag{2.39}$$

To check NP, the controllers and the bounds (2.30) for the case with zero diagonal

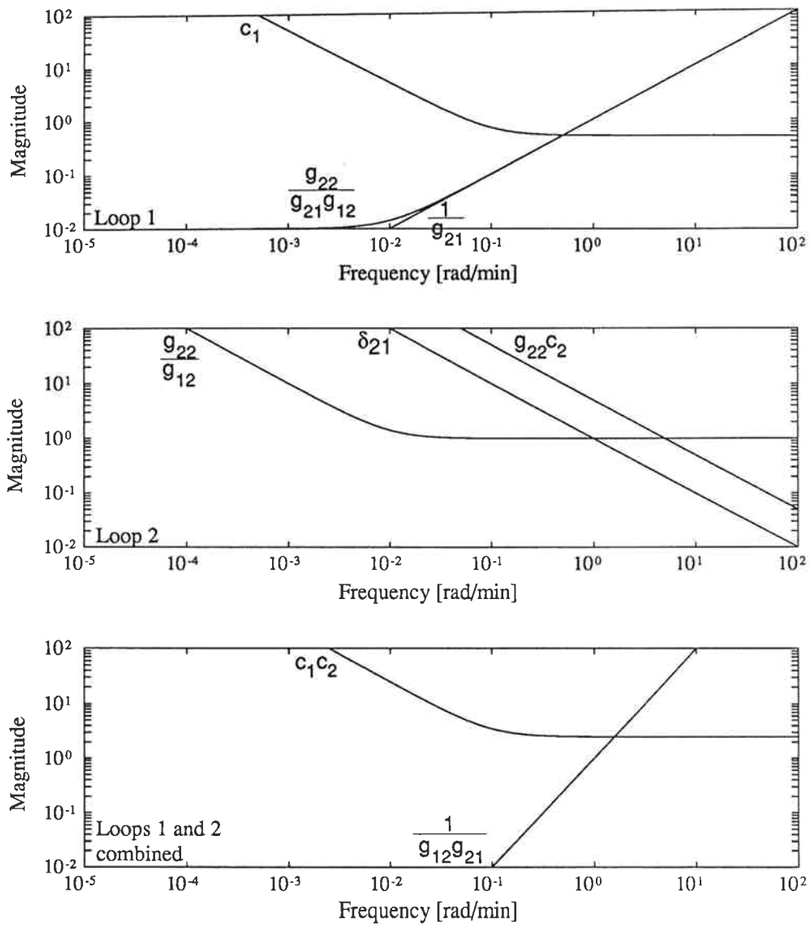


Figure 2.6: Bounds on controller and loop gains for Example 4.

elements are shown in Fig. 2.6. The bounds in Fig. 2.6 indicate that interactions put no serious limitations on achievable performance. In Fig. 2.7 we show responses to changes in setpoints r_1 and r_2 and in disturbance z_1. In the simulations a first order filter with a time constant of one minute is used for both setpoints, and a one minute time delay in the measurement of y_1 and a 0.1 minute time delay in manipulated variable u_2 are approximated with first order Padé approximations. The observed control performance is satisfactory, although there is an undesirable interaction from setpoint r_2 to output

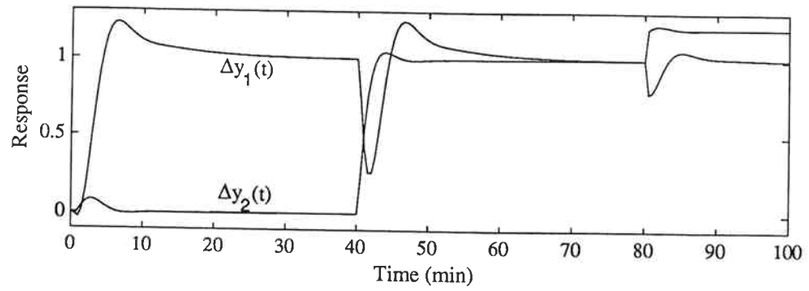


Figure 2.7: Responses for Example 4 to unit step changes in r_1 at $t = 0$, r_2 at $t = 40$ and in z_1 at $t = 80$.

y_1. This interaction cannot be predicted from Fig. 2.6, as Eq. (2.21) does not hold in the crossover region where the interactions occur.

2.6 Conclusions on Decentralized Control

In the paper we have derived bounds on the designs of the individual loops which when satisfied yield performance (NP) of the overall system (with all loops closed). For setpoint tracking the bounds are given by the performance relative gains, $|\gamma_{ii}|$ (Eq. 2.24), and for disturbance rejection by the closed-loop disturbance gains, $|\delta_{ik}|$ (Eq. 2.25). The bounds are tight (necessary and sufficient) at low frequencies where $S \approx (GC)^{-1}$. It is desirable that the bounds are as small as possible because a large bound requires a large bandwidth in loop i. Since stability of the individual loops is desired this may be impossible if $g_{ii}(s)$ contains time delays, neglected or uncertain dynamics, or rhp-zeros.

Importantly, these bounds depend on the model of the process only, that is, are independent of the controller. This means that frequency-dependent plots of γ_{ii} and δ_{ik} may be used to evaluate the achievable closed-loop performance (controllability) under decentralized control. Plants with small values of these measures are preferred. Furthermore, the values of δ_{ik} may tell the engineer which disturbance k will be most difficult to handle using feedback control. This may pinpoint the need for using feedforward control, or for modifying the process. For example, in process control adding a feed buffer tank will dampen the effect of disturbances in flowrate, temperature or composition. Plots of δ_{ik} may be used to tell if a tank is necessary and what holdup (residence time) would be needed.

The bounds may also be used to obtain a first guess of the controller parameters. However, as the derivation of the bounds depends on approximations which are valid at low frequencies only, undesirable effects may occur at frequencies around the closed loop bandwidth. Thus the behavior of the closed-loop system must be checked using other methods, and the controllers possibly redesigned.

Acknowledgements. Support from NTNF is gratefully acknowledged.

Nomenclature (see also Section 4.1)

D_1, D_2 - diagonal matrices
$e = y - r$ - vector of offsets
$g_{ij} = [G]_{ij}$ - ij'th element of G
$g_{dik} = [G_d]_{ik}$ - ik'th element of G_d
G - plant transfer matrix
$\tilde{G}$ - matrix consisting of diagonal elements of G
G^{ij} - G with row i and column j removed
r - vector of reference outputs (setpoints)
u - vector of manipulated inputs
w_{di} - performance weight for disturbance rejection in loop i.
w_{ri} - performance weight for setpoint following in loop i.
y - vector of outputs
$Y = \frac{g_{12}g_{21}}{g_{11}g_{22}}$ - Rijnsdorp or Balchen interaction measure for 2×2 system
z - vector of disturbances

Greek letters
$\delta_{ik} = g_{ii}[G^{-1}G_d]_{ik} = [\tilde{G}G^{-1}G_d]_{ik}$ - Closed Loop Disturbance Gain (CLDG)
$\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G)$ - condition number
$\gamma^*(G) = \min_{D_1 D_2} \gamma(D_1 G D_2)$ - optimal (minimized) condition number
$\gamma_{ij} = g_{ii}[G^{-1}]_{ij} = [\tilde{G}G^{-1}]_{ij}$ - Performance Relative Gain
Γ - Matrix of Performance Relative Gains (PRGA)
$\lambda_i(G)$ - i'th eigenvalue of matrix G
$\lambda_{ij}(G) = g_{ij}[G^{-1}]_{ji}$ - ij'th element in RGA-matrix Λ
Λ - RGA matrix
ω - frequency
ω_B - closed loop bandwidth

Norms
$\rho(A) = \max_i |\lambda_i(A)|$ - spectral radius
$\bar{\sigma}(A)$ - maximum singular value or spectral norm ($= \|A\|_{i2}$ - induced 2-norm)
$\underline{\sigma}(A) = 1/\bar{\sigma}(A^{-1})$ - minimum singular value
$\|A\|_1 = \Sigma_{i,j}|a_{ij}|$ - 1-norm
$\|A\|_2 = (\Sigma_{ij}|a_{ij}|^2)^{0.5}$ - 2-norm (Euclidean norm)
$\|A\|_e = \max_{ij} |a_{ij}|$ - e-norm (magnitude of largest element in A).
$\|G\|_\infty = \sup_\omega \bar{\sigma}(G)$-$H^\infty$ norm of $G(j\omega)$.
$\|A\|_{i1} = \max_j \Sigma_i |a_{ij}|$ - induced 1-norm ("largest column sum")
$\|A\|_{i\infty} = \max_i \Sigma_j |a_{ij}|$ - induced ∞-norm ("largest row sum")
$\|A\|_m = 2\max\{\|A\|_{i1}, \|A\|_{i\infty}\}$

Subscripts
i - index for outputs or loops
j - index for manipulated inputs or setpoints
k - index for disturbances

References

[1] Arkun, Y. (1987). Dynamic block relative gain and its connection with the performance and stability of decentralized structures. *Int. J. Control*, **46**, 4, 1187-1193.

[2] Arkun, Y. (1988). Relative Sensitivity: A Dynamic Closed-Loop Interaction Measure and Design Tool. *AIChE Journal*, **34**, 4, 672-675.

[3] Balchen, J. G. (1958). *Lecture notes*, Norwegian Institute of Technology, Trondheim, Norway (In Norwegian).

[4] Balchen, J. G. and Mumme, K. (1988). *Process Control. Structures and Applications*, Van Nostrand Reinhold, New York.

[5] Bristol, E. H. (1966). On a new measure of interactions for multivariable process control. *IEEE Trans. Automat. Control*, **AC-11**, 133-134.

[6] Bristol, E. H. (1978). Recent results on interactions in multivariable process control. Paper at AIChE Annual Meeting, Chicago, IL.

[7] Bristol, E. H. (1981). The right half plane'll get you if you don't watch out. *Joint American Control Conference (JACC)*, paper TA-7A.

[8] Grosdidier, P., Morari, M. and Holt, B. R. (1985). Closed-Loop Properties from Steady-State Gain Information. *Ind. Eng. Chem. Fundam.*, **24**, 221-235.

[9] Grosdidier, P. (1990) Analysis of Interaction with the Singular Value Decomposition, *Comp. chem. Engng.*, **14**, 6, pp. 687-689.

[10] Manousiouthakis, V., Savage, R. and Arkun, Y. (1986). Synthesis of Decentralized Process Control Structures Using the Concept of Block Relative Gain, *AIChE Journal*, **32**, 6, 991-1003.

[11] McAvoy, T. J. (1981). Connection between Relative Gain and Control Loop Stability and Design. *AIChE Journal*, **27**, 4, 613-619.

[12] McAvoy, T. J. (1983). *Interaction Analysis*, Instrument Society of America, Research Triangle Park, USA.

[13] Minto, K. D. and Nett, C. N. (1989). A Quantitative Approach to the Selection and Partitioning of Measurements and Manipulations for the Control of Complex Systems. Presentation at the American Control Conference, Pittsburgh, PA.

[14] Nett, C. N. (1987). Presentation at the American Control Conference, Minneapolis, MN.

[15] Nett, C. N. and Manousiouthakis, V. (1987). Euclidian Condition and Block Relative Gain: Connections, Conjectures and Clarifications. *IEEE Trans Aut. Control*, **AC-32**, pp. 405-407.

[16] Nett, C. N. and Uthgenannt, J. A. (1988). An explicit formula and optimal weight for the 2-block structured singular value interaction measure. *Automatica*, **24**, 2, 261-265.

[17] Rijnsdorp, J. E. (1965). Interaction in Two-variable Control Systems for Distillation Columns-I. *Automatica*, **1**, pp. 15-28.

[18] Seeborg, D. E., Edgar, T. F. and Mellichamp, D. A. (1989). *Process Dynamics and Control*, John Wiley & Sons, New York.

[19] Shinskey, F. G. (1967). *Process Control Systems*, McGraw-Hill, New York.

[20] Shinskey, F. G. (1984). *Distillation Control*, 2nd Edition, McGraw Hill, New York.

[21] Skogestad, S. and Morari, M. (1987). Effect of Disturbance Directions on Closed Loop Performance, *Ind. Eng. Chem. Res.*, **26**, 2029-2035.

[22] Skogestad, S. and Morari, M. (1987). Implications of Large RGA Elements on Control Performance, *Ind. Eng. Chem. Res.*, **26**, 11, 2323-2330.

[23] Skogestad, S. and Morari, M. (1988). Variable selection for decentralized control. Paper at *AIChE Annual Meeting*, Washington DC.

[24] Skogestad, S., Morari, M. and Doyle, J. C. (1988). Robust Control of Ill- Conditioned Plants: High Purity Distillation, *IEEE Trans. Autom. Control*, **33**, 12, 1092-1105.

[25] Skogestad, S. and Morari, M. (1989). Robust Performance of Decentralized Control Systems by Independent Designs, *Automatica*, **25**, 1, 119-125.

[26] Skogestad, S. and Hovd, M. (1990). Use of Frequency-dependent RGA for COntrol Structure Selection. *Proc. American Control Conference*, San Diego, CA.

[27] Tung, L. S. and Edgar, T. F. (1981). Analysis of Control-Output Interactions in Dynamic Systems, *AIChE Journal*, **27**, 4, 690-693.

[28] Witcher, M. F. and McAvoy, T. J. (1977). Interacting Control Systems: Steady State and Dynamic Measurement of Interaction, *ISA Transactions*, **16**, 3, 35-41.

[29] Yu, C.-C. and Fan, M. K. H. (1990). Decentralized Integral Controllability and D-Stability, *Chem. Eng. Sci.*, **45**, 11, pp.3299-3309.

[30] Yu, C.-C. and Luyben, W. L. (1987). Robustness with Respect to Integral Controllability, *Ind. Eng. Chem. Res.*, **26**, 1043-1045.

Appendix 1. Proof of Theorem 1. Consider Eq.(2.4) as a function of frequency, ie., let $s = j\omega$. Since $\lim_{s\to\infty} \lambda_{ij}(s)$ is finite and different from zero, Eq.2.4 may be written as a fraction of two polynomials in s where the numerator polynomial and denominator polynomial are of the same order. The phase change in $\lambda_{ij}(j\omega)$ as ω goes from 0 to ∞

must then be caused by RHP-poles or zeros in $g_{ij}(s)$, $det(G^{ij}(s))$ or $det(G(s))$ and the theorem follows.

Appendix 2. Consider a non-singular plant transfer function matrix G, and assume that neither G nor the diagonal controller C has a pole on the imaginary axis at the frequency in question. We have $[G^{-1}]_{ij} = (-1)^{i+j}\det(G^{ji})/\det(G)$ where G^{ji} denotes the matrix G with row j and column i removed. Correspondingly, $S_{ji} = (-1)^{i+j}\det(I+GC)^{ji}/\det(I+GC)$

Proposition: ($[G^{-1}]_{ij} = 0$ and (I^{ji} and G^{ji} have zero gain in the same input direction)) $\Rightarrow S_{ij} = [(I+GC)^{-1}]_{ij} = 0$.

Proof: If I^{ji} and G^{ji} have zero gain in the same input direction, $(GC)^{ji}$ will have zero gain in the same direction, as C is diagonal. Thus, $\det(I^{ji}+(GC)^{ji}) = \det(I+GC)^{ji} = 0 \Rightarrow S_{ij} = 0$.

Appendix 3. Transfer function matrices for Example 2. The transfer function matrices $G(s)$ and $G_d(s)$ can be calculated from the formulae
$G(s) = C(sI-A)^{-1}B + D$ and $G_d(s) = C(sI-A)^{-1}B_d + D_d$.

$$A = \begin{pmatrix} -5.161e-3 & 0 & 0 & 0 & 0 \\ 0 & -7.366e-2 & 0 & 0 & 0 \\ 0 & 0 & -1.829e-1 & 0 & 0 \\ 0 & 0 & 0 & -4.620e-1 & 9.895e-1 \\ 0 & 0 & 0 & -9.895e-1 & -4.620e-1 \end{pmatrix} \tag{2.40}$$

$$B = \begin{pmatrix} -6.296e-2 & 6.236e-2 \\ 5.481e-3 & -1.719e-2 \\ 3.041e-3 & -1.078e-2 \\ -1.856e-2 & -1.393e-2 \\ -1.229e-1 & -5.608e-3 \end{pmatrix} \tag{2.41}$$

$$C = \begin{pmatrix} -7.223 & -5.170 & 3.836 & -1.633e-1 & 1.121 \\ -8.913 & 4.728 & 9.876 & 8.425 & 2.186 \end{pmatrix} \tag{2.42}$$

$$D = D_d = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.43}$$

$$B_d = \begin{pmatrix} -9.364e-3 & -1.333e-2 \\ 1.960e-2 & 8.018e-3 \\ 3.266e-3 & -2.116e-2 \\ -2.827e-2 & 5.319e-3 \\ -6.784e-3 & 2.719e-3 \end{pmatrix} \tag{2.44}$$

Chapter 3

Controllability Analysis for Unstable Processes

Morten Hovd Sigurd Skogestad*
Chemical Engineering, University of Trondheim-NTH
N-7034 Trondheim, Norway

Abstract

We study the controllability of unstable processes, with emphasis on selection of measurements and manipulated variables, and the pairing problem for decentralized control. The well known pairing criteria based on the Niederlinski Index and the steady state Relative Gain Array (RGA) have been generalized to open loop unstable plants. Right Half Plane (RHP) zeros in individual transfer function elements may make it practically impossible to stabilize the individual loops, and in such cases these pairing criteria are not very helpful. We have found RGA $\approx I$ in the bandwidth region to indicate good pairings also when paired elements have significant RHP zeros, but have found the Direct Nyquist Array (DNA) to perform poorly as an indicator of good pairings. In some cases it is preferable to avoid pairings giving narrow Gershgorin bands.

*Author to whom correspondence should be addressed. E-mail: SKOGE@KJEMI.UNIT.NO, phone: 47-7-594154, fax: 47-7-594080

3.1 Introduction

In engineering practice, a system is called controllable if it is possible to achieve the specified aims of the control, whatever these may be ([19], p. 171). Unfortunately, in standard state-space control theory the term "controllability" has a rather limited definition in terms of Kalman's *state* controllability, which mainly has to do with realization theory. State controllability will only be considered briefly in this paper, and it will be clear from the context whenever the term controllability refers to a state and not to a plant.

For an unstable plant we must use feedback for stabilization. Thus, whereas the presence of RHP-zeros put an upper bound on the the allowed bandwidth, the presence of RHP-poles put a lower bound. It is also clear that it may be difficult to stabilize an unstable plant if there are RHP-zeros or time delays - "the system goes unstable before we are able to observe what is happening". This qualitative statement is quantified by the results on sensitivity relationships below.

In most of the paper we assume that a decentralized controller is used, as such controllers are very common in the chemical process industry. A significant amount of work has been done on the choice of pairings for decentralized control of stable plants, e.g. [1], [4], [13], [18] and [20], to reference a few. The concept of Decentralized Integral Controllability (DIC) [20], is not relevant for unstable plants, as the system must necessarily become unstable when the controller gains are sufficiently reduced. In the paper we show how stability based pairing criteria fail or must be interpreted differently when the number of RHP poles in $G(s)$ and $\tilde{G}(s)$ differ ($\tilde{G}(s)$ consists of the diagonal elements of $G(s)$). In contrast, we demonstrate the success of the frequency dependent RGA for unstable plants.

Notation. $G(s)$ is the $n \times n$ transfer function of the process, with the ij'th element denoted $g_{ij}(s)$. $\tilde{G}(s)$ is the transfer function matrix consisting of the diagonal elements of $G(s)$. The controller is denoted $C(s)$, with individual elements $c_i(s)$ for the case of decentralized control. In the following, the Laplace variable s will be dropped where it is not needed for clarity. The sensitivity function is given by $S = (I + GC)^{-1}$ and the complementary sensitivity function by $H = I - S = GC(I + GC)^{-1}$. Similarly, for decentralized control the sensitivity functions and complementary sensitivity functions of the individual loops can be collected in diagonal matrices, giving $\tilde{S} = (I + \tilde{G}C)^{-1}$ and $\tilde{H} = I - \tilde{S} = \tilde{G}C(I + \tilde{G}C)^{-1}$. The following relationships are also used

$$(I + GC) = (I + E_H\tilde{H})(I + \tilde{G}C) \tag{3.1}$$

$$E_H = (G - \tilde{G})\tilde{G}^{-1} \tag{3.2}$$

3.2 Controllability Measures

In this section we review some proposed controllability measures, and generalize some of the controllability measures to unstable plants. In the paper we make use of the multivariable Nyquist theorem, and we will therefore state it here

Theorem 6 *Let the map of the Nyquist D contour under* $\det(I + G(s)C(s))$ *encircle the origin* n_C *times in the clockwise direction. Let the number of open loop unstable poles of* $G(s)C(s)$ *be* n_U*. Then the closed loop system is stable if and only if* $n_C = -n_U$*.*

Proof: The theorem has been proved several times, see [12].

3.2.1 State Controllability and Observability

It is well known (e.g. [11]) that only states that are both observable and controllable can be stabilized by feedback control. It is therefore necessary that the selection of manipulated and measured variables result in all unstable states being controllable and observable.

We believe that in most cases an engineer with a good understanding of the process will intuitively choose measurements and manipulated variables such that this requirement will be fulfilled. Nevertheless, it does make sense to check that all unstable states are controllable and observable. Preferably, the controllability and observability of the unstable states should not rely on one single manipulated variable or measurement, as the system will then necessarily become unstable if this manipulated variable/measurement fails.

3.2.2 Sensitivity Relationships and RHP-zeros

Assume a plant $G(s)$ with a RHP-pole at $s = p$ is stabilized by feedback control such that $S(s)$ and $H(s)$ are stable. Then $S(s)$ must have a RHP zero at $s = p$ (follows since $H(s) = G(s)C(s)S(s)$ is stable while $G(s)$ has a RHP pole at $s = p$). This shows that the presence of an RHP-pole imposes restrictions on the closed-loop system in addition to the requirement of stabilization. The approach of first stabilizing the plant with some simple controller, and then proceed as if the RHP-pole never existed, as is proposed by some authors (e.g., [15]), is therefore flawed.

To quantify the effect of the RHP-poles on the closed-loop system we shall consider the sensitivity integral relationships of Freudenberg and Looze [2, 3] which extend the Bode integral relationship to plants with RHP-poles and RHP-zeros. Let us first consider Single Input Single Output (SISO) systems, and suppose that $G(s)C(s)$ is rational and has at least two more poles than zeros. Let $G(s)C(s)$ have n_U poles (including multiplicities) in the open right half plane, at locations $p_i, \quad i = 1, 2, \ldots, n_U$. Then, if the closed loop system is stable, the sensitivity function must satisfy:

No RHP zero:

$$\int_0^\infty \ln |S(j\omega)| d\omega = \pi \sum_{i=1}^{n_U} Re[p_i] \tag{3.3}$$

One real RHP zero at $s = z$:

$$\int_0^\infty \ln |S(j\omega)| W(z,\omega) d\omega = \pi \sum_{i=1}^{n_U} \ln \left| \frac{\bar{p}_i + z}{p_i - z} \right| \tag{3.4}$$

$$W(z,\omega) = \frac{2z}{z^2 + \omega^2}$$

$\bar{p}_i$ denotes the complex conjugate of p_i. Eq. (3.3) and Eq. (3.4) show that we need $|S(j\omega)| > 1$ over some range of frequencies, which means that the effect of disturbances is actually amplified at these frequencies. A RHP transmission zero of $G(s)$ limits the achievable bandwidth of the plant regardless of the type of controller used (e.g, [16]). This is confirmed by (3.4) where the shape of the weight $W(z,\omega)$ (equals $2/z$ at low frequencies and falls off with a -2 slope from $\omega = z$) implies that essentially all the positive area for $\ln |S(j\omega)|$ has to be at frequencies lower than the RHP-zero, and there will have to be a peak $|S(j\omega)| > 1$ which will become increasingly large as the crossover frequency approaches z.

Since with no RHP poles the integrals in (3.3) and (3.4) equal zero, wee see that the presence of RHP poles increase the area for which $|S| > 1$. We also see that the peak of $|S|$ will approach infinity if $p \to z$. In practice, this means that $|p_i|$ must be smaller than $|z|$ in order to stabilize the plant. In contrast, in the absence of RHP zeros, RHP poles do not impose any practical performance limitations (in terms of peaks in $|S|$), as the frequency range where $|S| > 1$ as required by Eq. (3.3) may be arbitrarily large and is only limited by high frequency roll-off considerations.

For Multiple Input Multiple Output (MIMO) systems the situation is not so clear, although some useful insight exists. For a $n \times n$ system with no RHP-zeros we get

$$\sum_{i=1}^{n} \int_0^\infty \ln \sigma_i[S(j\omega)] d\omega = \pi \sum_{i=1}^{n_U} Re[p_i] \tag{3.5}$$

The difference compared to SISO systems is that the relationship involves the sum of the log magnitudes of the singular values, suggesting that it may be possible to trade off sensitivity properties in different directions. For MIMO systems with a RHP transmission zero at z, we have

$$\int_0^\infty \ln \bar{\sigma}[S(j\omega)] W(z,\omega) d\omega \geq 0 \tag{3.6}$$

Note that this result depends on the presence of RHP-zeros only, and does not show the combined effects of RHP poles and zeros similar to (3.4) for SISO systems. The reason results for MIMO systems are so much harder to find is the issue of plant directionality, as illustrated by the following simple examples

$$G_1(s) = f(s) \begin{bmatrix} 1 & 0 \\ 0 & \frac{s-z}{s-p} \end{bmatrix}$$

$$G_2(s) = f(s) \begin{bmatrix} s-z & 0 \\ 0 & \frac{1}{s-p} \end{bmatrix}$$

where $f(s)$ is an arbitrary stable minimum phase rational transfer function with at least three more poles than zeros. G_1 and G_2 both have one RHP transmission zero and one RHP pole at the same frequencies. However, in G_1 the RHP pole and transmission zero lie in the same direction and may cause a serious performance limitation if p approaches z (recall (3.4)), whereas in G_2 the RHP pole and transmission zero lie in directions at right angles to each other and only the RHP zero itself causes a limitation.

Jacobsen and Skogestad [7] have studied the combined effect of RHP zeros (time delays) and RHP poles for distillation columns.

Implications of RHP zeros on the selection of controlled and manipulated variables. One should attempt to choose inputs and outputs such that RHP transmission zeros are avoided. If an RHP transmission zero cannot be avoided, it should preferably be at as high a frequency as possible, and lie in a plant direction such that it interferes as little as possible with the control of any RHP poles.

Similar considerations apply for decentralized control when a paired *element* g_{ii} has a RHP zero that is not a transmission zero of the plant. In many cases such RHP zeros will disappear when the other loops are closed, and there is then no fundamental bandwidth restriction in channel i. However, we know that the zeros of $1 + g_{ii}c_i$ will approach the zeros of g_{ii} for high gain feedback. A choice will therefore have to be made between individual loop stability and system performance. This dilemma can only be avoided if the selection of controlled and manipulated variables makes it possible to choose a pairing for which none of the paired elements have RHP zeros within the desired bandwidth.

3.2.3 Decentralized Fixed Modes

Wang and Davison [22] showed that it may be impossible to move some modes by decentralized feedback, even if all states are both controllable from the inputs and observable from the outputs. Wang and Davison termed such modes "decentralized fixed modes". Lunze [11] gives a good explanation of a necessary and sufficient condition for the existence of decentralized fixed modes. Consider a system described by

$$\dot{x} = Ax + Bu; \quad y = Cx + Du$$

The simplest way to prove that an eigenvalue of A does not correspond to a decentralized fixed mode is to try with an arbitrary constant feedback matrix K with the structure of the pairing in consideration. A mode which is fixed for any constant decentralized feedback is also fixed for dynamic decentralized feedback [22]. When selecting input and outputs for decentralized control, one must clearly ensure that there exists at least one pairing for which all decentralized fixed modes are stable. This will usually not be difficult, as it suffices to ensure that the unstable state is both controllable and observable in one individual channel [11].

Example 1: Consider the plant

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -10 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -8 \end{bmatrix} x + \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x \end{aligned} \tag{3.7}$$

which yields

$$G(s) = \begin{bmatrix} \frac{1}{s+10} & \frac{2(s+4)}{(s+10)(s2)} \\ 0 & \frac{1}{s+8} \end{bmatrix} \tag{3.8}$$

We first try the pairing $y_1 - u_1$, $y_2 - u_2$, with a constant feedback matrix $K = \text{diag}\{k_1, k_2\}$ and computing the closed loop autotransition matrix

$$A + BKC = \begin{bmatrix} -10 + k_1 & k_1 & k_2 \\ 0 & 2 & k_2 \\ 0 & 0 & -8 + k_2 \end{bmatrix} \tag{3.9}$$

The eigenvalue at 2 is unaffected by feedback, and it is therefore impossible to stabilize the system with this pairing. This is not surprising, by simple inspection of Eq. (3.8) it is clear that decentralized feedback with the pairing $y_1 - u_1$, $y_2 - u_2$ cannot give feedback around the unstable state corresponding to the pole at $s = 2$. With the opposite pairing, $y_1 - u_2$, $y_2 - u_1$, we have no decentralized fixed modes, and the system can be stabilized by decentralized feedback.

There are cases where the existence of decentralized fixed modes and the choice of pairing for decentralized control is not as obvious as for Example 1.

Example 2: Consider the plant

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -8 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} x + \begin{bmatrix} -10 & 3 \\ 1 & 2 \\ 0 & 6 \\ 0 & -4 \end{bmatrix} u \\ y &= \begin{bmatrix} 2 & 10 & 3 & 0 \\ 3 & 12 & 0 & -4 \end{bmatrix} x \end{aligned} \tag{3.10}$$

corresponding to the transfer function matrix

$$G(s) = \begin{bmatrix} \frac{-10s+40}{(s+2)(s+8)} & \frac{44s^2+300s-56}{(s+2)(s+8)(s-2)} \\ \frac{-18s+36}{(s+2)(s+8)} & \frac{49s^2+238s-584}{(s+2)(s+8)(s-4)} \end{bmatrix} \tag{3.11}$$

It can be shown that with the pairing indicated by Eq. (3.11) ($y_1 - u_1$, $y_2 - u_2$), the mode corresponding to the pole at $s = 2$ is fixed, whereas for the opposite pairing the mode corresponding to the pole at $s = 4$ is fixed. This process is thus impossible to stabilize by decentralized feedback, even though all states are both controllable and observable.

3.2.4 The Niederlinski Index

The Niederlinski Index [18] is defined as

$$N_I = \frac{\det G(0)}{\det \tilde{G}(0)} \tag{3.12}$$

For *stable* plants it has been shown that if all the individual loops are stable and have integral action, a necessary condition for the stability of the overall system is that $N_I > 0$ [4]. This result also holds for unstable plants if we assume that the number of RHP poles in G and $\tilde{G}$ are the same, but this assumption generally does not hold. However, we present a generalized Niederlinski index criterion for unstable plants:

Theorem 7 *Assume:*
1. The transfer function GC is strictly proper.
2. The controller C is diagonal, has integral action in all channels and is otherwise stable.
3. The number of open loop unstable poles in G is n_U.
4. The map of the Nyquist D contour under $det(I+\tilde{G}C) = det\tilde{S}^{-1}$ encircles the origin $\tilde{n}_C$ times in the clockwise direction.
Then, a necessary condition for the overall system to be stable is

$$\text{sign}\{N_I\} = \text{sign}\{(-1)^{-n_U-\tilde{n}_C}\} \tag{3.13}$$

Remark: The individual loops (i.e. $\tilde{S}$) may or may not be stable. If we require the individual loops to be stable, then $\tilde{n}_C = -\tilde{n}_U$, where $\tilde{n}_U$ is the number of open loop unstable poles in $\tilde{G}$.

Proof: Let the map of the Nyquist D contour under $\det(I + E_H\tilde{H})$ encircle the origin n_E times in the clockwise direction. Thus, from Eq. (3.1) we have $n_C = \tilde{n}_C + n_E$ and we get from Thm. 1 that the overall system is stable if and only if $n_E = -n_U - \tilde{n}_C$. $E_H\tilde{H}$ is strictly proper, as GC is strictly proper. $\tilde{H}(0) = I$ because of the requirement for integral action in all channels, regardless of whether $\tilde{H}(s)$ is stable. We therefore have (see Corollary 1.1 in [5])

$$\lim_{s\to\infty}(I + E_H(s)\tilde{H}(s)) = I \tag{3.14}$$

$$\lim_{s\to 0}(I + E_H(s)\tilde{H}(s)) = \lim_{s\to 0} G(s)\tilde{G}^{-1}(s) \tag{3.15}$$

Then the map of $\det(I + E_H\tilde{H})$ starts at N_I (for $s = 0$) and ends at 1 (for $s = \infty$). For stability this map must have $n_E = -n_U - \tilde{n}_C$ encirclements of the origin, and we must require N_I to be positive if n_E is even and N_I to be negative if n_E is odd.

3.2.5 The Relative Gain Array

The Relative Gain Array (RGA) was first introduced by Bristol [1]. It is defined at any frequency as

$$\Lambda = G \times [G^{-1}]^T \tag{3.16}$$

where the sign $\times$ denotes element by element multiplication (Hadamard product). The ij'th element of Λ is denoted[1] λ_{ij}.

[1]The ij'th relative gain λ_{ij} should not be confused with λ_i which denotes the i'th eigenvalue.

To simplify notation, we will in the following consider loop 1, without loss of generality (the generalization to loop k is trivial). Introduce $G' = \text{diag}\{g_{11}, G^{11}\}$, where G^{11} is obtained from G by removing row 1 and column 1. Let G' have n'_U RHP poles. Note that n'_U may be different for different loops.

Theorem 8 *Under assumptions 1-3 of Thm. 7, a necessary condition for simultaneously obtaining*
a) Stability of the closed loop system
b) Stability of loop 1 by itself
c) Stability of the system with loop 1 removed
is that

$$\text{sign}\{\lambda_{11}(0)\} = \text{sign}\{(-1)^{-n_U+n'_U}\} \tag{3.17}$$

Proof: Follows by substituting G' for $\tilde{G}$ in the proof of Thm. 7, and noting that $\det GG'^{-1} = 1/\lambda_{11}$. This theorem generalizes the widely used RGA pairing criterion ([1], [4]) to unstable plants.

For 2×2 plants $N_I = 1/\lambda_{11}(0)$ but for larger systems these measures contain different information.

3.2.6 The Direct Nyquist Array

The Direct Nyquist Array (DNA) [19] is simply an array of polar plots of the elements $g_{ij}(s)$ of the plant transfer function matrix $G(s)$. The usefulness of the DNA technique stems from Gershgorins theorem, which states that the eigenvalues λ_i of a matrix $G(s)$ must lie within the union of circles $|\lambda_i - g_{ii}(s)| < r_i(s)$ where $r_i(s) = \sum_{i=1,i\neq j}^{n} |g_{ij}(s)|$. The circles $r_i(s)$ are superimposed on the plots of the diagonal elements of G. As the frequency changes, the circles will move (and their radius will change), thus forming n bands of circles, known as *Gershgorin bands*. Jensen et al. [8] suggest to use the DNA also to obtain pairings for decentralized control, preferring pairings giving Gershgorin bands which are narrow compared to the magnitude of the diagonal elements.

The following derivation shows why a plot of the Gershgorin bands of $G(S)$ is useful for closed loop stability: Consider first the eigenvalues of $(I + GC)$, which are located within the Gershgorin bands of $(I+GC)$. Since $\det(I+GC) = \prod_i \lambda_i(I+GC)$ it follows from Thm. 1 that *if the number of encirclements of the Gershgorin bands of $(I + GC)$ equals the number of unstable poles, n_U, and none of these bands include the origin, the closed loop system is stable.* For decentralized control, the centers of the circles which make up the Gershgorin bands of $(I + GC)$ are given by $(I + \tilde{G}C)$. When G and $\tilde{G}$ have different numbers of RHP poles and the individual loops are stable, the numbers of encirclement of the origin of the Nyquist D contours under $\det(I + GC)$ and $\det(I + \tilde{G}C)$ must differ. Since the centers of the Gershgorin bands of $I + GC$ are given by $(I + \tilde{G}C)$, some of the Gershgorin bands of $(I + GC)$ *must include* the origin for the overall system to be stable. Furthermore, it will then be undesirable that the Gershgorin bands of G are narrow, as this will make the stability margins for the individual loops and the overall system small.

To derive the relationship between the Gershgorin bands of $(I + GC)$ and G normalize the elements in G with the diagonal elements, and get

$$G = \begin{bmatrix} 1 & \frac{g_{12}}{g_{22}} & \frac{g_{13}}{g_{33}} & \cdots \\ \frac{g_{21}}{g_{11}} & 1 & \frac{g_{23}}{g_{33}} & \cdots \\ \vdots & & \ddots & \vdots \end{bmatrix} \text{diag}\{g_{ii}\} \tag{3.18}$$

This should be compared to

$$(I + GC) = (I + E_H\tilde{H})(I + \tilde{G}C) = \tag{3.19}$$

$$\begin{bmatrix} 1 & \tilde{h}_2\frac{g_{12}}{g_{22}} & \tilde{h}_3\frac{g_{13}}{g_{33}} & \cdots \\ \tilde{h}_1\frac{g_{21}}{g_{11}} & 1 & \tilde{h}_3\frac{g_{23}}{g_{33}} & \cdots \\ \vdots & & \ddots & \vdots \end{bmatrix} \text{diag}\{1 + g_{ii}c_i\}$$

We see that the width of the Gershgorin bands (relative to the magnitude of the diagonal elements) will be the same for G and $(I + GC)$ at frequencies below the bandwidth (where $\tilde{h}_i \approx 1$). At frequencies beyond the bandwidth (where $\tilde{h}_i < 1$) $(I + GC)$ will have narrower Gershgorin bands than G. Only in the bandwidth region, where peaks in $\tilde{h}_i$ may occur, can the Gershgorin bands of $(I + GC)$ be wider than the Gershgorin bands of G. Thus, the widths of the Gershgorin bands of G (relative to the magnitude of the diagonal elements) can be used as estimates of the widths of the Gershgorin bands of $(I + GC)$. In this paper, we use the diagonal similarity transform of Mees [14] in order to reduce the conservativeness associated with the location of the eigenvalues of G. The diagonal elements of this transformation matrix are the elements of the left eigenvalue corresponding to the Perron root of $G\tilde{G}^{-1}$.

3.2.7 The SSV Interaction Measure

The Structured Singular Value Interaction Measure (SSV-IM)[5] is the structured singular value (μ) of E_H with respect to a stable perturbation matrix with the same structure as $\tilde{H}$. The map of $\det(I + E_H\tilde{H})$ cannot encircle the origin if $\rho(E_H\tilde{H}) < 1 \quad \forall\omega$, which is satisfied if $\bar{\sigma}(\tilde{H})\mu(E_H) < 1 \quad \forall\omega$. The last relationship follows since the least conservative way to "split up" $\rho(E_H\tilde{H})$ is to use the structured singular value. Since these relationships are useful only when we require that $\det(I + E_H\tilde{H})$ should *not* encircle the origin (i.e. $-n_U - \tilde{n}_C = 0$), we see that they generally apply only when the individual loops are stable and G and $\tilde{G}$ have the same number of RHP poles.

3.2.8 The Performance RGA

The Performance Relative Gain Array (PRGA) has recently been introduced [21, 6]. It is derived from simple manipulations with the transfer function from setpoints r to offset $e = y - r = -Sr$: At low frequencies ($\omega < \omega_B$) where $\tilde{H} \approx I$, we have

$$(I + E_H\tilde{H}) \approx I + E_H = G\tilde{G}^{-1} = \Gamma^{-1}; \quad \omega < \omega_B \tag{3.20}$$

and we derive from Eq. (3.1) that

$$e \approx -\tilde{S}\Gamma r; \qquad \omega < \omega_B \tag{3.21}$$

where the PRGA matrix is

$$\Gamma = \tilde{G}G^{-1} \tag{3.22}$$

Despite their different use this definition points to the similarity between the PRGA and the Inverse Nyquist Array. The PRGA is dependent on scaling of the outputs, and must be recomputed for new pairings. Note, however, that the diagonal elements of the PRGA matrix are equal to the diagonal elements of the RGA matrix, and are hence independent of scaling. If G is scaled such that the maximum acceptable magnitude of the individual offsets e_i is unity, the PRGA matrix gives approximate bandwidth requirements and loop gain requirements at frequencies below the bandwidth, as the loop gain in loop i should be larger in magnitude than the magnitude of any element in row i of the PRGA matrix, and small PRGA elements are therefore preferred. The use and the limitations of the PRGA is discussed more thoroughly in [6].

3.2.9 High Frequency RGA and Stability

The encirclements of the origin of $\det(I+GC)$ caused by the RHP poles of G, will occur in the frequency region corresponding to the RHP poles of G. For practical systems the bandwidth region will lie at frequencies significantly higher than the RHP poles of G, and we therefore do not want any additional encirclements in the bandwidth region. The poles of $\tilde{G}$ lie at the same locations as the poles of G, and we therefore want $\det(I + \tilde{G}C)$ also to avoid encirclements of the origin in the bandwidth region. We thus want $\det(I + GC)$ and $\det(I + \tilde{G}C)$ to behave similarly in the bandwidth region. If G is triangular then $\det(I + E_H\tilde{H}) = 1$ and $\det(I + GC) = \det(I + \tilde{G}C)$, and in this case we also have that the RGA matrix $\Lambda = I$ and the PRGA matrix Γ is triangular with diagonal elements equal to 1. We therefore prefer pairings which give $\Lambda \approx I$ in the bandwidth region. This agrees with the results of Nett and coworkers ([17]). In summary, we want the PRGA elements to be small at low frequencies, and in the bandwidth region we desire the PRGA close to triangular with diagonal elements close to 1.

3.3 Examples

3.3.1 Example 3

Consider the system $G(s) = C(sI - A)^{-1}B + D$, with

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}; \quad B = \begin{bmatrix} 5 & -8 \\ 4 & 10 \\ 2 & -8 \end{bmatrix}$$
$$C = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \quad D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.23}$$

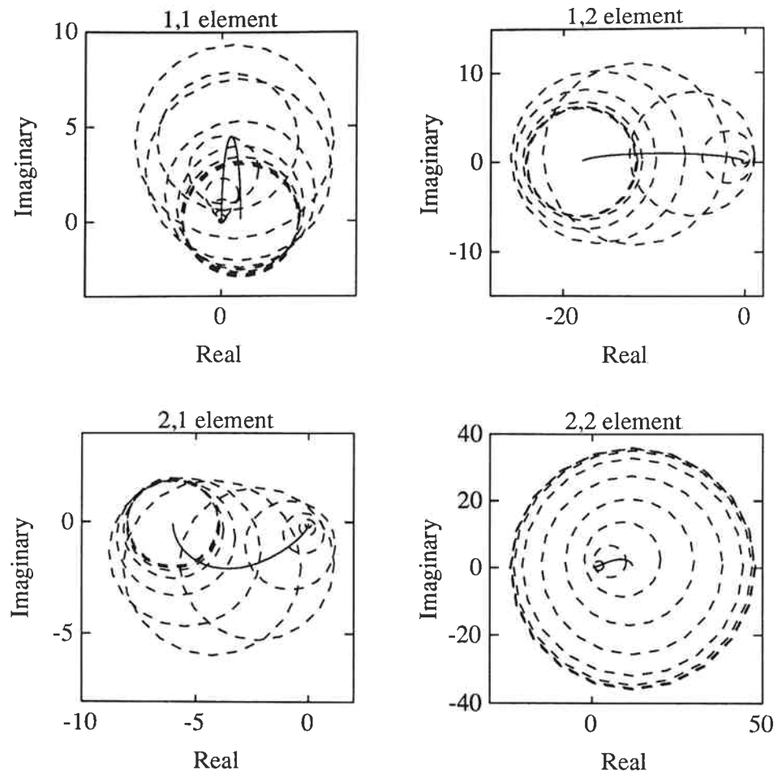


Figure 3.1: DNA with Gershgorin bands for Example 3.

The transfer function matrix G has one unstable pole at frequency 1 [rad/min], and we must therefore expect the closed loop bandwidth to be at least 1 [rad/min]. Note that $\tilde{G}$ has two unstable poles for both pairings, because the RHP pole appears in all four elements of G. Neither $G(s)$ nor any of its elements have RHP zeros. $G(0)$ is given by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -18 \\ -6 & 12 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{3.24}$$

The pairing ($y_1 - u_1$, $y_2 - u_2$) indicated by Eq. (3.24) we term "pairing 1", and the opposite pairing we term "pairing 2". The Niederlinski Index is -8 for pairing 1 and 0.89 for pairing 2. Thm. 7 tells us that a negative Niederlinski Index is necessary if we require both loops in addition to the overall system being stable and we therefore have to choose pairing 1. If we use the steady state RGA in accordance with Thm. 8, we would arrive at the same conclusion.

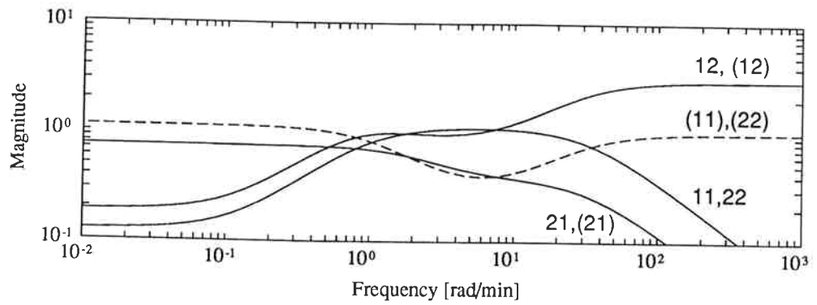


Figure 3.2: PRGA for Example 3. Solid lines denote pairing 1, and dashed lines and labels in parenthesis denote pairing 2

The DNA of pairing 1 and pairing 2 are shown in Fig. 3.1, with Gershgorin bands superimposed on the $1,1$ and $2,2$ elements for pairing 1, and on elements $1,2$ and $2,1$ for pairing 2. We see that all the Gershgorin bands include the origin. However, pairing 1 gives much wider Gershgorin bands relative to the magnitude of the diagonal elements than pairing 2. We explained above that when G and $\tilde{G}$ have different number of RHP poles and stability of the individual loops is desired, narrow Gershgorin bands are undesirable. The DNA thus gives a weak indication that pairing 1 is preferable to pairing 2.

The PRGA for pairing 1 and pairing 2 are shown in Fig. 3.2[2]. With a closed loop bandwidth in the region $1 - 30$ [rad/min] (approximately) the PRGA indicate that pairing 1 is preferable (The PRGA is close to triangular and the RGA$\approx I$). For pairing 1, using the controllers $c_1(s) = -\frac{s+1}{s}$, $c_2(s) = -\frac{(s+1)(0.1s+1)}{s(0.01s+1)}$ we find that both loops and the overall system is stable, and the predictions using the Niederlinski Index and the PRGA are shown to hold. This example also demonstrates that for unstable systems there is no reason to avoid pairings corresponding to negative steady state RGA values.

3.3.2 Example 4: Polypropylene Reactor

This example illustrates the problems encountered when we are not able to stabilize the individual loops. We consider the polypropylene reactor control example studied by Lie [9, 10]. A schematic outline of the process is shown in Fig. 3.3. The monomer feed enters into a stirred tank reactor containing a slurry of monomer, catalyst, cocatalyst, polymer and some impurities. The reaction is exothermic, causing some of the slurry components to vaporize. The vapor leaving the reactor is transferred to an accumulator vessel. Heat is removed from the system by condensing parts of the vapor leaving the reactor, before it enters the accumulator. Heat removal is adjusted by adjusting a split range valve which determines what fraction of the vapor leaving the reactor is passed

[2]We assume the outputs in Eq. (3.23) to be properly scaled. For the PRGA plot for pairing 2, $G(s)$ has been rearranged to bring the paired elements on the diagonal.

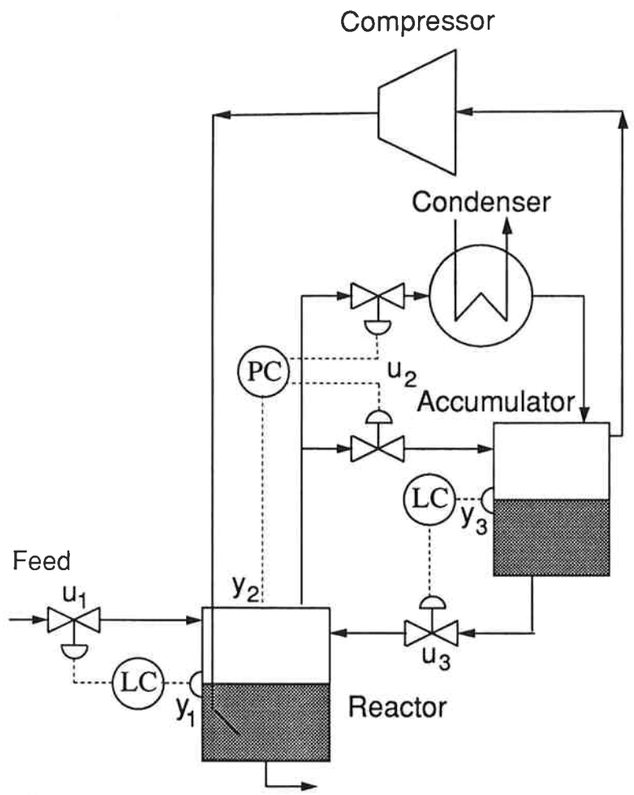


Figure 3.3: Schematic outline of the process in Example 4, with pairing 1 selected for control.

through the condenser. The liquid in the accumulator is returned to the reactor, and the vapor from the accumulator is compressed and bubbled through the reactor slurry. This results in a 3×3 plant model $G(s)$ with seven states as given in [9, 10]. The inputs and outputs are

y_1 - reactor slurry level $(0-1)$
y_2 - reactor pressure (gauge pressure in atmospheres)
y_3 - accumulator liquid level $(0-1)$.
u_1 - monomer feed flowrate (kg/h).
u_2 - split range valve position $(0-1)$.
u_3 - accumulator to reactor liquid flowrate (kg/h).

We scale the outputs such that a magnitude of 1 for the scaled outputs correspond to

deviations: $y_1 = 0.05$, $y_2 = 1.0$atm and $y_3 = 0.10$. The plant G has a pair of RHP poles at $s = 0.685 \pm 0.688j$, one pure integrator (the accumulator level) and no RHP transmission zeros. The reactor holdup becomes unstable because of the overhead condensation loop. All elements of G except g_{11} have RHP zeros at frequencies close to the RHP poles (e.g. g_{12} has a RHP zero at 2.16, g_{22} have a pair of complex RHP zeros at $0.10 \pm 0.90j$).

As we are considering the use of low-order controllers (typically PI or PID controllers), we cannot expect loops with a RHP zero at frequencies around or lower than the frequency of the RHP pole to be stable. This makes it difficult to apply the Niederlinski Index (Thm. 7) or the steady state RGA (Thm. 8) for this example. Although six different pairings is possible for a 3×3 system, we will here only consider the two pairings that pair y_1 with u_1.[3] There is a strong incentive for pairing y_1 and u_1; as g_{11} has no RHP zero we are able to stabilize the system with only this loop closed. Readers are referred to [9] for a more thorough discussion of all possible pairings. Thus, for this example, the pairing $y_1 - u_1$, $y_2 - u_2$, $y_3 - u_3$ is termed pairing 1, and the pairing $y_1 - u_1$, $y_2 - u_3$, $y_3 - u_2$ is termed pairing 2. Pairing 1 is shown in Fig. 3.3. The RGA plot is not shown here, but the diagonal elements equal the diagonal elements of the PRGA shown in Fig. 3.4. The figure clearly shows that pairing 1 is preferable, as the PRGA values for the paired elements are closer to one in the bandwidth region for pairing 1 than for pairing 2[4]. The PRGA also shows that for both pairings, the interaction leads to increased loop gain requirements at low frequencies for loops 2 and 3.

For this problem, we can only expect loop 1 to be stable by itself, the map of the Nyquist D contour under $1 + g_{11}c_1$ will therefore encircle the origin twice in the anticlockwise direction. For pairing 1, g_{22} and g_{33} both have two RHP poles and two RHP zeros within the desired bandwidth. The two RHP zeros will most likely result in two unstable poles of $1 + g_{22}c_2$ and $1 + g_{33}c_3$. $1 + g_{22}c_2$ and $1 + g_{33}c_3$ must therefore be expected to add no net encirclements of the origin, and narrow Gershgorin bands would therefore be a *sufficient* criterion for the viability of pairing 1. For pairing 2, g_{23} and g_{32} both have two RHP poles and one RHP zero within the desired bandwidth. $1 + g_{23}c_2$ and $1 + g_{32}c_3$ must therefore be expected to each add one anticlockwise encirclement of the origin. Wide Gershgorin bands are thus *necessary* for the viability of pairing 2. From Fig. 3.5 we see that the Gershgorin bands are very wide for both pairings[5]. Thus, the DNA plot does not tell that pairing 1 is a good pairing, nor does it give any reason for discarding pairing 2.

Based on the plots of the PRGA (or the frequency dependent RGA), we choose pairing 1. This conclusion is consistent with the simulations of Lie [10] and with industrial practice.

[3] For this example the steady state RGA is positive and close to I for another pairing, $y_1 - u_2$, $y_2 - u_3$, $y_3 - u_1$. This pairing defies common sense. For example, it means controlling the accumulator level with the monomer feed flowrate to the reactor. However, the desired bandwidth is about 10rad/h, and in this frequency range the RGA indicates that pairing 1 is preferable.

[4] For pairing 2 $G(s)$ has been rearranged to bring the paired elements for pairing 2 on the diagonal.

[5] For the Gershgorin band for the 1, 1 element the similarity transform for pairing 1 is used.

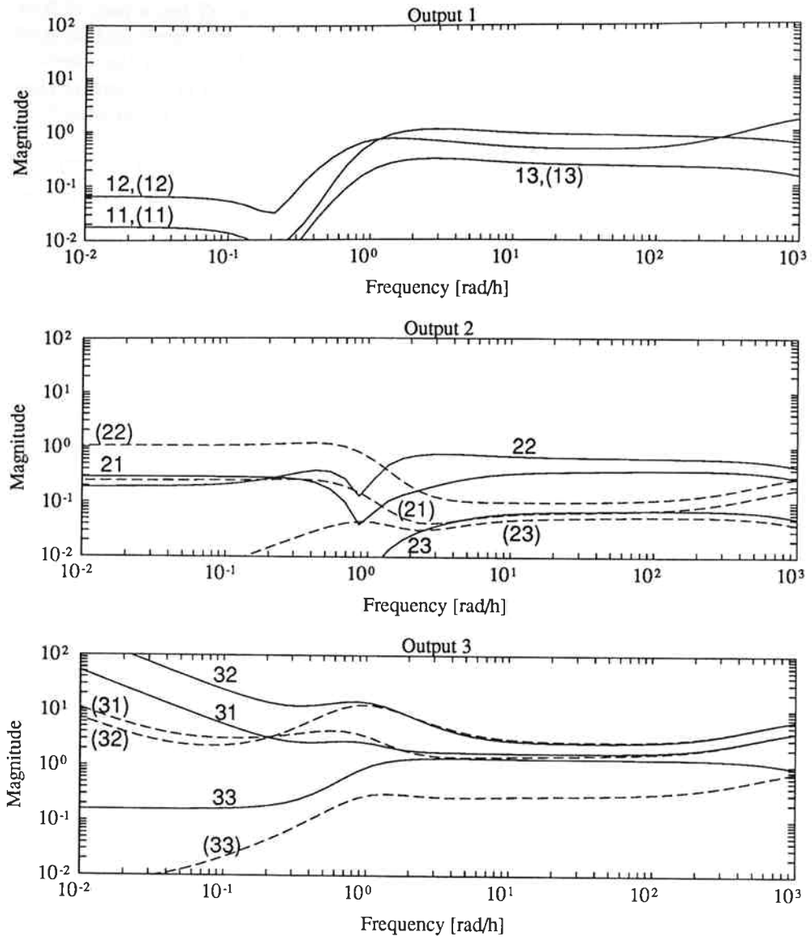


Figure 3.4: PRGA for Example 4. Solid lines denote pairing 1, and dashed lines and labels in parenthesis denote pairing 2.

Nomenclature

A - Matrix in the state space description of a transfer function matrix.
B - Matrix in the state space description of a transfer function matrix.
C - Controller transfer function matrix; matrix in the state space description of a transfer function matrix.
D - - Matrix in the state space description of a transfer function matrix.

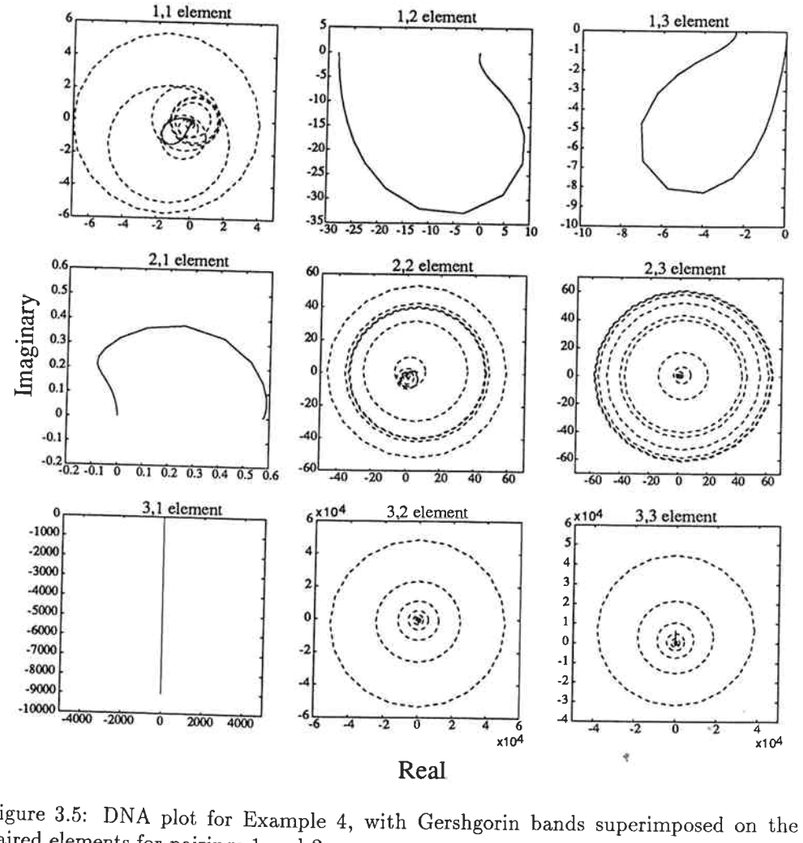


Figure 3.5: DNA plot for Example 4, with Gershgorin bands superimposed on the paired elements for pairings 1 and 2.

c_i - ii'th element of controller C (for diagonal C).
E_H - $(G - \tilde{G})\tilde{G}^{-1}$.
G - Plant transfer function matrix.
G^{11} - G with row 1 and column 1 deleted.
$\tilde{G}$ - Matrix consisting of the diagonal elements of G.
G' - diag$\{g_{11}, G^{11}\}$.
g_{ij} - ij'th element of G.
H - complementary sensitivity function $GC(I + GC)^{-1}$.
$\tilde{H}$ - matrix of complementary sensitivity functions for the individual loops $\tilde{G}C(I + \tilde{G}C)^{-1}$.
h_i - complementary sensitivity function for loop i, ii'th element of $\tilde{H}$.
K - diagonal, constant feedback controller matrix.

N_I - Niederlinski Index, $\det G(0)/\det \tilde{G}(0)$.
n - plant dimension ($n \times n$).
n_C - number of encirclements of the origin by the map of the Nyquist D contour under $\det(I + GC)$.
n_E - number of encirclements of the origin by the map of the Nyquist D contour under $\det(I + E_H \tilde{H})$.
$\tilde{n}_C$ - number of encirclements of the origin by the map of the Nyquist D contour under $\det(I + \tilde{G}C)$.
n_U - number of open loop unstable poles of GC.
$\tilde{n}_U$ - number of open loop unstable poles of $\tilde{G}C$.
n'_U - number of open loop unstable poles of $G'C$.
$\tilde{n}_U$ - number of open loop unstable poles of $\tilde{G}C$.
p_i - Location of pole in the open right half plane.
S - sensitivity function $(I + GC)^{-1}$.
$\tilde{S}$ - matrix of sensitivity functions for the individual loops $(I + \tilde{G}C)^{-1}$.
s - Laplace variable.
s_i - complementary sensitivity function for loop i, ii'th element of $\tilde{S}$.
u - vector of manipulated variables.
W - weight in the integral in Eq. (3.4) and Eq. (3.6).
x - vector of states.
y - vector of outputs.
z - Location of zero in the open right half plane.

Greek symbols:

Γ - Performance Relative Gain Array matrix $\tilde{G}G^{-1}$. Λ - Relative Gain Array matrix $G \times [G^{-1}]^T$.
λ_{ij} - ij'th element of Λ.
λ_i - i'th eigenvalue of a matrix.
μ - structured singular value.
ρ - spectral radius (magnitude of largest eigenvalue) of a matrix.
σ - singular value of a matrix.
$\bar{\sigma}$ - largest singular value of a matrix.
ω - frequency.
ω_B - bandwidth frequency.

References

[1] Bristol, E. H., (1966). On a New Measure of Interaction for Multivariable Process Control, *IEEE Trans. Autom. Control*, **AC-11**, pp. 133-134.

[2] Freudenberg, J. S. and Looze, D. P. (1985). Right Half Plane Poles and Zeroes and Design Tradeoffs in Feedback Systems. *IEEE Trans. Autom. Control*, **AC-30**, 6, pp. 555-565.

[3] Freudenberg, J. S. and Looze, D. P. (1988). *Frequency Domain Properties of Scalar and Multivariable Feedback Systems.* Lecture Notes in Control and Information Sciences 104, Springer Verlag, Berlin, Germany.

[4] Grosdidier, P., Morari, M. and Holt, B. R. (1985). Closed-Loop Properties from Steady-State Gain Information. *Ind. Eng. Chem. Fundam.*, **24**, pp. 221-235.

[5] Grosdidier, P. and Morari, M. (1986). Interaction Measures for Systems Under Decentralized Control. *Automatica*, **22**, 3, pp. 309-319.

[6] Hovd, M. and Skogestad, S. (1992). Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. To appear in *Automatica.*

[7] Jacobsen, E. W. and Skogestad, S. (1991). Control of Unstable Distillation Columns. *Proc. American Control Conference*, Boston, Mass., pp. 773-778.

[8] Jensen, N., Fisher, D. G. and Shah, S. L. (1986). Interaction Analysis in Multivariable Control Systems. *AIChE Journal*, **32**, 6, pp. 959-970.

[9] Lie, B. (1990). Control Structures for Polymerization Processes Applied to Polypropylene Manufacturing. *Dr. Ing. Thesis*, University of Trondheim-NTH.

[10] Lie, B. and Balchen, J. G. (1992). A Comparison of Strategies for Control of a Polypropylene Reactor, *Preprints IFAC Symposium DYCORD+'92*, College Park, Maryland, Apr. 1992, 265-270.

[11] Lunze, J. (1992). *Feedback Control of Large-Scale Systems*, Prentice-Hall, New York.

[12] Maciejowski, J. M. (1989). *Multivariable Feedback Design.* Addison-Wesley, Wokingham, England.

[13] McAvoy, T. J., (1983). *Interaction Analysis*, ISA Monograph, Research Triangle Park, North Carolina.

[14] Mees, A. J. (1981). Achieving Diagonal Dominance. *Systems & Control Letters*, **1**, 3, pp. 155-158.

[15] Morari, M., Grimm, W., Oglesby, M. J. and Prosser, I. D. (1985). Design of Resilient Processing Plants VII. Design of Energy Management System for Unstable Reactors - New Insights, *Chem. Eng. Sci.*, **40**, pp. 187-198.

[16] Morari, M. and Zafiriou, E., (1989). *Robust Process Control*, Prentice-Hall, Englewood Cliffs, New Jersey.

[17] Minto, K.D. and Nett, C. N. (1989). A Quantitative Approach to the Selection and Partitioning of Measurements and Manipulators for the Control of Complex Systems. Presentation at *1989 American Control Conference*, Minneapolis.

[18] Niederlinski, A. (1971). A Heuristic Approach to the Design of Linear Multivariable Interacting Control Systems. *Automatica*, **7**, pp. 691-701.

[19] Rosenbrock, H. H. (1970). *State-Space and Multivariable Theory.* Nelson, London, England.

[20] Skogestad, S. and Morari, M. (1988). Variable Selection for Decentralized Control. *AIChE Annual Meeting*, Washington DC, paper 128c. Also to appear in *Modelling, Identification and Control* (1992).

[21] Skogestad, S., and Hovd, M. (1990). Use of Frequency-Dependent RGA for Control Structure Selection. *Proc. American Control Conference*, San Diego, California, pp. 2133-2139.

[22] Wang, S.-H. and Davison, E. J. (1973). On the Stabilization of Decentralized Control Systems. *IEEE Trans. Autom. Control*, **AC-18**, 5, pp. 473-478.

Chapter 4

A Procedure for Regulatory Control Structure Selection with Application to the Fluid Catalytic Cracking Process

M. Hovd and S. Skogestad*

Chemical Engineering
University of Trondheim, NTH
N-7034 Trondheim, Norway

Abstract

The paper outlines a methodology for control structure selection for the regulatory control system. Control structure selection consists of the selection and pairing of manipulated and measured variables, and involves use of tools such as the existence of right half plane (RHP) transmission zeros, the relative gain array (RGA), the performance relative gain array (PRGA), and the closed loop disturbance gain (CLDG). The objectives of the regulatory control level and its interaction with the higher levels in the control hierarchy is discussed in detail. The regulatory control system for the Fluid Catalytic Cracking process is used as an example. Several authors have found the Kurihara control structure to be preferable to the conventional control structure. The reason for this is shown to be that RHP transmission zeros limit the achievable bandwidth for the conventional control structure. However, it is shown that two other control structures have better controllability characteristics than both the conventional and the Kurihara control structures. The sensitivity of the measurement selection and variable pairing with respect to changes in the operating point and parametric uncertainty is also studied.

*Author to whom correspondence should be addressed. E-mail: SKOGE@KJEMI.UNIT.NO, phone: 47-7-594154, fax: 47-7-591410.

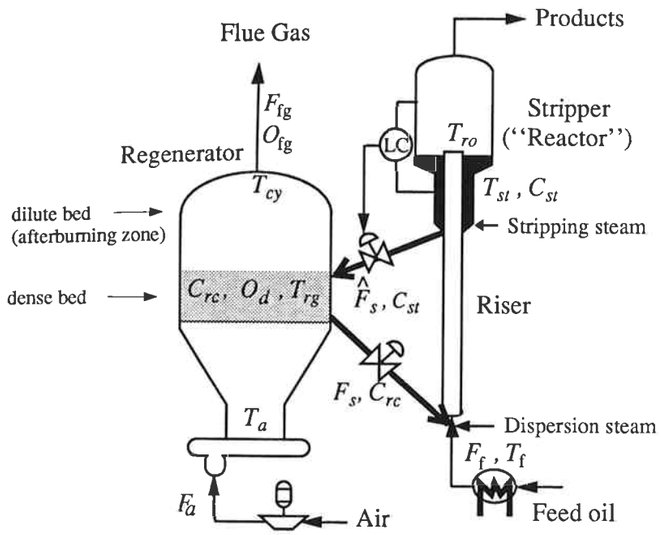


Figure 4.1: Schematic overview of an FCC plant.

4.1 Introduction

In the chemical industries, the lowest level in the control system is virtually always a *regulatory control* level, which keeps a set of measurements at setpoint which are determined by higher levels in the control hierarchy. The higher levels in the control hierarchy depend on a regulatory control system that performs well. The performance of the regulatory control system can be strongly affected by the control structure used, and control structure selection is therefore an important issue in the design of the regulatory control system. In a recent book Rijnsdorp [21] spends a full chapter discussing control structure selection. However, his approach is rather qualitative. In this paper we apply a more quantitative method for control structure selection to the control of of the reactor-regenerator complex in a Fluidized Catalytic Cracking (FCC) unit.

The FCC process is an important process in refineries for upgrading heavy hydrocarbons to more valuable lighter products. Both decentralized controllers and more complex model predictive controllers are used to control the FCC process. However, when model predictive control (MPC) is used, it is usually applied on top of a highly decentralized regulatory control level and sends setpoint changes to the individual loops. Thus, it is important also in this case that the regulatory control level is well designed.

A schematic overview of the FCC process is shown in Fig. 4.1. Feed oil is contacted with hot catalyst at the bottom of the riser, causing the feed to vaporize. The cracking reactions occur while the oil vapor and catalyst flow up the riser. As a byproduct

of the cracking reactions coke is formed and is deposited on the catalyst, thereby reducing catalyst activity. The catalyst and products are separated in the stripper, which for historical reasons is still often called the reactor [1]. Steam is supplied to the stripper in order to remove volatile hydrocarbons from the catalyst. The catalyst is then returned to the regenerator where the coke is burnt off in contact with air. The combustion of coke in the regenerator provides the heat needed for feed vaporization and the endothermic reaction in the riser.

The issue of regulatory control structure selection has been discussed by several authors, e.g. Kurihara [12], Lee and Weekman [15] and Lee and Groves [14]. In this paper we provide a more quantitative analysis. Specifically, the existence of right half plane (RHP) zeroes and the frequency dependent relative gain array (RGA), performance relative gain array (PRGA), and closed loop disturbance gain (CLDG) are used for control structure selection, and we study the effect of structural and parametric uncertainty in the models and uncertainty in the manipulated variables on the choice of control structure for decentralized control. The use of frequency dependent RGA, PRGA and CLDG are explained in [25, 10].

4.2 The Regulatory Control Problem

The *overall control objective* is to maintain safe operation while keeping the operating conditions close to the economically optimal conditions. This objective is commonly achieved using a hierarchical control system, with different tasks assigned to each level in the hierarchy. In this paper we consider what is typically the lowest level in this control hierarchy, the regulatory control level. The objective of this level is generally to facilitate smooth operation and not to optimize objectives related to profit, which is done at higher levels. Usually, this is a decentralized control system which keeps a set of measurements at given setpoints. This is a cascaded control system where the value of these setpoints are determined by the operator or by higher levels in the control hierarchy. Note that also the regulatory control system itself may include cascaded loops. For example, one often cascades the valve position to a flow measurement such that flow becomes the manipulated input rather than the valve position. In the following, the terms "regulatory control system" and "lower-level control system" will be used as synonyms.

At the intermediate level (supervisory control level) there may be a model based system that uses a multivariable process model to calculate how the plant should be operated to optimize some objective. An important feature of the supervisory control level may be to take into account constraints in the controlled and manipulated variables, by coordinating the use of the manipulated variables, that is, use additional degrees of freedom to avoid reaching constraints for the manipulated variables used by the lower-level control system. The top level in the control hierarchy is a plant wide optimization. This optimization is usually steady state and is performed offline

[1] With less active catalysts, the residence time in the reactor was needed to perform the cracking. With modern catalysts, the catalyst and products must be separated as quickly as possible at the riser outlet to prevent overcracking.

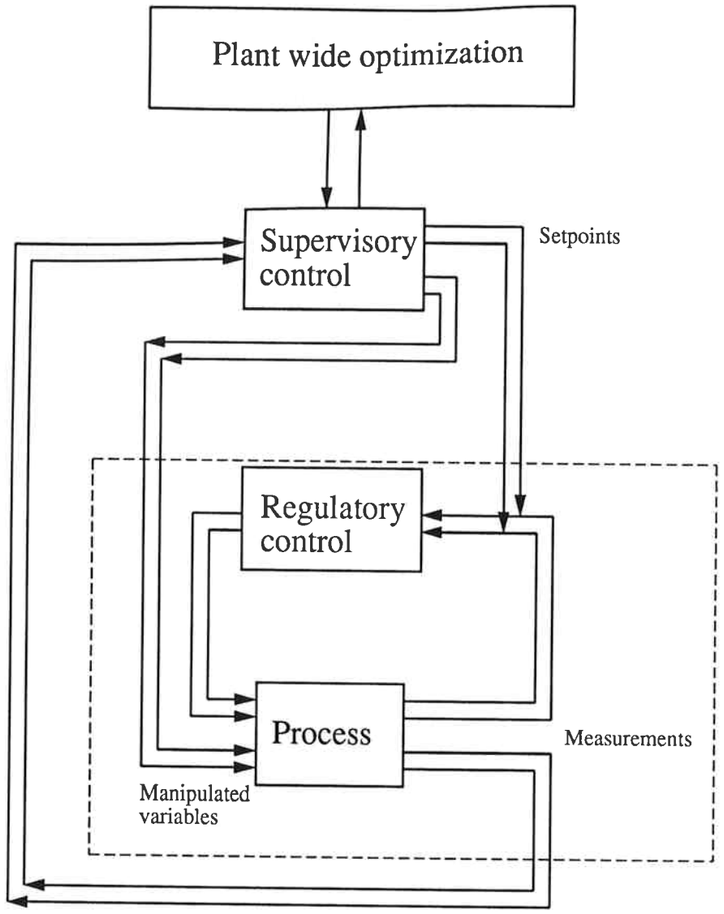


Figure 4.2: Schematic representation of a hierarchical control system.

at regular intervals. A schematic representation of such a control hierarchy is depicted in Fig. 4.2. Note that we have not in this figure included functions related to logic control (startup/shutdown) and safety systems. These are of course important, but need not be considered during normal operation.

Although the implementation may be done in many different ways, and even on the same control system, it is still important to distinguish between the various control levels due to their different objectives. The largest amount of money is probably made at the higher levels, but the lower level must function properly to realize the benefits

of the higher levels. Although seemingly obvious, this is often not understood.

The regulatory control system should fulfill the following objectives:

O1. It should provide a sufficient quality of control to enable a trained operator to keep the plant running safely without use of the higher levels in the control system. This sharply reduces the need for providing costly backup systems for the higher levels of the control hierarchy in case of failures.

O2. It should be simple to understand and tune. Thus, in most cases simple decentralized control loops are used at this level. There are of course cases for which interactions are so strong that multivariable control may be needed at this level. However, very simple schemes are then preferred to compensate for interactions, such as ratios, sums, etc.

O3. It should make it possible to use simple (at least in terms of dynamics) models at the higher level. We want to use relatively simple models because of reliability and the prohibitive costs involved in obtaining and maintaining a detailed dynamic model of the plant, and because complex dynamics will add to the computational burden on the higher level control system. This may be achieved by having a regulatory control level at the bottom of the control hierarchy. This may also reduce the effect of model uncertainty and provide for local linearization, for example, by using a cascade on a valve to avoid the nonlinear valve characteristics.

O4. It should make it possible to use longer sampling intervals at the higher levels of the control hierarchy in order to reduce the need for computing power at this level. Preferably, the time scales of the lower-level and higher-level control system should be separated such that response of the lower-level control system, as seen from the higher level, is almost immediate.

As a consequence of the objectives listed above, the following more specific objectives for the regulatory control system arise:

O5. It should provide for fast control when this is needed for some variables.

O6. It must be able to follow the setpoints set by the higher levels in the control hierarchy. The setpoints of the lower loops are the manipulated variables for the higher levels in the control hierarchy, and we want to be able to change these variable as directly and with as little interaction as possible. Otherwise, the higher level will need a model of the dynamics and interactions of the outputs from the lower level control system.

O7. It should provide for local disturbance rejection. This follows from O6, since we want to be able to keep the controlled variables in the regulatory control system at their setpoints. As disturbances we must also include the "unused" manipulated variables (additional degrees of freedom) which are adjusted directly by the higher levels of the control system.

O8. It should be designed such that the remaining control problem does not contain unnecessary performance limitations such as RHP-zeros, large RGA elements, or strong sensitivity to disturbances. The "remaining control problem" is the control problem as seen from the higher level which has as manipulated inputs the "unused" manipulated inputs and the setpoints to the lower-level control system. By "unnecessary" is meant limitations that do not exist in the original problem formulation without the lower-level control system in place.

In this paper we will primarily consider objectives 5, 6 and 7. These objectives are related to the "controllability" of the lower level control system. Objective 2 is automatically fulfilled since we will only consider a fully decentralized lower level control system.

Control Structure Selection

To fulfill the objectives for the regulatory control system listed above, one must perform a control structure selection. This involves making the following structural decisions:

D1. Outputs y: selection of controlled variables (control objectives for the regulatory control system).

These variables include *primary* and *secondary* controlled variables. The primary outputs are often easy to select as they are variables which are important to control in themselves, also in terms of the overall control objective. Typically, these include variables for which reasonably fast control is needed (see Objective 5 above), such as liquid levels, certain temperatures and pressures. The secondary outputs are usually easily measured variables which in them selves are not important to control, but which are selected to meet objectives 1-3 above. Typically, the secondary variables include temperatures and pressures at selected locations in the process. The problem of selecting the output variables for the regulatory control system is therefore closely related to the issue of measurement selection.

D2. Inputs u: selection of manipulated variables for the regulatory control system.

These selected inputs will be a subset of all possible manipulated inputs, and the remaining "unused" variables will be manipulated inputs available for the higher levels.

D3. Pairing of the chosen controlled and manipulated variables for decentralized control.

The choice of pairing will influence the effect of interactions and disturbances, as well as the systems ability to tolerate failure of one or more loops in the decentralized control system.

For decision D1, selection of outputs, the FCC example presented in this paper provides a very good example. Here a relevant issue is whether to control the regenerator temperature, T_{rg}, or the riser outlet temperature, T_{ro} as secondary controlled outputs for the regulatory control system (see Fig. 4.1).

Distillation column control provides an excellent example of the importance of selecting appropriate inputs (decision D2) to use for the lower-level control system, which in this case is the level control system. It is well known that closing the level loops with the "LV-configuration" (corresponding to having reflux L and boilup V as the remaining unused inputs for composition control) may make the remaining composition control problem very difficult because of serious interactions (resulting in large RGA-values, see e.g. [26]). Note that the lower level control system for the LV-configuration meets essentially all of the objectives mentioned above, except objective 8.

4.3 Measures for Evaluating Controllability

The measures used in this paper for evaluation controllability are outlined in this section. Additional measures also exist, see e.g. [31, 29].

Right Half Plane Transmission Zeros. A right half plane (RHP) transmission zero of $G(s)$ limits the achievable bandwidth of the plant. This holds regardless of the type of controller used (e.g., [18]). The reason is that with a RHP transmission zero the controller can not invert the plant and perfect control is impossible. Thus plants with RHP transmission zeros within the desired bandwidth should be avoided.

In the multivariable case a RHP transmission zero of $G(s)$ does not imply that the matrix elements, $g_{ij}(s)$, have RHP zeros. Conversely, the presence of RHP zeros in the elements does not necessarily imply a RHP transmission zero of $G(s)$. If we use a multivariable controller then RHP zeros in the elements do *not* imply any particular problem. However, if decentralized controllers are used, then we generally avoid pairing on elements with "significant" RHP zeros (RHP zeros close to the origin), because otherwise this loop may go unstable if left by itself (with the other loops open).[2]

Relative Gain Array. The relative gain array (RGA) has found widespread use as a measure of interaction and as a tool for control structure selection for single-loop controllers. It was first introduced by Bristol [1]. It was originally defined at steady-state, but it may easily be extended to higher frequencies [2]. Shinskey [23, 24] and several other authors have demonstrated practical applications of the RGA. Important advantages with the RGA is that it depends on the plant model only and that it is scaling independent. For $n \times n$ plants $G(s)$ the RGA matrix can be computed frequency-by-frequency ($s = j\omega$) using the formula

$$\Lambda(s) = G(s) \times (G^{-1}(s))^T \tag{4.1}$$

where the $\times$ symbol denotes element by element multiplication (Hadamard or Schur product). An important use of the RGA is that pairing on negative *steady-state* relative gains should be avoided [5]. The reason is that with integral control this yields instability of either 1) the overall system, 2) the individual loop, or 3) the remaining system when the loop in question is removed. It is also established that plants with

[2] Usually the main reason for using a decentralized control system in the first place is to allow for loops to be operated independently.

large RGA-values, in particular at high frequencies, are fundamentally difficult to control irrespective of the controller used (poor controllability) [28]. On the other hand, if the RGA elements corresponding to the paired inputs and outputs are small in the bandwidth region, this indicates possible stability problems when using decentralized control [10, 11].

PRGA. One inadequacy of the RGA (e.g. [17], p. 166) is that it only measures two-way interactions (e. g. $\Lambda = I$ for a triangular plant), and it may therefore indicate that interactions are not a problem when significant one-way coupling exist. To overcome this problem we introduce the frequency dependent performance relative gain array (PRGA). The PRGA matrix is defined as

$$\Gamma(s) = \tilde{G}(s)G(s)^{-1} \tag{4.2}$$

where $\tilde{G}(s)$ is the matrix consisting of only the diagonal elements of $G(s)$, i.e., $\tilde{G} = diag\{g_{ii}\}$. The matrix Γ was originally introduced at steady-state by Grosdidier [7] in order to understand the effect of directions under decentralized control. The elements of Γ are given by

$$\gamma_{ij}(s) = g_{ii}(s)[G^{-1}(s)]_{ij} = \frac{g_{ii}(s)}{g_{ji}(s)}\lambda_{ji}(s) \tag{4.3}$$

Note that the diagonal elements of RGA and PRGA are identical, but otherwise PRGA does not have all the nice algebraic properties of the RGA. PRGA is independent of *input* scaling, that is, $\Gamma(GD_2) = \Gamma(G)$, but it depends on output scaling. This is reasonable since performance is defined in terms of the magnitude of the outputs.

Closed Loop Disturbance Gain. A disturbance measure closely related to the PRGA, the closed loop disturbance gain (CLDG), was recently introduced by Skogestad and Hovd [25]. For a disturbance k and an output i, the CLDG is defined by

$$\delta_{ik}(s) = g_{ii}(s)[G(s)^{-1}G_d(s)]_{ik} \tag{4.4}$$

A matrix of CLDG's may be computed from

$$\Delta = \{\delta_{ik}\} = \tilde{G}G^{-1}G_d \tag{4.5}$$

The CLDG is scaling dependent, as it depends on the expected magnitude of disturbances and outputs. Actually, this is reasonable since CLDG is a performance measure, which generally are scaling dependent.

The use of PRGA and CLDG: Performance Relationships for Decentralized Control. The following derivation follows [25, 10]. Assume the controller $C(s)$ is diagonal with entries $c_i(s)$. (see Fig. 4.3). This implies that after the variable pairing has been determined, the order of the elements in y and u has been arranged so that the plant transfer matrix $G(s)$ has the elements corresponding to the paired variables on the main diagonal. Let $y(s)$ denote the output response for the overall system when all loops are closed and let $e(s) = y(s) - r(s)$ denote the output error. The closed loop response becomes

$$e(s) = -S(s)r(s) + S(s)G_d(s)d(s); \quad S = (I + GC)^{-1} \tag{4.6}$$

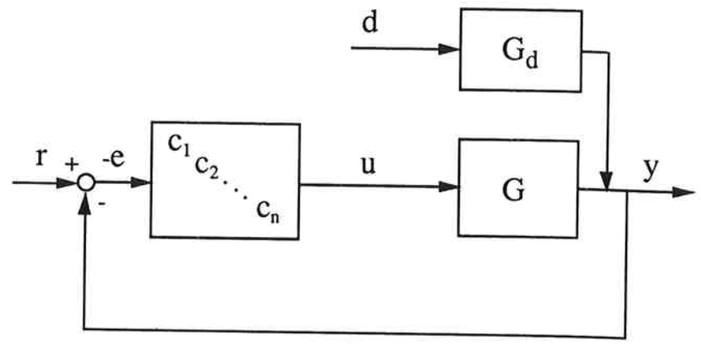


Figure 4.3: Block diagram of controller and plant.

where $S(s)$ is the sensitivity function for the overall system, and $d(s)$ denotes the disturbances. The Laplace variable s is often omitted to simplify notation.

At low frequencies ($\omega < \omega_B$) we usually have $S \approx (GC)^{-1} = C^{-1}\tilde{G}^{-1}\tilde{G}G^{-1} = (\tilde{G}C)^{-1}\tilde{G}G^{-1} \approx \tilde{S}\tilde{G}G^{-1}$ and we get

$$e \approx -\tilde{S}\tilde{G}G^{-1}r + \tilde{S}\tilde{G}G^{-1}G_d d; \quad \omega < \omega_B \tag{4.7}$$

Here we recognize the PRGA ($\tilde{G}G^{-1}$) and the CLDG ($\tilde{G}G^{-1}G_d$). When we consider the effect of a setpoint change r_j and a disturbance d_k on the offset e_i this gives

$$e_i \approx -\frac{\gamma_{ij}}{g_{ii}c_i}r_j + \frac{\delta_{ik}}{g_{ii}c_i}d_k; \quad \omega < \omega_B \tag{4.8}$$

From (4.8) we see that the ratio $\gamma_{ij}/(g_{ii}c_i)$ gives the magnitude of the offset in output i to a unit setpoint change for output j. This ratio should preferably be small. That is, on a conventional magnitude Bode plot (log-log), the curve for $|\gamma_{ij}|$ should lie below $|g_{ii}c_i|$ at frequencies where we want small offsets. From (4.8) we see that the ratio $\delta_{ik}/(g_{ii}c_i)$ gives the magnitude of the offset in output i to a unit disturbance d_k. That is, the curve for $|\delta_{ik}|$ should lie below $|g_{ii}c_i|$ at frequencies where we want the offsets less than 1 in magnitude. A plot of $|\delta_{ik}(j\omega)|$ will give useful information about which disturbances k are difficult to reject.

Assume that G and G_d have been scaled such that 1) the expected disturbances, $|d_k(j\omega)|$, are less than or equal to one at all frequencies, and 2) the outputs, y_i are such that the allowed errors, $|e_i(j\omega)|$, are less than or equal to one. *In this case the frequency where $|\delta_{ik}(j\omega)|$ crosses one, directly corresponds to the minimum bandwidth needed in loop i to reject disturbance k.* It is preferable that this frequency is low in order to avoid stability problems for the individual loops.

Limitations of (4.8). The main limitation with (4.8) is that it applies only to lower and intermediate frequencies. Furthermore, the issue of stability is not addressed. Another limitation is the assumption that all diagonal elements in $G(s)$ are nonzero. In particular, it is obvious that relations involving $|g_{ii}c_i|$ in the denominator are not meaningful when $g_{ii} = 0$. These issues are addressed by Hovd and Skogestad [10].

4.3.1 Summary of Controllability Rules

Let us at this point summarize some results we shall use in this paper:

Rule 1. Avoid plants (designs) with RHP transmission zeros within the desired bandwidth (i. e. RHP transmission zeros at low frequencies are bad).

Rule 2. Avoid plants (designs) with large RGA-values (in particular at frequencies near cross-over). This rule applies for any controller, not only to decentralized control [28].

Rule 3. Avoid *pairings ij* with negative values of the steady-state RGA, $\lambda_{ij}(0)$ [5].

Rule 4. Prefer *pairings ij* where $g_{ij}(s)$ puts minimal restrictions on the achievable bandwidth for this loop, that is, avoid pairings with RHP-zeros in $g_{ij}(s)$. The rule follows from (4.8) above in order to satisfy performance and at the same time have stability of the individual loop. Rule 4 is the conventional rule of pairing on variables "close to each other".

Rule 5. For decentralized control avoid control structures (an entire *set* of *pairings*) with large values of $|\delta_{ik}|$ or $|\gamma_{ij}|$ in the crossover region, and in particular if the achievable bandwidth for the corresponding loop i is restricted (because of $g_{ii}(s)$, see rule 4) (the rule follows from (4.8) above).

Rule 6. For decentralized control prefer pairings with RGA values close to 1 in the crossover region.

4.4 The FCC process

4.4.1 FCC Operating Modes

The FCC process can be operated in two distinct modes, the partial combustion mode and the complete combustion mode. The emphasis in this paper will be on the partial combustion mode.

In the partial combustion mode, large amounts of both CO and CO_2 are formed when the coke is burnt in the regenerator. The CO rich regenerator flue gas can be sent to a CO boiler where high pressure steam is produced. However, if there are significant amounts of oxygen leaving the regenerator dense bed, this will react with the CO to form CO_2 in the zone above the regenerator dense bed, or in the regenerator cyclones and downstream piping. This "afterburning" is a strongly exothermic reaction, and since there is relatively little mass in this zone a large temperature rise will occur if there are significant amounts of oxygen present. It is therefore necessary to control the afterburning to avoid violating metallurgical temperature limits for the regenerator cyclones or downstream piping.

In the complete combustion mode, little CO leaves the regenerator dense bed because excess quantities of air are supplied so that most of the CO formed by the combustion of coke is oxidized to CO_2 within the regenerator dense bed. Special catalysts which promote the oxidation of CO to CO_2 may also be used [20]. When operating in

the complete combustion mode afterburning is therefore not such a serious concern. However, it is not always possible to operate an FCC unit in the complete combustion mode, especially if the feed oil has a large coke production tendency. There is also an economic incentive for operating in the partial combustion mode, as the heat recovered in the CO boiler is valuable.

4.4.2 FCC Model

The models used in this work derive in the main part from the model proposed by Lee and Groves [14] for the partial combustion mode. This model augments the regenerator model of Errazu and coworkers [4] with the riser model of Shah and coworkers [22]. Details about the model are given in Appendix 1.

Riser Model

A static model is used for the riser. We use an ideal plug flow model and the three lump kinetic scheme of Weekman and Nace [30], where the feed is gas oil, which can crack to gasoline or light gases/coke. The static riser model is used to compute

- T_{ro} — temperature at the riser outlet.
- C_{sc} — mass fraction of coke on catalyst at riser outlet.

Regenerator Model

When modeling the regenerator it is common to assume that the temperature and the amount of coke on the catalyst is uniform throughout the regenerator dense bed. In the model used in this paper, oxygen is also assumed to be uniformly distributed, as Errazu and coworkers [4] found that this assumption allows operational data can be described well.

This yields a third order model for the regenerator, with the following states:

- C_{rc} — mass fraction of coke on regenerated catalyst.
- T_{rg} — regenerator dense bed temperature.
- O_d — mole fraction of oxygen in the gas leaving the dense bed.

To compute the regenerator cyclone temperature T_{cy}, we represent the afterburning of CO to CO_2 in the dilute phase in the regenerator by using a simple equation taken from [12]

$$T_{cy} = T_{rg} + c_t O_d \tag{4.9}$$

For Eq. (4.9) to be reasonable, there must be an excess of CO over O_2 in the gas leaving the regenerator dense bed, that is, Eq. (4.9) is only valid in the partial combustion mode. Note that when using Eq. (4.9), controlling $\Delta T_{rg} = T_{cy} - T_{rg}$ (the temperature rise from regenerator dense bed to regenerator cyclones) is equivalent to controlling O_d (the amount of oxygen leaving the regenerator dense bed).

Some authors use the assumption that the oxygen moves in perfect plug flow through the dense bed (e.g. [12], [13]). In the discussion we will comment on how the assumptions about oxygen flow pattern affect our conclusions.

Stripper Model

The stripper is modelled as a mixing tank, and yields two additional state variables:

C_{st} — mass fraction of coke on catalyst in the stripper.

T_{st} — temperature in stripper.

The catalyst holdup in the stripper W_{st} is assumed to be kept constant by perfect control. This means that the flowrate of spent catalyst from the stripper to the regenerator, $\hat{F}_s$, equals the flowrate of regenerated catalyst, F_s.

Complete Combustion Mode

For the complete combustion mode the same model is used, except that some of the parameter values are adjusted, as discussed in Appendix 1.

4.4.3 Constraints in FCC operation

The optimal operating point for an FCC usually lies at one or several constraints. The control structure which allows operation closest to the constraints is therefore preferable. The location of the optimal operating point, and consequently the importance of the different constraints can vary depending on the feed characteristics and the desired product split. Different control structures may thus be preferable at different operating points, but it is not realistic to expect the control structure to be reconfigured when the operating conditions are changed.

Common constraints include:

- Maximum regenerator cyclone temperature (T_{cy}) constraint. This constraint is usually important in the partial combustion mode, and is determined by the metallurgical properties of the cyclones.
- Minimum flue gas oxygen concentration (O_{fg}) constraint. This constraint is important in the complete combustion mode, as a sufficient concentration of oxygen in the flue gas ensures virtually complete combustion of CO to CO_2 within the regenerator dense bed, and therefore ensures that afterburning is avoided. Furthermore, sufficient oxygen in the flue gas also ensures that the concentration of CO in the flue gas is within environmental limits.
- Maximum wet gas compressor capacity. The wet gas compressor is situated downstream of the FCC unit, and compresses the wet gases produced in the FCC for transportation to downstream gas treatment plants.
- Maximum air blower capacity (F_a). The air blower provides the air needed for the combustion in the regenerator.

Implications for regulatory control: Constrained outputs (measurements) should be selected as controlled outputs for the regulatory control system, thus enabling operation close to these constraints. This means that there is an argument for selecting the regenerator cyclone temperature T_{cy} as a controlled variable in the partial combustion mode, and for selecting the flue gas oxygen concentration O_{fg} as a controlled variable in the complete combustion mode.

Manipulated inputs that are prone to reach constraints should be avoided in the regulatory control system. The feed flowrate F_f strongly influences the wet gas production, and should be avoided as a manipulated variable for plants operating close to the wet gas compressor constraint. Similarly, the air blower capacity constraint is an argument for avoiding the use of the air flowrate F_a as a manipulated variable for regulatory control. Nevertheless, in papers on regulatory control of the FCC process (e.g. [19, 9]) F_a is consistently used as a manipulated variable. For plants operating close to the air blower capacity constraint, the supervisory control level must then ensure that this constraint is not encountered, for instance by changing reaction conditions or feed composition such that less coke is formed.

For a more complete description of the constraints encountered in the operation of FCC's, see [8].

4.5 Controllability Analysis for the FCC process

4.5.1 Scope of the Controllability Analysis

The following sections will address:

- Choice of controlled variables. How does the choice of controlled variables affect controllability?
- Pairing of controlled and manipulated variables for decentralized control.
- Effect of operating point. Is the control system sensitive changes in the operating conditions?
- Sensitivity to parametric uncertainty. Do changes in parameter values lead to different conclusions?
- Sensitivity to input uncertainty. The actual moves in the manipulated variables will not be exactly equal to those calculated by the controller. Does this influence performance?
- Disturbance rejection. Using the CLDG explained in section 3, we will investigate the effect of disturbances on the FCC when decentralized control is used.
- Effect of model features. Which features of the models are important for making decisions about the control of the FCC process?

These points will be examined for both the partial combustion and complete combustion modes.

In the following we assume that a decentralized control system is used and that the transfer function matrices have been arranged to give the paired elements on the diagonal. The word "RGA" or "RGA's" will refer to the diagonal elements of the RGA-matrix of $G(s)$. Note that the RGA's are identical when $G(s)$ is a 2×2 matrix.

4.5.2 Variable Classification

Independent Variables (u's and d's). We will consider the following six independent variables:

- F_s - The flowrate of regenerated catalyst entering the riser[3].
- F_a - The flowrate of air to the regenerator.
- k_c - The coke production rate factor (i. e., feed oil composition)
- T_f - The feed oil temperature.
- F_f - The total feed oil flowrate (both fresh and recycled oil).
- T_a - The air temperature.

There are actually a few more manipulated variables, but we have assumed that these are already used by the regulatory control level to control holdup and pressure. These variables include (see Fig. 4.1):

- $\hat{F}_s$ - Spent catalyst flowrate.
- F_{fg} - Flue gas flowrate.
- W_{wg} - Wet gas compressor throughput[4].

$\hat{F}_s$ is used to control the catalyst holdup in the stripper, whereas F_{fg} is used to control the regenerator pressure P_{rg}. W_{wg} indirectly controls the stripper pressure P_{st}. Since we have assumed the loops to be closed, we should actually have included the pressures P_{rg} and P_{st} as disturbances for the model we are considering. However, we have not done this in order to keep the model simple. In practice the pressures P_{rg} and P_{st} may have to be adjusted when the catalyst slide valve position is changed, in order to avoid reversal of the catalyst flow (F_s). This is taken care of by the supervisory control system.

Other possible independent variables which are not included here are the fraction of dispersion steam, λ, and the stripping steam flowrate (not included in the model), as these are known to have relatively little effect provided they are above some minimum threshold values necessary for the proper operation of the plant[5].

All the six independent variables above may be used as manipulated variables (u's) for control, but in most cases we will only use two: the regenerated catalyst flowrate, F_s and the air flowrate F_a.

[3]The true manipulated variable is actually the regenerated catalyst slide valve position. We use the regenerated catalyst flowrate as a manipulated variable in order to keep the model simple.

[4]W_{wg} is not shown in Fig. 4.1, as the wet gas compressor is situated downstream of the distillation column receiving the reaction products.

[5]A minimum amount of dispersion steam is needed for good and rapid mixing of feed and catalyst at the riser entrance. A minimum amount of stripping steam is needed for effective stripping of volatile hydrocarbons from the spent catalyst.

The remaining independent variables may then be regarded as disturbances (d's) in the regulatory control system. The variables k_c, T_f and F_f are all related to the oil feed. k_c is the coke production rate factor and depends on the feed composition. (Immediately downstream of the FCC there is a distillation column which separates the products from the cracking reactions. The heavy fraction from the distillation column, "slurry", has a large coke producing tendency. The coke production rate factor can therefore be changed indirectly by changing the amount of slurry which is recycled to the riser). The air temperature T_a is generally a disturbance since there is usually no air preheater.

Measurements. Typically, the following measurements are available:
Partial combustion mode:

- T_{ro} - The riser outlet temperature.
- T_{rg} - The regenerator dense bed temperature.
- T_{cy} - The regenerator cyclone temperature.

Complete combustion mode:

- T_{ro} - The riser outlet temperature.
- T_{rg} - The regenerator dense bed temperature.
- $O_{fg} \approx O_d$ - The oxygen concentration in the flue gas

Controlled Variables (y's) (Structural decision D1.)
It is not obvious what controlled variables should be used for the regulatory control system. In some implementations only T_{ro} is controlled by the regulatory control system (see [8]). However, in this work we shall first consider controlling three variables: T_{ro}, T_{rg}, and T_{cy} (partial combustion mode) or O_{fg} (complete combustion mode). The justification for trying to control three variables is as follows:

The product distribution is determined by the reaction conditions inside the riser, which are therefore very important for the economic performance of the FCC process. Both T_{ro} and T_{rg} are related to the conditions in the riser. There is an incentive to control *both* T_{ro} and T_{rg}, as the product distribution depends both on the temperature inside the riser and the catalyst to oil ratio. For a given value of T_{ro}, a high value for T_{rg} implies a low catalyst to oil ratio.

The need to control afterburning in the partial combustion mode and to avoid afterburning in the complete combustion mode should be obvious. Specifically, the controlled variables considered in this work are:
Partial combustion mode:
Primary variable: T_{cy} or $\Delta T_{rg} = T_{cy} - T_{rg}$.
Secondary variables: T_{ro} and T_{rg}
Complete combustion mode:
Primary variable: O_{fg}.
Secondary variables: T_{ro} and T_{rg}.

4.6 Analysis of the Partial Combustion Mode

All the numerical values in this section are for Case P1 in Table 4.1, unless otherwise is stated.

4.6.1 Control Structure Alternatives

3×3 Control Systems. As discussed above we may want to control three outputs.

$$y^3 = (T_{ro}, T_{cy}, T_{rg})^T \tag{4.10}$$

We will consider the following rwo sets of manipulated variables (Structural decision D2.):

$$u_1^3 = (F_s, F_a, k_c)^T \tag{4.11}$$

$$u_2^3 = (F_s, F_a, T_f)^T \tag{4.12}$$

Here we assume that the unit is operating at some maximum capacity limit such that F_f is not available as a manipulated variable. In the following we assume that the individual inputs and outputs are numbered according to their position in y^3, u_1^3, and u_2^3, and that the transfer function matrix $G(s)$ has been arranged accordingly (such that, e.g. the 1,2 element of $G(s)$ is the transfer function from F_a to T_{ro}).

When computing the RGA matrices, we find that the only pairings giving positive steady state RGA values are those indicated by Eqs. (4.10)-(4.12). We thus find that when using the inputs in u_1^3, the pairing T_{ro}-F_s, T_{cy}-F_a, T_{rg}-k_c should be used. Likewise, when using the inputs in u_2^3, the pairing T_{ro}-F_s, T_{cy}-F_a, T_{rg}-T_f should be used. These RGA's are shown as functions of frequency in Fig. 4.4a for u_1^3, and in Fig. 4.4b for u_2^3. We see that u_1^3 gives RGA values reasonably close to 1 at steady state, but the small RGA's for loops 2 and 3 in the desired bandwidth region (approximately 1 [rad/min]) indicate control problems. The use of u_2^3 gives rise to unfavorably large steady state RGA's, and in the desired bandwidth area the RGA's are small for loops 1 and 3.

2×2 Control Systems: Selection of Controlled Variables. The above results indicate that with the available manipulated variables, the 3 × 3 control problem is not well suited for high performance control with a decentralized controller. *We will therefore in the following consider the 2×2 problem.*

Restricting our attention to the 2×2 control problem allows us to compare and contrast our results with the results of previous authors, and will give a good illustration of the effect measurement selection can have on plant controllability. Furthermore, the output variables T_{ro} and T_{rg} are strongly coupled, and good control of one will in many cases also limit the offset in the other variable. There is thus probably little need to control both T_{ro} and T_{rg} in order to fulfill the objectives for the lower level control system given above. In the following we will use

$$u = (F_s, F_a)^T \tag{4.13}$$

as manipulated variables. These two manipulated variables are always used for 2×2 control systems for FCC's, because of their strong and direct effect on the process

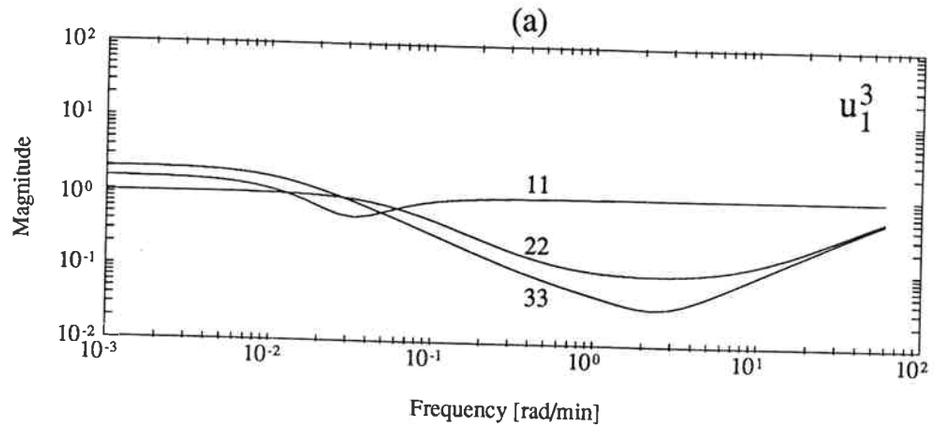


Figure 4.4: RGA for the 3 × 3 control problem in the partial combustion mode. a) y^3-u_1^3, with pairing T_{ro}-F_s, T_{cy}-F_a, T_{rg}-k_c. b) y^3-u_2^3, with pairing T_{ro}-F_s, T_{cy}-F_a, T_{rg}-T_f (Case P1, Table 4.1).

conditions. Wolff et al. [31] also found from an analysis of the disturbance sensitivities that this is a good choice. The remaining independent variables k_c, T_f, F_f and T_a will in the following be considered as disturbances. We will investigate the following five choices of controlled variables:

1. $y_C = (T_{ro}, \Delta T_{rg})^T$. This corresponds to the *conventional control structure*, which uses F_s to control T_{ro} and F_a to control ΔT_{rg}. It appears to have been described first by Pohlenz (1963), but the name is due to Kurihara (1967).

2. $y_K = (T_{rg}, \Delta T_{rg})^T$. This corresponds to the *Kurihara control structure* (Kurihara, 1967), which uses F_a to control T_{rg} and F_s to control ΔT_{rg}.

3. $y_{AK} = (T_{rg}, T_{cy})^T$. We have chosen to call this the *alternative Kurihara* control structure, as the elements of y_{AK} are linear combinations of the elements of y_K.

4. $y_H = (T_{ro}, T_{cy})^T$. We call this the *Hicks control structure* as it was first proposed by Hicks [9], who used F_s to control T_{ro} and F_a to control T_{cy}.

5. $y_R = (T_{ro}, T_{rg})^T$. We will name this the *riser-regenerator control structure.*

4.6.2 Scaling of Variables

The PRGA and CLDG depend on scaling, and the variables must therefore be scaled appropriately to allow easy interpretation of the PRGA and CLDG. The PRGA depends on the scaling of the outputs, and the CLDG depends on the scaling of outputs and disturbances.

For the partial combustion mode, the transfer functions are scaled such that output errors, $e_i = y_i - r_i$ of magnitude 1 correspond to

1. Riser exit temperature, T_{ro}: 3 K
2. Regenerator cyclone temperature, T_{cy}: 2 K
3. Regenerator dense bed temperature, T_{rg}: 3 K

and such that a disturbance d_k of magnitude 1 corresponds to

1. Feed oil temperature, T_f: 5 K
   (when considered as a disturbance)
2. Air temperature, T_a: 5 K
3. Feed oil flowrate, F_f: 4 kg/s (ca. 10%)
4. Feed oil composition, expressed by the coke production rate factor k_c: 2.5 % relative to its original value (when k_c is considered as a disturbance)

4.6.3 RHP Transmission Zeros

1) Conventional Control Structure, $y_C = (T_{ro}, \Delta T_{rg})^T$.
For this choice of controlled variables we get a transfer function matrix with a RHP transmission zero at 0.018 [rad/min]. This RHP transmission zero will seriously limit the achievable bandwidth, and only slow control is possible. This choice of controlled variables will therefore be discarded.

2) Kurihara Control Structure, $y_K = (T_{rg}, \Delta T_{rg})^T$.
For this choice of controlled variables we get a RHP transmission zero 0.19 [rad/min]. This is one decade higher than for the conventional control structure, and is thus a much less severe restriction on the achievable bandwidth. This result is in accordance with the results of Kurihara [12] and Lee and Groves [14], who found the Kurihara control structure to be preferable to the conventional control structure.

3) Alternative Kurihara Control Structure, $y_{AK} = (T_{rg}, T_{cy})^T$.
For this choice we get the same result as for the Kurihara control structure, a RHP transmission zero at 0.19 [rad/min]. This is as expected, since the measurements for the alternative Kurihara control structure are linear combinations of the measurements for the Kurihara control structure ($T_{cy} = T_{rg} + \Delta T_{rg}$).

4) Hicks Control Structure, $y_H = (T_{ro}, T_{cy})^T$.
For this choice of controlled variables we get no RHP transmission zero (see also Table 4.1). This means that there is no inherent bandwidth limitation in the model, and the bandwidth will be limited solely by unmodeled effects and uncertainty.

5) Riser-regenerator control structure $y_R = (T_{ro}, T_{rg})^T$.
Also for this choice of controlled variables there is no RHP transmission zero.

The existence and location of RHP transmission zeros is a fundamental measure of controllability, as a RHP transmission zero will limit the achievable bandwidth for any type of controller. Our analysis of RHP zeros clearly indicate that one either should choose $y_H = (T_{ro}, T_{cy})^T$ or $y_R = (T_{ro}, T_{rg})^T$ as controlled variables.

It may be argued that controlling T_{cy} and T_{ro} is the best choice with respect to the constraints:

- Controlling T_{cy} avoids exceeding the metallurgical temperaure limit in the regenerator cyclones.
- The amount of wet gases produced is a strong function of the riser exit temperature, T_{ro}. Controlling T_{ro} therefore helps ensuring that the wet gas compressor operating limits are not exceeded. Controlling T_{rg} instead of T_{ro} would not have the same beneficial effect on the operation of the wet gas compressor.

When using $y_R = (T_{ro}, T_{rg})^T$ as controlled variables, one must ensure that T_{cy} does not drift above its metallurgical constraint. This may be done with a (slower) control loop from T_{cy} to the setpoint for T_{rg}, or by including T_{cy} as a controlled variable for the supervisory control system. Based on the results in this section, we will in the following concentrate on the Hicks control structure, but also state briefly results for the riser-regenerator control structure, which also has the potential for high performance control.

4.6.4 Pairing of Controlled and Manipulated Variables

For the Hicks control structure, the steady state RGA for the pairing $T_{ro} - F_s$, $T_{cy} - F_a$, proposed by Hicks [9], is about 0.5. In Fig. 4.5 the PRGA for this pairing is shown. We see that the PRGA is relatively small for all frequencies, and approaches identity for frequencies higher than 0.1 [rad/min]. As the desired bandwidth is above 0.1 [rad/min], the interaction between the loops can be expected to be small for the pairing $T_{ro} - F_s$, $T_{cy} - F_a$.

For the riser-regenerator control structure, the RGA is close to 1 at steady state for the pairing $T_{ro} - F_s$, $T_{rg} - F_a$. The frequency dependent PRGA (not shown) also indicates that this pairing gives little interaction between the loops.

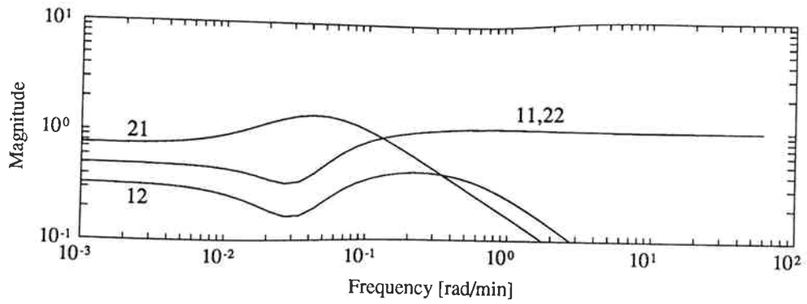


Figure 4.5: PRGA for the Hicks control structure with the pairing $T_{ro} - F_s$, $T_{cy} - F_a$ (Case P1, Table 4.1).

4.6.5 Disturbances

The effect of disturbances has been investigated using the closed loop disturbance gain (CLDG) explained above. Note that it is not meaningful to use the CLDG to directly compare different choices of controlled output variables, as the CLDG contains no information about uncontrolled outputs. For this reason, we have chosen to include the CLDG's only for the Hicks control structure.

Based on the results in the previous sections, we will here only consider the pairing $T_{ro} - F_s$, $T_{cy} - F_a$. The frequency-dependent CLDG's shown in Fig. 4.6 predict that disturbance $k = 3$ in the feed oil flowrate, F_f, is most difficult to reject, followed by disturbances 1 (in T_f) or 4 (in k_c). Disturbance 2 (in T_a) appears to have very little effect. δ_{13}, the CLDG for the effect of F_f on T_{ro} does not roll of at high frequencies. Some high frequency effect of F_f on T_{ro} must therefore be expected. This suggests the use of feedforward from F_f to F_s used together with feedback control for good disturbance rejection at high frequencies, unless F_f is controlled such that only slow changes in this variable can occur. The predictions based on the CLDG's, which are independent of the controller, are verified by closed loop simulations using two PI controllers in Fig. 4.7. Since it is mainly unmodelled effects that limit the achievable bandwidth for the Hicks control structure, we have chosen to tune each loop to give a closed loop bandwidth for the loop of approximately one minute. This results in the following PI controllers

$$c_1(s) = 1.23\frac{0.56s+1}{0.56s} \tag{4.14}$$

$$c_2(s) = 0.5\frac{s+1}{s} \tag{4.15}$$

These controller tunings result in a sensitivity function S with essentially no peak.

Let us now consider the regenerator temperature, T_{rg}, which is uncontrolled in the Hicks control structure. From the simulation in Fig. 4.7 we see that the effects of disturbances on T_{rg} are relatively small, certainly much better than for the open loop system. Nevertheless, we see from Fig. 4.7 that the disturbances combine to produce an

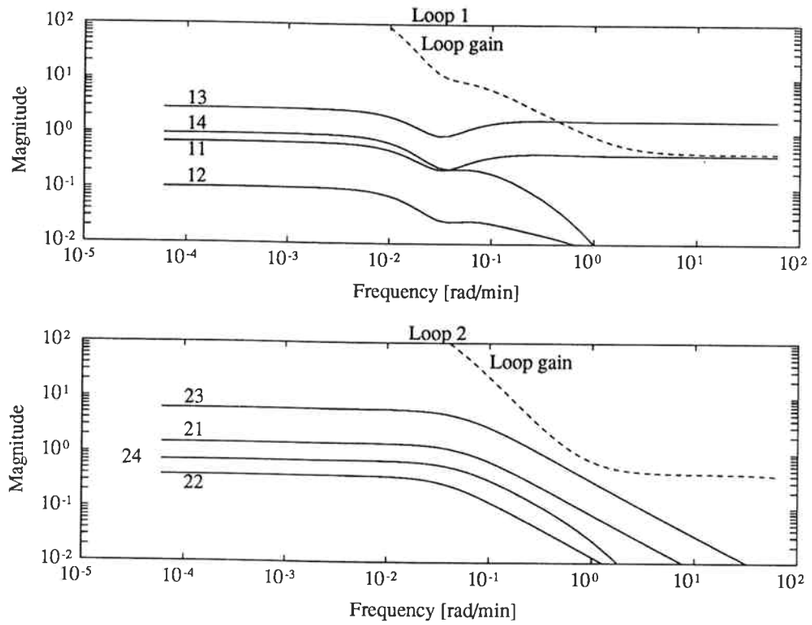


Figure 4.6: Closed loop disturbance gains, δ_{ik}, for the Hicks control structure. i denotes output and k disturbance. The loop gains resulting from the controller used in Fig. 4.7 are shown with dashed lines. (Case P1).

offset in T_{rg} of about 4.5 K. Note, however, that there is no hard constraint associated with T_{rg}. Thus if control of T_{rg} is desired (e.g. using k_c), it can be made sufficiently slow, such that the control of T_{rg} does not interfere significantly with the control of T_{ro} and T_{cy}.

For the riser-regenerator control structure, the CLDG's (not shown) give similar results as for the Hicks control structure. However, we emphasize that with the riser-regenerator control structure T_{cy} is uncontrolled, and measures *must* be taken to ensure that T_{cy} does not exceed its maximum temperature constraint.

4.7 Discussion

In this section we discuss the effects of changes in the operating point, sensitivity to uncertainty in the parameters, and the effects of model features.

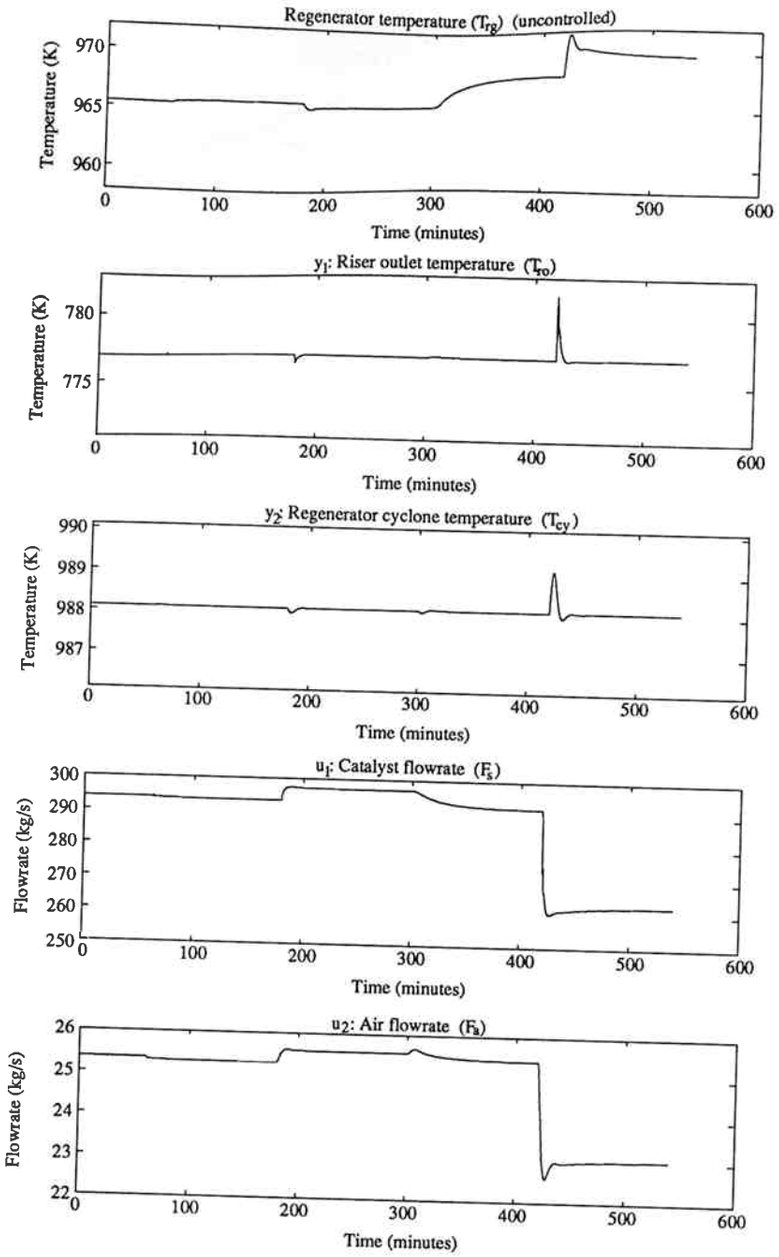


Figure 4.7: Closed loop simulation of the effect of disturbances for the Hicks control structure using the PI controllers in eqs. (4.14)-(4.15). A 5 K step increase in T_a occurs at 60 minutes, 5 K decrease in T_f at 180 minutes, 2.5% increase in k_c at 300 minutes and a 4 kg/s decrease in F_f occurs at 420 minutes. (Case P1).

	Case P1	Case P2
T_{rg}	965.4 K	966.6 K
T_{ro}	776.9 K	770.6 K
T_{cy}	988.1 K	997.4 K
C_{rc}	5.207×10^{-3}	3.578×10^{-3}
$G_H(0)$	$\begin{pmatrix} 0.5587 & 10.16 \\ -0.5577 & 10.35 \end{pmatrix}$	$\begin{pmatrix} 0.3893 & 10.83 \\ -0.7606 & 8.22 \end{pmatrix}$
RGA(0)	0.505	0.280
RHP zeros [rad/min]		
Multivariable	–	–
In elements	$\begin{pmatrix} - & - \\ - & - \end{pmatrix}$	$\begin{pmatrix} - & - \\ - & - \end{pmatrix}$

Table 4.1: Operating Points used for the partial combustion mode. Control-related data is for the Hicks control structure (– denotes that no RHP-zero is present at frequencies below 100 [rad/min]).

4.7.1 Effect of Changes in the Operating Point

The transfer function matrix G_H for the Hicks structure is defined by

$$\begin{bmatrix} T_{ro} \\ T_{cy} \end{bmatrix} = G_H \begin{bmatrix} F_s \\ F_a \end{bmatrix} \tag{4.16}$$

Although a number of operating points ("cases") have been studied, our findings can be illustrated by the two cases in Table 4.1. Case P1 was studied above, and corresponds to an operating point with a catalyst to oil ratio of 7.0, and Case P2 corresponds to a catalyst to oil ratio of 6.7.

The results on *selection of controlled variables* appear unaffected by changes in the operating point, as we have not found any operating point with a RHP transmission zero for the transfer function matrix with $y_H = (T_{ro}, T_{cy})^T$ or $y_R = (T_{ro}, T_{rg})^T$ as controlled variables. The RHP transmission zeros found for the other possible choices of controlled variables are also found when the operating point is changed, although their locations do vary somewhat.

The *choice of pairing* also appears insensitive to changes in the operating point, as the RGA at steady state is positive for the operating points studied, and is close to 1 in our desired bandwidth region, for both the Hicks and riser-regenerator control structures.

4.7.2 Sensitivity to Parametric Uncertainty

The objective of this discussion is to investigate whether small errors in the parameters can have consequences for control performance, and whether parametric uncertainty therefore should be considered in the process of control structure selection. Due to the

large number of parameters in the models, the sensitivity to parameter uncertainty has not been exhaustively researched.

The results on *selection of controlled variables* appear unaffected by parametric uncertainty also. We have not found any parameter change which causes a RHP transmission zero with $y_H = (T_{ro}, T_{cy})^T$ or $y_R = (T_{ro}, T_{rg})^T$ as controlled variables. Again, the RHP transmission zeros found for the other possible choices of controlled variables are also found when parameters are changed, although their locations do vary somewhat.

In general, we have found the *choice of pairing* also to be insensitive to changes in the parameters both for the Hicks and the Grosdidier control structures. Cases has been found for which the steady state RGA indicates that the pairing for the Hicks control structure should be changed. However, on closer scrutiny these cases appear to be unrealistic, since disturbances have such a large effect that the plant would be virtually inoperable.

Lee and Weekman [15] claim that the control structure selection for the FCC process can be very sensitive to the model structure and the parameter values used. Our results appear relatively insensitive to errors in the parameter values for the particular model structure used. The sensitivity to the model structure will be discussed below.

4.7.3 Effect of Model Features

Afterburning.

Inclusion of the simple model for afterburning in the partial combustion mode given in Eq. (4.9) allows considering T_{cy} as a controlled variable. The results in this paper demonstrate how controllability in the partial combustion mode is improved when T_{cy} is chosen as a controlled variable instead of $\Delta T_{rg} = T_{cy} - T_{rg} = c_t O_d$. The afterburning model is therefore important. The results have proved to be relatively insensitive to the value of c_t, and since afterburning is known to be a very fast phenomenon it appears to be little need for modelling afterburning in more detail.

Air Flow Pattern in the Regenerator.

Errazu and coworkers [4] found that the behavior of the regenerator is well described by a model which assumes that the oxygen in the regenerator dense bed is uniformly distributed. Other authors, e. g. Kurihara [12] and Krishna and Parkin [13] assume the air to move in plug flow through the regenerator. We have studied the effect of this assumption on our results.

Kurihara assumes the oxygen leaving the regenerator dense bed to be at a pseudo-steady state, and provides the following equation

$$O_d = O_{in} exp\left[\frac{-WP_{rg}/F_a}{[1.06 \times 10^{10}/F_a^2 + 1/(k_{or} exp(-E_{or}/RT_{rg})C_{rc})]}\right] \tag{4.17}$$

R_{cb}, the rate of coke combustion is found from a mass balance with the oxygen consumed:

$$R_{cb} = \frac{F_a}{M_a}(O_{in} - O_d)\frac{4(1+\sigma)}{(1+\sigma)n + 2 + 4\sigma}M_c \tag{4.18}$$

The values of E_{or} and P_{rg} are taken from Denn [3], whereas the value for k_{or} has been adjusted to give a steady state close to Case P1 in Table 4.1[6].

The results for the Hicks control structure in the partial combustion mode are summarized for two cases in Table 4.2. In Case K1 the gain from F_a to T_{cy} is negative, and we have a negative RGA. The immediate effect of increasing F_a will be to increase O_d, and hence also T_{cy}. However, T_{rg} will also increase, and because O_d in Eq. (4.17) is a very strong function of T_{rg}, O_d eventually decreases to a value below its original value. We therefore get a negative steady state gain from F_a to T_{cy}. In Case K2 in Table 4.2 the value of O_d is lower because of the higher T_{rg}. There is therefore less scope for further reduction of O_d, and the steady state gain from F_a to T_{cy} is positive, and we get a positive RGA.

Kurihara [12] states that the Hicks control structure is "incomplete" from a safety point of view, because it has incomplete feedback information with respect to the states T_{rg} and C_{rc}, which he states are the variables which govern FCC safety. The argument is supported by a closed loop simulation of the Hicks control structure in which the system goes unstable after the air blower saturates. The controller tunings used in the simulations are not provided, but these results are probably explained by the negative RGA found for Case K1 in Table 4.2 which means that one of the control loops must

[6]It should be noted that Kurihara ignored the presence of hydrogen in the coke. Clearly, for the mass balance in (4.18) the presence of hydrogen in the coke, represented by the parameter n, is important. The omission of hydrogen in the mass balance is repeated by Denn [3] in his presentation of the Kurihara model when he states that a ratio of CO_2 to CO of unity results in a value of $C_1 = 2$ in his equation (5.62a). The right hand side of eq. (5.62a) also need to be multiplied by $32/M_a$ to be correct.

	Case K1	Case K2
T_{rg}	966.1 K	983.2 K
T_{ro}	777.3 K	783.7 K
T_{cy}	987.3 K	993.1 K
C_{rc}	5.039×10^{-3}	4.409×10^{-3}
$G_H(0)$	$\begin{pmatrix} 2.907 & 36.72 \\ -5.562 & -45.73 \end{pmatrix}$	$\begin{pmatrix} 0.695 & 12.72 \\ -1.125 & 5.306 \end{pmatrix}$
RGA(0)	-1.87	0.205
RHP zeros [rad/min]		
Multivariable	–	–
In elements	$\begin{pmatrix} - & - \\ - & 0.025 \end{pmatrix}$	$\begin{pmatrix} - & - \\ - & - \end{pmatrix}$

Table 4.2: Operating points assuming plug flow of air in the regenerator for the Hicks control structure in the partial combustion mode. (– denotes that no RHP-zero is present at frequencies below 100 [rad/min]).

be unstable for the overall closed loop system to be stable [5]. In Kurihara's simulation example it must therefore be the loop F_s-T_{ro} which is unstable, and the system therefore becomes unstable when the other input (F_a, the air blower) saturates. In our opinion the assumption of plug flow of air in the regenerator used by Kurihara is naive, as strong backmixing occurs in the regenerator. These results do however demonstrate the validity of Lee and Weekmans [15] claim that the control structure selection for the FCC is sensitive to the model structure used.

Model Reduction.

All the calculations above were based on a five state model. However, the dynamics of the oxygen in the regenerator and the temperature and coke concentration in the stripper are much faster than the dynamics of the temperature and coke concentration in the regenerator. We have found that the FCC model can be reduced to only two states by setting $\frac{dO_d}{dt}$, $\frac{dT_{st}}{dt}$ and $\frac{dC_{st}}{dt}$ equal to zero. The error introduced by this model reduction is minor, the most important effect being that the RHP transmission zero for the Kurihara control structure appears at a slightly higher frequency for the two state model than for the five state model. The zeros for the individual model elements are also affected, but we have found no instance where this will affect the conclusions for control structure selection. To provide the reader with a simple model, a state space realization of the two state model for Case P1 in Table 4.1 is given in Appendix 2.

4.7.4 Sensitivity to Input Uncertainty

In this paper, only decentralized controllers are considered. Decentralized controllers are known to be relatively insensitive to input uncertainty (uncertainty in the actuators). The low RGA values found indicate that more complex controllers (e.g. decouplers) will also be insensitive to input uncertainty [28].

4.7.5 Complete Combustion Mode

The control structure selection for the FCC in the complete combustion mode has been studied using the same procedure as demonstrated above for the partial combustion mode. We will here only include the main results. Numerical results are for Case C1 in Table 4.3, unless otherwise stated.

Control Structure Alternatives

Analysis of the RGA's shows that the 3×3 control problem with $y^3 = (T_{ro}, O_{fg}, T_{rg})^T$ as controlled variables is not well suited for decentralized control. In the following we will therefore consider the 2×2 control problem. We use $u = (F_s, F_a)^T$ as manipulated variables, and consider the remaining independent variables k_c, T_f, F_f and T_a as disturbances. The following three choices of controlled variables will be considered:

1. $y_{CC} = (T_{ro}, O_{fg})^T$. We term this the *conventional control structure for the complete combustion mode.*

2. $y_{KC} = (T_{rg}, O_{fg})^T$. We will call this the *Kurihara control structure for the complete combustion mode*, although Kurihara [12] only considered the partial combustion mode.

3. $y_R = (T_{ro}, T_{rg})^T$. We will consider the riser-regenerator control structure also for the complete combustion mode.

Scaling of Variables

We use the same scalings as for the partial combustion mode. For the oxygen concentration in the flue gas, O_{fg}, the transfer functions are scaled such that an offset of magnitude 1 corresponds to 0.1 mole%.

RHP Transmission Zeros

1) Conventional Control Structure, $y_{CC} = (T_{ro}, O_{fg})^T$
For this choice of controlled variables we obtain no RHP transmission zero.

2) Kurihara Control Structure, $y_{KC} = (T_{rg}, O_{fg})^T$
For this choice of controlled variables we obtain a RHP transmission zero at the frequency 0.40 [rad/min].

3) Riser-regenerator Control Structure , $y_R = (T_{ro}, T_{rg})^T$
A RHP transmission zero is found at the frequency 0.013 [rad/min].

We will therefore only consider the conventional control structure in the following.

Pairing of Controlled and Manipulated Variables

The RGA for the pairing T_{ro}-F_s, O_{fg}-F_a is positive and quite small (2-3) at steady state for almost all operating points investigated.

The only exception is found in a region of high regenerator temperatures and low concentration of oxygen in the flue gas. In this region the RGA at low frequencies is negative and larger (ca. -9), whereas the RGA approaches 1 at high frequencies also in this region. This operating region should be avoided, as the low oxygen concentration in the flue gas indicates that the oxygen concentration in the dense bed is insufficient to convert all CO to CO_2 within the dense bed, and afterburning may therefore result. This problem is discussed further below.

The PRGA (not shown) also indicates that the pairing T_{ro}-F_s, O_d-F_a is preferable.

Disturbances

We find from the CLDG's (not shown) that a disturbance in F_f will have some high frequency effect on T_{ro}, the other disturbances can be rejected provided both loops have bandwidths of around 1 [rad/min]. At low frequencies, somewhat higher loop gains are required in the complete combustion mode than for the Hicks control structure in the partial combustion mode.

	Case C1	Case C2
T_{rg}	998.4 K	1012.3 K
T_{ro}	788.0 K	797.3 K
O_d	1.038×10^{-2}	3.098×10^{-3}
C_{rc}	9.645×10^{-3}	2.527×10^{-3}
$G_{CC}(0)$	$\begin{pmatrix} 29.98 & -11.23 \\ -0.001875 & 0.01210 \end{pmatrix}$	$\begin{pmatrix} -11.23 & 84.18 \\ 0.004131 & -0.02798 \end{pmatrix}$
RGA(0)	2.38	-9.3
RHP zeros [rad/min]		
Multivariable	–	0.014
In elements	$\begin{pmatrix} - & 0.018 \\ - & - \end{pmatrix}$	$\begin{pmatrix} - & 0.0007 \\ 1 & 0.08 \end{pmatrix}$

Table 4.3: Operating points used for the complete combustion mode, with control-related data for the conventional control structure (– denotes that no RHP-zero is present at frequencies below 100 [rad/min]).

Effect of Changes in the Operating Point

Although a number of operating points ("cases") have been studied, our findings can be illustrated by the results in Table 4.3.

Case C1 in Table 4.3 shows a typical operating point. Case C2 shows an unstable operating point, found in a region of high regenerator temperatures and low concentration of oxygen in the flue gas. We concluded above that this operating region should be avoided. The transfer function matrix G_C has a RHP pole at 7×10^{-4} [rad/min]. At this operating point there is also a pair of complex RHP transmission zeros in G_C at a frequency of 0.14 [rad/min]. The system is easily stabilized, e. g. by feedback from T_{ro} to F_s, but fast control of both T_{ro} and O_{fg} will not be possible in this region. However, the drift into the unstable region is slow and a well designed control system should easily avoid this region.

4.8 Conclusion on FCC Controllability Analysis

Partial Combustion Mode.
A favorable selection of controlled variables is critical for good control of the FCC process. Both the conventional control structure, which has T_{ro} and ΔT_{rg} as controlled variables, and the Kurihara control structure, which has T_{rg} and ΔT_{rg} as controlled variables, have RHP transmission zeros which limit the achievable bandwidth. On the other hand, we get no RHP transmission zeros with the Hicks control structure, which has T_{ro} and T_{cy} as controlled variables, or with the riser-regenerator control structure, which has T_{ro} and T_{rg} as controlled variables. Either the Hicks control structure or the riser-regenerator control structure should therefore be chosen. The best pairing for the Hicks control structure is T_{ro}-F_s, T_{cy}-F_a, and for the riser-regenerator control structure

the best pairing is T_{ro}-F_s, T_{rg}-F_a. These pairings result in RGA's which are positive at steady state and close to 1 in the frequency range around the desired closed loop bandwidth.

The Hicks control structure has the advantage over the riser-regenerator control structure that it controls T_{cy}, which often is constrained in the partial combustion mode. With hindsight the choice of the Hicks control structure might appear obvious, but the fact that the socalled conventional control structure was predominant for many years ([19, 15]) makes it clear that it is not so obvious. To our knowledge, this paper is the first to demonstrate quantitatively how controllability of the FCC process is affected by the choice of controlled variables.

We argue that Kurihara's [12] claim that the Hicks control structure is unsatisfactory for safety reasons does not hold. Since T_{ro} is controlled in the Hicks control structure, it should also be easier for operators to accept than the Kurihara control structure. According to Lee and Weekman [15], the Kurihara control structures failure to control T_{ro} leads to resistance from the operators. The Hicks control structure corrects this shortcoming of the Kurihara control structure without having to use the somewhat more complex structure proposed in [15].

Complete Combustion Mode.
T_{ro} and O_d should be chosen as controlled variables, as this choice of controlled variables give no RHP transmission zero. In contrast, choosing T_{rg} and O_d or T_{ro} and T_{rg} as controlled variables gives RHP transmission zeros which limit the achievable bandwidth. The best pairing is T_{ro}-F_s, O_d-F_a, as this corresponds to positive steady state RGA values and to RGA values close to 1 around the desired closed loop bandwidth.

Disturbances.
Both in the partial combustion mode and in the complete combustion mode the effect of disturbances in the feed oil rate on the riser outlet temperature is most difficult to reject. Fortunately, the feed oil flowrate is usually controlled, but our results demonstrate that any deliberate change in this variable should be made slowly, or feedforward from F_f to F_s should be used.

Sensitivity to Parametric Uncertainty and Changes in the Operating Point.
Our results on measurement selection and variable pairing appear relatively insensitive to parametric uncertainty and changes in the operating point.

Effect of Model Features.
1) The FCC process can be described fairly well with a second order model.
2) A model of the afterburning in the partial combustion mode is essential for a proper selection of measurements.
3) The negative steady state RGA obtained at some operating points in the partial combustion mode when assuming plug flow of air in the regenerator means that one of the control loops must be unstable on its own (at the same operating points) in order to have a stable closed loop system. However, in our opinion the assumption of plug flow of air in the regenerator is a poor one.

Acknowledgement. The authors would like to thank Pierre Grosdidier for the many helpful comments and suggestions he has made during the preparation of this paper.

Nomenclature

FCC Models:

Nominal parameter values are given in parentheses. When different values are used for the partial and complete combustion modes, the values are given as (partial combustion mode/complete combustion mode).

C_{cat} - Catalytic coke produced in riser, mass fraction
C_{rc} - Coke on regenerated catalyst, mass fraction
C_{sc} - Coke on catalyst leaving riser, mass fraction
C_{st} - Coke on catalyst in stripper, mass fraction
c_{pa} - Heat capacity of air ($1.074kJ/kgK$)
c_{pc} - Heat capacity of catalyst ($1.005kJ/kgK$)
c_{po} - Heat capacity of oil ($3.1355kJ/kgK$)
c_{pD} - Heat capacity of steam ($1.9kJ/kgK$)
$[COR]$ - Catalyst to oil ratio on a mass basis
c_t - Factor in Eq. (4.9) ($5555K$/molefraction)
E_{cf} - Activation energy for coke formation ($41.79kJ/mole/20.00kJ/mole$)
E_{cb} - Activation energy for coke combustion assuming uniformly distributed oxygen in the regenerator ($158.59kJ/mole$).
E_f - Activation energy for the cracking of gas oil ($101.5kJ/mole$)
E_g - Activation energy for the cracking of gasoline
($112.6kJ/mole$)
E_{or} - Activation energy for coke combustion assuming plug flow of air in the regenerator ($146.4kJ/mole$).
F_a - Flowrate of air to the regenerator ($25.35\frac{kg}{s}/28.0kg/s$)
F_f - Feed oil flowrate ($40.63kg/s$)
F_{fg} - Flue gas flowrate
F_s - Flow rate of regenerated catalyst ($294kg/s$)
$\hat{F}_s$ - Flowrate of spent catalyst (assumed $= F_s$).
k_1 - Reaction rate constant for the total rate of cracking of gas oil ($9.6 \times 10^5 s^{-1}$)
k_2 - Reaction rate constant for the rate of cracking of gas oil to gasoline ($7.2 \times 10^5 s^{-1}$)
k_3 - Reaction rate constant for the rate of cracking of gasoline to light gases/carbon ($4.22 \times 10^5 s^{-1}$)
k_c - Reaction rate constant for the production of coke ($0.019s^{-1}/0.0093s^{-1}$)
k_{cb} - Reaction rate constant for coke combustion assuming uniformly distributed oxygen in the regenerator ($2.077 \times 10^8 s^{-1}$).
k_{or} - Reaction rate constant for coke combustion assuming plug flow of air in the regenerator ($58.29m^2/(sN)$)
M_a - Molecular weight of air (28.8544)
M_c - Bulk molecular weight of coke (14)
m - Factor for the dependence of the initial catalyst
activity on C_{rc} (80)
n - Number of moles of hydrogen per mole of carbon
in the coke (2)
N - Exponent for the dependence of C_{cat} on C_{rc} (0.4/0.0)

O_d - Concentration of oxygen in gas leaving regenerator dense bed, molefraction
O_{in} - Concentration of oxygen in air to regenerator (0.2136 molefraction)
R - Universal gas constant
R_{cb} - Rate of coke combustion (kg/s)
P_{rg} - Regenerator pressure ($172000N/m^2$)
P_{st} - stripper pressure
T_0 - Temperature at riser entrance
T_a - Temperature of air to the regenerator
T_{cy} - Regenerator cyclone temperature
T_{st} - Temperature in stripper
T_f - Feed oil temperature
T_{ro} - Temperature at riser outlet
$T(z)$ - Temperature at elevation z in the riser
t_c - Residence time in riser ($9.6s$)
W - Holdup of catalyst in regenerator ($176000kg$)
W_a - Holdup of air in the regenerator ($20kmol$)
W_{st} - Holdup of catalyst in stripper ($17500kg$)
W_{wg} - Wet gas compressor throughput
y_f - Mass fraction of gas oil
y_g - Mass fraction of gasoline
z - Dimensionless distance along riser
α - Catalyst deactivation constant ($0.12s^{-1}$)
α_2 - Fraction of the gas oil that cracks which cracks to gasoline, $k_2/k_1 = 0.75$.
ΔH_{cb} - Heat of combustion of coke ($kJ/kmol$)
ΔH_f - Heat of cracking ($506.2kJ/kg$)
λ - Mass flowrate of dispersion steam / mass flowrate of feed oil (0.035)
σ - Molar ratio of CO_2 to CO in the regenerator dense bed
ϕ_0 - Initial catalyst activity at riser entrance
Θ - Dimensionless temperature at position z in riser

Control Analysis:

$C(s)$ - Diagonal controller transfer function matrix
$c_i(s)$ - Controller element for output i
$d(s)$ - Vector of disturbances.
$e(s) = y(s) - r(s)$ - Vector of output errors
$G(s)$ - Process transfer function matrix
$G_{CC}(s)$ - Transfer function matrix for the conventional control structure in the complete combustion mode
$G_H(s)$ - Transfer function matrix for the Hicks control structure in the partial combustion mode
$G_d(s)$ - Disturbance transfer function matrix
$g_{ij}(s)$ - ij'th element of $G(s)$
$g_{dik}(s)$ - ik'th element of $G_d(s)$
$r(s)$ - Reference signal for outputs
$S(s)$ - Sensitivity function $S = (I + GC)^{-1}$

$u(s)$ - Vector of manipulated inputs.
$y(s)$ - vector of outputs
$\Delta(s)$ - Closed loop disturbance gain matrix
$\delta_{ik}(s)$ - ij'th element of $\Delta(s)$
$\Lambda(s)$ - Relative gain matrix
$\lambda_{ij}(s)$ - ij'th element of $\Lambda(s)$
ω - Frequency
ω_B - Closed loop bandwidth

References

[1] Bristol, E. H. (1966). On a new Measure of Interactions for Multivariable Process Control. *IEEE Trans. Automat. Control*, AC-11, 133-134.

[2] Bristol, E. H. (1978). Recent Results on Interactions in Multivariable Process Control. Paper at *AIChE Annual Meeting*, Chicago, IL.

[3] Denn, M. M. (1986). *Process Modeling.* John Wiley & Sons, Inc., New York.

[4] Errazu, A. F., DeLasa, H. I. and Sarti, F. (1979). A Fluidized Bed Catalytic Cracking Regenerator Model, Grid Effects. *Can. J. Chem. Eng.*, 57, 191-197.

[5] Grosdidier, P., Morari, M. and Holt, B. R. (1985). Closed Loop Properties from Steady State Gain Information. *Ind. Eng. Chem. Fundam.*, 24, 221-235.

[6] Grosdidier, P. (1990). Personal communication.

[7] Grosdidier, P. (1990). Analysis of Interaction Direction with the Singular Value Decomposition. *Computers chem. Engng.*, 6, 687-689.

[8] Grosdidier, P., Mason, A., Aitolahti, A., Heinonen, P. and Vanhamäki, V. (1992). FCC Unit Reactor-Regenerator Control. Submitted to *Computers chem. Engng.*, preprint.

[9] Hicks, R. C., Worrell, G. R. and Durney, R. J. (1966). Atlantic Seeks Improved Control; Studies Analog-Digital Models. *The Oil and Gas Journal*, Jan. 24, 97-105.

[10] Hovd, M. and Skogestad, S. (1992). Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. To appear in *Automatica.*

[11] Hovd, M. and Skogestad S. (1992) Controllability Analysis for Unstable Processes. Presented at *IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

[12] Kurihara, H. (1967). Optimal Control of Fluid Catalytic Cracking Processes. *Ph.D. Thesis* MIT.

[13] Krishna, A. S. and Parkin, E. S. (1985). Modeling the Regenerator in Commercial Catalytic Cracking Units. *Chem. Eng. Prog.*, 81, 4, 57-62.

[14] Lee, E. and Groves, F. R. (1985). Mathematical Model of the Fluidized Bed Catalytic Cracking Plant. *Trans Soc. Comp. Sim.*, 2, 219-236.

[15] Lee, W. and Weekman, V. W. (1976). Advanced Control Practice in the Chemical Process Industry: A View from Industry. *AIChE Journal*, 22, 27-38.

[16] Ljungquist, D. (1990). Online Estimation in Nonlinear State Space Models with Application to Catalytic Cracking. *Dr. ing. Thesis*, Norwegian Institute of Technology.

[17] McAvoy, T. J. (1983). *Interaction Analysis*, Instrument Society of America, Research Triangle Park, USA.

[18] Morari, M. and Zafiriou, E. (1989). *Robust Process Control.* Prentice Hall, Englewood Cliffs, NY.

[19] Pohlenz, J. B., (1963). How Operational Variables Affect Fluid Catalytic Cracking. *Oil and Gas Journal*, April 1, pp. 124-143.

[20] Rheaume, L., Ritter, R. E., Blazek, J. J. and Montgomery, J. A. (1976). Two new Carbon Monoxide Oxidation Catalysts get Commercial Tests, *Oil and Gas Journal*, May 24.

[21] Rijnsdorp, J. E. (1991). *Integrated Process Control and Automation*, Elsevier, New York.

[22] Shah, Y. T., Huling, G. P., Paraskos, J. A. and McKinney, J. D. (1977). A Kinematic Model for an Adiabatic Transfer Line Catalytic Cracking Reactor. *Ind. Eng. Chem. Proc. Des. Dev.*, 16, 89-94.

[23] Shinskey, F. G. (1967). *Process Control Systems*, McGraw Hill, New York.

[24] Shinskey, F. G. (1984). *Distillation Column Control*, 2nd Edition, McGraw Hill, New York.

[25] Skogestad, S. and Hovd, M. (1990). Use of Frequency Dependent RGA for Control Structure Selection. *American Control Conference*, San Diego CA.

[26] Skogestad, S., Lundström, P. and Jacobsen, E. w. (1990). Selecting the Best Distillation Control Configuration. *AIChE J.*, **36**, pp. 753-764.

[27] Skogestad, S. and Morari, M. (1987a). Effect of Disturbance Direction on Closed Loop Performance. *Ind. Eng. Chem. Res.*, 26, 2029-2035.

[28] Skogestad, S. and Morari, M. (1987b). Implications of Large RGA Elements on Control Performance. *Ind. Eng Chem. Res.*, 26, 11, 2323-2330.

[29] Skogestad, S. and Wolff, E. A. (1992). Controllability Measures for Disturbance Rejection. *Preprints IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

[30] Weekman, V. W. and Nace, D. M. (1970). Kinetics of Catalytic Cracking Selectivity in Fixed, Moving and Fluid Bed Reactors. *AIChE Journal*, 16, 397-404.

[31] Wolff, E. A., Skogestad, S., Hovd, M. and Mathisen, K. W. (1992). A Procedure for Controllability Analysis. *Preprints IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

Appendix 1. Details of the FCC Model.

Riser Model

Material balance for gas oil:

$$\frac{dy_f}{dz} = -K_1 y_f^2 [COR]\Phi t_c \tag{4.19}$$

Material balance for gasoline:

$$\frac{dy_g}{dz} = (\alpha_2 K_1 y_f^2 - K_3 y_g)[COR]\Phi t_c \tag{4.20}$$

where

$$K_1(\Theta) = k_1 exp(\frac{-E_f}{RT_0(1+\Theta)}) \tag{4.21}$$

$$K_3(\Theta) = k_3 exp(\frac{-E_g}{RT_0(1+\Theta)}) \tag{4.22}$$

$$\Theta = (T(z) - T_0)/T_0 \tag{4.23}$$

$$\Phi = \phi_0 exp(-\alpha t_c [COR] z) \tag{4.24}$$

$$\phi_0 = 1 - mC_{rc} \tag{4.25}$$

Here $K_1 y_f^2 [COR]$ represents the kinetics for the cracking of gas oil and $K_3 y_g [COR]$ the kinetics for cracking of gasoline. Φ represents the deactivation of the catalyst caused by coke deposition, of which ϕ_0 represents the reduction in catalyst activity caused by the coke remaining on the catalyst after regeneration. t_c is the residence time in the riser, and $\alpha_2 = k_2/k_1$ is the fraction of the cracked gas oil which cracks to gasoline.

The catalyst to oil ratio $[COR]$, which was omitted in [14], is reintroduced into Φ in order to be consistent with the original model of Shah and coworkers [22]. A correlation taken from Kurihara [12] is used to estimate the amount of coke produced.

$$C_{cat} = k_c \sqrt{\frac{t_c}{C_{rc}^N} exp(\frac{-E_{cf}}{RT_{ro}})} \tag{4.26}$$

The amount of coke on the catalyst leaving the riser is thus

$$C_{sc} = C_{rc} + C_{cat} \tag{4.27}$$

The energy balance yields:

$$\frac{d\Theta}{dz} = \frac{\Delta H_f F_f}{T_0(F_s c_{ps} + F_f c_{po} + \lambda F_f c_{pd})} \frac{dy_f}{dz} \tag{4.28}$$

Regenerator Model

The regenerator is described by the following equations. Balance for coke on regenerated catalyst:

$$W\frac{d}{dt}C_{rc} = F_s(C_{st} - C_{rc}) - R_{cb} \tag{4.29}$$

Energy balance:

$$Wc_{ps}\frac{d}{dt}T_{rg} = T_{st}F_s c_{ps} + T_a F_a c_{pa} - T_{rg}(F_s c_{ps} + F_a c_{pa}) - \Delta H_{cb} R_{cb}/M_c \tag{4.30}$$

where ΔH_{cb} depends both on the temperature and σ, the ratio of CO_2 to CO produced [4]. Here n denotes the average coke composition CH_n. The concentration of oxygen in the regenerator dense bed is given by a material balance

$$W_a\frac{d}{dt}O_d = \frac{F_a}{M_a}(O_{in} - O_d) - \frac{(1+\sigma)n + 2 + 4\sigma}{4(1+\sigma)}\frac{R_{cb}}{M_c} \tag{4.31}$$

The rate of coke combustion is given by

$$R_{cb} = k_{cb} exp(\frac{-E_{cb}}{RT_{rg}}) O_d C_{rc} W \tag{4.32}$$

Stripper Model

The flowrate of stripping steam is small compared to the flowrates of catalyst and feed oil, and the effect of the steam on the heat balance of the stripper is therefore ignored. Assuming the stripping is effective, the only effect of the stripper will be to introduce a lag between the riser outlet and the catalyst return to the regenerator. This lag is modeled using an ideal mixing tank, and the balances for coke and energy yield the following equations.

$$W_{st}\frac{d}{dt}C_{st} = F_s(C_{sc} - C_{st}) \tag{4.33}$$

$$W_{st}c_{ps}\frac{d}{dt}T_{st} = F_s c_{ps}(T_{ro} - T_{st}) \tag{4.34}$$

Parameter Values

The parameter values used are given in the nomenclature section. The values used are taken from Ljungquist [16], who has slightly modified the values given by Lee and Groves. The value of c_t is taken from [12].

Complete Combustion Mode

The above model has been adjusted to describe the complete combustion mode as follows:

1. The ratio of CO_2 to CO produced in the regenerator, σ, is fixed to a high value, such that there is an excess of O_2 over CO in the gas leaving the regenerator dense bed. Then the oxygen concentration in the flue gas, O_{fg}, will be approximately equal to the oxygen concentration in the gas leaving the regenerator dense bed, O_d.

2. The air rate to the regenerator, the coke production rate (k_c) and the feed oil temperature have been adjusted in order to achieve energy balance in the desired operating region.

3. Finally the model parameters have been adjusted to obtain the same signs for the steady state gains as observed by Grosdidier [6]:

$$\begin{bmatrix} T_{ro} \\ O_d \end{bmatrix} = \begin{bmatrix} + & - \\ - & + \end{bmatrix} \begin{bmatrix} F_s \\ F_a \end{bmatrix} \tag{4.35}$$

These signs for the steady state gains should be obtained in the region of the following operating conditions [7]:
Flue gas oxygen concentration: 0.4-1.4 mole%
Regenerator dense bed temperature: 989-1009 K
In order to obtain the desired signs for the steady state gains the activation energy for coke formation, E_{cf}, has been reduced, and the exponent expressing the dependency of the coke production on the amount of coke on regenerated catalyst, N, has been set to zero[7].

Appendix 2. Linear Two State Model for the FCC.

A linear two state model for the FCC in the partial combustion mode (Case P1 in Table 4.1), obtained by linearization and model reduction of the five state nonlinear model in Appendix 1:

$$\frac{dx}{dt} = Ax + Bu + Ed \tag{4.36}$$
$$y = Cx + Du + Fd \tag{4.37}$$

$$A = \begin{bmatrix} -2.55 \times 10^{-2} & 1.51 \times 10^{-6} \\ 227 & -4.10 \times 10^{-2} \end{bmatrix} \tag{4.38}$$

$$B = \begin{bmatrix} 3.29 \times 10^{-6} & -2.60 \times 10^{-5} \\ -2.80 \times 10^{-2} & 7.80 \times 10^{-1} \end{bmatrix} \tag{4.39}$$

$$C = \begin{bmatrix} 1.32 \times 10^{3} & 0.559 \\ -4.42 \times 10^{3} & 0.538 \\ 0 & 1 \end{bmatrix} \tag{4.40}$$

[7]This seems reasonable since FCC's operating in the complete combustion mode commonly achieve very good regeneration with very small amounts of coke on regenerated catalyst [20]. For these low amounts of coke on regenerated catalyst, Eq. (4.26) would predict both a very large coke production and a very strong dependence of the coke production on the amount of coke on regenerated catalyst. This does not appear realistic, and we have therefore set $N = 0$ in our studies of the complete combustion mode of operation.

$$D = \begin{bmatrix} 0.362 & 0 \\ 0 & 0.877 \\ 0 & 0 \end{bmatrix} \tag{4.41}$$

$$E = \begin{bmatrix} 6.87 \times 10^{-7} & 0 & -7.06 \times 10^{-6} & 3.53 \times 10^{-2} \\ 2.47 \times 10^{-2} & 9.24 \times 10^{-3} & -2.54 \times 10^{-1} & 0 \end{bmatrix} \tag{4.42}$$

$$F = \begin{bmatrix} 0.246 & 0 & -0.253 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.43}$$

$$y = \begin{bmatrix} T_{ro} \\ T_{cy} \\ T_{rg} \end{bmatrix}; \quad x = \begin{bmatrix} C_{rc} \\ T_{rg} \end{bmatrix}; \quad u = \begin{bmatrix} F_s \\ F_a \end{bmatrix}; \quad d = \begin{bmatrix} T_f \\ T_a \\ F_f \\ k_c \end{bmatrix}$$

Chapter 5

Design of Robust Decentralized Controllers

Morten Hovd Sigurd Skogestad*
Chemical Engineering, University of Trondheim, NTH,
N-7034 Trondheim, Norway.

Abstract

The procedure for independent design of robust decentralized controllers proposed by Skogestad and Morari [20] is improved by requiring the controller to be a decentralized Internal Model Control (IMC) type controller. It is shown how to find bounds on the magnitude of the IMC filter time constantssuch that robust stability or performance is guaranteed. This allows the use of real perturbation blocks for modeling the uncertainty associated with the controllers. In contrast, Skogestad and Morari [20] found bounds on the sensitivity functions and complementary sensitivity functions for the individual loops, and therefore allowed a much larger class of designs, resulting in more conservative conditions.

The concept of Robust Decentralized Detunability is introduced. If a system is Robust Decentralized Detunable, any subset of the loops can be detuned independently and to an arbitrary degree without endangering robust stability. A simple test for Robust Decentralized Detunability is developed for systems controlled by a decentralized IMC controller.

The problem of sequential design of robust decentralized controllers is also addressed. It is shown how to include into the design problem for loop k a simple estimate the effect of closing subsequent loops will have on loop k and the loops that have already been closed.

*To whom correspondence should be addressed. FAX: +47-7-594080, e-mail: skoge@kjemi.unit.no

5.1 Introduction

Decentralized control remains popular in the chemical process industry, despite developments of advanced controller synthesis procedures leading to full multivariable controllers. Some of the reasons for the continued popularity of decentralized control are:

1. Decentralized controllers are easy to implement.
2. They are easy for operators to understand.
3. The operators can be allowed to retune the controllers to take account of changing process conditions (as a result of 2 above).
4. Some measurements or manipulated variables may fail. Tolerance of such failures are more easily incorporated into the design of decentralized controllers than full controllers.
5. The control system can be brought gradually into service during process startup and taken gradually out of service during shutdown.

The design of a decentralized control system consists of two main steps:

a) Control structure selection, that is, choosing manipulated inputs and controlled outputs, and pairing inputs and outputs.

b) Design of each single-input single-output (SISO) controller.

In this paper we will consider b), and assume that a) has already been done (e.g. by using the tools in [8, 9]). Standard controller synthesis algorithms (e.g. H_2 or H_∞ synthesis) lead to multivariable controllers, and cannot handle requirements for controllers with a specified structure. Instead, some practical approaches to the design of decentralized controllers have evolved:

- Parameter optimization.
- Sequential design [3, 14, 16].
- Independent design [20].

We discuss all these three approaches to the design of decentralized controllers, with reference to the potential advantages of decentralized control listed above. Parameter optimization will be considered only briefly, whereas we discuss and illustrate the

unique problems associated with sequential design in more detail. New results on independent design are presented which represent improvements over the existing design procedure. Throughout this work we will use the structured singular value (see below) as the measure of control quality.

A system is said to have robust stability if it is stable regardless of whatever uncertainty is contained within the system. Because of items 4 and 5 above, we would like the system to remain stable if any subset of the control loops are out of service, or if the individual controllers have been detuned. Furthermore, we would like this stability to be a *robust* property, We define such systems to be Robust Decentralized Detunable:

Definition 1 *A closed loop system is said to be Robust Decentralized Detunable if each controller element can be detuned independently by an arbitrary amount without endangering robust stability.*

Decentralized detunability for a given controller should not be confused with decentralized integral controllability (DIC), which is a property of the plant only. DIC implies that there for a given plant *exists* a decentralized controller with integral action in all channels that is decentralized detunable.

5.2 Notation

In this paper, $G(s)$ will denote the plant, which is assumed to be of dimension $n \times n$. The matrix consisting of the diagonal elements of $G(s)$ is denoted $\tilde{G}(s)$, and $g_{ij}(s)$ is the ij'th element of $G(s)$. The reference signal (setpoint) is denoted r, manipulated inputs are denoted u and outputs are denoted y. If disturbances are present, $G_d(s)$ denotes the (open loop) transfer function from disturbances d to outputs y. Throughout this work, all controllers are assumed to be completely decentralized. The decentralized conventional feedback controller is denoted $C(s)$, with i'th diagonal element $c_i(s)$ (Fig. 5.1a). Likewise, the decentralized IMC controller is denoted Q, with i'th diagonal element $q_i(s)$ (Fig. 5.1b). The controllers $C(s)$ and $Q(s)$ are related by

$$C(s) = Q(s)(I - \tilde{G}(s)Q(s))^{-1} \tag{5.1}$$

The sensitivity function is $S(s) = (I + G(s)C(s))^{-1}$ and the $H(s) = I - S(s) = G(s)C(s)(I + G(s)C(s))^{-1}$ is the complementary sensitivity function. The sensitivity functions and complementary sensitivity functions for the individual loops are collected in the diagonal matrices $\tilde{S}(s) = (I - \tilde{G}(s)C(s))^{-1}$ and $\tilde{H}(s) = \tilde{G}(s)C(s)(I - \tilde{G}(s)C(s))^{-1}$. Note that the diagonal elements of $\tilde{S}(s)$ and $\tilde{H}(s)$ do not equal the diagonal elements of $S(s)$ and $H(s)$, respectively. The i'th element on the diagonal of $\tilde{S}$ and $\tilde{H}$ are $\tilde{s}_i$ and $\tilde{h}_i$, respectively.

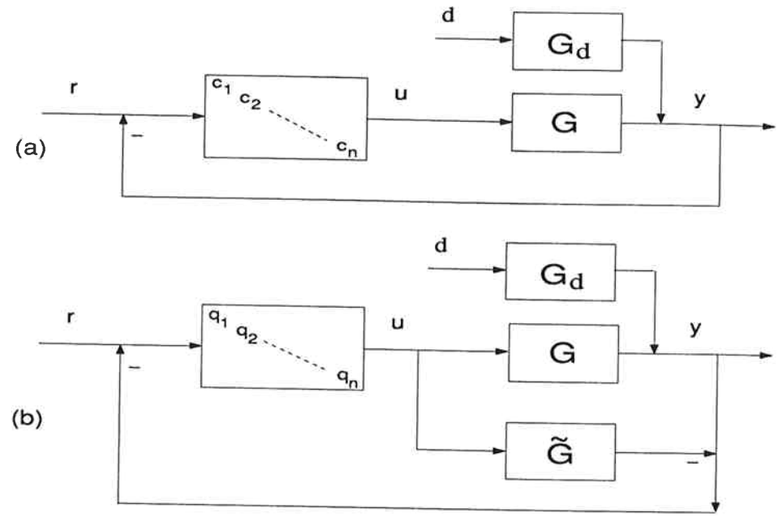


Figure 5.1: Block diagram of feedback systems. (a) Conventional decentralized controller. (b) Decentralized IMC controller.

5.3 Robust Control and the Structured Singular Value

Since no model is a perfect representation of the system, the control system stability and performance should be little affected by the uncertainties of the model. In this paper we use the structured singular value, μ, introduced by Doyle [5], as a measure of the robustness of feedback systems. Within the μ framework, one accepts that it is the impossible to find a perfect model, and instead require information about the structure, location and estimates of the magnitude of the model uncertainties.

In Fig. 5.2 we have drawn an example of a feedback system with uncertainty in the inputs and outputs[1], represented by the perturbation blocks Δ_I and Δ_O, respectively. Note that the individual perturbations can be restricted to have a certain structure.

[1]Many other types of uncertainties possible, see [5] for details on how to represent different uncertainties with perturbation blocks.

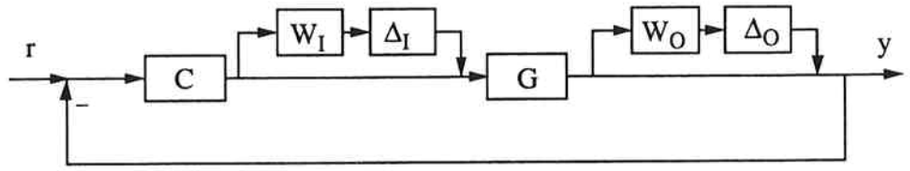


Figure 5.2: Block diagram for feedback system with uncertainty in the inputs and outputs.

For instance, as individual inputs and outputs usually do not affect each other, both Δ_I and Δ_O are assumed to be diagonal. The weights W_I and W_O are frequency-dependent and normalize the maximum magnitude of Δ_I and Δ_O to unity.

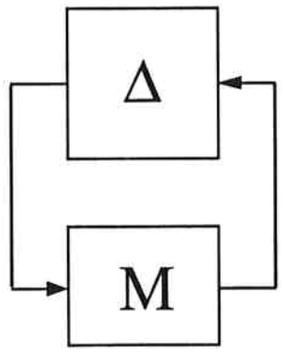


Figure 5.3: Feedback system rearranged into a perturbation block Δ and an interconnection matrix M.

Any block diagram with uncertainties represented by perturbation blocks can be rearranged into the $M - \Delta$ structure of Fig. 5.3, if external inputs and outputs are neglected. In Fig. 5.3, Δ is a block diagonal matrix with the perturbation blocks of the original block diagram on the diagonal, and M contains all the other blocks in the block diagram (plant, controller, weights). For the specific case in Fig. 5.2, we have that

$$\Delta = \text{diag}\{\Delta_I, \Delta_O\}; \qquad M = \begin{bmatrix} -W_I CG(I+CG)^{-1} & -W_I C(I+GC)^{-1} \\ W_O G(I+CG)^{-1} & -W_O GC(I+GC)^{-1} \end{bmatrix}$$

Provided M is stable (the system has Nominal Stability, NS) and Δ is norm bounded and stable (stable perturbation blocks), it follows from the Nyquist stability criterion [5] that the overall system is stable provided $\det(I - M\Delta) \neq 0 \quad \forall\Delta, \forall\omega$. In this case the system is said to have Robust Stability (RS). The structured singular value is

defined such that

$$\mu_\Delta^{-1} = \min_\delta\{\delta | \det(I - M\Delta) = 0 \quad \text{for some } \Delta, \bar{\sigma}(\Delta) \leq \delta\} \tag{5.2}$$

If weights are used to normalize the maximum value of the largest singular value of Δ to unity ($\bar{\sigma}(\Delta) = 1$) at all frequencies, like in Fig. 5.2, the system will remain stable for any allowable perturbation Δ provided $\mu_\Delta(M) < 1$.

Doyle [5] showed that performance can be analyzed in the μ framework by considering an equivalent stability problem of larger dimension. In this paper we use a performance specification of the type $\bar{\sigma}(W_P S_p) \leq 1 \quad \forall\omega$ where S_p is the worst sensitivity function (S) made possible by the perturbation blocks. This performance specification can be incorporated in the μ framework by closing the loop from outputs to output disturbances with the performance weight W_P and a *full* perturbation block Δ_P. If $\mu_\Delta(M) \leq 1 \quad \forall\omega$ (after normalizing the magnitude of the perturbation blocks) and M is stable for the corresponding $M - \Delta$ structure of increased dimension (in our specific example, $\Delta = \text{diag}\{\Delta_I, \Delta_O, \Delta_P\}$), the system is said to have Robust Performance (RP), as the performance specification is fulfilled for all the possible model uncertainties.

To simplify notation, we will use "$\mu(M)$" in the meaning $\sup_\omega \mu_\Delta(M)$. Doyle and Chu [6] proposed an algorithm for the synthesis of controllers which minimizes μ, known as $D - K$ iteration. However, $D - K$ iteration results in full controllers, and the problem of synthesizing μ-optimal decentralized controllers has not been solved.

5.4 Parameter Optimization

When using parameter optimization, an *à priori* parametrization of the controller and the chosen measure of control quality (in our case μ) is optimized with respect to the controller parameters, using some optimization routine. Controller design using parameter optimization is easily formulated in a computer program and often gives satisfactory designs (e.g. [21]). However, the optimization is not necessarily convex, and problems with local minima may be encountered. Another problem is that, since all the loops are assumed to be in service at each step in the optimization, advantages 4 and 5 in the introduction may not be achieved. One must therefore check specifically whether these advantages are achieved after the optimization is finished. More importantly, the parameter optimization approach gives no guidelines for how to achieve advantages 4 and 5 if analysis of a proposed controller shows that they are not achieved.

5.5 Independent Design

Independent design of robust decentralized controllers was introduced by Skogestad and Morari [20]. It is based on Theorem 1 in [19], which we state here:

Theorem 9 *Let the μ interconnection matrix M be written as a lower Linear Fractional Transformation (LFT) of the transfer function matrix T*

$$M = F_l(N,T) = N_{11} + N_{12}T(I - N_{22}T)^{-1}N_{21} \tag{5.3}$$

and let k be a given constant. Assume $\mu_\Delta(N_{11}) < 1$ and $det(I - N_{22}T) \neq 0$ then

$$\mu_\Delta(M) \leq 1 \tag{5.4}$$

if

$$\bar{\sigma}(T) \leq c_T \tag{5.5}$$

where c_T solves

$$\mu_{\tilde{\Delta}} \begin{bmatrix} N_{11} & N_{12} \\ c_T N_{21} & c_T N_{22} \end{bmatrix} = 1 \tag{5.6}$$

and $\tilde{\Delta} = \text{diag}\{\Delta, T\}$

Proof: See [19].

The condition $\mu_\Delta(M)$ is typically the RP condition we want to satisfy, and T is some important transfer function which depends on the controller. Skogestad and Morari [20] uses Thm. 9 to find bounds on the sensitivity function and complementary sensitivity functions for the individual loops (i.e. $T = \tilde{S}$ and $T = \tilde{H}$ are used). The bounds on $\tilde{S}$ and $\tilde{H}$ can be combined over different frequency ranges. Thus, if either the bound on $\tilde{S}$ or the bound on $\tilde{H}$ is fulfilled for all loops at all frequencies, then $\mu_\Delta(M) \leq 1$ is achieved.

The rationale behind Thm. 9 is to treat the transfer functions (T) as a "class of possible designs" (i.e. as uncertainty), and finds bounds on the magnitude of this fictitious uncertainty which guarantees that $\mu_\Delta(M) \leq 1$. One is faced with finding controllers such that the bounds on the transfer functions are fulfilled. It is therefore important for the success of independent design that T introduces as little additional uncertainty as possible. It turns out that parametrizing the class of possible designs as $T = \tilde{S}$ and $T = \tilde{H}$ are not ideal for this purpose.

5.5.1 Example 1

Consider Example 1 in Chiu and Arkun [3]:

$$G(s) = \begin{bmatrix} \frac{1.66}{39s+1} & \frac{-1.74e^{-2s}}{4.4s+1} \\ \frac{0.34e^{-s}}{8.9s+1} & \frac{1.4e^{-s}}{3.8s+1} \end{bmatrix} \tag{5.7}$$

There is independent input uncertainty with input uncertainty weight $W_I(s) = 0.07I_2$, and the performance requirement is given by the performance weight $W_p(s) = 0.25\frac{7s+1}{7s}I_2$

Chiu and Arkun [3] attempted independent design for this example, using $T = \tilde{S}$ and $T = \tilde{H}$, but were unable to find a controller which fulfilled the resulting bounds. In [3] it was therefore claimed that independent design can not be performed for this example. We will however demonstrate below that independent design can be performed for this example, by parametrizing the class of possible designs within the framework of Internal Model Control.

5.5.2 Independent Design with Decentralized IMC Controllers

We use the Internal Model Control (IMC) technique [7] to parametrize the individual controller elements, and select T not as a transfer function, but rather as a parametrization of the tuning constant ϵ_i in the IMC controller. Our approach is similar to that of Lee and Morari [10], but we use ϵ_i as the parameter rather than the filter f_i. The relationship between the elements q_i of the IMC controller and the elements c_i of the conventional controller is given by

$$c_i = q_i(1 - g_{ii}c_i)^{-1} \tag{5.8}$$

In the IMC design procedure [15], q_i has the form

$$q_i = \hat{g}_{ii}^{-1} f_i \tag{5.9}$$

where $\hat{g}_{ii}$ is the minimum phase part of g_{ii}, and f_i is a low pass filter used to make q_i realizable and to detune the system for robustness. In order to simplify the exposition, we will assume the plant G to be open loop stable, and use a low pass filter of the form

$$f_i = \frac{1}{(\epsilon_i s + 1)^{n_f}} \tag{5.10}$$

That is, the f_i is taken to be a low pass filter of order n_f, consisting of n_f *identical* first order low pass filters in series. For details on IMC design, and on filter form for unstable systems, the reader is referred to Morari and Zafiriou [15].

Choice of T for Independent Design. After fixing n_f, the only thing which remains uncertain in the IMC technique is the value of ϵ_i. To fulfill performance requirements at low frequencies, the closed loop system must be sufficiently fast, which means that the filter time constant ϵ must be smaller than a certain value. On the other hand, the closed loop system must be sufficiently detuned to avoid robustness problems at higher frequencies, thus requiring ϵ to be larger than a certain value, meaning that $1/\epsilon$ must be smaller than some value. We will therefore use Thm. 9 to find bounds on ϵ and $e_i \stackrel{\text{def}}{=} 1/\epsilon$ which can be combined over different frequency ranges. Since we are using a specific control structure the class of possible designs is much smaller than if we use Thm. 9 to find bounds on $\tilde{S}$ and $\tilde{H}$. Bounds on $\tilde{S}$ and $\tilde{H}$ are therefore potentially much more conservative.

To derive conditions on ϵ_i and e_i that guarantee $\mu_\Delta(M) \leq 1$ we will proceed as follows: First we parametrize the μ interconnection matrix M as an LFT of the IMC filter F and then as an LFT of the "uncertainty" in the filter time constant. We refer the readers to [19] or [15] for details on how to find the LFT's needed in Fig. 5.4. Below we will only elaborate on how to express f_i as an LFT of the "uncertainty" associated with ϵ_i or e_i. We then solve Eq. (5.6) at each frequency point to find the desired bound. Note that it is sufficient at each frequency to satisfy the bound either for ϵ_i or for $e_i = 1/\epsilon_i$, but we must of course use the same bound for ϵ_i (and e_i) at all frequencies. Note that although $\hat{g}_{ii}^{-1}$ in Eq. (5.9) will normally not be realizable, its frequency response is easily calculated. Also note that since we work with the frequency response, we will have to check à posteriori for the (internal) stability of the μ interconnection matrix.

First Order Low Pass Filters. Consider first the case $n_f = 1$. We then have $f_i = 1/(\epsilon_i s + 1)$. The objective is to find the allowable ranges for ϵ_i and $e_i = 1/\epsilon_i$ that at each frequency guarantee $\mu(M) \leq 1$. Since we do not allow negative values for ϵ_i we should not write $|\epsilon_i| \leq c_\epsilon$. Instead write

$$\epsilon_i = \frac{\epsilon^*}{2}(1 + \Delta_\epsilon) \qquad |\Delta_\epsilon| \leq 1 \tag{5.11}$$

$$e_i = \frac{e^*}{2}(1 + \Delta_e) \qquad |\Delta_e| \leq 1 \tag{5.12}$$

and fix $c_\epsilon = 1$ and $c_e = 1$ in Eq. (5.6). The μ interconnection matrix in Eq. (5.6) then depends only on ϵ^* or e^* (depending on whether we try to find bounds on ϵ_i or e_i). Note that all quantities, including Δ_ϵ and Δ_e are real. In order to use Thm. 9 we now

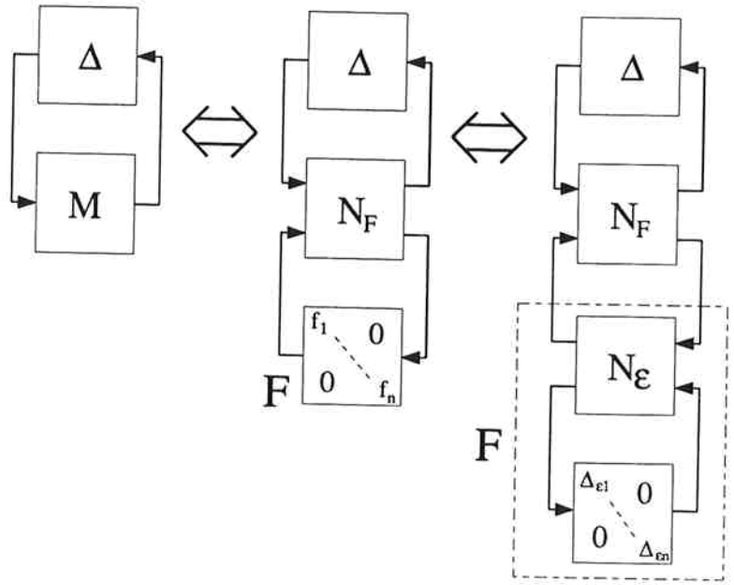


Figure 5.4: The interconnection matrix M expressed as an LFT of the IMC filter F and as an LFT of the "uncertainty" associated with the filter time constants.

need to write f_i as an LFT of Δ_ϵ and Δ_e. For $T = \Delta_\epsilon$ we have

$$f_i = \frac{1}{\frac{\epsilon^*}{2}(1 + \Delta_\epsilon)s + 1} \tag{5.13}$$

which may be written as an LFT, $f_i = \mathrm{F}_l(N_{\epsilon_i}, \Delta_\epsilon)$, with

$$N_{\epsilon_i} = \frac{1}{\frac{\epsilon^*}{2}s + 1}\begin{bmatrix} 1 & -1 \\ \frac{\epsilon^*}{2}s & \frac{-\epsilon^*}{2}s \end{bmatrix} \tag{5.14}$$

Similarly, for $T = \Delta_e$, we have

$$N_{e_i} = \frac{1}{\frac{e^*}{2} + s}\begin{bmatrix} \frac{e^*}{2} & \frac{e^*}{2} \\ s & \frac{-e^*}{2} \end{bmatrix} \tag{5.15}$$

This shows how to express an individual filter element f_i as an LFT of the real "uncertainty" in the filter time constant in that filter element. The LFT for the overall IMC filter $F = \mathrm{diag}\{f_i\}$ is then just a simple diagonal augmentation of the corresponding blocks of the LFT for the individual filter elements. For example, let N_{11}^i denote the N_{11} block for the LFT of element i. The block N_{11} for the LFT of the overall IMC filter will then be given by $N_{11} = \mathrm{diag}\{N_{11}^i\}$.

We use $|\Delta_\epsilon| \leq 1$ and $|\Delta_e| \leq 1$ and correspondingly fix $c_\epsilon = 1$ and $c_e = 1$ in Eq. (5.6). The required bounds for ϵ_i and e_i are then found by iterating on the value of ϵ^* or e^*

(as appropriate) until $\mu_{\tilde{\Delta}} = 1$ (the matrix N in Eq. (5.6) will depend on ϵ^* when we iterate on ϵ^*, and depend on e^* when we iterate on e^*). Denote the values of ϵ^* and e^* which give $\mu_{\tilde{Delta}} = 1$ ϵ_s^* and e_s^*, respectively. For a fixed frequency, we are then guaranteed that $\mu_\Delta(M) \leq 1$ provided

$$\epsilon_i \leq \epsilon_s^* \, \forall i; \quad \text{or} \tag{5.16}$$

$$e_i \leq e_s^* \, \forall i \iff \epsilon_i \leq \frac{1}{e_s^*} \, \forall i \tag{5.17}$$

Note that although the ϵ_i's are independent, we get the same bound for all ϵ_i.

The iterations are performed using only positive real values for ϵ^* and e^*. Note that $\mu_{\tilde{\Delta}}$ in Eq. (5.6) is non-decreasing with increasing values of ϵ^* or e^*. This follows from the fact that we use $|\Delta_\epsilon| \leq 1$ and $|\Delta_e| \leq 1$, therefore the set of possible values for $\epsilon_i = \frac{\epsilon^*}{2}(1 + \Delta_\epsilon)$ for any fixed value of ϵ^* contains all the possible values for ϵ_i for any smaller ϵ^* (and similarly for e_i and e^*). A very simple iteration scheme, e.g. bisection, can therefore be used.

Higher Order Low Pass Filters. In IMC design, one will often use filters of order higher than one. We therefore need to be able to express the higher order filters as LFT's of Δ_ϵ and Δ_e. For this we can use the rules for series interconnection of linear dynamical systems. First note that $G(s) = C(sI - A)^{-1}B + D$ may be written as an LFT of $\frac{1}{s}I$, with

$$N_{11} = D; \quad N_{12} = C; \quad N_{21} = B; \quad N_{22} = A$$

The formulae for series interconnection $G = G_1 G_2$ of dynamical systems $G_1(s) = C_1(sI - A_1)^{-1}B_1 + D_1$ and $G_2(s) = C_2(sI - A_2)^{-1}B_2 + D_2$ are (e.g. [13]):

$$\begin{aligned} A &= \begin{bmatrix} A_1 & 0 \\ B_2C_1 & A_2 \end{bmatrix} \\ B &= \begin{bmatrix} B_1 \\ B_2D_1 \end{bmatrix} \\ C &= \begin{bmatrix} D_2C_1 & C_2 \end{bmatrix} \\ D &= D_2D_1 \end{aligned}$$

The formulae for series interconnection of dynamical systems can be used directly to express an n_f'th order low pass filter as LFT's of diag$\{\Delta_{\epsilon 1}, \cdots, \Delta_{\epsilon n_f}\}$ and diag$\{\Delta_{e1}, \cdots, \Delta_{en_f}\}$. As we will normally use the same time constant for all first order factors of the n_f'th order filter, we will have $\Delta_{\epsilon 1} = \Delta_{\epsilon 2} = \cdots = \Delta_{\epsilon n_f}$ and $\Delta_{e1} = \Delta_{e2} = \cdots = \Delta_{en_f}$, and we have repeated scalar, real "uncertainty" associated with the filter in each IMC controller element.

Note that although we have repeated scalar "uncertainties" for each individual filter element, the filter time constants may differ in different filter elements, and the "uncertainties" in different filter elements are therefore independent. For a plant of dimension $n \times n$ we therefore end up with n repeated scalar uncertainty blocks for the IMC filter, each of these blocks being repeated n_f times[2].

5.5.3 Independent Design Procedure

With the preliminaries above, we can now propose an independent design algorithm:

1. Find the matrices N_ϵ, expressing the μ interconnection matrix M as an LFT of Δ_ϵ, and the matrix N_e, expressing M as an LFT of Δ_e. N_ϵ will depend on the value of ϵ^*, and N_e will depend on the value of e^*, and we must therefore recompute N_ϵ and N_e for every new value of ϵ^* and e^*, respectively.

2. We get

   $$\mu(M) \leq 1$$

   if

   $$0 \leq \epsilon_i \leq e_s^* \quad \forall i \tag{5.18}$$

   where ϵ_s^* solves

   $$\mu(N_\epsilon) = 1 \tag{5.19}$$

   Similarly, let e_s^* solve $\mu(N_e) = 1$, giving the bound

   $$0 \leq e_i \leq e_s^* \quad \forall i \quad \Longleftrightarrow 1/e_s^* \leq \epsilon_i \quad \forall i \tag{5.20}$$

3. From 2 and Thm. 9 we know that $\mu(M) < 1$ for the range of values of ϵ which at *all* frequencies is *either* within the range of values in Eq. (5.18) *or* within the range of values in Eq. (5.20).

4. Choose a value of ϵ within the range of values found in point 3, and verify the stability of M for this choice of ϵ. [3]

[2] One may use low pass filter of different orders in the different filter elements, in which case the value of n_f will differ for different filter elements.

[3] For any value of ϵ within the range found in point 3, the map under the Nyquist D-contour of $\det(I - M\Delta)$ will encircle the origin the same number of times. Thus, if M is found to be unstable in Step 4, it is "robustly unstable".

If we are successful in Steps 3 and 4, the controller design is completed. Since we have both real and complex perturbations, Step 2 requires μ calculations for mixed real and complex perturbations [24], which is still a research topic. However, the existing μ software has proved to be acceptable in many cases.

5.5.4 Examples

Example 1 (continued)

Consider again Example 1 studied above. For this problem we choose a second order low pass filter in each element of the decentralized IMC controller. Since we have a 2×2 system, this will add two real, repeated scalar perturbations, each repeated twice. Solving Eq. (5.6), we obtain the results in Fig. 5.5. We see that values of ϵ between

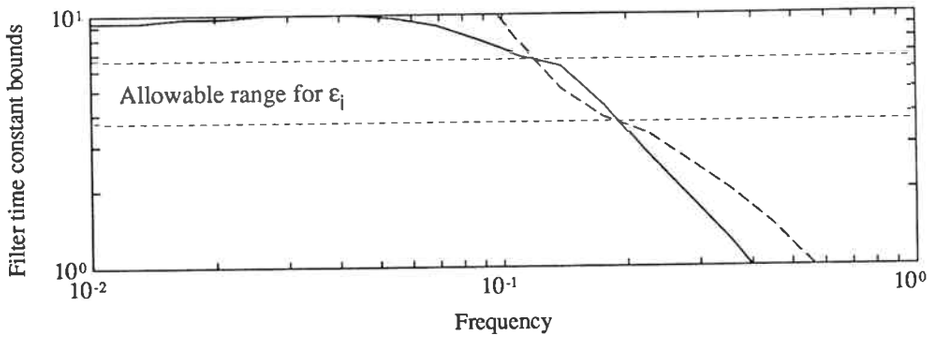


Figure 5.5: Filter time constant bounds for Example 1. Solid: ϵ_s^* (upper bound). Dashed: $1/\epsilon_s^*$ (lower bound).

3.7 and 6.6 are at all frequencies either below the upper bound or above the lower bound. Choosing $\epsilon = 5$ for both loops, it is easily verified that the system is nominally (internally) stable. We have thus completed an independent design for this example.

Example 2

Here we consider Example 2 in [3].

$$G(s) = \begin{bmatrix} \frac{0.66}{6.7s+1} & \frac{-0.61}{8.4s+1} & \frac{-0.005}{9.06s+1} \\ \frac{1.11}{3.25s+1} & \frac{-2.36}{5s+1} & \frac{-0.01}{7.09s+1} \\ \frac{34.7}{8.15s+1} & \frac{46.2}{10.9s+1} & \frac{0.87(11.61s+1)}{(3.89s+1)(18.8s+1)} \end{bmatrix} \tag{5.21}$$

In this example only robust stability is considered, with independent, multiplicative input uncertainty with uncertainty weight $W_I(s) = 0.13\frac{5s+1}{0.25s+1}$. In [3] it is found that

independent design using Thm. 9 with $T = \tilde{H}$ and $T = \tilde{S}$ cannot be used to design a robust controller for this example. Since the process is stable and only multiplicative uncertainty is considered, this clearly illustrates the shortcomings of that method.

As for Example 1, a second order low pass filter is used in each diagonal element of the IMC controller. This will add three real, repeated scalar perturbations, each repeated twice. From Step 2 in the independent design procedure we obtain the results

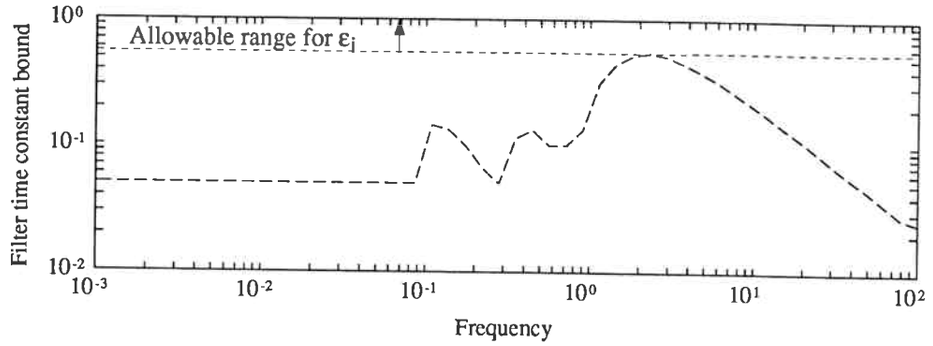


Figure 5.6: Lower bound on filter time constant $(1/\epsilon_s^*)$ for Example 2.

in Fig. 5.6. From Fig. 5.6 we see that any value of ϵ larger than 0.55 will be acceptable. Choosing $\epsilon = 1$ for all loops, we find that the system is stable. We thus find that the system will be robustly stable for any value of $\epsilon_i > 0.55$. In general we want ϵ to be small for a faster nominal response.

For both Example 1 and Example 2, Chiu and Arkun [3] were unable to perform an independent design, using the procedure of Skogestad and Morari [20]. This demonstrates the importance of introducing as little conservatism as possible in the description of the uncertainty associated with the controllers when performing an independent design.

Robust Decentralized Detunability in the IMC Framework

In the IMC framework, controllers are detuned by increasing the filter time constants. We have thus found for Example 2 above that the loops can be detuned independently of each other, without endangering robust stability, provided all loops have $\epsilon_i > 0.55$. Thus the closed loop system in Example 2 with $\epsilon_i > 0.55$ in all loops is found to be robust decentralized detunable according to Definition 1. After removing the performance requirement from Example 1 and redoing the calculations for robust stability, we find that it is robust decentralized detunable provided $\epsilon_i > 0.16$ for both loops.

A requirement for Robust Decentralized Detunability is that the individual loops are

stable. A decentralized IMC controller as parametrized in Eq. (5.9) will make the individual loops stable, which in most cases is an advantage. However, integral action is inherent in IMC controllers, and integral action and stability of the individual loops is known to be incompatible with stability of the overall system for certain plants. We would like to emphasize Step 4 in the Independent design procedure, that nominal stability must be checked explicitly for one value of ϵ within the bounds found. The Niederlinski Index criterion [17] gives a necessary condition for obtaining stability both of the individual loops and the overall system when there is integral action in all channels. The Niederlinski Index criterion has recently been generalized to open loop unstable plants [9]. Let the number of Right Half Plane (RHP) poles in G be n_U (including multiplicities), and the number of RHP poles in $\tilde{G}$ be $\tilde{n}_U$. Note that in general $\tilde{n}_U \neq n_U$. If all the individual loops are stable, a necessary condition for the stability of the overall system is that

$$\text{sign}\{N_I\} = \text{sign}\{\frac{\det G(0)}{\det \tilde{G}(0)}\} = \text{sign}\{(-1)^{-n_U+\tilde{n}_U}\} \tag{5.22}$$

Thus, before attempting to perform an independent design, one should check that overall stability can be achieved with integral action in all channels and having stable individual loops.

Example 3

Consider the process

$$G(s) = \begin{bmatrix} \frac{5}{20s^2+12s+1} & \frac{8}{20s^2+12s+1} \\ \frac{6}{40s^2+12s+1} & \frac{2}{40s^2+12s+1} \end{bmatrix} \tag{5.23}$$

with independent actuator uncertainty with uncertainty weight $W_I(s) = 0.2\frac{10s+1}{s+1}I_2$. Since this plant is stable and the Niederlinski Index is negative, $N_I = -3.8$, we know that we cannot have the individual loops stable and at the same time achieve overall system stability. Nevertheless, we proceed with independent design, and choose third order low pass filters for both loops. We find that Step 3 in the independent design procedure indicates that any value of $\epsilon > 4$ (approximately) will give robust stability (figure omitted). Calculating μ for $\epsilon = 5$ for both loops, we do indeed obtain a value of $\mu < 1$ at all frequencies. The reason, which we find in Step 4 in the independent design procedure, is that the overall system is nominally unstable. The μ test merely tells us that this instability is a robust property. For other cases, it may not be this easy to tell *à priori* that the overall system will be unstable with the individual loops stable.

5.5.5 Conclusions on Independent Design

We have proposed a parametrization of the class of allowable decentralized designs which is based on the following four key steps:

1. Use an IMC controller design for each loop.
2. Select the filter time constant ϵ_i as the "uncertain" parameter.
3. Parametrize ϵ_i and $e_i = 1/\epsilon_i$ such that only positive values are allowed.
4. Obtain bounds on both ϵ_i and e_i that guarantee robust stability/performance.

We have found that:

- The result of considering only decentralized IMC controllers with a specified filter structure, is that the set of possible controller designs considered is much smaller than the set of possible controller designs when trying to find bounds on $\tilde{S}$ and $\tilde{H}$, and the resulting bound are therefore less conservative.
- One can derive a bound on the IMC filter time constants which ensures that the system is robust decentralized detunable.
- It is critical that real perturbations are used for the parametrization of ϵ_i and e_i. μ software capable of handling real perturbations is therefore needed.

The bounds obtained are common to all the filter elements, and it is not obvious how to take advantage of the possibility of having differing filter time constants in the different filter elements. However, one may easily use constant ratios between the filter time constants in the independent design procedure (e.g. choosing $\epsilon_1 = \epsilon^*$, $\epsilon_2 = 10\epsilon^*$, etc.).

If independent design fails in the first step with our improved independent design procedure, the "uncertainty" associated with the filter time constants can be reduced even further by assuming all filter time constants to have fixed values relative to each other (e.g. assuming all filter time constants in *all* filter elements to be equal). This may be termed "simultaneous design". For a plant of size $n \times n$ and low pass filters of order n_f, this will reduce the "uncertainty" associated with the filter time constants from n real scalar uncertainties repeated n_f times, to one real scalar uncertainty repeated $n \times n_f$ times.

The independent design procedure proposed here can also be applied to other types of controllers, for example, one can find bounds on the ratio of gain to integral time (k/T_i) for PID controllers. However, decentralized IMC controllers have only one tuning parameter, and are therefore preferable for our independent design procedure.

One can easily use parameter optimization to find the filter time constants ϵ_i that minimize $\mu(M)$. Independent design has the advantage of providing a *range* of values for which robust stability/performance is fulfilled, and can also guarantee that the system is robust decentralized detunable.

5.6 Sequential Design

Sequential design of decentralized controllers was introduced in the control literature by Mayne [14], but it is probably fair to say that it has always been the most common way of designing decentralized controllers in industry. Sequential design involves closing and tuning one loop at the time. This is the advantage of sequential design: each step in the design procedure may be considered as a single-input single-output (SISO) control problem. The loops that have already been designed are (assumed to be) kept in service when closing and tuning subsequent loops. However, if the subsequent closing of other loops makes a loop perform badly, the engineer must go back and redesign a loop that has been closed earlier. Thus sequential design may involve iteration.

If the loops of a decentralized control system have been designed in the order $1, 2, \cdots, k, k+1, \cdots, n$ without having to redesign any controller element, and stability has been achieved after the design of each loop, sequential design will automatically ensure a limited degree of failure tolerance: The system will remain stable if *all* the loops $k+1, \cdots, n$ are to fail or be taken out of service simultaneously. Similarly, during startup the system will be stable if the loops are brought into service in the same order as they have been designed.

Sequential design of decentralized controllers has been addressed by several authors (e.g. [2, 3, 14, 16, 23]). However, the only published procedure for robust (in terms of μ) sequential design of decentralized controllers appear to be the one due to Chiu and Arkun [3, 4].

5.6.1 The Robust Sequential Design Procedure of Chiu and Arkun

The sequential design procedure of Chiu and Arkun [3, 4] involves performing the *independent* design procedure of Skogestad and Morari [20] at each step in the design. One loop is then closed with tuning parameters in accordance with the bounds obtained from independent design. A new independent design is then performed with this loop closed, and so on. If the independent design procedure fails (the bounds conflict and Eq. (5.5) cannot be satisfied) in the first step, Chiu and Arkun propose to close a

sufficient number of loops to enable independent design to be performed. No guidelines are given for how to choose tuning parameters for these loops that have to be closed prior to the application of independent design.

Another weakness with their design method is that the controller design for loop k is done by finding bounds for all remaining loops $k-n$, which guarantee robust stability or performance if the controllers in *all* loops $k-n$ fulfill the bounds. Closing *only* loop k with a controller fulfilling the bounds found, will not necessarily mean that the subsystem consisting of loops $1-k$ is stable.

The improved independent design procedure presented in Section 5.5 can readily be incorporated into the sequential design procedure of Chiu and Arkun. However, even with this improvement we may encounter all the problems mentioned above. There is therefore a need for a robust sequential design procedure starting from one single loop rather than a procedure involving a large number of loops simultaneously.

5.6.2 Problems Unique to Sequential Design

In sequential design, three problems arise which are not encountered using independent design or parametric optimization:

1. The final controller design, and thus the control quality achieved, may depend on the order in which the controller in the individual loops are designed.

2. The optimal design for the controller in loop k depends on the design of the controllers in *all* the other loops, some of which are still not designed.

3. The individual elements of the plant transfer function G may contain right half plane (RHP) zeros that do not correspond to RHP transmission zeros of G

The conventional rule for dealing with problem 1 is to close the fast loops first, the reason being that the loop gain and phase in the bandwidth region of the fast loops is relatively insensitive to the tuning of the slower loops. While this argument is reasonable for *loop* k, *output* k may still be sensitive to the tuning of the controller in a slower loop l, if u_l has a large effect on y_k.

We will attempt to reduce the severity of problem 2 by using simple estimates of how the undesigned loops will affect the output of the loop to be designed.

We require the system to be stable after the closing of each loop, but it may not be possible to close the fast loop (k) first, if the corresponding transfer function element has a significant RHP zero that is not a transmission zero of the plant G. However, such RHP zeros in the individual elements of G may disappear when the other loops

are closed (as the RHP zero is not an transmission zero), and it may therefore be possible to achieve fast control in loop k if the controller for this loop is designed at a later stage. This is illustrated in Example 4 below.

5.6.3 Preliminaries

In the following, we will assume without loss of generality that the loops are closed (and controllers designed) in the order $1, 2, \cdots, k, k+1, \cdots$, and that the loop to be designed is k. Let G_k denote the submatrix of dimension $k \times k$ in the upper left corner of G. Introduce $\hat{G}_k = \text{diag}\{G_k, g_{ii}\}$, $i = k+1, k+2, \cdots, n$, $\hat{S}_k = (I + \hat{G}_k C)^{-1}$ and $\hat{H}_k = I - \hat{S}_k = \hat{G}_k C (I + \hat{G}_k C)^{-1}$. We then have

$$S = \hat{S}_k (I + E_k \hat{H}_k)^{-1} \tag{5.24}$$

where $E_k = (G - \hat{G}_k)\hat{G}_k^{-1}$. When performing sequential design, one should keep in mind that the effective transfer function from u_k to y_k can change when subsequent loops are closed. This is due to the interaction between the loops. We see that rows 1 to k of $(I + E_k \hat{H}_k)^{-1}$ expresses how interaction affects loops 1 to k, and can be considered as an *input weight* to $S_k = (I + G_k C_k)^{-1}$.

For the special case $k = 1$ we have $\hat{G}_1 = \tilde{G}$, $\hat{S}_1 = \tilde{S}$ and $\hat{H}_1 = \tilde{H}$ (see Notation). Recall that $\tilde{S}$ and $\tilde{H}$ consist of the closed loop transfer functions of the individual loops.

Loop Gain Requirements for Setpoint Following and Disturbance Rejection

Consider the feedback system in Fig. 5.1a. Assume that the plant transfer function G and the disturbance transfer function G_d are scaled such that the largest tolerable offset in any controlled variable has magnitude 1 and the largest individual disturbance expected has magnitude 1 at any frequency. Prior to designing the first loop we have $k = 1$ and $\hat{G}_1 = \tilde{G} = \text{diag}\{g_{ii}\}$ in Eq. 5.24. Eventually all the loops will be closed and at low frequencies we will have $\tilde{h}_i \approx 1 \forall i$. We use this information to predict the overall response in terms of the individual loop responses. Consider only frequencies below the bandwidths of all the loops ($\tilde{h}_i \approx 1 \ \forall i$ and $(I + E_1 \tilde{H})^{-1} \approx \tilde{G} G^{-1}$) and find

$$e = y - r = -Sr + SG_d d \approx -\tilde{S}\Gamma r + \tilde{S}\Gamma G_d d; \quad \omega < \omega_B \tag{5.25}$$

where $\Gamma = \tilde{G} G^{-1} = \{\gamma_{ij}\}$ is known as the Performance Relative Gain Array (PRGA) [8] and $\Gamma G_d = \{\delta_{ik}\}$ is known as the Closed Loop Disturbance Gain (CLDG). Thus $|\gamma_{ij}(j\omega)|$ gives the loop gain requirement at frequency ω for a change in setpoint j

to cause an acceptably small offset in output i. Likewise, $|\delta_{ik}(j\omega)|$ gives the loop gain requirement in loop i for rejecting disturbance k. The PRGA and CLDG are introduced in [8, 22], and a more detailed explanation of their uses can be found therein. Here we will use the PRGA for two purposes:

1. Determine the *order* of loop closing (closing first loops that are required to be fast).
2. Estimate loop gain requirements for counteracting interactions and disturbances, thereby finding an estimate of the complementary sensitivity functions ($\tilde{h}_i$'s) for the loops that are not closed.

5.6.4 Sequential Design Procedure

The proposed sequential design procedure is outlined here. We assume that the design specifications include a performance requirement of the type

$$\bar{\sigma}(W_p S [I \ G_d]) < 1 \ \forall \omega$$

i.e. we want to optimize *robust performance* (for some specified model uncertainty) both with respect to setpoint changes and disturbances[4]. Note that S can be expressed in terms of $\hat{S}_k$ as shown in Eq. (5.24). Obviously, we can only have a performance requirement for an output where we have a controller. For this reason, define W_{pk} as the matrix of dimension $k \times k$ consisting of the upper left corner of W_p. Likewise, define W_{1k} as the matrix consisting of the first k rows of $(I + E_k \hat{H}_k)^{-1}[I \ G_d]$, using an estimate of $\hat{H}_k$.

Our sequential design procedure is then for step k to design a SISO controller that minimizes $\sup_\omega \bar{\sigma}(W_{pk} S_k W_{1k})$. *If model uncertainty is included the problem is to minimize the structured singular value of some matrix, in which* W_{1k} *is used as an input weight for performance.*

Comments to the Sequential Design Method:

Step 1. When designing loop 1, we have $\hat{G}_k = \tilde{G} = \text{diag}\{g_{ii}\}$; $i = 1, \cdots, n$. An estimate of $\tilde{H} = \text{diag}\{\tilde{h}_i\}$ is needed for calculating $(I + E_k \tilde{H})^{-1}$. The loop gain requirements given above in terms of the PRGA and CLDG are helpful for this purpose, as will be demonstrated in the examples. W_{p1} consists of the first element on the main diagonal of W_p, and W_{11} is the first row of $(I + E_k \tilde{H})^{-1}[I \ G_d]$. For a plant of dimension $n \times n$ and with n_d disturbances, the perturbation block for performance will thus be of dimension $(n + n_d) \times 1$.

[4] If disturbances are not considered, an empty matrix can be substituted for G_d.

Step k. Here $\hat{G}_k = \text{diag}\{G_k, g_{ii}\}$; $i = k+1, \cdots, n\}$. Controllers for loops $1, \cdots, k-1$ have now been found, and $\hat{H}_k$ is estimated to be $\hat{H}_k = \text{diag}\{H_{k-1}, \tilde{h}_i\}$; $i = k, \cdots, n$, where $\tilde{h}_i$ is the original estimate of the complementary sensitivity functions for loop i. Here W_{1k} consists of the k first rows of $(I + E_k\tilde{H})^{-1}[I\ G_d]$, and the perturbation block for performance is of dimension $(n + n_d) \times k$.

Step n. The controllers in all the other loops have been designed, and we therefore have $W_{1n} = [I\ G_d]$, and the perturbation block for performance is of dimension $(n + n_d) \times n$.

Design Method for the Controllers in the Individual Loops

A choice has to be made as to what design method should be used for designing the controllers in the individual loops. We will consider μ-synthesis and parametric optimization.

μ synthesis is relatively fast, but it results in controllers with a high number of states. The number of states in controller element c_i will be equal to the number of states in the μ interconnection matrix for the design problem, plus the number of states in the D-scaling matrices that are used to scale the interconnection matrix. Controller element c_i will become a part of the interconnection matrix when designing subsequent loops, and the number of states will therefore accumulate. Model reduction for the reduction of the number of states in the controller is therefore necessary. In our experience, performance may suffer severely when the number of states in each controller element is reduced to a number normally considered acceptable for process control (typically three states or less).

Parameter optimization is relatively slow on the computer. The controller has to be parametrized à priori, e.g. using a PID structure, and the achievable control quality is dependent on the controller parametrization. However, the number of states in the controller is fixed, and no model reduction is necessary. In this work parametric optimization is therefore chosen for the design of the individual loops.

Iteration

Iteration (redesigning loops) is in general undesirable, both because it is time consuming and because one is no longer guaranteed the limited degree of failure tolerance normally associated with sequential design when one has to resort to iteration. One objective with our procedure is that the estimate of W_{1k} (using the estimate of $\hat{H}_k$) should reduce the need for iteration. However, it will of course be possible to reduce the value of μ further using iteration, but the improvement has been small for the examples we have considered.

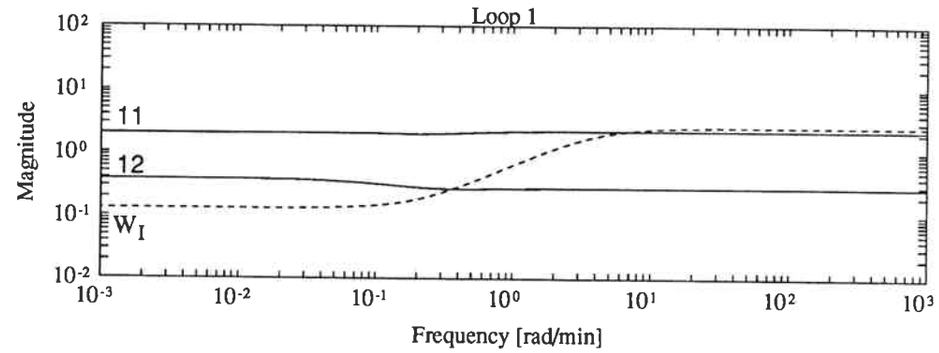


Figure 5.7: Performance relative gain array and uncertainty weight for Example 2. Element 13 is smaller than 10^{-2} at all frequencies.

5.6.5 Examples

The sequential design procedure outlined above will be demonstrated in two examples.

Example 2 (continued)

Consider again Example 2 from [3], and add the performance requirement $\bar{\sigma}(W_p S) < 1$. This should be satisfied for all possible plants allowed by the input uncertainty. We choose the performance weight

$$W_p(s) = w_p(s)I; \quad w_p(s) = 0.4\frac{\tau_{cl}s + 1}{\tau_{cl}}$$

The objective is to make the system as fast as possible in a robust sense, by minimizing τ_{cl} subject to $\mu_{RP} \leq 1$ (μ_{RP} meaning μ for robust performance).

We choose to pair on the diagonal elements of G as done previously. The PRGA for this example is shown in Fig. 5.7, together with the uncertainty weight. PRGA elements larger than 1 imply interactions, and the figure shows that there is severe interaction from loops 1 and 2 into loop 3. The loop gain in loop 3 must consequently be high at the frequencies where the feedback in loops 1 and 2 is effective. This means that the bandwidth in loop 3 has to be higher than the bandwidths in loops 1 and 2. The bandwidth in loop 3 will be limited by the input uncertainty. When attempting to minimize τ_{cl}, we must therefore minimize the difference in bandwidths between loop 3 and the two other loops. From Fig. 5.7 and Eq. (5.24) (with $\tilde{H}$ substituted for $\hat{H}_k$) we also see that we want $\tilde{h}_1$ and $\tilde{h}_2$ to roll off quickly at frequencies beyond their respective loop bandwidths, as this will reduce the interactions from loops 1 and 2 into loop 3 in the frequency range above the bandwidths of loops 1 and 2. The initial estimates for the complementary sensitivity functions for the individual loops are therefore chosen to be of the form

$$\tilde{h}_i(s) = \frac{1}{(\frac{s}{\omega_i} + 1)^2} \tag{5.26}$$

Estimating $\tilde{h}_i$ to be a first order low pass filter would imply that loops 1 and 2 would roll off more slowly beyond their respective bandwidths, and the interactions from loops 1 and 2 into loop 3 would therefore make it necessary to have a larger difference in bandwidths between loop 3 and loops 1 and 2. The approach taken here to minimize τ_{cl} may be considered somewhat naive: τ_{cl} is minimized (subject to $\mu \leq 1$) during the controller design for each loop. However, as the same τ_{cl} in the end will apply for all outputs, we choose ω_i in Eq. (5.26) to be consistent with τ_{cl}, that is, we specify a $\tilde{h}_i = 1 - \tilde{s}_i$ such that $|w_p(j\omega)\tilde{s}_i(j\omega)| < 1 \forall \omega$. We thus choose $\omega_1 = 1/\tau_{cl}$. From the PRGA's in Fig. 5.7 we see that $|\gamma_{31}| \approx 5|\gamma_{32}|$ at frequencies around 1 and assuming

the magnitude of the loop gain in loop 3 to have a slope of -2 on a log-log plot this indicates $\omega_2/\omega_1 = 2.2$. Loop 3 is the fast loop, and considering the uncertainty weight we therefore fix $\omega_3 = 1 rad/min$. It is clear from Fig. 5.7 that loop 3 must be closed first, and probably loop 2 second as there is some interaction from loop 1 to loop 2, especially at high frequencies.

The controller parametrization is chosen to be

$$c_i(s) = k\frac{T_1 s + 1}{T_1 s}\frac{T_2 s + 1}{10T_2 s + 1} \tag{5.27}$$

Note that this is not on the PID form since the pole in the last term is at a lower frequency ($s = 0.1/T_2$) than the zero. This controller parametrization allows loops 1 and 2 to roll off quickly beyond their respective bandwidths, whereas the loop gain in loop 3 can increase rapidly over one decade at frequencies slightly lower than the loop bandwidth.

Step 1: Loop 3. W_{11} is the third row of $(I + E_k\hat{H}_k)^{-1}$, and there is one 1×1 perturbation block for the input uncertainty, and one 3×1 perturbation block for the performance specification. Iterating on τ_{cl} (and changing ω_1 and ω_1 correspondingly, as explained above) $\mu = 0.992$ is obtained for $\tau_{cl} = 8.5$, and the corresponding controller is

$$c_3(s) = 84.9\frac{4.70s + 1}{4.70s}\frac{4.01s + 1}{40.1s + 1} \tag{5.28}$$

Step 2: Loop 2. $\tilde{h}_3$ is updated in $\hat{H}_k$, and W_{12} is the second and third rows of $(I + E_k\hat{H}_k)^{-1}$. There is one diagonal perturbation block of dimension 2×2 for the input uncertainty, and a 3×2 perturbation block for performance. $\mu = 0.998$ is obtained for $\tau_{cl} = 11$, and we find

$$c_2(s) = -0.079\frac{1.32s + 1}{1.32s}\frac{0.186s + 1}{1.86s + 1} \tag{5.29}$$

Step 3: Loop 1. Now all loops are included in the design problem (and there are no disturbances present), consequently $W_{13} = I_3$, and there is one diagonal 3×3 perturbation block for the input uncertainty and a full 3×3 perturbation block for performance. $\mu = 1.000$ is obtained for $\tau_{cl} = 18$ and

$$c_1(s) = 0.04\frac{0.385s + 1}{0.385s}\frac{0.898s + 1}{8.98s + 1} \tag{5.30}$$

In comparison, the best decentralized controller found using simultaneous parametric optimization with the same controller parametrization gave $\mu = 1$ for $\tau_{cl} = 16$.

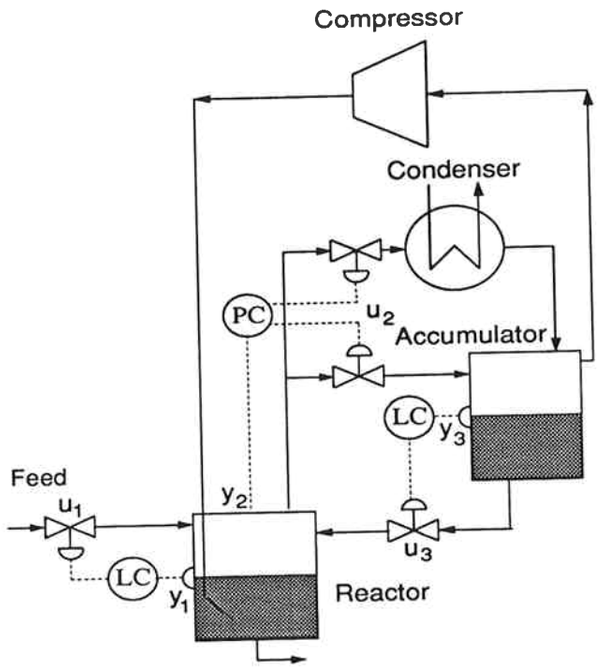


Figure 5.8: Schematic outline of the process in Example 4.

Example 4

We consider the polypropylene reactor studied by Lie and Balchen [11, 12] and Hovd and Skogestad [9]. A schematic outline of the process is shown in Fig. 5.8. There are three inputs, three outputs and four disturbances. The process has a complex pair of RHP poles, but no RHP transmission zeros. However, there are RHP zeros in all elements of G except in g_{11} at frequencies close to the frequency corresponding to the RHP poles. A more detailed description of the process and details on the scalings used are given in the appendix, together with a state space description of the process. Only input uncertainty is considered, and the uncertainty weight is

$$W_I(s) = w_I(s)I; \quad w_I(s) = 0.2\frac{\frac{1}{24}s+1}{\frac{1}{240}s+1}$$

which reflects a steady state uncertainty of 20% and a maximum neglected time delay of 0.5 minute. The performance requirement, in terms of the scaled outputs and disturbances, is $\bar{\sigma}(W_p S[I\ G_d]) < 1$ at all frequencies and for all uncertainties allowed by the uncertainty weight. The performance weight is given by

$$W_p(s) = w_p(s)I; \quad w_p(s) = 0.4\frac{0.2s+1}{0.2s}$$

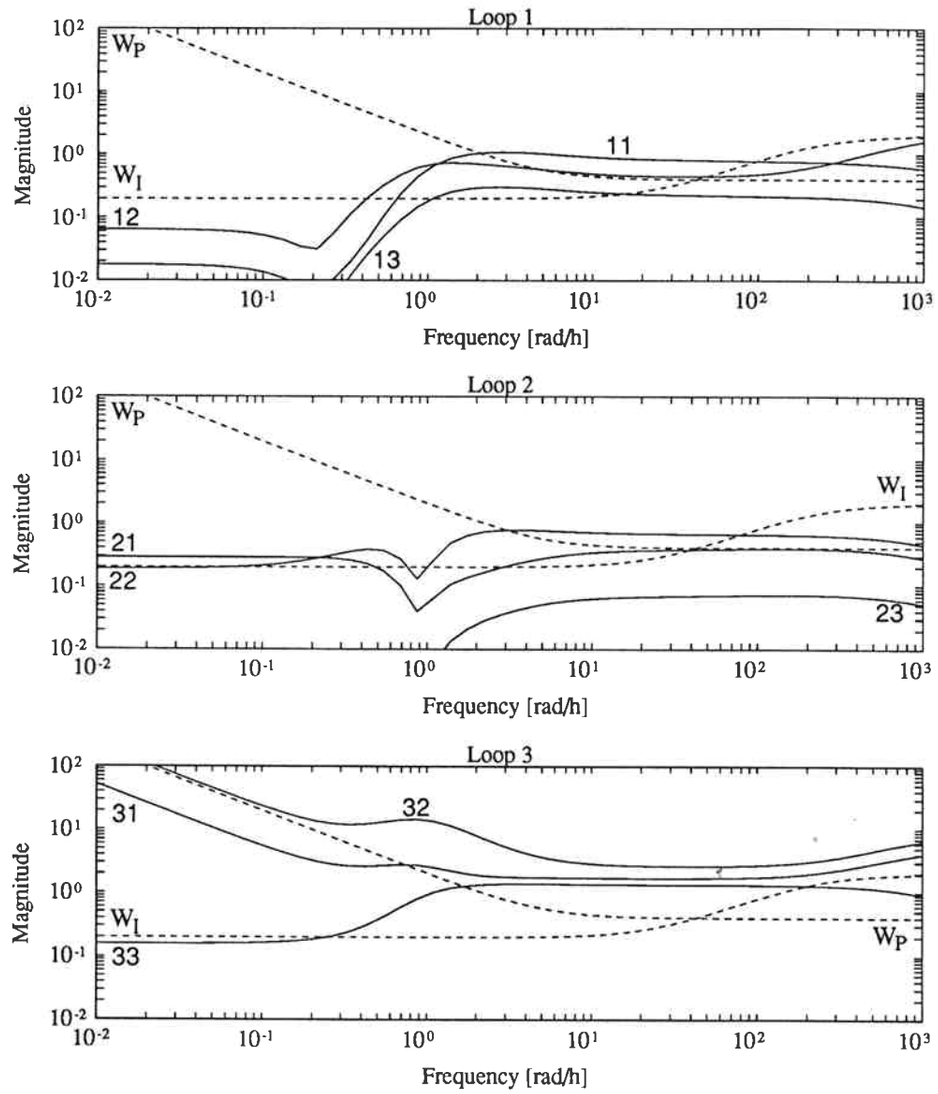


Figure 5.9: Performance relative gain array, uncertainty weight and performance weight for Example 4.

We will use decentralized control with pairings as indicated in Fig. 5.8. This pairing was found to be preferable in [11, 9], and corresponds to industrial practice. The PRGA and CLDG for this process are shown in Fig. 5.9 and Fig. 5.10 respectively, together with the uncertainty and performance weights. Fig. 5.9 shows that the interaction is relatively modest in the bandwidth region, except for loop 3 where the interaction

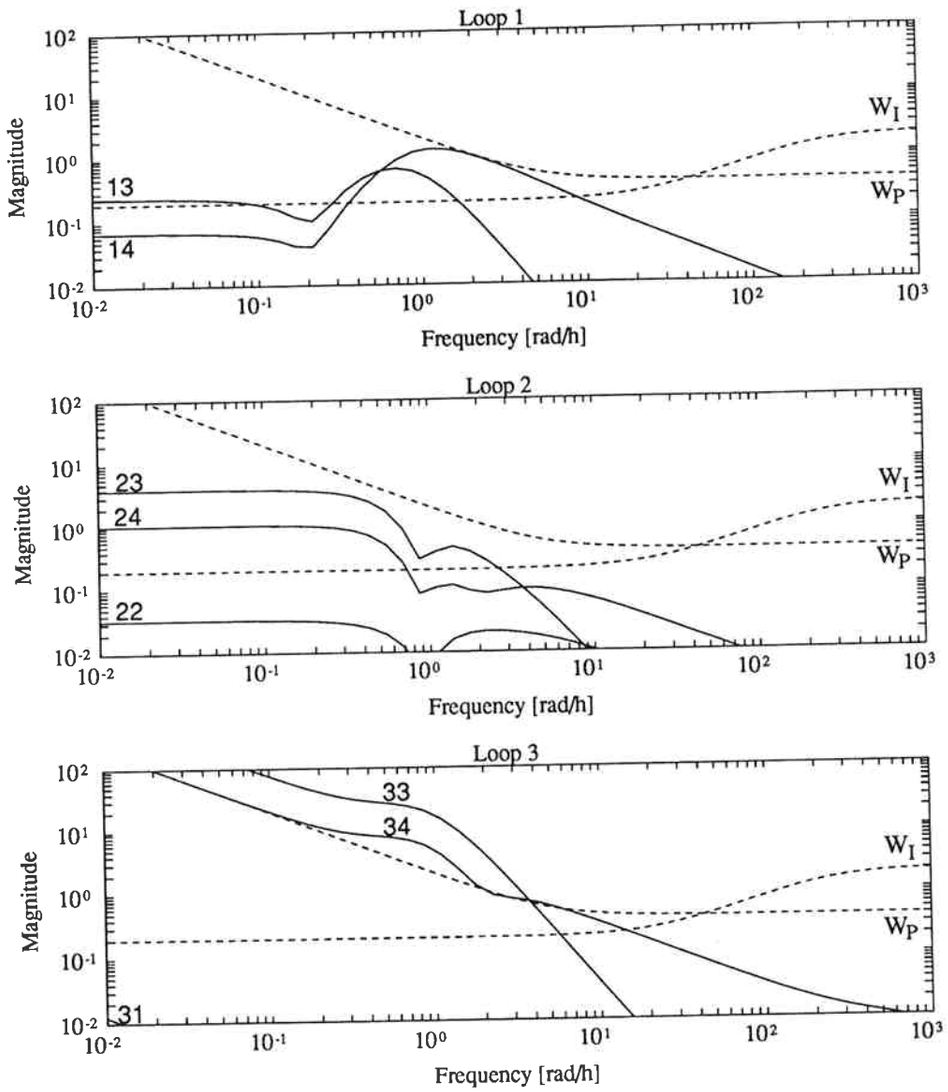


Figure 5.10: Closed loop disturbance gains, uncertainty weight and performance weight for Example 4. Elements of the CLDG not shown are smaller than 10^{-2} at all frequencies.

between the loops causes increased loop gain requirement at low frequencies. From Fig. 5.10 we see that loop 3 is also most severely affected by the disturbances. Loop 3 must therefore be the fastest loop, and loop 1 is chosen as the second fastest loop as it is somewhat more affected by disturbances than loop 2 at frequencies around 2

[rad/min]. The uncertainty and performance weights give upper and lower bounds for the bandwidths of the individual loops. We choose to spread out the bandwidths of the individual loops between these bounds, and choose the following initial estimated for the complementary sensitivity functions for the individual loops:

$$\tilde{h}_1 = \frac{1}{0.04s+1}; \quad \tilde{h}_2 = \frac{1}{0.2s+1}; \quad \tilde{h}_3 = \frac{1}{0.015s+1}$$

PI controllers, $c_i(s) = k\frac{T_i s+1}{T_i s}$ are used for all loops. Because of RHP zeros in individual elements and subsystems, the design sequence for this example must be 1-2-3, i.e., we are required to design the fastest loop last. In contrast to Example 2, the performance weight for this example is fixed, and we therefore minimize μ at each stage in the design.

Step 1: Loop 1. W_{11} is the first row of $(I + E_2\tilde{H})^{-1}[I\ G_d]$, and there is one 1×1 perturbation block for the input uncertainty and a 7×1 perturbation block for performance. The minimum value of μ found is 0.6363, and the corresponding (unscaled) controller $c_1(s) = 2.31 \cdot 10^5 \cdot \frac{0.962s+1}{0.962s}$.

Step 2: Loop 2. $\hat{H}_2$ is updated, using the controller found for loop 1. W_{12} consists of the first two rows of $(I + E_2\hat{H}_2)^{-1}[I\ G_d]$, there is one diagonal 2×2 perturbation block for the input uncertainty and a full 7×2 perturbation block for performance. The minimum value of μ found is 1.09, and the corresponding controller is $c_2(s) = 4.99 \cdot 10^{-2} \cdot \frac{0.148s+1}{0.148s}$.

Step 3: Loop 3. Now all loops are included in the design problem, consequently $\hat{G}_3 = G$ and $W_1 = [I\ G_d]$. There is a diagonal 3×3 perturbation block for the input uncertainty and a full 7×3 perturbation block for performance. The minimum value of μ found is 0.89, and the corresponding controller is $c_3(s) = -1.27 \cdot 10^5 \cdot \frac{0.131s+1}{0.131s}$.

In comparison, with the best decentralized PI controller found using simultaneous parametric optimization μ improved only marginally, from 0.89 to 0.86. The best multivariable controller found using μ-synthesis gave $\mu = 0.65$.

The fact that a lower value of μ is achieved after closing all loops that when closing loops 1 and 2 only, illustrates that the design problem may become easier as more loops are brought into service, and that the estimates of W_{1k} used may be conservative. However, it appears that the controller found using sequential design is relatively close to the optimal for a decentralized PI controller.

5.6.6 Discussion on Sequential Design

1. Sequential design using parameter optimization consists of much smaller optimization problems than designing all loops simultaneously by parameter optimization, and designing one loop at the time is therefore preferable. Also, parameter optimization for all loops simultaneously does not guarantee the limited degree of failure tolerance that is associated with sequential design. For the examples studied in this paper, the sequential design procedure presented in this paper achieved a quality of control that is not significantly poorer than that achieved using parameter optimization for all loops simultaneously.

2. The idea of using a simplified estimate of the effect of closing the other loops is not new. Balchen and Mummé ([1], Appendix C) derive an estimate the transfer function between input u_i and output y_i when all the other loops are closed, using an estimate of the complementary sensitivity function for the other loops. In [1] this is used to find pairings for decentralized control.

   Other uses can also be considered using the estimated transfer function from u_i to y_i when the other loops are closed. For example, one may use the Ziegler-Nichols tuning rules using this estimate of the transfer function. Since the Ziegler-Nichols tuning rules are very simple, the loops can be redesigned with little effort, thus reducing the problem that the initial estimate of the complementary sensitivity function for the other loops may well be poor in the bandwidth region.

3. It is easier to estimate the complementary sensitivity function for the individual loops than to estimate the controller in the individual loops. This holds especially at low frequency, where control is almost perfect, and we know that $\tilde{h}_i \approx 1$.

4. The idea of using an estimate of the effect of closing the other loops is not confined to the H_∞ or μ framework, it may also be used with other norms, e.g. H_2.

5. Many multivariable controllers consist of simple pre- and/or post-compensator and have the main dynamics in a diagonal matrix. The compensators are often designed to counteract interactions at a given frequency (e.g. steady-state decoupling), and there will be interactions at other frequencies. Both the independent design and the sequential design procedures in this paper may be used for designing the diagonal matrix of such multivariable controllers. Note that decentralized controllers are known to be relatively robust (but the performance may be poor even nominally), when using non-diagonal compensators the issue of robustness is more important [18].

5.6.7 Conclusions on Sequential Design

Sequential design should not be based on independent design, but rather on considering one loop at the time.

We propose a sequential design procedure, starting from the individual loops. It is shown that including the appropriate rows of $(I + E_k\hat{H}_k)^{-1}[I\ G_d]$, using an estimate of $\hat{H}_k$, has the effect of including an estimate of the effect that the loops that are still open will have on the closed loops and the loop that is closed at step k in the procedure.

Acknowledgement The authors wish to tank M. P. Newlin and P. M. Young of the California Institute of Technology for access to their software for calculating μ with mixed real and complex perturbations.

References

[1] Balchen J. G. and Mummé, K. I. (1988). *Process Control. Structures and Applications*, Van Nostrand Reinhold, New York, USA.

[2] Bernstein, D. S. (1987). Sequential Design of Decentralized Dynamic Compensators Using the Optimal Projection Equations. *Int. J. Control*, **46**, pp. 1569-1577.

[3] Chiu, M.-S. and Arkun, Y. (1991). A Methodology for Sequential Design of Robust Decentralized Control Systems. Submitted to *Automatica*, preprint.

[4] Chiu, M.-S. and Arkun, Y. (1991). A Methodology for Sequential Design of Robust Decentralized Control Systems. *Proc. ACC*, Boston, Massachusetts, pp. 1836-1841.

[5] Doyle, J. C., Wall, J. E. and Stein, G. (1982). Performance and Robustness Analysis for Structured Uncertainty. *Proc. IEEE Conf. Decision Contr.*, Orlando, FL., pp. 629-636.

[6] Doyle, J. C. and Chu, C.-C. (1985). Matrix Interpolation and H_∞ Performance Bounds. *Proc. ACC*, Boston, MA, pp. 129-134.

[7] Garcia, C. and Morari, M (1985). Internal Model Control - 2. Design Procedure for Multivariable Systems. *Ind. Eng. Chem. Proc. Des. & Dev.*, **24**, pp. 472-484.

[8] Hovd, M. and Skogestad, S. (1992). Simple Frequency-Dependent Tools for Control System Analysis, Structure Selection and Design. To appear in *Automatica*.

[9] Hovd, M. and Skogestad, S. (1992). Controllability Analysis for Unstable Processes. *Proc. IFAC Workshop on Interactions Between Process Design and Process Control*, London, England, September 1992.

[10] Lee, J. H. and Morari, M. (1991). Robust Measurement Selection. *Automatica*, **27**, 3, pp. 519-527.

[11] Lie, B. (1990). Control Structures for Polymerization Processes Applied to Polypropylene Manufacturing. *Dr. Ing. Thesis*, University of Trondheim-NTH.

[12] Lie, B. and Balchen, J. G. (1992). A Comparison of Strategies for Control of a Polypropylene Reactor, *Preprints IFAC Symposium DYCORD+'92*, College Park, Maryland, Apr. 1992, 265-270.

[13] Maciejowski, J. M. (1989). *Multivariable Feedback Design.* Addison-Wesley, Wokingham, England.

[14] Mayne, D. Q. (1973). The Design of Linear Multivariable Systems. *Automatica*, **9**, pp. 201-207.

[15] Morari, M. and Zafiriou, E., (1989). *Robust Process Control*, Prentice-Hall, Englewood Cliffs, New Jersey.

[16] Nett, C. N. and Uthgenannt, (1988). An Explicit Formula and An Optimal Weight for the 2-Block Structured Singular Value Interaction Measure. *Automatica*, **24**, pp. 261-265.

[17] Niederlinski, A. (1971). A Heuristic Approach to the Design of Linear Multivariable Interacting Control Systems. *Automatica*, **7**, pp. 691-701.

[18] Skogestad, S. and Morari, M. (1987). Implications of Large RGA-Elements on Control Performance, *Ind. Eng. Chem. Res.*, **26**, 11, pp. 2323-2330.

[19] Skogestad, S. and Morari, M. (1988). Some New Properties of the Structured Singular Value. *IEEE Trans. Autom. Control*, **33**, 12, pp. 1151-1154.

[20] Skogestad, S. and Morari, M. (1989). Robust Performance of Decentralized Control Systems by Independent Design. *Automatica*, **25**, 1, pp.119-125.

[21] Skogestad, S. and Lundström, P. (1990). Mu-optimal LV-control of Distillation Columns. *Computers chem. Engng.*, **14**, 4/5, pp. 401-413.

[22] Skogestad, S. and Hovd, M., (1990) Use of Frequency-dependent RGA for Control Structure Selection. *Proc. American Control Conference*, San Diego, CA., pp. 2133-2139.

[23] Viswanadham, N. and Taylor, J. H. (1988). Sequential Design of Large-Scale Decentralized Control Systems. *Int. J. Control*, **47**, pp. 257-279.

[24] Young, P. M., Newlin, M. P. and Doyle, J. C. (1991). μ Analysis with Real Parametric Uncertainty. *Proc. 30th CDC*, Brighton, England, pp. 1251-1256.

Appendix. Description of the Process in Example 4

The monomer feed enters into a stirred tank reactor containing a slurry of monomer, catalyst, cocatalyst, polymer and some impurities. The reaction is exothermic, causing some of the slurry components to vaporize. The vapor leaving the reactor is transferred to an accumulator vessel. Heat is removed from the system by condensing parts of the vapor leaving the reactor, before it enters the accumulator. Heat removal is adjusted by adjusting a split range valve which determines what fraction of the vapor leaving the reactor is passed through the condenser. The liquid in the accumulator is returned to the reactor, and the vapor from the accumulator is compressed and bubbled through the reactor slurry. This results in a 3×3 plant model $G(s)$ with seven states. The inputs and outputs are

y_1 - reactor slurry level $(0-1)$
y_2 - reactor pressure (gauge pressure in atmospheres)
y_3 - accumulator liquid level $(0-1)$.
u_1 - monomer feed flowrate (kg/h).
u_2 - split range valve position $(0-1)$.
u_3 - accumulator to reactor liquid flowrate (kg/h).

Four disturbances are considered:

d_1 - monomer feed temperature $(^\circ C)$
d_2 - cooling water flowrate through heat exchanger (kg/h)
d_3 - catalyst mass feed flowrate (kg/h)
d_4 - recycle flow of unreacted monomer (kg/h)

We scale the outputs such that a magnitude of 1 for the scaled outputs correspond to offsets: $y_1 = 0.05$, $y_2 = 1.0$atm and $y_3 = 0.10$. Likewise, the disturbances are scaled such that a magnitude of 1 for the scaled disturbances correspond to: $d_1 = 20^\circ C$, $d_2 = 10000kg/h$, $d_3 = 3kg/h$ and $d_4 = 1000kg/h$.

State space description of the plant in Example 4.

The description is given as $y(s) = G(s)u(s)+G_d(s)d(s) = [C(sI-A)^{-1}B+D]u(s)+[C(sI-A)^{-1}B_d+D_d]d(s)$.

$$A = \begin{bmatrix} 0 & 0 & 0 & -3.08e+03 \\ 4.30e-04 & -4.71e-01 & 0 & 0 \\ 3.49e-01 & 0 & -4.71e-01 & 0 \\ 0 & 0 & 0 & -7.50e+00 \\ 0 & 0 & 0 & 3.08e+03 \\ 0 & 0 & 0 & 3.57e+01 \\ 0 & -3.81e+02 & 0 & -3.53e+02 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 3.17e+03 & 1.00e+00 \\ 0 & 0 & 0 \\ 0 & 0 & -1.00e+00 \\ 0 & 8.49e+00 & -1.96e-02 \\ 0 & -3.17e+03 & 0 \\ 0 & -6.77e+01 & 0 \\ 0 & 0 & -1.00e+00 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.00e+00 & 0 & 1.00e+00 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1.45e-03 & 0 & -4.79e-04 \\ 0 & 0 & -1.00e+00 \\ 0 & 2.73e+03 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_d = \begin{bmatrix} 0 & 0 & 0 & -1.00e+00 \\ 0 & 0 & 1.00e+00 & -9.15e-04 \\ 0 & 0 & 0 & -7.40e-01 \\ 2.99e-05 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2.02e-04 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 7.49e-05 & 0 & 3.29e-05 & 1.23e-02 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.86e-01 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5.39e-04 & 5.75e-03 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad D_d = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Chapter 6

Robust control of systems consisting of symmetrically interconnected subsystems

M. Hovd S. Skogestad*

Chemical Engineering
University of Trondheim, NTH
N-7034 Trondheim, Norway

Abstract

This paper is concerned with robust control of systems consisting of n similar interacting subsystems. The transfer function matrices for these systems are block circulant matrices. For H_∞-optimal control, we show that we can simplify controller synthesis by considering two systems of the same dimension as the subsystems instead of the overall system. This leads to a dramatic reduction in dimension for systems consisting of many subsystems, enabling us to optimize the H_∞ criterion in n directions, and not only in the *worst* direction. The structured singular value, μ, is shown to be independent of the structure of the uncertainty for cases with single-input single-output (SISO) subsystems and only one kind of uncertainty. For MIMO systems with subsystems of dimension $n_o \times n_i$ or cases with several full or repeated uncertainty blocks, we are able to reduce the system size from $n \cdot n_o \times n \cdot n_i$ to a block diagonal system with two blocks of dimension $n_o \times n_i$. In the case of SISO subsystems, we derive limits on the magnitude of the interaction within which the system is guaranteed to be decentralized integral controllable (DIC).

*Author to whom all correspondence should be addressed. E-mail: SKOGE@KJEMI.UNIT.NO, phone: 47-7-594154, fax: 47-7-591410.

6.1 Introduction

Whenever possible, one should make use of any special properties of the system in order to simplify control system analysis or design. In this paper we study systems consisting of symmetrically interconnected subsystems, and show how controller design based on H_∞, H_2 or μ (structured singular value) may be simplified.

Systems consisting of similar subsystems in parallel which interact with each other occur quite often in practice, for instance in distribution networks, when there are parallel units (e.g. reactors, compressors, heat exchangers) in a chemical plant [14, 15] or for electric power plants operating in parallel [13]. With n nominally identical subsystems in parallel the $n \times n$ transfer matrix of the plant may be written

$$G(s) = \begin{bmatrix} g(s) & i(s) & i(s) & \dots & i(s) \\ i(s) & g(s) & i(s) & \dots & i(s) \\ i(s) & i(s) & g(s) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & i(s) \\ i(s) & i(s) & \dots & i(s) & g(s) \end{bmatrix} \tag{6.1}$$

where the diagonal elements $g(s)$ denote the transfer function of the individual subsystems, and the offdiagonal elements $i(s)$ denote the interactions. For MIMO subsystems, both $g(s)$ and $i(s)$ are matrices. We have not found any name for the matrix $G(s)$ in the literature, but we shall refer to it as a *block parallel matrix* in the following, as we believe that transfer function matrices of the form of $G(s)$ in Eq. (6.1) occur predominantly for nominally identical, interacting processes in parallel.

This type of system is termed a *Symmetrically Interconnected System* by Sundareshan and Elbanna [17]. Sundareshan and Elbanna studies conditions for controllability and observability of the system and solutions to the matrix Riccati and Lyapunov equations. They also give a methodology for the synthesis of a decentralized controller and an alternative "almost decentralized' control structure. No robustness analysis is performed. Lunze [13] studies robust stability of symmetrically interconnected systems. Lunze's model formulation can take account of a variety of different uncertainties and model errors. However, Lunze does not account for the structure of the uncertainty, which can lead to very conservative results, as we illustrate in this paper.

6.2 Notation

In this paper, $G(s)$ denotes the plant, which is assumed to consist of n subsystems in parallel, each subsystem of dimension $n_o \times n_i$. The plant $G(s)$ therefore has the

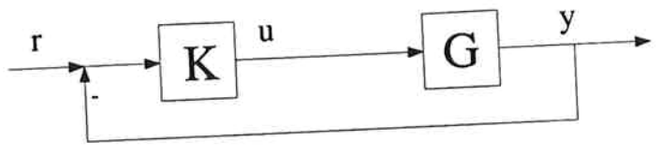


Figure 6.1: Block diagram of a feedback system.

dimension $n \cdot n_o \times n \cdot n_i$. The controller is denoted $K(s)$, and $S(s) = (I + G(s)K(s))^{-1}$ denotes the sensitivity function. The reference signal is denoted r, the manipulated inputs u, and the controlled outputs y. A block diagram of a feedback system is shown in Fig. 6.1. Circulant and block circulant matrices are denoted $C(s)$, and parallel and block parallel matrices are denoted $P(s)$. The Laplace variable s will often be suppressed to simplify the notation. Eigenvalues are denoted λ, with two subscripts: a letter referring to the matrix of which λ is an eigenvalue, and a number to distinguish the different eigenvalues of a matrix. Thus λ_{X1} is the first eigenvalue of the matrix X. The blocks on the diagonal of a matrix that is transformed to be block diagonal (see Eq. (6.22)) are denoted γ; subscripts are used for γ in the same way as for λ. The matrix M is in general a matrix consisting of blocks which are block circulant, but M is also used for the matrix in the design objective, i.e. we want to optimize the H_∞ norm or structured singular value of M. The matrix N is the matrix in the design objective (M) expressed as a linear fractional transformation (LFT) of the controller K. The matrix $\tilde{A}$ means that a matrix A has been transformed, such that it consists of blocks which are block diagonal.

6.3 Examples of Processes Consisting of Symmetrically Interconnected Subsystems

It is quite common to have units in parallel, either because one single unit would be too large or to add flexibility. Typical examples of parallel units include reactors, heat exchangers and compressors, and some examples may be found in the books of Shinskey [14, 15]. Similar examples are found for electric power generation plants [13].

For a system consisting of SISO subsystems, we define the degree of interaction at steady state

$$a(0) = i(0)/g(0) \tag{6.2}$$

Example 1. *Flow splitting.* One example is the control of flow in parallel streams from a single source as shown in Fig. 6.2 ([14], p. 201), where the manipulated inputs are the n valve positions and the controlled outputs are the n flows. Opening valve 1 causes q_1 (flow 1) to increase and q_2 (flow 2) to decrease because of the consequent

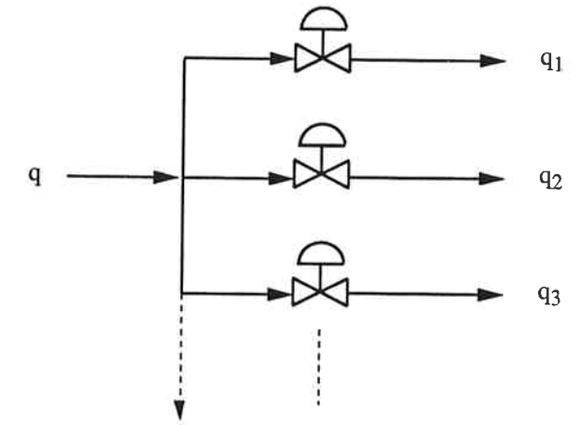


Figure 6.2: Splitting into parallel streams.

reduction in header pressure p. If there are two parallel streams the steady-state value of a is expected to lie between -1 and 0. The value of 0 would be obtained if the source was a large tank such that the header pressure was unaffected by increasing flow 1, and the value of -1 would be obtained if the source was a pump with constant total flow q. For n parallel streams from a single source similar arguments yield at steady state

$$-1/(n-1) \le a(0) \le 0 \tag{6.3}$$

The lower bound is obtained by considering constant total flow. In this case a change Δq_1 in flow 1 would yield $\Delta q_2 = \Delta q_3 = \cdots = \Delta q_n = -\Delta q_1/(n-1)$. A value less than the lower bound $-1/(n-1)$ would imply that the total flow is reduced by opening a valve and is unlikely in a practical situation.

Example 2. *Parallel reactors with combined precooling.* In Fig. 6.3 is shown n identical mixing tank reactors in parallel, with a common precooler. The cooling medium comes from a single source which is split into n streams and then completely evaporated by heat exchange with the reactors. The streams are then combined and this stream is superheated by precooling the reactor feed. At steady state all temperatures and flows in the parallel streams are assumed equal. Consider the transfer matrix $G(s)$ between the valve positions q_i and the reactor temperatures T_i. By neglecting the dynamics of the evaporator and the superheater, the model $G(s)$ can be shown [16] to be of the

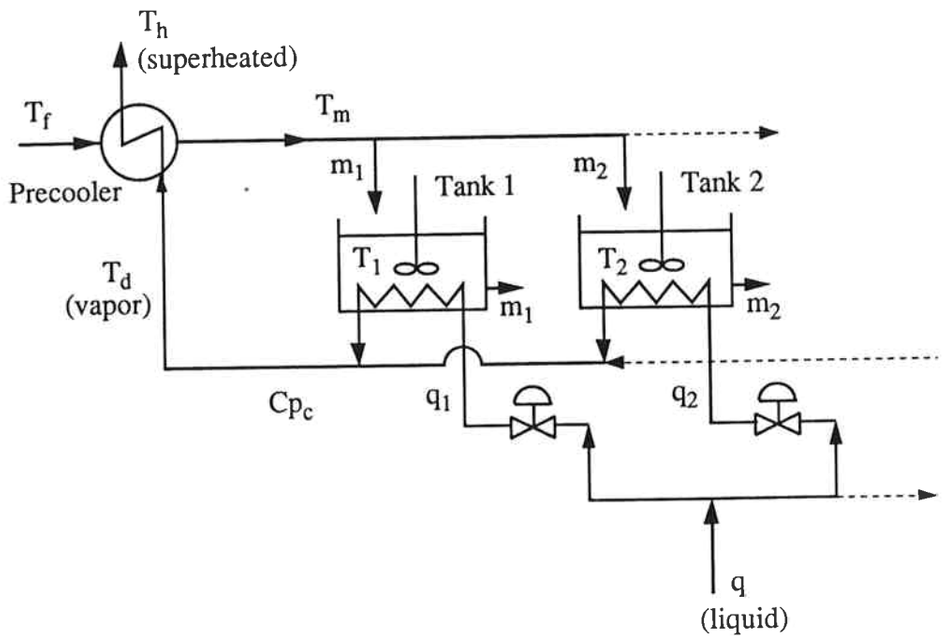


Figure 6.3: Cooling system for parallel reactors (Example 2).

form

$$G(s) = \frac{k}{\tau s + 1} \begin{bmatrix} 1 & a & a & \dots & a \\ a & 1 & a & \dots & a \\ a & a & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a \\ a & a & \dots & a & 1 \end{bmatrix} \tag{6.4}$$

where k and a are real constants and τ is the time constant for holdup in each of the reactors. Based on physical arguments we have

$$-1/(n-1) \leq a \leq 1 \tag{6.5}$$

The lower limit is obtained by considering the case with no precooler and assuming constant total flow (recall the parallel flow example above). The upper limit is obtained by considering the case with no evaporator, and thus no heat exchange taking place in the tanks. In this case the cooling streams are split and then recombined without changing their temperatures, and an increase in any cooling stream will affect all reactor temperatures equally and we have $a = 1$.

Note that $G(s)$ is singular both for $a = -1/(n-1)$ ($\text{rank}(G) = n-1$) and for $a = 1$ ($\text{rank}(G) = 1$). Obviously, when G is singular, independent control of the controlled

outputs is not possible. An example of this is found in Braatz et al. [2], who consider the cross directional control of a coating process for which $a = -1/(n-1)$.

We have searched for other examples of parallel processes with a outside the bounds given by Eq. (6.5), but have been unable to find any.

In many cases the individual subsystems may have several inputs and outputs, for example, if there are reactors in parallel for which one wants to control both temperature and some concentration(s), using cooling and flow(s) as manipulated variables.

In the example above the processes were parallel but the control objectives (e.g. to keep the individual reactor temperatures constant) were otherwise decoupled. There are also interesting examples of control of parallel processes where the control objectives are coupled. For example, consider again the stream split example in Fig. 6.2 and assume the control objective is to keep the total flow $\sum_i q_i$ at a fixed value q_s subject to the constraint that the individual flows should be equal, that is, $q_1 = q_2 = \cdots = q_n$. The requirement of equal flows may be needed, for example, to avoid overheating of tubes in a burner with parallel passes ([15], p. 104). Since the equal flow objectives yield $n-1$ objectives of the form $q_2 - q_1 = 0$, there are a total of n control objectives which may be collected in the output vector y, but note that the overall transfer function from inputs (valve positions) to outputs y is no longer an example of a symmetrically interconnected system as defined in Eq. (6.1).

6.4 Results from Matrix Theory

6.4.1 Systems with SISO Subsystems

Consider the parallel matrix P, which is a complex square $n \times n$ matrix of the form

$$P = k \begin{bmatrix} 1 & a & a & \cdots & a \\ a & 1 & a & \cdots & a \\ a & a & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a \\ a & a & \cdots & a & 1 \end{bmatrix} \tag{6.6}$$

where both a and k in general can be functions of frequency.

Classification of the Matrix P

The parallel matrix P is a symmetric matrix which belongs to a subclass of circulant matrices. The general form of a circulant matrix C is:

$$C = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_n & c_1 & \cdots & c_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_2 & c_3 & c_4 & \cdots & c_1 \end{bmatrix} \tag{6.7}$$

It can easily be seen that if $c_1 = k$ and $c_2 = c_3 = \cdots = c_n = ak$ then $P = C$. The results in this section on circulant matrices are from [1] and [5]. Circulant matrices belong to the class known as Toeplitz matrices, as all elements along any one diagonal are identical.

Properties of the Matrix P

Eigenvalues, Eigenvectors and Singular Value Decomposition. Introduce $v_l = exp(\frac{2\pi(l-1)i}{n})$ where $i = \sqrt{-1}$ and $l = 1, \cdots, n$. That is, v_l is a root of the equation $v^n = 1$, and we have

$$1 + v_l + v_l^2 + \cdots + v_l^{n-1} = 0 \text{ for } v_l \neq 1 \tag{6.8}$$

From the theory of circulant matrices we know that the eigenvalues of the circulant matrix C are given by:

$$\lambda_{Cl} = c_1 + c_2 v_l + c_3 v_l^2 + \cdots + c_n v_l^{n-1} \tag{6.9}$$

This means that the parallel matrix P has eigenvalues λ_{Pl} given by the formula:

$$\lambda_{Pl}/k = 1 + a(v_l + v_l^2 + \cdots + v_l^{n-1}) \tag{6.10}$$

From Eqs. (6.8) and (6.10) we see that the matrix P will have at most two distinct eigenvalues[1] which are given by

$$\lambda_{P1} = (1 + (n-1)a)k \tag{6.11}$$

$$\lambda_{P2} = \lambda_{P3} = \cdots = \lambda_{Pn} = (1-a)k \tag{6.12}$$

The eigenvector corresponding to λ_{Pl} is:

$$m_l = [1 \; v_l \; v_l^2 \; \cdots \; v_l^{n-1}]^T \tag{6.13}$$

[1] This is a key property that we make use of in this paper.

n	F_n	R_n
$n=2$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
$n=3$	$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -0.5(1+i\sqrt{3}) & -0.5(1-i\sqrt{3}) \\ 1 & -0.5(1-i\sqrt{3}) & -0.5(1+i\sqrt{3}) \end{bmatrix}$	$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 & \sqrt{2} & 0 \\ 1 & -0.5\sqrt{2} & 0.5\sqrt{6} \\ 1 & -0.5\sqrt{2} & -0.5\sqrt{6} \end{bmatrix}$
$n=4$	$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}$	$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Table 6.1: Alternative choices for the eigenvector matrix for $n \times n$ parallel matrices.

Since v_l can take n distinct values, P will always have a complete set of eigenvectors, and will thus always be diagonalizable. In fact, all circulant matrices of the same order have the same eigenvectors, and are therefore diagonalized by the same eigenvector matrix, the Fourier matrix. The Fourier matrix of order n is given by [5]

$$F^H = \frac{1}{\sqrt{n}}[m_1 \; m_2 \; \cdots \; m_n] \tag{6.14}$$

F is unitary ($FF^H = F^H F = I$), and we have for any circulant matrix C

$$C = F^H \Lambda_C F; \qquad \Lambda_C = \text{diag}\{\lambda_{C1}, \cdots, \lambda_{Cn}\} \tag{6.15}$$

Furthermore, we have for the singular value decomposition $C = U\Sigma_C V^H$ that the singular values $\sigma_l = \mid \lambda_l \mid$, and $V = F^H$, the eigenvector matrix, and $U = F^H D$, where $D = \text{diag}\{d_l\}$, $d_l = \lambda_l / \mid \lambda_l \mid = \lambda_l/\sigma_l$.

Diagonalization of Parallel Matrices Using Real Eigenvector Matrices. In general, the eigenvectors corresponding to repeated eigenvalues are not unique. Thus, for *parallel* matrices, any linearly independent set of $n-1$ vectors arising from linear combinations of $m_2, \cdots, m_n$ can be used instead of $m_2, \cdots, m_n$. We may make use of this property to obtain real eigenvector matrices, R, for parallel matrices. In the theory that follows it will be required that R is *orthogonal*. In Table 6.1 we show the Fourier matrix and one choice for the corresponding real orthogonal eigenvector matrix R for $n = 2$, $n = 3$ and $n = 4$. For diagonalization one has $\Lambda_P = XPX^{-1}$ where $X = F$ (Fourier matrix) or $X = R$ (real, orthogonal matrix).

For systems consisting of symmetrically interconnected subsystems, R may be used instead of F throughout this paper. Furthermore, for realization of controllers we require the use of the real matrix R, as will become clear below. When generalizing

the results of this paper to symmetric circulant plants, we have to impose further restrictions on R in addition to requiring it to be a real, orthogonal matrix. The choice of R for symmetric circulant matrices is explained in a later section.

Combinations of Parallel Matrices

If A and B are parallel matrices of the same dimension and k_i a scalar, then A^T, A^H, $k_1A + k_2B$, AB, $\sum_i k_i A^i$ are parallel matrices and A and B commute, that is, $AB = BA$. Note that A^{-1} is also a parallel matrix.

For example, if a both the process G and the controller C are parallel transfer function matrices, the sensitivity function $S = (I + GC)^{-1}$ and the complementary sensitivity function $H = I - S$ are both parallel matrices.

Matrices Consisting of Circulant Blocks

Consider a matrix M consisting of $m_1 \times m_2$ blocks, each block being a circulant matrix of order n. We shall call the class of such matrices $CB_{m_1,m_2,n}$. The many results from the theory of circulant matrices do not hold for matrices consisting of circulant blocks. However, we can find some results which will prove helpful.

- If M_1 and M_2 both belong to the class $CB_{m_1,m_2,n}$ and α_1 and α_2 are scalars, then $\alpha_1 M_1 + \alpha_2 M_2$ also belongs to the class $CB_{m_1,m_2,n}$ and M_1^H belongs to the class $CB_{m_2,m_1,n}$. If M_1 belongs to the class $CB_{m_1,m_2,n}$ and M_2 belongs to the class $CB_{m_2,m_1,n}$ then M_1M_2 belongs to the class $CB_{m_1,m_1,n}$.

- For "diagonalization" we have

$$\tilde{M}_1 = (I_{m_1} \otimes F_n) M_1 (I_{m_2} \otimes F_n)^H \tag{6.16}$$

where $\otimes$ denotes the Kronecker product, and $\tilde{M}_1$ is a matrix with the same block structure as M_1, each block in $\tilde{M}_1$ being the (diagonal) eigenvalue matrix of the corresponding block in M_1. This is illustrated by an example. Consider

$$M_1 = \begin{bmatrix} C_1 & C_2 & C_3 \\ C_4 & C_5 & C_6 \end{bmatrix} \tag{6.17}$$

where $C_1, C_2, \cdots, C_6$ all are circulant matrices of order n. M_1 then belongs to the class $CB_{2,3,n}$. We then have

$$M_1 = (I_2 \otimes F_n)^H \tilde{M}_1 (I_3 \otimes F_n) \tag{6.18}$$

$$(I_2 \otimes F_n)^H = \begin{bmatrix} F_n^H & 0 \\ 0 & F_n^H \end{bmatrix} \tag{6.19}$$

$$(I_3 \otimes F_n) = \begin{bmatrix} F_n & 0 & 0 \\ 0 & F_n & 0 \\ 0 & 0 & F_n \end{bmatrix} \tag{6.20}$$

$$\tilde{M}_1 = \begin{bmatrix} \Lambda_{C1} & \Lambda_{C2} & \Lambda_{C3} \\ \Lambda_{C4} & \Lambda_{C5} & \Lambda_{C6} \end{bmatrix} \tag{6.21}$$

where Λ_{Ci} is the (diagonal) eigenvalue matrix of block C_i.

6.4.2 Systems with MIMO Subsystems

Block Circulant and Block Parallel Matrices

The matrix C in Eq. (6.7) is *block circulant* if $c_1, c_2, \cdots, c_n$ all are blocks of dimension $n_o \times n_i$. For such a block circulant matrix C we can generate a block diagonal matrix

$$\tilde{C} = (F_n \otimes I_{n_o}) C (F_n \otimes I_{n_i})^H \tag{6.22}$$

where $\tilde{C} = \text{diag}\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$, and $\gamma_1, \gamma_2, \cdots, \gamma_n$ all have dimension $n_o \times n_i$ and can be calculated from the blocks of C using

$$\begin{bmatrix} \gamma_{C1} \\ \gamma_{C2} \\ \vdots \\ \gamma_{Cn} \end{bmatrix} = (\sqrt{n} F_n \otimes I_{n_o})^H \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{6.23}$$

Proof: Follows from the proof of Theorem 5.6.4 in [5], by setting $B_k = I_{n_o} A_{k+1} I_{n_i}$.

If $c_2 = c_3 = \cdots = c_n$ then we term the matrix C *block parallel*, and we have

$$\gamma_{C1} = c_1 + (n-1)c_2 \tag{6.24}$$

$$\gamma_{C2} = \gamma_3 = \cdots = \gamma_n = c_1 - c_2 \tag{6.25}$$

In this way, a symmetrically interconnected system C consisting of n units in parallel can be decomposed into one distinct subprocess γ_{C1} and $n-1$ equal subprocesses γ_{C2}.

Combinations of Block Circulant Matrices

If A is a block circulant matrix with $n \times n$ blocks, each of size $n_o \times n_i$ we have

- A^H and A^T are block circulant matrices with $n \times n$ blocks of size $n_i \times n_o$.

- If A^{-1} exists, it is a block circulant matrix.
- If B is block circulant, consisting of $n \times n$ blocks of size $n_i \times n_b$, then AB is a block circulant matrix with blocks of size $n_o \times n_b$.
- In general, block circulant matrices do not commute, $AB \neq BA$.

Matrices Consisting of Blocks which are Block Circulant

Consider a matrix N consisting of $m_r \times m_c$ blocks, each block being a block circulant matrix with $n \times n$ subblocks. Subblocks belonging to the same column of main blocks must have the same number of columns. Let n_i^c denote the number of columns of the subblocks in column c of main blocks. Likewise, subblocks belonging to the same row of main blocks must have the same number of rows, and we use n_o^r to denote the number of rows in the subblocks of blocks in row r of main blocks. To illustrate, consider

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{6.26}$$

Let M_{11}, M_{12} M_{21} and M_{22} all be block circulant matrices consisting of $n \times n$ subblocks, and let the subblocks of M_{ij} have dimension $n_o^{ij} \times n_i^{ij}$. Then $n_o^{11} = n_o^{12} = n_o^1$, $n_o^{21} = n_o^{22} = n_o^2$, $n_i^{11} = n_i^{21} = n_i^1$ and $n_i^{12} = n_i^{22} = n_i^2$, but n_o^1 may be different from n_o^2, and n_i^1 may be different from n_i^2. Introduce the matrices

$$\mathcal{F}_\mathcal{L} = \text{diag}\{F_n \otimes I_{n_o^r}\} \tag{6.27}$$

$$\mathcal{F}_\mathcal{R} = \text{diag}\{F_n \otimes I_{n_i^c}\} \tag{6.28}$$

We then have that N can be "diagonalized" by the following transformation

$$\tilde{N} = \mathcal{F}_\mathcal{L} N \mathcal{F}_\mathcal{R}^H \tag{6.29}$$

where $\tilde{N}$ is a matrix consisting of blocks which are block diagonal, where each block of $\tilde{N}$ can be calculated from the corresponding block of N using Eq. (6.23).

6.5 Realization of Full Block Parallel Controllers

In the following we will consider the use of *block parallel* controllers for the control of plants described by block parallel transfer function matrices. We assume that the controller design has resulted in that the distinct subprocess γ_{G1} of the plant G is controlled by the controller γ_{K1} and the $n-1$ identical subprocesses are controlled by $n-1$ controllers all equal to γ_{K2}, and that γ_{K1} and γ_{K2} both have dimension

$n_i \times n_o$. The controller K will then be a block parallel matrix with diagonal blocks $k_{ii} = [\gamma_{K1} + (n-1)\gamma_{K2}]/n$ and offdiagonal blocks $k_{ij} = [\gamma_{K1} - \gamma_{K2}]/n$.

In order to not increase the number of states in the controller unnecessarily, the controller should be realized as

$$K(s) = (R_n \otimes I_{n_i})^T \tilde{K}(s)(R_n \otimes I_{n_o}) \tag{6.30}$$

where $\tilde{K} = \text{diag}\{\gamma_{K1}, \gamma_{K2}, \cdots, \gamma_{K2}\}$ and R_n is the real eigenvector matrix for a parallel matrix of dimension $n \times n$ (recall Section 6.4.1 and Table 6.1). A block diagram for $K(s)$ is shown in Fig. 6.4.

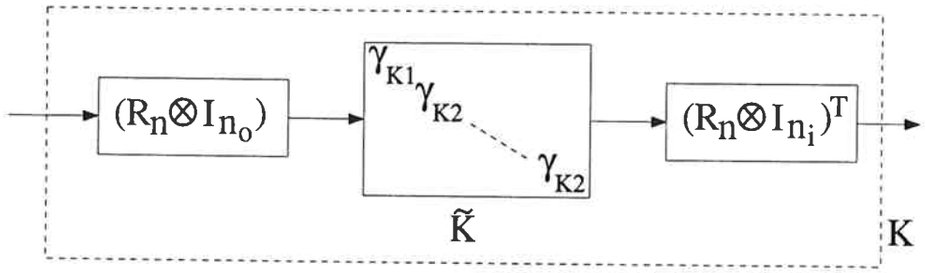


Figure 6.4: Realization of a full block parallel controller K.

6.6 H_∞ Control

In this section we will use H_∞ theory for analyzing robustness of control systems and for synthesizing controllers meeting predefined robustness criteria. *Throughout Sections 6.6 and 6.7 we will assume that the plant can be described by block parallel transfer function matrices.* We will also assume that all performance requirements are the same for all subsystems. The uncertainties are assumed to be the located in the same positions and to have the same magnitudes for the different subsystem, as the subsystems are designed to be nominally identical. *The uncertainty and performance weights are therefore also assumed to be block parallel matrices.*

All H_∞ syntheses, μ-calculations and μ-syntheses in this section are performed with MuSyn Inc.'s μ-Analysis and Synthesis Toolbox for MATLAB^{TM}.

6.6.1 Optimal H_∞ Control

H_∞ control theory can be used for designing controllers which ensures that the closed loop system satisfies singular value loop shaping specifications. For example, the stan-

dard "mixed sensitivity" H_∞ problem is to minimize

$$\|M\|_\infty = \left\| \begin{bmatrix} W_O H \\ W_P S \end{bmatrix} \right\|_\infty \tag{6.31}$$

where $S = (I + GK)^{-1}$ is the sensitivity function and $H = GK(I + GK)^{-1}$ the complementary sensitivity function. This corresponds to simultaneously trying to optimize robust stability with respect to *output* uncertainty and nominal performance.

Problem: In the following, consider the design of a controller in order to minimize the H_∞ norm of some matrix M (not necessarily of the form used in Eq. (6.31)). We assume that the plant $G(s)$ is a block parallel matrix consisting of n units in parallel, each block of $G(s)$ having dimension $n_o \times n_i$. Likewise, we assume that all weights are block parallel matrices with blocks of dimension compatible with the dimension of the blocks of $G(s)$. Furthermore, we assume that the matrix M consists of $m_r \times m_c$ blocks, each block of dimension $n \cdot n_o^r \times n \cdot n_i^c$ (e.g. for the matrix M in Eq. (6.31) we have $m_r = 2$ and $m_c = 1$).

Theorem 10 *The H_∞-optimal controller for this problem can be obtained by designing two H_∞-optimal controllers for the two systems corresponding to the "plants" γ_{G1} and γ_{G2}.*

Proof:

1. Express the matrix $M(s)$, whose H_∞ norm is to be minimized, as a Linear Fractional Transformation (LFT) of the controller $K(s)$ (see Fig. 6.5a).

$$M(s) = N_{11}(s) + N_{12}(s)K(s)\left[I - N_{22}(s)K(s)\right]^{-1} N_{21}(s) \tag{6.32}$$

   N_{11}, N_{12}, N_{21} and N_{22} consist of blocks which are block parallel.

2. Premultiplication or postmultiplication of $M(s)$ by unitary matrices will not change the singular values, and will thus leave the H_∞ norm unchanged (Fig. 6.5b). We use the matrices $\mathcal{F}_\mathcal{L}$ and $\mathcal{F}_\mathcal{R}$ defined in Eqs. (6.27) and (6.28)

$$\begin{aligned}
\tilde{M} &= \mathcal{F}_\mathcal{L} M \mathcal{F}_\mathcal{R}{}^H && (6.33)\\
&= \tilde{N}_{11} + \tilde{N}_{12}\tilde{K}\left[I - \tilde{N}_{22}\tilde{K}\right]^{-1}\tilde{N}_{21} && (6.34)\\
\tilde{N}_{11} &= \mathcal{F}_\mathcal{L} N_{11} \mathcal{F}_\mathcal{R}{}^H && (6.35)\\
\tilde{N}_{12} &= \mathcal{F}_\mathcal{L} N_{12}(F_n \otimes I_{n_i})^H && (6.36)\\
\tilde{N}_{21} &= (F_n \otimes I_{n_o}) N_{21} \mathcal{F}_\mathcal{L}{}^H && (6.37)\\
\tilde{N}_{22} &= (F_n \otimes I_{n_o}) N_{22}(F_n \otimes I_{n_i})^H && (6.38)
\end{aligned}$$

Thus, since N_{11}, N_{12}, N_{21} and N_{22} all consist of blocks which are block parallel, $\tilde{N}_{11}$, $\tilde{N}_{12}$, $\tilde{N}_{21}$ and $\tilde{N}_{22}$ all consist of blocks which are block diagonal, the first subblock in each block being distinct and the other $n-1$ subblocks equal.

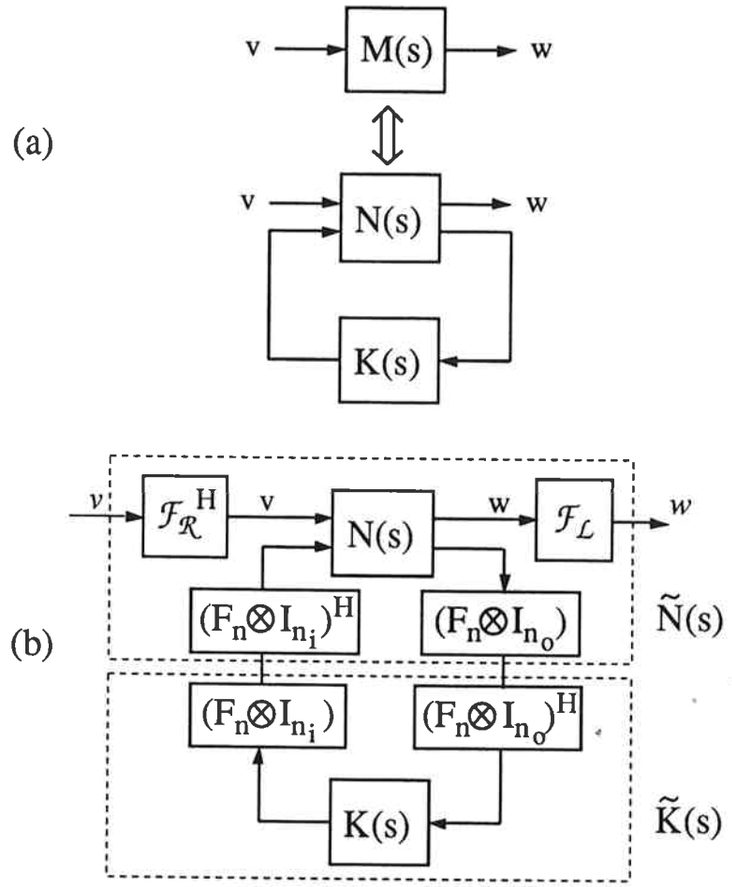


Figure 6.5: (a) Expressing $M(s)$ as a linear fractional transformation of the controller $K(s)$. (b) Pre- and postmultiplication of $M(s)$ by unitary matrices.

This implies that the controller design has been decoupled into n non-interacting design subproblems, one of these design subproblems being distinct, and $n-1$ design subproblems being equal. Thus, the controller $\tilde{K}$ can be designed by designing one controller for the distinct subproblem, and one controller for one of the $n-1$ subproblems. QED.

Remark 1: To see why the structure of $\tilde{N}$ given in Eqs. (6.33)-(6.38) implies that the controller synthesis problem can be decomposed into n non-interacting synthesis

subproblems, consider Fig. 6.6. The matrix at the top of Fig. 6.6 may represent $\tilde{N}$

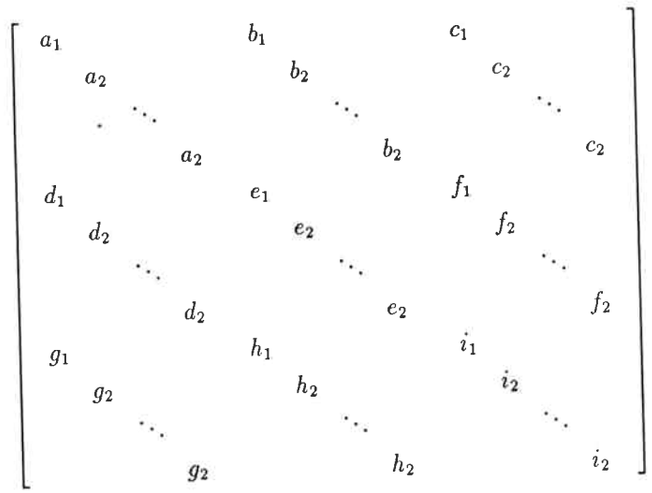


Figure 6.6: Top: $\tilde{N}$ for a case with 3×3 main blocks. Bottom: $\tilde{N}$ permuted to have the n independent synthesis subproblems along the main diagonal.

for a case where $\tilde{N}$ has 3×3 main blocks. After permutations (row and column interchanges) we get the matrix at the bottom of Fig. 6.6, from which it is apparent that the controller design problem consists of n independent subproblems, one distinct and $n-1$ identical.

Remark 2: The same also holds for the H_2-optimal problem, since the Frobenius

norm is also unitarily invariant.

Remark 3: The controller $\tilde{K}$ in Fig. 6.5 will be block diagonal, the first block on the diagonal being distinct and the $n-1$ other blocks equal. Consequently,

$$K(s) = (F_n \otimes I_{n_i})^H \tilde{K}(s)(F_n \otimes I_{n_o}) = (R_n \otimes I_{n_i})^T \tilde{K}(s)(R_n \otimes I_{n_o}) \tag{6.39}$$

will be a block parallel matrix. We see that instead of designing a $n \cdot n_i \times n \cdot n_o$ dimensional controller for a system with a $n \cdot n_o \times n \cdot n_i$ dimensional plant, we can design two controllers of dimension $n_i \times n_o$, each controller designed for a system with a plant of dimension $n_o \times n_i$.

Remark 4: Note that this corresponds to optimizing the H_∞ objective for both the systems corresponding to γ_{G1} and γ_{G2}. If $m_c = n_i = 1$ or $m_r = n_o = 1$, *all* directions in the H_∞ criterion are optimized. In contrast, the controller Doyle et al. [9] terms the "central solution" to the H_∞ synthesis problem only optimizes the *worst* direction in the overall H_∞ criterion. In general, the solution to the H_∞ controller synthesis problem is non-unique [9], since many controllers will achieve the optimum H_∞ norm in the worst direction, while doing equally well or better in the other directions. How to minimize the peak values of the singular values corresponding to directions other than the worst direction is a line of research, called "super-optimal" H_∞ control [10]. We have here found a class of problems where the solution to the super-optimal H_∞ control problem is very simple.

Remark 5: The interconnection matrix $\tilde{N}(s)$ in Fig. 6.5b has the same number of states as $N(s)$ in Fig. 6.5a. The number of states in the controller $\tilde{K}$ will therefore equal the number of states in a controller based on regular H_∞ synthesis. That is, for this class of problems super-optimality does not require a controller with a higher number of states.

Example 3. Consider the process

$$G(s) = \frac{1}{(20s+1)(100s+1)} \begin{bmatrix} -0.25s+1 & 0.75s \\ 0.75s & -0.25s+1 \end{bmatrix} \tag{6.40}$$

with eigenvalues

$$\lambda_{G1}(s) = \frac{0.5s+1}{(20s+1)(100s+1)}; \quad \lambda_{G2}(s) = \frac{-s+1}{(20s+1)(100s+1)} \tag{6.41}$$

We want to minimize the H_∞ norm of M in Eq. (6.31), with weights $W_O(s) = w_O(s)I$: $w_O(s) = 0.2\frac{5s+1}{0.5s+1}$ and $W_P(s) = w_P(s)I$; $w_P(s) = 0.5\frac{2s+1}{2s}$. Using the approach outlined above we may solve this problem by considering the two SISO "plants" $\lambda_{G1}(s)$ and

$\lambda_{G2}(s)$. H_∞ synthesis for these two "plants" gave H_∞ norms of 0.56 and 0.89, respectively. State space descriptions of the resulting "controllers" λ_{K1} and λ_{K2} are given in the Appendix. The "controllers λ_{K1} and λ_{K2} may be combined into a regular controller with eight states using Eq. (6.30). Conventional H_∞ synthesis for the overall system also gave a controller with eight states which achieved a H_∞ norm of 0.89. However, in this case the peak of the singular value corresponding to the "good" direction was 0.73 (for the central solution), whereas our controller gave 0.56. In Fig. 6.7 is shown the responses to setpoint changes for the two cases. We also show the response for an inverse-based controller with gives peak values of 0.89 for *both* singular values (see section on Inverse-based controllers below). When the setpoint enters in the "bad" direction, the three responses are indistinguishable, whereas when the setpoint enters in the "good" direction it is clear that the controller synthesized using our method is superior and the inverse-based controller is worst.

Example 4. Consider the 8×12 plant

$$G(s) = \begin{bmatrix} G_1(s) & G_2(s) & G_2(s) & G_2(s) \\ G_2(s) & G_1(s) & G_2(s) & G_2(s) \\ G_2(s) & G_2(s) & G_1(s) & G_2(s) \\ G_2(s) & G_2(s) & G_2(s) & G_1(s) \end{bmatrix} \tag{6.42}$$

where $G_1(s)$ and $G_2(s)$ are transfer function matrices with two outputs ($n_o = 2$) and three inputs ($n_i = 3$) with state space realizations $G_l(s) = C_l(sI - A_l)^{-1}B_l + D_l$, $l = 1, 2$, and

$$A_1 = A_2 = \begin{bmatrix} -0.05 & 0 & 0 \\ 0 & -0.1 & 0 \\ 0 & 0 & -0.2 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 2 & 3 & 9 \\ 6 & 9 & 4 \\ 8 & 0 & 1 \end{bmatrix}; \quad B_2 = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 1 & 0 \\ -2 & 1 & 0 \end{bmatrix}$$

$$C_1 = C_2 = \begin{bmatrix} 0.35 & 0.60 & 0.20 \\ 0.45 & -0.20 & 2.40 \end{bmatrix}$$

$$D_1 = D_2 = \begin{bmatrix} -1.25 & 0 & 0 \\ 0 & -1.25 & -1.25 \end{bmatrix}$$

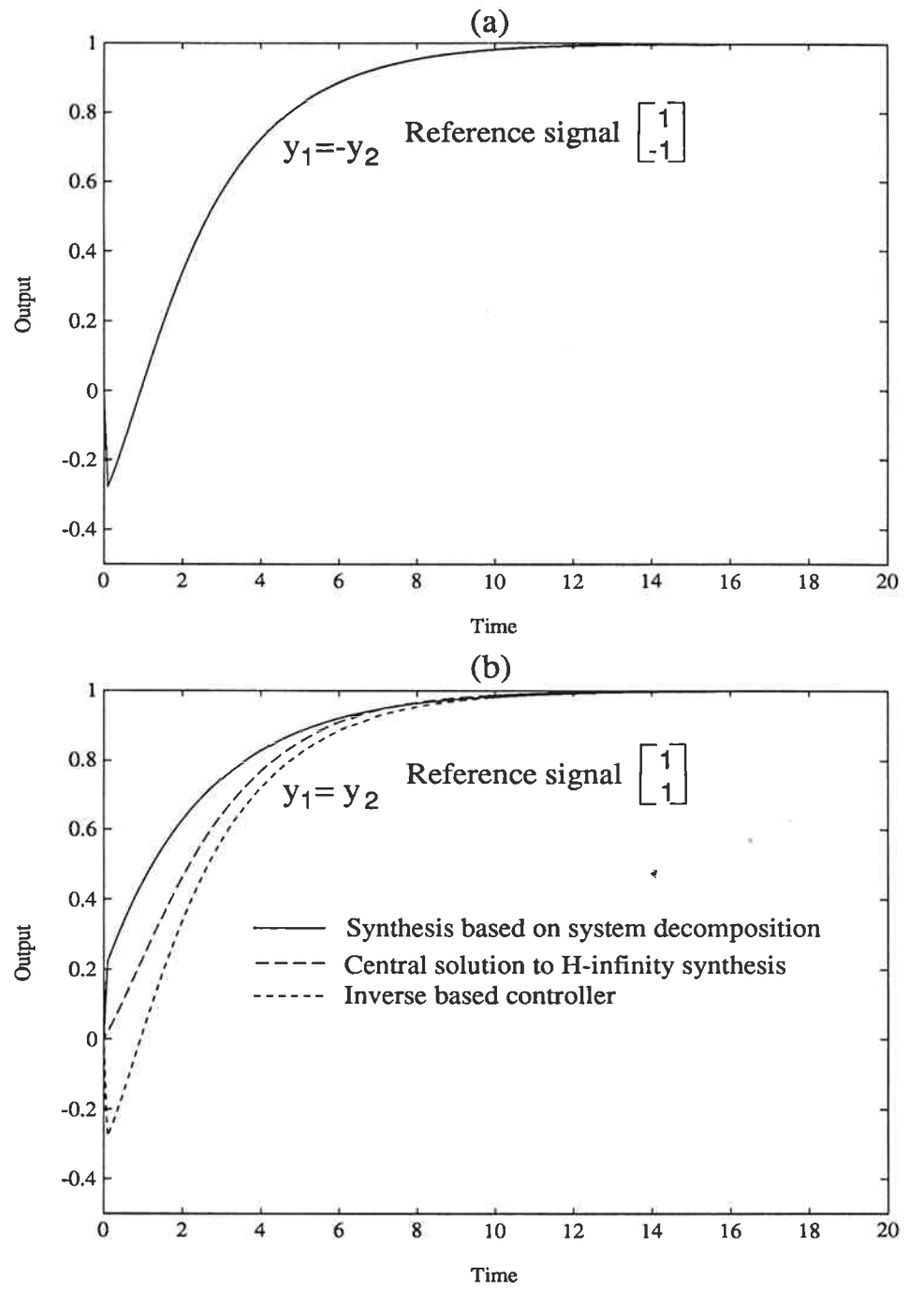


Figure 6.7: Response of the closed loop system to setpoint changes entering in the "bad" (a) and "good" (b) directions in Example 3.

The design criterion is to minimize the infinity-norm of

$$M = \begin{bmatrix} W_o GK(I+GK)^{-1} \\ W_p(I+GK)^{-1} \\ W_u K(I+GK)^{-1} \end{bmatrix} \tag{6.43}$$

with weights

$$\begin{aligned} W_o &= 0.2\frac{4s+1}{0.4s+1}I_8 \\ W_p &= \text{diag}\left\{ \begin{bmatrix} 0.5\frac{2.5s+1}{2.5s} & 0 \\ 0 & 0.5\frac{0.3s+1}{0.3s} \end{bmatrix} \right\} \\ W_u &= 0.1I_{12} \end{aligned}$$

We decompose the plant $G(s)$ into γ_{G1} and $\gamma_{G2} = \gamma_{G3} = \gamma_{G4}$, each of dimension 2×3, using Eq. (6.22), and design one 3×2 controller γ_{K1} for the system corresponding to γ_{G1}, and one 3×2 controller γ_{K2} for the system corresponding to γ_{G2}. For both these design subproblems a H_∞-norm of 0.91 was achieved, and the same H_∞-norm of 0.91 was achieved for the overall system after calculating the controller K from γ_{K1} and γ_{K2} according to Eq. (6.30). The best value of the H_∞-norm achieved when attempting H_∞ synthesis on the overall plant was 0.99. This demonstrates weaknesses in the synthesis software we have available[2], and that controller synthesis becomes simpler when the system is decomposed into problems with fewer states and of lower dimension. H_∞ synthesis for the overall plant gives an interconnection matrix with 28 states, whereas the number of states in each synthesis subproblem is 7 after decomposition of the plant. The number of states in the final controller is 28 for both cases. State space descriptions of the "controllers" γ_{K1} and γ_{K2} are given in the Appendix.

6.6.2 Inverse-based Controllers

Consider a parallel plant $G(s)$ consisting of SISO subsystems, and the special case when the following conditions hold:

C1. λ_{G1} and λ_{G2} have the same RHP zeros.

C2. All weights are scalar times identity matrices.

C3. G and K only appear as products of each other in the problem statement (as in Eq. (6.31))

[2]The μ-tools and Robust Control toolboxes for MATLAB.

C4. The plant eigenvalues have the same pole excess.

Then if we design a controller λ_{K1} for the plant eigenvalue λ_{G1}, the H_∞ norm that was obtained for the SISO "system" corresponding to λ_{G1} can be obtained for the overall system by choosing

$$\lambda_{K2} = \cdots = \lambda_{Kn} = \lambda_{K1}\lambda_{G1}/\lambda_{G2} \tag{6.44}$$

The result will be an inverse-based controller of the type

$$K(s) = k(s)G^{-1}(s) \tag{6.45}$$

thus effectively transforming the $n \times n$ H_∞ design problem to n identical SISO problems. Condition C1 ensures that any RHP zeros in the plant eigenvalues cancel in Eq. (6.44), such that the controller is stable. If λ_{G2} has a larger pole excess than λ_{G1}, calculating λ_{K2} from Eq. (6.44) may result in an improper controller which is impossible to realize. Condition C4 ensures that this is not a problem. The conditions C1-C4 ensure that the same H_∞ norm is achievable for the subproblems corresponding to λ_{G1} and λ_{G2}. Thus, the H_∞ controller will be unique if the solution to the SISO subproblems corresponding to λ_{G1} and λ_{G2} are unique. Whereas condition C2 will normally hold for SISO subsystems in parallel, condition C3 may well be violated, e.g. if M contains a term like $W_u K(I+GK)^{-1}$ corresponding to a bound on the closed loop transfer function from reference signal to manipulated variables.

There may exist an inverse-based controller achieving the optimal H_∞ norm also when conditions C1-C4 are not fulfilled (or when the plant consists of MIMO subsystems), and in many cases this inverse-based controller can be found by first synthesizing the controller corresponding to the most difficult system direction and then use Eq. (6.44) to find the controller corresponding to the other direction. For such cases the H_∞-optimal controller will not be unique.

Example 5. To illustrate, consider Example 2 in Section 2, with four reactors in parallel ($n = 4$), and choose the values $k = 1$, $\tau = 100$ and $a = 0.7$. From Eq. (6.11) and Eq. (6.12) we then have that the plant eigenvalues are

$$\lambda_{G1} = \frac{3.1}{100s+1}; \quad \lambda_{G2} = \frac{0.3}{100s+1}$$

We use M as in Eq. (6.31), with weights $W_P = w_P I$; $\{w_P = 0.5\frac{10s+1}{10s}\}$ and $W_O = w_O I$; $\{w_O = 0.2\frac{5s+1}{0.5s+1}\}$. Thus conditions C1-C4 are all fulfilled. The magnitude of the weights w_P and w_O are shown together with the magnitude of the plant eigenvalues (=singular values) in Fig. 6.8.

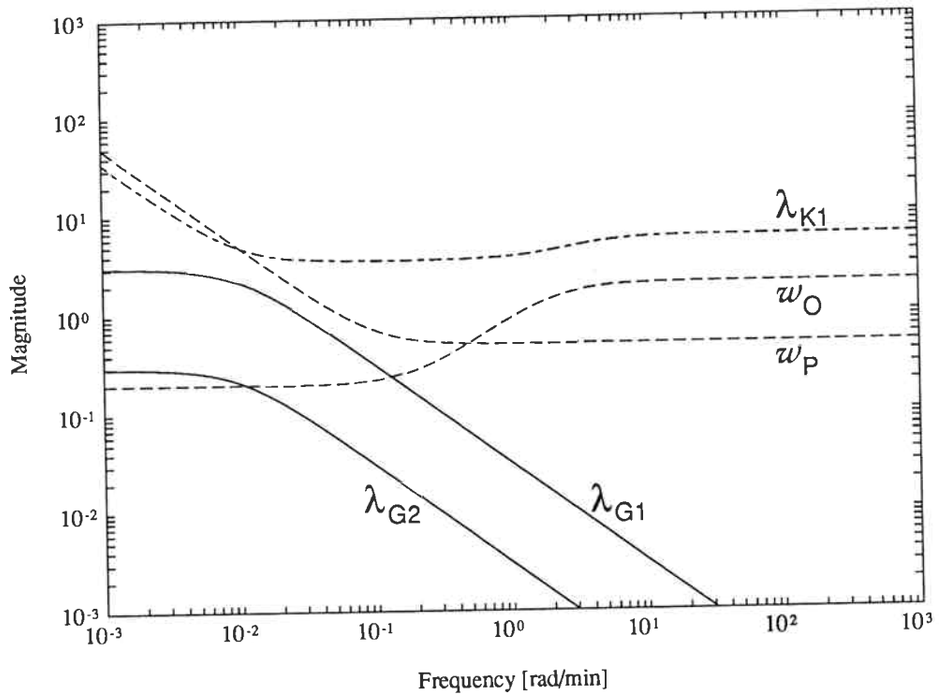


Figure 6.8: Magnitudes of plant eigenvalues λ_{G1} and λ_{G2}, performance weight w_P, output weight w_O and first controller eigenvalue λ_{K1} for Example 5.

An H_∞-optimal controller λ_{K1} was designed for the SISO "plant" λ_{G1} according to Eq. (6.31) (with weights w_P and w_O substituted for W_P and W_O, respectively), achieving a H_∞ norm of 0.50. The magnitude of λ_{K1} is also shown in Fig. 6.8, and a state space description is given in the Appendix. After calculating λ_{K2} from λ_{K1} according to Eq. (6.44), and combining the two controllers λ_{K1} and λ_{K2} according to Eq. (6.30) to find the controller K for the full plant, we found that the same H_∞ norm was achieved for the full system as for the SISO case. Furthermore, we found numerically that the controller K was the same as the controller resulting from H_∞ synthesis for the full 4×4 plant.

Example 3 continued. We see that condition C1 above is violated, as λ_{G2} has a zero at $s = 1$, whereas λ_{G1} has no RHP zero. Calculating λ_{K2} from λ_{K1} using Eq. (6.44) gave an internally unstable closed loop system. Stability of the closed loop system was achieved when calculating λ_{K1} from λ_{K2} using Eq. (6.44), but in this case the peak value of the singular value was 0.89 in *both* directions.

6.7 μ-optimal Control

Consider the feedback system in Fig. 6.9. Assuming that M and Δ are stable, the structured singular value [6], μ, is defined such that $\mu^{-1}(M)$ is the smallest value of the maximum singular value of Δ for which the system $(I + M\Delta)^{-1}$ can become unstable. Note that the value of $\mu(M)$ depends on the *structure* of the perturbation block Δ. In Fig. 6.10 is shown an example of a system that can be rewritten in

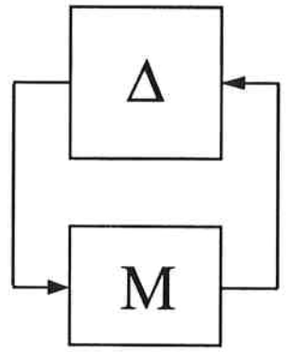


Figure 6.9: M-Δ structure for μ calculation.

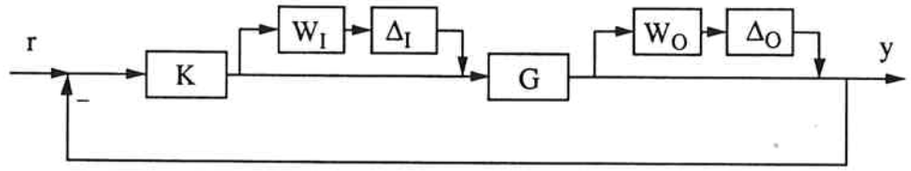


Figure 6.10: Block diagram of plant with uncertainties.

the M-Δ structure of Fig. 6.9 (after neglecting external input r and external output y). Δ_I and Δ_O are blocks representing the uncertainties in the system, and for this example $\Delta = \text{diag}\{\Delta_I, \Delta_O\}$. The weights W_I and W_O normalize the uncertainties such that Δ_I and Δ_O have a maximum singular value less than or equal to unity at all frequencies. Thus the system will remain stable for any allowed perturbation provided the system is nominally stable (i.e. the system is stable if the uncertainty blocks are removed) and $\mu(M) < 1$. This property is termed Robust Stability (RS). Robust Performance (RP) may also be addressed in the μ framework, by considering an equivalent stability problem of increased dimension. We will use a performance specification of the type $\bar{\sigma}(W_P S_p) < 1, \forall\omega$, where S_p is the worst sensitivity function made possible by the uncertainties and W_P is a performance weight. This performance specification is included in the μ framework by closing the loop from offset to reference

signal with the performance weight W_P and a *full* perturbation block Δ_P, and using $\Delta = \text{diag}\{\Delta_I, \Delta_O, \Delta_P\}$.

It is possible to calculate $\mu(M)$ exactly only in some special cases, but reasonably tight upper and lower bounds are readily available. We will use the following properties of μ:

$$\rho(M) \leq \mu(M) \leq \bar{\sigma}(M) \tag{6.46}$$

$$\mu(M) \leq \bar{\sigma}(D_l M D_r^{-1}) \tag{6.47}$$

Eq. (6.46) holds for any complex valued perturbation Δ, but μ may be lower than the lower limit for real perturbations. D_l and D_r are real positive matrices with a structure such that $D_r^{-1}\Delta D_l = \Delta$. For example, $D = dI$ if Δ is a full matrix, and D is a full matrix if $\Delta = \delta I$ (repeated scalar uncertainty). If all blocks in Δ are square, $D_l = D_r$. The upper bound on the value of μ in Eq. (6.47) is usually quite tight [7], and the standard D-K iteration procedure for μ-optimal controller synthesis used in MATLAB involves finding the controller for which the upper bound in Eq. (6.47) is minimized [8]. D-K iteration involves two steps:

Step 1: Find D-scales (D_l, D_r) by computing the upper bound on μ in Eq. (6.47).

Step 2: Scale the controller design problem with the D-scales found in Step 1, and design an H_∞-optimal controller for the scaled design problem.

Although convergence is not guaranteed D-K iteration appears to work well [8]. Note that although the magnitude bound for any given uncertainty is the same for all subsystems, the uncertainties do allow the individual subsystems to differ from each other. Furthermore, we put no particular restriction on the location or type of the uncertainties—additive, multiplicative, inverse multiplicative uncertainty blocks and any other type of uncertainty description which can be recast into the M-Δ structure in Fig. 6.9 can be handled.

6.7.1 Robust Stability for SISO Subsystems with One Uncertainty Block for Each Subsystem

In this section we consider RS with SISO subsystems and only one source of uncertainty in each subsystem, and the perturbation block Δ and the interconnection matrix M are both square matrices of dimension $n \times n$.

μ Analysis: The interconnection matrix M will be a parallel matrix. We know from Section 6.4.1 that for parallel matrices the magnitude of the eigenvalues equal the

magnitude of the singular values, and from Eq. (6.46) we see that the value of μ will equal the spectral radius. Thus for this problem the value of μ is completely insensitive to the structure of the uncertainty for any complex uncertainty, we get the same value for μ regardless of whether the uncertainty block is full, diagonal or is a repeated scalar block.

μ Synthesis: The observation that the value of μ is insensitive to the structure of the uncertainty can be used to simplify the design of a robust controller. We get identical results if we choose Δ to be a full matrix, and the solution is therefore identical to the H_∞ solution considered previously. The controller design can be performed by designing controllers for two SISO systems, one SISO system for λ_{G1} of the plant, and another SISO system for $\lambda_{G2}, \cdots, \lambda_{Gn}$. Note that synthesizing a controller for minimizing μ for RS is only meaningful if the system (including uncertainties) can become open loop unstable, as otherwise $K = 0$ is an optimal choice.

6.7.2 μ Analysis and Synthesis for Systems with More Than One Uncertainty Block

For systems with more than one uncertainty block, or for systems consisting of MIMO subsystems, the interconnection matrix M will in general not be parallel (or circulant), but it will be a matrix consisting of blocks which are block parallel (recall Section 6.4.2). For matrices consisting of blocks which are block parallel, the magnitudes of the eigenvalues and the singular values will in general differ, and the upper and lower limits in Eq. (6.46) will not be particularly helpful. However, the structure of the problem is such that significant simplifications to analysis and synthesis for the upper bound on μ can be achieved. Henceforth we will therefore only consider the upper bound on μ in Eq. (6.47), that is, consider the problem

$$\min_K(\inf_{\mathbf{D}} \bar{\sigma}(D_l M D_r^{-1})) = \inf_{\mathbf{D}}(\min_K \bar{\sigma}(D_l M D_r^{-1})) \tag{6.48}$$

where $\mathbf{D}$ denotes the allowable set of matrices D_l and D_r. The "inner problem" on the right of Eq. (6.48), $\min_K \bar{\sigma}(D_l M D_r^{-1})$, is an H_∞ problem, and we shall see that this H_∞ problem is similar to those studied in Section 6.6.1. Below we shall consider D_l and D_r fixed and consider $\min_K \bar{\sigma}(D_l M D_r^{-1})$.

We start from the conventional M-Δ block structure of Fig. 6.9, and express M as an LFT of the controller K in the same way as in Fig. 6.5a. Clearly, an identity matrix can be inserted anywhere in the block diagram without altering it. We therefore insert $I = \mathcal{F}_\mathcal{R}{}^H \mathcal{F}_\mathcal{R}$ and $I = \mathcal{F}_\mathcal{L}{}^H \mathcal{F}_\mathcal{L}$ (if M is not a square matrix, the dimensions of these two

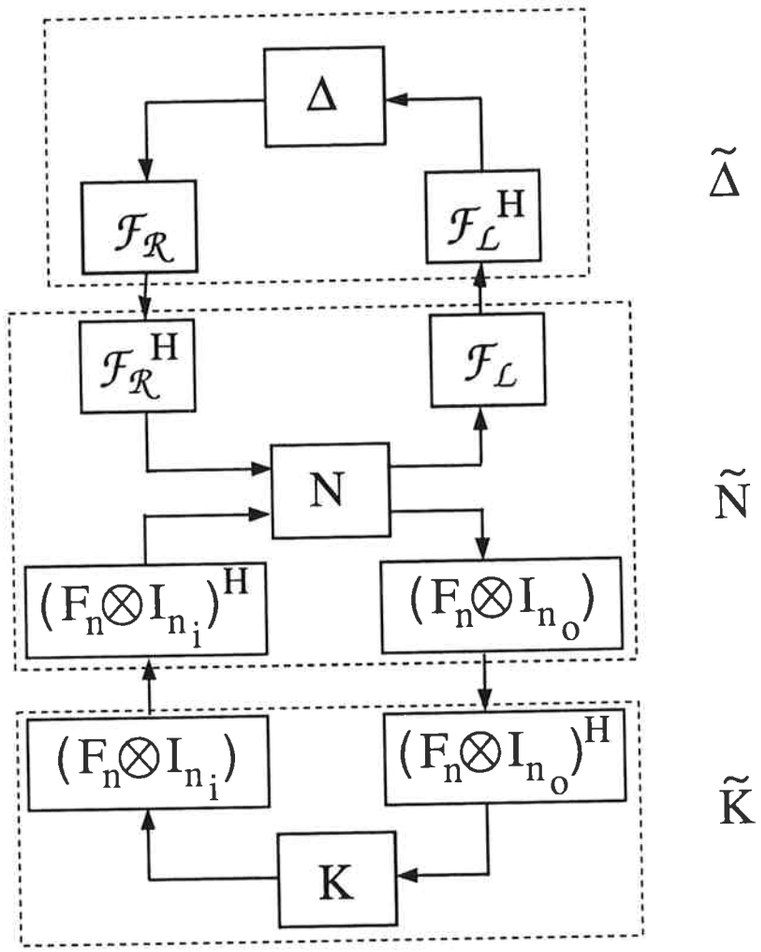


Figure 6.11: Expressing M as an LFT of K, and transforming to obtain $\tilde{\Delta}$, $\tilde{N}$ and $\tilde{K}$.

identity matrices will differ), $I = (F_n \otimes I_{n_i})^H (F_n \otimes I_{n_i})$, and $I = (F_n \otimes I_{n_o})^H (F_n \otimes I_{n_o})$ in the block diagram, see Fig. 6.11. Thus, instead of studying M and Δ, we can study $\tilde{N}$, $\tilde{K}$ and $\tilde{\Delta} = \mathcal{F}_\mathcal{R} \Delta \mathcal{F}_\mathcal{L}^H$. $\tilde{N}$ has the same structure as in Section 6.6, and thus consists of n independent design subproblems, one distinct and one repeated $n-1$ times. Repeated scalar blocks in Δ will remain as repeated scalar blocks in $\tilde{\Delta}$, and full blocks in Δ will remain as full blocks in $\tilde{\Delta}$. However, independent, diagonal blocks in Δ will become full blocks in $\tilde{\Delta}$[3]. This loss of structure in the uncertainty description is the

[3]The full blocks of $\tilde{\Delta}$ which correspond to independent, diagonal blocks of Δ, will have a certain structure which cannot be utilized in the μ framework.

price we have to pay for the resulting simplifications in μ analysis and synthesis. Note, however, that if Δ does not contain any independent diagonal perturbation blocks, we have introduced no conservatism by transforming to the $\tilde{M}$-$\tilde{\Delta}$ structure. We show next how transforming to the $\tilde{M}$-$\tilde{\Delta}$ structure simplifies μ analysis and synthesis.

1. Consider first the case when $\tilde{\Delta}$ contains *only* full blocks. The corresponding scaling matrices D_l and D_r are then diagonal matrices of the type $D = \text{diag}\{d_i I_i\}$ where I_i is an identity matrix of dimension corresponding to the dimension of the ith block of $\tilde{\Delta}$. Express $D_l \tilde{M} D_r^{-1}$ as an LFT of $\tilde{K}$, and call the resulting interconnection matrix $\tilde{N}_D$. We see that $\tilde{N}_D$ has the same structure as $\tilde{N}$; $\tilde{N}_D$ consists of blocks which are block diagonal, the first subblock on the diagonal of each block being distinct and the $n-1$ other subblocks being identical. We therefore find that $\tilde{N}_D$ consiste of n independent design subproblems, one distinct and one repeated $n-1$ times. For clarity, recall that permutations do not change singular values, and recall Fig. 6.6. The matrix at the top of Fig. 6.6 has the same structure as $\tilde{N}_D$.

2. Now consider the case when $\tilde{\Delta}$ contains both full and repeated scalar perturbation blocks. We want to show that, as in the full block case, $\tilde{N}_D$ consists of only two distinct design subproblems, one distinct subproblem and one repeated $n-1$ times. To see this, let $\tilde{\Delta}$ have the structure $\tilde{\Delta} = \text{diag}\{\tilde{\Delta}_f, \tilde{\Delta}_r\}$, where $\tilde{\Delta}_f$ contains the full block uncertainties and $\tilde{\Delta}_r$ contains the repeated scalar blocks. Now, absorb the repeated scalar blocks $\tilde{\Delta}_r$ into $\tilde{M}$ to give the structure M_f-$\tilde{\Delta}_f$. Thereafter, apply the D-scales D_{fl} and D_{fr} corresponding to $\tilde{\Delta}_f$ to $\tilde{M}_f$, thus obtaining $D_{fl} \tilde{M}_f D_{fr}^{-1}$. Expressing $D_{fl} \tilde{M}_f D_{fr}^{-1}$ as an LFT of $\tilde{K}$ we get $\tilde{N}_{Df}$. Although $\tilde{N}_{Df}$ includes the repeated scalar perturbation blocks, it will be a matrix consisting of blocks which are block diagonal, the first subblock on the diagonal of each main block being distinct and the $n-1$ other subblocks being identical. We have therefore found that N_{df} also consists of only two distinct design subproblems, one distinct and one repeated $n-1$ times. This is true irrespective of the particular values of $\tilde{\Delta}_r$, and the same must therefore hold for the interconnection matrix $\tilde{N}_D$ corresponding to $D_l \tilde{M} D_r^{-1}$.

For the calculation of the upper bound on $\mu(\tilde{M})$, we only need the two first subblocks on the diagonal of each main block of $\tilde{M}$. However, unlike the H_∞ case, we cannot consider γ_{G1} and γ_{G2} independently, because of the full uncertainty blocks (i.e. the values corresponding to the full uncertainty blocks in D_l and D_r must be the same for γ_{G1} and γ_{G2}). This follows from Eq. (6.48) where we see that for the full uncertainty

blocks the minimization over **D** gives the same "D-scales" for all subproblems. If there is only one full block, the subproblems may be considered independently, as the D-scale corresponding to the full block can be normalized to one for all subproblems. Of course, when there is only repeated scalar uncertainties the subproblems may also ber considered independently.

To recapitulate, we have found that if we approximate any independent diagonal uncertainty blocks with full uncertainty blocks, we obtain the following results:

Step 1 in D-K Iteration (μ analysis). We can calculate the upper bound on μ in Eq. (6.47) and obtain the D-scaling matrices D_l and D_r by considering a block diagonal plant $\tilde{G} = \text{diag}\{\gamma_{G1}, \gamma_{G2}\}$ instead of the full plant G.

Step 2 in D-K Iteration (H_∞ synthesis for the scaled problem). The controller synthesis problems for γ_{G1} and γ_{G2} can be considered independently (the scaled problem will also fall within the class of problems considered in the section on H_∞-optimal control).

That is, Step 2 in D-K iteration may be solved as two independent problems, while in Step 1 we have to combine the two controllers γ_{K1} and γ_{K2} before finding $\mu(M)$ and the D-scales. However, note that the size of the problem is also reduced in Step 1 relative to a regular μ optimization where the block parallel nature of the problem is not taken into account. If we have a large number of subsystems, the size of the μ synthesis and analysis problems can thus be reduced dramatically.

If and when D-K iteration converges, the full controller K can thereafter be found from $\tilde{K}$ using Eq. (6.30). The value of μ for the interconnection matrix M found from G and K is guaranteed to be equal to or lower than the value of μ achieved in the controller synthesis[4].

Remark 1. Intuitively, it is not surprising that the we can keep the structure of the repeated scalar uncertainty blocks when we consider $\tilde{G} = \text{diag}\{\gamma_{G1}, \gamma_{G2}\}$ instead of G, as repeated scalar uncertainty is the only type of uncertainty which can not ruin the block parallel structure of G.

Remark 2. When the subsystems of the plant have multiple inputs and multiple outputs, the uncertainty weights must be chosen carefully to try to minimize the conservativeness introduced by assuming full uncertainty blocks, as scaling problems must otherwise be expected.

[4]We may have introduced conservatism in the design by approximating independent diagonal uncertainty blocks by full uncertainty blocks. $\mu(M)$ may therefore be lower than the μ-value achieved in the controller synthesis.

To illustrate this problem, consider a hypothetical system consisting of subsystems with two uncertain inputs. The relative uncertainty in each input is 10% and the two inputs are independent. The range of operation for input u_1 is $0-10$ and the range of operation for input u_2 is $0-1$. This uncertainty can be modeled using $\tilde{u} = (I_2 + 0.1I_2\Delta)u$, where Δ is a diagonal 2×2 perturbation matrix, $\bar{\sigma}(\Delta) \leq 1$. However, to be able to simplify analysis and synthesis in the way described above, we are forced to make Δ a full 2×2 matrix. Considering the ranges of operation of the two inputs, the uncertainty should then be modeled as

$$\tilde{u} = (I_2 + \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix} \Delta \begin{bmatrix} 0.1 & 0 \\ 0 & 1 \end{bmatrix})u \tag{6.49}$$

Thus, the fictitious uncertainty going from u_1 to $\tilde{u}_2$ is no larger than the physically motivated uncertainty between u_2 and $\tilde{u}_2$.

Remark 3. For some problems the assumption of full uncertainty blocks can be conservative, as some uncertainty blocks may be diagonal. However, it will in many cases be worthwhile to initially synthesize a controller assuming full uncertainty blocks, as the reduction in system size and in the number of independent uncertainty blocks makes the synthesis task much simpler. The controller thus found can be used as a starting point for the synthesis of a controller using the actual uncertainty structure, if additional improvement in robust performance is necessary.

Example 6. Consider the same process as in Example 5, but with input and output uncertainty as shown in Fig. 6.10. We also included a performance specification expressed as $\bar{\sigma}(W_P S_p) < 1, \forall\omega$. The output uncertainty Δ_O and input uncertainty Δ_I are both assumed to be diagonal, with uncertainty weights $W_O(s) = \text{diag}\{0.2\frac{2.5s+1}{0.25s+1}\}$ and $W_I(s) = \text{diag}\{0.2\frac{5s+1}{0.5s+1}\}$. We also include the performance specification as $\bar{\sigma}(W_P S_p) < 1, \forall\omega$. The performance specification is equivalent to closing the loop from offset to reference signal with a full perturbation block ("uncertainty block") Δ_P and the performance weight $W_P(s)$. The performance weight used here is $W_P(s) = \text{diag}\{0.5\frac{10s+1}{10s}\}$. Thus, without taking into account the parallel nature of the problem, for this problem we have two diagonal 4×4 uncertainty blocks and one full 4×4 performance block, and we get a 12×12 μ interconnection matrix:

$$M = \begin{bmatrix} -W_I KG(I+KG)^{-1} & -W_I K(I+GK)^{-1} & W_I K(I+GK)^{-1} \\ W_O G(I+KG)^{-1} & -W_O GK(I+GK)^{-1} & W_O GK(I+GK)^{-1} \\ -W_P G(I+KG)^{-1} & -W_P (I+GK)^{-1} & W_P (I+GK)^{-1} \end{bmatrix} \tag{6.50}$$

Using our method we may design a parallel controller, by assuming all the uncertainty blocks to be full and reducing the system to get one 6×6 μ interconnection matrices for

Step 1 in D-K iteration and two 3×3 H_∞ problems for Step 2 in D-K iteration. Using this procedure we were able find a controller resulting in a μ-value of 0.93. The state space representation of the eigenvalues of this controller are given in the Appendix.

Thereafter we tried to improve the controller design by using the true structure for the uncertainties, with the input and output uncertainties being diagonal blocks of dimension 4×4 and using M in Eq. (6.50). However, we found that this increased the complexity of the controller synthesis problem so much that we were unable to improve the design. The best controller we were able to find gave a μ-value of 0.96. This result shows that there are numerical problems as the μ-optimal controller should obviously give a μ-value of 0.93 or lower.

6.8 Comparison with Previous Work

Lunze [13] studies robust stability of symmetrically interconnected systems. Lunze makes full use of the structure of the system for assessing nominal stability of the overall system. However, he does not make use of any structure in the uncertainty description, as he uses independent element-by-element uncertainties, with a common magnitude bound on the uncertainty for all elements. He then rewrites the system in a structure similar to the M-Δ structure of Section 6.7, and his robustness criterion is that the Perron root of $M\Delta$ should be less than one. This method can be very conservative. This will be illustrated by considering the example in Section 7.1 of [13].

Example 7. In this example we study the stability of a multi-area power system with different nominal values of power generation ([13], Section 7.1). The block diagram of the system is shown in Fig. 6.12, where $g(s)$ is the scalar transfer function of an individual power station, the block diagram of which is shown in Fig. 6.13. Thus, all the individual power stations are assumed to have the same dynamics. Furthermore, all the power stations are assumed to have the same amount of rotating mass, T. $T_0 = nT$ is the total amount of rotating mass in the system. For further information about the model we refer to [13] and the references contained therein. The parameter k_{si} is adjusted proportionally to the nominal value of power generation of the ith power station, $k_{si} = -15$ corresponds to a 200 MW power station. The Δ-blocks in Fig. 6.12 represent this adjustment of the parameters k_{si}, we thus have real uncertainties. $k_{li} = -2.5$ is used for all subsystems, and $T = 15.3$ is assumed, as this value of T agrees with the intermediate results of Lunze (the value $T = 16.25$ which Lunze claims to use, does not agree with his intermediate results). Furthermore, the state space representation of the subsystems used by Lunze does not agree with the transfer function model given

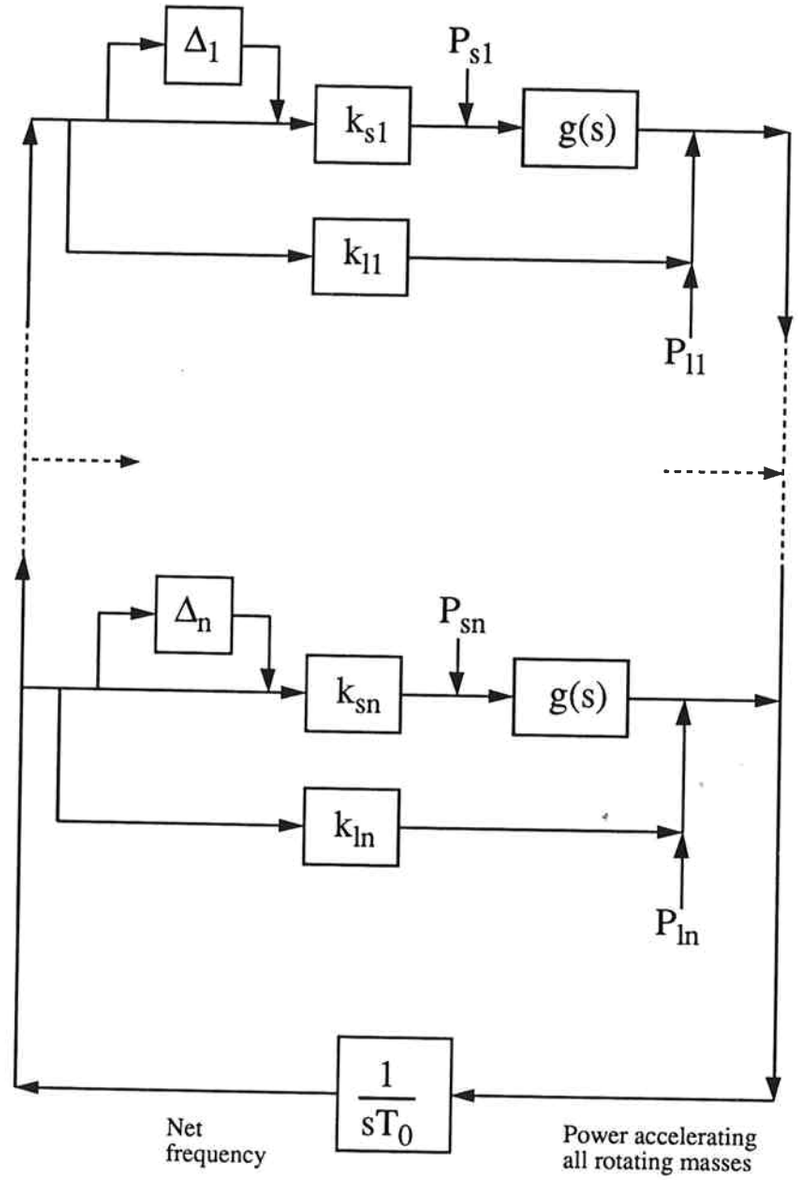


Figure 6.12: Block diagram of the multi-area power system studied by Lunze [13].

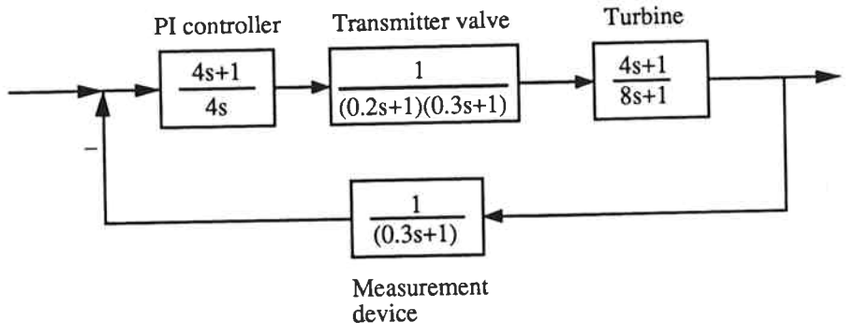


Figure 6.13: Block diagram of an individual power station corresponding to transfer function $g(s)$ in Fig. 6.12.

in Fig. 6.13 (nor does it agree with Figure 4b in [13]). We will use the transfer function description in Fig. 6.13 for $g(s)$ in Fig. 6.12. With these modifications, Lunzes method finds that the overall system is stable provided

$$\frac{1}{n}\sum_{i=1}^{n}\{|\Delta_i|\} < 0.805 \tag{6.51}$$

for k_{si} having the nominal value of -15.[5]

Rewriting the system in Fig. 6.12 in the M-Δ structure, we find

$$\Delta_{n\times n} = \mathrm{diag}\{\Delta_i\} \tag{6.52}$$

$$M = F^H \begin{bmatrix} \lambda_{M1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix} F \tag{6.53}$$

$$= \frac{1}{\sqrt{n}}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \lambda_{M1} \frac{1}{\sqrt{n}}\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \tag{6.54}$$

$$\lambda_{M1} = \frac{g(s)k_{si}/sT}{1-(g(s)k_{si}+k_{li})/sT} \tag{6.55}$$

The reason for all the zero eigenvalues of M is that the only feedback occurs through the SISO integration $\frac{1}{sT_0}$, and M can therefore have only one nonzero eigenvalue. Clearly,

[5] Lunze [13], using apparently erroneous data, found that stability is ensured for $\frac{1}{n}\sum_{i=1}^{n}|\Delta_i| < 0.735$.

because M has only one nonzero eigenvalue, this M-Δ structure can be rewritten as

$$\tilde{\Delta}_{1\times 1} = \frac{1}{n}\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} diag\{\Delta_i\} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{6.56}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\{\Delta_i\} \tag{6.57}$$

$$\tilde{M}_{1\times 1} = \lambda_{M1} \tag{6.58}$$

This is the same as the M-Δ structure for a system consisting of only one subsystem, which is stable for any $\tilde{\Delta}$ such that $-240.7 < k_{si}(1+\tilde{\Delta}) < 2.5$. Values of $k_{si}(1+\tilde{\Delta})$ larger than the upper limit gives positive feedback, and hence instability because of the integration, whereas values lower than the lower limit gives a magnitude larger than one at ω_{180} in a Bode plot. Thus any combination of Δ_i's such that $\frac{2.5-k_{si}}{k_{si}} < \frac{1}{n}\sum_{i=1}^{n}\Delta_i < \frac{-240.7-k_{si}}{k_{si}}$ will give a stable overall system. For the nominal value of $k_{si} = -15$ used by Lunze, this gives $-7/6 < \frac{1}{n}\sum_{i=1}^{n}\{\Delta_i\} < 15.0$. This is significantly less conservative than the result using Lunze's method, which guarantees stability for $-0.805 < \frac{1}{n}\sum_{i=1}^{n}\{\Delta_i\} < 0.805$. From Fig. 6.12 we see that in addition we must require that the individual power stations $g(s)$ are stable (which they are), as there otherwise is a possibility that the outputs from the individual power stations grow unboundedly, while the signal entering the integration $\frac{1}{sT_0}$ remains bounded.

Clearly, careful examination of Fig. 6.12 reveals that this example deals with a SISO stability problem, as there is only one feedback path in the system. Knowledge of the mathematics used above is therefore not required here. However, this example clearly illustrates the importance of taking the structure of the uncertainties into account. Another disadvantage with the work of Lunze is that his robustness analysis appears to give little help in the synthesis of controllers. Lunze suggests using a decentralized controller with identical elements on the diagonal. Lundström et al. [12] found such controllers to not necessarily be optimal for symmetrically interconnected systems, even when considering decentralized controllers only.

6.9 Decentralized Control

In this section, we will only consider single input, single output (SISO) systems in parallel, i.e. we will assume that the plant transfer function matrix is a parallel matrix.

The only type of controller which is both decentralized and circulant, is a controller with all the diagonal elements equal. If $a \neq 0$, the distinct eigenvalue λ_{G1} will differ from

the other eigenvalues of the plant, and a circulant and decentralized controller which is optimal for the control of λ_{G1} will not be optimal for the other plant eigenvalues. Thus, when designing decentralized controllers for systems consisting of similar interacting subsystems, one should consider allowing the controller elements to differ. This issue has been discussed previously by Lundström et al. [12].

However, although the theory of circulant matrices is of limited use for the design of decentralized controllers, we can use this body of theory to prove the existence of a plant property which indicates that decentralized controllers can be useful for the control of systems consisting of similar interacting subsystems.

6.9.1 Decentralized Integral Controllability

Decentralized Integral Controllability (DIC) is a system property which means that there exists a decentralized controller with integral action for which the proportional gain in the individual loops can be reduced independently of each other by an arbitrary amount without upsetting stability. That is, for a system which is DIC, there exists a diagonal controller $K(s) = \frac{1}{s} diag\{k_i\}\bar{K}(s)$, where $\bar{K}(s)$ is a diagonal transfer function matrix and the k_i's are variable gains, such that the closed loop system is stable for any k_i provided $0 \leq k_i \leq k_i^*$; $\forall i$. For this problem DIC can be tested using the steady state gain matrix only, since the steady state gain matrix is symmetric, and therefore has real eigenvalues[6].

In this section we will assume the system to be open loop stable, as this is obviously a prerequisite for DIC.

The special properties of the parallel matrix P makes conditions for DIC easy to find. We will use Lemma 7.2 in [4] which gives a sufficient condition for what is known as D-stability. Any matrix which is D-stable and whose principal submatrices are all D-stable is DIC.

Lemma 7.2, [4].
The matrix A is D-stable if:
$(I + A)^{-1}$ *exists and there exists a diagonal* $D > 0$ *such that*
$\| D(I - A)(I + A)^{-1}D^{-1} \|_2 < 1$

Assume without loss of generality that the elements on the main diagonal of $G(0)$ are positive. Clearly the plant is D-stable if $\| (I - G(0))(I + G(0))^{-1} \|_2 < 1$. The matrix $(I - G(0))(I + G(0))^{-1}$ is parallel, with 2-norm (largest singular value) equal to the

[6] For problems where the steady state gain matrix have eigenvalues on the imaginary axis, it is not known whether DIC can be tested using the steady state gain matrix only [4].

magnitude of the largest eigenvalue. The magnitude of the largest eigenvalue is given by

$$\rho = \max\left\{\left|\frac{1-\lambda_{G1}}{1+\lambda_{G1}}\right|, \left|\frac{1-\lambda_{G2}}{1+\lambda_{G2}}\right|\right\} \tag{6.59}$$

where λ_{G1} is the first (distinct) eigenvalue of $G(0)$ and λ_{G2} equals the repeated eigenvalue of $G(0)$. It is clear that $\rho < 1$ provided *both* λ_{G1} and λ_{G2} are positive (at steady state both λ_{G1} and λ_{G2} must be real). We find that λ_{G1} and λ_{G2} are both positive, and D-stability is thus ensured, provided

$$\frac{-1}{n-1} < a(0) < 1 \tag{6.60}$$

where $a(0)$ is defined in Eq. (6.2). Analyzing the principal submatrices (which are also parallel matrices) the same way, we find that they are also all D-stable for values of a within the limits given in Eq. (6.5).

D-stability of $G(0)$ and its principal submatrices is a sufficient condition for DIC. A necessary condition for DIC is that the elements on the main diagonal of the steady state RGA matrix should be positive for the steady state gain matrix and all its principal submatrices [4]. To use this condition we need some further properties of the parallel matrix P in Eq. (6.6):

Inverse. The inverse of P is a parallel matrix with

$$[P^{-1}]_{ii} = \frac{[1+(n-2)a]}{[1+(n-1)a](1-a)}\frac{1}{k} \tag{6.61}$$

Proof: Can be found from the eigenvalue decomposition.

Relative Gain Array (RGA) [3]. From the above result follows directly

$$RGA_{ii} = p_{ii}[P^{-1}]_{ii} = \frac{[1+(n-2)a]}{[1+(n-1)a](1-a)} \tag{6.62}$$

From Eq. (6.62) it is clear that this necessary condition for DIC is fulfilled only for values of $a(0)$ within the limits given in Eq. (6.60). We therefore conclude that

$$\text{DIC} \Longleftrightarrow \frac{1}{n-1} < a(0) < 1 \tag{6.63}$$

the other eigenvalues of the plant, and a circulant and decentralized controller which is optimal for the control of λ_{G1} will not be optimal for the other plant eigenvalues. Thus, when designing decentralized controllers for systems consisting of similar interacting subsystems, one should consider allowing the controller elements to differ. This issue has been discussed previously by Lundström et al. [12].

However, although the theory of circulant matrices is of limited use for the design of decentralized controllers, we can use this body of theory to prove the existence of a plant property which indicates that decentralized controllers can be useful for the control of systems consisting of similar interacting subsystems.

6.9.1 Decentralized Integral Controllability

Decentralized Integral Controllability (DIC) is a system property which means that there exists a decentralized controller with integral action for which the proportional gain in the individual loops can be reduced independently of each other by an arbitrary amount without upsetting stability. That is, for a system which is DIC, there exists a diagonal controller $K(s) = \frac{1}{s} diag\{k_i\}\bar{K}(s)$, where $\bar{K}(s)$ is a diagonal transfer function matrix and the k_i's are variable gains, such that the closed loop system is stable for any k_i provided $0 \leq k_i \leq k_i^*$; $\forall i$. For this problem DIC can be tested using the steady state gain matrix only, since the steady state gain matrix is symmetric, and therefore has real eigenvalues[6].

In this section we will assume the system to be open loop stable, as this is obviously a prerequisite for DIC.

The special properties of the parallel matrix P makes conditions for DIC easy to find. We will use Lemma 7.2 in [4] which gives a sufficient condition for what is known as D-stability. Any matrix which is D-stable and whose principal submatrices are all D-stable is DIC.

Lemma 7.2, [4].
The matrix A is D-stable if:
$(I + A)^{-1}$ *exists and there exists a diagonal* $D > 0$ *such that*
$\| D(I - A)(I + A)^{-1}D^{-1} \|_2 < 1$

Assume without loss of generality that the elements on the main diagonal of $G(0)$ are positive. Clearly the plant is D-stable if $\| (I - G(0))(I + G(0))^{-1} \|_2 < 1$. The matrix $(I - G(0))(I + G(0))^{-1}$ is parallel, with 2-norm (largest singular value) equal to the

[6] For problems where the steady state gain matrix have eigenvalues on the imaginary axis, it is not known whether DIC can be tested using the steady state gain matrix only [4].

magnitude of the largest eigenvalue. The magnitude of the largest eigenvalue is given by

$$\rho = \max\left\{\left|\frac{1-\lambda_{G1}}{1+\lambda_{G1}}\right|, \left|\frac{1-\lambda_{G2}}{1+\lambda_{G2}}\right|\right\} \tag{6.59}$$

where λ_{G1} is the first (distinct) eigenvalue of $G(0)$ and λ_{G2} equals the repeated eigenvalue of $G(0)$. It is clear that $\rho < 1$ provided *both* λ_{G1} and λ_{G2} are positive (at steady state both λ_{G1} and λ_{G2} must be real). We find that λ_{G1} and λ_{G2} are both positive, and D-stability is thus ensured, provided

$$\frac{-1}{n-1} < a(0) < 1 \tag{6.60}$$

where $a(0)$ is defined in Eq. (6.2). Analyzing the principal submatrices (which are also parallel matrices) the same way, we find that they are also all D-stable for values of a within the limits given in Eq. (6.5).

D-stability of $G(0)$ and its principal submatrices is a sufficient condition for DIC. A necessary condition for DIC is that the elements on the main diagonal of the steady state RGA matrix should be positive for the steady state gain matrix and all its principal submatrices [4]. To use this condition we need some further properties of the parallel matrix P in Eq. (6.6):

Inverse. The inverse of P is a parallel matrix with

$$[P^{-1}]_{ii} = \frac{[1+(n-2)a]}{[1+(n-1)a](1-a)}\frac{1}{k} \tag{6.61}$$

Proof: Can be found from the eigenvalue decomposition.

Relative Gain Array (RGA) [3]. From the above result follows directly

$$RGA_{ii} = p_{ii}[P^{-1}]_{ii} = \frac{[1+(n-2)a]}{[1+(n-1)a](1-a)} \tag{6.62}$$

From Eq. (6.62) it is clear that this necessary condition for DIC is fulfilled only for values of $a(0)$ within the limits given in Eq. (6.60). We therefore conclude that

$$\text{DIC} \Longleftrightarrow \frac{1}{n-1} < a(0) < 1 \tag{6.63}$$

6.10 Generalization to Block Circulant Processes with a Symmetric Block Structure

The results in this paper on H_∞-optimal control and μ analysis and synthesis are easily generalized to processes described by block circulant transfer function matrices with a symmetric block structure, i.e. transfer function matrices of the form

$$G(s) = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 & \cdots & c_3 & c_2 \\ c_2 & c_1 & c_2 & c_3 & c_4 & \cdots & c_3 \\ c_3 & c_2 & c_1 & c_2 & c_3 & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & c_3 & c_2 & c_1 & c_2 & c_3 \\ c_3 & \cdots & c_4 & c_3 & c_2 & c_1 & c_2 \\ c_2 & c_3 & \cdots & c_4 & c_3 & c_2 & c_1 \end{bmatrix} \tag{6.64}$$

Note that it is only when the individual blocks $c_1, c_2, c_3, \cdots$ are symmetric that $G(s)$ is symmetric. If $c_1, c_2, c_3, \cdots$ are of dimension 1×1, $G(s)$ is termed a symmetric circulant matrix. Symmetric circulant matrices occur for example in the cross directional control in paper manufacturing [11, 18], if edge effects are neglected.

The generalization of H_∞ synthesis and μ analysis and synthesis from block parallel plants to block circulant plants with a symmetric block structure is straight forward. Let k be the number of independent blocks γ_{Gi} in Eq. (6.23). In general, if n is an even number, $k = n/2 + 1$, and if n is and odd number, $k = (n-1)/2 + 1$. Then:

1. The design of an H_∞-optimal controller (including the second 'K' step in D-K iteration in μ synthesis) for a block circulant process with a symmetric block structure involves k independent synthesis subproblems for the k "plants" $\gamma_{G1}, \cdots, \gamma Gk$.

2. The first 'D' step in D-K iteration (μ analysis), involves μ analysis for the plant $\tilde{G} = \text{diag}\{\gamma_{G1}, \cdots, \gamma_{Gk}\}$.

3. However, one has to exercise some care when finding the real matrix R used for controller realization in Eq. (6.30). Recall from Section 6.4.1 that only linear combinations of eigenvectors corresponding to *identical* eigenvalues may be used to find real eigenvectors. If the matrix F^H is used as the eigenvector matrix for a symmetric circulant matrix, the first eigenvalue of the matrix is distinct, and the corresponding eigenvector is the first column of F^H, which is real. Similarly, if n is an even number, eigenvalue number $(n/2) + 1$ is also distinct, and the corresponding eigenvector is column $(n/2)+1$ of F^H, which is real. All the other eigenvalues of a symmetric circulant matrix occur in pairs, eigenvalue number p and eigenvalue number $n + 2 - p$ being identical (for $p = \{2, 3, \cdots, \nu\}$, where $\nu = (n+1)/2$ if n is odd and $\nu = n/2$ if n is even). Columns p and $n + 2 - p$ of R^T must therefore result from linear combinations of columns p and $n + 2 - p$ of F^H. In addition, we require all columns of R to be real, mutually orthogonal and of unit length. Let r_l denote column l of R^T, and m_l denote column l of F^H (see Eqs. (6.13-6.14). Note that m_p is the complex conjugate of m_{n+2-p}, and that m_p and m_{n+2-p} are orthogonal. The matrix R will fulfill all the requirements above if we choose:

$$\begin{aligned} r_1 &= m_1 \\ r_{(n/2)+1} &= m_{(n/2)+1} \quad \text{if } n \text{ is an even number} \\ r_p &= \frac{1}{\sqrt{2}}(m_p + m_{n+2-p}) \\ r_{n+2-p} &= \frac{i}{\sqrt{2}}(m_p - m_{n+2-p}) \quad \text{for } p = \{2, 3, \cdots, \nu\} \end{aligned} \tag{6.65}$$

6.11 Conclusions

H_∞ Control: For symmetrically interconnected systems, instead of considering a plant of dimension $n \cdot n_o \times n \cdot n_i$ we can consider two plants, each of dimension $n_o \times n_i$. H_∞-synthesis will then result in a block parallel controller which optimizes the H_∞ criterion in n directions. For plants with SISO subsystems, the controller will have the structure of an SVD controller.

The Structured Singular Value, μ: For cases with SISO subsystems and one source of uncertainty in each subsystem, the value of μ is shown to be insensitive to the structure of the uncertainty.

In the general case, for the first step in D-K iteration (μ analysis) the plant size may be reduced to a block diagonal plant with two blocks of dimension $n_o \times n_i$ on the diagonal. The two blocks on the diagonal of the plant can be considered separately if the perturbation matrix for the transformed system ($\tilde{\Delta}$)) contains no more than one full block. For the second step of D-K iteration, these two blocks may be considered independently. However, this reduction in plant size requires that any diagonal, independent uncertainty blocks are approximated by full uncertainty blocks.

Decentralized Control: A necessary and sufficient condition for Decentralized Integral Controllability of symmetrically interconnected systems is proven to be that the

steady state interaction parameter $a(0)$ must be within the limits given in Eq. (6.60). We have been unable to find any physical system with interaction parameter $a(0)$ outside these limits.

Extension to Block Circulant Processes with a Symmetric Block Structure: The results in this paper are easily generalized to plants described by transfer function matrices which are block circulant with a symmetric block structure. The details of how to generalize the results are given above.

References

[1] Bellman, R. (1970). *Introduction to Matrix Analysis*, McGraw-Hill, New York.

[2] Braatz, R. D., Tyler, M. L., Morari, M., Pranckh, F. R. and Sartor, L. (1992). Identification and Cross-Directional Control of Coating Processes: Theory and Experiments. *Proc. ACC*, Chicago, IL, pp. 1556-1560.

[3] Bristol, E. H. (1966). On a New Measure of Interactions for Multivariable Process Control, *IEEE Trans. Autom. Control*, **AC-11**, 133-134.

[4] Campo, P. and Morari, M. (1992). Achievable Closed Loop Properties of Systems Under Decentralized Control; Conditions Involving the Steady State Gain, submitted to *IEEE Trans. Autom. Control*, preprint.

[5] Davis, P. J. (1979). *Circulant Matrices*, Wiley, New York.

[6] Doyle, J. C., Wall, J. E. and Stein, G. (1982). Performance and Robustness Analysis for Structured Uncertainty. *Proc. IEEE Conf. Decision Contr.*, Orlando, FL., pp. 629-636.

[7] Doyle, J. (1982). Analysis of Feedback Systems with Structured Uncertainties. *IEE Proc.*, Pt. D, **129**, 6, pp. 242-250

[8] Doyle, J. C. and Chu, C.-C. (1985). Matrix Interpolation and H_∞ Performance Bounds. *Proc. ACC*, Boston, MA, pp. 129-134.

[9] Doyle, J. C., Glover, K., Khargonekar, P. and Francis, B. (1989). State-Space Solutions to Standard H_2 and H_∞ Control Problems. *IEEE Trans. Autom. Control*, **AC-34**, 8, pp. 831-847.

[10] Gu, D.-W., Tsai, M. C. and Postlethwaite, I. (1990). Improved Formulae for the 2-Block Super-Optimal Solution, *Automatica*, **26**, 2, pp. 437-440.

[11] Laughlin, D. L., Morari, M. and Braatz, R. D. (1992). Robust Performance of Cross-Directional Basis-Weight Control in Paper Machines. Submitted to *Automatica*, preprint.

[12] Lundström, P., Skogestad, S., Hovd, M. and Wang, Z.-Q. (1991). Non- Uniqueness of Robust H_∞ Decentralized PI Control. *Proc. ACC*, Boston, MA, pp. 1830-1835.

[13] Lunze, J. (1989). Stability Analysis of Large-Scale Systems Composed of Strongly Coupled Similar Subsystems, *Automatica*, **25**, 4, pp. 561-570.

[14] Shinskey, F. G. (1979). *Process Control Systems*, 2nd Edition, McGraw-Hill, New York.

[15] Shinskey, F. G. (1984). *Distillation Control*, 2nd Edition, McGraw-Hill, New York.

[16] Skogestad, S., Lundström, P. and Hovd, M. (1989). Control of Identical Parallel Processes. Presented at AIChE Annual Meeting, San Francisco, CA, Paper no. 167 Ba.

[17] Sundareshan, M. K. and Elbanna, R. M. (1991). Qualitative Analysis and Decentralized Controller Synthesis for a Class of Large-scale Systems with Symmetrically Interconnected Subsystems, *Automatica* **27**, 2, pp. 383-388.

[18] Wilhelm, R. G. Jr. and Fjeld, M. (1983). Control Algorithms for Cross Directional Control: The State of the Art. *Preprints 5th IFAC PRP Conference*, Antwerp, Belgium, pp. 139-150.

Appendix. State space descriptions of controllers found in the examples.

A			
-4.993E−07	-1.633E−04	2.858E−02	3.357E−06
-1.853E−04	-0.1051	-7.644E+02	3.392E−02
2.929E−02	7.995E+02	-2.908E+03	-0.5641
-3.220E−06	-6.248E−02	0.2264	-2.000
B	C^T	D	
0.6859	0.6972	0	
1.436E+02	1.454E+02		
-2.420E+04	-2.419E+04		
1.891	-2.346		

Table 6.2: State space description of λ_{K1} for the controller found in Example 3.

A			
-5.009E−07	-2.046E−02	-1.183E−05	3.129E−05
-2.001E−02	1.294E+03	-6.275E+02	3.182
1.352E−05	6.276E+02	-1.430E−04	1.056E−03
-2.966E−05	-3.182	1.276E−03	-1.611
B	C^T	D	
0.5429	0.5314	0	
1.353E+04	1.353E+04		
-5.061	4.495		
16.61	-16.61		

Table 6.3: State space description of λ_{K2} for the controller found in Example 3.

diag{ A }	B	
-231.439	0.0134	2.9209
-2.470	1.2090	-2.1960
-0.823	-0.3021	0.5568
-1.064	-0.3085	0.0152
-0.176	-0.1196	-3.4066
-1E−08	1.82E−04	4.6785
-2E−09	0.7891	3.1E−09
C^T		
0.0565	-41.2565	37.2301
0.0455	-0.4676	0.3644
0.1665	-0.1291	-0.0588
0.0720	-0.1922	0.0877
-0.0221	0.0114	-0.0050
-0.0331	-0.0253	0.0350
-0.0395	-0.0237	0.0410
D		
	-0.0125	0
	0	-6.25E−03
	0	-6.25E−03

Table 6.4: State space description of γ_{K1} for the controller found in **Example 4.**

diag{ A }	B	
-3591.5	0.0107	-20.3237
-266.7	0.0330	-24.2771
-130.4	0.0257	15.3908
-1.120	0.0663	$1.031E{-}03$
-0.0521	0.0734	-0.3711
$-3.9E{-}08$	-3.0E-10	-0.6377
$-4.7E{-}09$	-0.2745	-6.5E-11
C^T		
-1241	316.2	-171.4
53.76	41.55	38.59
-18.46	46.87	21.23
$-9.744E{-}03$	0.2968	0.1639
$1.219E{-}02$	$-2.804E{-}03$	$4.267E{-}03$
$-3.593E{-}02$	$2.020E{-}02$	$2.295E{-}03$
$6.100E{-}03$	$-1.288E{-}02$	$-4.801E{-}03$
D		
	0	0
	0	0
	0	0

Table 6.5: State space description of γ_{K2} for the controller found in Example 4.

A			B
$-1.00E{-}07$	$-1.73E{-}03$	$1.29E{-}06$	0.187
$-1.73E{-}03$	$-4.44E{+}05$	$1.11E{+}03$	$1.63E{+}03$
$-1.30E{-}06$	$-1.11E{+}03$	-0.60	1.20
C			D
0.188	$1.62E{+}03$	-1.21	0

Table 6.6: State space description of λ_{K1} for the controller found in Example 5.

A		
Sub-diagonal	Main diagonal	Super-diagonal
	$-9.692E{-}08$	0
0	$-1.509E{-}01$	0
0	$-1.021E{+}00$	0
0	$-9.600E{+}00$	$9.525E{+}00$
$-9.526E{+}00$	$-9.600E{+}00$	0
0	$-1.338E{+}02$	0
0	$-1.110E{+}04$	
B	C^T	D
$-1.944E{-}01$	$-5.222E{-}01$	0
$-7.293E{-}02$	$9.074E{-}00$	
$2.644E{-}01$	$1.218E{+}01$	
$-2.228E{+}00$	$-4.392E{+}01$	
$-1.747E{+}00$	$-2.743E{+}01$	
$2.460E{+}01$	$-2.472E{+}03$	
$-1.633E{+}01$	$-3.098E{+}05$	

Table 6.7: State space description of λ_{K1} for the controller found in Example 6.

A		
Sub-diagonal	Main diagonal	Super-diagonal
	-9.975E−08	0
0	-1.535E−02	0
0	-3.325E−01	0
0	-4.099E−01	0
0	-2.192E−00	9.257E−01
-9.257E−01	-2.192E+01	0
0	-1.645E+01	0
0	-1.250E+02	0
0	-1.987E+02	
B	C^T	D
4.426E−01	6.897E−01	0
-8.813E−02	1.931E−01	
2.441E−01	-1.079E−00	
6.614E−01	1.413E+01	
5.825E−02	-5.937E+00	
5.437E−01	-5.180E+00	
-3.673E−01	-6.958E+01	
-1.435E+02	9.608E+02	
-1.437E+02	-1.541E+03	

Table 6.8: State space description of $\lambda_{K2} = \lambda_{K3} = \lambda_{K4}$ for the controller found in Example 6.

Chapter 7

On the Structure of the Robust Optimal Controller for a Class of Problems

Morten Hovd, Richard D. Braatz* and Sigurd Skogestad†
Chemical Engineering
University of Trondheim, NTH
N-7034 Trondheim, Norway

Abstract

In this paper we investigate the structure of the robust optimal controller for a class of control problems investigated by many researchers. The robust optimal controller for a problem in this class is an SVD controller. This finding may be used to simplify the controller synthesis (K) part of the D-K iteration procedure used for synthesizing μ-optimal controllers.

Conditions for when the optimal controller in general has the structure of an SVD controller are discussed, focusing on the issues of realizability of the transformed interconnection matrix and whether the transformation makes the structure of the perturbation block (Δ) more conservative.

*currently at California Institute of Technology, Chemical Engineering 210-41, Pasadena, CA 91125, USA
†To whom correspondence should be addressed. FAX: +47-7-594080, e-mail: skoge@kjemi.unit.no

7.1 Introduction

In this paper we investigate the structure of the robust optimal controller for a class of control problems investigated by many researchers. An example of a control problem in this class is given by the robust controller design problem for a distillation column studied previously by Skogestad et al. [22].

The nominal plant for this problem is given by

$$G(s) = \frac{1}{75s+1}\begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \tag{7.1}$$

which has a condition number of 141.7 and a RGA-value of 35.5 at all frequencies. This model is an excellent example for demonstrating the problems with ill-conditioned plants and has been studied by many researchers [17, 3, 24].

For this problem, the relative magnitude of the uncertainty in each of the manipulated variables is given by

$$w_I(s) = 0.2(5s+1)/(0.5s+1) \tag{7.2}$$

The robust performance specification is that $\|w_P S_p\|_\infty < 1$ where

$$w_P = 0.5(10s+1)/10s \tag{7.3}$$

and S_p is the worst sensitivity function possible with the given bounds on the uncertainty in the manipulated variables.

This robust controller design problem is easily captured in the framework of the structured singular value, μ [6]. The resulting μ condition for Robust Performance (RP) becomes:

$$\text{RP} \iff \mu_\Delta(M) < 1 \quad \forall\omega \tag{7.4}$$

$$M = \begin{bmatrix} -W_I KSG & W_I KS \\ W_P SG & -W_P S \end{bmatrix}; \quad \Delta = \text{diag}\{\Delta_I, \Delta_P\} \tag{7.5}$$

where Δ_I is a *diagonal* 2×2 perturbation block, Δ_P is a full 2×2 perturbation block and

$$W_I = w_I I_2 \quad \text{and} \quad W_P = w_P I_2$$

Skogestad et al. [22] designed a controller giving a value of $\mu = 1.067$. Freudenberg [10] used another design method to find a controller with $\mu = 1.054$, and also used this problem as an example in [4]. Lundström et al. [17] used the latest state-space H_∞ software [1] to design a controller with $\mu = 0.978$. In a somewhat altered form,

this robust controller design problem has been considered by Yaniv and Barlev [24], and was used as a benchmark for the 1991 CDC [3].

Engstad [9] found numerically that the controller obtained by Lundström et al. [17] has the structure of an SVD controller. We prove that the μ-optimal controller is an SVD controller for this robust controller design problem. This suggests that the controller obtained by Lundström et al. [17] is very near μ-optimal. The resulting analysis suggests how the design problem can be simplified prior to applying D-K iteration (or H_∞-synthesis) for finding the controller. The simplified design problem is equivalent to the original design problem provided this contains only full and/or multiplicative repeated scalar perturbation blocks. We then describe the class of problems for which the optimal controller is an SVD controller.

7.2 Background

SVD Controller The plant $G(s)$ which can be decomposed into $G(s) = U\Sigma_G(s)V^H$, where

$$\begin{aligned} \Sigma_G(s) &= \frac{1}{75s+1}\begin{bmatrix} 1.9721 & 0 \\ 0 & 0.0139 \end{bmatrix} \\ U &= \begin{bmatrix} 0.6246 & -0.7809 \\ 0.7809 & 0.6246 \end{bmatrix}; \; V = \begin{bmatrix} 0.7066 & -0.7077 \\ -0.7077 & -0.7066 \end{bmatrix} \end{aligned} \tag{7.6}$$

U and V are unitary matrices. This is the singular value decomposition of the plant $G(s)$[1]. We define an SVD controller for the plant $G(s)$ to have the form

$$K(s) = V\Sigma_K(s)U^H \tag{7.7}$$

where $\Sigma_K(s)$ is a diagonal matrix.

SVD controllers have been studied previously by Hung and MacFarlane [13] and Lau et al. [14]. However, in both of these references the SVD structure is essentially used to counteract interactions at one given frequency, as the problems considered are such that U and V change with frequency. In this paper we consider problems for which U and V are constant at all frequencies and can be chosen to be real. Restricting our attention to these cases allow us to address the optimality of the SVD controller for robust control.

[1] With the slight modification that the dynamic term, $1/(75s+1)$, is multiplied into the singular value matrix Σ_G, thus giving the singular values *phase*.

Robust Performance The goal of any controller design is that the overall system is stable and satisfies some minimum performance requirements. These requirements should be satisfied at least when the controller is applied to the *nominal* plant, that is, we require nominal stability and nominal performance.

In practice the real plant G_p is not equal to the model G. The term *robust* is used to indicate that some property holds for a set Π of possible plants G_p as defined by the uncertainty description. In particular, by *robust stability* we mean that the closed loop system is stable for all $G_p \in \Pi$. By *robust performance* we mean that the performance requirements are satisfied for all $G_p \in \Pi$. Performance is commonly defined in robust control theory using the H_∞-norm of some transfer function of interest.

Definition 2 *The closed loop system exhibits nominal performance if*

$$\|\Psi\|_\infty \equiv \sup_\omega \overline{\sigma}(\Psi) < 1 \tag{7.8}$$

Definition 3 *The closed loop system exhibits robust performance if*

$$\|\hat{\Psi}\|_\infty \equiv \sup_\omega \overline{\sigma}(\hat{\Psi}) < 1, \quad \forall G_p \in \Pi \tag{7.9}$$

For example, for rejection of disturbances at the plant output, Ψ would be the weighted sensitivity

$$\begin{aligned} \Psi &= W_1 S W_2, \quad S = (I + GK)^{-1} \\ \hat{\Psi} &= W_1 S_p W_2, \quad S_p = (I + G_p K)^{-1} \end{aligned} \tag{7.10}$$

In this case, the input weight W_2 is often equal to the disturbance model. The output weight W_1 is used to specify the frequency range over which the sensitivity function should be small and to weigh each output according to its importance. The value K is the transfer function of the controller.

Doyle [6] derived the *structured singular value*, μ, to test for robust performance. To use μ we must model the uncertainty (the set Π of possible plants G_p) as norm bounded perturbations (Δ_i) on the nominal system. Through weights each perturbation is normalized to be of size one:

$$\|\Delta_i\|_\infty \leq 1 \tag{7.11}$$

The perturbations, which may occur at different locations in the system, are collected in the block-diagonal matrix Δ_U (the U denotes uncertainty)

$$\Delta_U = \text{diag}\,\{\Delta_i\} \tag{7.12}$$

and the system is arranged to match the left block diagram in Figure 7.1. The interconnection matrix M in Figure 7.1 is determined by the nominal model (G), the size and nature of the uncertainty, the performance specifications, and the controller (K).

For notational convenience in this section we assume M and each Δ_i are square (analogs to the definitions and theorems in this section hold in the nonsquare case [16]). We assume each Δ_i is complex. For the example studied in this paper, these assumptions hold. The definition of μ is:

Definition 4 *Let $M \in \mathbf{C}^{n\times n}$ be a square complex matrix and define the set $\mathbf{\Delta}$ of block-diagonal perturbations by*

$$\mathbf{\Delta} \equiv \left\{ \text{diag}\,\{\Delta_1, \cdots, \Delta_l, \delta_1 I_{r_1}, \cdots, \delta_m I_{r_m}\} \,\middle|\, \Delta_i \in \mathbf{C}^{p_i \times p_i}, \delta_j \in \mathbf{C}, \sum_{i=1}^{l} p_i + \sum_{j=1}^{m} r_j = n \right\} \tag{7.13}$$

Then $\mu_{\mathbf{\Delta}}(M)$ (the structured singular value with respect to the uncertainty structure $\mathbf{\Delta}$) is defined as

$$\mu_{\mathbf{\Delta}}(M) \equiv \begin{cases} 0 \text{ if there does not exist } \Delta \in \mathbf{\Delta} \text{ such that } \det(I + M\Delta) = 0 \\ \left[\min_{\Delta\in\mathbf{\Delta}} \{\overline{\sigma}(\Delta) \,|\det(I + M\Delta) = 0\}\right]^{-1} \quad \text{otherwise} \end{cases} \tag{7.14}$$

Partition M in Fig. 7.1 to be compatible with $\Delta = \text{diag}\{\Delta_U, \Delta_P\}$:

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{7.15}$$

The following are tests for robust stability and robust performance [6]

Theorem 11 *The closed loop system exhibits robust stability for all $\|\Delta_U\|_\infty \leq 1$ if and only if the closed loop system is nominally stable and*

$$\mu_{\mathbf{\Delta}_U}(M_{11}(j\omega)) < 1 \qquad \forall\omega \tag{7.16}$$

Theorem 12 *The closed loop system exhibits robust performance for all $\|\Delta_U\|_\infty \leq 1$ if and only if the closed loop system is nominally stable and*

$$\mu_{\mathbf{\Delta}}(M(j\omega)) < 1 \qquad \forall\omega \tag{7.17}$$

where $\mathbf{\Delta} = \text{diag}\{\Delta_U, \Delta_P\}$, and Δ_P is a full square matrix with dimension equal to the number of outputs (the subscript P denotes performance).

Multiple performance objectives can be tested similarly using block-diagonal Δ_P. Note that the issue of robust stability is simply a special case of robust performance.

It is a key idea that μ is a *general* analysis tool for determining robust performance. Any system with uncertainty adequately modeled as in (7.11) can be put into $M - \Delta_U$ form, and robust stability and robust performance can be tested using (7.16) and (7.17). Standard programs calculate the M and Δ [1], given the transfer functions describing the system components and the location of the uncertainty and performance blocks Δ_i.

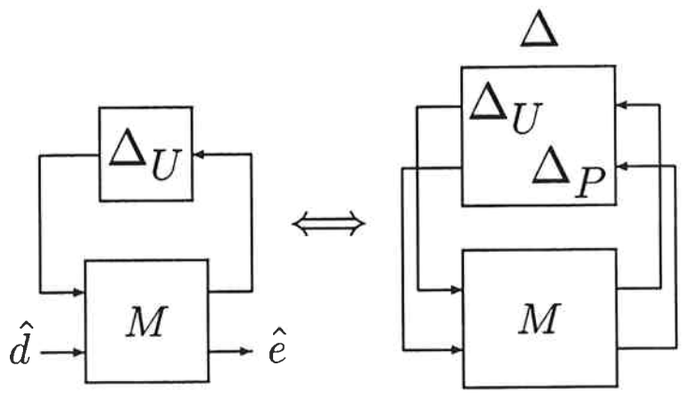


Figure 7.1: Robust Performance and the $M - \Delta$ block structure

Computation of μ μ with complex Δ is commonly calculated through upper and lower bounds. First define two subsets of $\mathbf{C}^{n\times n}$

$$\mathbf{Q} = \left\{Q \in \mathbf{\Delta} : Q^H Q = I_n\right\} \tag{7.18}$$

where Q^H is the conjugate transpose of Q and I_n is the $n \times n$ identity matrix, and

$$\mathbf{D} = \{\text{diag}\,[d_i I_i] : \dim(I_i) = \dim(\Delta_i), d_i \text{ positive real scalar}\} \tag{7.19}$$

then [6]

$$\max_{Q\in\mathbf{Q}} \rho(QM) \leq \mu_{\mathbf{\Delta}}(M) \leq \inf_{D\in\mathbf{D}} \overline{\sigma}\left(DMD^{-1}\right) \tag{7.20}$$

A result of Doyle [6] is that the lower bound, $\max_{Q\in\mathbf{Q}} \rho(QM)$, is always equal to $\mu_{\mathbf{\Delta}}(M)$. Unfortunately, the maximization is not convex, and computing the global maximum of such functions is in general difficult. In contrast, the computation of the upper bound is convex. However, the upper bound is not necessarily equal to μ except when the number of complex Δ-blocks is ≤ 3. The upper and lower bounds are almost always within a percent or so for real problems [19], so for engineering purposes μ never has to be calculated exactly.

Controller Synthesis M is a function of the controller K. The H_∞-optimal control problem is to find a stabilizing K which minimizes $\sup_\omega \overline{\sigma}(M(K))$. The state-space approach for solving the H_∞-control problem is described in [8].

The D-K iteration method (often called μ-synthesis) is an *ad hoc* method which attempts to minimize the tight upper bound of μ in (7.20), i.e. it attempts to solve

$$\min_K \inf_{D\in\mathbf{D}} \sup_\omega \overline{\sigma}\left(DM(K)D^{-1}\right) \tag{7.21}$$

The approach in D-K iteration is to alternatively minimize $\sup_\omega \overline{\sigma}\left(DM(K)D^{-1}\right)$ for either K or D while holding the other constant. For fixed D, the controller synthesis is solved via H_∞-optimization. For fixed K, the quantity is minimized as a convex optimization. The resulting D as a function of frequency is fitted with an invertible stable minimum-phase transfer function and wrapped back into the nominal interconnection structure. This increases the number of states of the scaled M, which leads the next H_∞-synthesis step to give a higher order controller. The iterations stop after $\sup_\omega \overline{\sigma}\left(DM(K)D^{-1}\right)$ is less than 1 or is no longer diminished. The resulting high-order controller can usually be reduced significantly using standard model reduction techniques [1]. Though this method is not guaranteed to converge to a global minimum, it has been used extensively to design robust controllers and seems to work well [7].

7.3 The Structure of the μ-Optimal Controller

In Section 7.3.1 we give the structure of the optimal controller found by Lundström et al. in [17]. In Section 7.3.2 we show that the μ-optimal controller is an SVD controller when Δ_I is a full uncertainty block. In Section 7.3.3 we explain how this also allows us to derive a simplified D-K iteration design procedure in which the synthesis part (K) can be solved as two decoupled subproblems. The issue of the structure of Δ_I is treated in Section 7.4.

7.3.1 The Structure of the Controller Found by Lundström in [17]

In Fig. 7.2 the elements of $\tilde{K}(s) = V^H K(s) U$ for the controller found by Lundström in [17] are shown. We see that the diagonal elements of $\tilde{K}(s)$ are much larger than the offdiagonal elements. Engstad [9] showed numerically that removing the offdiagonal elements of $\tilde{K}(s)$ does not alter the value of μ, and thus showed that the optimal controller found by Lundström in [17] can be chosen to have the structure of an SVD controller.

7.3.2 Analysis of the Optimal Control Problem

Here we analyze the control problem for full block Δ_I as used by Lundström et al. in [17]. We prove that the μ-optimal controller must be an SVD controller.

In Fig. 7.3 and 7.4 we give equivalent block diagrams for the $M - \Delta$ structure in (7.5). In Fig. 7.3a the original feedback system with the uncertainty block Δ_I and

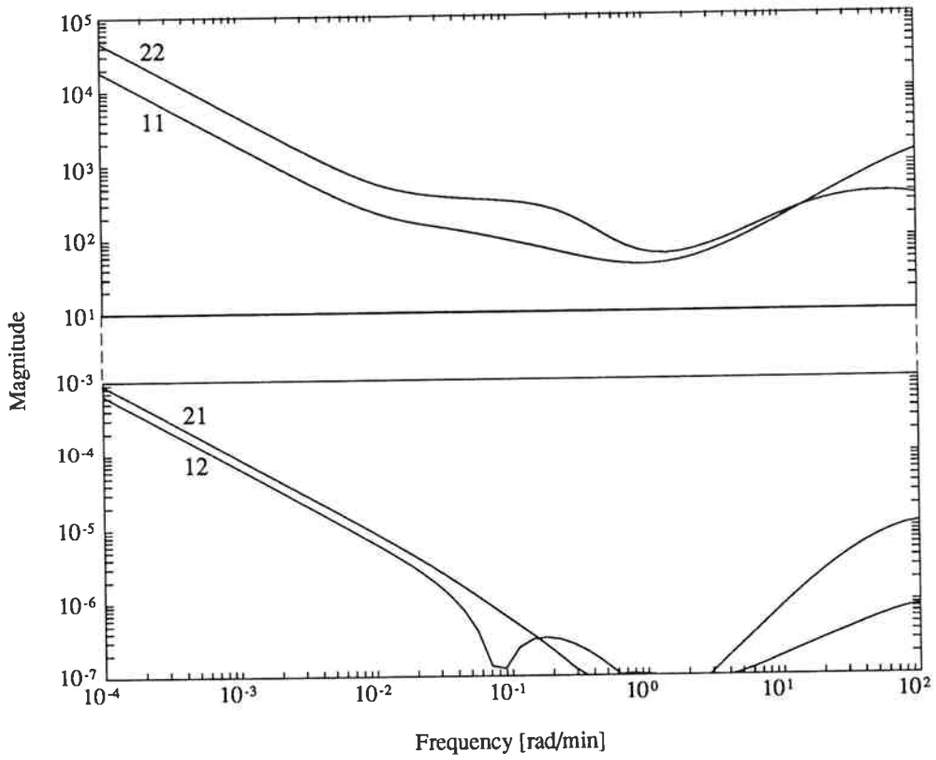


Figure 7.2: Magnitude plot of the elements of $\tilde{K} = V^H K U$ for the controller K found by Lundström in [17].

the performance block Δ_P is shown[2]. In Fig. 7.3b the corresponding $M - \Delta$ structure is shown, and in Fig. 7.3c the weights W_I and W_P have been factored out, and M is expressed as a linear fractional transformation (LFT) of the controller K. Clearly, an identity matrix can be inserted in any channel between the blocks in Fig. 7.3c without altering the problem. In Fig. 7.4a we have inserted identities in four different places (e.g. $UU^H = U^H U = I$). U and V are given by the non-standard singular value decomposition of the plant G in Eq. (7.6), and Σ_G is the corresponding singular value matrix for which the singular values have phase (and are realizable). Note that $\text{diag}\{V^H, U^H\}$ commutes with $\text{diag}\{W_I, W_P\}$, since both W_I and W_P are scalar times identity matrices. The blocks within each dashed box in Fig. 7.4a are combined to form

[2]The minus sign in front of W_P may appear odd. This is due to Lundström [17] defining offset e as measurement y - reference signal r. Including the performance specification by closing the loop from offset to reference signal then introduces the minus in front of W_P. Note, however, that removing this minus sign and substituting $-\Delta_P$ for Δ_P does not change the system.

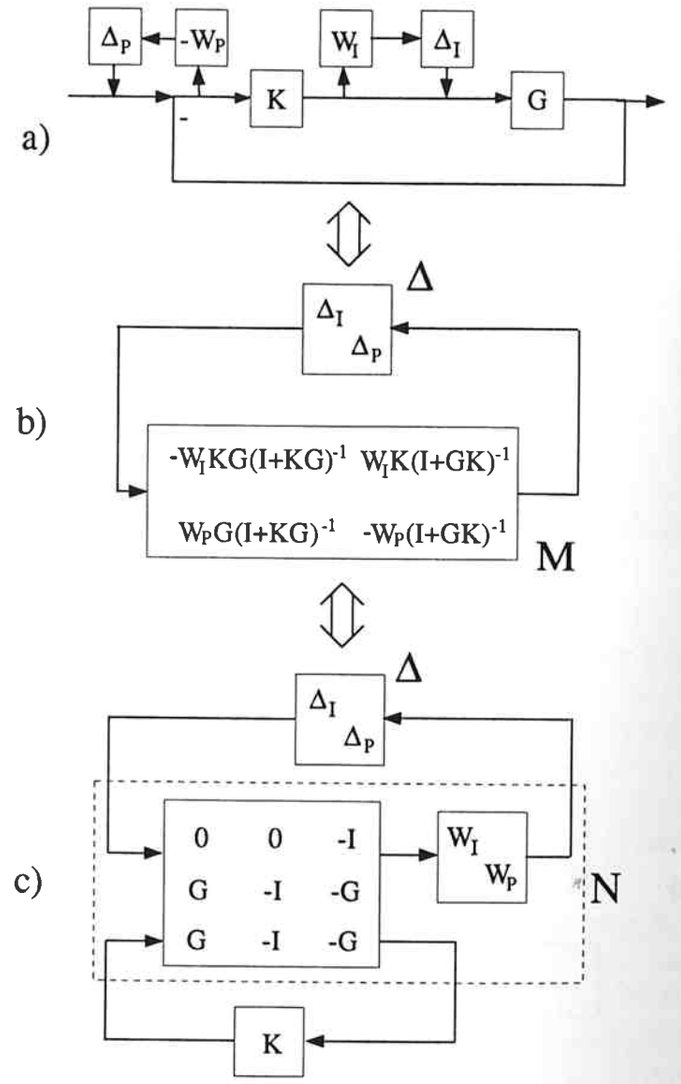


Figure 7.3: a) The feedback system with perturbation blocks. b) The corresponding $M - \Delta$ structure of the synthesis problem. c) Expressing M as a linear fractional transformation of the controller K.

$\tilde{\Delta}$, $\tilde{N}$, and $\tilde{K}$, which we show in Fig. 7.4b. In Fig. 7.4c we show the corresponding conventional feedback system with perturbation blocks.

Now we consider the structure of the transformed system in Fig. 7.4b. All the blocks in $\tilde{N}$ in Fig. 7.4b are diagonal. The transformed performance block $\tilde{\Delta}_P$ is a full

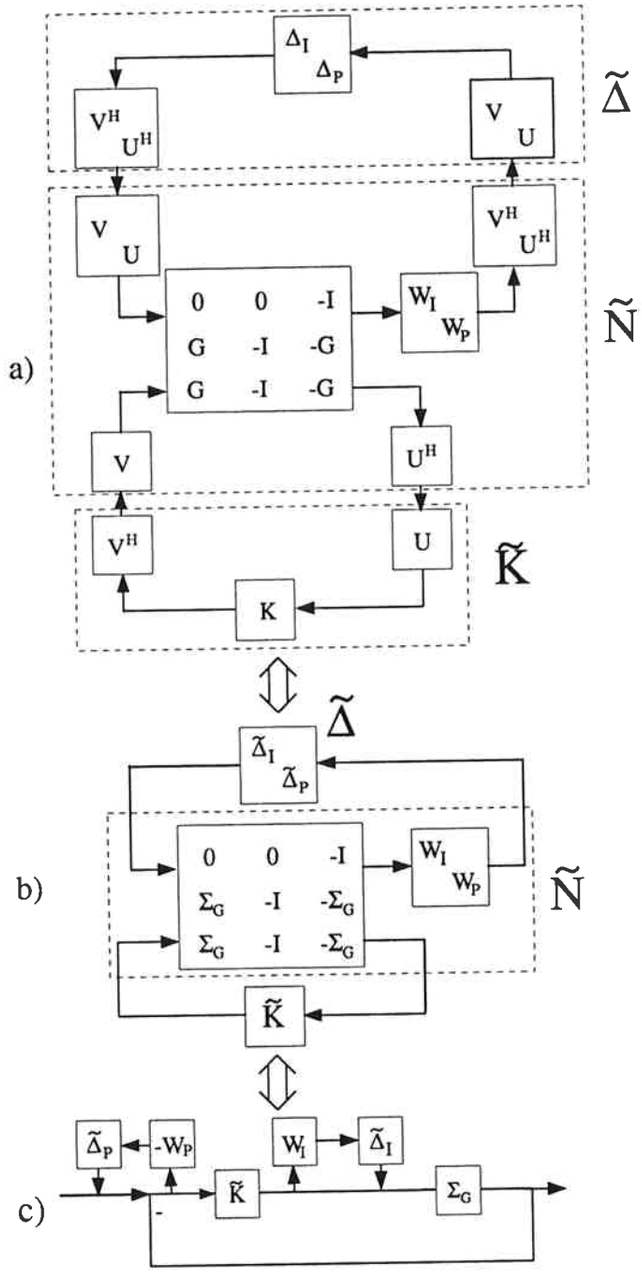


Figure 7.4: a) Inserting identity matrices (unitary matrices multiplied by their respective conjugate transposes) into the problem. b) The resulting transformed feedback system $\tilde{M} - \tilde{\Delta}$. c) The corresponding feedback system with "plant" Σ_G and controller $\tilde{K}$.

block. With Δ_I diagonal, the transformed uncertainty block $\tilde{\Delta}_I = V^H \Delta_I V$ in $\tilde{\Delta}$ would be a full block, with a certain structure that cannot be utilized in the μ framework. We will return to the issue of structure of Δ_I in Section 4.1, for now we will assume Δ_I to be full with no additional structure, as was also assumed by Lundström during the design of the controller in [17].

Now consider the robust optimal control problem, in which we desire to minimize μ:

$$\min_K \mu_\Delta(M(K)) = \min_{\tilde{K}} \mu_{\tilde{\Delta}}(\tilde{M}(\tilde{K})) = \min_{\tilde{K}} \inf_{D \in \mathbf{D}} \sup_\omega \overline{\sigma}\left(D\tilde{M}(\tilde{K})D^{-1}\right) \quad (7.22)$$

$$= \inf_{D \in \mathbf{D}} \min_{\tilde{K}} \sup_\omega \overline{\sigma}\left(D\tilde{M}(\tilde{K})D^{-1}\right) \quad (7.23)$$

The first equality holds because the transformed system is equivalent to the original system. The second equality holds because the number of uncertainty blocks is ≤ 3. Because the perturbations are full block, the D-scales are diagonal. Since we also have that every block in $\tilde{N}$ is diagonal (see Fig. 7.4b), the controller synthesis with given D-scales

$$\min_{\tilde{K}} \sup_\omega \overline{\sigma}\left(D\tilde{M}(\tilde{K})D^{-1}\right) \quad (7.24)$$

consists of (in this case two) *completely decoupled synthesis subproblems, each subproblem involving a SISO "plant"*, since the optimal controller may be obtained from Eq. (7.23) when the D-scales are adjusted in an outer loop. The resulting robust optimal controller $\tilde{K}$ must therefore be diagonal, and the overall controller $K = V\tilde{K}U^H$ is an SVD controller ($\Sigma_K = \tilde{K}$ in Eq. (7.7)). Note that the fact that the robust optimal controller is an SVD controller does not imply that it is an inverse-based controller, i.e. we do not have to choose $\tilde{K} = k\Sigma_G^{-1}$.

To see why the fact that the D-scales are diagonal and all blocks of $\tilde{N}$ are diagonal implies that the controller synthesis problem can be decomposed into n non-interacting synthesis subproblems, consider Fig. 7.5. We express $D\tilde{M}D^{-1}$ as a LFT of $\tilde{K}$, and term the resulting interconnection matrix $\tilde{N}_D$. This matrix $\tilde{N}_D$ will consist of blocks which are diagonal. The matrix at the top of Fig. 7.5 may represent $\tilde{N}_D$. After permutations (row and column interchanges) we get the matrix at the bottom of Fig. 7.5, from which it is apparent that the controller design problem consists of two independent subproblems.

Finally, note that a simple way to realize that the optimal controller is on the SVD form is to go directly from Fig. 7.3a to Fig. 7.4c by introducing $G = U\Sigma_G V^H$ and assuming the weights W_I and W_P are scalar times identity matrices and that Δ_I and Δ_P are full blocks. From Fig. 7.4c it is clear (except for the full $\tilde{\Delta}_I$ and $\tilde{\Delta}_P$) that the design in terms of $\tilde{K}$ involves a set of SISO subproblems.

$$\tilde{N}_D = \begin{bmatrix} a_1 & & b_1 & & c_1 & \\ & a_2 & & b_2 & & c_2 \\ d_1 & & e_1 & & f_1 & \\ & d_2 & & e_2 & & f_2 \\ g_1 & & h_1 & & i_1 & \\ & g_2 & & h_2 & & i_2 \end{bmatrix}$$

$$\updownarrow \text{ Permutations}$$

$$\begin{bmatrix} a_1 & b_1 & c_1 & & & \\ d_1 & e_1 & f_1 & & & \\ g_1 & h_1 & i_1 & & & \\ & & & a_2 & b_2 & c_2 \\ & & & d_2 & e_2 & f_2 \\ & & & g_2 & h_2 & i_2 \end{bmatrix}$$

Figure 7.5: Top: $\tilde{N}_D$. Bottom: $\tilde{N}_D$ permuted to have the two independent synthesis subproblems along the main diagonal.

7.3.3 Consequences for D-K iteration

Now consider performing the D-K iteration design procedure to the transformed system to try to determine the robust optimal controller. The controller synthesis part (K) of D-K synthesis consists of two completely decoupled synthesis subproblems, each subproblem involving a SISO "plant". This holds also after applying the D-scales from the robustness analysis (D) part of D-K synthesis, since the D-scales also consist of diagonal blocks. When using (7.7) to find the controller K from the diagonal Σ_K, we see that the resulting controller will have the structure of an SVD controller.

However, since $\tilde{\Delta}$ contains full blocks, the same D-scales must be applied to both synthesis subproblems. This is seen from Eq. (7.23) where we iterate on the D-scales in an outer loop. The robustness analysis (D) part of D-K iteration must therefore be performed simultaneously for both subproblems, i.e. we must consider the diagonal matrix Σ_G for robustness analysis, and not its diagonal elements separately.

Performing D-K iteration on the transformed system will converge faster and is numerically better conditioned than on the original system. This is both because the H_∞ subproblems are smaller than the original problem, and because the algorithm will be initialized with a controller which has the correct (optimal) directionality.

7.4 Discussion

7.4.1 The Structure of Δ_I for Robust Performance

We now show numerically that the robust performance of the μ-optimal controller is insensitive to the structure of Δ_I for the specific example used in this paper.

In Fig. 7.6 we give the robust performance μ plots for the controller of Lundström et al. for both when Δ_I is full block and when Δ_I consists of independent scalar blocks. The plots are indistiguishable, i.e. robust performance is independent of the structure of Δ_I for the SVD controller. Notice the flatness of the μ plots; it is well-known that the μ-optimal controller has the property that the optimal $\mu(M(j\omega))$ is constant, except at very high frequencies where μ must approach $|w_P|$ for proper controllers. Since the μ plot for the controller of Lundström et al [17] is very flat, and the controller is an SVD controller [9], we expect that this controller is very nearly μ-optimal.

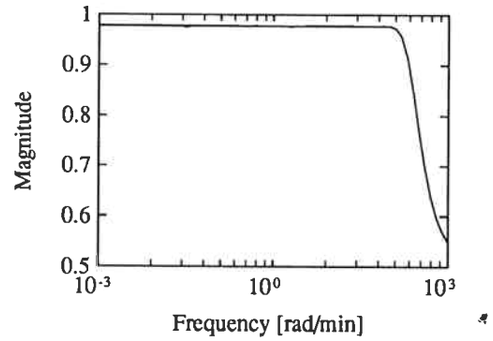


Figure 7.6: Comparison of robust performance μ plots with Δ_I equal to full block and to a diagonal block.

We emphasize that in general it does not hold that the robustness of SVD controllers is insensitive to the structure of the individual blocks. For example, in [22] Skogestad et al. used an inverse-based controller (i.e., a kind of SVD controller) for distillation column control with the DV-configuration, and found that the value of μ for robust performance is strongly dependent on whether Δ_I is a full or a diagonal uncertainty block. The transfer function matrix for the DV-configuration is found by postmultiplying $G(s)$ in Eq. (7.1) by

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \tag{7.25}$$

7.4.2 The Structure of the Uncertainty Block for Robust Stability

We now show that the structure of the uncertainty block is unimportant in determining robust stability for cases with only one multiplicative or inverse multiplicative uncertainty provided that the controller is an SVD controller. To do this, we will need the following result, where A^H is the conjugate transpose of A.

Theorem 13 (μ for Normal Matrices) *Assume M is normal (i.e. $M^H M = M M^H$), then $\mu_\Delta(M) = \rho(M) = \overline{\sigma}(M)$ irrespective of the structure of Δ (provided Δ is complex).*

Proof: Result follows directly from (7.20) and that $\rho(M) = \overline{\sigma}(M)$ for normal matrices. QED.

This theorem states that the value for μ is independent of the structure of Δ provided that the M matrix is normal. This result has proven useful for studying the robust control of cross-directional paper manufacturing [15] and coating processes [2], and for parallel processes [12, 20].

We now apply this theorem to show that for SVD controllers the robust stability for the system under study is independent of the structure of the uncertainty block for cases with only one multiplicative or inverse multiplicative uncertainty block.

Theorem 14 (Structure of the uncertainty block) *Assume that the weights are scalar times identity matrices and that an SVD controller is used. For cases with only one multiplicative or inverse multiplicative uncertainty block, the robust stability for the system is independent of the structure of the uncertainty block.*

Proof: In general, for cases with one uncertainty block the μ interconnection matrix for robust stability, M_{11}, can be expressed as $M_{11} = W_1 T W_2$. We are considering cases where W_1 and W_2 are scalar times identity matrices. There are only two different types of multiplicative uncertainty, input and output multiplicative uncertainty. Likewise, inverse multiplicative uncertainty at the plant input and output are the only two different types of inverse multiplicative uncertainty. In Table 7.1 we give the transfer function matrix T for these four cases. The proof follows from Thm. 13 by substituting $K = V \Sigma_K U^H$ and $G = U \Sigma_G V^H$ into M_{11}. Noting that W_1 and W_2 will commute with any square matrix, and that diagonal matrices always commute, proving that $M_{11}^H M_{11} = M_{11} M_{11}^H$ only involves trivial (but tedious) algebra. QED.

In the example used in this paper, Δ_I is a multiplicative input uncertainty. We therefore conclude from Thm. 14 that for robust stability the structure of Δ_I does not

Type of uncertainty	T
Multiplicative input uncertainty	$-KG(I+KG)^{-1}$
Multiplicative output uncertainty	$-GK(I+GK)^{-1}$
Inverse multiplicative input uncertainty	$(I+KG)^{-1}$
Inverse multiplicative output uncertainty	$(I+GK)^{-1}$

Table 7.1: Transfer function matrix T ($M_{11} = W_1 T W_2$) for different types of multiplicative and inverse multiplicative uncertainties.

matter, since an SVD controller is used. Note that this is not necessarily true when the controller is not an SVD controller.

7.4.3 Generalization of the Results

It is of interest to determine for which class of controller synthesis problems the μ-optimal controller has the structure of an SVD controller. It is easier to consider when the "μ-upper bound" optimal controller has the structure of an SVD controller. Since the D-K iteration procedure uses the upper bound when designing the controller, and the upper bound is within 1-2% of μ for all practical problems to date, considering the optimally in terms of the upper bound is not restrictive.

A practical requirement is that the transformed interconnection matrix (corresponding to $\tilde{N}$ in Fig. 7.4) must be realizable, in order to enable the use of standard state-space based H_∞-synthesis algorithms, e.g. [8]. In order to be specific, we give the following three classes of systems for which the transformed interconnection matrix will always be realizable:

1. the plant is described by scalar dynamics multiplied by a constant matrix, with scalar times identity weights,
2. the plant is parallel, with parallel weights, and
3. the plant is symmetric circulant, with symmetric circulant weights.

See [11] for the definitions of parallel and circulant matrices. Theorem 14 can be generalized to hold also for Cases 2 and 3, this involves only trivial algebra, and will not be done here. In general, it will be possible to find a realization of the transformed interconnection matrix provided:

- U and V are real, and
- the weights used have a specific structure. Denote the unitary block diagonal matrix used to transform the interconnection matrix by X and the block diagonal

matrix of weights W. For example, at the right of Fig. 7.4a $X = \text{diag}\{V^H, U^H\}$ and $W = \text{diag}\{W_I, W_P\}$. We require that $W = X^H \Sigma_W X$ if W premultiplies the rest of the μ interconnection matrix (as in Fig. 7.3), and that $W = X \Sigma_W(s) X^H$ if W postmultiplies the rest of the μ interconnection matrix. Here Σ_W is a diagonal matrix for which the diagonal elements are realizable transfer functions.

The distillation column example studied in this paper is in Class 1. Although it is a relatively poor model, distillation column models given by scalar dynamics multiplied by a constant matrix have been used by numerous researchers (for example, see the references listed in [21]). Restricting the weights to be scalar times identity matrices, as is done for Class 1, is a relatively severe restriction, since it, for example, means that one output of the plant cannot be weighted more than any other output in the performance weight. However, in many cases this problem may be circumvented by scaling the plant.

Nominally identical units in parallel with interactions, for example pumps, compressors or heat exchangers in parallel, are described by parallel models. Hovd and Skogestad [12] have studied the robust optimal control of these systems in detail.

It is easy to show that such models are diagonalized by a real Fourier matrix.[3] To show how this diagonalization works on a simple example, a 2×2 parallel process has the model

$$G(s) = \begin{bmatrix} a(s) & b(s) \\ b(s) & a(s) \end{bmatrix} \tag{7.26}$$

The 2×2 Fourier matrix, which diagonalizes $G(s)$, is

$$F = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{7.27}$$

Applying this to $G(s)$ gives

$$\Sigma_G = F G(s) F^T = \begin{bmatrix} a(s) + b(s) & 0 \\ 0 & a(s) - b(s) \end{bmatrix} \tag{7.28}$$

Paper machines [15, 23] and coating processes [2] have been approximated by symmetric circulant models. Though this class is more general than the class of parallel models, symmetric circulant models are diagonalized by the same real Fourier matrix.

[3]The standard Fourier matrix of [5] has pairs of columns which are complex conjugates of each other. A real Fourier matrix is defined by replacing each of these pairs of columns by an appropriate scaling to the addition and subtraction of the columns (see [11] for details).

Whether the μ-optimal controller is an SVD controller also depends on the location and structure of the perturbation blocks.[4] The μ-optimal controller will be a SVD controller only if the transformations of the interconnection matrix involved do not make the structure of the resulting perturbation matrix $\tilde{\Delta}$ more conservative than the structure of the original perturbation matrix Δ. If the original problem contains only full perturbation blocks and/or multiplicative (or inverse multiplicative) repeated scalar perturbation blocks, the structure of $\tilde{\Delta}$ will equal the structure of Δ. On the other hand, if the problem contains diagonal or additive repeated scalar blocks, the structure of $\tilde{\Delta}$ may be more conservative than the structure of Δ. For the special case when the plant is described by a normal transfer function matrix (such as for parallel or symmetric circulant plants), additive repeated scalar perturbation blocks do not make the structure of $\tilde{\Delta}$ any more conservative than the structure of Δ.

Even when the structure of the transformed perturbation matrix $\tilde{\Delta}$ is more conservative than the structure of the original perturbation matrix Δ, it may still be useful to perform D-K iteration on the transformed system to get initial D-scales for performing D-K iteration on the original system.

For the cases of H_2- and H_∞-optimal control there is no perturbation block (Δ) in the problem, and considerations about the structure of the perturbation block therefore do not apply. However, both the H_2 and H_∞ norms are invariant under unitary transformations, and the optimal controller will have the structure of an SVD controller provided $\tilde{N}$ is realizable, as discussed above for the μ-optimal controller.

7.4.4 Other Uses of the SVD Controller Structure

Design of controllers with a low number of states

μ synthesis using D-K iteration is known for resulting in controllers with many states. The SVD structure can also be used for designing controllers with a low number of states. Using V as a pre-compensator and U^H as a post-compensator, we are left with n SISO controllers to design for a plant of dimension $n \times n$. In this way, Engstad [9] managed to design a controller with four states for the distillation column used as an example in this paper which gave a value of $\mu = 1.036$.

[4]Recall that the performance specifications can be written in terms of performance perturbation blocks. Thus without loss of generality we can speak only of the locations of the perturbation blocks and not on the locations of the transfer functions of interest for performance.

Use as a Controllability Measure

The SVD controller structure can be used for obtaining a simple lower bound on the achievable value for μ. The frequency response of the μ interconnection matrix can be decomposed frequency-by-frequency. At each frequency the simplified D-K iteration can be used. Clearly, in this case state-space based algorithms cannot be used for the controller synthesis part (K) of D-K iteration, but because each design subproblem at a fixed frequency only involves finding one complex scalar, finding this complex scalar should be relatively simple. This frequency-by-frequency approach is not likely to lead to any realizable controller, since issues such as causality and phase-gain relationships are ignored. Instead, the resulting value for μ will be a lower bound on the structured singular value obtainable by any realizable controller, and may therefore be regarded as a controllability measure.

Since state-space based synthesis algorithms cannot be used anyway in the controller synthesis part of D-K iteration, we are for this case not concerned with the realizability of the decomposed μ interconnection matrix. This controllability measure can therefore be calculated also for processes for which U and V are both complex and varying with frequency.

In [16] it is also suggested to use the upper bound on μ on a frequency-by-frequency basis as a controllability measure. Use of the SVD structure in the calculation of the upper bound will simplify the calculations involved.

7.5 Conclusions

We have shown that the robustness of an SVD controller is insensitive to the structure of the uncertainty for cases with only one uncertainty block, if the uncertainty block is multiplicative or inverse multiplicative. For the distillation control problem in [22], we further showed that the μ-optimal controller has the structure of an SVD controller. This finding may be used to simplify the controller synthesis (K) part of the D-K iteration procedure used for synthesizing μ-optimal controllers.

Conditions for when the optimal controller in general has the structure of an SVD controller have been discussed, focusing on the issues of realizability of the transformed interconnection matrix $\tilde{N}$ (see Fig. 7.4b) and whether the transformation makes the structure of the perturbation block (Δ) more conservative.

Theree cases have been given for which the compensation given by unitary transformations of the SVD controller handles the multivariable effects in an optimal manner, provided the problem contains no independent diagonal perturbation blocks (or such

blocks are approximated by full blocks). Tuning for robustness and performance may for these cases be obtained by a simple diagonal controller. These results can be used to simplify D-K iteration for the synthesis of robust optimal controllers for these three classes of problems, and also show how the SVD controller structure can be used to find robust controllers with only a few adjustable parameters and a low number of states.

Acknowledgements. The authors thank Petter Lundström for providing Fig. 7.6, and Pål Kristian Engstad for giving us the initial idea of looking at the optimality of the SVD controller for robust control. The second author thanks the American-Scandinavian Association of Los Angeles and the Norway-America Association in Oslo, Norway for support during the author's stay in Trondheim during the summer of 1992; and he thanks the Fannie and John Hertz Foundation for their continued support.

References

[1] Balas, G., Doyle, J. C., Glover, K., Packard, A. and Smith, R. (1991). μ-analysis and Synthesis Toolbox. User's guide. MUSYN Inc., Minneapolis, USA.

[2] Braatz, R. D., Tyler, M. L., Morari, M., Pranckh, F. R. and Sartor, L. (1992). Identification and Cross-directional Control of Coating Processes: Theory and Experiments. *Proc. American Control Conference*, Chicago, USA, pp. 3021-3025.

[3] *Proceedings of the 30th Conference on Decision and Control*, (1991), Brighton, England, Session T2-7.

[4] Chen, J. and Freudenberg, J. S. (1992). Robust Performance with Respect to Diagonal Input Uncertainty. *IEEE Trans. Autom. Control*, **37**, 5, pp. 658-662.

[5] Davis, P. J. (1979). *Circulant Matrices.* Wiley & Sons, New York, USA.

[6] Doyle, J. C. (1982). Analysis of Feedback Systems Under Structured Uncertainty. *IEE Proc.*, **129**(D), pp. 242-250.

[7] Doyle, J. C. (1985). Structured Uncertainty in Control System Design. *Proc. of the 24th IEEE Conf. on Decision and Control*, Ft. Lauderdale, Florida, USA., pp. 260-265.

[8] Doyle, J. C., Glover, K., Khargonekar, P. and Francis, B. (1989). State-space Solutions to Standard H_2 and H_∞ Control Problems. *IEEE Trans. Autom. Control*, **34**, 8, pp. 831-847.

[9] Engstad, P. K. (1991). Comparison of Robustness Between Different Control Algorithms. *Diploma Thesis*, Dept. of Engineering Cybernetics, University of Trondheim-NTH.

[10] Freudenberg, J. S. (1989). Analysis and Design for Ill-conditioned Plants. Part 2: Directionally Uniform Weightings and an Example. *Int. J. Control*, **49**, 3, pp. 873-903.

[11] Hovd, M. (1992). Studies on Control Structure Selection and Design of Robust Decentralized and SVD Controllers. *Dr. Ing. Thesis*, University of Trondheim-NTH.

[12] Hovd, M. and Skogestad, S. (1992). Robust Control of Systems Consisting of Symmetrically Interconnected Subsystems. *Proc. American Control Conference*, Chicago, USA, pp. 3021-3025.

[13] Hung, Y. S. and MacFarlane, A. G. J. (1982). *Multivariable Feedback: A Quasi-Classical Approach*, Lecture Notes in Control and Information Sciences, 40, Springer Verlag, Berlin, Germany.

[14] Lau, H., Alvarez, J. and Jensen K. F. (1985). Synthesis of Control Structures by Singular Value Analysis: Dynamic Measures of Sensitivity and Interaction. *AIChE Journal*, **31**, 3, pp. 427-439.

[15] Laughlin, D. L., Morari, M. and Braatz, R. D. (1992). Robust Performance of Cross-directional Basis-weight Control in Paper Machines. *Automatica*, preprint.

[16] Lee, J. H., Braatz, R. D., Morari, M. and Packard, A. (1992). Screening Tools for Robust Control Structure Selection. *Automatica*, preprint; also AIChE Annual Meeting, Los Angeles, 1991.

[17] Lundström, P., Skogestad, S. and Wang, Z.-Q. (1991). Performance Weight Selection for H-infinity and μ-control Methods. *Trans. Inst. MC*, **13**, 5, pp. 241-252.

[18] Lundström, P. (1992). Personal communication.

[19] Packard, A. (1988). What's New with μ: Structured Uncertainty in Multivariable Control. *PhD Thesis*, University of California, Berkeley.

[20] Skogestad, S., Lundström, P. and Hovd, M. (1989). Control of Identical Parallel Processes. *AIChE Annual Meeting*, Paper 167Ba, San Francisco, USA.

[21] Skogestad, S. and Morari, M. (1987). The Dominant Time Constant for Distillation Columns. *Comput. Chem. Engng.*, **11**, 6, pp. 607-617.

[22] Skogestad, S., Morari, M. and Doyle, J. C. (1988). Robust Control of Ill-conditioned Plants: High Purity Distillation. *IEEE Trans. Autom. Control*, **33**, 12, pp.1092-1105.

[23] Wilhelm, R. G. Jr. and Fjeld, M. (1983). Control Algorithms for Cross Directional Control: The State of the Art. *Preprints 5th IFAC PRP Conference*, Antwerp, Belgium, pp. 139-150.

[24] Yaniv, O. and Barlev, N. (1990). Robust Non Iterative Synthesis of Ill-conditioned Plants. *Proc. American Cont. Conf.*, San Diego, USA, pp. 3065-3066.

Chapter 8

Final Discussion, Conclusions and Directions for Future Work

8.1 Discussion

This thesis deals with different aspects of "structure" in control systems. It starts with discussing whether imposing structure on a control system is a sensible thing to do. It is argued that for most plants it is sensible to use a control system with three layers. Starting from the bottom these layers are[1]:

1. The regulatory control level.
2. The supervisory control level.
3. Plant wide optimization level.

It is found that there is a further need for imposing structure *within* the regulatory control system. In practice, the regulatory control level is therefore highly decentralized, consisting of single loops. In each loop, a manipulated variable is used to keep a controlled variable at a setpoint which is determined by the higher levels in the control hierarchy.

Having established the need for imposing a specific structure on the control system, the thesis then deals with how to select a sensible structure for the regulatory control level. The control structure selection problem for the regulatory control level consists of the selection of controlled and manipulated variables and (since the regulatory control level uses mainly decentralized control) the pairing of controlled and manipulated variables.

[1]There are also other layers of decision making, but in this thesis these other layers are not considered to be a part of the control system

The tools for control structure selection for the regulatory control level are then applied to the riser-regenerator section of the FCC process. It is demonstrated that proper control structure selection for the regulatory control of the FCC process is critical. Specifically, the choice of controlled variables is shown to determine the achievable bandwidth of the regulatory control level.

After having determined what would be a sensible control structure, one also has to be able to design a controller for that control structure. The design of decentralized controllers is therefore studied, starting from the premise that the control structure has been determined à priori. Standard controller synthesis algorithms (e.g. H_2- or H_∞-synthesis) cannot handle a requirement for a specific structure for the controller. Instead three practical approaches to the design of decentralized controllers have evolved:

- Parameter optimization.
- Independent design.
- Sequential design.

The thesis discusses the advantages and drawbacks of these three different design approaches, and presents new results on independent and sequential design.

In some cases a specific structure for the controller does not arise because of any à priori decision, but rather because of the structure of the controller design problem itself. In the latter part of this thesis, it is shown how the SVD controller structure arises naturally for certain classes of controller design problems[2] for H_2-, H_∞-, and μ-optimal control. It is thereafter shown how the knowledge of the structure of the optimal controller can be used to simplify controller synthesis.

In summary, this thesis deals with four different aspects of structure in control systems:

1. The need for structuring the control system.
2. How to determine a sensible structure for the regulatory control level.
3. The design of decentralized (i.e., structured) controllers.
4. A specific controller structure arising from the structure of the controller design problem itself.

[2]When the individual units have multiple inputs and/or outputs, the controllers found in Chapter 6 are strictly speaking not SVD controllers, but the controller structure is very similar.

8.2 Conclusions

This thesis has dealt with control structure selection and design of robust controllers. The main contributions of this thesis are summarized below:

- **Chapter 2.** The relationship between the sign of the RGA elements and RHP zeros is clarified in Thm. 1. From Thm. 2 it is clear that the RGA is a measure of the sensitivity to uncertainty in the individual elements. The proof of Thm. 2 is new, and it is much simpler than the original proof of Yu and Luyben [17].

  The Performance Relative Gain Array (PRGA) and the Closed Loop Disturbance Gain (CLDG) are introduced. For decentralized control the PRGA and CLDG give loop gain requirements for setpoint following and disturbance rejection, respectively, for frequencies below the bandwidths of the individual loops. The PRGA and CLDG depend on on the plant alone, and may therefore be used to evaluate controllability under decentralized control. Specifically, pairings of inputs and outputs giving large elements of the PRGA and/or the CLDG in the bandwidth region should be avoided. The CLDG may also be used to pinpoint the need for feedforward control or modification of the process to reduce the effect of disturbances.

- **Chapter 3.** The Niederlinski index pairing criterion and the steady-state RGA pairing criterion have been generalized to hold for open loop unstable plants. It is found that for unstable processes it may be preferable to choose pairings giving negative values for the Niederlinski index and/or the steady state RGA. The chapter discusses the use of several tools for control configuration selection, and comments on how they must be applied differently or not be used at all for open loop unstable plants. It is demonstrated that it is advantageous to select pairings which around the closed loop bandwidth result in a RGA matrix that is close to identity (or, equivalently, a PRGA matrix that is close to triangular, with diagonal elements close to one).

- **Chapter 4.** The control structure selection for the regulatory control level is first studied in general and a set of specific objectives for regulatory control are introduced. Subsequently, the riser-regenerator section of a FCC plant is studied, with emphasis on the partial combustion mode of operation. It is demonstrated that the selection of controlled variables for the regulatory control level has a strong effect on the controllability of this plant. Specifically, with the model used in this thesis it is shown that the conventional control structure (Pohlenz, [13])

has a RHP transmission zero which severely restricts the achievable bandwidth. The Kurihara control structure (Kurihara, [8]) has a RHP transmission zero at a frequency that is one decade higher than the RHP transmission zero for the conventional control structure, and the achievable bandwidth for the Kurihara control structure is therefore less severly restricted. However, neither the Hicks control structure (Hicks, [6]) or the riser-regenerator control structure have any RHP transmission zero, and one of these two control structures are therefore preferred over the conventional or Kurihara control structures. For the complete combustion mode of operation, the conventional control structure has no RHP transmission zero, and is therefore preferable to the alternative control structures studied. Feed flowrate disturbances will affect the riser outlet temperature at high frequencies, and thus only slow changes in the feed flowrate should be made. The other disturbances studied can be satisfactorily rejected by the regulatory control system. The findings in Chapter 4 are shown to be insensitive to changes in the operating point and parametric uncertainty.

- **Chapter 5.** Skogestad and Morari's independent design procedure [15] for design of robust decentralized controllers has been extended and made less conservative. Specifically, it is shown how to find bounds on the filter time constant for a decentralized IMC controller such that robust stability and/or performance can be guaranteed. In contrast, Skogestad and Morari derived bounds on the sensitivity function and complementary sensitivity function for the individual loops, which results in a much larger class of possible designs and hence is more conservative.

  Key steps in the new procedure, which can also be applied to other controller parametrizations, are:

  1. Treat the controller parameters as real design uncertainty.
  2. Ensure that only positive parameters are allowed by the uncertainty description.
  3. Find bounds on the controller parameters such that robust stability or robust performance is guaranteed.

  The concept of Robust Decentralized Detunability is introduced. If a decentralized controller is used to control a plant, and any of the loops in the system can be detuned by an arbitrary amount without endangering robust stability, then the system is said to be Robust Decentralized Detunable. It is shown how to

find bounds on the filter time constants for a decentralized IMC controller which ensure Robust Decentralized Detunability.

For sequential design of decentralized controllers it is shown how one can incorporate into the controller design problem for loop k a simple estimate of the effect of closing the loops that are still open. For the examples studied, in which robust control in terms of μ is considered, this improvement of the sequential design procedure gives controllers that are only slightly inferior to the best decentralized controller (with the same controller parametrization) that has been found.

The two design procedures outlined above both include model uncertainty explicitly, and both may also be used to design multivariable controllers which contain a diagonal part, e.g. SVD controllers (Chapter 7). In this case it is crucial to include the model uncertainty to avoid poorly conditioned controllers (e.g., decouplers).

- **Chapter 6.** In this chapter the control of plants consisting of n interacting units that are nominally identical. For H_2- and H_∞-optimal control, it is shown how the controller synthesis problem can be decomposed into two independent synthesis problems of smaller dimension. If each individual unit has n_i inputs and n_o outputs, each of these two design problems involve a "plant" of dimension $n_o \times n_i$. Since the overall plant is of dimension $n \cdot n_o \times n \cdot n_i$, the reduction in size for the synthesis subproblems is large if the number of units n is large. For H_∞-optimal control, the resulting overall controller will be super-optimal, since it optimizes the H_∞ objective in n directions.

  For μ-optimal control, the "D" part of D-K iteration (μ analysis) involves a block diagonal "plant" with two blocks, each block of dimension $n_o \times n_i$. In some cases these two blocks can be considered independently. For the "K" part of D-K iteration (controller synthesis) the two blocks of the block diagonal "plant" can always be considered independently. However, this reduction in problem size for D-K iteration requires that any independent diagonal perturbation blocks are approximated by full blocks, which may introduce conservatism.

  If $n_i = n_o = 1$, the controllers which minimize the H_2- and H_∞- norms both have the structure of an SVD controller, as has the controller which minimizes the upper bound on μ ($\mu(M) \leq D_l M D_r^{-1}$).

  For decentralized control, conditions for Decentralized Integral Controllability are derived for the case $n_i = n_o = 1$.

- **Chapter 7.** For the distillation control problem in [14] it has been shown numerically [5] that the robust optimal controller[3] has the structure of an SVD controller, and this is proven in general for certain classes of problems. It is shown that provided there are no independent, diagonal uncertainty blocks in the problem (or if they are approximated by full perturbation blocks), the robust optimal controller is an SVD controller for the following three classes of problems:

  1. The plant is described by scalar dynamics multiplied by a constant matrix, with scalar times identity weights.
  2. The plant is parallel, with parallel weights.
  3. The plant is symmetric circulant, with symmetric circulant weights.

  In these cases the compensation given by unitary transformations of the SVD controller handles the multivariable effects in an optimal manner, provided the problem contains no independent scalar uncertainty blocks (or such blocks are approximated by full blocks). Tuning for robustness and performance may for these cases be obtained by a simple diagonal controller. The results in Chapter 7 can be used to simplify D-K iteration for the synthesis of robust optimal controllers for these three classes of problems, and also show how the SVD controller structure can be used to find robust controllers with only a few adjustable parameters and a low number of states.

  The robustness of an SVD controller is shown to be insensitive to the structure of the uncertainty for cases with only one uncertainty block, provided this uncertainty block is either multiplicative or inverse multiplicative.

8.3 Directions for Future Work

8.3.1 Structuring of the Control System

A more mathematical "proof" or motivation is needed for why "simplicity of structure" is preferable for the control system. The "proof" would have to quantify the cost of obtaining and maintaining a model of the plant (cost of information), the cost of making changes to the control system, the reliability of the control system, and possibly other factors. Such a "proof" will require a lot of work and may be difficult to find for the general case.

[3] The controller which minimizes the upper bound on μ ($\mu(M) \leq D_l M D_r^{-1}$) is termed the robust optimal controller.

This can be seen as a part of the more general problem of obtaining the mathematical framework needed to get good mathematicians involved in important engineering problems.

8.3.2 Control Structure Selection

Further work is still needed to make the task of control structure selection easier. Two specific areas for further work are:

- **Simplification of Controllability Measures.** Recent work aims to take into account the effects of uncertainty [11, 12] and constraints in the inputs or outputs [16]. Ideally, controllability measures should be easy to calculate, in order to make it possible to test a large number of alternatives with relatively little effort. Unfortunately, many of the criteria in [11, 12, 16] require heavy computations. It would therefore be advantageous to either develop simpler criteria or to find closed form solutions to the calculation of the measures involved. The test for the existence of a controller with integral action which fulfills the robust performance requirement, presented by Braatz et al. [2], is one example of such a simple measure.

- **Inherent Bandwidth Limitations with Decentralized Control.** Individual elements of the transfer function matrix can contain RHP zeros that do not correspond to a RHP transmission zero of the plant. These RHP zeros in the individual elements may disappear when the other loops in the control system are closed. In this case the RHP zero in the element does not cause any inherent bandwidth limitation for the loop.

  Likewise, closing some loops may introduce RHP zeros in the transfer function element for a loop that is still to be closed. Bristol [3] claims that pairing on a negative RGA element will cause the corresponding transfer function element to become nonminimum phase (either with a RHP zero or a RHP pole) when the other loops are closed. This claim is not correct, as trivial counterexamples can be found. To make this claim correct, one will at least require additional assumptions about the speed of the direct effect of a control move relative to the indirect effect (via the other loops).

  A better understanding of the role of RHP zeros in individual transfer function elements when using decentralized control would therefore be valuable.

8.3.3 Controllability Case Studies

Controllability case studies should be performed, in order to investigate how design modifications may improve controllability. Controllability analysis should be used not only for control structure selection, but also during process design to ensure that the process that is designed has acceptable controllability. This point is becoming more important, as modern processes are designed with tight integration between units, and the resulting designs are prone to result in control problems with a high degree of interaction.

8.3.4 μ_{RP} as a Controllability Measure

Lee et al. [12] suggest using the frequency-by-frequency upper bound on μ_{RP}, $\min_K \bar{\sigma}(DMD^{-1})$, as a controllability measure[4]. Alternatively, one may avoid D-K iteration and minimize $\mu_{RP}(M(K))$ directly provided a relatively simple parametrization of the controller is used. Specifically, it has been shown in Chapter 7 that the SVD structure is optimal in many cases, and this may be used as one such controller parametrization.

8.3.5 Controllability Analysis — Consequences for Supervisory Control

Controllability analysis may reveal that the regulatory control level is sensitive to changes in some variables that are used as manipulated variables by the supervisory control level. This will be an argument for ensuring that the supervisory control level in the control system change these manipulated variables slowly. However, limiting the rate of change of the manipulated variables will of course also limit the achievable bandwidth for the supervisory control level. This means that there may be a tradeoff in the design of the supervisory control level, and it is not clear how this tradeoff should be made.

8.3.6 Controllability Analysis for the FCC

The scope of the controllability analysis should be enlarged, by including the downstream fractionation column and wet gas compressor, and other units with strong interactions with the riser-regenerator section of the FCC process. Increasing the scope of the controllability analysis will make it more realistic, and the result of such an analysis would be very interesting. Such a controllability analysis would require substantial modeling work, or access to unpublished models possibly held by operating companies

[4] μ_{RP} denotes the structured singular value for Robust Performance

or liCence holders. However, it is the experience of this author that operating companies and licence holders for the FCC process are very secretive, and it therefore does not appear that anyone from academia would be given access to and/or be allowed to publish the necessary information.

8.3.7 Multiple Steady States for the FCC

In the literature on steady state multiplicity in FCCs, e.g. [7, 10, 4, 1], a number of conflicting views are presented, and the authors base their analysis on different models and assumptions. An attempt to clarify the situation with respect to steady state multiplicity would therefore be worthwhile.

8.3.8 SVD Controllers and Other "Compensator-based" Controllers

Further investigations into the design of SVD controllers and other "compensator-based" controllers are needed. The SVD controller structure or other "compensator-based" controller structures should be tried on selected examples, and the diagonal part can be designed using the methods described in Chapter 5. In this case it is important to include uncertainty in the design problem, since the overall controller is not diagonal.

8.3.9 Robust Decentralized Control

Tests for Robust Decentralized Detunability for conventional decentralized controllers (e.g., PI or PID controllers) should be developed, along the lines of the test for Robust Decentralized Detunability for decentralized IMC controllers that is described in Chapter 5. This is not quite as straight forward as it may appear, since setting the proportional gain to zero in a controller with integral action will remove the feedback around the integrator, which will then be on the limit of instability. The problem formulation must avoid this problem, which is only of a mathematical nature and does not correspond to a physical problem.

8.3.10 Cross-directional Control of Paper Machines

The direct application of the theory of Chapters 6 and 7 to the cross-directional control of paper machines requires that edge effects are ignored. Further work is needed to include the edge effects in the problem, or to study the effect of the error that is introduced by neglecting the edge effects. The assumption of repeated scalar uncertainty in the actuators, which is implicitly used by Laughlin et al. [9], should be avoided.

References

[1] Arandes, J. M. and de Lasa, H. I. (1992). Simulation and Multiplicity of Steady States in Fluidized FCCUs. *Chem. Engng. Sci.*, **47**, pp. 2535-2540.

[2] Braatz, R. D. and Morari, M. and Lee, J. H. (1991). Necessary/Sufficient Loop-shaping Bounds for Robust Performance. *AIChE Annual Meeting*, Los Angeles, California, paper 154d.

[3] Bristol, E. H. (1981). The right half plane'll get you if you don't watch out. *Joint American Control Conference (JACC)*, paper TA-7A.

[4] Edwards, W. M. and Kim, H. N. (1988). Multiple Steady States in FCC Unit Operations. *Chem. Engng. Sci.*, **43**, pp. 1825-1830.

[5] Engstad, P. K. (1991). Comparison of Robustness Between Different Control Algorithms. *Diploma Thesis*, Dept. of Engineering Cybernetics, University of Trondheim-NTH.

[6] Hicks, R. C., Worrell, G. R. and Durney, R. J. (1966). Atlantic Seeks Improved Control; Studies Analog-Digital Models. *The Oil and Gas Journal*, Jan. 24, 97-105.

[7] Iscol, L. (1970). The Dynamics and Stabilityt of a Fluid Catalytic Cracker. *Proc. Joint Automatic COntrol Conference*, Atlanta, Georgia, pp. 602-607.

[8] Kurihara, H. (1967). Optimal Control of Fluid Catalytic Cracking Processes. *Ph.D. Thesis* MIT.

[9] Laughlin, D. L., Morari, M. and Braatz, R. D. (1992). Robust Performance of Cross-Directional Basis-Weight Control in Paper Machines. Submitted to *Automatica*, preprint.

[10] Lee, W. and Kugelman, A. M. (1973). Number of Steady-state Operating Points and Local Stability of Open-loop Fluid Catalytic Cracker. it Ind. Eng. Chem. Proc. Des. Dev., **12**, pp. 197-204.

[11] Lee, J. H. and Morari, M. (1991). Robust Measurement Selection. *Automatica*, **27**, 3, pp. 519-527.

[12] Lee, J. H., Braatz, R. D., Morari, M. and Packard, A. (1992). Screening Tools for Robust Control Structure Selection. *Automatica*, preprint; also AIChE Annual Meeting, Los Angeles, 1991.

[13] Pohlenz, J. B., (1963). How Operational Variables Affect Fluid Catalytic Cracking. *Oil and Gas Journal*, April 1, pp. 124-143.

[14] Skogestad, S., Morari, M. and Doyle, J. C. (1988). Robust Control of Ill-conditioned Plants: High Purity Distillation. *IEEE Trans. Autom. Control*, **33**, 12, pp.1092-1105.

[15] Skogestad, S. and Morari, M. (1989). Robust Performance of Decentralized Control Systems by Independent Design. *Automatica*, **25**, 1, pp.119-125.

[16] Skogestad, S. and Wolff, E. A. (1992). Controllability Measures for Disturbance Rejection. *Preprints IFAC Workshop on Interactions between Process Design and Process Control*, London, England.

[17] Yu, C.-C. and Luyben, W.L. (1987). Robustness with Respect to Integral Controllability, *Ind. Eng. Chem. Res.*, **26**, pp. 1043-1045.

Appendix A

Model Requirements for Model Predictive Control

M. Hovd *, J. H. Lee†, M. Morari‡
Chemical Engineering Department
210-41 California Institute of Technology
Pasadena, CA 91125, USA

Abstract

A method for deriving truncated step response models for Model Predictive Control (MPC) is derived and demonstrated. Truncated step response models can result in a dramatic reduction in the number of step response coefficients needed, and can therefore significantly reduce the computational load on the control system. The truncated models are derived in a way which attempts to minimize the robustness degradation caused by the modeling error introduced by truncation. The truncation of the step response model only affects the state estimation part of the MPC. The algorithm for finding the control move is not affected as long as the prediction horizon is kept smaller than the truncation time.

*Present address: Chemical Engineering Department, Norwegian Institute of Technology, N-7034 Trondheim, Norway.

†Present address: Department of Chemical Engineering, Auburn University, Auburn, AL 36849-5127

‡To whom all correspondence should be addressed: phone (818)356-4186, fax (818)568-8743, e-mail MM@IMC.CALTECH.EDU

A.1 Introduction

Many chemical plants have outputs whose response times differ widely. In conventional Model Predictive Control (MPC), using a step response model, it is necessary to use a model which represents the behaviour of the slow outputs with sufficient accuracy. Thus the model must be able to predict plant behaviour far into the future. However, the presence of fast outputs necessitates the use of a short sampling time in order to enable good control of these outputs. As a result, a high-order model is needed to describe the plant, leading to a large computational load on the control system, and possibly causing computational delays. The same problem arises if the responses of all outputs are dominated by a large time constant, but fast closed loop response is desired.

MPC has traditionally been used only for asymptotically stable plants, with ad hoc corrections being required for integrating plants. Recent developments have extended the use of MPC to integrating and unstable plants in a consistent manner [1]. These developments are used in this paper to find models which resolve the dilemma described above. In the models found in this paper the the step response model is truncated at a time such that the fast modes have settled, and a very simple state space realization is used to describe the remaining dynamics.

A.2 Truncated Models

The truncated models used in this work are formulated as follows:

$$\tilde{\mathcal{Y}}(k) = M^{\tau}\tilde{\mathcal{Y}}(k-1) + \mathcal{S}\Delta u(k-1) + T\Delta w(k-1) \tag{A.1}$$

$$\tilde{y}(k) = N\tilde{\mathcal{Y}}(k) \tag{A.2}$$

where

$$\tilde{\mathcal{Y}}(k) = [\tilde{y}(k|k), \cdots, \tilde{y}(k+n-1|k), \tilde{y}_S, \tilde{y}_D]^T \tag{A.3}$$

$$M^{\tau} = \begin{bmatrix} 0 & I_{n_y} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n_y} & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_{n_y} & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & I_{n_y} & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & I_{n_y} & C_T & C_D \\ 0 & 0 & 0 & \cdots & 0 & 0 & A_T & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & A_D \end{bmatrix} \tag{A.4}$$

$$N = [\, I_{n_y} \;\; 0 \;\; \cdots \;\; 0 \,] \tag{A.5}$$

$$\mathcal{S} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_{n-1} \\ S_n \\ B_T \\ 0 \end{bmatrix}; \quad T = \begin{bmatrix} S_{D1} \\ S_{D2} \\ \vdots \\ S_{Dn-1} \\ S_{Dn} \\ 0 \\ B_D \end{bmatrix} \tag{A.6}$$

where

$$S_i = \begin{bmatrix} S_{1,1,i} & S_{1,2,i} & \cdots & S_{1,n_u,i} \\ S_{2,1,i} & S_{2,2,i} & \cdots & S_{2,n_u,i} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n_y,1,i} & S_{n_y,2,i} & \cdots & S_{n_y,n_u,i} \end{bmatrix}; \quad (i = 1, \cdots, n) \tag{A.7}$$

The element $S_{k,l,i}$ is the ith step response coefficient for the effect of a step in manipulated variable l on output k. The S_{Di}'s are step response coefficient matrices for the disturbances, and correspond to the S_i's for manipulated variables. The change in the manipulated variable at time k is given by $\Delta u(k) = u(k) - u(k-1)$, and $\tilde{y}(k)$ is the current process output (for MIMO systems $\tilde{y}(k)$ is a vector). The dynamic states of the system are contained in the vector $\tilde{\mathcal{Y}}(k)$. Each dynamic state $\tilde{y}(\ell|k)$ has a special interpretation: it is the expected process output at time ℓ assuming the manipulated variable does not change at present or in the future (*i.e.*, $\Delta u(k+j) = 0$ for $j \geq 0$) and that no disturbances enter at present or in the future (*i.e.*, $w(k+j) = 0$ for $j \geq 0$). The vectors $\tilde{y}_S$ and $\tilde{y}_D$ contain dummy states through which the slow effects of the change in the manipulated variables and the effects of the disturbances, respectively, enter. Note that the upper left corner of M^{τ} equals M^S in [1]. The fast dynamics are therefore described by M^S and $S_1, \cdots, S_n$, with the slow dynamics being described by A_T, B_T and C_T. A_T, B_T and C_T constitute a state space description of the slow dynamics, and describe the effect of changes in the manipulated variables on the outputs after the truncation time n sampling intervals into the future. The matrix A_T is the autotransition matrix, containing the discrete pole locations for the slow dynamics, and B_T and C_T are input and output matrices, respectively. Similarly, A_D, B_D and C_D constitute a state space description of the slow effects of the disturbances.

First Order Dynamics for the Truncated Step Response. Although A_T, B_T and C_T in principle can describe a system of any order, we will from here on assume that the truncated part of the step response is modelled using first order responses for

each plant element. This results in the following equations:

$$A_T = \begin{bmatrix} a_{11} & & & \\ & a_{21} & & \\ & & \ddots & \\ & & & a_{n_y.n_u} \end{bmatrix} \tag{A.8}$$

$$B_T = \begin{bmatrix} \begin{bmatrix} b_{11} \\ \vdots \\ b_{n_y 1} \end{bmatrix} & & & \\ & \ddots & & \\ & & \begin{bmatrix} b_{1 n_u} \\ \vdots \\ b_{n_y.n_u} \end{bmatrix} \end{bmatrix} \tag{A.9}$$

$$C_T = \begin{bmatrix} \overbrace{I_{n_y} \quad I_{n_y} \quad \cdots \quad I_{n_y}}^{n_u} \end{bmatrix} \tag{A.10}$$

This means that the residual dynamics of the ij'th element of the transfer function is modelled as $b_{ij}/(z - a_{ij})$. In equations (A.8), (A.9) and (A.10) the responses for the slow dynamics are completely decoupled from each other. However, it is important to realize that fitting the coefficients in A_T and B_T corresponding to each output and manipulated variable independently, without taking into account the coefficients used to fit the other responses, may not give a good model for the overall system, especially for ill-conditioned plants. This is because directions in the plant will then be neglected. The values of all a_{ij} and b_{ij} should therefore be fitted simultaneously.

A.3 State Estimation

The model for the disturbances in the previous section is very general. However, we will in this paper restrict ourselves to consider only disturbances which can be described by stable or integrating transfer functions of first order, as this will allow a very simple parametrization of the optimal Kalman filter gain.

First Order Disturbance Dynamics. We assume that the disturbances $w(k)$ are white noise. The following choices for S_D, A_D, B_D and C_D will then give first order disturbance dynamics.

$$S_{Di} = 0 \ \forall i \tag{A.11}$$

$$C_D = I_{n_y} \tag{A.12}$$

$$B_D = I_{n_y} \tag{A.13}$$

A_D in M^τ is a diagonal matrix

$$A_D = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{n_y} \end{bmatrix} \tag{A.14}$$

$$0 \leq \alpha_i \leq 1 \tag{A.15}$$

With M^τ as given above, the effect of disturbances on the outputs will be of first order. Hence, $\alpha_i = 1$ implies a ramp disturbance at the i'th output (white noise integrated twice) and $\alpha_i = 0$ implies a step disturbance at the i'th output (white noise integrated once).

This model of the disturbance dynamics allows the following simple parametrization of the Kalman filter gain K:

$$K = \begin{bmatrix} \mathcal{I} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} f_{1a} & & \\ & \ddots & \\ & & f_{n_a} \end{bmatrix} + \begin{bmatrix} 0 \\ I_{n_y} \\ A_D + I_{n_y} \\ \vdots \\ \sum_{\ell=0}^{n-2} A_D^\ell \\ 0 \\ A_D^{n-1} \end{bmatrix} \begin{bmatrix} f_{1b} & & \\ & \ddots & \\ & & f_{n_b} \end{bmatrix} \tag{A.16}$$

$$f_{ib} = \frac{\alpha_i f_{ia}^2}{1 + \alpha_i - \alpha_i f_{ia}} \tag{A.17}$$

$$\mathcal{I} = \begin{bmatrix} \overbrace{I_{n_y} \quad I_{n_y} \quad \cdots \quad I_{n_y}}^{n} \end{bmatrix}^T \tag{A.18}$$

The optimal state estimate is thus:

$$\begin{aligned} \bar{\mathcal{Y}}(k|k) &= (M - KNM)\bar{\mathcal{Y}}(k-1|k-1) \\ &+ K\hat{y}(k) + (I - KN)\mathcal{S}\Delta u(k-1) \end{aligned} \tag{A.19}$$

where $\hat{y}(k)$ denotes the measurement at time k. It is worth noting that as a result of Eq. (A.17) the Kalman filter contains only one adjustable parameter for each controlled variable. It is therefore feasible to tune the Kalman filter online.

A.4 Prediction

The dynamic states of the optimal estimators developed in the previous section represent the current and future outputs assuming all current and future inputs are zero (*i.e.*, $\Delta u(k+j) = 0$ for $j \geq 0$). The predictive controller computes the best current and future control moves based on the prediction of future outputs. Then future outputs can be expressed in terms of current and $(m-1)$ future inputs through the following equation:

$$\bar{\mathcal{Y}}_p^m(k+1|k) = M_p \bar{\mathcal{Y}}(k|k) + \mathcal{S}_p^m \Delta\mathcal{U}(k) \tag{A.20}$$

where

$$\mathcal{S}_p^m = \begin{bmatrix} S_1 & 0 & \cdots & \cdots & 0 \\ S_2 & S_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ S_m & S_{m-1} & \cdots & \cdots & S_1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ S_p & S_{p-1} & \cdots & \cdots & S_{p-m+1} \end{bmatrix} \tag{A.21}$$

$$M_p = \begin{bmatrix} I_{p\cdot n_y \times p\cdot n_y} & 0 \end{bmatrix} M^T \tag{A.22}$$

$$\Delta\mathcal{U}(k) = \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \\ \vdots \\ \vdots \\ \Delta u(k+m-1) \end{bmatrix} \tag{A.23}$$

The notation $\bar{\mathcal{Y}}_p^m(k+1|k)$ denotes the predicted future outputs up to time $k+p$ for constant inputs starting at time $k+m$, based on the measurements up to time k. Hence, we allow the flexibility of setting the number of future input moves m $(1 \leq m \leq p)$ differently from the output prediction horizon p. The equation (A.20) provides the "optimal" prediction of the future outputs based on the current measurements since $\bar{\mathcal{Y}}(k|k)$ is the optimal estimate of the states representing the current and future process outputs assuming $\Delta u(k+j) = 0, v(k+j+1) = 0$ and $w(k,j) = 0 \ \ \forall j \geq 0$ [2], where $v(k+j+1)$ is the measurement noise at time $k+j+1$.

A.5 Feedback Control

We adopt the following quadratic optimization objective (used in QDMC [4]):

$$\min_{\Delta\mathcal{U}(k)} ||\Gamma[\bar{\mathcal{Y}}_p^m(k+1|k) - \mathcal{R}(k+1)]||^2 + ||\Lambda\Delta\mathcal{U}(k)||^2 \tag{A.24}$$

$\mathcal{R}(k+1) = [r(k+1), \cdots, r(k+p)]^T$ is the future output reference vector. Γ and Λ are weighting matrices that are chosen to be diagonal for most cases. This optimization problem can be cast into the following least-squares problem:

$$\begin{aligned} \begin{bmatrix} \Gamma\mathcal{S}_p^m \\ \Lambda \end{bmatrix} \Delta\mathcal{U}(k) &= \begin{bmatrix} \Gamma & 0 \\ 0 & I_{m\cdot n_u} \end{bmatrix} \begin{bmatrix} \mathcal{R}(k+1) - M_p\bar{\mathcal{Y}}(k|k) \\ 0 \end{bmatrix} \\ &\equiv \begin{bmatrix} \Gamma\mathcal{E}(k+1|k) \\ 0 \end{bmatrix}_2 \end{aligned} \tag{A.25}$$

The solution to this least-squares problem is

$$\Delta\mathcal{U}(k) = \{(\mathcal{S}_p^m)^T\Gamma^T\Gamma\mathcal{S}_p^m + \Lambda^T\Lambda\}^{-1}(\mathcal{S}_p^m)^T\Gamma^T\Gamma\mathcal{E}(k+1|k) \tag{A.26}$$

The current control move is implemented:

$$\Delta u(k) = [\ I_{n_u} \ \ 0 \ \ \cdots \ \ 0\ ]\Delta\mathcal{U}(k) \tag{A.27}$$

The controller can be interpreted as a state-observer-based compensator since

$$\Delta u(k) = K_{MPC}(\mathcal{R}(k+1) - M_p\bar{\mathcal{Y}}(k|k)) \tag{A.28}$$

where

$$K_{MPC} = [\ I_{n_u} \ \ 0 \ \ \cdots \ \ 0\ ]\{(\mathcal{S}_p^m)^T\Gamma^T\Gamma\mathcal{S}_p^m + \Lambda^T\Lambda\}^{-1}(\mathcal{S}_p^m)^T\Gamma^T\Gamma \tag{A.29}$$

From the above it is clear that the error introduced by truncating the model only affects the state estimation, and will not affect the prediction or feedback control provided the prediction horizon, p is less than the number of step response coefficients used, n. It will often be sufficient to choose the prediction horizon shorter than the truncation time, especially if the residual dynamics are monotonic.

A.6 Requirements for Model Accuracy

The requirements for model accuracy will vary with frequency, depending on performance requirements and model uncertainty (which should not be confused with the

error in the model deliberately introduced by truncation). For instance, any error in the magnitude of the steady state gain less than 100% can be handled by a controller with integral action, provided the sign of the steady state gain is correct. At frequencies far beyond the closed loop bandwidth, even larger errors can be tolerated, as performance requirements are lax and the loop gain is very small. The need for high model accuracy therefore occurs primarily at intermediate frequencies. A procedure for choosing the a_{ij}'s and b_{ij}'s in equations (A.8) and (A.9) therefore needs to take the different accuracy requirements at different frequencies into account.

The measure adopted in this work is the μ performance weighted plant reduction measure of Rivera and Morari [5].

$$\| D_s N_{12} \hat{S} (P - \hat{P}) \hat{P}^{-1} \hat{H} N_{21} D_s^{-1} \|_\infty \tag{A.30}$$

Here D_s is the $D-$scale used in the calculation of the structured singular value [3] for robust performance for the full plant. N_{12} and N_{21} are the $1,2$ and $2,1$ blocks, respectively, of an affine parametrization of the μ interconnection matrix M_μ in terms of $\hat{S}$.

$$M_\mu = N_{11} + N_{12} \hat{S} N_{21} \tag{A.31}$$

$\hat{S}$ and $\hat{H}$ are the sensitivity function and complementary sensitivity function using the reduced plant model $\hat{P}$.

$$\hat{S} = (I + \hat{P}C)^{-1}; \quad \hat{H} = I - \hat{S} = \hat{P}C(I + \hat{P}C)^{-1} \tag{A.32}$$

P is the full model and C is the controller. Details on how to find affine parametrizations of M_μ in terms of $\hat{S}$ can be found in [6].

The use of $\hat{S}$ and $\hat{H}$ in the above measures implies knowledge of both the controller and the reduced order plant. Knowledge of the controller is also necessary in order to find the μ interconnection matrix M_μ which is needed to find the $D-$scale. Usually neither the controller nor the reduced order plant will be known a priori, and therefore $\hat{S}$ and $\hat{H}$ will not be known. We will therefore initially design a controller for P (the full model) which fulfills the robust performance requirements in terms of the structured singular value. The resulting controller can be used to calculate S and H, which are used instead of $\hat{S}$ and $\hat{H}$ in Eq. (A.30). The inaccuracy thus introduced is usually minor and can be tolerated since the final controller design will be checked by calculating the structured singular value for the closed loop system. Numerical experience suggests that Eq. (A.30) works well with S and H substituted for $\hat{S}$ and $\hat{H}$. This is reasonable as the controller design for the full order model reflects performance requirements and bandwidth limitations. The result will then be that the measure (A.30) will weight

errors at the same frequency similarly even if S and H are substituted for $\hat{S}$ and $\hat{H}$, provided S and H are approximately the same as $\hat{S}$ and $\hat{H}$.

If increased accuracy in the use of (A.30) is desired, iteration can be performed. That is, first a reduced order model is found by using S and H instead of $\hat{S}$ and $\hat{H}$ in (A.30). This reduced order model can be used to design a new controller, which is used both to update D_s, $\hat{S}$ and $\hat{H}$. These updated estimates can be used to find a new reduced model, and so on. Note that the $D-$scales will change for each iteration since the controller changes.

A.7 Heuristic Rules

Choosing the Truncation Time. The point of truncation should be chosen after the last inflection point on the step response. For MIMO systems the truncation time should be chosen after the last inflection point in any of the step responses. This will make it easier to find a first order model which fits the high frequency behaviour of the true plant well. Model accuracy at steady state can always be achieved by fitting the steady state value of the step response, but this will often lead to large model errors at frequencies which are important for feedback control.

In general, higher model accuracy is achievable the later the step response is truncated, as the fast modes will have settled more completely. The final choice for the truncation time will therefore have to weigh the model accuracy needed against the resulting number of step response coefficients.

Initial Guesses for a_{ij} and b_{ij}. A good model fit at high frequencies can be obtained if the values for a_{ij} and b_{ij} are chosen so that starting from the value of the step response at time n, the model also fits the step responses at times $n+1$ and $n+2$. With such choices for a_{ij} and b_{ij} the error in the step response will accumulate slowly (unless the second derivative of the step response changes very fast), and the error will therefore occur mainly at lower frequencies. Thus if a good model fit at high frequencies is considered important, an initial estimate of a_{ij} and b_{ij} can be obtained from:

$$b_{ij} = S_{n+1,ij} - S_{n,ij} \tag{A.33}$$

$$a_{ij} = (S_{n+2,ij} - S_{n+1,ij})/b_{ij} \tag{A.34}$$

These initial guesses are exact if the residual dynamics are exactly first order. However, a minimum requirement for low frequency accuracy is that the determinant of the steady state gain has the correct sign. If the sign of the determinant of the steady state gain of the model is wrong, the resulting closed loop system will be unstable,

because of the integral action inherent in MPC. Using the estimates in Eq. (A.33) and Eq. (A.34) gives no guarantee that the sign of the determinant of the steady state gain is correct.

We emphasize that the values for a_{ij} and b_{ij} found from equations (A.33) and (A.34) should only be considered as reasonable initial estimates, and a method which weighs the accuracy requirements at different frequencies should be used to determine the final values of the a_{ij}'s and b_{ij}'s. However, the criterion we have used to determine the best model, (A.30), is not convex. Minimizing (A.30) one may therefore only find a local minimum, and reasonable initial values for the a_{ij}'s and b_{ij}'s are therefore useful.

A.8 Example

To demonstrate the use of truncated models in MPC, we will apply such a model to the control of a distillation column. The true plant is chosen to be column A, model F2 used by Skogestad and Lundström [7]. The model is given by:

$$\begin{aligned} dy_D &= \frac{87.8}{1+194s}dL + (\frac{1.4}{1+15s} - \frac{87.8}{1+194s})dV \\ dx_B &= \frac{108.2}{1+194s}g_L(s)dL - (\frac{1.4}{1+15s} + \frac{108.2}{1+194s})dV \end{aligned} \tag{A.35}$$

where $g_L(s)$ represents the liquid flow dynamics and is given by

$$g_L(s) = \frac{1}{[1+0.492s]^5} \tag{A.36}$$

The uncertainty considered by Skogestad and Lundström [7] is uncertainty in the manipulated variables, with a diagonal uncertainty weight

$$W_i(s) = 0.2\frac{5s+1}{0.5s+1}I_2 \tag{A.37}$$

The performance weight is

$$W_p(s) = 0.5\frac{10s+1}{10s}I_2 \tag{A.38}$$

The $\mu-$optimal PI controller found by Skogestad and Lundström is

$$\begin{bmatrix} dL \\ dV \end{bmatrix} = \begin{bmatrix} 0.14\frac{2.74s+1}{2.74s} & 0 \\ 0 & 0.62\frac{13.1s+1}{13.1s} \end{bmatrix} \tag{A.39}$$

With this controller $\mu = 0.94$ was obtained, and the performance requirements are therefore fulfilled (a $\mu-$value less than 1 means that the closed loop system will fulfill the performance requirements specified by $W_p(s)$ for all uncertainties allowed by $W_i(s)$).

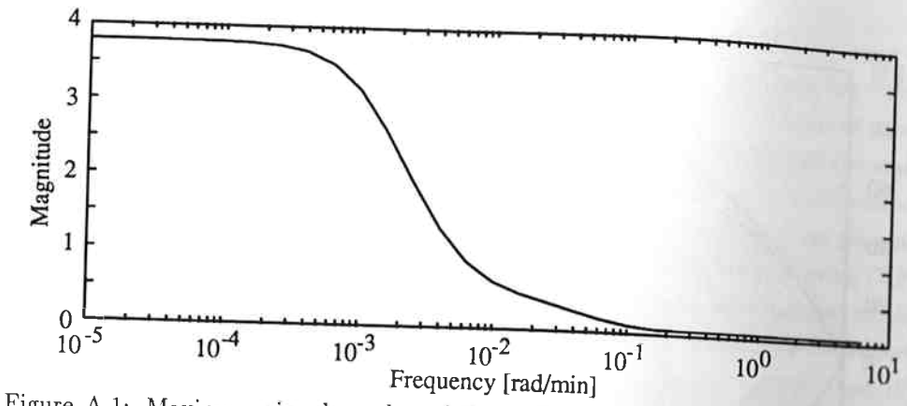


Figure A.1: Maximum singular value of the multiplicative error for the distillation control example.

The magnitude of $W_i(s)$ reaches 1 at a frequency of 1 rad/min. To ensure that sampling does not limit the achievable bandwidth, a sampling interval of 0.5 minutes is chosen. The 1,2 element of the step response has an inflection point at approximately 16 minutes. It was therefore decided to truncate the model at time 17.5 minutes, i.e. to use 35 step response coefficients per plant element. H and S are calculated using the controller in Eq. (A.39).

A constrained optimization algorithm was used to find the A_T and B_T which minimizes (A.30). The values found are:

$$A_T = \text{diag}\begin{pmatrix} 0.9962 & 0.9965 & 0.9967 & 0.9966 \end{pmatrix} \tag{A.40}$$

$$B_T = \begin{bmatrix} 0.2071 & 0 \\ 0.2473 & 0 \\ 0 & -0.1913 \\ 0 & -0.2504 \end{bmatrix} \tag{A.41}$$

The maximum singular value of the corresponding multiplicative error $(P-\hat{P})\hat{P}^{-1}$ is shown in Fig. A.1. The resulting model step responses can be compared to the step responses of the true plant in Fig. A.2

The error at low frequencies is larger than what is commonly considered acceptable. However, stability at low frequencies is not impaired because the directions of the model at low frequencies closely match the directions of the plant, it is only the gains are significantly different. The model reduction measure (A.30) clearly emphasizes accuracy at high frequencies for this case. The mode corresponding to the time constant of 15 minutes has not settled at the truncation time of 17.5 minutes. An accurate model fit at all frequencies using first order dynamics in all elements is therefore not possible.

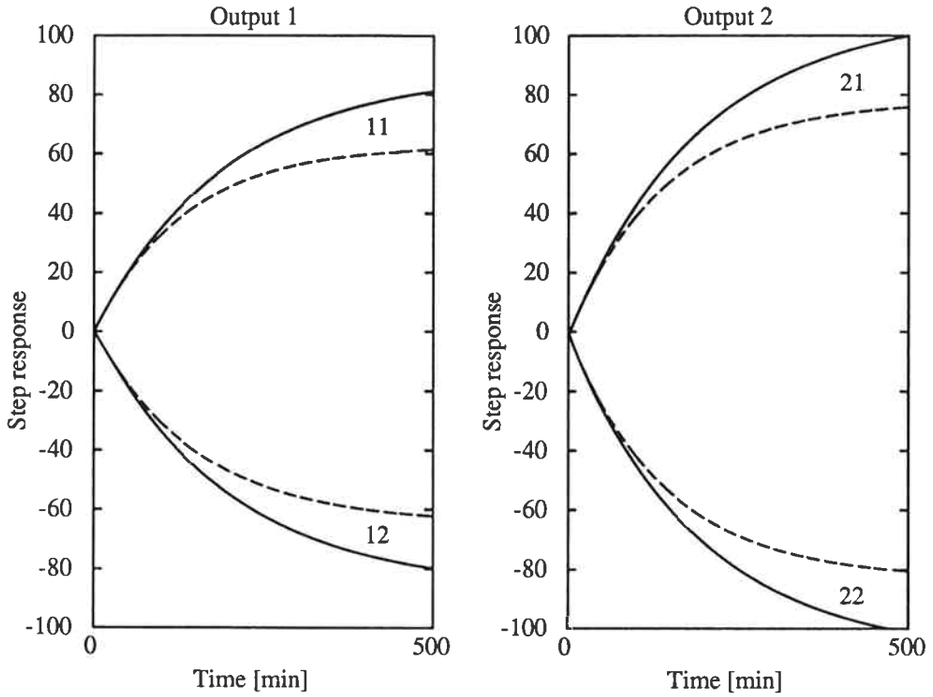


Figure A.2: Step response for the plant (solid lines) and the model (dashed lines).

If we had chosen to truncate at a later time, e.g. 50 minutes, it should be possible to find four first order responses for the residual dynamics which would result in a close model fit at all frequencies.

The best MPC controller tuning parameters and filter parameters that we have found are:

Controller		Filter	
Variable	Weight	Parameter	Value
y_D	17.69	f_{ay_D}	0.1589
x_B	29.02	f_{ax_B}	0.1423
L	0.0003446	f_{by_D}	0.00238
V	0.02715	f_{bx_B}	0.00193

The original example by Skogestad and Lundström [7] did not contain any information about disturbances. We therefore decided to consider disturbances which affect the outputs similarly to the manipulated variables. Thus $\alpha_{y_D} = \alpha_{x_B} = 0.9965$ was chosen. This will make the disturbance dynamics emulate the error arising from the

input uncertainty. The resulting closed loop system was found to have a $\mu-$value of 0.91. This is a slight improvement in the structured singular value from the value found in [7], but may not by itself be sufficient to justify the added complexity of using model predictive control instead of diagonal PI controllers. However, one of the main advantages for using MPC is the ease with which constraints can be accommodated. This advantage does not show up in the $\mu-$analysis. When constraints are present, the controller action must be calculated by solving a Quadratic Programming (QP) problem instead of the least squares solution used in section 5. The solution of the QP problem will, as shown in section 5, not be affected by the model error introduced by truncating the step response model if the prediction horizon is shorter than the truncation time.

An observant reader may note that the values for f_{by_D} and f_{bx_B} do not correspond to the values found from Eq. (A.17). It is by no means surprising that the value for f_b which minimizes the error in the state estimate is not the value which optimizes robust performance, as the optimization criteria are different.

The best MPC controller found by calculating f_{by_D} and f_{bx_B} from Eq. (A.17) had a $\mu-$value of 1.11, meaning that the performance specified by $W_p(s)$ cannot be guaranteed for all cases. Since the example contains no information about the disturbances, one may consider using α_{y_D} and α_{x_B} as variables in the optimization, while still calculating f_{by_D} and f_{bx_B} from Eq. (A.17). For this example it turns out that this only helps a little, as the resulting $\mu-$value obtained is 1.06.

The $\mu-$values found in this work may not be the true optimal values. There are two reasons for this:

- The μ optimization is not convex, and we may therefore only have found local optima.

- Our optimization program is not able to handle integer optimization variables. We therefore had to fix the number of moves and the prediction horizon *a priori*. Based on a few initial calculations these were chosen to be:
  Number of moves: 5
  Prediction horizon: 6
  A more thorough search for the optimal number of moves and prediction horizon were not considered necessary for this example.

MPC without a truncated model would need a sampling interval of about 17 minutes in order to use the same number of coefficients and describe the step response up to three times the main time constant. The best $\mu-$value we have been able to obtain with this

sampling interval is 4.88, which demonstrates the improvement in performance for the same number of coefficients that can be obtained by using truncated models.

A.9 Conclusion

A method for finding truncated models for MPC has been derived and demonstrated. These truncated models result in a dramatically reduced number of step response coefficients needed to model the system, leading to a significant decrease in the computational load on the control system. The model error introduced by truncation need not affect system robustness or performance. An important advantage of MPC, the easy handling of constraints, is unaffected provided the prediction horizon is kept shorter than the truncation time.

References

[1] Lee, J. H., Morari, M. and Garcia, C. E. (1992). State Space Interpretation of Model Predictive Control. Accepted for publication in *Automatica*, preprint.

[2] Åström, K. J. and Wittenmark, B. (1984). *Computer Controlled Systems. Theory and Practice*, Prentice Hall, Englewood Cliffs, N. J., USA.

[3] Doyle, J. C, Wall, J. E. and Stein, G. (1982). Performance and Robustness Analysis for Structured Uncertainty, *Proc. IEEE Conf. Decision Contr.*, Orlando, FL., pp. 629-636.

[4] Garcia, C. E. and Morshedi, A. M. (1984). Quadratic Programming Solution of Dynamic Matrix Control (QDMC), *Proc. Am. Control Conf.*, San Diego, California, USA.

[5] Rivera, D. E. and Morari, M. (1992). Plant and Controller Reduction Problems for Closed-Loop Performance, *IEEE Trans. Autom. Control*, **37**, 3, pp. 398-404.

[6] Skogestad, S. and Morari, M. (1988) Some New Properties of the Structured Singular Value, *IEEE Trans. Autom. Control*, **33**, 12, pp. 1151-1154.

[7] Skogestad, S. and Lundström, P. (1990) Mu-optimal LV-control of Distillation Columns, *Computers chem. Engng.*, **14**, 4/5, pp. 401-413.

Control strategy
- plant
- sensor ↔ actuator connection
- why and how to control!
- control structure