

**STUDIES ON ROBUST CONTROL
OF
DISTILLATION COLUMNS**

Thesis by
Sigurd Skogestad

*In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy*

**California Institute of Technology
Pasadena, California**

1987

(Submitted January 26, 1987)

ACKNOWLEDGEMENTS

I am grateful to my advisor, Manfred Morari, for creating a most stimulating environment during my stay at Caltech. His persistent insistence on taking a rigorous approach to any problem, which not always appealed to the engineer in me ("Why waste time on proving something which seems obvious?"), was an important factor which made this work possible. I will try to bring this attitude with me in my future work, Manfred.

John Doyle arrived at Caltech just when I was starting my research. His theoretical results and insights about feedback theory provided the basis for a major part of this thesis.

When I first started at Caltech in September 1983, I shared office with the late Costa Economou who was most helpful in answering all my questions about feedback control.

On a less serious note, Pierre Grosdidier and I had lots of fun together. Professionally, Pierre's knowledge of linear algebra and decentralized control were very helpful. I also appreciated the friendship and good times I had with my other office and lunch mates: Tony Skjellum, Claudio Scali, Lionel Laroche, Evangelos Zafiriou, Chris Webb and Dan Laughlin.

My stay at Caltech was made possible by financial support from Norsk Hydro and by a *utdanningsstipend* from the Norwegian Institute of Technology (NTH).

STUDIES ON ROBUST CONTROL OF DISTILLATION COLUMNS

by

Sigurd Skogestad

ABSTRACT

Distillation is undoubtedly the most important unit operation in chemical engineering. During design a significant effort is normally put into steady-state optimization of the column with respect to its size, feed location and reflux ratio. However, operating the column close to this optimal point requires reasonably tight control of the product compositions. This is usually not achieved in industrial practice due to stability problems. Improved strategies for distillation control offer a viable means for significant economic savings as compared to the existing *ad hoc* techniques. This thesis addresses robust control of distillation columns in the face of model-plant mismatch caused by model uncertainty, nonlinearity and changes in operating conditions. The robust control paradigm, introduced by Doyle and coworkers, is used as the basis for controller design and analysis. An important tool is the Structured Singular Value (SSV) which enables the evaluation of a plant's achievable control performance. This provides a consistent basis for comparing controllers and design alternatives. Achievable performance is also related to other commonly used measures such as the RGA and the condition number.

Physical insight is used to derive low-order column models which address the issues most important for feedback control. It is shown that the dynamic behavior can be explained in terms of the fundamental difference between external and internal flows. This difference manifests itself both at steady-state and in the dynamic response. Furthermore, the initial response, which is of principal importance for feedback control, is affected much less by changes in the operating conditions than

is the steady-state response. The initial response is even less markedly affected when logarithmic compositions are used.

An important issue in distillation control is which two of the possible five manipulated inputs should be selected for composition control; each configuration may yield entirely different control performance. Issues which must be addressed include model uncertainty and dynamic response as well as rejection of flow disturbances by the level loops.

Finally, a design method for robust decentralized controllers, which generalizes the SSV-interaction measure of Grosdidier and Morari, is introduced. Each loop is designed independently such that robust performance of the overall system is guaranteed.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
CHAPTER I: INTRODUCTION	1
CHAPTER II: DESIGN OF RESILIENT PROCESSING PLANTS: EFFECT OF	
MODEL UNCERTAINTY ON DYNAMIC RESILIENCE	10
Introduction	10
Uncertainty, Stability and Performance	14
Single Perturbations (Unstructured Uncertainty)	19
Multiple Perturbations ("Structured" Uncertainty)	31
The Condition Number as a Sensitivity measure	47
Conclusions	52
References	55
Appendix	59
CHAPTER III: EFFECT OF DISTURBANCE DIRECTIONS ON	
CLOSED LOOP PERFORMANCE	67
Introduction	69
Singular Value Decomposition	71
Effect of Disturbance Direction on Manipulated Variables	73
Effect of Disturbance Direction on Closed Loop Performance	78
Example: LV-Distillation Column	84
References	92
CHAPTER IV: ROBUST CONTROL OF ILL-CONDITIONED PLANTS:	
HIGH-PURITY DISTILLATION	93
Introduction	95
Distillation Column Example	98
Robustness Analysis with μ	112

μ -analysis of the Distillation Column	116
Uncertainty Modelling	129
Conclusions	143
References	144
Appendix	145

CHAPTER V: IMPLICATIONS OF LARGE RGA-ELEMENTS ON

CONTROL PERFORMANCE	147
Introduction	149
Relationship between the RGA and the Condition Number	150
The RGA and Model Uncertainty	152
Choice of Controller Structure	161
Large RGA-Elements are Bad News	163
Examples	164
References	174

CHAPTER VI: SOME NEW PROPERTIES OF

THE STRUCTURED SINGULAR VALUE	176
Introduction	178
How to find N	178
The Structured Singular Value μ	181
New Properties for μ	182
Examples	186
References	189
Appendix	190

CHAPTER VII: ROBUST PERFORMANCE OF DECENTRALIZED

CONTROL SYSTEMS BY INDEPENDENT DESIGNS	192
Introduction	194
Nominal Stability (of H and S)	199

Robust Performance	202
Numerical Example	210
Conclusion	214
References	217
Appendix	218

CHAPTER VIII: A SYSTEMATIC APPROACH TO

DISTILLATION COLUMN CONTROL	220
Introduction	222
Modelling of the Column	226
Choice of Control Configuration	244
Controller Design / Implementation	256
References	258

CHAPTER IX: SHORTCUT MODELS FOR DISTILLATION COLUMNS -

I. STEADY- STATE BEHAVIOR	259
Introduction	262
Case Study Examples	264
A Simple Expression for S	266
Why Are Distillation Columns Nonlinear ?	271
Steady-State Gains from Expression for S	273
A New Formula for the Optimal Feed Location	279
Extensions to Multicomponent Mixtures	281
Conclusions	282
References	284
Appendix	285

CHAPTER X: SHORTCUT MODELS FOR DISTILLATION COLUMNS -

II. DYNAMIC COMPOSITION RESPONSE	287
Introduction	289

Derivation of Expression for τ_c	292
A Simple Formula for τ_c	303
Discussion and Conclusion	308
References	313
Appendix	314

CHAPTER XI: UNDERSTANDING THE DYNAMIC BEHAVIOR OF

DISTILLATION COLUMNS	315
Introduction	317
A Simplified Dynamic Model Based on Internal and External Flows	322
Singular Values, The Condition Number and the RGA	325
Distillation Column Examples	329
High-Frequency Behavior (Initial Response)	335
Validity of First-Order Model at High Frequency	340
Ill-Conditioning at High Frequency	345
Estimation of τ_2	349
The Effect of Reboiler and Condenser Holdup	350
Simplified Model Including Flow Dynamics	353
Discussion/Conclusion	356
References	358
Appendix	360

CHAPTER XII: CONTROL CONFIGURATION SELECTION FOR

DISTILLATION COLUMNS	362
Introduction	364
The Distillation Column from a System Point of View	365
A Simplified Approach	375
Differences between Control Configurations	385
Conclusions	404

References	406
Appendix	407
CHAPTER XIII: LV-CONTROL OF A HIGH-PURITY DISTILLATION	
COLUMN	409
Introduction	411
The Distillation Column	415
Control Theory	420
Formulation of the Control Problem	424
Results for Operating Point A	428
Effect of Nonlinearity (Results for Operating Point C)	435
Conclusions	441
References	442
Appendix	443
CHAPTER XIV: CONCLUSIONS	446

Chapter I

INTRODUCTION

Distillation is undoubtedly the most important unit operation in chemical engineering. During design a significant effort is normally put into steady-state optimization of the column with respect to its size, feed location and reflux ratio. However, operating the column close to this optimal point requires reasonably tight control of the product compositions. This is usually not achieved in industrial practice due to stability problems. Improved strategies for distillation control offer a viable means for significant economic savings as compared to the existing *ad hoc* techniques. A system is "robust" if it is insensitive to model-plant mismatch. This thesis addresses robust control of distillation columns in the face of model-plant mismatch caused by model uncertainty, nonlinearity and changes in operating conditions. The thesis is naturally divided into three parts:

- A. Robust control
- B. Dynamic and steady-state behavior of distillation columns
- C. Robust control of distillation columns

This introduction is organized accordingly.

A. Robust Control. Feedback control is used to control processes despite unmeasured disturbances and model-plant mismatch without which feedforward control would suffice. Another important factor which limits the achievable performance of a system is the presence of RHP-zeros (inverse responses, time delays) in the plant. A good theory for feedback control should address these three issues in a direct manner.

Control theory is a new science. The early work in the 1930's and 40's provided a theory for single-loop (SISO) plants based on an input-output description using the frequency domain. This theory addressed all three above-mentioned issues to some degree. However, the extension of these results to multivariable (MIMO) systems did not prove to be straightforward. During the 60's the "optimal" control theory based on a state-space system description in the time domain was developed.

MIMO systems are handled readily within this framework. The design techniques suggested, for example, the Linear Quadratic Gaussian (LQG) controller, treated disturbances as stochastic processes. These methods represented significant theoretical breakthroughs, but in hindsight probably more so in optimization than in feedback theory. The names "optimal" and "modern" which were associated with this theory proved to be deceiving, since they engendered the belief that all problems in linear control theory had been solved. However, the slow acceptance of optimal control theory in industrial practice eventually led to the realization that the theory had serious shortcomings (Horowitz,1975). For one thing, "classical" control theory, which had proven to work in practice for SISO systems, had no clear link to the optimal control theory. Secondly, it became obvious during the 70's that optimal control did not address the issue of model uncertainty at all. Furthermore, important concepts such as RHP-zeros and bandwidth were obscured by the state-space formalism. This finally led in the late 70's and early 80's to the development of a robust linear control theory (e.g., Doyle, 1984) which directly addressed the problem of model uncertainty and also provided a bridge between "classical" frequency domain theory and the "modern" state space theory. The formulation of the control problem is accomplished in the frequency domain. This includes the definition of performance specifications (H_∞) and the quantification of model uncertainty. However, all computations are carried out within the state-space formalism, and the LQG-theory (Riccati equations, etc.) remains useful for solving the numerical problems which arise.

Robust Performance (RP) is satisfied if the performance specifications are also met for the "worst case" plant as defined by the uncertainty bounds. The Structured Singular Value (μ) introduced by Doyle (1982) provides a non-conservative means for testing whether Robust Performance (RP) is satisfied:

$$RP \Leftrightarrow \mu(N) < 1$$

where the matrix N depends on the plant model, controller, performance specifications and uncertainty bounds. In the thesis, this theory is detailed and specific applications to chemical engineering problems (including distillation control) are presented. New results which link quantities such as the Relative Gain Array (RGA) and the condition number to RP are also provided.

Linear control theory is used throughout this thesis; only recently has a viable theory which account for model uncertainty been introduced, and no such theory as yet exists for nonlinear systems. A good understanding of linear systems is clearly a prerequisite for attacking the vastly more complicated behavior of nonlinear systems.

B. Dynamic and steady-state behavior of distillation columns. A precondition for applying any control theory is that a model of the physical system is available. In particular, this model should include the characteristics of the plant which are most important for feedback control. For distillation columns, this includes the presence of RHP-zeros caused by the flow dynamics and the plant's ill-conditioned character. Furthermore, to use the μ -theory, the sources of model uncertainty must be identified and bounds on these must be quantified. For this step it is very important to have a clear physically motivated picture of the process, so that the "uncertainty" (which sometimes includes model-plant mismatch caused by nonlinearity or changes in operating conditions) can be treated in a non-conservative manner. Therefore, a main objective of this thesis is to develop simple dynamic models for distillation columns.

The mathematical treatment of distillation columns is also a new "science," and its development has followed an analogous pattern to that of control theory. The early work of the 1920's and 30's treated the steady-state behavior using graphical methods (McCabe and Thiele) and simple short-cut models (Gilliland). Restrictive assumptions made these methods inaccurate for certain design calculations,

but they provided invaluable insight into the steady-state behavior of distillation columns. The methods are difficult to apply for multicomponent and nonideal mixtures, and no simple methods for predicting the dynamic behavior were developed. With the inception of the digital computer in the late 50's and early 60's, the need for simple methods was no longer considered important; the governing nonlinear differential equations for distillations columns are easily formulated, and virtually any column could therefore be simulated. Consequently, most of the research effort was directed away from simple analytical methods. Still, some progress was made in the 60's toward the *understanding* of the dynamic behavior of distillation columns (Moczek et al., 1963, Wahl and Harriot, 1970). It was established that the dynamic response is dominated by one large time constant which is roughly the same regardless of where a disturbance or input is introduced or where composition is measured. Yet, the generally held belief remains that the dynamic behavior of distillation columns is difficult to predict. This is exemplified by the following quote from Shinskey (1984) (the industry standard book on distillation control): "The only general relationship that seems to apply to the dominant time constant is that it is proportional to the total liquid volume divided by the column feed rate." This is an immediate observation because this essentially determines the scale factor for time in the problem. One important objective of this research has been to provide relationships which enable understanding and prediction of the dynamic behavior of distillation columns without the need for reliance on simulations, and which may also be used as an integral part of a control study.

C. Robust Control of Distillation Columns. Most two-product distillation column can be described as a 5×5 plant, but the control system design is usually simplified by means of the following procedure:

1. Choose two manipulated inputs for composition control (corresponding to a specific control "configuration").

2. Design a level and pressure control system (usually three SISO controllers).
3. Design a 2×2 controller for composition control.

Step 1 is probably the most important one, yet almost no guidelines are available in the literature. The only exception is the work of Shinskey (1984) who bases his analysis mainly on steady-state RGA-values. From Shinskey's book the reader is led to believe that the RGA is just an interaction measure which indicates the difficulties one can expect when tuning single loops of multivariable systems. But if this were true then the RGA recommendations regarding control configuration selection would be archaic today, now that the implementation of multivariable control systems is quite straightforward. Practical evidence suggests, however, that an RGA evaluation should be very useful for the design of multivariable control systems. Thus, while the RGA has proven to be very useful tool for *categorizing* experience, it has not helped to *explain* the observed phenomena. A goal of this research is to fill the apparent gap between the theory and the industrial experience as represented by Shinskey (1984).

Another issue which has been obscure is the use of ratios between flows, e.g. $\frac{L}{D}$ or $\frac{V}{B}$, as manipulated inputs for composition control. It is well-known (Shinskey, 1984) that the RGA-values are affected, but apart from this, what is their real effect? Do they result in a plant which is more linear? The results in the thesis show that this is not the case, but that a main feature is that they provide improved flow disturbance rejection.

Once the choice of control configuration is made, the design of the control system is not too difficult. First, the level control system has to be designed and guidelines for this step are already available in the literature (e.g., Shinskey, 1984). Secondly, the 2×2 controller for composition control must be designed. The academic literature (e.g., Weichedel and Mcavoy, 1980) have perennially discussed whether or not to use "decouplers" as part of the controller. However, to address this problem

rigorously, model uncertainty has to be considered. In this thesis, new results are presented (partly in terms of the RGA) which clearly demonstrate those cases for which decouplers may be used to good effect and cases for which robust performance will be poor.

A final question which merits an answer is the following: Distillation columns are strongly nonlinear, but can they be adequately controlled using a linear controller? Industrial experience indicates that the answer is "yes." This is indeed confirmed by the results of this thesis. In particular, it is shown that the initial response (which is of main importance for feedback control) does not depend strongly on operating conditions. This claim is supported both by a general analysis of the dynamic behavior of distillation columns and by a particular study of the robust control of a high-purity column. It is found that the initial response is even less markedly affected by nonlinearity when logarithmic compositions are used.

Thesis Overview.

This introduction motivates the need for a systematic approach to distillation column control which yields a control system with robust performance, that is, a system which performs satisfactory also in the presence of model-plant mismatch. Important steps include modelling, selection of the control configuration as well as robust controller design. This approach is presented in this thesis as follows:

Part A. Robust Control: Chapter II discusses the issue of robust stability and how model uncertainty limits the achievable performance. The notion of disturbance directions, which is important for multivariable systems (and in particular for ill-conditioned plants), is treated in Chapter III. Subsequently, Chapter IV discusses general problems of controlling ill-conditioned plants, indicates the applicability of the μ -theory for the analysis and controller design for such plants in face of model uncertainty, and demonstrates the usefulness of this approach via a simplified distillation column example. The RGA is used extensively in the literature to screen

design alternatives with respect to their expected control performance: Some new relationships between model uncertainty and large RGA-elements are presented in Chapter V which justify this measure's widespread use. Chapter VI provides some new properties of μ ; the application of these results to the design of decentralized controllers yielding robust performance is demonstrated in Chapter VII.

Part B. Dynamic and steady-state behavior of distillation columns: This part begins with Chapter VIII which provides an overview of the remainder of the thesis and also shows how to evaluate the steady-state gains for various configurations. Chapter IX discusses the use of the separation factor for the estimation of steady-state gains. Chapter X and XI consider the dynamic column behavior and present important new insights. It is shown that the dynamic behavior can be explained in terms of the fundamental difference between external and internal flows. This difference manifests itself both at steady-state and in the dynamic response. Furthermore, the initial response, which is of principal importance for feedback control, is affected much less by changes in the operating conditions than is the steady-state response. The initial response is even less markedly affected when logarithmic compositions are used.

Part C. Robust control of distillation columns: Chapter XII addresses the issues of control configuration selection. It is shown that ratio configurations yield complex multivariable controllers which in some cases provide improved flow disturbance rejection. However, their linearizing effect on the plant does not prove significant. Finally, in Chapter XIII a more realistic study of the distillation example of Chapter IV is presented. Chapter XIII shows that the robust control of a high-purity distillation column over a wide range of operating conditions is in indeed possible with a single linear controller. Logarithmic compositions were used in this example to reduce the effect of nonlinearity. The simulation results as well as the μ -analysis presented in this chapter confirm the general results regarding

the dynamic behavior of distillation columns of Chapter XI, and also support the control theory developed previously in Part A.

References.

Doyle, J. C., Analysis of Feedback Systems with Structured Uncertainties, IEE Proc., Pt. D, **129**, 242-247 (1982)

Doyle, J. C., Lecture Notes, ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis, MN, Oct. 8-10 (1984)

Horowitz, I. and U. Shaked, Superiority of Transfer Function over State-Variable Methods in Linear Time-Invariant Feedback System Design, IEEE Trans. Automat. Contr., **AC-20**, 84-97 (1975).

Moczek, J. S., R. E. Otto and T. J. Williams, Approximation Model for the Dynamic Response of Large Distillation Units, Proc. 2nd IFAC Congress, Basel (1963).

Shinskey, F. G., Distillation Control, 2nd Ed., McGraw-Hill, New York (1984).

Wahl, E.F. and P. Harriot, Understanding and Prediction of the Dynamic Behavior of Distillation Columns, Ind. & Eng. Chemistry Proc. Des. & Dev., **9**, 396-407 (1970).

Weischedel, K. and T. J. McAvoy, Feasibility of Decoupling in Conventionally Controlled Distillation Columns, Ind. & Eng. Chemistry Fundam., **19**, 379-384 (1980).

Chapter II

DESIGN OF RESILIENT PROCESSING PLANTS: EFFECT OF MODEL UNCERTAINTY ON DYNAMIC RESILIENCE

DESIGN OF RESILIENT PROCESSING PLANTS
EFFECT OF MODEL UNCERTAINTY ON DYNAMIC RESILIENCE

Sigurd Skogestad
Manfred Morari

Chemical Engineering, 206-41
California Institute of Technology
Pasadena, California 91125
(818)356-4186

Chem. Eng. Sci. Ms #CES 1724
Submitted July, 1985
Revised September 1986
Updated January 1987

Abstract

The achievable quality of control for a particular system (its dynamic resilience) is limited by the nonminimum phase characteristics of the plant, constraints on the manipulated variables and model uncertainty. Model uncertainty requires that the controller be detuned and performance be sacrificed. The goal of this paper is to quantify this well-known qualitative statement.

The closed-loop system must remain stable for all possible plants as defined by the uncertainty description. This robust stability requirement is used to derive simple bounds on the nominal performance for some specific cases. These bounds are relatively easy to evaluate and should be effective tools for screening alternative designs in terms of their resilience characteristics. The RGA and the minimized condition number are accurate measures with respect to element uncertainty, provided the relative errors of the transfer matrix elements are independent (uncorrelated) and have similar magnitude bounds.

I. INTRODUCTION

Most chemical plants are designed on the basis of steady state considerations, and the control system is designed separately in a subsequent stage of the project. This separation is acceptable provided that there exist suitable design-stage methods which can assess the "controllability" of the plant. That is, it must be determined a priori whether the design of a control system offering "reasonable" closed-loop response will subsequently be feasible. Until recently, such methods were not available. As a result, the expected performance often was not achieved in the operating plant. In some instances, a minor change at the initial design stage could have resulted in a "controllable" plant.

Previously, the controllability assessment has been based on simulations. This approach is complex and requires a complete dynamic model of the plant. Usually a number of case studies are performed with different choices of inputs, disturbances, operating conditions, controller structures and controller parameters. All those choices could bias the controllability assessment in an erroneous manner.

Morari (1983) suggested making the problem of controllability assessment independent of the controller selection problem. This is done by finding a plant's best achievable closed-loop control performance for all possible constant parameter linear controllers. This target, the upper bound on the achievable closed loop performance, is defined as the plant's dynamic resilience. Thus, "dynamic resilience" is an expression of the plant's inherent limitation on the closed-loop system's dynamic response which is not biased by specific choices of controllers.

The limitations imposed by non-minimum phase elements and constraints have been discussed in quantitative detail by Morari (1983) and Holt and Morari (1985). Fundamentally, perfect control can only be achieved if the plant is invertible. Non-minimum phase elements (Right Half Plane (RHP) zeros and time delays) make it impossible to invert the plant and retain (internal) stability of the closed-loop system. The effect of constraints on performance is also related to a plant's closeness to singularity. If the minimum singular value of a plant \tilde{P} , ($\sigma(\tilde{P})$) is small then the plant is nearly singular. This means that the plant has a very small gain for a particular input direction. To achieve tight control, the controller would have to provide very large input signals in this direction, possibly violating input size constraints.

The objective of this paper is to study the effects of model uncertainty on dynamic resilience. Model uncertainty requires that the controller be detuned and performance be sacrificed. The primary goal is to quantify this well-known qualitative statement by deriving expressions relating achievable closed loop performance and uncertainty.

The first (and most important) step is to quantify the model uncertainty. This is usually not a trivial problem, and very misleading results may arise if an inappropriate uncertainty description is used. Another goal of this paper is to demonstrate some of these pitfalls. Therefore, the design engineer encounters a difficult situation: simple achievable performance bounds may be obtained with a crude uncertainty description but such bounds are often misleading. On the other hand, a detailed description of the

model uncertainty is needed to find more meaningful bounds. Such descriptions are normally not available. A first step in resolving this dilemma is to identify for specific problem classes (e.g. distillation columns) the sources of model uncertainty which are likely to cause complications. The engineer can then concentrate on these when quantifying the uncertainty. Some of the examples in this paper will be helpful in this respect.

II. UNCERTAINTY, STABILITY AND PERFORMANCE

1. Model Uncertainty: Causes and Definition.

The linear time invariant models used throughout this paper describe the actual plant dynamics only approximately.

1. All real processes are nonlinear. In this paper, linear transfer functions are used to represent the plant and some "uncertainty" is introduced by linearizing the nonlinear plant at various operating points. This may lead to a linear model with "uncertain" coefficients.
2. In other cases the process may be represented quite accurately by linear models. However, different operating conditions can lead to changes of the parameters in the linear model. For example, increased throughput/flowrates usually result in smaller deadtimes and time constants.
3. Consequently, in many cases parts of the "uncertainty" are known accurately. However, there will always exist "true" uncertainties even though the underlying process is essentially linear: The model parameters are never known exactly and, at high frequencies, even the model order is unknown.

Definition of Model Uncertainty: We assume the plant P is linear and time invariant, but that its exact mathematical description is unknown. However, it is known to be in a specified "neighborhood" of the "nominal" system, whose mathematical "model" \tilde{P} is available. This neighborhood will be denoted the "uncertainty set"; it defines the "set of possible plants" Π . In some cases the uncertainty set Π may include a finite number of plants. However, in most cases we will define Π in terms of norm-bounded perturbations on \tilde{P} , and the set Π becomes infinite.

2. The Effect of Model Uncertainty

Before discussing how uncertainty limits the achievable performance (dynamic resilience), a digression on why feedback is used for control is of interest. Obviously, for stable plants in the absence of uncertainty, feedforward control would be sufficient. Feedback is used to control a plant despite unmeasured disturbances and model uncertainty. One particular example is the application of integral action in order to achieve perfect steady state control. Without exact knowledge of the steady state gain, perfect control may be achieved through feedback.

However, even though high gain feedback can be used to reduce the effect of uncertainty, it is intuitively obvious that there must be a limit to the extent that uncertainty can be tolerated before the system must be detuned and performance sacrificed. Thus uncertainty may impose limitations on the achievable performance (dynamic resilience). Here, quantitative effects of uncertainty on closed-loop performance will be found. First, additional terminology is required:

Performance: "Performance" is the quality of the closed-loop response. Typically, the error signal (e) should be small for the expected disturbances (d) and reference signals (r) (Fig. 1). The sensitivity function (S) describes the relationship between r, d and e.

$$e = S(r-d), S = (I+PC)^{-1} \quad (1)$$

In order to have "good" performance, S has to be "small." In this paper, the magnitude of S is measured using the singular value $\bar{\sigma}$. At a given frequency ω , $\bar{\sigma}(S(j\omega))$ represents the "worst" amplification ($\|e\|_2 / \|r-d\|_2$) of (r-d). By "worst" we mean that r-d is in the direction giving rise to the largest amplification. A typical performance specification is

$$\bar{\sigma}(S) \leq 1/|w_p| \quad \forall \omega \quad (2)$$

where $w_p(S)$ is a weight which is used to define what responses are acceptable. The complementary sensitivity function H will also be used to measure performance. H is defined by

$$\begin{aligned} H &= I - S \\ \text{or} \quad H &= PC(I + PC)^{-1} \end{aligned} \quad (3)$$

H relates the output y to the reference signal r

$$y = Hr \quad (4)$$

It is desirable to have $H \approx I$. $\bar{\sigma}(H) < 1$ at some frequency implies that tight control ($H \approx I$) is not possible. \tilde{S} and \tilde{H} are used to denote the nominal ($P=\tilde{P}$) sensitivity and complementary sensitivity functions.

Nominal stability (N.S.): The nominal closed loop system (with no uncertainty) is stable.

Nominal Performance (N.P.). The nominal closed loop system \tilde{S} (with no uncertainty) satisfies the performance specification (2).

Robust stability (R.S.): The closed loop system is stable for all plants in the "uncertainty set."

Robust performance (R.P.): The closed loop system satisfies the performance specification (2) for all plants in the "uncertainty set."

In the context of 'how' uncertainty affects performance, there are at least three problems of interest:

Problem 1: The effect of the robust stability requirement on nominal performance: How does the stability requirement for all plants in the uncertainty set limit the nominal performance?

Problem 2: The effect of the robust performance specification on nominal performance: If we specify that some particular performance requirement has to be satisfied for all plants in the uncertainty set, how does this bound the nominal performance?

Problem 3: Achievable robust performance: Design the best possible controller; what is the best achievable performance by all plants in the uncertainty set?

In Problem 1 and 2, a "lower bound" on robust performance is specified (for Problem 1 this "lower bound" is simply the requirement of stability), and we are considering effect on the nominal performance. The goal is to derive some simple bounds on the nominal system which, when satisfied, give the desired robust performance. These bounds are intended to assist the engineer in designing a controller for the nominal system such that the specified performance for all plants in the uncertainty set is achieved.

In Problem 3, there is no in particular concern for the performance of the nominal system. In this case, the problem is to find the "upper bound" on robust performance using any linear

controller. This situation is addressed by Doyle (1984) and actually involves finding the optimal controller; this " μ -synthesis" is a complicated mathematical and numerical problem which will not be addressed here.

This paper will be concerned mainly with Problem 1. Problem 2 may be formulated trivially as a special case of Problem 1 provided that the appropriate norm (H_∞) is used to define performance (Doyle, (1982b, 1984)). Problem 1 is important in itself for the case when the plant is "operating" most of the time close to its nominal point, but with occasional plant perturbations. In this case performance may not be important when perturbations occur provided that the system remains stable. Furthermore, for Problem 1 it will be possible to derive reasonably simple bounds on the achievable nominal performance. Simplicity is desired in order for the engineer to gain insight into 'why' a particular design is sensitive to uncertainty.

Two approaches may be taken in order to find bounds on nominal performance imposed by robust stability (Problem 1):

1. A performance related transfer function which is to be bounded (for example, $\bar{\sigma}(\tilde{H})$ or $\bar{\sigma}(\tilde{S})$) is selected. This requires that the uncertainty be expressed in terms of a specific single perturbation ("unstructured" uncertainty) as discussed in Section III. The bounds derived using unstructured uncertainty are generally conservative since the actual uncertainty rarely "fits" into a single norm-bounded perturbation.
2. A reasonably "tight" description of the uncertainty is chosen. This is done by identifying more precisely where the uncertainty occurs in the system, such as by considering uncertainty in the

model parameters. This generally leads to an uncertainty description with multiple perturbations (Δ_i 's). By assuming norm bounds on these (e.g. $\bar{\sigma}(\Delta_i) \leq 1$) uncertainties, it is possible to derive non-conservative conditions for robust stability using the Structured Singular Value, μ . This approach is due to Doyle and coworkers (1982, 1984) and is considered in Section IV. One disadvantage of this procedure is that the resulting conditions are not in terms of a simple bound on $\bar{\sigma}(\tilde{H})$ or $\bar{\sigma}(\tilde{S})$, but involve $\mu(N)$ where N may be a complicated function of \tilde{S} and \tilde{H} .

A number of conditions in this paper are stated as both necessary and sufficient for robust stability, but it is stressed that necessity is only meaningful if the assumed uncertainty is an accurate ("tight") description of the true uncertainty.

III. SINGLE PERTURBATIONS (UNSTRUCTURED UNCERTAINTY)

In this section, the uncertainty which occurs at different parts of the system will be lumped into one single perturbation L . In most cases this will correspond to "unstructured" uncertainty. (More precisely, "unstructured" uncertainty means that several sources of uncertainty are described with a single perturbation which is a "full" matrix of the same size as the plant P).

Let $P \in \Pi$ be any member of the set of possible plants Π , and let $\tilde{P} \in \Pi$ denote the nominal model of the plant. To describe unstructured uncertainty the following four single perturbations are commonly used: additive (L_A), multiplicative input (L_I), multiplicative output (L_O) and inverse multiplicative output (L_S) perturbations (Fig. 2)

$$P = \tilde{P} + L_A \quad \text{or} \quad L_A = P - \tilde{P} \quad (5a)$$

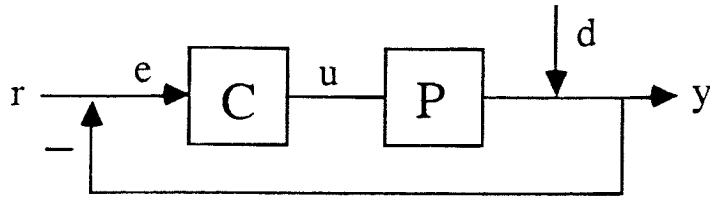


Figure 1. Feedback system with controller C and plant P .

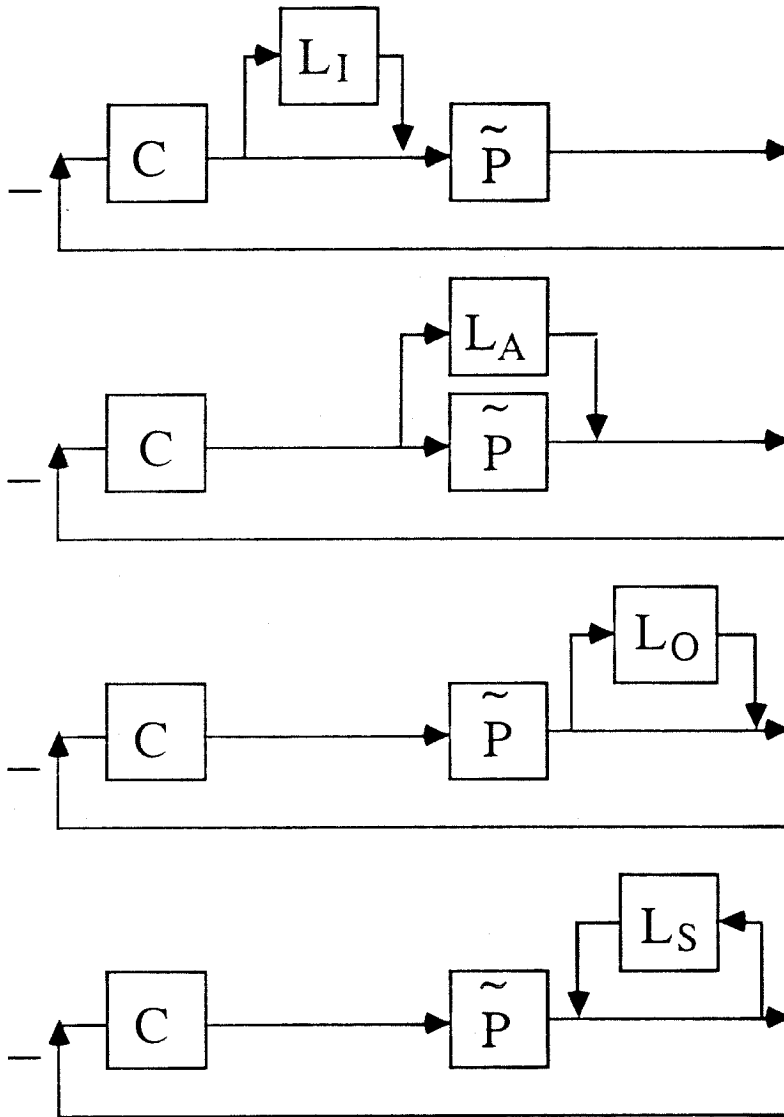


Figure 2. Four common uncertainty descriptions involving single perturbations:

Input multiplicative uncertainty (L_I); Additive uncertainty (L_A);
Output multiplicative uncertainty (L_O); Output inverse
multiplicative uncertainty (L_S).

$$P = \tilde{P}(I+L_I) \quad \text{or} \quad L_I = \tilde{P}^{-1}(P-\tilde{P}) \quad (5b)$$

$$P = (I+L_O)\tilde{P} \quad \text{or} \quad L_O = (P-\tilde{P})\tilde{P}^{-1} \quad (5c)$$

$$P = (I+L_S)^{-1}\tilde{P} \quad \text{or} \quad L_S = (P-\tilde{P})P^{-1} \quad (5d)$$

Additional examples are given by Doyle et al. (1982). The conditions derived for robust stability will be different depending on which single perturbation is chosen to describe the uncertainty.

1. Simple bounds on $\bar{\sigma}(\tilde{H})$, $\bar{\sigma}(\tilde{H}_I)$ and $\bar{\sigma}(\tilde{S})$

In each of the cases above the magnitude of the perturbation L may be measured in terms of a bound on $\bar{\sigma}(L)$

$$\bar{\sigma}(L) \leq \ell(\omega) \quad \forall \omega \quad (6)$$

where

$$\ell(\omega) = \max_{P \in \Pi} \bar{\sigma}(L)$$

The bound $\ell(\omega)$ can also be interpreted as a scalar weight on a normalized perturbation $\Delta(s)$

$$L(s) = \ell(s)\Delta(s), \quad \bar{\sigma}(\Delta) \leq 1 \quad \forall \omega \quad (7)$$

The magnitude bound $\ell(\omega)$ will not generally constitute a tight description of the "real" uncertainty. This means that the set of plants satisfying (7) will be larger than the original set Π .

Output Multiplicative Uncertainty. The sensitivity function S has to be stable for all $P \in \Pi$. Using the identity

$$S = \tilde{S}(I+(P-\tilde{P})\tilde{P}^{-1}\tilde{H})^{-1} = \tilde{S}(I+L_O\tilde{H})^{-1} \quad (8)$$

and the Nyquist stability condition, the following robust stability bound is derived.

Theorem 1. Bound in terms of $\bar{\sigma}(\tilde{H})$ (Doyle and Stein, 1981, Postlethwaite and Foo, 1985)

Assume the nominal system is closed loop stable, that is, assume in particular that \tilde{H} is stable. Let Π be any set of plants such that

P and \tilde{P} have the same number of RHP (unstable) poles. Then robust stability is guaranteed

$$\text{if } \bar{\sigma}(\tilde{H}) \leq 1/\ell_0(\omega) \text{ where } \ell_0(\omega) = \max_{P \in \Pi} \bar{\sigma}(L_0) \quad (9)$$

(Condition (9) is necessary and sufficient for robust stability if it is assumed that all plants satisfying $\bar{\sigma}(L_0) \leq \ell_0(\omega)$ actually occur (Doyle and Stein, 1981)).

The robust stability condition (9) can always be satisfied for open loop stable systems since $\tilde{H} = 0$ (no feedback) is always possible. However, good disturbance rejection and good command following require $\tilde{H} \approx I$ (i.e., $\bar{\sigma}(\tilde{H}) \approx 1$). Condition (9) says that the system has to be "detuned" ($\bar{\sigma}(\tilde{H}) < 1$) at frequencies where $\ell_0(\omega) \geq 1$. This is reasonable since $\ell_0(\omega) > 1$ for some ω implies that the plant can have zeros on both sides of the imaginary axis; it is well known that RHP-zeros limit the achievable performance.

Input Multiplicative Uncertainty. In this case a theorem similar to Theorem 1 is obtained, but with \tilde{H} replaced \tilde{H}_I (Postlethwaite, 1985):

$$\text{R.S. if } \bar{\sigma}(\tilde{H}_I) \leq 1/\ell_I(\omega), \quad \ell_I(\omega) = \max_{P \in \Pi} \bar{\sigma}(L_I) \quad (10)$$

where

$$\tilde{H}_I = C(I + \tilde{P}C)^{-1}\tilde{P} = \tilde{P}^{-1}\tilde{H}P \quad (11)$$

\tilde{H}_I is the nominal closed loop transfer function as seen from the input of the plant. It is desirable to have this transfer function close to I in order to reject disturbances affecting the inputs to the plant. However, since performance is usually measured at the output of the plant it may be of interest to use (10) in order to derive a bound in terms of \tilde{H} . To derive this bound the inequality

$$\bar{\sigma}(\tilde{H}_I) = \bar{\sigma}(\tilde{P}^{-1}\tilde{H}P) \leq \bar{\sigma}(\tilde{P}^{-1})\bar{\sigma}(\tilde{H})\sigma(\tilde{P}) = \gamma(\tilde{P})\bar{\sigma}(\tilde{H})$$

is used; the bound is:

$$\text{R.S. if } \bar{\sigma}(\tilde{H}) \leq \frac{1}{\gamma(\tilde{P})} \frac{1}{\ell_I(\omega)} \quad \forall \omega \quad (12)$$

Here $\gamma(\tilde{P}) = \bar{\sigma}(\tilde{P})/\underline{\sigma}(\tilde{P})$ is the condition number of the plant. (12) has been used to introduce the condition number as a stability sensitivity measure with respect to input uncertainty (Morari, 1983), but this is misleading. The condition number enters the stability condition (12) mainly as the result of the conservative step introduced by going from an input (Eq. (10)) to an output uncertainty description (Eq. (12)). For $\gamma(\tilde{P})$ large, (12) may be arbitrarily conservative even though the uncertainty is tightly described in terms of a norm-bounded input uncertainty such that (10) is both necessary and sufficient. However, even though (12) is misleading and the system is stable, input uncertainty usually does cause control problems when $\gamma(\tilde{P})$ is large. As shown by Morari and Doyle (1986), robust performance (measured at the output of the plant) is usually poor in such cases (even though the nominal performance may be excellent). Output uncertainty does not lead to the same performance problems, and this indicates why input uncertainty is of more concern than output uncertainty for ill-conditioned plants.

Inverse multiplicative output uncertainty. Using the identity

$$S = \tilde{S}(I+L_S\tilde{S})^{-1}\tilde{P}P^{-1} = (I+\tilde{S}L_S)^{-1}\tilde{S}\tilde{P}P^{-1} \quad (13)$$

and the inverse Nyquist stability condition, the following theorem may be derived.

Theorem 2. Bound in terms of \tilde{S} (Postlethwaite, 1985)

Assume the nominal system is closed loop stable, that is, assume in particular that \tilde{S} is stable. Let Π be any set of plants such that P and \tilde{P} have the same number of RHP zeros. Then robust stability is

guaranteed if

$$\bar{\sigma}(\tilde{S}) \leq \frac{1}{\ell_S(\omega)} \quad \text{where} \quad \ell_S(\omega) = \max_{P \in \Pi} \bar{\sigma}(L_S) \quad (14)$$

For minimum phase systems (no time delays or RHP zeros), the nominal sensitivity function \tilde{S} may be arbitrary small ("perfect control") and (14) can always be satisfied. Therefore, Condition (14) seems to imply that for minimum phase systems arbitrarily good performance (\tilde{S} small) is possible regardless of how large the uncertainty is. This is claimed by Postlethwaite, but is not quite true. The pitfall is that any real system has to be strictly proper, and $S = I$ and $\tilde{S} = I$ as $\omega \rightarrow \infty$ must be required. Consequently, to satisfy (14) it is necessary that $\bar{\sigma}(L_S) = \bar{\sigma}(P - \tilde{P})P^{-1} \leq 1$ as $\omega \rightarrow \infty$ for all possible P . This condition is usually violated in practice, because the order of the actual plant is higher than that of the model.

Theorems 1 and 2 prescribe two fundamentally different ways of handling uncertainty: To guarantee robust stability Theorem 1 prescribes that the system be detuned (low gain), while Theorem 2 prescribes that the control be tightened (high gain). In practice, it is desirable to combine the two approaches: By tightening the control at low frequencies better performance is obtained. Eventually, at higher frequencies, the system has to be detuned to guarantee robust stability. In fact, Postlethwaite (1985) has shown that it is possible to combine Theorem 1 and 2 over different frequency ranges. However, the bounds are still conservative since there is no "tight" description of the uncertainty. A better approach is to derive tighter uncertainty descriptions in the first place and then derive robust

stability bounds (Section IV).

2. Input Uncertainty for Distillation Column (Example)

Conditions (9) and (10) indicate that the system has to be detuned such that $\bar{\sigma}(\tilde{H}) < \frac{1}{\ell_0(\omega)}$ (or $\bar{\sigma}(\tilde{H}_I) < \frac{1}{\ell_I(\omega)}$) in order to guarantee robust stability. However, because of the conservativeness introduced by using unstructured uncertainty, these conditions are generally only sufficient for robust stability; the detuning indicated may be much larger than what is actually necessary. This is illustrated conveniently through an example.

Consider the distillation column described in Table 1 where the overhead composition is to be controlled at $y_D = 0.99$ and the bottom composition at $x_B = 0.01$ using the distillate D and boilup V as manipulated inputs. By linearizing the nonlinear model at steady state and by assuming that the dynamics may be approximated by a first order response with time constant $\tau = 75$ min, the following linear model is derived (Skogestad, 1986):

$$\tilde{P} = \frac{1}{\tau s + 1} \begin{bmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{bmatrix}$$

A simple decentralized control system with two PI controllers is chosen

$$C(s) = \frac{1 + \tau s}{s} \begin{bmatrix} -0.15 & 0 \\ 0 & -7.5 \end{bmatrix} \quad (15)$$

This controller gives acceptable nominal performance, and can be shown (Section IV) to give Robust Stability when there is relative uncertainty of magnitude $w_I(s)$ on each manipulated variable:

$$w_I(s) = 0.2 \frac{5s+1}{0.5s+1}$$

This implies a relative uncertainty of up to 20% in the low frequency

Binary separation, constant molar flows, feed liquid

Relative volatility $\alpha = 1.5$

No. of theoretical trays $N = 50$

Feed tray location $N_F = 21$

Feed rate and composition $F = 1 \text{ kmol/min}, z_F = 0.5$

Product compositions $y_D = 0.99, x_B = 0.01$

Product rates $D = B = 0.5 \text{ kmol/min}$

Computed from steady state model

Reflux rate $L = 2.71 \text{ kmol/min} (1.39 L_{\min})$

Linearized steady state gains.

LV-configuration:

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \begin{bmatrix} dL \\ dV \end{bmatrix}$$

DV-configuration:

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{bmatrix} \begin{bmatrix} dD \\ dV \end{bmatrix}$$

Table 1. Data for distillation column example.

range which increases at high frequencies, reaching a value of 1 at $\omega \approx 1 \text{ min}^{-1}$. This increase with frequency allows for a time delay of about 1 minute, and may represent the effect of the flow dynamics which were neglected when developing the model. This relative uncertainty can be written in terms of two scalar multiplicative perturbations Δ_D and Δ_V .

$$\begin{aligned} dD &= (1+w_I(s)\Delta_D)dD_C, \quad |\Delta_D| \leq 1 \quad \forall \omega \\ dV &= (1+w_I(s)\Delta_V)dV_C, \quad |\Delta_V| \leq 1 \quad \forall \omega \end{aligned} \quad (16)$$

(dD and dV are the actual inputs, while dD_C and dV_C are the desired values of the flow rates as computed by the controller). (16) can be approximated by an "unstructured" single perturbation $L_I = W_I \Delta_I$ (Δ_I is a "full" 2x2 matrix)

$$\begin{bmatrix} dD \\ dV \end{bmatrix} = (I+W_I(s)\Delta_I) \begin{bmatrix} dD_C \\ dV_C \end{bmatrix}, \quad \bar{\sigma}(\Delta_I) \leq 1 \quad \forall \omega \quad (17)$$

with $\ell_I(\omega) = |W_I(j\omega)|$. (10) indicates that Robust Stability is guaranteed if $\bar{\sigma}(\tilde{H}_I) \leq 1/\ell_I(\omega) \quad \forall \omega$. However, from Fig. 3 it is seen that this condition is violated over a wide frequency range, despite the fact that the system is known to be robustly stable. The reason for the conservativeness of condition (10) in this instance is that the use of unstructured uncertainty (17) includes plants not included in the "true" uncertainty description (16). These problems may be avoided by using the structured singular value $\mu(\tilde{H}_I)$ as discussed in Section IV.2.

3. Integral Control and Robust Stability

Because of the importance of integral control in the context of process control we will derive specifically conditions under which controllers with integral action can be designed in the presence of uncertainty. We will keep the uncertainty as general as possible. To

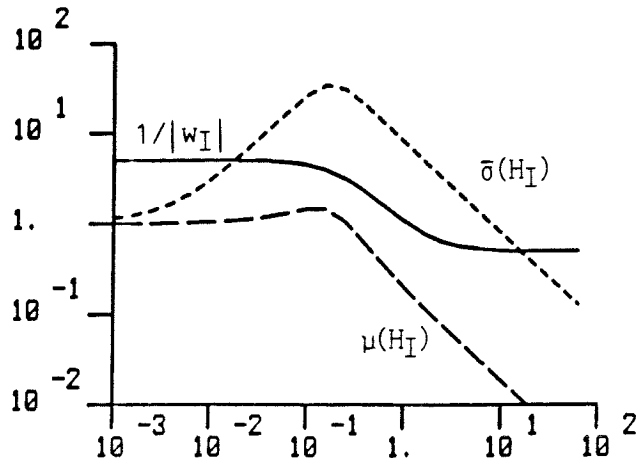


Figure 3. Robust stability for the distillation column with diagonal input uncertainty is guaranteed since $\mu(H_I) \leq 1/|w_I| \quad \forall \omega$. The use of unstructured uncertainty and $\bar{\sigma}(H_I)$ is conservative, and would require the system to be detuned to guarantee robust stability.

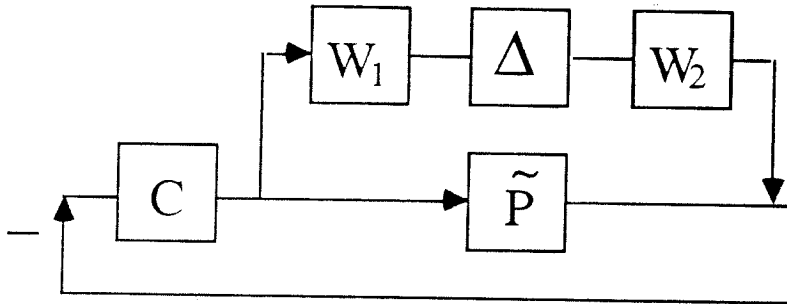


Figure 4. System with weighted additive uncertainty. Rearranging this system to fit Fig. 6 gives $M = W_1 C (I + \tilde{P} C)^{-1} W_2$.

this end define Π_A as the set of plants which is generated by a single weighted additive norm bounded perturbation (Fig. 4)

$$\Pi_A = \{P: P = \tilde{P} + L_A\}, \quad L_A = W_2 \Delta_A W_1, \quad \bar{\sigma}(\Delta_A) \leq 1, \quad \forall \omega \quad (18)$$

This is a generalization of (7) because the weights W_1 and W_2 are allowed to be matrices. Π_A includes additive uncertainty (5a) ($W_1 = I_A$, $W_2 = I$), multiplicative input uncertainty (5b) ($W_1 = I$, $W_2 = \tilde{P} l_I$) and multiplicative output uncertainty (5c) ($W_1 = \tilde{P} l_O$, $W_2 = I$) as special cases.

A necessary and sufficient condition for "perfect control" and robust stability will be stated first. Note that "perfect control" ($\tilde{H} = I, \forall \omega$) is clearly not possible for real systems which must be strictly proper, (i.e., $\tilde{H} \rightarrow 0$ as $\omega \rightarrow \infty$), but the notion of "perfect control" is nevertheless useful.

Theorem 3. Perfect Control ($\Pi = \Pi_A$). Assume: \tilde{P} is minimum phase, all plants have the same number of unstable poles. Robust stability and "perfect control" ($\tilde{H} = I$) may be achieved

$$\text{iff } \det(P \tilde{P}^{-1}) \neq 0 \quad \forall \omega, \quad \forall P \in \Pi_A \quad (19)$$

This theorem implies that perfect control is possible if and only if none of the plants P in the set Π_A have zeros on the $j\omega$ -axis (i.e., $\det P \neq 0$). From the proof and (5) it is clear that $\det P = 0 \quad \forall \omega$ is equivalent to

Additive uncertainty: $\lambda_A < \underline{\sigma}(\tilde{P}) \quad \forall \omega$

Multiplicative input uncertainty: $\lambda_I < 1 \quad \forall \omega$

Multiplicative output uncertainty: $\lambda_O < 1 \quad \forall \omega$

The necessity of condition (19) is obvious since perfect control ($S = \tilde{S} = 0$) is never possible for plants with RHP zeros. Regarding sufficiency, it is clear from Theorem 2 that for minimum phase plants, perfect control is always possible in principle. The search for zeros is restricted to the $j\omega$ -axis as the result of the particular norm bounded uncertainty assumed, it implies that zeros cannot appear in the RHP without crossing the $j\omega$ -axis. Theorem 3 offers little that is new; it is stated mainly as a means to prove Theorem 7 in Section V. The following conditions for integral control are more interesting.

Theorem 4A. Integral Control. Let Π be any set of plants such that P and \tilde{P} have the same number of RHP (unstable) poles. Also, assume that PC and $\tilde{P}C$ are strictly proper. Then robust stability and integral control ($\tilde{H}(0)=I$) may be achieved

$$\text{only if } \det(P(0)\tilde{P}(0)^{-1}) > 0 \quad \forall P \in \Pi \quad (20)$$

Theorem 4A implies that for stable plants, integral control is never possible if the sign of the plant, expressed in terms of $\det P(0)$, changes. This is a direct generalization of the result for SISO systems. Note that Theorem 4A does not apply to cases where the pole may cross the $j\omega$ -axis. As an example, let $\tilde{P} = \frac{1}{s+a}$ and $P = \frac{1}{s-a}$ ($a>0$). These plants may be stabilized using a single controller with integral action (e.g. $C = \frac{k(s+a)}{s}$, $k>a$) despite of the fact that they do not satisfy condition (20).

For the special case when Π is of the norm-bounded form Π_A (18), condition (19) is both necessary and sufficient:

Theorem 4B. Integral Control($\Pi = \Pi_A$). Assume all plants $P \in \Pi_A$ are stable and that PC and $\tilde{P}C$ strictly proper. For controllers with integral control robust stability may be achieved

$$\text{if and only if } \det(P(0)\tilde{P}(0)^{-1}) > 0 \quad \forall P \in \Pi_A \quad (21)$$

The sufficiency of condition (21) follows mainly from the assumed norm-bounded additive perturbation, and also from the fact that at frequencies $\omega \neq 0$, the robust stability condition (similar to (10)) may always be satisfied by detuning the system (provided the plant is stable). Note that condition (21) does not imply robust stability if Π is not on the form Π_A . As an example, consider the set Π consisting of the two plants $\tilde{P} = \text{diag}\{1,1\}$ and $P = \text{diag}\{-1,-1\}$. Since this

corresponds to two SISO plants where the gains change sign, integral control is not possible.

IV. MULTIPLE PERTURBATIONS ("STRUCTURED" UNCERTAINTY)

In this section, we will describe the uncertainty in a "structured" manner by actually trying to identify the sources and locations of uncertainty in the system. This usually leads to an uncertainty description with multiple perturbations (Δ_i). These perturbations may correspond to uncertainty in the model parameters, uncertainty with respect to the manipulated variables (input or actuator uncertainty) and the outputs (measurement uncertainty), etc. By using such a mechanistic approach, we can norm-bound each perturbation (e.g. $||\Delta_i|| < 1$) without introducing too much additional conservativeness and get a "tight" description of the uncertainty set.

However, we should not necessarily describe the uncertainty as rigorously as possible. Rather, we should take the engineer's approach and describe the uncertainty as rigorously as necessary. This means some of the sources of uncertainty (occurring at different places of the system) should be lumped into an "unstructured" multiplicative perturbation, for example, if this does not add too much conservativeness. This leads to a practical uncertainty description: Some sources of uncertainty are described in a "structured" manner (e.g., parametric uncertainty), while the rest (usually uncertain high-frequency dynamics) is lumped into a single "unstructured" perturbation (see Reactor Example below).

The main objective in this section is to familiarize the reader with the work of Doyle (1982, 1984). The results are presented without further motivation; subsequent examples illustrate how these

results may be used in practice.

1. General Theory

Consider the uncertainty as perturbations on the nominal system. Each perturbation Δ_i is assumed to be a stable and norm-bounded transfer matrix

$$\bar{\sigma}(\Delta_i) < 1 \quad \forall \omega \quad (22)$$

Weighting matrices are used to normalize the uncertainty such that the bound is one at all frequencies; that is, the actual perturbation L_i is written

$$L_i = W_2 \Delta_i W_1 \quad (23)$$

If Δ_i represents a real parameter variation we may restrict Δ_i to be real, but in general Δ_i may be any stable rational transfer matrix satisfying (22). The choice of the singular value $\bar{\sigma}$ as the norm for bounding Δ_i is not arbitrary, but is needed to obtain the necessity in the theorems which follow.

The perturbations (uncertainties) which may occur at different places in the feedback system (e.g., Fig. 5), can be collected and placed into one large block diagonal perturbation matrix

$$\Delta = \text{diag} \{ \Delta_1, \dots, \Delta_n \} \quad (24)$$

for which we have

$$\bar{\sigma}(\Delta) < 1 \quad \forall \omega \quad (25)$$

The blocks Δ_i in (24) can have any size and may also be repeated. For example, repetition is needed in order to handle correlations between the uncertainties in different elements. The nominal closed loop system with no uncertainty ($\Delta=0$) is assumed to be stable. The perturbations (uncertainty) give rise to stability problems because of the "additional" feedback paths created by the uncertainty. This is

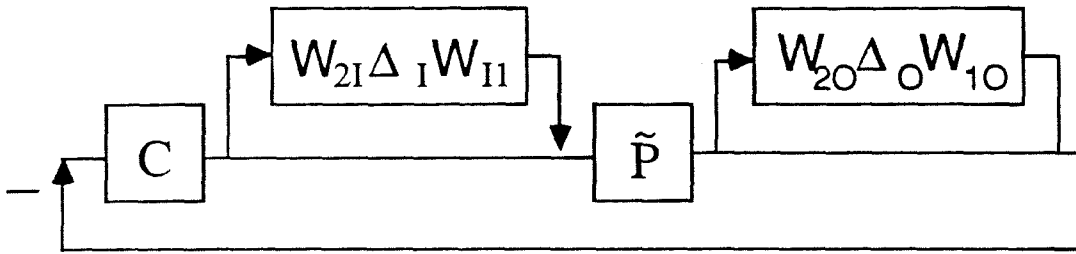


Figure 5. System with weighted multiplicative input and output uncertainty. Rearranging this system to fit Fig. 6 gives M as in Eq. (29).

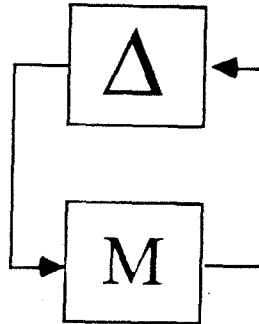


Figure 6. Interconnection structure for studying effect of uncertainty on stability. $\Delta = \text{diag}\{\Delta_1, \dots, \Delta_n\}$.

shown explicitly by writing the uncertainty as perturbations on the nominal system in the form (M Δ -structure) shown in Fig. 6. M is the nominal closed loop system "as seen from" the various uncertainties, and is stable since the nominal system is assumed stable. More precisely, M is the interconnection matrix giving the nominal transfer functions from the output of the perturbations Δ_i to their inputs. Constructing M is conceptually straightforward, but may be tedious for specific problems.

We want to derive conditions on M in order to guarantee robust stability. It may be shown (Doyle et al., 1982) that for a nominally ($\Delta=0$) stable system, robust stability is equivalent to the stability of the M Δ -structure in Fig. 6. This system is stable if and only if $\det(I+\Delta M)$ does not encircle the origin as s traverses the Nyquist D contour for all possible Δ . Because the perturbations are norm bounded, (i.e. all Δ 's satisfying (25) are allowed) this is equivalent to

$$\det(I+\Delta M) \neq 0 \quad \forall \omega, \quad \forall \Delta, \quad \bar{\sigma}(\Delta) < 1 \quad (26)$$

$$\Leftrightarrow \rho(\Delta M) \leq 1 \quad \forall \omega, \quad \forall \Delta, \quad \bar{\sigma}(\Delta) < 1$$

Condition (26) by itself is not very useful since it is only a yes/no condition which must be tested for all possible perturbations Δ . What is desired is a condition on the matrix M, preferably on some norm of M. This is supplied by the following theorem.

Theorem 5. Necessary and Sufficient Condition for Robust Stability

(Doyle et al., 1982). Assume the nominal system ($\Delta=0$) is stable. Then the closed loop system (Fig. 6) is stable for all Δ , $\bar{\sigma}(\Delta) < 1$ if and only if

$$\mu(M) \leq 1 \quad \forall \omega \quad (27)$$

Theorem 5 may be interpreted as a "generalized small gain theorem" applied to (26) which also takes the structure of Δ into account. The function μ , called the Structured Singular Value (SSV), is defined in order to get the tightest possible bound on M such that (26) is satisfied. A more precise definition of μ and some of its properties are given in Appendix 1. It is important to note that $\mu(M)$ depends both on the matrix M and on the structure of the perturbations Δ . $\mu(M)$ is a generalization of the spectral radius $\rho(M)$ and maximum singular value $\bar{\sigma}(M)$ in that $\mu(M) = \rho(M)$ when the perturbation Δ is totally structured ($\Delta = \delta I, |\delta| \leq 1$), and $\mu(M) = \bar{\sigma}(M)$ when the perturbation is unstructured (Δ is a full matrix). Note that the matrix M is a function of the nominal system only, and the condition $\mu(M) \leq 1$ limits the possible nominal transfer functions.

At this point, it is not apparent that the uncertainty description (22)-(25), does indeed provide a useful framework for handling uncertainty. Furthermore, it is not clear how to find the matrix M . Hopefully this will become clearer through the examples below.

2. Input uncertainty for distillation column (Example)

It is now possible to derive a less conservative robust stability test for the distillation column example. Previously, we assumed that the input uncertainty was unstructured, but now Δ_I in (17) may be restricted to be a diagonal matrix which results in a tight description of the uncertainty. The interconnection matrix $M = w_I(s)\tilde{H}_I$ and, from Theorem 5,

$$\text{R.S. iff } \mu(\tilde{H}_I) \leq 1/|w_I(j\omega)| = 1/l_I(\omega) \quad \forall \omega$$

where $\mu(\tilde{H}_I)$ is computed with respect to the diagonal matrix Δ_I . From Fig. 3 we see that this condition is satisfied and robust stability is

guaranteed with the chosen controller (15).

3. Simultaneous Input and Output Multiplicative Uncertainty

Consider the system in Fig. 5 which has both input and output multiplicative uncertainty with respect to the model of the plant \tilde{P} . The possible plants are given by

$$\begin{aligned} P &= (I+L_O)\tilde{P}(I+L_I) \\ L_I &= W_{2I}\Delta_I W_{1I}, \quad \bar{\sigma}(\Delta_I) < 1 \quad \forall \omega \\ L_O &= W_{2O}\Delta_O W_{1O}, \quad \bar{\sigma}(\Delta_O) < 1 \quad \forall \omega \end{aligned} \quad (28)$$

The perturbation block Δ_I represents the multiplicative input uncertainty. If its source is uncertainty with respect to the manipulated variables, then

$$\Delta_I: \text{ diagonal, } W_{1I} = \text{diag}\{w_{1I}\}, \quad W_{2I} = I$$

where w_{1I} represents the relative uncertainty on each manipulated input.

The block Δ_O represents the multiplicative output uncertainty. If its source is uncertainty or neglected deadtimes involved in one or more of the measurements, then

$$\Delta_O: \text{ diagonal, } W_{1O} = \text{diag}\{w_{0I}\}, \quad W_{2O} = I$$

w_{0I} represents the relative uncertainty for each measurement. These sources of input and output uncertainty are present in any plant. Δ_I and Δ_O are restricted to be diagonal matrices, since there is little reason to assume that the actuators or measurements influence each other. However, some of the unmodelled dynamics in the plant \tilde{P} itself, which has cross terms, may be approximated by lumping them into Δ_I or Δ_O , thus making either one of them a "full" matrix.

To examine the constraints on the nominal system imposed by the robust stability requirement for this uncertainty description, let $\Delta =$

$\text{diag}\{\Delta_I, \Delta_O\}$ and rearrange the system in Fig. 5 into the form in Fig. 6.

The interconnection matrix M becomes:

$$M = \begin{bmatrix} -W_{1I}C\tilde{P}(I+C\tilde{P})^{-1}W_{2I} & -W_{1I}C(I+\tilde{P})^{-1}W_{2O} \\ W_{1O}\tilde{P}(I+C\tilde{P})^{-1}W_{2I} & -W_{1O}\tilde{P}C(I+\tilde{P}C)^{-1}W_{2O} \end{bmatrix}$$

$$= \begin{bmatrix} W_{1I} & \\ & W_{1O} \end{bmatrix} \begin{bmatrix} -\tilde{P}^{-1}\tilde{H}\tilde{P} & -\tilde{P}^{-1}\tilde{H} \\ \tilde{S}\tilde{P} & -\tilde{H} \end{bmatrix} \begin{bmatrix} W_{2I} \\ W_{2O} \end{bmatrix} \quad (29)$$

and robust stability is guaranteed for all Δ such that $\bar{\sigma}(\Delta) < 1$ if and only if $\mu(M) \leq 1, \forall \omega$. μ is computed with respect to the structure of Δ which in turn depends on the structure assumed for Δ_I and Δ_O . Note that conditions (9) and (10) in Section IV are special cases of (29) when the weights are assumed to be scalar, Δ_I and Δ_O are "full" matrices, and either $\Delta_I = 0$ or $\Delta_O = 0$. However, this only applies for stable Δ 's; an unstable Δ_O (or Δ_I) may be allowed in condition (9) (or (10)), while only stable Δ 's were allowed when deriving (29).

4. Simultaneous Parametric and Unstructured Uncertainty (Reactor Example)

Consider a perfectly mixed batch reactor where an exothermic reaction is taking place. The reaction temperature T is controlled using the temperature T_C of the fluid in the cooling jacket (the fluid in the cooling jacket may be boiling, and T_C may be adjusted by changing the pressure). A heat balance for the batch reactor gives

$$C_p \dot{T} = (-\Delta H_r)r - UA(T-T_C)$$

where

T	reactor temperature (K)
T_C	coolant temperature (K)
r	reaction rate (function of T) (mol/s)
ΔH_r	heat of reaction (negative constant) (J/mol)
C_p	total heat capacity of fluid in reactor (J/K)
UA	overall heat transfer coefficient (J/sK)

Linearizing the reaction rate

$$r = r^0 + k_T dT$$

results in a linear transfer function from T_c to T

$$dT(s) = \frac{UA/C_p}{s+a} dT_c(s) \quad (30)$$

where

$$a = \frac{UA - (-\Delta H_r)k_T}{C_p} \quad (31)$$

Two sources of uncertainty will be considered for the linear model (30): 1) The effect of nonlinearity expressed as uncertainty in the pole location a . 2) Neglected high-frequency dynamics.

Pole Uncertainty (Δ_S). Most of the terms in (31) are nearly constant, except for $k_T = \partial r / \partial T$ which is a strong function of temperature (operating point). From (31) we see that the reactor may be open loop stable ($a > 0$) at low temperatures where k_T is small, and unstable at high temperatures where the reaction is more temperature sensitive. To describe the effect temperature has on a , let

$$|a - \tilde{a}| < r_a \tilde{a}$$

\tilde{a} : nominal pole location

r_a : relative "uncertainty" in a (real constant)

If $r_a > 1$, the plant may change between stability and instability.

Equivalently, the possible a 's may be written in terms of a norm-bounded perturbation Δ_S

$$a = \tilde{a}(1 + r_a \Delta_S), \quad |\Delta_S| < 1, \Delta_S \text{ real} \quad (32)$$

and this may be written as an inverse multiplicative perturbation

$(I + w_S \Delta_S)^{-1}$ on the plant.

$$\frac{1}{s+a} = \frac{1}{s+\tilde{a}} \frac{1}{1+w_S(s)\Delta_S}, \quad w_S(s) = \frac{r_a}{1+s/\tilde{a}} \quad (33)$$

Neglected Dynamics (modelled as unstructured output uncertainty (Δ_O)).

A plant always has some unknown uncertainty, mainly at higher frequencies, which cannot be modelled in a "structured" manner (using parametric uncertainty, etc.). These sources are most conveniently modelled as multiplicative uncertainty; in this case we choose to use output multiplicative uncertainty ($I + w_O \Delta_O$). Physically, this uncertainty may include neglected (and unknown) dynamics for changing the cooling temperature T_c (if T_c is manipulated indirectly with pressure), neglected actuator dynamics (the valve used to control pressure) and neglected dynamics introduced by the heat capacity of the walls. A conservative choice for $w_O(s)$ is found by approximating the neglected dynamics as an effective time delay, and choosing $|w_O(j\omega)| \approx 1$ at the frequency where the phase lag represented by the neglected dynamics reaches 60° ($|1 - e^{j\theta}| = 1$ for $\theta = 60^\circ$).

A block diagram representation of the uncertainty is depicted in Fig. 7. Note that both blocks (Δ_S and Δ_O) are in general needed: We cannot lump the pole uncertainty (Δ_S) into the output uncertainty (Δ_O) if the pole is allowed to cross the $j\omega$ -axis. This would result in $|w_O(j\omega)| \rightarrow \infty$ at $\omega = 0$. Similarly, we cannot lump the output uncertainty into the pole uncertainty. The reason is that the inverse multiplicative uncertainty description (Δ_S) cannot be used to model neglected or uncertain RHP zeros (this would require an unstable perturbation Δ_S). It is therefore not suited for handling neglected high frequency dynamics which most certainly include RHP zeros (one simple example is the "dead band" on any valve).

Combining the two scalar perturbations into one block perturbation $\Delta = \text{diag}\{\Delta_S, \Delta_O\}$ and rearranging Fig. 7 to match Fig. 6

gives the following interconnection matrix:

$$M = \begin{bmatrix} w_S \tilde{S} & -w_0 \tilde{H} \\ w_S \tilde{S} & -w_0 \tilde{H} \end{bmatrix} \quad (34)$$

From Theorem 5, robust stability is guaranteed

$$\begin{aligned} & \text{iff } \mu(M) \leq 1 \quad \forall \omega \\ & \text{or } \text{iff } |w_S \tilde{S}| + |w_0 \tilde{H}| \leq 1 \quad \forall \omega \end{aligned} \quad (35)$$

Because of the identity $\tilde{H} + \tilde{S} = 1$, this bound is impossible to satisfy if $|w_S|$ and $|w_0|$ are both "large" (that is, close to one or larger) over the same frequency range. For $r_a > 1$ the pole may cross the $j\omega$ -axis, and $|w_S| > 1$ for $\omega < \omega^* = \tilde{a} \sqrt{r_a^2 - 1}$ and $|w_S| < 1$ for $\omega > \omega^*$.

In that situation, Robust Stability is guaranteed only if the "unstructured" relative uncertainty given in terms of $|w_0(j\omega)|$ reaches one at a frequency higher than ω^* .

If pole uncertainty were the only source of uncertainty ($w_0=0$), the robust stability bound would be $|\tilde{S}| \leq |w_S|$. Since the plant is minimum phase, this bound could always be satisfied by increasing the gain and making \tilde{S} small, regardless of the size of r_a .

In summary, the pole location uncertainty is handled by "tightening" the control at low frequencies. Indeed, \tilde{S} small ("tight" control) is needed in order to stabilize an unstable plant. However, to realize robust stability in face of the uncertain high-frequency dynamics, it is necessary to detune the system and make \tilde{H} small ($\tilde{S} \approx 1$) at frequencies where $w_0(\omega)$ is larger than one. The implication of this result for process design is that we cannot stabilize an unstable plant if there are RHP-zeros or model uncertainty in the same frequency range as the location of the unstable pole.

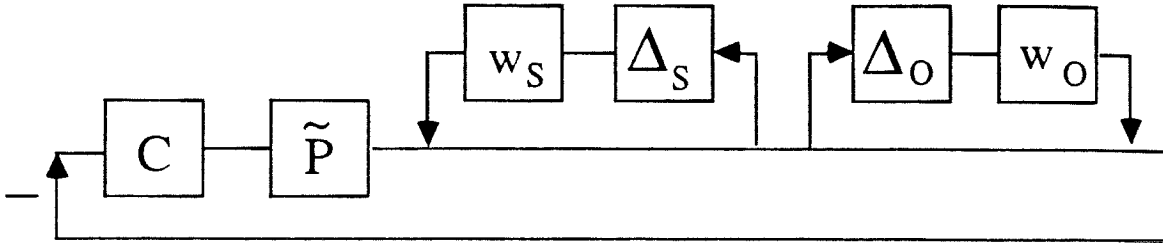


Figure 7. Block diagram representation of uncertainty for Reactor Example. Δ_O represents the neglected (and uncertain) high-frequency dynamics. Δ_S represents the pole uncertainty (changes between stability and instability are possible if $|w_S(j\omega)| > 1$ at some frequency). Rearranging this system to fit Fig. 6 gives M as in Eq. (34).

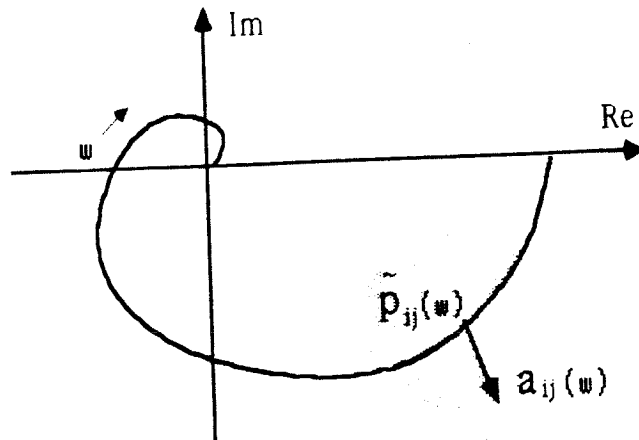


Figure 8. Additive element uncertainty: $|p_{ij} - \tilde{p}_{ij}| < a_{ij}(\omega)$. The disk represents the set of possible $p_{ij}(j\omega)$ at a given frequency.

5. Independent Uncertainty in the Transfer Matrix Elements

In many cases the uncertainty is most easily described in terms of uncertainties on the individual transfer matrix elements. This kind of uncertainty description may arise from an experimental identification of the system. No claim is made that this uncertainty description is a good representation of how the uncertainty actually occurs, but it is included as a possibly useful description in some cases.

The simplest form of element uncertainty arises from the assumption that each element p_{ij} in the plant P is independent, but confined to a disk with radius $a_{ij}(\omega)$ around \tilde{p}_{ij} in the Nyquist plane (Fig. 8), i.e.

$$|p_{ij} - \tilde{p}_{ij}| < a_{ij}(\omega) \quad \forall \omega \quad (36)$$

This corresponds to treating each element as an independent SISO plant with additive uncertainty of size $a_{ij}(\omega)$. Multiplicative (relative) uncertainty r_{ij} on the elements may also be written in the form (36) by using

$$a_{ij}(\omega) = r_{ij}(\omega) |\tilde{p}_{ij}| \quad (37)$$

The main limitation of the uncertainty description (36) is that correlations between the elements cannot be handled (potentially very conservative as shown for the distillation example in Section V.3).

Defining the complex perturbation, Δ_{ij} , (36) becomes

$$p_{ij} - \tilde{p}_{ij} = \Delta_{ij} a_{ij} \quad , \quad |\Delta_{ij}| < 1 \quad (38)$$

Or equivalently, in matrix form

$$P - \tilde{P} = \begin{bmatrix} \Delta_{11}a_{11} & \Delta_{12}a_{12} & \cdots \\ \Delta_{21}a_{21} & & \\ & \ddots & \\ & & \Delta_{nn}a_{nn} \end{bmatrix} \quad (39)$$

Introducing weighting matrices E and L it is possible to rewrite (39) in terms of the "large" diagonal perturbation matrix Δ_E

$$P - \tilde{P} = E\Delta_E L \quad (40)$$

where $E \in \mathbb{R}^{n \times n^2}$, $L \in \mathbb{R}^{n^2 \times n}$ and $\Delta_E \in \mathbb{C}^{n^2 \times n^2}$ are defined as

$$E = [I \ I \ \dots \ I], \quad L = \begin{bmatrix} \underline{a}_1 & & \\ & \underline{a}_2 & \\ & & \underline{a}_n \end{bmatrix}, \quad \underline{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{bmatrix} \quad (41)$$

$$\Delta_E = \text{diag}\{\Delta_{11}, \Delta_{21}, \dots, \Delta_{nn}\}, \quad |\Delta_{ij}| < 1$$

A block diagram representation of (40) is given by Fig. 4 with $W_2 = E$ and $W_1 = L$. This system may be rearranged into the form in Fig. 6 with the interconnection matrix $M = LC(I + \tilde{P}C)^{-1}E = L\tilde{P}^{-1}\tilde{H}E$. From Theorem 5 follows the necessary and sufficient condition for robust stability:

$$\mu(L\tilde{P}^{-1}\tilde{H}E) \leq 1 \quad \forall \omega \quad (42)$$

where μ is computed with respect to the diagonal matrix Δ_E . In principle, this condition may be used to generate all nominal closed loop transfer matrices \tilde{H} for which the closed loop system is robustly stable. Alternatively, it may be used to check whether a particular design meets the robust stability requirement. However, at the design stage (when dynamic resilience is to be determined), \tilde{H} is not known, but rather the restrictions on \tilde{H} as imposed by the uncertainty are of interest. In order to obtain an explicit bound on \tilde{H} from (42), assume that $\tilde{H} = \tilde{h}I$, that is, assume the nominal response is decoupled with identical responses. From (42) follows:

$$\text{R.S. } (\tilde{H} = \tilde{h}I) \quad \text{iff} \quad \sigma(\tilde{H}) = |\tilde{h}| \leq \frac{1}{\mu(L\tilde{P}^{-1}E)} \quad \forall \omega \quad (43)$$

Again, this bound shows that the system has to be detuned and

performance be sacrificed when the uncertainty is large, that is, in this case when $\mu(L\tilde{P}^{-1}E) > 1$. $\mu(L\tilde{P}^{-1}E)$ is a measure of the dynamic resilience which takes into account both the size of the uncertainty and the sensitivity of the plant to uncertainty. In Section V.3, a numerical example incorporating condition (43) is provided.

Upper bounds for $\mu(L\tilde{P}^{-1}\tilde{H}E)$. Alternative sufficient robust stability conditions for this uncertainty description have been derived by Kouvaritakis and Latchman (1985) and by Kantor and Andres (1983). Kouvariakis and Latchman's (1985) condition provides a tight upper bound on μ

$$\mu(L\tilde{P}^{-1}\tilde{H}E) \leq \min_{D_1, D_2} \frac{\bar{\sigma}(D_1 A D_2)}{\underline{\sigma}(D_1 \tilde{H}^{-1} \tilde{P} D_2)} \stackrel{\text{def}}{=} k^*(A, \tilde{H}^{-1} \tilde{P}) \quad (44)$$

Here $A = \{a_{ij}\}$ and D_1 and D_2 are diagonal matrices with real, positive entries. Kouvaritakis and Latchman (1985) claim that (44) is an equality, but their proof is wrong (Doyle, 1986). However, the bound is tight in most cases and is useful since it is easier to compute than $\mu(L\tilde{P}^{-1}\tilde{H}E)$. Another upper bound which is even easier to compute, but is more conservative, is given by Kantor and Andres (1983)

$$\mu(L\tilde{P}^{-1}\tilde{H}E) \leq \rho(A|\tilde{P}^{-1}\tilde{H}|) \quad (45)$$

The spectral radius of the positive matrix $A|\tilde{P}^{-1}\tilde{H}|$ is easily computed as the Perron-Frobenius root of the matrix.

Special case: Equal relative uncertainty. Consider the special case when all the elements have the same relative uncertainty r , i.e.

$$A = r|\tilde{P}| \quad (46)$$

Assuming $\tilde{H} = \tilde{h}I$ and using (44) and (45), the robust stability condition (42) becomes

$$\text{R.S. if } |\tilde{h}| < 1/r k^*(|\tilde{P}|, \tilde{P}) \quad \forall \omega \quad (47)$$

$$\text{if } |\tilde{h}| < 1/r \rho(|\tilde{P}||\tilde{P}^{-1}|) \quad \forall \omega \quad (48)$$

These conditions obviously also hold for the case when the relative uncertainties are different provided that r is replaced by the largest relative uncertainty in any element, r_{\max}

$$r_{\max}(\omega) = \max_{ij} r_{ij}(\omega) \quad (49a)$$

where

$$r_{ij} = \max_{P \in \pi} \left| \frac{p_{ij} - \tilde{p}_{ij}}{\tilde{p}_{ij}} \right| \quad (49b)$$

Note that $k^*(|\tilde{P}|, \tilde{P})$ can be viewed as a minimized condition number. A more thorough discussion of this result appears in Section V.

V. THE CONDITION NUMBER AS A SENSITIVITY MEASURE

This section discusses the use of the condition number $\gamma(\tilde{P})$ as a sensitivity measure with respect to uncertainty. A plant is called ill-conditioned if the condition number $\gamma(\tilde{P})$ is high. Physically, this means that the gain of the plant is strongly dependent on the input direction (see Notation). We give two interpretations to the condition number:

1. The minimized condition number, $\gamma^*(\tilde{P})$ is a stability sensitivity measure with respect to independent uncertainty on the elements with similar relative uncertainty.
2. $\gamma(\tilde{P})$ is a robust performance sensitivity measure with respect to input uncertainty (as discussed following Eq. (12)).

1. $\gamma^*(\tilde{P})$ as a sensitivity measure

It has been argued previously in a somewhat qualitative manner (Grosdidier, et al., 1985) that for robust stability the minimized condition number $\gamma^*(\tilde{P})$ is a measure of sensitivity with respect to model uncertainty. Furthermore, there is a direct relationship between large elements in the Relative Gain Array (RGA) and $\gamma^*(\tilde{P})$ (Grosdidier, et al., 1985), and large elements in the RGA are often claimed to indicate sensitivity to model uncertainty. It will be shown that the minimized condition number $\gamma^*(\tilde{P})$ and the RGA are useful measures with respect to element uncertainty, but only if the relative errors of the transfer matrix elements are independent and have similar magnitude bounds. This proves to be a restrictive assumption in many cases.

Express the uncertainty in terms of the largest relative uncertainty, r_{\max} , in any of the transfer matrix elements (49). This uncertainty description is independent of scaling.

Theorem 6. Condition number criterion. Let Π be any set of plants such that P and \tilde{P} have the same number of RHP (unstable) poles, and define r_{\max} as in (49). Assume the nominal response is decoupled, $\tilde{H} = \text{diag}\{\tilde{h}_i\}$ and assume the system is nominally stable. Then robust stability is guaranteed

$$\text{if } |\tilde{h}_i| < \frac{1}{r_{\max} \gamma_a^*(\tilde{P})} \quad \forall \omega, \quad \forall i \quad (50)$$

which is satisfied

$$\text{if } |\tilde{h}_i| < \frac{1}{r_{\max} \sqrt{n} \gamma^*(\tilde{P})} \quad \forall \omega, \quad \forall i \quad (51)$$

$\gamma^*(\tilde{P})$ is the minimized condition number and $\gamma_a^*(\tilde{P})$ is the minimized "absolute" condition number as defined in the Notation. The minimized condition numbers $\gamma^*(\tilde{P})$ and $\gamma_a^*(\tilde{P})$ are similar in magnitude since (Lemma 2, Appendix 2)

$$\gamma_a^*(\tilde{P})/\sqrt{n} \leq \gamma^*(\tilde{P}) \leq \gamma_a^*(\tilde{P})$$

Condition (50) in Theorem 6 is very similar to condition (47) involving $k^*(|\tilde{P}|, \tilde{P})$, but there are two differences:

1. Condition (50) also holds when the decoupled nominal responses are not identical.
2. Condition (50) is less conservative since $\gamma_a^*(\tilde{P}) \leq k^*(|\tilde{P}|, \tilde{P})$ (use $\bar{\sigma}(|D_1 \tilde{P} D_2|) \leq \bar{\sigma}(D_1 |\tilde{P}| D_2)$).

By comparing (47), (48), (50) and (51) the following chain of inequalities is obtained

$$\gamma^*(\tilde{P}) \leq \gamma_a^*(\tilde{P}) \leq k^*(|\tilde{P}|, \tilde{P}) \leq \rho(|\tilde{P}| |\tilde{P}^{-1}|) \quad (52)$$

Condition (50) is clearly conservative if the individual relative uncertainties on the elements, r_{ij} , are different in magnitude. However, from the discussion following Eq. (44), the bound is expected to be tight when the relative error bounds are equal; in fact, the bound is the tightest possible for 2x2 plants.

Theorem 7 (2x2): Assume $\tilde{H} = \tilde{h}I$ and $r_{ij} = r_{\max}$. Then condition (50) in Theorem 6 is necessary and sufficient for robust stability.

In particular, Theorem 7 implies that, for the case of equal relative element uncertainty (using the nomenclature from Section IV.5),

$$(2x2): \mu(L\tilde{P}^{-1}E) = r\gamma_a^*(\tilde{P}) \quad (53)$$

and for higher order systems $r\gamma_a^*(\tilde{P})$ is expected to give a tight upper bound on $\mu(L\tilde{P}^{-1}E)$.

Improved condition at steady state. The uncertainty description above assumes that each transfer matrix element is given by

$$p_{ij} = \tilde{p}_{ij}(1+r_{ij}\Delta_{ij}), \quad |\Delta_{ij}| < 1$$

where Δ_{ij} is a complex scalar. This may be reasonable at non-zero frequencies, but does not make any physical sense at steady state ($\omega=0$) where \tilde{P} , P and Δ_{ij} are real. Theorem 7 may therefore be conservative at $\omega = 0$ where complex perturbations cannot occur. Fortunately, for 2x2 systems it turns out that we can derive a tight condition by replacing $\gamma_a^*(\tilde{P})$ by $\gamma^*(\tilde{P})$. Indeed, for the case of equal relative uncertainty,

$$(2x2): \mu_{\text{real}}(L\tilde{P}^{-1}E) = r\gamma^*(\tilde{P}) \quad (\omega=0) \quad (54)$$

Theorem 4B and (54) may be combined into the following theorem.

Theorem 8. (2x2) Integral Control ($\tilde{H}(0)=I$)

Assume the uncertainties of the elements in $\tilde{P}(0)$ are independent and real and have equal relative magnitude bounds r . Then for open loop stable systems, robust stability and integral control may be achieved

$$\text{iff } \gamma^*(\tilde{P}(0)) < 1/r \quad (55)$$

If the magnitude bounds on the relative uncertainties are not equal, and r is replaced by r_{\max} , Theorem 8 provides a sufficient condition for robust stability and integral control. Theorem 8 is unique to 2x2 systems; numerical examples show no such relationship for systems of higher dimensions.

Theorem 7 and 8 give very clear interpretations of the minimized condition numbers as sensitivity measures: $\gamma^*(\tilde{P}(0))$ and $\gamma_a^*(\tilde{P}(j\omega))$ are accurate measures of sensitivity only if the plant uncertainties are given in terms of independent (uncorrelated) norm-bounded elements with equal relative error bounds. For other uncertainty structures the minimized condition number may be a very misleading sensitivity measure, and bounds on the uncertainties such as (55) may be arbitrarily conservative. This will be illustrated by a subsequent example.

2. Relationship to the RGA

A relationship between $\gamma^*(\tilde{P})$ and the induced 1- and ∞ -norms of the RGA has been conjectured by Grosdidier et al. (1985):

$$\gamma^*(\tilde{P}) \leq 2 \max[\|RGA\|_1, \|RGA\|_\infty] \quad (56)$$

Numerical examples show that this bound does not hold for systems of dimension 4x4 or higher. However, for 2x2 systems (56) holds even with $\gamma_a^*(\tilde{P})$, and a stronger result is:

Theorem 9 (2x2): $\gamma_a^*(\tilde{P}) \leq \|RGA\|_1 \quad (57)$

Note that for 2x2 systems $\|RGA\|_1 = 2\|RGA\|_{i1} = 2\|RGA\|_{i\infty}$. Numerical examples for 3x3 and 4x4 systems support the following extension to systems with higher dimensions:

$$\text{Conjecture 1 (nxn): } \gamma_a^*(\tilde{P}) \leq \|RGA\|_1 + k(n) \quad (58)$$

with $k(2) = 0$, $k(3) \approx 1$ and $k(4) \approx 2$.

The use of the function $k(n)$ was suggested by Nett (1986). For real matrices and high condition numbers, $\|RGA\|_1$ approaches $\gamma_a^*(\tilde{P})$. The bound (58) appears to be most conservative for small condition numbers. Note that these relationships also hold for the frequency dependent RGA if it is defined as in the Notation.

Theorems 6 and 7 and Conjecture 1 provide at least a partial explanation of why ill-conditioned multivariable systems with large RGA should already be avoided at the design stage: When γ_a^* (or γ^*) or equivalently $\|RGA\|_1$ is large, then the performance measured in terms of $|\tilde{h}|$ is very restricted (c.f. (50)) even if the model uncertainty r_{\max} is small.

3. Integral Control of High Purity Distillation Column (Example)

This example will illustrate that the stability bounds (50) and (55) can be extremely conservative if the element uncertainties are not independent. Once again consider the distillation column of Table 1, but this time with reflux L and boilup V as the manipulated inputs. The steady state gain matrix is

$$\tilde{P}(0) = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix}$$

and

$$(RGA)_{11} = 35.07, \|RGA\|_1 = 138.275, \gamma^*(\tilde{P}) = \gamma_a^*(\tilde{P}) = 138.268, \gamma(\tilde{P}) = 141.7$$

From the high condition number $\gamma^*(\tilde{P})$, one might conclude that the plant may become singular for very small perturbations. This would be true if the uncertainty had the form of independent element errors, but not necessarily otherwise. To illustrate this point consider conditions for using integral control ($\tilde{H}(0)=I$) under two different assumptions about the uncertainty.

Case 1: The elements are assumed independent and norm bounded with equal relative error r . From Section IV.5, Eq. (43) implies that robust stability with integral control may be achieved iff $\mu(L\tilde{P}^{-1}E) < 1$ ($\omega=0$) where μ is computed with respect to the real perturbation matrix Δ_E . Here:

$$E = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, L = r \begin{bmatrix} 0.878 & 0 \\ 1.082 & 0 \\ 0 & 0.864 \\ 0 & 1.096 \end{bmatrix}$$

$$L\tilde{P}^{-1}E = r \begin{bmatrix} 35.07 & -27.65 & 35.07 & -27.65 \\ 34.07 & -27.65 & 34.07 & -27.65 \\ 43.22 & -34.07 & 43.22 & -34.07 \\ 43.22 & -35.07 & 43.22 & -35.07 \end{bmatrix}$$

which gives $\mu_{\text{real}}(L\tilde{P}^{-1}E) = 138.268 r$ which is equal to $r\gamma^*(P)$ as expected from (54). The upper bound $\rho(\|\tilde{P}\|\|\tilde{P}^{-1}\|)$ on $\gamma^*(\tilde{P})$ (52) happens to give the same result, i.e. $\rho(\|\tilde{P}\|\|\tilde{P}^{-1}\|) = 138.268$. Consequently, robust stability with integral action is possible iff $r < \frac{1}{\gamma^*(P)} = 0.0072$. In practice, the variation in each element (mainly due to nonlinearities) is much larger than 0.7%, and integral control does not seem to be possible for this distillation column.

Case 2: A more realistic uncertainty description for this high purity distillation column is the following additive uncertainty (Skogestad, 1986)

$$P - \tilde{P} = \begin{bmatrix} d & -d \\ -d & d \end{bmatrix}$$

which may be written in terms of one real scalar Δ -block

$$P - \tilde{P} = W_2 \Delta W_1, \quad W_2 = |d| \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad W_1 = \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad |\Delta| < 1$$

This highly structured uncertainty is mainly due to the material balance constraints which cannot be violated. Using Theorem 5, robust stability and integral control ($\tilde{H}(0)=I$) are possible iff $\mu_{\text{real}}(W_1 \tilde{P}^{-1} W_2) \leq 1$ ($\omega=0$). Here $W_1 \tilde{P}^{-1} W_2 = 0 \cdot |d|$ and therefore robust stability and integral control are possible for any value of d and the elements may even change sign without causing stability problems. Thus, despite the high condition number, the system is not at all sensitive to this physically-motivated model error.

VI. Conclusions

To guarantee robust stability, model uncertainty requires feedback controllers be detuned and performance be sacrificed. To what extent detuning proves necessary depends on the size of the uncertainty as well as the sensitivity of the plant.

I) General Case: The Structured Singular Value $\mu(M)$ is by definition the best measure of the effect of uncertainty on performance:

$$\text{Robust stability iff} \quad \mu(M) \leq 1 \quad \forall \omega \quad (27)$$

However, here the issue is not control system design but rather process design. From this viewpoint, systems whose closed loop stability and performance are very sensitive to model error are undesirable because they are either impossible to control or require that enormous effort be put into the design of the control system. Condition (27) assumes that a control system has already been designed

and is therefore unsuitable for screening purposes at the design stage. If additional assumptions are made on the type of model uncertainty and the control structure, achievable performance can be related directly to characteristics of the system itself. In the following summary it is assumed that the nominal closed loop system is decoupled ($\tilde{H}=\tilde{H}I$) with identical responses. This proves to be a reasonable assumption at low frequencies, and leads to the least conservative bounds.

II) Uncorrelated Element Uncertainty: $P - \tilde{P} = E\Delta L$ (40)

$$\Delta = \text{diag}\{\Delta_i\}, \bar{\sigma}(\Delta_i) < 1$$

Robust stability iff

$$|\tilde{h}| < \frac{1}{\mu(L\tilde{P}^{-1}E)} \quad \forall \omega \quad (43)$$

III) Uncorrelated element uncertainty with similar relative errors:

Each element: $P_{ij} = \tilde{P}_{ij}(1+r_{ij}\Delta_{ij}), |\Delta_{ij}| < 1$

Largest relative error: $r_{\max} = \max_{ij} r_{ij}$

1. Robust stability if

$$|\tilde{h}| < \frac{1}{r_{\max}\gamma_a^*(\tilde{P})} \quad \forall \omega \quad (50)$$

2. 2x2 systems, $r_{ij} = r \quad \forall i,j, \text{ complex } \Delta_{ij}$

Robust stability iff (Theorem 7)

$$|\tilde{h}| < \frac{1}{r\gamma_a^*(\tilde{P})} \quad \forall \omega$$

3. 2x2 systems, $r_{ij} = r \quad \forall i,j, \text{ real } \Delta_{ij}$

Integral control and Robust stability may be achieved iff

$$|\tilde{h}(0)| = 1 < \frac{1}{r\gamma^*(\tilde{P}(0))} \quad (55)$$

The minimized condition number γ^* (or γ_a^*) or equivalently the RGA is a reliable indicator of closed-loop sensitivity to element

uncertainty only if the relative errors of the transfer matrix elements are independent (uncorrelated) and have similar magnitude bounds.

Acknowledgements

We are thankful to Pierre Grosdidier and John C. Doyle for numerous useful discussions and remarks. Financial support from the National Science Foundation, the Department of Energy and Norsk Hydro is gratefully acknowledged.

References

- Doyle, J. C. and G. Stein. Multivariable Feedback Design: Concepts for a Classical Modern Synthesis. IEEE Trans. Autom. Control, AC-26, 1,4-16 (1981).
- Doyle, J. C. Analysis of Feedback Systems with Structured Uncertainties, Inst. Elect Engrs. Proc., Pt. D., 129, D(6), 242-247 (1982).
- Doyle, J. C., J. E. Wall and G. Stein. Performance and Robustness Analysis for Structured Uncertainty, Proc. IEEE Conference on Decision and Control, Orlando, FL (1982).
- Doyle, J. C. Lecture Notes, ONR/Honeywell Workshop on Advances in Multivariable Control, Minneapolis, MN, Oct. 8-10 (1984).
- Doyle, J. C. Personal communication (1986).
- Grosdidier, P., M. Morari and B. R. Holt. Closed Loop Properties from Steady State Gain Information. Ind. Eng. Chem. Fund., 24, 221-235 (1985).
- Holt, B. R. and M. Morari. Design of Resilient Processing Plants V - The Effect of Deadtime on Dynamic Resilience. Chem. Eng. Sci., 40, 7, 1229-1237 (1985a).
- Holt, B. R. and M. Morari, Design of Resilient Processing Plants VI - The Effect of Right-Half-Plane Zeros on Dynamic Resilience, Chem. Eng. Sci., 40, 59-74 (1985b).
- Kantor, J. C. and R. N. Andres, Characterization of "Allowable Perturbations" for Robust Stability, IEEE Trans. Autom. Control, AC-28, 1, 107-109 (1983).
- Kouvaritakis, B. and M. Latchman, Necessary and Sufficient Stability

- Criterion for Systems with Structured Uncertainties: The Major Principal Direction Alignment Principle, *Int. J. Control*, 42, 3, 575-598 (1985).
- Morari, M. Design of Resilient Processing Plants III - A General Framework for the Assessment of Dynamic Resilience. *Chem. Eng. Sci.*, 38, 1881-1891 (1983).
- Morari, M. and J. Doyle, A Unifying Framework for Control System Design under Uncertainty and Its Implication for Chemical Process Control, in *Chemical Process Control - CPC III* (M. Morari & T. J. McAvoy, eds.) CACHE & Elsevier, Amsterdam 5-52 (1986).
- Nett, C. N., personal communication (1986).
- Postlethwaite, I. and Y. K. Foo, Robustness With Simultaneous Pole and Zero Movement Across the $j\omega$ -Axis, *Automatica*, 21, 4, 433-443 (1985).
- Skogestad, S. and M. Morari. Understanding Distillation, in preparation (1986).
- Stone, B. F. Best Possible Ratios of Certain Matrix Norms, *Numerische Mathematik*, 4, 114-116 (1962).

Notation

$C(s)$ - rational transfer matrix of fixed-parameter controller

$P(s)$ - nxn square rational transfer matrix of actual plant =
 $\{p_{ij}\}$

$\tilde{P}(s)$ - nxn square rational transfer matrix of nominal plant =
 $\{\tilde{p}_{ij}\}$

Π - set of all possible plants, i.e. $P \in \Pi$ and $\tilde{P} \in \Pi$

$|G|$ - matrix G with all elements replaced by their absolute value

$\|G\|_1 = \sum_{i,j} |g_{ij}|$ - 1-norm of matrix G

$\|G\|_2 = \left[\sum_{i,j} |g_{ij}|^2 \right]^{1/2}$ - 2-norm or Frobenius-norm of matrix G .

$\|G\|_{i1} = \max_j \sum_{i=1}^n |g_{ij}|$ - induced 1-norm ("max column sum")

$\|G\|_{i\infty} = \max_i \sum_{j=1}^n |g_{ij}|$ - induced ∞ -norm ("max row sum")

RGA = $Gx(G^{-1})^T$ where x denotes element-by-element multiplication
(also called the Schur or Haddemard product)

$\rho(G)$ - spectral radius of G , i.e. magnitude of largest eigenvalue

$\bar{\sigma}(G)$ - maximum singular value or spectral norm of the transfer matrix G , which at each frequency is equal to the induced 2-norm

$$\bar{\sigma}(G(j\omega)) = \sup_u \frac{\|Gu\|_2}{\|u\|_2} (j\omega) = \|G\|_{i2}$$

$\underline{\sigma}(G)$ - minimum singular value

$$\underline{\sigma}(G(j\omega)) = \min_u \frac{\|Gu\|_2}{\|u\|_2} (j\omega)$$

We have the property $\underline{\sigma}(G) = 1/\bar{\sigma}(G^{-1})$

$\gamma(G)$ = $\bar{\sigma}(G)/\underline{\sigma}(G)$ - condition number

$\gamma_a(G)$ = $\bar{\sigma}(|G|)/\underline{\sigma}(G)$ - absolute condition number

$\gamma^*(G)$ - minimized condition number, $\gamma^*(G) = \min_{D_1, D_2} \gamma(D_1 G D_2)$, where D_1 and D_2 are diagonal matrices with real, positive entries.

For G , 2x2 and real:

$$\gamma^*(G) = \begin{cases} \frac{1+\kappa^{1/2}}{|1-\kappa^{1/2}|} & \kappa > 0 \\ 1 & \kappa \leq 0 \end{cases} \quad (\text{Grosdidier, 1985})$$

$\gamma_a^*(G)$ - minimized absolute condition number,

$$\gamma_a^*(G) = \min_{D_1, D_2} \frac{\bar{\sigma}(|D_1 G D_2|)}{\underline{\sigma}(D_1 G D_2)}$$

$$(2 \times 2): \quad \gamma_a^*(G) = \frac{1+|\kappa|^{1/2}}{|1-\kappa|^{1/2}} \quad (\text{Appendix 2})$$

$\kappa(G)$ - Rijnsdorps interaction measure for 2x2 plant

$$\kappa(G) = \frac{g_{12}g_{21}}{g_{11}g_{22}}$$

$\mu(G)$ - structured singular value (see Appendix 1).

The Laplace variable s or $j\omega$ is omitted in most cases.

Appendix 1. The SSV μ and its properties

Definition (Doyle, 1982). The function $\mu(M)$, called the structured singular value (SSV) is defined at each frequency such that $\mu^{-1}(M)$ is equal to the smallest $\bar{\sigma}(\Delta)$ needed to make $(I+\Delta M)$ singular, i.e.

$$\mu^{-1}(M) = \min_{\delta} \{ \delta | \det(I+\Delta M) = 0 \text{ for some } \Delta, \bar{\sigma}(\Delta) \leq \delta(\omega) \} \quad (A1-1)$$

Δ is a block diagonal perturbation matrix. $\mu(M)$ depends on the matrix M and the structure of the perturbations Δ . The definition of μ may be extended by restricting Δ to a smaller set, e.g., Δ real. The above definition is not in itself useful for computing μ since the optimization problem implied by it does not appear to be easily solvable. Fortunately, Doyle (1982) has proven several properties of μ which makes it more useful for applications.

Properties of μ (Doyle, 1982)

1. The following bounds exist for μ :

$$\rho(M) \leq \mu(M) \leq \bar{\sigma}(M) \quad (A1-2)$$

$\mu(M) = \rho(M)$ in the case $\Delta = \delta I$. $\mu(M) = \bar{\sigma}(M)$ in the case Δ is "unstructured", i.e., Δ is a full matrix.

2. Let \mathcal{U} be the set of all unitary matrices with the same structure as Δ , then

$$\max_{U \in \mathcal{U}} \rho(MU) = \mu(M) \quad (A1-3)$$

This optimization problem is in general not convex.

3. Let \mathcal{D} be the set of real positive diagonal matrices $D = \text{diag}\{d_i I_i\}$ where the size of each block (size of I_i) is equal to the size of the blocks Δ_i . Then for 3 or fewer blocks

$$\min_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) = \mu(M) \quad (A1-4)$$

For 4 or more blocks numerical evidence suggests that (A1-4) gives a tight upper bound on $\mu(M)$. A good estimate for the

scaling matrix D is found by minimizing $\|DMD^{-1}\|_2$ (the Frobenius norm).

4. $\mu(\alpha M) = |\alpha| \mu(M)$, α is a scalar.
5. For real matrices M with real, non-repeated perturbations, the search in (A2-3) may be performed with real matrices U only, and only the cornerpoints ("+1") need to be considered. For (20) and (21) in Theorem 4 this implies that only cornerpoints for the possible $P(0)$'s need to be checked.

Appendix 2. Proof of Theorems

Proof of Theorem 3: The proof uses condition (26) from Section IV which applies to any stable norm bounded perturbation. The interconnection matrix M for the norm bounded additive set of plants, Π_A , is (rearrange Fig. 4 to get Fig. 6)

$$M = W_1 C(I + \tilde{P}C)^{-1} W_2 = W_1 \tilde{P}^{-1} \tilde{H} W_2 \quad (A2-1)$$

Using $P - \tilde{P} = W_2 \Delta_A W_1$ we find $\det(I + \Delta_A M) = \det(I + \Delta_A W_1 \tilde{P}^{-1} \tilde{H} W_2) = \det(I + W_2 \Delta_A W_1 \tilde{P}^{-1} \tilde{H}) = \det(I + (P - \tilde{P}) \tilde{P}^{-1} \tilde{H})$, and assuming nominal stability and using (26) it is found that robust stability is guaranteed

$$\text{iff} \quad \det(I + (P - \tilde{P}) \tilde{P}^{-1} \tilde{H}) \neq 0 \quad \forall P \in \Pi_A \quad (A2-2)$$

Theorem 3 follows from (A2-2) by assuming $\tilde{H} = I$.

Proof of Theorem 4A: Applying the Nyquist stability condition to (8) we see that closed loop stability requires that the image of

$$\det(I + (P - \tilde{P}) \tilde{P}^{-1} \tilde{H}) \quad (A2-3)$$

does not encircle the origin as s traverses the Nyquist D contour for any $P \in \Pi$. For the case of integral control, the image starts from $(\omega=0) \det P(0) \tilde{P}(0)^{-1}$. Using the strictly-proper assumption, the image ends at $(\omega=\infty) \det I = 1$. The image of (A2-3) will therefore always encircle the origin if it starts on the negative real axis and the system will be unstable.

Proof of Theorem 4B:

Necessity: Follows from Theorem 4A

Sufficiency: For this uncertainty description robust stability is guaranteed (Theorem 5)

$$\text{iff} \quad \mu(W_1 \tilde{P}^{-1} \tilde{H} W_2) \leq 1 \quad \forall \omega \quad (A2-4)$$

$$\text{iff} \quad \det(I + (P - \tilde{P}) \tilde{P}^{-1} \tilde{H}) \neq 0 \quad \forall \omega, \quad \forall P \in \Pi_A \quad (A2-2)$$

$\omega = 0$: (A2-2) with $\tilde{H}(0) = I$ is satisfied if (21) is satisfied.

$\omega > 0$: For stable plants, it is always possible to select a controller such that $\tilde{H} = \tilde{h}I$ and $|\tilde{h}| \leq 1/\mu (W_2 \tilde{P}^{-1} W_1) \quad \forall \omega$, i.e., such that (A2-4) holds. This proves that by assuming (21), (A2-4) is satisfied for all ω , and robust stability can always be achieved.

Proof of Theorem 6: Consider any set of plants Π , such that all $P \in \Pi$ have the same number of RHP poles. Assuming nominal stability, the Nyquist stability condition applied to (8) implies that robust stability is guaranteed if and only if the image of $\det(I + (P - \tilde{P})\tilde{P}^{-1}\tilde{H})$ does not encircle the origin as s traverses the Nyquist D-contour for all $P \in \Pi$. A sufficient condition for robust stability using the small gain theorem is therefore

$$\rho((P - \tilde{P})\tilde{P}^{-1}\tilde{H}) < 1 \quad \forall \omega, \quad \forall P \in \Pi \quad (A2-5)$$

Here the spectral radius ρ is invariant under similarity transformations. In particular, let D_1 and D_2 be real diagonal "scaling" matrices P and \tilde{P} . Then for any $P \in \Pi$ we have

$$\begin{aligned} \rho((P - \tilde{P})\tilde{P}^{-1}\tilde{H}) &= \rho(D_1(P - \tilde{P})D_2D_2^{-1}\tilde{P}^{-1}D_1^{-1}D_1\tilde{H}D_1^{-1}) \leq \bar{\sigma}(D_1(P - \tilde{P})D_2D_2^{-1}\tilde{P}^{-1}D_1^{-1}D_1\tilde{H}D_1^{-1}) \\ &\leq \frac{\bar{\sigma}(D_1(P - \tilde{P})D_2)\bar{\sigma}(D_1\tilde{H}D_1^{-1})}{\underline{\sigma}(D_1\tilde{P}D_2)} = \frac{\bar{\sigma}(D_1(P - \tilde{P})D_2)}{\underline{\sigma}(D_1\tilde{P}D_2)} \bar{\sigma}(\tilde{H}) \quad (\text{for } \tilde{H} = \text{diag}\{\tilde{h}_i\}) \end{aligned} \quad (A2-6)$$

$$\leq r_{\max} \frac{\bar{\sigma}(|D_1\tilde{P}D_2|)}{\underline{\sigma}(D_1\tilde{P}D_2)} \bar{\sigma}(\tilde{H}) \quad (\text{Lemma 1}) \quad (A2-7)$$

Combining (A2-5) and (A2-7) and choosing the scalings D_1 and D_2 to get the least conservative bound, R.S. is guaranteed

$$\text{if} \quad \min_{D_1, D_2} \bar{\sigma}(\tilde{H}) r_{\max} \frac{\bar{\sigma}(|D_1\tilde{P}D_2|)}{\underline{\sigma}(D_1\tilde{P}D_2)} < 1 \quad \forall \omega \quad (\tilde{H} = \text{diag}\{\tilde{h}_i\}) \quad (A2-8)$$

which is equivalent to

$$\bar{\sigma}(\tilde{H}) < \frac{1}{r_{\max} \gamma_a^*(\tilde{P})} \quad (\tilde{H} = \text{diag}\{\tilde{h}_i\}) \quad (50)$$

This proves condition (50). Condition (51) follows directly from (50)

by applying Lemma 2 below.

Lemma 1

Consider any set of plants Π . Then,

$$\max_{P \in \Pi} (\bar{\sigma}(D_1(P-\tilde{P})D_2)) \leq r_{\max} \bar{\sigma}(|D_1\tilde{P}D_2|) \quad (A2-9)$$

and equality applies if the set Π is norm bounded with independent elements and all elements have the same relative uncertainty.

Proof:

Let A be the matrix which bounds each element in $P - \tilde{P}$

$$|P - \tilde{P}| \leq A(\omega) \quad \forall P \in \Pi \quad (A2-10)$$

Then

$$\max_{P \in \Pi} \bar{\sigma}(P - \tilde{P}) \leq \bar{\sigma}(A) \leq r_{\max} \bar{\sigma}(|\tilde{P}|) \quad (A2-11)$$

(The first inequality is an equality if all P satisfying (A2-10) may occur in practice). The last inequality follows trivially since $A \leq r_{\max} |\tilde{P}|$. It will be an equality if the relative uncertainty bounds of the elements are equal. To derive (A2-9), note that the relative errors r_{ij} and r_{\max} are unchanged by applying the diagonal scalings D_1 and D_2 to the plant.

Comment: Note that the bound involves $|D_1\tilde{P}D_2|$ and not $D_1|\tilde{P}|D_2$. The last would be more conservative since for D_1 and D_2 real and positive $\bar{\sigma}(|D_1\tilde{P}D_2|) \leq \bar{\sigma}(D_1|\tilde{P}|D_2)$

Lemma 2

Let G be a matrix of size $n \times n$. Then

$$\bar{\sigma}(G) \leq \bar{\sigma}(|G|) \leq \sqrt{n} \bar{\sigma}(G) \quad (A2-12)$$

Proof: The following property is proved by Stone (1962)

$$\frac{1}{\sqrt{n}} \|G\|_2 \leq \bar{\sigma}(G) \leq \|G\|_2$$

Using the obvious property $||G||_2 = |||G|||_2$ we get

$$\bar{\sigma}(|G|) \leq ||G||_2 \leq \sqrt{n} \bar{\sigma}(G) \quad \text{QED}$$

Proof of Theorem 7:

Let \tilde{P} be a nonsingular 2x2 transfer matrix and consider the case of independent elements with equal relative errors r .

$$P = \begin{bmatrix} \tilde{p}_{11}(1+r\Delta_{11}) & \tilde{p}_{12}(1+r\Delta_{12}) \\ \tilde{p}_{21}(1+r\Delta_{21}) & \tilde{p}_{22}(1+r\Delta_{22}) \end{bmatrix}, \quad |\Delta_{ij}| < 1 \quad \forall \omega \quad (\text{A2-13})$$

Comparing (50) with (43) we see that Theorem 7 holds if it can be proved for this uncertainty description that

$$(2 \times 2): \quad \mu(L\tilde{P}^{-1}E) = r\gamma_a^*(\tilde{P}) \quad (53)$$

Since both $\mu(L\tilde{P}^{-1}E)$ and $r\gamma_a^*(\tilde{P})$ scale linearly with r , (53) is equivalent to the following statement

$$\mu(L\tilde{P}^{-1}E) \leq 1 \Leftrightarrow r\gamma_a^*(\tilde{P}) \leq 1$$

Note from (43) that $\mu(L\tilde{P}^{-1}E) \leq 1$ is a condition for having "perfect control" ($\tilde{H}=I$). Then using Theorem 3, which applies to robust stability and "perfect control," we get

$$\mu(L\tilde{P}^{-1}E) \leq 1 \Leftrightarrow \det P(j\omega) \neq 0, \quad \forall \omega$$

Theorem 7 will therefore be correct if for each frequency

$$r\gamma_a^*(\tilde{P}) \leq 1 \Leftrightarrow \det(P(j\omega)) \neq 0, \quad \forall \omega \quad (\text{A2-14})$$

i.e. if it can prove the following statement:

"At each frequency the smallest r which makes $\det P = 0$ is $r = 1/\gamma_a^*(\tilde{P})$ " (A2-15)

Define at each frequency

$$\kappa = \frac{\tilde{p}_{12}\tilde{p}_{21}}{\tilde{p}_{11}\tilde{p}_{22}} = |\kappa| e^{j\phi}$$

and use

$$\det P = 0 \text{ iff } (1+r\Delta_{11})(1+r\Delta_{22}) = \kappa(1+r\Delta_{12})(1+r\Delta_{21}) \quad (\text{A2-16})$$

The smallest r which satisfies (A2-16) is found for $\Delta_{11} = \Delta_{22} = \Delta_1$, $\Delta_{12} = \Delta_{21} = \Delta_2$ and

$$(1+r\Delta_1)^2 = \kappa(1+r\Delta_2)^2$$

Introduce $\Delta_1 = e^{j\phi_1}$, $\Delta_2 = e^{j\phi_2}$, where ϕ_1 and ϕ_2 are free to be chosen, to find:

$$1-|\kappa|^{1/2}e^{j\phi/2} = r|\kappa|^{1/2}e^{j(\frac{\phi}{2}+\phi_2)} - r e^{j\phi_1}$$

The left hand side is fixed. Using geometrical arguments it is evident that the smallest r satisfying this expression is found when ϕ_1 and ϕ_2 are chosen such that the two terms on the right hand side are aligned and in the direction of the left hand side:

$$\begin{aligned} |1-|\kappa|^{1/2}e^{j\phi/2}| &= |1-\kappa|^{1/2}| = r(|\kappa|^{1/2}+1) \\ r &= \frac{|1-\kappa|^{1/2}}{1+|\kappa|^{1/2}} \end{aligned} \quad (\text{A2-17})$$

The derivation of the expression for $\gamma_a^*(\tilde{P})$ is very tedious but straightforward and follows the derivation for $\gamma^*(\tilde{P})$ (Grosdidier et al., 1985). This derivation shows that r given in (A2-17) is equal to $1/\gamma_a^*(\tilde{P})$ which proves (A2-15) and thus proves the theorem.

Proof of Theorem 8:

The proof is similar to that of Theorem 7. The set of plants is again given by (A2-13) but the perturbations are assumed to be real ($-1 < \Delta_{ij} < 1$) and all the elements in \tilde{P} are also assumed to be real. As for the proof of Theorem 7, (54) is proved if we can prove the following statement is proved:

"The smallest r which makes $\det P = 0$ (\tilde{P} and Δ_{ij} real) is $r = 1/\gamma^*(\tilde{P})$ "

(A2-18)

It is necessary to find the smallest r which satisfies (A2-16) when Δ_{ij} is real.

Case 1: $\kappa < 0$. In this case (A2-16) cannot be satisfied for any $r < 1$, but it may clearly be satisfied if $r = 1$ (e.g. choose $\Delta_{12} = -1$ and $\Delta_{11} = -1$). Consequently, the smallest r which makes $\det P = 0$ in this case is $r = 1$, and since $\gamma^*(\tilde{P}) = 1$ for $\kappa < 0$ (Grosdidier, 1985) we have $r = 1/\gamma^*(\tilde{P})$.

Case 2a: $\kappa > 1$. Only cornerpoints of (A2-16) need to be checked (see Appendix 1). Then it is obvious that the smallest r which satisfies (A2-16) for $\kappa > 1$ is the solution of (choose $\Delta_{11} = \Delta_{22} = 1$, $\Delta_{12} = \Delta_{21} = -1$)

$$(1+r)^2 = \kappa(1-r)^2$$

which has as its smallest root $r = (\sqrt{\kappa}-1)/(\sqrt{\kappa}+1) = |1-\sqrt{\kappa}|/(1+\sqrt{\kappa})$

Case 2b: $0 < \kappa < 1$. The smallest r which satisfies (A2-16) in this case is a solution of

$$(1-r^2) = \kappa(1+r)^2$$

which has as its smallest root $r = (1-\sqrt{\kappa})/(1+\sqrt{\kappa}) = |1-\sqrt{\kappa}|/(1+\sqrt{\kappa})$.

$\gamma^*(\tilde{P})$ is given in the Notation. From this it is evident that r is equal to $1/\gamma^*(\tilde{P})$ also for $\kappa > 0$ and this proves statement (54).

QED

Proof of Theorem 9:

For 2x2 systems the RGA becomes

$$\text{RGA} = \begin{bmatrix} \lambda_{11} & 1-\lambda_{11} \\ 1-\lambda_{11} & \lambda_{11} \end{bmatrix}, \quad \lambda_{11} = \frac{\tilde{P}_{11}\tilde{P}_{22}}{\tilde{P}_{11}\tilde{P}_{22}-\tilde{P}_{12}\tilde{P}_{21}} = \frac{1}{1-\kappa} \quad (\text{A2-19})$$

$$\|\text{RGA}\|_1 = 2(|\lambda_{11}| + |1-\lambda_{11}|) = 2 \frac{1+|\kappa|}{|1-\kappa|} \quad (\text{A2-20})$$

Using the expression for γ_a^* (A2-17)

$$\gamma_a^* = \frac{1+|\kappa|^{1/2}}{(|\kappa|-2|\kappa|^{1/2}\cos\phi/2+1)^{1/2}} \leq \frac{1+2|\kappa|^{1/2}+|\kappa|}{|1-\kappa|} \leq \|\text{RGA}\|_1, \quad \text{QED}$$

Chapter III

EFFECT OF DISTURBANCE DIRECTIONS ON CLOSED LOOP PERFORMANCE

EFFECT OF DISTURBANCE DIRECTIONS ON CLOSED LOOP PERFORMANCE

Sigurd Skogestad

Manfred Morari*

Chemical Engineering, 206-41

California Institute of Technology

Pasadena, California 91125

(818)356-4186

submitted to I&EC Process Design & Dev.

May 1986

Abstract

The effectiveness of disturbance suppression in a multivariable control system can depend strongly on the direction of the disturbance. The "disturbance condition number" is introduced to quantify the effect of disturbance direction on closed loop performance. As an example a two point composition control system for a distillation column is analyzed for various disturbances and setpoint changes.

*To whom all correspondence should be addressed

I. Introduction

Disturbance rejection is often the main objective of process control. For multivariable systems, usually each disturbance affects all the outputs. As an example consider a distillation column. A feed composition disturbance corresponding to an increased amount of light component in the feed leads to an increase of both product compositions y_D and x_B . (Here y_D and x_B correspond to the mole fraction of light component in the top and bottom product). In this paper we define as "disturbance direction" the direction of the system output vector resulting from a specific disturbance. As we will show some disturbance directions may be easily counteracted by the control system, while others may not. The aim of this paper is to develop simple measures which may be used to tell how the disturbances are "aligned" with the plant and thus how well they can be rejected.

Consider the linear control system in Fig. 1. The process model is

$$\begin{aligned} \mathbf{y}(s) &= \mathbf{G}(s)\mathbf{m}(s) + \mathbf{G}_d(s) \mathbf{z}(s) \\ &= \mathbf{G}(s)\mathbf{m}(s) + \mathbf{d}(s) \end{aligned} \quad (1)$$

where \mathbf{y} is the output vector, \mathbf{m} is the manipulated input vector and \mathbf{d} represents the effect of the disturbances \mathbf{z} on the outputs. The transfer matrix $\mathbf{G}(s)$ is the process model, and $\mathbf{G}_d(s)$ is the disturbance model expressing the relationship between the physical disturbances z_i and their effect on the output. For a distillation column the components of $\mathbf{z} = (z_1, \dots, z_i, \dots, z_n)^T$ may correspond to disturbances in feed rate, feed composition, boilup rate, etc. The column vector \mathbf{g}_{di} of \mathbf{G}_d represents the disturbance model for each disturbance z_i . The effect of a particular disturbance z_i on the process output is \mathbf{d}_i ,

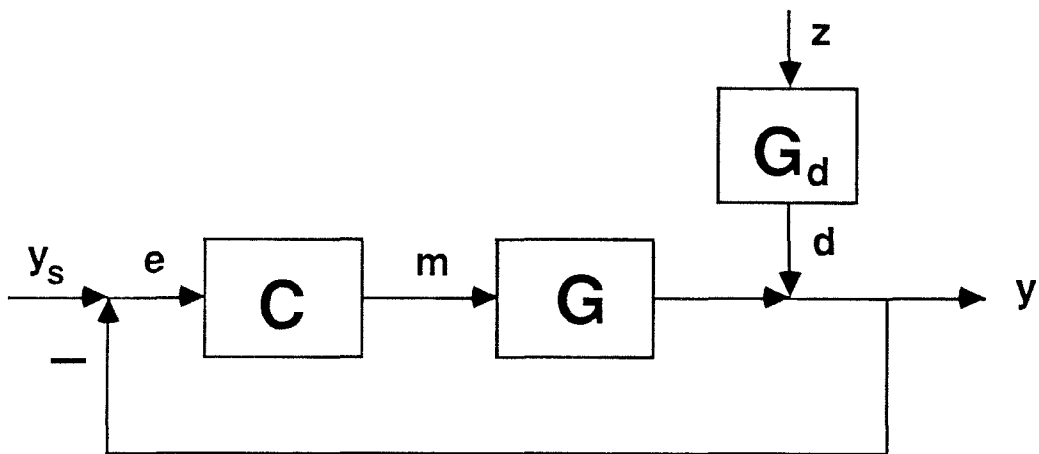


Figure 1. Block diagram of linear control system. The physical disturbances z have the effect $d = G_d z$ on the outputs.

$$\mathbf{d}_i = \mathbf{g}_{di} z_i \quad (2)$$

The direction of the vector \mathbf{d}_i will be referred to as the direction of disturbance i . The overall effect of all disturbances z_i on the output is \mathbf{d} ,

$$\mathbf{d} = \sum_i \mathbf{d}_i = \sum_i \mathbf{g}_{di} z_i = \mathbf{G}_d \mathbf{z} \quad (3)$$

In most cases we will consider the effect of one particular disturbance z_i . To simplify notation we will usually drop the subscript i , and $\mathbf{d} = \mathbf{g}_d z$ will then denote the effect of this single disturbance $z_i = z$ on the outputs. We will also be referring to \mathbf{d} as a "disturbance", although in general it represents the effect of the physical disturbance.

We will consider two different effects of disturbance directions. One is in terms of the magnitude of the manipulated variables \mathbf{m} needed to cancel the influence of the disturbance on the process output completely at steady state. It is independent of the controller \mathbf{C} . This may be used to identify problems with constraints at steady state. However, the issue of constraints at steady state is not really a control problem, but rather a plant design problem. Any well designed plant should be able to reject disturbances at steady state. The second and most important effect of disturbance directions is on closed loop performance. Here we mean by performance the behavior of the controlled outputs \mathbf{y} in the presence of disturbances.

II. Singular Value Decomposition

Throughout this paper we will make use of the Singular Value Decomposition (SVD) of a matrix (Klema and Laub, 1980). Any complex $n \times n$

matrix \mathbf{A} can be written in the form

$$\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^H \quad (4)$$

where \mathbf{U} and \mathbf{V} are unitary matrices ($\mathbf{U}^H = \mathbf{U}^{-1}$) and Σ is a diagonal matrix with real nonnegative entries

$$\Sigma = \text{diag} \{ \sigma_j \} \quad (5)$$

The superscript H denotes complex conjugate transpose. The set of $\{ \sigma_j \}$ are the singular values of \mathbf{A} , and we have

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_n \geq 0$$

The number of nonzero singular values is equal to the rank of the matrix \mathbf{A} . If the matrix \mathbf{A} is nonsingular all singular values are greater than zero, and this will be assumed in the following. The maximum singular value $\sigma_1 = \bar{\sigma}$ and the minimum singular value $\sigma_n = \underline{\sigma}$ are of particular interest because of the properties

$$\max_{\mathbf{v} \neq 0} \frac{||\mathbf{A}\mathbf{v}||_2}{||\mathbf{v}||_2} = \bar{\sigma}(\mathbf{A}) \quad (6a)$$

and

$$\min_{\mathbf{v} \neq 0} \frac{||\mathbf{A}\mathbf{v}||_2}{||\mathbf{v}||_2} = \underline{\sigma}(\mathbf{A}) \quad (6b)$$

Here $||\cdot||_2$ denotes the Euclidean vector norm.

$$||\mathbf{x}||_2 = \left(\sum_i x_i^2 \right)^{1/2}$$

Consequently, $\bar{\sigma}$ corresponds to the largest amplification of the matrix \mathbf{A} and $\underline{\sigma}$ to its smallest amplification. The matrix \mathbf{U} consists of the left singular vectors $\{\mathbf{u}_j\}$, $||\mathbf{u}_j||_2 = 1$, and the matrix \mathbf{V} of the right singular vectors $\{\mathbf{v}_j\}$, $||\mathbf{v}_j||_2 = 1$. For each right singular vector \mathbf{v}_j we have

$$\mathbf{A}\mathbf{v}_j = \sigma_j(\mathbf{A}) \mathbf{u}_j \quad (7)$$

and in particular for the singular vectors associated with the maximum

and minimum singular value

$$\mathbf{A}\bar{\mathbf{v}} = \bar{\sigma}(\mathbf{A})\bar{\mathbf{u}} \quad (8a)$$

$$\mathbf{A}\underline{\mathbf{v}} = \underline{\sigma}(\mathbf{A})\underline{\mathbf{u}} \quad (8b)$$

$\bar{\mathbf{v}}(\mathbf{A})$ therefore corresponds to the direction of the input with the largest amplification, and $\underline{\mathbf{v}}(\mathbf{A})$ to the direction with the smallest amplification. Furthermore

$$\mathbf{A}^{-1} = \mathbf{V} \Sigma^{-1} \mathbf{U}^H \quad (9)$$

which is the SVD of \mathbf{A}^{-1} , but with the order of the singular values reversed. Let $\ell = n-j+1$. It then follows from (9)

$$\sigma_j(\mathbf{A}^{-1}) = 1/\sigma_\ell(\mathbf{A}) \quad (10a)$$

$$\mathbf{u}_j(\mathbf{A}^{-1}) = \mathbf{v}_\ell(\mathbf{A}) \quad (10b)$$

$$\mathbf{v}_j(\mathbf{A}^{-1}) = \mathbf{u}_\ell(\mathbf{A}) \quad (10c)$$

and in particular

$$\bar{\sigma}(\mathbf{A}^{-1}) = 1/\underline{\sigma}(\mathbf{A}) \quad (11a)$$

$$\bar{\mathbf{u}}(\mathbf{A}^{-1}) = \underline{\mathbf{v}}(\mathbf{A}) \quad (11b)$$

$$\underline{\mathbf{u}}(\mathbf{A}^{-1}) = \bar{\mathbf{v}}(\mathbf{A}) \quad (11c)$$

III. Effect of Disturbance Direction on Manipulated Variables

Assume the disturbance model and the process model have been scaled such that at steady state $-1 \leq z_i \leq 1$ corresponds to the expected range of each disturbance and $-1 \leq m_j \leq 1$ corresponds to the acceptable range for each manipulated variable. For process control $m_j = -1$ may correspond to a closed valve and $m_j = 1$ to a fully open valve. The steady state process model is

$$\mathbf{y} = \mathbf{G}_m + \mathbf{G}_d \mathbf{z} \quad (12)$$

For complete disturbance rejection ($\mathbf{y}=0$) we require

$$\mathbf{m} = -\mathbf{G}^{-1} \mathbf{G}_d \mathbf{z} \quad (13)$$

Let $\|\mathbf{x}\|_\infty$ denote the largest component of the vector \mathbf{x} . To avoid problems with constraints we have to require

$$\|\mathbf{m}\|_\infty \leq 1 \text{ for all } \|\mathbf{z}\|_\infty \leq 1$$

Mathematically this is equivalent to requiring

$$\|\mathbf{G}^{-1} \mathbf{G}_d\|_{i\infty} \leq 1 \quad (14)$$

$\|\mathbf{A}\|_{i\infty}$ is the induced ∞ -norm of the matrix \mathbf{A} which is equal to its largest row sum:

$$\|\mathbf{A}\|_{i\infty} = \max_i \left(\sum_j |a_{ij}| \right) \quad (15)$$

Whether (14) is violated and constraints cause problems depends both on the process model \mathbf{G} and the disturbance model \mathbf{G}_d . Even if $\|\mathbf{G}^{-1}\|_{i\infty}$ is "large", $\|\mathbf{G}^{-1} \mathbf{G}_d\|_{i\infty}$ can be "small" if \mathbf{G}_d is "aligned" with \mathbf{G}^{-1} in a certain manner. We will discuss this in more detail below.

The disturbance condition number

Even when constraints are not causing any problems, it is of interest to investigate the magnitude of the manipulated variable necessary to compensate for the effect of a disturbance. In this context it is more reasonable to use the Euclidean (2-) norm as a measure of magnitude because it "sums up" the deviations of all manipulated variables rather than accounting for the maximum deviation only (like the ∞ -norm). Consider a particular disturbance $\mathbf{d} = \mathbf{g}_d \mathbf{z}$. For complete disturbance rejection of this disturbance

$$\mathbf{m} = -\mathbf{G}^{-1} \mathbf{d} \quad (16)$$

The quantity

$$\| \underline{\mathbf{m}} \|_2 / \| \underline{\mathbf{d}} \|_2 = \| \underline{\mathbf{G}}^{-1} \underline{\mathbf{d}} \|_2 / \| \underline{\mathbf{d}} \|_2 \quad (17)$$

depends only on the direction of the disturbance $\underline{\mathbf{d}}$ but not on its magnitude. It measures the magnitude of $\underline{\mathbf{m}}$ needed to reject a disturbance $\underline{\mathbf{d}}$ of unit magnitude which enters in a particular direction expressed by $\underline{\mathbf{d}} / \| \underline{\mathbf{d}} \|_2$.

The "best" disturbance direction, requiring the least action by the manipulated variables, is that of the singular vector $\underline{\mathbf{u}}(\underline{\mathbf{G}})$ associated with the largest singular value of $\underline{\mathbf{G}}$.

$$\underline{\mathbf{d}} = \underline{\mathbf{v}}(\underline{\mathbf{G}}^{-1}) = \underline{\mathbf{u}}(\underline{\mathbf{G}})$$

In this case we find by using (11b)

$$\| \underline{\mathbf{G}}^{-1} \underline{\mathbf{d}} \|_2 / \| \underline{\mathbf{d}} \|_2 = \| \underline{\mathbf{G}}^{-1} \underline{\mathbf{v}}(\underline{\mathbf{G}}^{-1}) \|_2 = \underline{\sigma}(\underline{\mathbf{G}}^{-1}) = 1/\bar{\sigma}(\underline{\mathbf{G}}) \quad (18)$$

By normalizing Eq. (17) with this "best" disturbance, we obtain the following measure which we call the disturbance condition number

$$\gamma_d(\underline{\mathbf{G}}) = \frac{\| \underline{\mathbf{G}}^{-1} \underline{\mathbf{d}} \|_2}{\| \underline{\mathbf{d}} \|_2} \bar{\sigma}(\underline{\mathbf{G}}) \quad (19a)$$

$$= \frac{\| \underline{\mathbf{G}}^{-1} \underline{\mathbf{g}}_d \|_2}{\| \underline{\mathbf{g}}_d \|_2} \bar{\sigma}(\underline{\mathbf{G}}) \quad (19b)$$

It measures the magnitude of the manipulated variables needed to reject a disturbance in the direction $\underline{\mathbf{d}}$ relative to rejecting a disturbance with the same magnitude, but in the "best" direction.

The "worst" disturbance direction is

$$\underline{\mathbf{d}} = \underline{\mathbf{v}}(\underline{\mathbf{G}}^{-1}) = \underline{\mathbf{u}}(\underline{\mathbf{G}})$$

and in this case we get

$$\gamma_d(\underline{\mathbf{G}})_{\max} = \bar{\sigma}(\underline{\mathbf{G}}^{-1}) \bar{\sigma}(\underline{\mathbf{G}}) = \gamma(\underline{\mathbf{G}}) \quad (20)$$

where $\gamma(\underline{\mathbf{G}})$ is the condition number of the plant. It follows

$$1 \leq \gamma_d(\underline{\mathbf{G}}) \leq \gamma(\underline{\mathbf{G}})$$

and $\gamma_d(\underline{\mathbf{G}})$ may be viewed as a generalization of the condition number $\gamma(\underline{\mathbf{G}})$

of the plant, which also takes into account the direction of the disturbances. The disturbance condition number $\gamma_d(G)$ is clearly scaling dependent since $\gamma(G)$ is scaling dependent. We know that ill-conditioned plants ($\gamma(G)$ large) indicate control problems (Morari, 1983, Skogestad and Morari, 1986a). A large value of $\gamma(G)$ indicates a large degree of directionality in the plant G , which may have to be compensated for by the controller in order to get good response. We used "may" in the last sentence, because this also depends on the disturbance direction: If $\gamma_d(G)$ is small for all disturbances, then it really does not matter if $\gamma(G)$ is large.

In the next section we will look at closed loop performance and show explicitly the physical significance $\gamma_d(G)$ in this context. However, let us first look at another measure which has been suggested for measuring disturbance directionality.

The Relative Disturbance Gain

We will show that the Relative Disturbance Gain (RDG) introduced by Stanley et al. (1985) is similar to the disturbance measure $\gamma_d(G)$ defined above, but with a different normalization. One advantage of the RDG is that it is scaling independent, while $\gamma_d(G)$ is scaling dependent. On the other hand, the physical significance of the RDG is less clear than that of $\gamma_d(G)$.

For a particular disturbance z , the RDG, β_ℓ , is defined for each manipulated variable, m_ℓ , as

$$\beta_\ell = \frac{(\partial m_\ell / \partial z) y_j}{(\partial m_\ell / \partial z) y_{\ell, m_j \neq \ell}} \quad (21)$$

$(\partial m_\ell / \partial z) y_j$ is the change in manipulated variable m_ℓ needed for perfect disturbance rejection. $(\partial m_\ell / \partial z) y_{\ell, m_j \neq \ell}$ is the change in manipulated

variable m_ℓ needed for perfect disturbance rejection for the corresponding output y_ℓ , while keeping all other manipulated variables constant. To find the relationship between β_ℓ and $\gamma_d(\mathbf{G})$ use the following identities (Grosdidier, 1985)

$$\left[\frac{\partial m_\ell}{\partial \mathbf{z}} \right]_{y_j} = -(\mathbf{G}^{-1} \mathbf{g}_d)_\ell \quad (22)$$

$$\left[\frac{\partial m_\ell}{\partial \mathbf{z}} \right]_{y_\ell, m_{j \neq \ell}} = -((\mathbf{G}_{\text{diag}})^{-1} \mathbf{g}_d)_\ell \quad (23)$$

Here \mathbf{G}_{diag} denotes the matrix consisting of the diagonal elements in \mathbf{G} .

Using $\mathbf{d} = \mathbf{g}_d \mathbf{z}$, the definition of β_ℓ (21) may be rewritten as

$$\beta_\ell = \frac{(\mathbf{G}^{-1} \mathbf{d})_\ell}{((\mathbf{G}_{\text{diag}})^{-1} \mathbf{d})_\ell} \quad (24)$$

(24) is similar to the definition of $\gamma_d(\mathbf{G})$ in Eq. (19), but with the diagonal plant as the normalization factor. Note that $\beta_\ell = 1$ if \mathbf{G} is diagonal.

We can also define a RDG matrix for the case when we have several disturbances \mathbf{z} as

$$\text{RDG} = \mathbf{G}^{-1} \mathbf{G}_d \div (\mathbf{G}_{\text{diag}})^{-1} \mathbf{G}_d \quad (25)$$

where the division in this case denotes element by element division.

Note the resemblance with the Relative Gain Array (RGA) which may be defined as the matrix

$$\text{RGA} = \mathbf{G} \times (\mathbf{G}^{-1})^T \quad (26)$$

where \times denotes element by element (Schur) multiplication. The RGA is also scaling invariant.

Stanley et al. (1985) claim that the RDG can be used to investigate the effect of decoupling. However, Eq. (21) and (24) clearly show that the RDG is independent of the controller which may or may not include

decoupling. The meaning of (24) when G is replaced by GH , where H denotes a decoupler, is not clear. The variables m_ℓ will then be some internal variable in the controller with no direct physical significance.

Below we will derive an alternative physical interpretation for β_ℓ in terms of closed loop performance which retains its significance when G is replaced by GH in Eq. (24).

IV. Effect of Disturbance Direction on Closed Loop Performance

In Section III we derived measures of how the magnitude of the manipulated variables depends on the disturbance direction. In this section we will rederive these measures in terms of closed loop performance. This will give us a more powerful interpretation of these measures and will allow us to define dynamic measures and to include "decouplers".

The disturbance condition number

One objective of the control system (Fig. 1) is to minimize the effect of the disturbances on the outputs \mathbf{y} . Consider a particular disturbance $\mathbf{d}(s) = \mathbf{g}_d(s)z(s)$. The closed-loop relationship between this disturbance and the outputs is

$$\mathbf{y}(s) = (\mathbf{I} + \mathbf{G}(s)\mathbf{C}(s))^{-1} \mathbf{d}(s) = \mathbf{S}(s)\mathbf{d}(s) \quad (27)$$

$$\text{where} \quad \mathbf{S}(s) = (\mathbf{I} + \mathbf{G}(s)\mathbf{C}(s))^{-1} \quad (28)$$

Let $\|\mathbf{y}(j\omega)\|_2$ denote the Euclidean norm of \mathbf{y} evaluated at each frequency. The quantity

$$\alpha(\omega) = \|\mathbf{S}\mathbf{d}(j\omega)\|_2 / \|\mathbf{d}(j\omega)\|_2 \quad (29)$$

depends only on the disturbance direction but not on its magnitude. $\alpha(\omega)$ measures the magnitude of the output vector $\mathbf{y}(j\omega)$ resulting from a sinusoidal disturbance $\mathbf{d}(j\omega)$ of unit magnitude and frequency ω .

The "best" disturbance direction causing the smallest output deviation is that of the right singular vector $\underline{v}(S)$ associated with the smallest singular value $\underline{\sigma}(S)$ of S . By normalizing $\alpha(\omega)$ with this best disturbance we obtain the disturbance condition number of S^{-1}

$$\gamma_d(S^{-1}) = \frac{\|Sd\|_2}{\underline{\sigma}(S)\|d\|_2} (j\omega) \quad (30)$$

S^{-1} is used in the argument of γ_d for consistency with the previously defined $\gamma_d(G)$ in Eq. (19).

Again

$$1 \leq \gamma_d(S^{-1}) \leq \gamma(S) \quad (31)$$

At low frequencies where the controller gain is high we have

$$S(j\omega) \approx (GC(j\omega))^{-1} \quad (32)$$

In particular, this expression is exact at steady state ($\omega=0$) if we have integral action. Based on this approximation we derive the disturbance condition number of GC .

$$\gamma_d(GC) = \frac{\|(GC)^{-1}d\|_2}{\|d\|_2} \bar{\sigma}(GC)(j\omega) \quad (33)$$

As stated above this measure has physical significance only when $\underline{\sigma}(GC) \gg 1$. To avoid problems with evaluating the measure at $\omega = 0$ write

$$C(s) = k(s)H(s) \quad (34)$$

where $k(s)$ is a scalar transfer function which includes any integral action. $H(s)$ may be viewed as a "decoupler". We have

$$\gamma_d(GC) = \frac{\|(GH)^{-1}d\|_2}{\|d\|_2} \bar{\sigma}(GH)(j\omega) \quad (35)$$

To evaluate how the disturbance direction is aligned with the plant G itself, choose $H = I$ (i.e. the controller is $k(s)I$) and rederive the disturbance condition number of G

$$\gamma_d(G) = \frac{\|G^{-1}d\|_2}{\|d\|_2} \bar{\sigma}(G) \quad (36)$$

$\gamma_d(G)$ can be interpreted in terms of closed loop performance as follows: If a scalar controller $C = k(s)I$ is chosen (which keeps the directionality of the plant unchanged), then $\gamma_d(G)$ measures the magnitude of the output y for a particular disturbance d , compared to the magnitude of the output if the disturbance were in the "best" direction (corresponding to the large plant gain). If $\gamma_d(G) = \gamma(G)$, the disturbance has all its components in the "bad" direction corresponding to low plant gain and low bandwidth. If $\gamma_d(G) = 1$, the disturbance has all its components in the "good" direction corresponding to high plant gain and high bandwidth.

Though a large value of $\gamma_d(G)$ does not necessarily imply bad performance, it usually does. In principle we could choose a compensator C which makes $\gamma_d(GC) = 1$ for all disturbances. However, this controller often leads to serious robustness problems. Whether robustness problems arise or not depends very much on the type of model uncertainty which is encountered. For a detailed analysis the reader is referred to Morari & Doyle (1986) and Skogestad & Morari (1986a).

Decomposition of d along singular vectors

The objective here is to gain insight into the type of dynamic response which is to be expected when disturbances along a particular direction affects a system with a high degree of directionality ($\gamma(S)$ is "large"). The singular vectors $v_j(S)$ of S form an orthonormal basis. The disturbance vector d can be represented in terms of this basis

$$d = \sum_{j=1}^n (v_j(S)^T \cdot d) v_j(S) \quad (37)$$

Then the output y is described by

$$\mathbf{y}(j\omega) = \mathbf{S}\mathbf{d}(j\omega) \quad (38a)$$

$$= \sum_{j=1}^n \mathbf{S}\mathbf{v}_j(\mathbf{S}) (\mathbf{v}_j(\mathbf{S})^T \cdot \mathbf{d}) (j\omega) \quad (38b)$$

$$= \sum_{j=1}^n \sigma_j(\mathbf{S}) \mathbf{u}_j(\mathbf{S}) (\mathbf{v}_j(\mathbf{S})^T \cdot \mathbf{d})(j\omega) \quad (38c)$$

$$= \sum_{j=1}^n \sigma_j(\mathbf{S}) \mathbf{d}^j(j\omega) \quad (38d)$$

where we have defined the new "disturbance components"

$$\mathbf{d}^j = (\mathbf{v}_j(\mathbf{S})^T \cdot \mathbf{d}) \mathbf{u}_j(\mathbf{S}) \quad (39)$$

(38d) shows that the response to a particular disturbance can be viewed as the sum of responses to the disturbances \mathbf{d}^j passing through the scalar transfer function $\sigma_j(\mathbf{S})$. The magnitude of \mathbf{d}^j depends on the alignment of the disturbance \mathbf{d} with the singular vector $\mathbf{v}_j(\mathbf{S})$. The characteristics of the response (speed) to \mathbf{d}^j depend on $\sigma_j(\mathbf{S})$.

For the controller

$$\mathbf{C}(s) = k(s)\mathbf{H}(s) \quad (40)$$

with integral action in $k(s)$ the approximation

$$\mathbf{S}(j\omega) \approx \frac{1}{k} (\mathbf{GH})^{-1}(j\omega) \quad (41)$$

is valid for small ω . Defining $\ell = n - j + 1$ and using (10), (38d) becomes

$$\mathbf{y}(j\omega) = \sum_{\ell=1}^n \frac{1}{k\sigma_\ell(\mathbf{GH})} \tilde{\mathbf{d}}^\ell \quad (42)$$

where

$$\tilde{\mathbf{d}}^\ell = \mathbf{d}^{n-\ell+1} = (\mathbf{u}_\ell^T(\mathbf{GH}) \cdot \mathbf{d}) \mathbf{v}_\ell(\mathbf{GH}) \quad (43)$$

The magnitude of $\tilde{\mathbf{d}}^\ell$ is given by the component of \mathbf{d} in the direction of

the singular vector $\mathbf{u}_\ell(\mathbf{GH})$ and $\tilde{\mathbf{d}}^\ell$ affects the output along the direction of the singular vector $\mathbf{v}_\ell(\mathbf{GH})$. If the loop transfer matrix \mathbf{GH} has a high gain in this direction (i.e., $\sigma_\ell(\mathbf{GH})$ is large) then the control will be quick and good. If the gain is low the response will be slow and poor. If \mathbf{GH} is ill-conditioned ($\gamma(\mathbf{GH})$ large), the widely different response characteristics for different disturbance components will result in unusual overall system responses. These issues will become clearer from the example at the end of this paper.

Performance Interpretation of the RDG

The process response to a particular disturbance $\mathbf{d} = \mathbf{g}_d \mathbf{z}$ is given by

$$\mathbf{y}(s) = (\mathbf{I} + \mathbf{GC}(s))^{-1} \mathbf{d}(s) \quad (44)$$

Let the controller \mathbf{C} be given by

$$\mathbf{C}(s) = \mathbf{H}(s)\mathbf{K}(s) \quad (45)$$

where $\mathbf{K}(s)$ is diagonal and includes integral action in all channels.

Since at low frequencies $\sigma(\mathbf{GC})(j\omega) \gg 1$, (44) can be approximated by

$$\mathbf{y}(j\omega) \approx (\mathbf{GHK})^{-1} \mathbf{d}(j\omega) \quad (46)$$

and in particular for output y_ℓ

$$y_\ell(j\omega) = \frac{1}{k_\ell(j\omega)} [(\mathbf{GH})^{-1} \mathbf{d}(j\omega)]_\ell \quad (47)$$

Compare $y_\ell(j\omega)$ to the response that would occur if the off-diagonal elements in the system \mathbf{GH} were neglected:

$$\beta_\ell(\mathbf{GH})(j\omega) = \frac{[(\mathbf{GH})^{-1} \mathbf{d}(j\omega)]_\ell}{[((\mathbf{GH})_{\text{diag}})^{-1} \mathbf{d}]_\ell} (j\omega) \quad (48)$$

(48) gives a performance interpretation to the RDG and extends it to frequencies other than zero. More importantly this definition provides a justification for using RDG to evaluate the effect of decouplers \mathbf{H} .

Given:

Binary separation, liquid feed, constant molar flows

Relative volatility $\alpha = 1.5$

No. of theoretical trays $N = 40$

Feed tray location $N_F = 21$

Feed composition $x_F = 0.5$

Product compositions $y_D = 0.99, x_B = 0.01$

External flow rates $F = 1, B = 0.5, D = 0.5$

Computed:

Reflux ratio $L/D = 5.41$

Gains using L and -V as inputs (linearized tray-by-tray model):

$$\begin{bmatrix} \Delta y_d \\ \Delta x_B \end{bmatrix} = \begin{bmatrix} 0.878 & 0.864 \\ 1.082 & 1.096 \end{bmatrix} \begin{bmatrix} \Delta L \\ -\Delta V \end{bmatrix}$$

Table 1. Data for distillation column.

The normalization using $(\mathbf{GH})_{\text{diag}}$ makes the RDG scaling independent which might be viewed as an advantage over γ_d . (In particular, the RDG is the same for any diagonal controller \mathbf{K}). However, contrary to γ_d , a physical interpretation becomes difficult or impossible. If a disturbance does not affect y_ℓ at all one finds $\beta_\ell = \infty$. For example, for a full 2x2 system with

$$\mathbf{d}^T = [0 \quad 1]$$

we find (Stanley et al., 1985)

$$\text{RDG} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \infty \\ \lambda(\mathbf{GH}) \end{bmatrix}$$

where λ is the 1,1 element of the RGA of \mathbf{GH} . (Also note that if \mathbf{GH} were diagonal, then β_1 would be undefined for this specific \mathbf{d}). Consequently, β_ℓ may range in magnitude from $-\infty$ to ∞ , and contrary to $\gamma_d(\mathbf{G})$ the magnitude of β_ℓ by itself is not very informative.

V. Example: LV-Distillation Column

Consider the distillation column in Table 1 with L and $-V$ as manipulated variables and the product compositions y_D and x_B as controlled outputs. The steady state gain matrix is (Skogestad and Morari, 1986b)

$$\mathbf{G} = \begin{bmatrix} 0.878 & 0.804 \\ 1.086 & 1.096 \end{bmatrix} \quad (49)$$

We assume there will be no problems with constraints. We want to study how well the system rejects various disturbances using a diagonal controller $\mathbf{C}(s) = k(s) \cdot \mathbf{I}$. Since we are only concerned about the outputs (y_D and x_B), the scaling does not matter provided the outputs are scaled such that an output of magnitude one is equally "bad" for

	Disturbance z				Setpoint change	
	x_F	F	q_F	ΔV_d	y_{DS}	x_{BS}
d	$\begin{bmatrix} 0.881 \\ 1.119 \end{bmatrix}$	$\begin{bmatrix} 0.394 \\ 0.586 \end{bmatrix}$	$\begin{bmatrix} 0.868 \\ 1.092 \end{bmatrix}$	$\begin{bmatrix} 0.864 \\ 1.096 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
RDG (Eq. (24))	$\begin{bmatrix} -0.05 \\ 1.04 \end{bmatrix}$	$\begin{bmatrix} -6.05 \\ 6.01 \end{bmatrix}$	$\begin{bmatrix} 0.29 \\ 0.72 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 35.1 \\ -\infty \end{bmatrix}$	$\begin{bmatrix} \infty \\ 35.1 \end{bmatrix}$
$\gamma_d(G)$ (Eq. (19))	1.48	11.75	1.09	1.41	110.7	88.5

Table 2. Disturbance measures for distillation example.

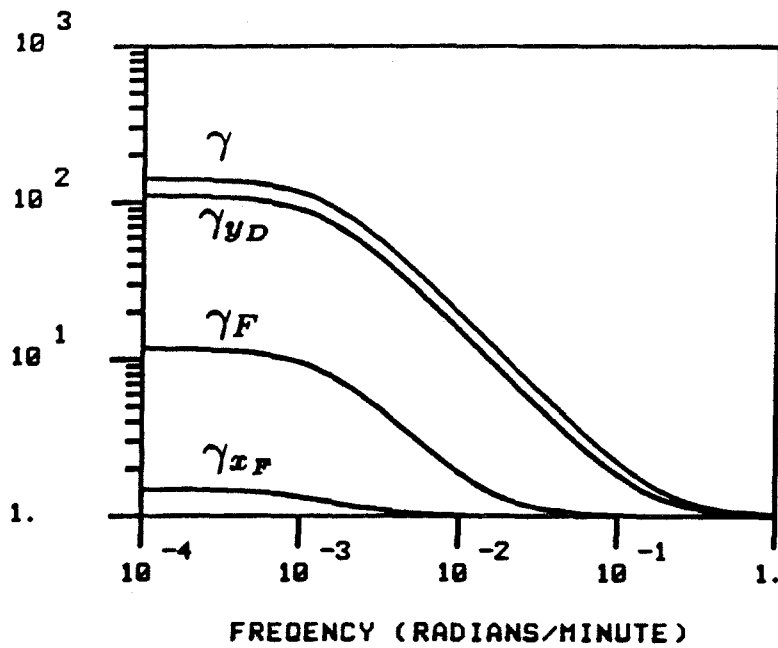


Figure 2. Disturbance condition number of S^{-1} for disturbances in feed rate F , feed composition x_F , and setpoint change in y_D . $C(s) = 0.1/s$ I.

both y_D and x_B . We have

$$\bar{\sigma}(G) = 1.972, \underline{\sigma}(G) = 0.0139, \gamma(G) = 141.7, \lambda_{RGA}(G) = 35.1$$

Consider disturbances z of unit magnitude in feed composition, x_F , feed flow rate, F , feed liquid fraction, q_F , and boilup rate, $-V_d$. The linearized steady state disturbance models are

$$d = g_d = \begin{bmatrix} 0.881 \\ 1.119 \end{bmatrix}, \begin{bmatrix} 0.394 \\ 0.586 \end{bmatrix}, \begin{bmatrix} 0.868 \\ 1.092 \end{bmatrix}, \begin{bmatrix} 0.864 \\ 1.096 \end{bmatrix} \quad (50a)$$

Also consider setpoint changes in y_D and x_B of magnitude one. These are mathematically equivalent to disturbances with

$$d = g_d = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (50b)$$

The steady state values of the RDG, $\beta_d(G)$, and the disturbance condition number, $\gamma_d(G)$, are given for these disturbances in Table 2. The disturbance condition number of S^{-1} , using the controller described below, is shown as a function of frequency in Fig. 2. From these data we see that disturbances in x_F , q_F and V are very well "aligned" with the plant, and there is little need for using a "decoupler" to change the directions of G . The feed flow disturbance is clearly the "worst" disturbance, but even it has its largest effect in the "good" direction.

A "decoupler" is clearly desirable if we want to follow setpoint changes which have a large component in the "bad" direction corresponding to low plant gains. However, a decoupler is not recommended for this distillation column because of severe robustness problems caused by uncertainty (Skogestad & Morari, 1986b). Therefore we cannot expect to get good setpoint tracking for this LV-configuration. Other configurations which are less sensitive to input uncertainty may be better (Skogestad & Morari, 1986b). If setpoint changes are of little or

no interest, the LV-configuration using a diagonal controller may be a good choice. The response to a feed rate disturbances is then expected to be somewhat sluggish because of the high value of $\gamma_d(G)$.

Time Responses

We will now confirm the predictions based on the data in Table 2 by studying some time responses. Assume the plant $G(s)$ has no dynamics, i.e., $G(s)$ is as given in Eq. (49) at all frequencies. This may seem unrealistic, but the dominating dynamics are often similar in all the elements of $G(s)$, and we can assume that these dynamics are exactly compensated for by the dynamics in the controller. This also assumes that the magnitude of the disturbances is small, such that a linear approximation with constant time constants for the column is valid. We use a diagonal controller of the form

$$C(s) = k(s)I$$

where $k(s)$ is a simple integrator with gain $0.1 \text{ (min}^{-1}\text{)}$

$$k(s) = \frac{0.1}{s}$$

(In practice $k(s)$ may be a PI-controller $k(s) = (1+Ts)/s$ with integral time T equal to the time constant of the distillation column).

Time domain simulations are shown for "disturbances" in x_F and F and for setpoint changes in y_D in Fig. 3-5. We have simulated all responses as step setpoint changes of size d (Eq. (50) to make comparisons easier. All simulations are linear, and readers who are concerned about nonphysical values for y_D and x_B may assume, for example, that the deviations, Δy_D and Δx_B , from the initial steady state, are in ppm. Dynamics have not been included in the "disturbances" for x_F and F , which is clearly unrealistic, but this has been done to make the

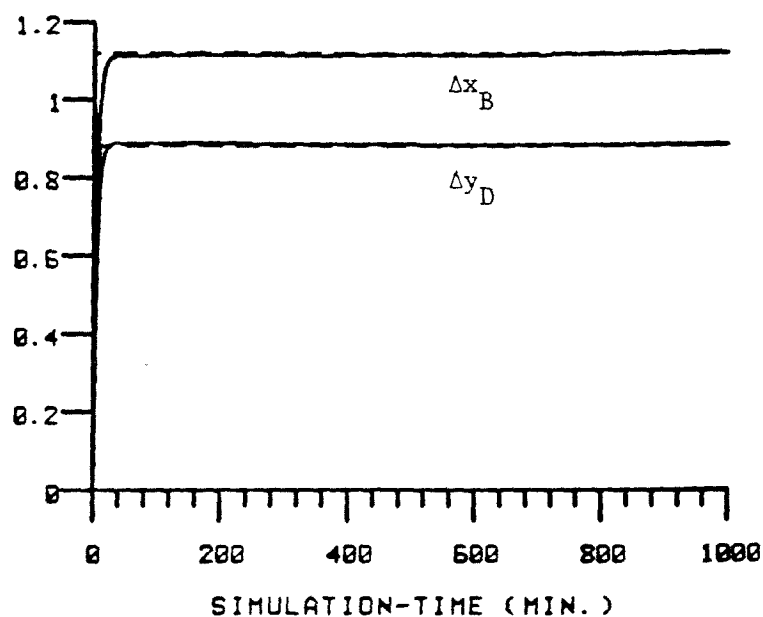


Figure 3. Step change in setpoint = $(0.881, 1.119)^T$.
(Closed loop response to "disturbance" in x_F .)

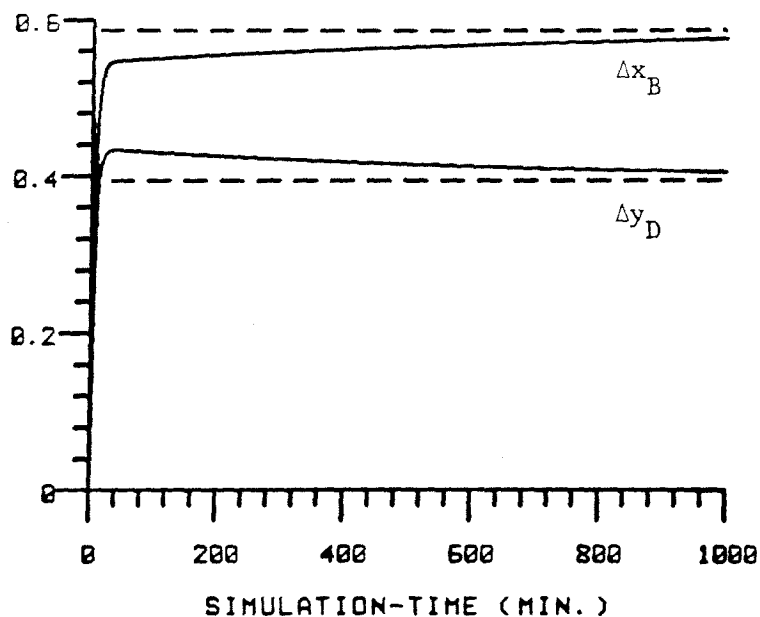


Figure 4. Step change in setpoint = $(0.394, 0.586)^T$.
(Closed loop response to "disturbance" in F .)

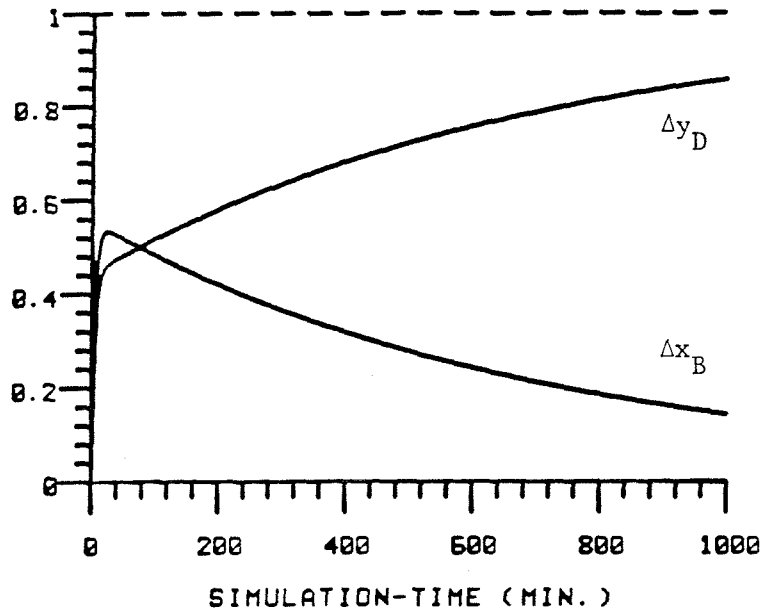


Figure 5. Step change in setpoint = $(1, 0)^T$.
(Closed loop response to setpoint change for y_D .)

	Disturbance		Setpoint change
	x_F	F	y_{DS}
d	$\begin{bmatrix} 0.881 \\ 1.119 \end{bmatrix}$	$\begin{bmatrix} 0.394 \\ 0.586 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
\tilde{d}^1 (Eq. (52a))	$\begin{bmatrix} 1.00 \\ 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.44 \\ 0.44 \end{bmatrix}$
\tilde{d}^2 (Eq. (52b))	$\begin{bmatrix} -0.008 \\ 0.008 \end{bmatrix}$	$\begin{bmatrix} -0.04 \\ 0.04 \end{bmatrix}$	$\begin{bmatrix} 0.05 \\ -0.05 \end{bmatrix}$

Table 3. \tilde{d}^1 and \tilde{d}^2 for distillation example.

example simpler. The time responses confirm what could be predicted based on the disturbance measures in Table 1 with respect to which disturbances are the worst. However, the measures in Table 1 give no direct way of predicting the shape of the responses. The responses are odd-looking and one might almost expect that the system is nonlinear. This is obviously not the case, and the response may in fact be easily explained by decomposing the disturbances along the singular vector directions of the closed loop system, as shown before. For each disturbance, the closed loop frequency response at low frequencies can be approximated by

$$\mathbf{y}(j\omega) \approx \frac{1}{k} \mathbf{G}^{-1} \mathbf{d}(j\omega)$$

By decomposing \mathbf{d} along the "directions" of \mathbf{G} as in Eqs. (42) and (43), we may write this response as the sum of two SISO responses

$$\mathbf{y}(j\omega) \approx \left[\frac{1}{k\bar{\sigma}(\mathbf{G})} \tilde{\mathbf{d}}^1 + \frac{1}{k\underline{\sigma}(\mathbf{G})} \tilde{\mathbf{d}}^2 \right] \quad (51)$$

where

$$\tilde{\mathbf{d}}^1 = (\bar{\mathbf{u}}^T \cdot \mathbf{d}) \bar{\mathbf{v}}(\mathbf{G}) \quad (52a)$$

$$\tilde{\mathbf{d}}^2 = (\underline{\mathbf{u}}^T \cdot \mathbf{d}) \underline{\mathbf{v}}(\mathbf{G}) \quad (52b)$$

Thus, each disturbance response will consist of two responses: one fast in the direction $\tilde{\mathbf{d}}^1$ and one slow in the direction of $\tilde{\mathbf{d}}^2$. The singular value decomposition $\mathbf{G} = \mathbf{U} \Sigma \mathbf{V}^H$ gives

$$\Sigma = \begin{bmatrix} \bar{\sigma} & 0 \\ 0 & \underline{\sigma} \end{bmatrix} = \begin{bmatrix} 1.972 & 0 \\ 0 & 0.01391 \end{bmatrix}$$

$$\mathbf{U} = [\bar{\mathbf{u}} \quad \underline{\mathbf{u}}] = \begin{bmatrix} -0.625 & -0.781 \\ -0.781 & 0.625 \end{bmatrix}$$

$$\mathbf{V} = [\bar{\mathbf{v}} \quad \underline{\mathbf{v}}] = \begin{bmatrix} -0.707 & -0.708 \\ -0.708 & 0.797 \end{bmatrix}$$

$\tilde{\mathbf{d}}^1$ and $\tilde{\mathbf{d}}^2$ are given in Table 3 for the cases simulated in Figs. 3-5.

The decomposition in Eq. (51) and (52) which applies at low frequencies, explains the actual responses very well: Initially there is a very fast response in the direction of $\bar{\mathbf{v}}^T = [-0.707 \ -0.708]$. This response has overall open loop transfer function $k\bar{o}(G) = 0.197/s$ corresponding to a first order response with time constant $1/0.197 = 5.1$ min. Added to this is a slow first order response with time constant $(0.1 \ \underline{g}(G))^{-1} = 720$ min in the direction of $\underline{\mathbf{v}}^T = [-0.708, 0.707]^T$.

Note that slow disturbance component $\tilde{\mathbf{d}}^2$ is the "error" at $t \approx 40$ min, because the fast response has almost settled at this time. As an example consider the disturbance in feed rate F (Fig. 4). At $t \approx 40$ min the deviation from the desired setpoint, $(0.394, 0.586)^T$, is approximately equal to $\tilde{\mathbf{d}}^2 = (-0.04, 0.04)^T$. Similarly, for the setpoint change in y_D (Fig. 5) the deviation from desired setpoint, $(1,0)^T$, at $t \approx 40$ min is approximately equal to $\tilde{\mathbf{d}}^2 = (0.55, -0.55)^T$.

References

- Grosdidier, P., personal communication (1985).
- Morari, M. A General Framework for the Assessment of Dynamic Resilience, Chem. Eng. Sci., 38 11, 1881-1891 (1983).
- Morari, M. and J. D. Doyle, A Unifying Framework for Control System Design under Uncertainty and Its Implications for Chemical Process Control. Third International Conference on Chemical Process Control, Asilomar, CA, January 12-17, 1986.
- Klema, V. C. and A. J. Laub. The Singular Value Decomposition: Its Computation and Some Applications, IEEE Trans. on Auto. Control, AC-25, 2, 164-176 (1980).
- Skogestad, S. and M. Morari. Effect of Model Uncertainty on Dynamic Resilience, Chemical Engineering Science (1986 a).
- Skogestad, S. and M. Morari. Understanding Distillation (in preparation, 1986b).
- Stanley, G., M. Marino-Galarraga and T. J. McAvoy. Shortcut Operability Analysis. 1. The Relative Disturbance Gain, Ind. Eng. Chem. Process Des. Dev., 24 4, 1181-1188 (1985).

Chapter IV

ROBUST CONTROL OF ILL-CONDITIONED PLANTS: HIGH-PURITY DISTILLATION

ROBUST CONTROL OF ILL-CONDITIONED PLANTS: HIGH-PURITY DISTILLATION

Sigurd Skogestad
Manfred Morari
John C. Doyle
Chemical Engineering, 206-41
California Institute of Technology
Pasadena, CA 91125
(818)356-4186

Submitted to IEEE Trans. Autom. Control
November 1986

Abstract

Ill-conditioned plants are generally believed to be difficult to control. Using a high-purity distillation column as an example, the physical reason for the poor conditioning and its implications on control system design and performance are explained. It is shown that an acceptable performance/robustness trade-off cannot be obtained by simple loop-shaping techniques (via singular values) and that a good understanding of the model uncertainty is essential for robust control system design. Physically motivated uncertainty descriptions (actuator uncertainty, nonlinearities) are translated into the H_∞ /Structured Singular Value framework, which is demonstrated to be a powerful tool to analyze and understand the complex phenomena. For the particular example the most effective solution to the control problem turns out to be an alternate choice of manipulated variables (inputs) for which very simple controllers yield robust high quality performance.

I. INTRODUCTION

It is well known that ill-conditioned plants cause control problems (Morari and Doyle, 1986, Skogestad and Morari, 1985). By ill-conditioned we mean that the gain of the plant is strongly dependent on the input direction, or equivalently that the plant has a high condition number

$$\gamma(G(j\omega)) = \bar{\sigma}(G(j\omega)) / \underline{\sigma}(G(j\omega)) \quad (1)$$

Here $\bar{\sigma}(G)$ and $\underline{\sigma}(G)$ denote the maximum and minimum singular values of the plant

$$\bar{\sigma}(G) = \max_{u \neq 0} \frac{\|Gu\|_2}{\|u\|_2}$$
$$\underline{\sigma}(G) = \min_{u \neq 0} \frac{\|Gu\|_2}{\|u\|_2}$$

$\|\cdot\|_2$ denotes the usual Euclidian norm. We also say that an ill-conditioned plant is characterized by strong "directionality" because inputs in directions corresponding to high plant gains are strongly amplified in the plant, while inputs in directions corresponding to low plant gains are not.

The main reason for the control problems associated with ill-conditioned plants is "uncertainty". Uncertainty in the plant model may have several origins:

1. There are always parameters in the linear model which are known only approximately. For the distillation column such parameters may be the relative volatility or the number of theoretical stages.
2. Measurement devices have imperfections. This may give rise to uncertainty on the manipulated inputs in a distillation column, since they are usually measured and adjusted in a cascade manner.
3. At high frequencies even the structure and the model order is

unknown, and the uncertainty will exceed 100% at some frequency.

4. The parameters in the linear model may vary due to nonlinearities or changes in the operating conditions. Examples of this are given in Section V.

For "tight control" of ill-conditioned plants the controller has to compensate for the strong directionality by applying large input signals in the directions where the plant gain is low, that is, a controller similar to G^{-1} in directionality is desirable. However, because of uncertainty, the direction of the large input may not correspond exactly to the low gain in the plant and the amplification of these large input signals may be much larger than expected from the model. This will result in large values of the controlled variables y (Fig. 1), leading to poor performance or even instability.

The concept of directionality is clearly unique to multivariable systems, and extensions of design methods developed for SISO systems are likely to fail for multivariable plants with a high degree of directionality. Furthermore, since the problems with ill-conditioned plants are closely related to how the uncertainty affects the particular plant, it is very important to model the uncertainty as precisely as possible. Most multivariable design methods (LQG, LQG/LTR, INA/DNA, IMC, etc.) do not explicitly take the uncertainty description into account, and these methods will in general not give acceptable designs for ill-conditioned plants.

A distillation column will be used as an example of an ill-conditioned plant. Here the product compositions are very sensitive to changes in the external flows (high gain in this direction), but quite insensitive to

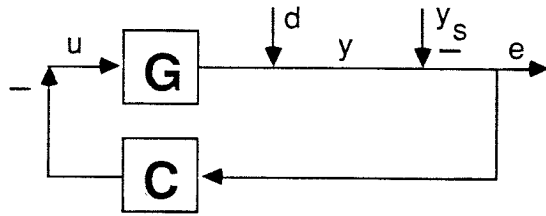


Figure 1. Classical linear feedback structure with error e as input to the controller. d represents the effect of the disturbance on the outputs y .

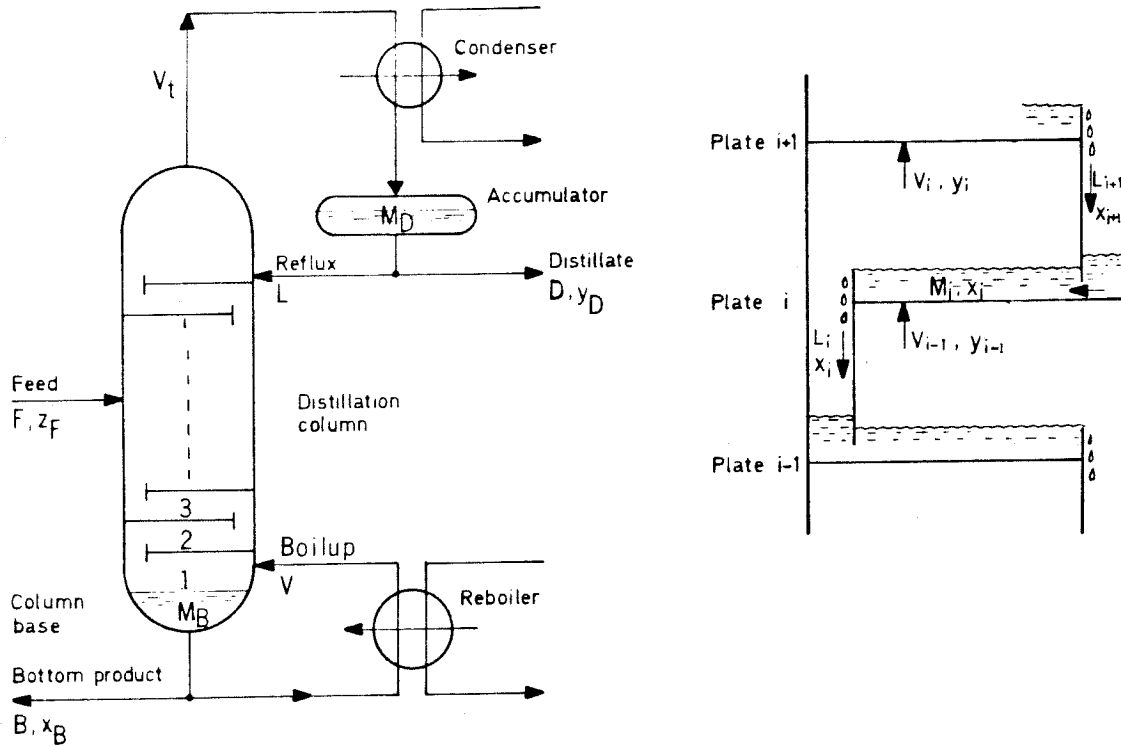


Figure 2. Two product distillation column with single feed and total condenser. Details are shown of the flows and holdups on a plate.

changes in the internal flows (low gain in this direction). Distillation columns are a major consumer of energy in the chemical industry, and there is a large potential for savings by maintaining tighter control of the product compositions. One interesting property of distillation columns is that the condition number may be arbitrary large if the purity of the products is sufficiently high. In this paper the main emphasis is on general properties of ill-conditioned plants, rather than on the control system design for a real distillation column.

II. DISTILLATION COLUMN EXAMPLE

Fundamentals of Distillation Control

The objective of a distillation column (Fig. 2) is to split the feed, F , which is a mixture of a light and a heavy component, into a distillate product, D , which contains most of the light component, and a bottom product, B , which contains most of the heavy component. The compositions z_F , y_D and x_B of these streams refer to the mole fractions of light component. Perfect separation would be obtained with $y_D = 1$ and $x_B = 0$. The driving force for this separation is the difference in volatility between the heavy (H) and light (L) component, which can be expressed by the relative volatility

$$\alpha = \frac{y_L/x_L}{y_H/x_H}$$

x - mole fraction in liquid

y - mole fraction in vapor in equilibrium with x

For a binary separation $y_H = 1 - y_L$ and $x_H = 1 - x_L$, and we get

$$y = \frac{\alpha x}{1 + (\alpha - 1)x} \quad (2)$$

(the subscript L is generally dropped for the light component). In a distillation column separation is improved over what can be obtained with one stage (Eq. 2), by stacking stages on top of each other as shown in Fig. 2.

In such a distillation column there are five controlled variables

- Vapor holdup (expressed by the pressure p)
- Liquid holdup in accumulator (M_D)
- Liquid holdup in column base (M_B)
- Top composition (y_D)
- Bottom composition (x_B)

and five manipulated inputs

- Distillate flow (D)
- Bottom flow (B)
- Reflux (L)
- Boilup (V) (controlled indirectly by the reboiler duty)
- Overhead vapor (V_t) (controlled indirectly by the condenser duty)

Because the composition dynamics are usually much slower than the flow dynamics, we will make the simplifying assumption of perfect control of holdup (i.e., p , M_D , M_B constant) and instantaneous flow responses. With these assumptions and using the mole fractions of the light component at each stage as state variables, we easily derive the nonlinear model shown in the Appendix. Different control configurations are obtained by choosing different inputs pairs (e.g., L and V) for composition control; the remaining three manipulated inputs are then determined by the requirement of keeping p , M_D and M_B under perfect control. Irrespective of the control configuration, the two operating variables corresponding to the high and low plant gain are, as we shall see, the external flows (product flow

rates, D and B) and the internal flows (which are changed by changing the reflux L and boilup V while keeping D and B constant).

Model of the Distillation Column

The distillation column described in Table 1 will be used as an example. The overhead composition is to be controlled at $y_D = 0.99$ and the bottom composition at $x_B = 0.01$. Consider first using reflux L and boilup V as manipulated variables for composition control, i.e.,

$$y = \begin{bmatrix} \Delta y_D \\ \Delta x_B \end{bmatrix}, u = \begin{bmatrix} \Delta L \\ \Delta V \end{bmatrix}$$

This choice is often made since L and V have an immediate effect on the product compositions y_D and x_B , respectively. By linearizing the steady state model and assuming that the dynamics may be approximated by a first order response with time constant $\tau = 75$ min, we derive the following linear model

$$\begin{bmatrix} dy_D \\ dx_D \end{bmatrix} = G_{LV} \begin{bmatrix} dL \\ dV \end{bmatrix}, \quad G_{LV} = \frac{1}{\tau s + 1} \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \quad (3)$$

This is admittedly a very crude model of this strongly nonlinear plant, but the model is simple and displays the main features of the distillation column behavior. The use of a low order model for this high order plant turns out to be a good approximation, since one time constant is usually dominating (Moczek, et al., 1965). In Section V we will consider the nonlinearities in more detail, and discuss how these may be treated as uncertainty on the linear model (3).

Singular Value Analysis of the Model

The condition number of the plant (3) is

$$\gamma(G_{LV}) = 141.7$$

Binary separation, constant molar flows

Relative volatility	$\alpha = 1.5$
No. of theoretical trays	$N = 50$
Feed tray location	$N_F = 21$
Feed rate and composition	$F = 1 \text{ kmol/min}, z_F = 0.5$
Fraction of liquid in feed	$q_F = 1.0$
Product compositions	$y_D = 0.99, x_B = 0.01$
Product rates	$D = B = 0.5 \text{ kmol/min}$
Tray holdup	$M_i = 0.5 \text{ kmol}, i = 2, 40$
Accumulator and column base holdup	$M_D = 32.1 \text{ kmol}, M_B = 11.1 \text{ kmol}$
<u>Computed at steady state from nonlinear model (Appendix)</u>	
Reflux rate	$L = 2.71 \text{ kmol/min} (1.39 \text{ L}_{\min})$
Boilup rate	$V = 3.21 \text{ kmol/min}$
Linearized steady state gains	

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \begin{bmatrix} dL \\ dV \end{bmatrix} + \begin{bmatrix} 0.394 \\ 0.586 \end{bmatrix} dF + \begin{bmatrix} 0.881 \\ 1.119 \end{bmatrix} dz_F$$

Table 1. Data for distillation column example.

which shows a high degree of directionality in the plant. More specific information about this directionality is obtained from the Singular Value Decomposition (SVD) of the steady state gain matrix

$$G = U \Sigma V^H$$

or equivalently since $V^H = V^{-1}$

$$G\underline{v} = \bar{\sigma}(G)\underline{u}$$

$$G\underline{v} = \underline{\sigma}(G)\underline{u}$$

where

$$\Sigma = \text{diag}(\bar{\sigma}, \underline{\sigma}) = \text{diag}(1.972, 0.0139)$$

$$V = [\underline{v} \quad \underline{v}] = \begin{bmatrix} 0.707 & 0.708 \\ -0.708 & 0.707 \end{bmatrix} \quad U = [\underline{u} \quad \underline{u}] = \begin{bmatrix} 0.625 & 0.781 \\ 0.781 & -0.625 \end{bmatrix}$$

The large plant gain, $\bar{\sigma}(G) = 1.972$, is obtained when the inputs are in the direction $\begin{bmatrix} dL \\ dV \end{bmatrix} = \underline{v} = \begin{bmatrix} 0.707 \\ -0.708 \end{bmatrix}$. Since

$$dB = -dD = dL - dV \quad (4)$$

this physically corresponds to the direction with the largest change in the external flows, D and B. From the direction of the output vector $\underline{u} =$

$\begin{bmatrix} 0.625 \\ 0.781 \end{bmatrix}$, we see that changes in the external flows move the outputs in the same direction, i.e., mainly affect the average composition $\frac{y_D + x_B}{2}$.

Any column with products of high purity is sensitive to changes in the external flows because the distillate rate D has to be about equal to the amount of light component in the feed. Any imbalance leads to large changes in the product compositions. Assume in our example that the distillate flow D is increased by 5% to 0.525 kmol/min. Since there is only 0.5 kmol/min of light component in the feed at least 0.025 kmol/min

of this has to be heavy component. The best attainable value for the top composition, even with total reflux, is then $y_D = 0.5/0.525 = 0.952$. This is far from the desired $y_D = 0.99$.

The low plant gain, $\underline{\sigma}(G) = 0.0139$, is obtained for inputs in the direction $\begin{bmatrix} dL \\ dV \end{bmatrix} = \underline{v} = \begin{bmatrix} 0.708 \\ 0.707 \end{bmatrix}$. From (4) observe that physically this corresponds to changing the internal flow only ($dB = dD \approx 0$), and from the output vector $\underline{u} = \begin{bmatrix} 0.781 \\ -0.625 \end{bmatrix}$, we see that the effect is to move the outputs in different directions, i.e., to change $y_d - x_B$. Thus, it takes a large control action to move the compositions in different directions and to make both products purer simultaneously.

The notion that some changes are more "difficult" than others is important, since it implies that some disturbances may be "easier" to reject than others. Let d be the effect of the disturbance on the outputs (Fig. 1), or let d represent a setpoint change. A disturbance d which has a direction close to \bar{u} , is expected to be "easy" to reject since it corresponds to the high plant gain. Similarly, a disturbance close to \underline{u} in direction is expected to be more difficult. The disturbance condition number, $\gamma_d(G)$, gives a more precise measure of how the disturbance is "aligned" with the directions of the plant (Skogestad and Morari, 1986a).

$$\gamma_d(G) = \frac{\|G^{-1}d\|_2}{\|d\|_2} \bar{\sigma}(G) \quad (5)$$

$\gamma_d(G)$ ranges in value between 1 and $\gamma(G)$. A value close to 1 indicates that the disturbance is in the "good" direction (\bar{u}) corresponding to the high plant gain, $\bar{\sigma}(G)$. A value close to $\gamma(G)$ indicates that the disturbance is in the "bad" direction (\underline{u}) corresponding to the low plant gain, $\underline{\sigma}(G)$. We

will consider the following two disturbances (actually setpoint changes) in the simulations

$$y_{S_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ with } \gamma_{d_1}(G) = 110.5$$

$$y_{S_2} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \text{ with } \gamma_{d_2}(G) = 12.3$$

y_{S_1} corresponds to a setpoint change in y_D only, and is seen to be a change with a large component in the "bad" direction. The direction of y_{S_2} corresponds to that of a feed flow rate disturbance (Table 1) and it is seen to have a large component in the "good" direction corresponding to the high plant gain.

Linear Closed Loop Simulations

Linear simulations of the distillation column using the model (3) will now be used to support the following three claims regarding ill-conditioned plants:

1. Inverse-based controllers are potentially very sensitive to uncertainty on the input variables.
2. Low condition-number controllers are less sensitive to uncertainty, but the response is strongly dependent on the disturbance direction.
3. Changing the plant may make the plant insensitive to uncertainty on the input variables.

1. Inverse-based controllers are potentially very sensitive to uncertainty on the input variables

The inverse-based controller

$$C_1(s) = \frac{k_1}{s} G_{LV}^{-1}(s) = \frac{k_1(1+75s)}{s} \begin{bmatrix} 39.942 & -31.487 \\ 39.432 & -31.997 \end{bmatrix}, k_1 = 0.7 \text{ min}^{-1} \quad (6)$$

may be derived by using the IMC design procedure with a first order filter

(Morari et al., 1987) or by using a steady state decoupler plus a PI controller. This controller should in theory remove all the directionality of the plant and give rise to a decoupled first order response with time constant 1.43 min. This is indeed confirmed by the simulations in Fig. 3A and Fig. 4A for the case with no uncertainty. In practice, the plant is different from the model, and for the simulations in Fig. 3B and 4B an error of 20% in the change of each of the manipulated inputs is assumed:

$$dL = 1.2 dL_C, dV = 0.8 dV_C \quad (7)$$

(dL and dV are the actual changes in the manipulated flow rates, while dL_C and dV_C are the desired values as computed by the controller). It is important to stress that this kind of diagonal input uncertainty, which stems from the inability to know the exact values of the manipulated variables, is always present, although the actual size of the uncertainty may vary. For the setpoint change in y_D (Fig. 3B) the simulated response with uncertainty differs drastically from the one predicted by the model, and the response is clearly not acceptable. The response is no longer decoupled, and Δy_D and Δx_B reach a value of about 6 before settling at their desired values of 1 and 0. The uncertainty has less deteriorating effect for the feed rate "disturbance" (Fig. 4B) which occurs mostly in the "good" direction.

There is a simple physical explanation for the observed poor response to the setpoint change in y_D . To accomplish this change, which occurs mostly in the "bad" direction corresponding to the low plant gains, the inverse-based controller generates a large change in internal flows ($dL + dV$), while trying to keep the changes in the external flows ($dB = -dD = dL - dV$) very small. However, uncertainty with respect to the values of dL and dV

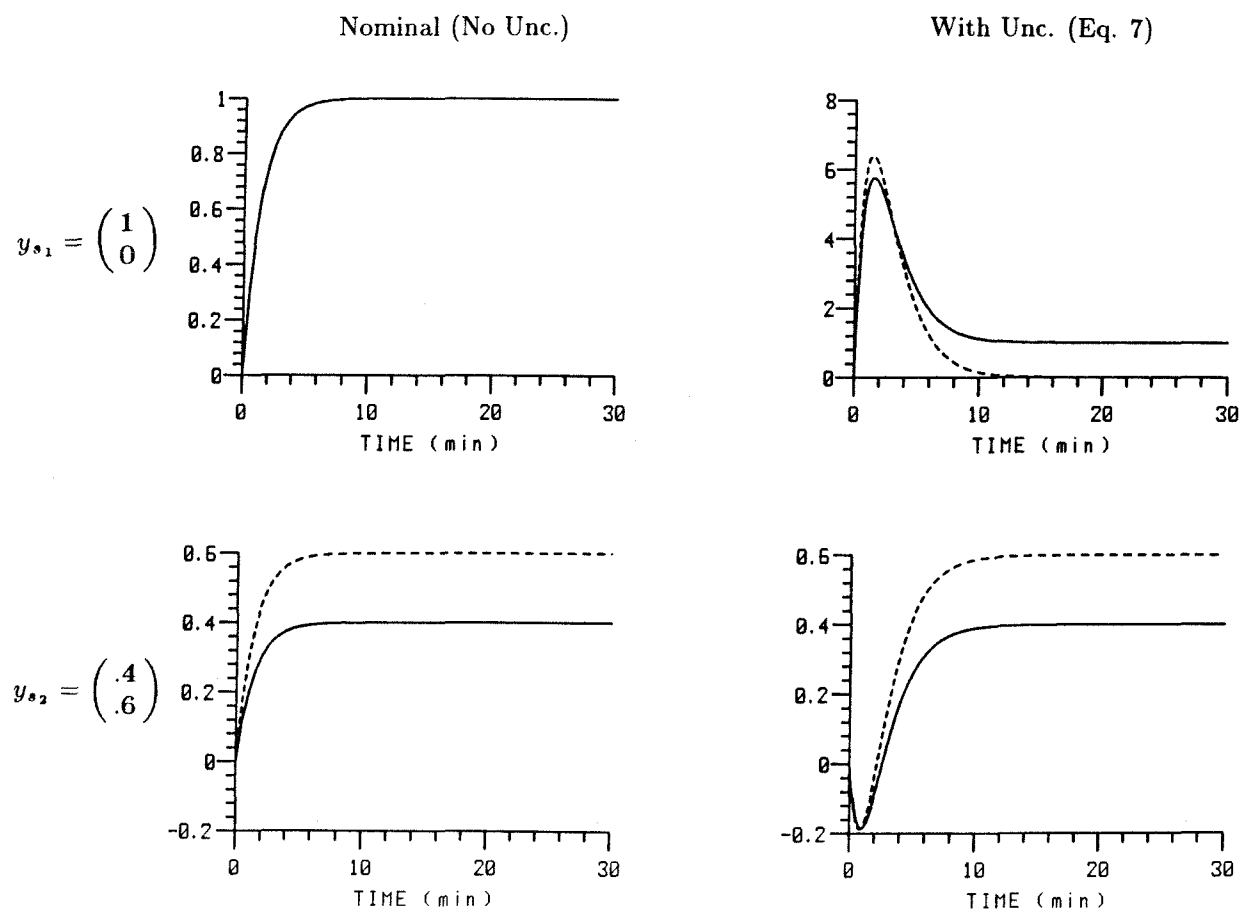


Figure 3 and 4. Closed loop responses Δy_D and Δx_B with inverse-based controller, $C_1(s)$, $k_1 = 0.7$ (time in minutes).

makes it impossible to keep $dL-dV$ small and the consequence is large changes in the external flows. This results in large changes in the compositions because of the high plant gain in this direction. This may be avoided by controlling D or B directly as shown below.

A more mathematical way of showing how the uncertainty changes the plant is as follows: Let the plant transfer model be $G(s)$ and let Δ_1 and Δ_2 represent the relative uncertainty for each manipulated input. Then the actual ("perturbed") plant is

$$G_p = G(I+\Delta) , \Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$$

With an inverse-based controller, $C(s) = c(s)G(s)^{-1}$, the loop transfer matrix becomes

$$G_p C = c(s)G(I+\Delta)G^{-1} = c(s)(I+G\Delta G^{-1}) \quad (8)$$

The error term

$$G_{LV}\Delta G_{LV}^{-1} = \begin{bmatrix} 35.1 \Delta_1 - 34.1 \Delta_2 & -27.7 \Delta_1 + 27.7 \Delta_2 \\ 43.2 \Delta_1 - 43.2 \Delta_2 & -34.1 \Delta_1 + 35.1 \Delta_2 \end{bmatrix} \quad (9)$$

is worst when Δ_1 and Δ_2 have different signs. With $\Delta_1 = 0.2$ and $\Delta_2 = -0.2$ (as used in the simulations, Eq. (7)) we find

$$G_{LV}\Delta G_{LV}^{-1} = \begin{bmatrix} 13.8 & -11.1 \\ 17.2 & -13.8 \end{bmatrix}$$

The elements in this matrix are much larger than one, and the observed poor response is not surprising.

2. Low condition number controllers are less sensitive to uncertainty, but the response is strongly dependent on the disturbance direction.

The poor response for the case with uncertainty in the example above was caused by the high condition-number controller which generates large

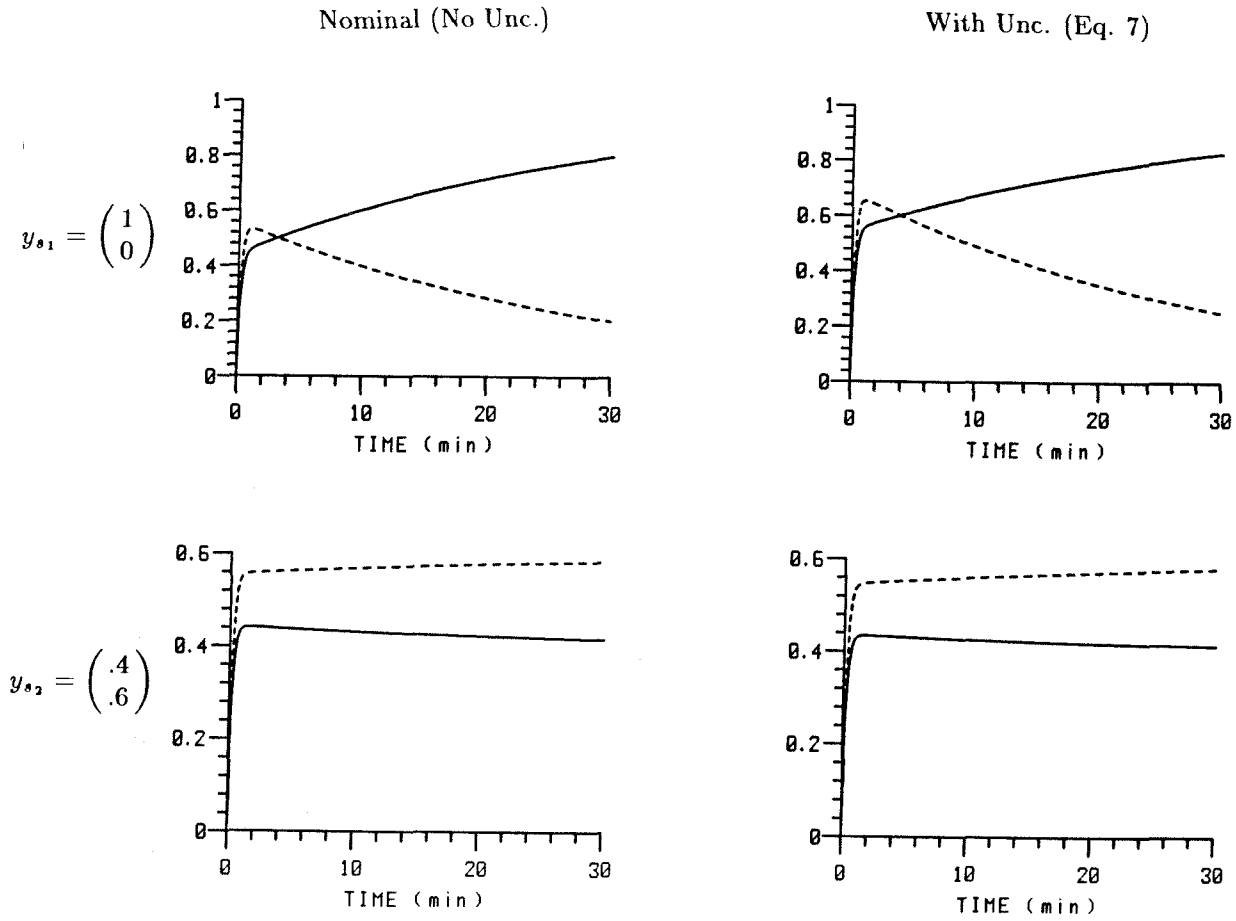


Figure 5 and 6. Closed loop responses Δy_D and Δx_B with diagonal controller, $C_2(s)$, $k_2 = 2.4$ (time in minutes).

input signals in the directions corresponding to the small plant gain. The simplest way to make the closed loop system less sensitive to the input uncertainty is to use a low condition number controller which does not have large gains in any particular direction. The problem with such a controller is that little or no correction is made for the strong directionality of the plant. This results in a closed loop response which depends strongly on the disturbance direction, as shown below. The diagonal controller

$$C_2(s) = \frac{k_2(75s+1)}{s} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad k_2 = 2.4 \text{ min}^{-1} \quad (10)$$

consists of two equal single loop PI controllers and has a condition number of one. As seen from the simulations in Fig. 5 and 6 the quality of the closed loop response depends strongly on the disturbance direction, but is only weakly influenced by uncertainty. The response to y_{S_1} is very sluggish, while the response to y_{S_2} is fast initially, but approaches the final steady state sluggishly. Note that a disturbance in the "good" direction

$$y_S = \underline{u} = \begin{bmatrix} 0.625 \\ 0.781 \end{bmatrix} \quad \text{with} \quad \gamma_d(G)=1$$

generates a first order nominal response with time constant $1/2.4 \cdot \bar{\sigma}(G) = 0.21 \text{ min}$. A disturbance in the "bad" direction

$$y_S = \underline{u} = \begin{bmatrix} 0.781 \\ -0.625 \end{bmatrix} \quad \text{with} \quad \gamma_d(G)=141.7$$

generates a first order nominal response with time constant $1/2.4 \cdot \underline{\sigma}(G) = 30 \text{ min}$. All other responses are linear combinations of these two extremes (Fig.5A and 6A).

3. Changing the plant may make the system insensitive to uncertainty on

the input variables

We already argued physically that the plant might be made less sensitive to uncertainty by controlling the external flows directly. Consider the case of distillate flow D and boilup V as manipulated variables ("direct material balance control") (Shinskey, 1984). Assuming perfect level and pressure control, i.e., $dL = dV - dD$, we have

$$\begin{bmatrix} dL \\ dV \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dD \\ dV \end{bmatrix} \quad (11)$$

and the following linear model is derived from (3)

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = G_{DV} \begin{bmatrix} dD \\ dV \end{bmatrix}$$

$$G_{DV} = G_{LV} \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} = \frac{1}{1+75s} \begin{bmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{bmatrix} \quad (12)$$

In practice the condenser level loop introduces a lag between the change in distillate flow, dD , and the reflux flow, dL (which is the input which actually affects the compositions), but this is neglected here. It is important to note that with (11) and without input uncertainty, identical responses may be obtained with the LV-plant (3) and the DV-plant (12) by using multivariable controllers. The plant (12) is also ill-conditioned; $\gamma(G_{DV}) = 70.8$. In this case the SVD yields

$$\Sigma = \begin{bmatrix} 1.393 & 0 \\ 0 & 0.0197 \end{bmatrix}, \quad V = \begin{bmatrix} -1.000 & -0.001 \\ -0.001 & 1.000 \end{bmatrix}, \quad U = \begin{bmatrix} 0.630 & 0.777 \\ 0.777 & -0.630 \end{bmatrix}$$

The high gain corresponds to an input $\begin{bmatrix} dD \\ dV \end{bmatrix}$ in the direction of $\nabla(G_{DV}) =$

$$\begin{bmatrix} -1.000 \\ -0.001 \end{bmatrix}, \text{ which, as expected, corresponds to a change in the external}$$

flows. The low gain again corresponds to a change in the internal flows ($dD \approx 0$). Note that in this case there is one manipulated variable (dD) which

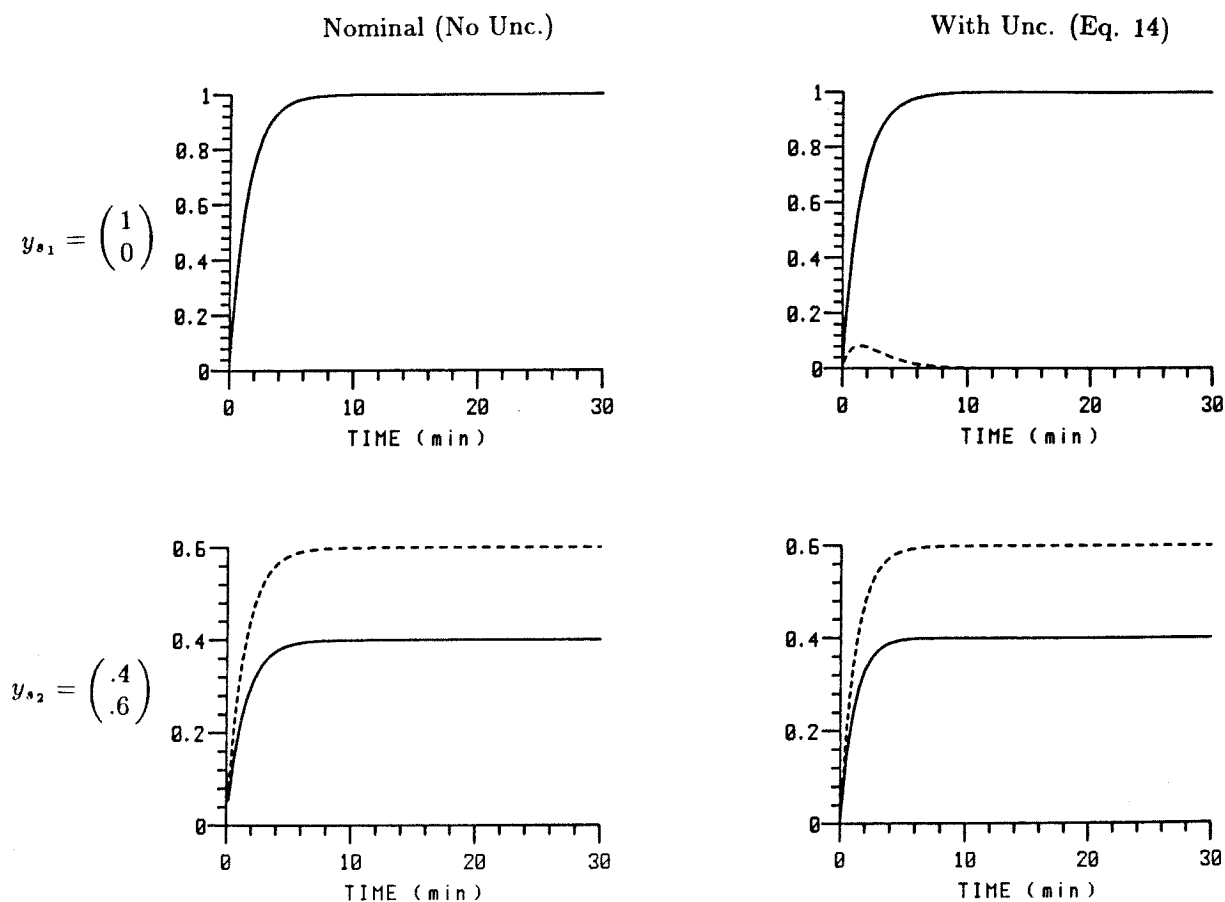


Figure 7 and 8. DV-configuration. Closed loop responses Δy_D and Δx_B with inverse-based controller, $C_s(s)$, $k_s = 0.7$ (time in minutes).

acts in the high gain direction, and another (dV) which acts in the low gain direction. This "decomposition" is significant, since uncertainty in dV, does not affect the external flows, dD.

To confirm that the system is much less sensitive to uncertainty in this case, consider the following inverse-based controller

$$C_3(s) = \frac{k_3}{s} G_{DV}(s)^{-1} = \frac{k_3(1+75s)}{s} \begin{bmatrix} -0.5102 & -0.5102 \\ 39.43 & -32.00 \end{bmatrix}, k_3 = 0.7 \text{ min}^{-1} \quad (13)$$

Without uncertainty this controller gives the same response as controller $C_1(s)$ with the LV-configuration. However, in this case the decoupled first order response with time constant 1.43 min is maintained also when there is uncertainty on the manipulated variables (Fig. 7 and 8). The following error with respect to dD and dV was used

$$dD = 1.2 dD_C, dV = 0.8 dV_C \quad (14)$$

From this example we see that ill-conditioned plants by themselves may not give performance problems if the uncertainty is appropriately aligned with the process. For the DV-configuration we find the error matrix $G\Delta G^{-1}$ in Eq. (8)

$$G_{DV\Delta G}^{-1}_{DV} = \begin{bmatrix} 0.45 \Delta_1 + 0.55 \Delta_2 & 0.45 \Delta_1 - 0.45 \Delta_2 \\ 0.55 \Delta_1 - 0.55 \Delta_2 & 0.55 \Delta_1 + 0.45 \Delta_2 \end{bmatrix}$$

and with $\Delta_1 = 0.2$ and $\Delta_2 = -0.2$ corresponding to (14)

$$G_{DV\Delta G}^{-1}_{DV} = \begin{bmatrix} -0.02 & 0.18 \\ 0.22 & 0.02 \end{bmatrix}$$

The elements in this matrix are small compared to one, and good performance is maintained even in the presence of uncertainty on each input. The nonzero off-diagonal elements explain why the response in Fig. 7B is not completely decoupled.

III. ROBUSTNESS ANALYSIS WITH μ

It is quite evident from the linear simulations above that multivariable systems exhibit a type of "directionality" which may make the closed loop response strongly dependent on the particular disturbance and model error assumed. One of the major weaknesses with the simulation approach is that it may be very difficult and time-consuming to find by trial-and-error the particular disturbance and model error which causes control problems. Therefore there is a need for a tool which solves the following robust performance problem in a more systematic manner:

Given a nominal plant, an uncertainty description, a set of possible disturbances and setpoint changes, a desired performance objective, and a controller: Will the "worst case" response satisfy the desired performance objectives?

If performance (allowed size of input and output signals) is defined using the H_∞ -norm and the uncertainty is described in terms of norm bounds in the frequency domain, this problem is solved fairly easily by computing Structured Singular Value μ (Doyle, 1982) of a particular matrix N at each frequency (Doyle et al., 1982). The elements in the matrix N are determined by the nominal model, the size and nature of the uncertainty, the performance specifications and the controller. Robust performance is guaranteed if and only if $\mu(N) < 1$ for all frequencies.

Some Definitions

Let us make a pause to define some of the terms used above more carefully. Achieving robust performance is clearly the ultimate goal of the controller design. However, it may be easier to solve this problem by first considering some subobjectives which have to be satisfied in order to achieve this:

Nominal Stability (NS): The model is assumed to be a reasonable approximation of the true plant. Therefore the closed loop system with the controller applied to the (nominal) plant model has to be stable.

Nominal Performance (NP): In addition to stability, the quality of the response should satisfy some minimum requirements - at least when the controller is applied to the plant model. For mathematical convenience we will define performance in terms of the weighted H_∞ -norm of the closed-loop transfer function between external inputs (disturbances and setpoints) and "errors" (may include $y_D - y_{DS}$, $x_B - x_{BS}$, manipulated inputs u , etc.). The simplest example of such a performance specification is a bound on the weighted sensitivity operator

$$\bar{\sigma}(W_{1P} S W_{2P}) \leq 1 \quad \forall \omega, S = (I + GC)^{-1} \quad (16a)$$

The input weight W_{2P} is often equal to the disturbance model. The output weight W_{1P} is used to specify the frequency range over which the errors are to be small and (if W_{1P} is not equal to $w_p(s)I$) which outputs are more important.

Robust Stability (RS): The closed loop system must remain stable for all possible plants as defined by the uncertainty description.

Robust Performance (RP): The closed loop system must satisfy the performance requirements for all possible plants as defined by the uncertainty description. As an example we may require (16a) to be satisfied when G is replaced by any of the possible perturbed plants G_p as defined by the uncertainty description.

$$\bar{\sigma}(W_{1P}(I + G_p C)^{-1} W_{2P}) \leq 1 \quad \forall \omega, \quad \forall G_p \quad (16b)$$

Most controller design methods (even "modern" optimal control, LQG), only address the problems of Nominal Stability and Nominal Performance.

The stability margins in the classical frequency domain design methods, are an attempt to address the Robust Stability problem, but these margins may be misleading and are a very indirect method.

Conditions for Robust Stability and Robust Performance

The definition of Robust Performance given above is of no value without simple methods to test if conditions like (16b) are satisfied for all possible perturbed plants G_p . Below we will state computationally useful conditions for RS and RP using the Structured Singular Value μ for the case when the uncertainty (the possible plants G_p) is modelled in terms of a set of norm-bounded perturbations on the nominal system. By use of weights each perturbation is normalized to be of size one:

$$\bar{\sigma}(\Delta_i) \leq 1 \quad \forall \omega$$

The perturbations, which may occur at different locations in the system, are collected in the diagonal matrix Δ

$$\Delta = \text{diag}\{\Delta_1, \dots, \Delta_n\}$$

and the system is rearranged to match the structure in Fig. 9. We will not go into detail on how this is done at this point. This will become clearer by studying the distillation column example in Section IV and V. The signal \hat{d} in Fig. 9 represents the external inputs (weighted disturbances and setpoint changes) affecting the system. The signal \hat{e} represents the weighted errors, or more generally the signals which are to be kept "small". The interconnection matrix N in Fig. 9 is a function of the nominal plant G , the controller C and the uncertainty weights. Performance weights are also absorbed into N such that the performance specifications involving \hat{e} and \hat{d} are normalized:

Robust Performance Specification: (Fig. 9)

$$\bar{\sigma}(E) \leq 1 \quad \forall \omega, \quad \forall \Delta \quad (17)$$

$$\text{where } \hat{e} = E\hat{d}, \quad E = N_{22} + N_{21}\Delta(I - N_{11}\Delta)^{-1}N_{12}$$

An example of such a performance specification is Eq. (16b). With these assumptions for the uncertainty and performance we have the following results (Doyle et al., 1982).

$$N.S. \Leftrightarrow N \text{ stable (internally)} \quad (18)$$

$$N.P. \Leftrightarrow \mu_{NP} = \sup_{\omega} \bar{\sigma}(N_{22}) < 1 \quad (19)$$

$$R.S. \Leftrightarrow \mu_{RS} = \sup_{\omega} \mu_{\Delta}(N_{11}) < 1 \quad (20)$$

$$R.P. \Leftrightarrow \mu_{RP} = \sup_{\omega} \mu(\Delta_{\Delta P})(N) < 1 \quad (21)$$

The quantities μ_{NP} , μ_{RS} and μ_{RP} represent the "μ-norms" and are introduced as a convenient notation. The conditions for N.P. and R.S. are necessary in order to satisfy the R.P. condition. Note that $\mu_{\Delta}(N_{11})$ is a function of both the matrix N_{11} and the structure of the uncertainty Δ . The Robust Performance condition (21) is computed with respect to the structure $\text{diag}(\Delta, \Delta_P)$, where Δ_P is a full matrix of the same size as N_{22} . The use of μ is less conservative than using any other matrix norm. In particular,

$$\mu_{\Delta}(N) \leq \bar{\sigma}(N)$$

and the equality holds only when Δ is a full matrix. The use and implications of conditions (19)-(21) will hopefully become clearer by studying how these results apply to the distillation column example.

IV. μ-ANALYSIS OF THE DISTILLATION COLUMN

Problem Definition

To study Robust Stability and Robust Performance of the distillation column using μ , the uncertainty and performance specifications must be defined. The same uncertainty and performance specifications will be assumed for the LV-configuration (3) and the DV-configuration (12). (In

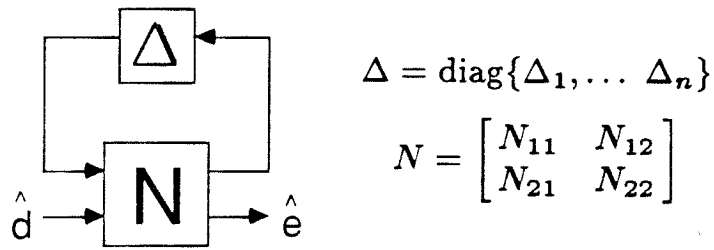


Figure 9. General representation of system with uncertainty Δ . \hat{d} represents weighted external inputs, \hat{e} represents weighted errors.

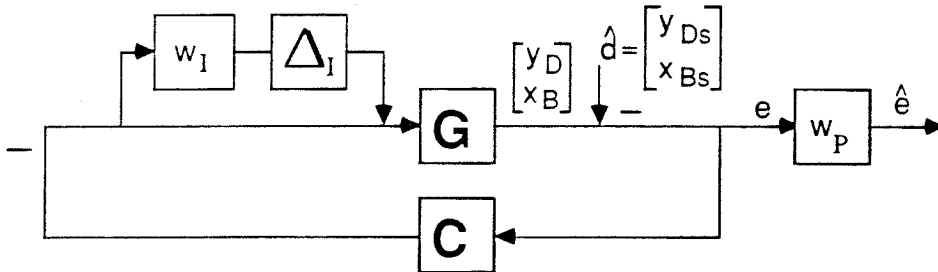


Figure 10. Block diagram of system with input uncertainty and with setpoints as external inputs. Rearranging this system to fit Fig. 9 gives N as in Eq. (27).

general, it is reasonable to use the same performance specifications, but the uncertainty may be different).

Uncertainty: The uncertainty with respect to the manipulated inputs which was discussed in Section II may be represented as multiplicative input uncertainty (Fig. 10)

$$G_p = G(I + w_I(s)\Delta_I), \quad \bar{\sigma}(\Delta_I) < 1 \quad \forall \omega \quad (22)$$

where $w_I(s)$ describes the magnitude of the relative uncertainty on each manipulated input

$$w_I(s) = 0.2 \frac{5s+1}{0.5s+1} \quad (23)$$

This implies an input error of up to 20% in the low frequency range as was assumed for the simulations. The uncertainty increases at high frequencies, reaching a value of one at about $\omega = 1 \text{ min}^{-1}$. The increase at high frequency may take care of the neglected flow dynamics. It allows for a time delay of about 1 minute in the responses between L and V and the outputs y_D and x_B . It may also represent neglected valve dynamics, dynamics for the heat transfer in the reboiler (for V), etc.

At first the uncertainty will be assumed to be unstructured, i.e., the perturbation matrix Δ_I is a full 2x2 matrix. This does not make much sense from a physical point of view, but is done for mathematical convenience. It will turn out that this assumption does not make any difference for the LV-configuration. The set of possible plants is now generated from Eq. (22) by allowing any 2x2 matrix Δ_I which satisfies $\bar{\sigma}(\Delta_I) \leq 1, \quad \forall \omega$.

Performance: We will consider the simple case with setpoint changes in (y_S) as external inputs and $e = y - y_S$ as errors. These signals are related through the sensitivity function

$$e = -S_p y_s, \quad S_p = (I + G_p C)^{-1}$$

y_s and e are related to the weighted external inputs (\hat{d}) and errors (\hat{e}) by

$$y_s = W_{2P} \hat{d}, \quad \hat{e} = W_{1P} e$$

and we have

$$\hat{e} = E \hat{d}, \quad E = -W_{1P} S_p W_{2P} \quad (24)$$

We choose to express the performance specifications through the weights

$$W_{2P} = I, \quad W_{1P} = w_p I, \quad w_p(s) = 0.5 \frac{10s+1}{10s} \quad (25)$$

The Robust Performance specification (17) then becomes

$$\bar{\sigma}(S_p) < 1/|w_p|, \quad \forall \omega \quad (26)$$

This bound on the sensitivity function S_p should be satisfied for all allowable G_p given by (22). The performance weight $w_p(s)$ (25) implies that we require integral action ($w_p(0) = \infty$) and allows an amplification of disturbances at high frequencies of at most a factor of two ($\lim_{\omega \rightarrow \infty} |w_p(i\omega)|^{-1} = 2$). A particular sensitivity function which exactly matches the performance bound (26) at low frequencies and satisfies it easily at high frequencies is $S = \frac{20s}{20s+1} I$. This corresponds to a first order response with time constant 20 min.

Performance and Stability Conditions

With the information given above the matrix N in the ΔN -structure (Fig. 9) becomes

$$N = \begin{bmatrix} -W_I C S G & W_I C S \\ W_P S G & -W_P S \end{bmatrix}, \quad S = (I + G C)^{-1} \quad (27)$$

This matrix is found from Fig. 10 by breaking the loops ($\Delta_I = 0$) at the input and output of the block Δ_I . As an example with $\Delta_I = 0$, the transfer function from the external inputs (\hat{d}) to the errors (\hat{e}) is

$$N_{22} = -w_p(I+GC)^{-1}$$

Similarly, the transfer function from \hat{d} to the input of Δ_I is

$$N_{12} = w_I C(I+GC)^{-1}$$

Conditions for Nominal Performance and Robust Stability are derived from (19) and (20) by using (27)

$$\text{N.P.} \quad \Leftrightarrow \quad \bar{\sigma}(S) \leq 1/|w_p| \quad \forall \omega, \quad S = (I+GC)^{-1} \quad (28)$$

$$\text{R.S.} \quad \Leftrightarrow \quad \bar{\sigma}(H_I) \leq 1/|w_I| \quad \forall \omega, \quad H_I = CG(I+CG)^{-1} = CSG \quad (29)$$

The condition for Robust Stability is expressed in terms of $\bar{\sigma}$ since Δ_I is assumed to be a full matrix. Note that S is the nominal sensitivity function at the output of the plant, while H_I is the closed loop transfer function as seen from the input of the plant. In some cases $GC = CG$ (in particular this is the case for the controllers $C_1(s)$, $C_2(s)$ and $C_3(s)$ in our examples) and we have $H_I = H$, where

$$H = GC(I+GC)^{-1} = I - S$$

is the closed loop transfer function as seen from the output of the plant. However, $H_I = H$ does not generally hold for multivariable systems. The Robust Performance specification (26) should be satisfied for all plants given by (22). From (21) one finds

$$\text{R.P.} \quad \Leftrightarrow \quad \mu(\Delta_{I\Delta_P})(N) \leq 1, \quad \forall \omega \quad (30)$$

Analysis of the LV-Configuration

The set of possible plants is given by (22) with $G = G_{LV}$ (3). We will analyze the LV-configuration for the inverse-based and the diagonal controller.

$$C_1(s) = c_1(s)G_{LV}^{-1}(s) \quad (31)$$

$$C_2(s) = c_2(s) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (32)$$

We will first consider the choices $c_1(s) = \frac{0.7}{s}$ and $c_2(s) = \frac{2.4(1+75s)}{s}$ used in the simulations in Section II, and then let

$$c_1(s) = \frac{k_1}{s}, \quad c_2(s) = \frac{k_2(1+75s)}{s} \quad (33)$$

and see if Robust Performance can be improved with other choices for k_1 and k_2 . Finally, we will consider the " μ -optimal" controller, $C_\mu(s)$, i.e., the controller which minimizes μ_{RP} . We found this controller through a software package which uses a somewhat simplified version of the μ -synthesis procedure described by Doyle (1985). The simplification involves only considering the upper left corner when minimizing the H_∞ -norm of Eq. (7.3) in Doyle's paper (1985). This means that the resulting controller is not necessarily optimal.

Nominal Performance and Robust Stability. One way of designing controllers which meet the N.P. and R.S. specifications is to use multivariable loop shaping (Doyle and Stein, 1981). For Nominal Performance, $\underline{g}(GC)$ must be above $|w_P|$ at low frequencies. For Robust Stability with input uncertainty, $\bar{\sigma}(CG)$ must lie below $1/|w_I|$ at high frequencies (Fig. 11).

For the inverse-based controller (31) we get $\bar{\sigma}(C_1G) = \underline{g}(GC_1) = |c_1|$ and it is trivial to choose a $c_1(s)$ to satisfy these conditions. The choice $c_1(s) = \frac{0.7}{s}$ which was used in the simulations gives a controller which has much better nominal performance than required, and which can allow about two times more uncertainty than assumed. This is also seen from Fig. 12 and 13 where the Nominal Performance and Robust Stability conditions (28) and (29) are displayed graphically.

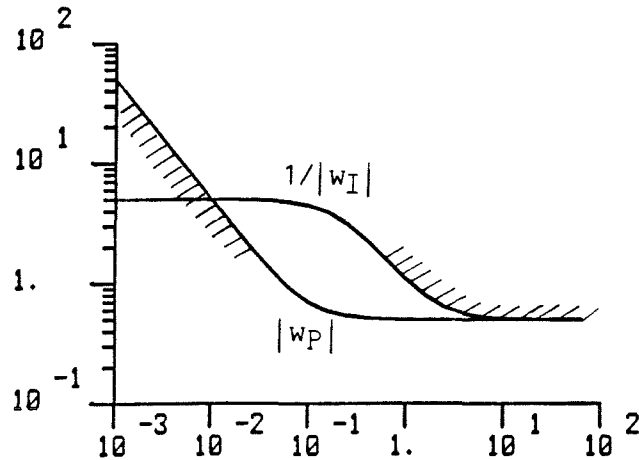


Figure 11. Multivariable loop shaping. For Nominal Performance, $\bar{\sigma}(\text{GC})$ must lie above $|w_P|$ at low frequencies. For Robust Stability with input uncertainty, $\bar{\sigma}(\text{CG})$ must lie below $1/|w_I|$ at high frequencies.

FIG12

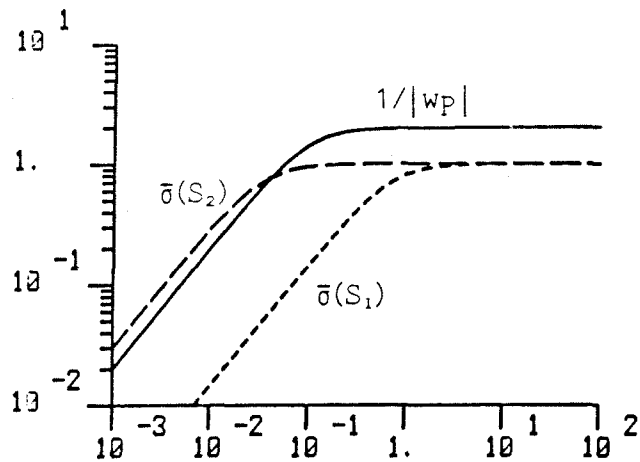


Figure 12. The inverse-based controller, $C_1(s)$, $k_1 = 0.7$ has much better Nominal Performance than required by the condition $\bar{\sigma}(S) \leq 1/|w_P|$, $\forall \omega$. The diagonal controller, $C_2(s)$, $k_2 = 2.4$, does not satisfy the N.P. condition at low frequency.

FIG13

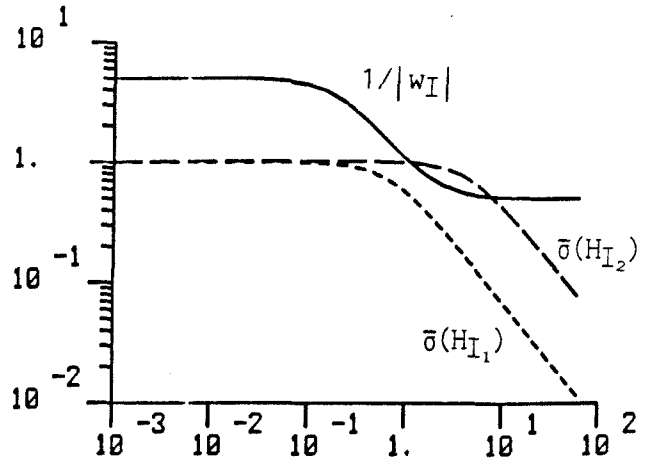


Figure 13. The inverse-based controller, $C_1(s)$, $k_1 = 0.7$ is guaranteed Robust Stability since $\bar{\sigma}(H_I) \leq 1/|w_I|$, $\forall \omega$. The diagonal controller, $C_2(s)$, $k_2 = 2.4$ will give an unstable system for some of the plants defined by (22), since the R.S. condition is not satisfied at all frequencies.

For the diagonal controller (32) we find $\bar{\sigma}(C_2G) = 1.972 |c_2|$ and $\underline{\sigma}(GC_2) = 0.0139 |c_2|$, and the difference between these two singular values is so large that no choice of c_2 is able to satisfy both N.P. and R.S. This is shown graphically in Fig. 12 and 13 for the choice $c_2(s) = \frac{2.4(1+75s)}{s}$.

Robust Performance. In the case with input uncertainty sufficient ("conservative") tests for Robust Performance in terms of singular values are easily derived:

$$R.P. \Leftrightarrow \gamma \cdot \bar{\sigma}(w_P S_I) + \bar{\sigma}(w_I H_I) \leq 1 \quad \forall \omega \quad (34a)$$

$$\text{or } R.P. \Leftrightarrow \bar{\sigma}(w_P S) + \gamma \cdot \bar{\sigma}(w_I H) \leq 1 \quad \forall \omega \quad (34b)$$

$$\text{or } R.P. \Leftrightarrow (1 + \sqrt{\gamma}) \cdot (\bar{\sigma}(w_P S) + \bar{\sigma}(w_I H_I)) \leq 1 \quad \forall \omega \quad (34c)$$

Here γ denotes the condition number of the plant or the controller (the smallest one should be used). These conditions indicate that the use of an ill-conditioned controller ($\gamma(C_1)=141.7$) may give very poor Robust Performance even though both the Nominal Performance ($\bar{\sigma}(w_P S) < 1$) and Robust Stability conditions ($\bar{\sigma}(w_I H_I) < 1$) are individually satisfied. If a controller with a low condition number ($\gamma(C_2)=1$) is used we see that we get R.P. for "free" provided we have satisfied N.P. and R.S. This is always the case for SISO systems and gives a partial explanation for why Robust Performance was never an important issue in the classical control literature. Furthermore, for SISO systems (34a) is necessary and sufficient for R.P.

Conditions (34) are very useful since they directly relate Robust Performance to N.P., R.S. and the condition number. However, (34) may be very conservative and in order to get a "tight" condition for R.P. the μ -condition (30) has to be used with N given by (27). μ for R.P. is plotted in Fig. 14 and 15 for the two controllers $C_1(s)$ and $C_2(s)$ used in the simulations. As expected, the inverse-based controller $C_1(s)$ is far from

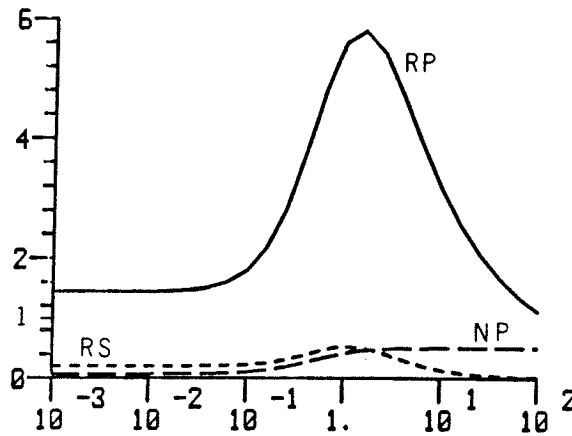


Figure 14.

μ -plots for inverse-based controller, $C_1(s)$, $k_1 = 0.7$. The system has very good performance when the plant is equal to the plant model, and is guaranteed stability for all plants given by (22), but robust performance is poor.

FIG15

FIG16

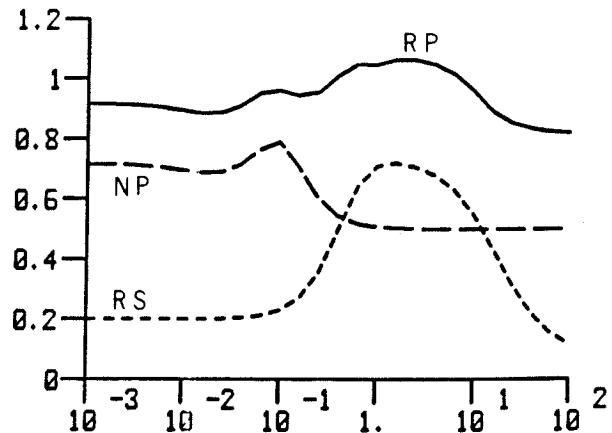
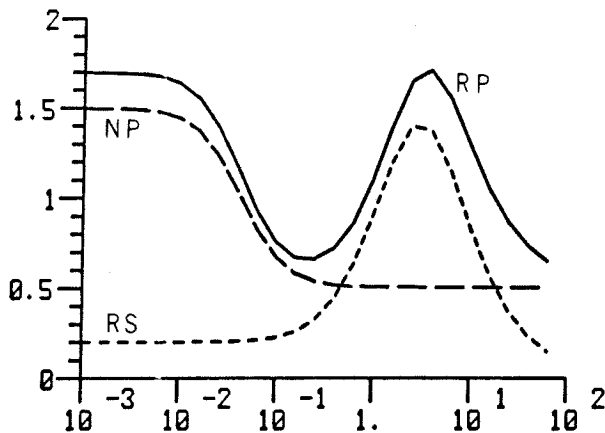


Figure 15.

μ -plots for diagonal controller, $C_2(s)$, $k_2 = 2.4$.

Figure 16.

μ -plots for " μ -optimal" controller, $C_\mu(s)$.

meeting the Robust Performance requirements (μ_{RP} is about 5.8), even though the controller was shown to achieve both N.P. and R.S. On the other hand, the performance of the diagonal controller $C_2(s)$ is much less affected by uncertainty. ($\mu_{RP} = 1.71$).

Optimizing k_1 and k_2 wrt. R.P. For the inverse-based controller the "optimal" value for k_1 is 0.14 corresponding to a value of μ_{RP} equal to 3.3 which still implies poor performance. This value for k_1 seems reasonable since it corresponds to a loop shape k_1/s which is further away from the R.S. constraint in Fig. 11.

For the PI-controller, the optimal gain is $k_2 = 2.4$, which is the value already used ($\mu_{RP} = 1.70$). It is not clear how low μ_{RP} can be made if $C(s)$ is only restricted to be diagonal (decentralized control); we were able to get μ_{RP} down to 1.42 by a trial-and-error procedure.

μ -Optimal Controller. The synthesis method (Doyle, 1985) used to design the " μ -optimal" controller gives controllers of very high order, but by employing model reduction, we were able to find a " μ -optimal" controller with 6 states. μ for R.P. for this controller is shown as a function of frequency in Fig. 16. (The μ -plot is not quite flat as it should be for the optimal case). The peak value for μ is 1.06, which means that this controller "almost" satisfies the Robust Performance condition. This value for μ_{RP} is significantly lower than for the diagonal PI controller, $C_2(s)$, and the time responses are also better as shown in Fig. 17 and 18. In particular, the approach to steady state is much faster. The state space realization of this μ -optimal controller is shown in Fig. 19. At low frequencies the controller is approximately

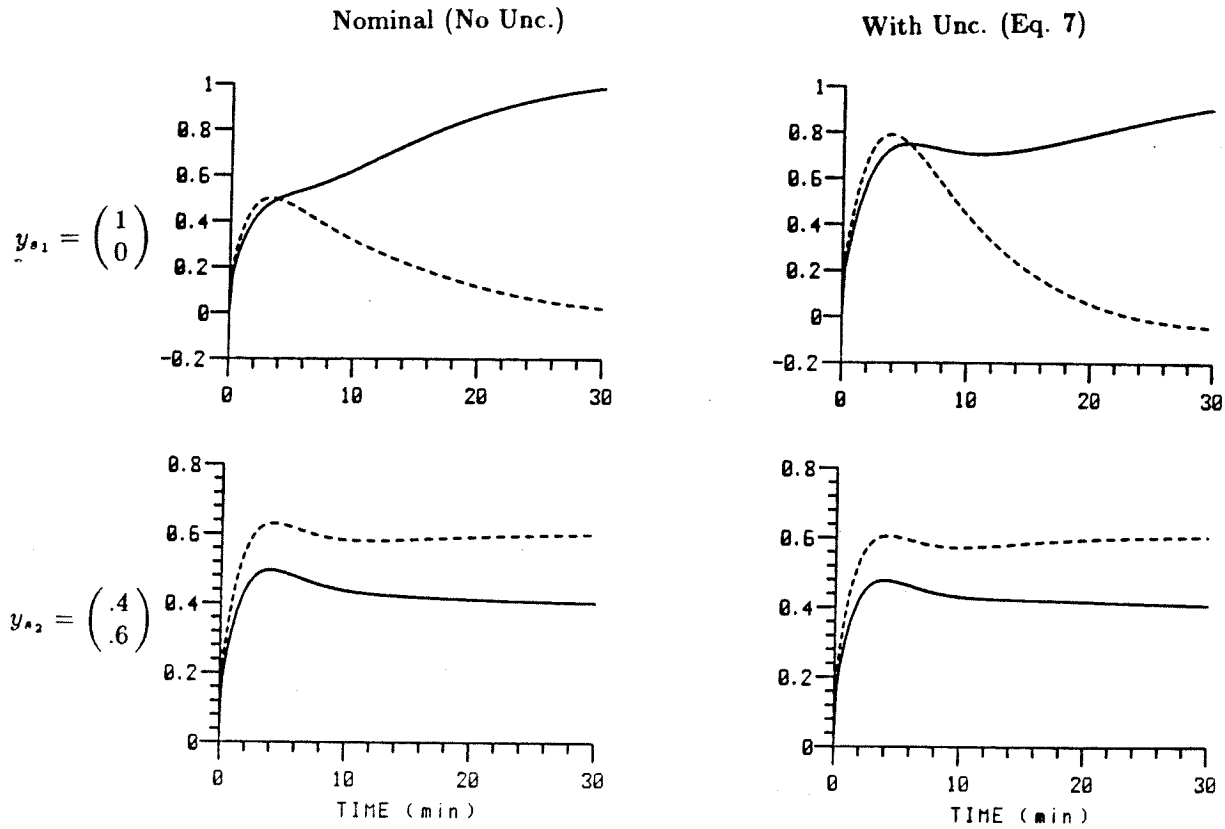


Figure 17 and 18. Closed-loop responses Δy_D and Δx_B with " μ -optimal" controller, $C_\mu(s)$ (time in minutes).

$$A = \begin{bmatrix} -1.002 \cdot 10^{-7} & 0 & 0 & 0 & 0 & 0 \\ 0 & -3.272 \cdot 10^{-6} & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.1510 & 0 & 0 & 0 \\ 0 & 0 & 0 & -9.032 & 0 & 0 \\ 0 & 0 & 0 & 0 & -538.2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -586.8 \end{bmatrix}$$

$$B = \begin{bmatrix} -65.13 & -90.09 \\ 72.24 & 90.31 \\ 5.492 & -4.394 \\ -90.86 & -113.6 \\ 1867 & -1494 \\ 672.2 & 840.3 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.6564 & 0.7171 & 4.949 & 5.033 & -1691 & -311.2 \\ 0.6555 & 0.5425 & 4.941 & -5.040 & -1689 & 311.6 \end{bmatrix}$$

$$D = \begin{bmatrix} 5866 & -3816 \\ 5002 & -4878 \end{bmatrix}$$

Figure 19. State space realization of " μ -optimal" controller, $C_\mu(s) = C(sI-A)^{-1}B + D$.

$$C_{\mu}(s) \approx \frac{(1+75s)}{s} \begin{bmatrix} 3.82 & -0.92 \\ 1.73 & -3.52 \end{bmatrix} \quad (\omega \leq 0.1) \quad (35)$$

The condition number at low frequency is 2.1, and the controller gives some compensation for the directionality of the plant ($\gamma(G) = 141.7$, while $\gamma(GC_{\mu}) = \gamma(C_{\mu}G) = 66.5$).

Structure of Δ_I . Note that Δ_I was assumed to be a "full" matrix in all the calculations above. It turns out that for this particular plant (3), the same values are found for μ_{RS} and μ_{RP} also when Δ_I is assumed to be diagonal, which is a more reasonable assumption from physical consideration (there is no reason to expect that the manipulated variables will influence each other). For the DV-configuration below. It is of crucial importance to model Δ_I as a diagonal matrix and not as a full matrix.

Analysis of the DV-Configuration

The set of possible plants is given by (22) with $G = G_{DV}$ (12), but with Δ_I restricted to be diagonal. We will again consider an inverse-based and a diagonal PI controller

$$C_3(s) = \frac{k_3}{s} G_{DV}^{-1}(s) \quad (13)$$

$$C_4(s) = \frac{1+75s}{s} \begin{bmatrix} -0.15 & 0 \\ 0 & -7.5 \end{bmatrix} \quad (36)$$

In the simulations in Section II we studied the controller $C_3(s)$ with $k_3 = 0.7$. For this controller the Nominal Performance and Robust Stability conditions are identical to those of controller $C_1(s)$ and the LV-configuration. However, based on the simulations and other arguments presented before, μ for Robust Performance is expected to be much better.

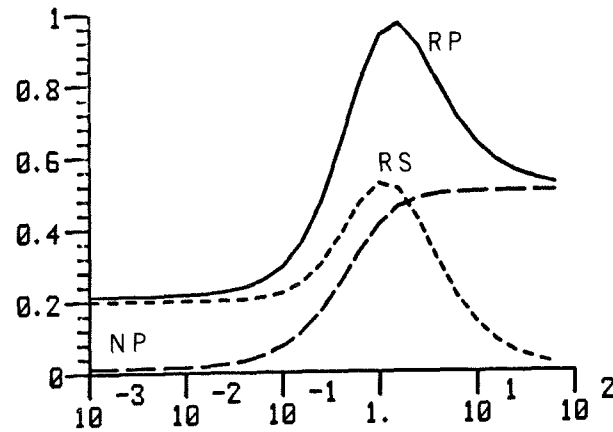


Figure 20. DV-configuration. μ -plots for inverse-based controller, $C_3(s)$, $k_3 = 0.7$.

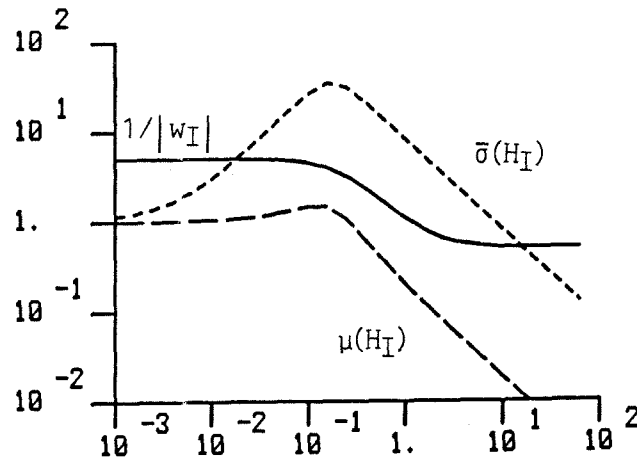


Figure 21. DV-configuration. Robust Stability for controller $C_4(s)$. Using $\bar{\sigma}(H_I)$ instead of $\mu(H_I)$ will be very conservative in this case, and one would mistakenly conclude that the system does not satisfy the R.S. condition $\mu(H_I) < 1/|w_I|$.

This is indeed the case, as seen from the μ -plots in Fig. 20. μ_{RP} is 0.965, which means that the performance criterion is satisfied for all possible model errors. The uncertainty block Δ_I was assumed to be diagonal. If Δ_I were full (which is not the case) the value of μ_{RP} is about 4.1. The reason for the high value in this case is that the off-diagonal elements in Δ_I introduce errors in D when V is changed.

Even lower values for μ are obtained by reducing the gain k_3 in $C_3(s)$ from 0.7 to 0.13. $k_3 = 0.13$ gives $\mu_{RP} = 0.63$. In fact, this controller seems to be very close to the μ -optimal controller for this plant, as we were not able to reduce μ_{RP} below this value by applying the software.

With $C_4(s)$ which consists of two PI-controllers, $\mu_{RP} = 1.15$. This is almost acceptable, although the value of μ_{RP} is significantly higher than for the inversed-based controller $C_3(s)$ with $k_3 = 0.13$. Thus a decentralized controller gives acceptable performance.

The potential conservativeness in using $\bar{\sigma}$ instead of μ is clearly illustrated by considering the Robust Stability test for this case (Fig. 21). Using $\mu_{\Delta_I}(H_I)$ (Δ_I diagonal) we see that the system satisfies the R.S. condition. However, by looking at $\bar{\sigma}(H_I)$ (or equivalently by computing μ with Δ_I a full matrix), we would erroneously expect the system to become unstable for very small errors on the inputs.

V. UNCERTAINTY MODELLING

In this section we will first discuss in somewhat general terms how to quantify uncertainty and then consider as an example, other sources than input uncertainty for the distillation column. In order to use the framework for analyzing systems with uncertainty outlined in Section III, we need to model the uncertainty as norm bounded perturbations. Since the

uncertainty structure is very problem dependent, it is difficult to give general methods for how to do this. However, the examples given below for the distillation column should be sufficient to show that most uncertainties occurring in process control can be modelled as norm bounded perturbations.

Choosing the Right Structure

It may be very important that the correct structure is chosen for the uncertainty description, i.e., that the uncertainty is modelled as it occurs physically. We will illustrate this by considering the following two examples:

- multiplicative uncertainty at the input (Fig. 22A) or at the output of the plant (Fig. 22B)
- output uncertainty as multiplicative (Fig. 22B) or inverse multiplicative uncertainty (Fig. 22C)

Choices of Multiplicative Uncertainty. The distillation column (and any other plant) has multiplicative uncertainty at the input of the plant. Simply shifting this uncertainty to the output of the plant (and using $w_0=w_I$) will, in general, give a completely different system. As an example, for the LV-configuration using controller $C_1(s)$ we found $\mu_{RP} = 5.78$ with the uncertainty at the input of the plant, but μ_{RP} is only 0.96 if this uncertainty is shifted to the output. Recall condition (34) which showed that with input uncertainty and using an ill-conditioned controller, Robust Performance might be poor even when the R.S. and N.P. conditions were satisfied individually. We do not have this problem when the uncertainty is at the output. In this case we get a R.P. condition similar to (34) but without the condition number

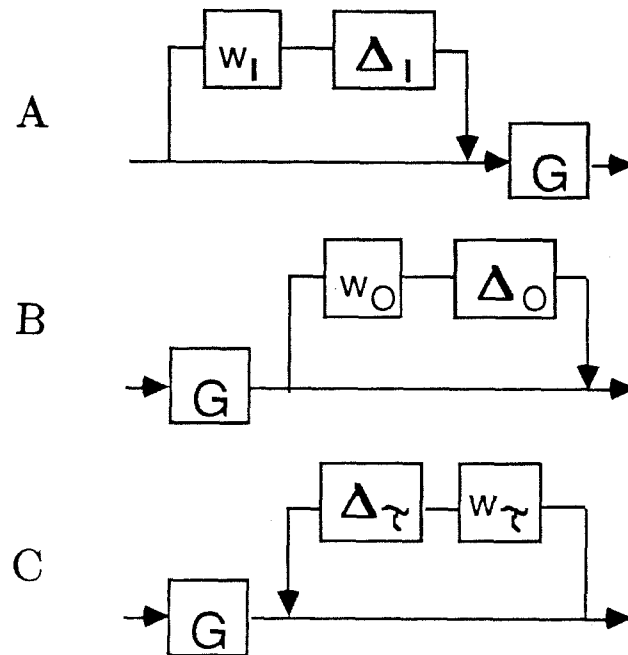


Figure 22.

Three common uncertainty representations. (A) Input multiplicative uncertainty; (B) Output multiplicative uncertainty; (C) Output inverse multiplicative uncertainty.

$$R.P. \leq \bar{\sigma}(w_p S) + \bar{\sigma}(w_o H) \leq 1 \quad \forall \omega$$

This illustrates that output uncertainty usually puts much less constraints on the design of the controller than input uncertainty, and for ill-conditioned plants one should be careful to model the uncertainty at the location where it is actually occurring.

Choices of Output Uncertainty. We will show below that parametric uncertainty in the time constant may be represented as inverse multiplicative uncertainty

$$(I + w_\tau \Delta_\tau)^{-1} G \quad (37)$$

Approximating it as the seemingly similar multiplicative uncertainty

$$(I + w_o \Delta_o) G \quad (38)$$

has drastically different implications. For Robust Stability (37) imposes a constraint on the sensitivity

$$\mu(S) \leq 1/|w_\tau|, \quad S = (I+GC)^{-1} \quad (39)$$

and (38) on the complementary sensitivity

$$\mu(H) \leq 1/(|w_o|), \quad H = GC(I+GC)^{-1} \quad (40)$$

(37) is best suited to describe pole variations while (38) is better for the modelling of zero variations. (37) cannot be used to describe uncertain high frequency dynamics. (38) cannot be used to model plants which have poles that can cross the $j\omega$ -axis.

Simplify if possible. The two examples above illustrated that it may be very important to model the uncertainties as they occur physically.

However, this is not always of crucial importance, and whenever possible the uncertainty description should be simplified by lumping various uncertainties into a single perturbation. There are two reasons for this:

1) Computations are simpler, 2) Introducing too many sources of uncertainty

may be very conservative since it becomes very unlikely for the "worst case" to occur in practice. In particular, the individual uncertainties may be correlated, and it may be impossible for the "worst case" to occur. This will be illustrated for the distillation column later.

Representing Nonlinearities as Uncertainty for the Distillation Column

In addition to the input uncertainty, the main source of "uncertainty" for the distillation column are nonlinearities. All the developments below are for the LV-configuration. However, because of (11) they also apply to the DV-configuration.

A simulation using the equations given in the Appendix and the input uncertainty (7) reveals that the system is unstable with the inverse-based controller $C_1(s)$ (6). Our linear analysis predicted Robust Stability, and the reason for the discrepancy is nonlinearities which were neglected. One way of handling nonlinearities within a linear framework is to treat them as uncertainty. This is clearly not a rigorous way of handling nonlinearities, but this approach is taken in lack of anything better.

Nonlinear open loop responses to large changes ($\pm 6.2\%$) and a small change (0.003%) in boilup V (keeping L constant) are shown in Fig. 23. These responses may be approximated by linear first order responses:

$$V + 6.2\%: \quad \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} \frac{-1.380}{73s+1} \\ \frac{-0.047}{14.5s+1} \end{bmatrix} dV \quad (41a)$$

$$V + 0.003\%: \quad \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} \frac{-0.933}{267s+1} \\ \frac{1.027}{210s+1} \end{bmatrix} dV \quad (41b)$$

$$V - 6.2\%: \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} \frac{0.045}{46s+1} \\ \frac{1.381}{43s+1} \end{bmatrix} dV \quad (41c)$$

Note that the smallest eigenvalue of the linearized plant corresponds to a time constant of 220 min. The value $\tau = 75$ min used in the nominal model (3) represents an average value of the time constants found in the nonlinear simulations.

Time constant uncertainty. From the following simple formula for estimating the linearized time constant (Skogestad and Morari, 1986b).

$$\tau_c = \frac{N_T M_i / F}{z \ln S} \quad (42)$$

$$z = \frac{D}{F} (1-y_D) y_D + \frac{B}{F} x_B (1-x_B)$$

$$\ln S = \ln \frac{y_D (1-x_B)}{x_B (1-y_D)}$$

We find that the time constant reaches its largest value when both products have equal purity ($x_B=1-y_D=0.01$), and this explains the observation that the time constant is large for small changes in V and much smaller for large changes. In our case $N_T = 41$, $M_i/F = 0.5$ min, $z = 0.01$, $\ln S = 9.19$ and we find $\tau_c = 223$ min. The new steady state reached by increasing V by 6.2% is $y_D = 0.71403$, $x_B = 0.000602$. For this operating point we find $\ln S = 8.33$, $z = 0.102$ and $\tau_c = 24$ min. The observed variations in the time constant may be captured with the following linear model

$$G(s) = \begin{bmatrix} \frac{1}{1+\tau_D s} & 0 \\ 0 & \frac{1}{1+\tau_B s} \end{bmatrix} G(0) \quad (43a)$$

$$\tau_D = \tau(1+r_\tau \Delta\tau_D), \quad |\Delta\tau_D| < 1 \quad \forall \omega \quad (43b)$$

$$\tau_B = \tau(1+r_\tau \Delta\tau_B), \quad |\Delta\tau_B| < 1 \quad \forall \omega \quad (43c)$$

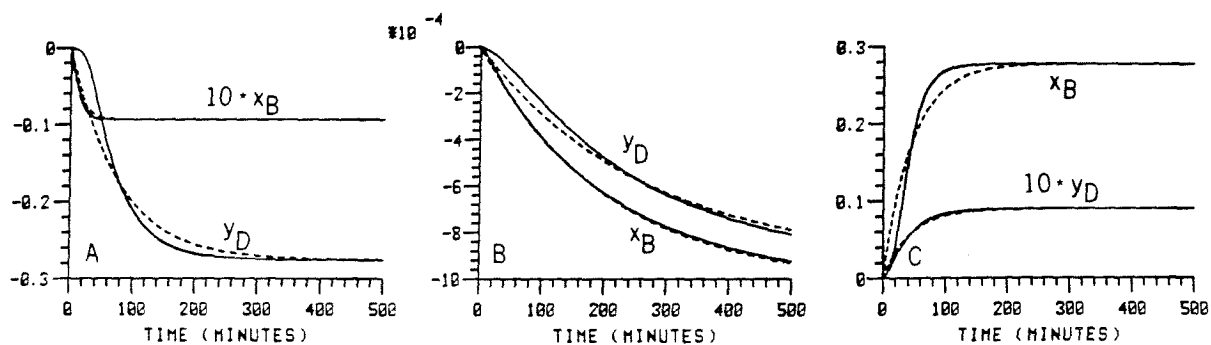


Figure 23. (—) Nonlinear open loop responses Δy_D and Δx_B for changes in boilup V (reflux L constant). (---) Approximation with linear first order response (Eq. 41).
A: $V + 6.2\%$, B: $V + 0.003\%$, C: $V - 6.2\%$.

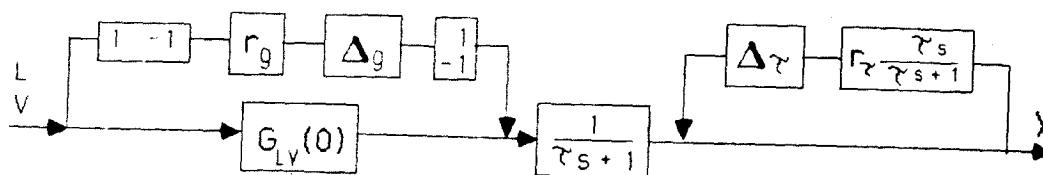


Figure 24. Block diagram representation of gain uncertainty and time constant uncertainty.

Here $\tau = 75$ min and r_τ is a constant expressing the relative uncertainty in the time constants. The scalars Δ_{τ_D} and Δ_{τ_B} are independent which allows for different values for τ_D and τ_B . Note that this model implies that both inputs always have the same time constant with respect to y_D and x_B . This is indeed what is observed when linear analysis is applied at different operating points, and this also applies to disturbances in F , z_F , etc. This pole uncertainty may be written in terms of an inverse multiplicative uncertainty at the output of the plant as shown in Fig. 24. It is fortunate that it occurs at the output since we know that the system is less sensitive to uncertainty at the output than at the input of the plant. Also note that this kind of inverse output uncertainty puts a constraint on the sensitivity function S , similar to a performance requirement. The Robust Stability test for this uncertainty alone is

$$\mu_{\Delta_\tau}(S) \leq 1/|w_\tau|, \quad w_\tau(s) = r_\tau \frac{\tau s}{\tau s + 1}$$

where Δ_τ is a diagonal 2x2 matrix. Clearly, we need $r_\tau < 1$ to satisfy this bound. It may seem strange that we have chosen the nominal value of τ to be 75 minutes, since it is clearly not possible to include even the linearized time constant (230 min) in the model (43). Recall, however, that we are trying to represent a nonlinear system. The linearized time constant only applies in a very small operating region, and as the system moves away from this steady state (maybe because of instability) the time constant will be small. It is therefore much more important to include the smallest value observed for the time constants in the approximation (43).

Gain uncertainty. We observe from (41) that the linearized gains vary

tremendously with operating conditions. However, the gains are clearly correlated and it is of crucial importance to take this into account to avoid a hopelessly conservative uncertainty description. If the elements in the steady state gain matrix (3) were assumed to be independent, the gain matrix would become singular for relative errors in each of the elements exceeding (Skogestad & Morari, 1985)

$$\frac{1}{\gamma^*(G_{LV})} = \frac{1}{138.3} = 0.007 \quad (44)$$

Here $\gamma^*(G)$ is the minimized condition number

$$\gamma^*(G) = \min_{D_1, D_2} \gamma(D_1 G D_2) \quad (45)$$

(D_1 and D_2 are diagonal matrices with real, positive entries). Physically we know that the distillation column will not become singular and a more structured uncertainty description is needed. Skogestad and Morari (1986c) have suggested that for small changes in D/B the variations in the steady state gains may be captured with additive uncertainty on the elements using a single perturbation Δ_g . For the LV-configuration

$$\begin{aligned} G_{LVp}^{(0)} &= G_{LV}^{(0)} + r_g \Delta_g \begin{bmatrix} 1 & -1 \\ -\frac{D}{B} & \frac{D}{B} \end{bmatrix} \\ &= G_{LV}^{(0)} + \begin{bmatrix} 1 \\ \frac{D}{B} \\ -\frac{D}{B} \end{bmatrix} r_g \Delta_g [1 \quad -1], \quad |\Delta_g| < 1 \end{aligned} \quad (46)$$

This model does not match our data (41) too well, where large variations in D/B are observed. However, under closed loop we do not expect large changes in D/B (though the changes in L and V individually may be large) and (46) with D/B = 1 will be used to represent the gain variations.

It is important to note that the additive uncertainty in (46) does not change the singular vectors ∇ and \underline{y} . A SVD of the perturbation matrix

$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ in (46) yields

$$\mathbf{v} = \begin{bmatrix} 0.707 \\ -0.707 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 0.707 \\ -0.707 \end{bmatrix}$$

The direction of the "input" singular vector \mathbf{v} is the same as that of the nominal plant (3), while the "output" singular vector \mathbf{u} is almost perpendicular to that of the nominal plant. This means that this source of nonlinearity is also "nice" in that it mainly changes the plant at the output. Physically this means that changes in the external flow (D and B) are always the changes with the largest effect, and this is exactly what we would expect from physical considerations, also for the nonlinear plant.

Choice of values for r_T and r_g . There is clearly a correlation between the variations in the time constant and gains which is not captured by the proposed uncertainty description (Fig. 24). However the main effect of both these uncertainties is to change the direction of the output singular vectors, $\underline{\mathbf{u}}$ and $\bar{\mathbf{u}}$. None of them add RHP-zeros. In order to make computations simpler and to avoid conservative results (by neglecting the correlation between Δ_T and Δ_g), a reasonable approach may therefore be to use only one of these uncertainties to describe the effect of nonlinearity. This is the approach taken here and we choose to use the time constant uncertainty only. One reason for not choosing the gain uncertainty, is that this introduces uncertainty at steady state, which will normally not be the case since the setpoints are not changed significantly.

The trajectory taken by the plant under closed loop may be very different from the open loop responses, and open loop data such as (41) may not be appropriate to determine the value of r_T . We therefore decided

to use nonlinear closed loop simulations to find an appropriate value for r_T . In particular, closed loop simulations which are at the limit to instability are convenient, since these results may easily be compared to values of μ_{RS} for the robust stability of the linear approximation. To determine the value of r_T the following procedure was used:

1. Nonlinear closed loop simulations were carried out for a large feed rate disturbance (+30%) with the inverse-based controller $C_1(s)$, $k_1 = 0.7$. The feed rate disturbance was chosen as the most difficult disturbance (Skogestad and Morari, 1986a) which would take the system furthest away from the nominal steady state. To make the elements in the matrix (9) large, the relative errors on the manipulated inputs L and V were chosen with different signs. These errors were increased until the system was at the limit to instability. The limiting values were

$$dL = 1.04 \text{ dL}, dV = 0.96 \text{ dV}_C$$

2. μ_{RS} was computed for the LV-configuration with 4% input uncertainty ($w_I=0.04$) and with various values for the relative uncertainty on the time constants, r_T . $r_T = 0.35$ was found to give $\mu_{RS} = 1$, i.e., correspond to a system at the limit to instability.

The value found for r_T using this procedure is clearly not the only possible (note that no error was assume in the gains), but hopefully represents a reasonable compromise between representing all possible plants and avoiding a very conservative uncertainty description.

Effect of additional uncertainty on μ_{RS} and μ_{RP} . With the additional time constant uncertainty (43) the interconnection matrix N becomes

$$N = \begin{bmatrix} -w_{ICS}G & -w_{ICS} & w_{ICS} \\ w_{TS}G & w_{TS} & w_{TS}(I-S) \\ w_{PS}G & w_{PS} & -w_{PS} \end{bmatrix} \quad (47)$$

$\mu(N)$ for R.P. (Table 2) is computed with respect to the structure $\text{diag}\{\Delta_I, \Delta_T, \Delta_P\}$ where Δ_I and Δ_P are "full" 2x2 matrices and Δ_T is a diagonal 2x2 matrix. For computational convenience the matrix Δ_T is assumed complex.

The inverse-based controller $C_1(s)$, is, of course, not robustly stable. (It was shown to be unstable with 4% input uncertainty and now there is 20%). μ_{RS} is increased from 0.53 to 4.77 by adding 35% time constant uncertainty. The μ -values for the diagonal controller, $C_2(s)$, and the "p-optimal", $C_\mu(s)$, are seen to be only weakly influenced by adding the pole uncertainty. Robust stability is still predicted for the μ -optimal controller. This is confirmed by nonlinear simulations.

To confirm that the gain uncertainty does not significantly change these results, similar calculations were also done with the nonlinearities represented as uncertainty on the gains (Table 3). Interestingly enough, it turns out that choosing $r_g = 0.35$, $r_p = 0$ gives very similar results as $r_g = 0$, $r_p = 0.35$ (Fig. 25). Furthermore, combinations of these uncertainties were found to add up approximately in a linear fashion with respect to the value of μ . This confirms that in this case, these two sources of uncertainty (pole and gain uncertainty) have a very similar effect on the plant, and that for computational simplicity we need to use only one of them. Similar results are found for the DV-configuration (Table 3), although the pole uncertainty is found to be worse than the gain uncertainty.

Input Uncertainty, $w_I = 0.2 \frac{5s+1}{0.5s+1}$

	$r_T=r_g=0$			$r_T=0.35$		$r_g=0.35$	
	μ_{NP}	μ_{RS}	μ_{RP}	μ_{RS}	μ_{RP}	μ_{RS}	μ_{RP}
LV-configuration							
Inverse-Controller, $C_1(s), k_1=0.7$	0.50	0.53	5.78	4.77	7.50	4.83	7.53
Diagonal PI, $C_2(s), k_2=2.4$	1.50	1.39	1.70	1.61	1.91	1.47	1.82
Optimal Inverse, $C_1(s), k_1 = 0.14$	0.50	0.20	3.29	2.60	4.18	2.62	4.19
" μ -optimal", $C_\mu(s)$	0.79	0.72	1.06	0.99	1.29	0.87	1.24
DV-configuration							
Inverse-controller, $C_3(s), k_3=0.7$	0.50	0.53	0.97	0.83	1.18	0.53	1.07
Diagonal PI, $C_4(s)$	0.81	0.37	1.14	0.85	1.61	0.61	1.45
Optimal Inverse, $C_3(s), k_3=0.13$	0.50	0.20	0.63	0.47	0.81	0.20	0.73

Table 2. Values of μ_{NP} , μ_{RS} and μ_{RP} for distillation column with diagonal input uncertainty, and effect on μ_{RS} and μ_{RP} (μ_{NP} is unchanged) by adding time constant uncertainty (r_T) and gain uncertainty (r_g).

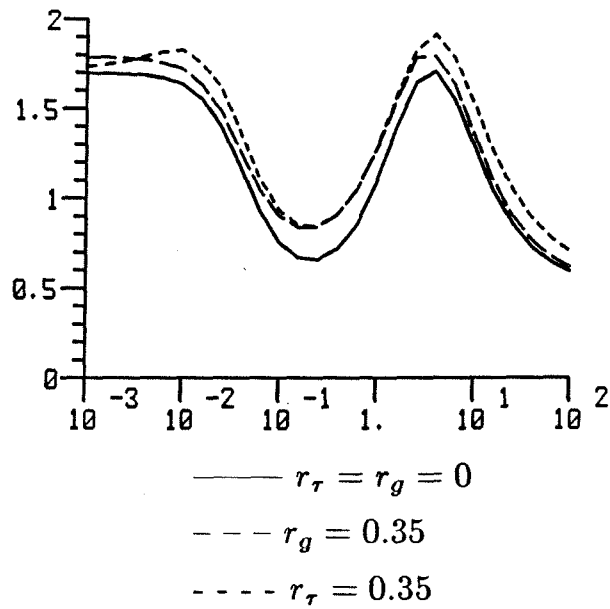


Figure 25.

μ -plots for R.P. for diagonal controller, $C_2(s)$, $k_2 = 0.7$. The addition of gain or time constant uncertainty is seen to have a similar and not too significant effect on the value of μ .

The conclusions with respect to the effect of the nonlinearities for the distillation column are

1. The main effect of the nonlinearities is to change the "directions" at the output of the plant.
2. Representing this effect by uncertainty in the time constant seems to be a good approach.
3. Since output uncertainty may be thought of as a output "disturbance", controllers which were found to give good Robust Performance in Section IV, are not affected very much by these nonlinearities.

VI. CONCLUSIONS

High purity distillation columns are inherently ill-conditioned because the product compositions are very sensitive to changes in the product flow rates. This may cause performance problems if the uncertainty changes the directionality of the plant. This is the case for the traditional LV-configuration, where input uncertainty on the manipulated variables changes the directionality at the input of the plant, and makes it impossible to use an inverse based controller ("decoupler"). For the DV-configuration (direct material balance) the input uncertainty poses no problem. The structured singular value, μ , was used as a tool to study the effect of uncertainty on stability and performance in a systematic manner.

It is clear from this example that even with a powerful tool like μ the control system design process consists of a sequence of iterative steps involving nonlinear modelling, and simulation as well as linear analysis and synthesis. In particular, it is nontrivial to arrive at a description of model "uncertainty" which captures the behavior of the real process and can be treated mathematically.

References

- Doyle, J. C. and G. Stein, Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis, IEEE Trans. Automat. Contr., AC-26, 1, 4-16 (1981).
- Doyle, J. C., Analysis of Feedback Systems with Structured Uncertainties, Proc. IEE, 129, D(6), 242-247 (1982).
- Doyle, J. C., Structured Uncertainty in Control System Design, 24th CDC, Ft. Lauderdale, FL (1985).
- Doyle, J. C., J. E. Wall and G. Stein, Performance and Robustness Analysis for Structured Uncertainty, 21st CDC, Orlando FL (1982).
- Morari and J. Doyle, A Unifying Framework for Control System Design under Uncertainty and Its Implication for Chemical Process Control, CPC 1986, Asilomar, CA, Jan. (1986).
- Morari, M., E. Zafiriou and C. G. Economou. "Robust Process Control", Springer Verlag, Berlin (1987).
- Moczek, J. S., R. E. Otto and T. J. Williams, Approximate Models for the Dynamic Response of Large Distillation Columns, Chem. Eng. Progr. Symp. Ser., No. 61, 136-146 (1965).
- Shinskey, F. G., "Distillation Control", 2nd Edition, McGraw-Hill (1984).
- Skogestad, S. and M. Morari, Design of Resilient Processing Systems. Effect of Model Uncertainty on Dynamic Resilience (submitted to Chem. Eng. Sci. (1985).
- Skogestad, S. and M. Morari, Effect of Disturbance Directions on Closed Loop Performance, submitted to Ind. Eng. Chem. Process Des. Dev. (1986a).
- Skogestad, S. and M. Morari, Shortcut Models for the Dynamic Behavior of Distillation Column, in progress (1986b).
- Skogestad, S. and M. Morari, Nonlinearity and Uncertainty in the Steady State Gain Matrix for Distillation Columns, in progress (1986c).
- Stein, G., Beyond Singular Values and Loop Shapes, 24th CDC, Ft. Lauderdale, FL (1985).

Appendix. Simplified nonlinear dynamical model of distillation column with total condenser.

Assumptions:

- Constant molar flows
 - No vapor holdup (immediate vapor response, $dV_{top} = dV_{btm}$)
 - Liquid holdup M_i constant (immediate liquid response, $dL_{top} = dL_{btm}$)
 - Vapor Liquid Equilibrium (VLE) and perfect mixing on each stage
 - Perfect level control in accumulator and column base, pressure constant
- N - no. of equilibrium (theoretical) stages including the reboiler
 $N_T = N + 1$ - total no. of stages including total condenser
 N_F - feed stage location

Material balance describing change in holdup of light component on each tray:

$i = 2, N$ ($i \neq N_F$, $i \neq N_F + 1$):

$$M_i \dot{x}_i = L_{i+1} x_{i+1} + V_{i-1} y_{i-1} - L_i x_i - V_i y_i$$

Above feed location, $i = N_F + 1$

$$M_i \dot{x}_i = L_{i+1} x_{i+1} + V_{i-1} y_{i-1} - L_i x_i - V_i y_i + F_V y_F$$

Below feed location, $i = N_F$

$$M_i \dot{x}_i = L_{i+1} x_{i+1} + V_{i-1} y_{i-1} - L_i x_i - V_i y_i + F_L x_F$$

Reboiler, $i = 1$

$$M_B \dot{x}_B = L_{i+1} x_{i+1} - V_i y_i - B x_B, \quad x_B = x_1$$

Total condenser, $i = N_T$

$$M_D \dot{x}_D = V_{i-1} y_{i-1} - L_i x_i - D x_D, \quad y_D = x_{N_T}$$

VLE on each tray

$$y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i}, \quad i = 1, N$$

Flow rates

$$i > N_F \text{ (above feed):} \quad L_i = L, \quad V_i = V + F_V$$

$$i \leq N_F \text{ (below feed):} \quad L_i = L + F_L, \quad V_i = V$$

$$F_L = q_F F, \quad F_V = F - F_L$$

$$D = V_{N_T-1} - L_{N_T} = V + F_V - L \text{ (accumulator holdup constant)}$$

$$B = L_2 - V_1 = L + F_L - V \text{ (column base holdup constant)}$$

Compositions x_F and y_F are found by solving the flash equations for the feed

$$Fz_F = F_L x_F + F_V y_F$$

$$y_F = \frac{\alpha x_F}{1 + (\alpha - 1)x_F}$$

Chapter V

IMPLICATIONS OF LARGE RGA-ELEMENTS ON CONTROL PERFORMANCE

IMPLICATIONS OF LARGE RGA-ELEMENTS ON CONTROL PERFORMANCE

Sigurd Skogestad

Manfred Morari

Chemical Engineering, 206-41

California Institute of Technology

Pasadena, CA 91125

(818)356-4186

Submitted to I&EC Process Design & Dev.

September 1986

Updated January 1987

Abstract:

Large elements in the RGA imply a plant which is fundamentally difficult to control:

1. The plant is very sensitive to uncorrelated uncertainty in the transfer matrix elements.
2. The closed loop system with an inverse-based controller is very sensitive to input uncertainty. With a diagonal controller the system is not sensitive to diagonal input uncertainty, but the controller does not correct for the strong directionality of the plant, and may therefore give poor performance even without uncertainty.

1. Introduction

Each element in the RGA is defined as the open loop gain divided by the gain between the same two variables when all other loops are under "perfect" control (Bristol, 1966)

$$\lambda_{ij} = \frac{(\partial y_i / \partial u_j)_{u_{k \neq j}}}{(\partial y_i / \partial u_j)_{y_{k \neq i}}} = \frac{\text{Gain all other loops open}}{\text{Gain all other loops closed}} \quad (1)$$

or equivalently (this definition may be used for any frequency)

$$\text{RGA} = \Lambda = \{\lambda_{ij}\} = G(s) \times G^{-1}(s)^T \quad (2a)$$

where \times denotes element-by-element multiplication. For 2x2 plants

$$\text{RGA} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 1-\lambda_{11} \\ 1-\lambda_{11} & \lambda_{11} \end{bmatrix}, \quad \lambda_{11} = 1 / (1 - \frac{g_{12}g_{21}}{g_{11}g_{22}}) \quad (2b)$$

For decentralized control it is preferable to choose pairings of inputs (j) and outputs (i) which have λ_{ij} close to one. Pairing of variables with negative values of λ_{ij} should be avoided since in this case the loop gains may change sign as the other loops are opened or closed (Grosdidier et al., 1985).

A value of λ_{ii} close to one also indicates that "interactions" are small if a SISO controller is used for this loop. This makes it possible to tune each loop individually and to get an acceptable (stable) response when all the loops are closed simultaneously: For 2x2 plants we can tune each loop individually with integral control, such that the stability of the overall system is guaranteed if $\lambda_{11} > 0.5$ (Grosdidier & Morari, 1986).

However, the RGA is used as a measure of control quality in a much wider sense than as a tool for choosing pairings for decentralized control. In particular, large elements in the RGA are suggested to imply a plant which is fundamentally difficult to control (Bristol,

1966; McAvoy, 1983, Grosdidier et al., 1985). The following identity supports this claim. Let $G = \{g_{ij}\}$ and $G^{-1} = \{\hat{g}_{ij}\}$. Then (Grosdidier et al., 1985)

$$d\hat{g}_{ji}/\hat{g}_{ji} = -\lambda_{ij} \frac{dg_{ij}}{g_{ij}} \quad (3)$$

This identity shows that the elements of the inverse (G^{-1}) are extremely sensitive to small changes in G if the RGA-elements are large. This seems to indicate that plants with large RGA-elements are very sensitive to modelling errors and this is indeed true as we show in the paper.

In this paper we want to answer the following two questions:

- A) Is a plant with large elements in the RGA always difficult to control?
- B) Is a plant with small elements in the RGA always easy to control (in the absence of other limitations on control performance, such as constraints and RHP zeros)?

We will look at the questions in the context of model uncertainty. Because of the extensive use of the RGA as a tool for evaluating control configurations for distillation columns (Shinskey, 1984), we will show through some examples how our results apply in this case.

2. Relationships between the RGA and the condition number

A plant is ill-conditioned if it has a high condition number. The condition number based on $\bar{\sigma}(\cdot)$ is

$$\gamma(G) = \bar{\sigma}(G)\bar{\sigma}(G^{-1}) = \bar{\sigma}(G)/\underline{\sigma}(G) \quad (4)$$

where the maximum singular value $\bar{\sigma}(G) = \max_u \|Gu\|_2 / \|u\|_2$ is the induced 2-norm. Bristol (1966) himself pointed out the close resemblance between the condition number and the RGA. Plants with

large elements in the RGA are always ill-conditioned as seen from (5) (Nett & Manousiouthakis, 1986)

$$\gamma(G) \geq \gamma^*(G) \geq ||\Lambda||_m^{-1} / \gamma^*(G) \geq ||\Lambda||_m - 1 \quad (5)$$

Here $\gamma^*(G)$ is the minimized scaled condition number

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

(D_1 and D_2 are diagonal "scaling" matrices with real, positive entries)

and

$$||\Lambda||_m = 2 \max \{ ||\Lambda||_{i1}, ||\Lambda||_{i\infty} \} \quad (6)$$

where $||\Lambda||_{i\infty}$ denote the induced 1- and ∞ -norms

$$||\Lambda||_{i1} = \max_j \sum_{i=1}^n |\lambda_{ij}| \quad (\text{"max column sum"})$$

$$||\Lambda||_{i\infty} = \max_i \sum_{j=1}^n |\lambda_{ij}| \quad (\text{"max row sum"})$$

From (5) we see that large elements in the RGA always imply a large value of $\gamma^*(G)$ and $\gamma(G)$. Since ill-conditioned plants are generally believed to cause control problems, (5) gives some justification to the claim that plants with large elements in the RGA are fundamentally difficult to control. Note that $\gamma(G)$ can be significantly larger than $\gamma^*(G)$, and the plant may therefore be ill-conditioned ($\gamma(G)$ large) even if all the elements in the RGA are small. In particular, 2x2 plants with an odd number of negative elements in G always have $\gamma^*(G) = ||\Lambda||_m = 1$, but $\gamma(G)$ may be arbitrary large. For example

$$G = \begin{bmatrix} 1 & -0.01 \\ 1 & 0.01 \end{bmatrix}, \quad \gamma(G) = 100, \quad \gamma^*(G) = ||\Lambda||_m = 1$$

There is a close relationship between $||\Lambda||$ and $\gamma^*(G)$. From (5) we know that $\gamma^*(G)$ is always large when there are large elements in the RGA. And, similarly, a large value of $\gamma^*(G)$ always implies large elements in

the RGA. This is seen from the following bound which applies to 2x2 plants (Grosdidier, 1985)

$$(2 \times 2) \quad \gamma^*(G) \leq ||\Lambda||_1 \quad (7)$$

and from the following conjecture for nxn plants (Skogestad et al., 1986, Nett & Manousiouthakis, 1986)

$$\gamma^*(G) \leq ||\Lambda||_1 + k(n) \quad (8)$$

with $k(2) = 0$, $k(3) = 1$ and $k(4) = 2$. Here

$$||\Lambda||_1 = \sum_{i,j} |\lambda_{ij}|$$

Note that for 2x2 plants the 1- and the "m"-norm of the RGA are identical

$$(2 \times 2): \quad ||\Lambda||_1 = ||\Lambda||_m = 2||\Lambda||_{i\infty} = 2||\Lambda||_{i1}$$

Combining (5) and (7) one shows that $||\Lambda||_1$ and $\gamma^*(G)$ are always close in magnitude (in particular when they are large).

$$(2 \times 2): \quad ||\Lambda||_1 - \frac{1}{\gamma^*(G)} \leq \gamma^*(G) \leq ||\Lambda||_1 \quad (9)$$

Consequently, for 2x2 plants $\gamma^*(G) \rightarrow ||\Lambda||_1$ as $||\Lambda||_1 \rightarrow \infty$, and numerical evidence suggests that this also holds for nxn plants.

This close relationship between $||\Lambda||_1$ and $\gamma^*(G)$ is important because $\gamma^*(G)$ is a sensitivity measure with respect to uncorrelated/independent errors in the transfer matrix elements (Skogestad & Morari, 1986a). This result is presented below.

3. The RGA and Model Uncertainty

3.1 Independent Relative Element Uncertainty

This result will introduce $||RGA||_1$ as a sensitivity measure with respect to independent uncertainty on the plant elements.

Result 1. (2x2) (Skogestad & Morari, 1986)

Assume each transfer matrix elements has a relative uncertainty of magnitude r , that is, the actual ("perturbed") plant is

$$G_P = \begin{bmatrix} g_{11}(1+r\Delta_{11}) & g_{12}(1+r\Delta_{12}) \\ g_{21}(1+r\Delta_{21}) & g_{22}(1+r\Delta_{22}) \end{bmatrix}, |\Delta_{ij}| < 1 \quad (10)$$

The uncertainties on each element are assumed to be independent, that is, there is no correlation between the Δ_{ij} 's. The plant G_P remains nonsingular at steady state ($\omega=0$) for any real perturbations $-1 \leq \Delta_{ij} \leq 1$ if and only if

$$r < \frac{1}{\gamma^*(G)} \quad (11)$$

which is satisfied if

$$r < \frac{1}{\| \Lambda \|_1}, \quad \| \Lambda \|_1 = 2|\lambda_{11}| + 2|1-\lambda_{11}| \quad (12)$$

(Condition (12) also holds for complex perturbations $|\Delta_{ij}| \leq 1$ and $\omega > 0$ (Skogestad & Morari, 1986a). Condition (11) does not hold in these cases).

Condition (11) is necessary and sufficient. Condition (12) is only sufficient, but it is also "tight" because of the close relationship between $\gamma^*(G)$ and $\|RGA\|_1$ shown in (9).

Conditions similar to (12) may be derived for nxn plants using conjecture (8) and Theorem 6 in Skogestad and Morari (1986a): The plant remains nonsingular at any frequency ω for complex relative errors of magnitude r on each element if

$$r < 1/(\| \Lambda \|_1 + k(n)) \quad (13)$$

The control implications of conditions (11)-(13) follow from that the fact that if a plant is singular at a certain frequency ω , then the plant has a zero on the $j\omega$ -axis. The presence of this RHP-zero limits the achievable control quality (Morari, 1983). In particular, it is

impossible to have integral control for a plant which may become singular at steady state ($\omega=0$) (Skogestad & Morari, 1986a). Consequently, if there are large elements in the RGA and $\|RGA\|_1$ is large, we can allow only very small uncertainties in the elements without having control problems.

The main restriction inherent in these results is the assumption of independent element uncertainty. Conditions (12) and (13) may be very conservative if the element uncertainties are correlated. For example, for distillation columns, even though $\|RGA\|_1$ is large and the elements in G may vary widely with operating conditions (r may be close to 1), it can be shown that the plant never becomes singular (Skogestad & Morari, 1986b). Therefore, for distillation column control, Result 1 does not "explain" why plants with large values of the RGA are difficult to control.

3.2 Uncertainty on each manipulated input

The following result will introduce the RGA as a measure of how performance is affected by uncertainty on each manipulated input. This result is of more general interest than Result 1, because uncertainty on the manipulated inputs is always present.

Result 2. Diagonal Input Uncertainty

Let Δ_i represent the relative uncertainty on the i 'th manipulated input, i.e., the perturbed plant is

$$G_p = G(I + \Delta_I), \quad \Delta_I = \text{diag}\{\Delta_i\} \quad (14)$$

The loop transfer matrix $G_p C$ may be written in terms of the nominal GC

$$G_p C = GC(I + C^{-1} \Delta_I C) \quad (15)$$

$G_p C$ is closely related to performance because of the identity (Fig. 1)

$$y = (I + G_p C)^{-1} d$$

For 2x2 plants the error term in (15) may be expressed in terms of the RGA of the controller C as follows

$$C^{-1}\Delta_I C = \begin{bmatrix} \lambda_{11}(C)\Delta_1 + \lambda_{21}(C)\Delta_2 & \lambda_{11}(C) \frac{c_{12}}{c_{11}} (\Delta_1 - \Delta_2) \\ -\lambda_{11}(C) \frac{c_{21}}{c_{22}} (\Delta_1 - \Delta_2) & \lambda_{12}(C)\Delta_1 + \lambda_{22}(C)\Delta_2 \end{bmatrix} \quad (16)$$

If the elements in the matrix (16) are large compared to 1 (in particular, the diagonal ones), the loop transfer matrix $G_p C$ will be very different from the nominal GC, and poor response or even instability is expected when $\Delta_I \neq 0$.

nxn Plants. For nxn plants it is easily shown that the diagonal elements of the error matrix $C^{-1}\Delta_I C$ may be written as a straightforward generalization of the 2x2 case

$$(C^{-1}\Delta_I C)_{ii} = \sum_{j=1}^n \lambda_{ji}(C)\Delta_j \quad (17)$$

That is, the diagonal element of $C^{-1}\Delta_I C$ depend on the column elements of $\Lambda(C)$ and the magnitude of the uncertainty.

Control Implications of Result 2. The following is clear from (16) and (17):

- i) Controllers with large elements in the RGA of the controller should always be avoided.

It should be added that it is the behavior of $G_p C$ around crossover ($||G_p C|| \approx 1$) which is of primary importance for the stability and performance of the closed-loop system. Therefore, control problems are expected if the RGA has large element in this frequency range.

Inverse-based controller.

To have "tight" control it is desirable to use an inverse-based controller $C(s) = c(s)G^{-1}(s)$ (where $c(s)$ is a scalar). With this controller the error term becomes

$$C^{-1}\Delta_I C = G\Delta_I G^{-1} = \begin{bmatrix} \lambda_{11}\Delta_1 + \lambda_{12}\Delta_2 & -\lambda_{11} \frac{g_{12}}{g_{22}} (\Delta_1 - \Delta_2) \\ \lambda_{11} \frac{g_{21}}{g_{11}} (\Delta_1 - \Delta_2) & \lambda_{21}\Delta_1 + \lambda_{22}\Delta_2 \end{bmatrix} \quad (18)$$

Here $\lambda_{ij} = \lambda_{ij}(G)$. Similarly, for $n \times n$ plants the diagonal elements in $(G\Delta_I G^{-1})$ depend on the row elements of $\Lambda(G)$:

$$(G\Delta_I G^{-1})_{ii} = \sum_{j=1}^n \lambda_{ij}(G)\Delta_j \quad (19)$$

Note that $\Lambda(G^{-1}) = \Lambda^T(G)$. This means that the RGA-elements of the inverse-based controllers are the same as for the plant. The RGA is independent of scaling, but the off-diagonal elements in (18) will depend on the scaling of the outputs. For the correct interpretation of these elements the plant outputs should be scaled such that an output deviation of magnitude 1 has equal significance for all outputs. (Comment: Similar results, but with g_{12}/g_{22} replaced by g_{21}/g_{22} and g_{21}/g_{11} replaced by g_{12}/g_{21} , may be derived for the case of output uncertainty and performance measured at the input of the plant. This case is generally of less interest).

Control Implications of (18) and (19)

- ii) An inverse-based controller should not be used for a plant with large elements in the RGA.
- iii) Inverse-based controllers may give poor response even if the

elements in the RGA are small. This may happen if g_{12}/g_{22} or g_{21}/g_{11} are large. One example is a triangular plant which always has $\lambda_{11} = 1$, but where the response using an inverse-based controller may display large "interactions" in the presence of uncertainty.

Diagonal Controller.

A diagonal controller always has $\lambda_{11}(C) = 1$ and the error term in (15) becomes

$$C^{-1}\Delta_I C = \Delta_I$$

Therefore the response is only weakly influenced by the presence of input uncertainty. However, it may be difficult to achieve a good nominal response when the controller is restricted to being diagonal (this may be the case even if λ_{11} is close to one provided the plant is nearly triangular). The diagonal controller gives limited correction for the "directionality" of the plant and $\gamma(GC)$ may be large. In this case the response depends strongly on the "disturbance direction": Let d represent the effect on the disturbance on the output. The response is poor for a disturbance d with a large disturbance condition number (Skogestad & Morari, 1986c).

$$\gamma_d(GC) = \frac{\| (GC)^{-1} d \|_2}{\| d \|_2} \bar{\sigma}(GC) \quad (20)$$

$\gamma_d(GC)$ ranges in value between 1 and $\gamma(GC)$. A value close to 1 indicates that the disturbance is in the "good" direction corresponding to the high loop gain, $\bar{\sigma}(GC)$. A value close to $\gamma(GC)$ indicates that the disturbance is in the "bad" direction corresponding to the low loop

gain, $\underline{\sigma}(GC)$. (For an inverse-based controller $\gamma_d(GC) = 1$ for all disturbances).

d may also represent the effect of a setpoint change. If arbitrary setpoint changes are allowed, then there exists a setpoint change y_s such that $\gamma_{y_s}(GC) = \gamma(GC)$. However, in many cases setpoint changes occur infrequently and we may accept a poor response for them. In this case it may not matter if $\gamma(GC)$ is large as long as $\gamma_d(GC)$ is small for all major disturbances.

Diagonal controllers do not generally correct for the directionality of the plant and $\gamma_d(GC)$ is large whenever $\gamma_d(G)$ is large ($\gamma_d(G)$ is the disturbance condition number of the plant). In particular, this is the case for plants with large RGA-elements. This is seen from the following identities

$$C \text{ diagonal: } \gamma(GC) \geq \gamma^*(G) \quad (21)$$

This follows since a diagonal controller merely corresponds to a scaling of the input to the plant. By applying (5) we derive

$$C \text{ diagonal: } \gamma(GC) \geq \|\Lambda\|_m - 1 \quad (22)$$

and we see that a plant with large RGA-values always will have $\gamma(GC)$ large, and will yield poor performance (at least if arbitrary setpoint changes are allowed).

One special case when a diagonal controller may yield acceptable performance for an ill-conditioned plant ($\gamma(G)$ large) is when the plant is naturally "decoupled" at the input ($V=I$). Write the Singular Value Decomposition (SVD) of G

$$G = U \Sigma V^H, \quad \Sigma = \begin{bmatrix} \sigma(G) & 0 \\ 0 & \underline{\sigma}(G) \end{bmatrix} \quad (23)$$

For the case $V = I$ (or more generally, V has only one nonzero element

in each row and column, which gives $V = I$ by rearranging the inputs) a diagonal controller can be found which removes most of the directionality in the plant: Choose $C(s) = c(s) \Sigma^{-1}$ to get $GC = c(s)U$ which has $\gamma_d(GC) = 1$ for all disturbances. Note, however, that the response is not decoupled (unless U is diagonal). Also note that for this case ($V = I$) $\gamma^*(G) = \gamma^*(U \Sigma) = 1$ and the elements in the RGA are all positive and less than 1

(choose $D_2 = \Sigma^{-1}$ and use $\gamma(U) = 1$ and from (5): $\|\Lambda\|_m \leq \gamma^*(G) + 1 = 2$, i.e. $\|\Lambda\|_{i\infty} \leq 1$)

Finding worst-case conditions from the RGA

It is of interest to know the worst-case combination of Δ_j 's (input uncertainty) to use in simulation studies. To obtain the worst case consider (17). If all Δ_j have the same magnitude ($|\Delta_j| < r_I$) then the largest possible magnitude (worst case) of any diagonal element is given by $r_I \|\Lambda(C)\|_{i1}$ ("max column sum"). To obtain this value the signs of the Δ_j 's should be the same as those in the column of $\Lambda(C)$ with the largest elements (if we are looking at $G\Delta_I G^{-1}$ in (19) then the worst case is found when the Δ_j 's have the same sign as those in the row of $\Lambda(G)$ with the largest elements).

Example

Consider a plant with steady-state gain matrix

$$G(0) = \begin{bmatrix} 1 & 0.1 & -2 \\ 1 & 2 & -3 \\ -0.1 & -1 & 1 \end{bmatrix} \quad (24)$$

The RGA is

$$\Lambda(G(0)) = \begin{bmatrix} -1.89 & -0.13 & 3.02 \\ 3.59 & 3.02 & -5.61 \\ -0.7 & -1.89 & 3.59 \end{bmatrix}$$

Assume Δ_1 , Δ_2 and Δ_3 have the same magnitude $|\Delta|$. The second row has the largest row sum ($\|\Lambda(G)\|_{1\infty} = 12.21$) and the worst combination of input uncertainty for an inverse-based controller is

$$\Delta_1 = \Delta_2 = -\Delta_3 = \Delta$$

We find

$$\text{diag}(G\Delta_I G^{-1}) = (-5.0\Delta, 12.2\Delta, -6.2\Delta)$$

Note that in this case we would arrive at the same worst case by considering row 1 and row 3. Because of this the worst case will always have Δ_1 and Δ_2 with the same sign and Δ_3 with different sign even if their magnitude is different.

We may in some cases arrive at a different conclusion by considering other frequencies. Also note that, unless an inverse-based controller is used, it is $\Lambda(C)$ rather than $\Lambda(G)$ which should be used to obtain the worst case.

4. Choice of Controller Structure

An important decision facing the engineer is the choice of the controller structure. Two extremes will be considered here

- diagonal controller
- inverse-based controller

The diagonal controller has advantages: It has fewer tuning parameters, is easier to understand and retune, and can be made failure tolerant more easily. These issues are not considered here. We want to decide which of the two choices above may result in the best multivariable controller. Based on the discussions above, Table 1 was prepared to assist the engineer in making this choice. The table should be used only as a rough guideline, since diagonal input uncertainty is the only source of uncertainty considered.

		$\max_d \gamma_d(G)$	
		Large	Small
$\ \Lambda\ _{i\infty}$	Large	----	Diagonal
	Small	Inverse-based (V=I: diagonal)	Inverse-based (diagonal)

Table 1. Guidelines for choice of best multivariable controller structure ("large" implies a comparison with one, typically >10).

Decoupling. One issue which has been extensively discussed in the chemical engineering literature is the use of decouplers for distillation columns (e.g., Luyben, 1970, Arkun et al., 1984). The idea of using a decoupler (D) is that the multivariable aspects are taken care of by the decoupler and tuning of the control system is reduced to a series of single loop problems. Let $K(s)$ denote these "single-loop" controllers. The overall controller C is

$$C(s) = DK(s) \quad (25)$$

Since $K(s)$ is diagonal the RGA of C and D are the same

$$\Lambda(C(s)) = \Lambda(D) \quad (26)$$

and if a steady-state decoupler is used (D is a constant matrix), the RGA of $C(s)$ will be the constant at all frequencies. The sensitivity of decouplers to decoupler errors has been discussed in the literature (e.g., Toijala (Waller) and Fagervik, 1972), and the observed sensitivity is easily explained from Result 1 (12). However, the most important reason for the robustness problems encountered with decouplers is probably input uncertainty. Recall from (17) that any controller with large RGA-elements is sensitive to input uncertainty. Decouplers generally have the same RGA-elements as the plant and should therefore not be used for plants with large RGA-elements. Let G_{diag} denote the matrix consisting of the diagonal elements in G. Then for the decouplers most commonly studied in the literature we find

$$\text{"Ideal Decoupling": } D = G^{-1} G_{\text{diag}}, \Lambda(C) = \Lambda^T(G) \quad (27)$$

$$\text{"Simplified Decoupling": } D = G^{-1}((G^{-1})_{\text{diag}})^{-1}, \Lambda(C) = \Lambda^T(G) \quad (28)$$

In both these cases the decoupler will lead to serious robustness problems if the plant has large RGA-elements. On the other hand, if "one-way" decoupling is used, then D is triangular and $\Lambda(C) = \Lambda(DK) = I$.

A "one-way" decoupler is therefore much less sensitive to input uncertainty.

5. Large RGA-elements are bad news

Let us now answer the two questions presented in the introduction.

A) Is a plant with large elements in the RGA always difficult to control?

Yes. This follows from Result 1 and 2. However, if the following conditions are satisfied, control may still be acceptable:

1. The transfer matrix elements are correlated and despite the large values in the RGA the plant is not likely to become singular, and
2. There exists a controller with small RGA-elements (e.g., a diagonal controller) which gives an acceptable response for all disturbances. This is the case if all disturbances are in the "good" direction, (i.e., $\gamma_d(G)$ is small despite the fact that $\gamma(G)$ is large).

Note that condition 2 implies that the plant is actually not ill-conditioned for the expected disturbances. We will give an example of such a case below (Diagonal controller for LV-distillation column).

B) Is a plant with small elements in the RGA always easy to control?

No. As seen from (17) an inverse-based controller results in serious "interactions" if there is input uncertainty and some of the offdiagonal elements are large. A diagonal controller gives large interactions even in the absence of uncertainty, if the plant is nearly triangular. (Consider, for example, the plant $G = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix}$ which has Λ

= I).

Let us also answer the following additional question:

C) Is a plant with a large condition number always difficult to control?

No. Based on the uncertainty descriptions investigated in this paper, the RGA rather than $\gamma(G)$ gives a measure of the plants sensitivity to uncertainty. We will show in an example below, that an inverse-based controller gives very good control for a plant with $\gamma(G) = 71$ even in the presence of uncertainty (Fig. 5).

6. Examples

The distillation column described in Table 2 is used as an example. The product compositions y_D and x_B are to be controlled by manipulating the reflux (L) and either the boilup (V) or the distillate flow (D). The column is assumed to have no dynamics. (This is, of course, not true. However, we make the crude assumption that the dynamics are given in terms of a single first order lag, which is exactly cancelled by a zero in the controller).

We show simulations for two different configurations of manipulated inputs.

- LV-configuration, $\gamma(G_{LV}) = 142$, $\lambda_{11}(G_{LV}) = 35$
- DV-configuration, $\gamma(G_{DV}) = 71$, $\lambda_{11}(G_{DV}) = 0.45$

and consider two controllers for each of these

- Inverse-based controller ($GC = I \cdot 0.7/s$)
- Diagonal controller

The gain of the controllers were adjusted to guarantee robust stability for relative uncertainty on each manipulated input with a magnitude bound

Binary separation, constant molar flows, feed liquid

Relative volatility	$\alpha = 1.5$
No. of theoretical trays	$N = 40$
Feed tray (1=reboiler)	$N_F = 21$
Feed composition	$z_F = 0.5$
Product compositions	$y_D^O = 0.99, x_B^O = 0.01$
Product rates	$D/F = B/F = 0.5$
Reflux rate	$L/F = 2.706$

Steady State Gain Matrices

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = G \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + E \begin{bmatrix} dF \\ dz_F \end{bmatrix}$$

LV-configuration

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} dL \\ dV \end{bmatrix} \quad G_{LV} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix}$$

DV-configuration

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} dV \\ dD \end{bmatrix} \quad G_{DV} = \begin{bmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{bmatrix}$$

Disturbance matrix

$$\text{(both configurations)} \quad E = \begin{bmatrix} 0.394 & 0.881 \\ 0.586 & 1.119 \end{bmatrix}$$

Table 2. Steady-state data for distillation column.

$$w_I(s) = 0.2 \frac{5s+1}{0.5s+1}$$

Robust stability is guaranteed if and only if (Skogestad & Morari, 1986a)

$$\mu(CG(I+CG)^{-1}) \leq 1/|w_I|, \quad \forall \omega \quad (29)$$

where the structured singular value μ is computed with respect to a diagonal matrix. Condition (29) is shown graphically in Fig. 2.

For each of these four systems the responses to two setpoint changes are shown

$$y_{s1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad y_{s2} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

The setpoint change y_{s1} has a large component in the "bad" direction corresponding to the low plant gain ($\gamma_d(G) = 110.7$ for the LV-configuration and $\gamma_d(G) = 54.9$ for the DV-configuration). y_{s2} has the same direction as a feed flow disturbance, and has $\gamma_d(G) = 11.8$ and 4.3 for the two configurations (Table 3).

The responses are shown both for the nominal case ($\Delta_I=0$) and with 20% relative uncertainty on each manipulated input

$$\Delta_I = \begin{bmatrix} 0.2 & 0 \\ 0 & -0.2 \end{bmatrix}$$

which give the following error terms (17) for GC when a inverse-based controller is used:

$$(G\Delta_I G^{-1})_{LV} = \begin{bmatrix} 35.1\Delta_1 - 34.1\Delta_2 & -27.7(\Delta_1 - \Delta_2) \\ 43.2(\Delta_1 - \Delta_2) & -34.1\Delta_1 + 35.1\Delta_2 \end{bmatrix} = \begin{bmatrix} 13.8 & -11.1 \\ 17.2 & -13.8 \end{bmatrix}$$

$$(G\Delta_I G^{-1})_{DV} = \begin{bmatrix} 0.45\Delta_1 + 0.55\Delta_2 & 0.45(\Delta_1 - \Delta_2) \\ -0.55(\Delta_1 - \Delta_2) & 0.55\Delta_1 + 0.45\Delta_2 \end{bmatrix} = \begin{bmatrix} -0.02 & 0.18 \\ -0.22 & 0.02 \end{bmatrix}$$

The simulations illustrate the following points:

- An inverse-based controller gives poor response when λ_{11} is large

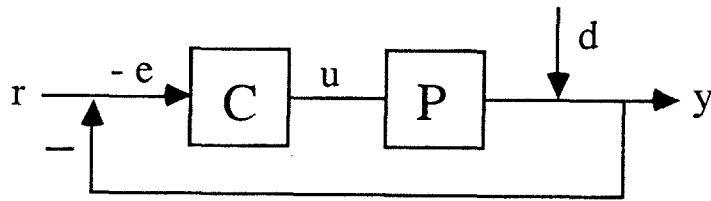


Figure 1. Classical feedback structure. d represents the effect of the disturbance on the output.

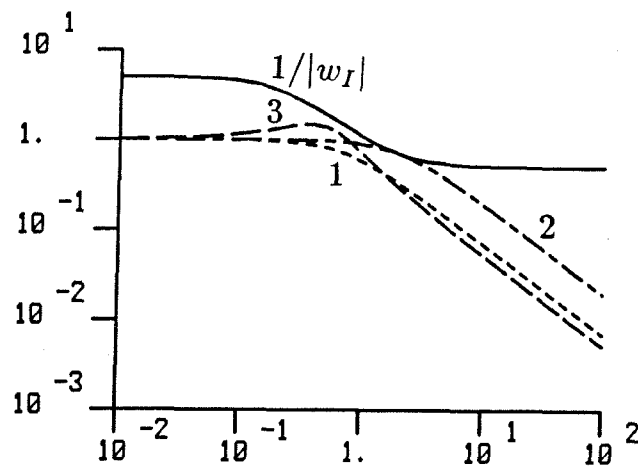


Figure 2. The controllers satisfy the Robust Stability condition (22). $\mu(CG((I+CG)^{-1}))$ is shown for 1: Inverse-controller for LV- and DV-configurations; 2: Diagonal LV-controller; 3: Diagonal DV-controller

	LV-	DV-
	Config.	Config.
RGA; $\lambda_{11} =$	35.1	0.45
$\ \Lambda \ _1 =$	138.3	2
Condition no., $\gamma(G) =$	141.7	70.8
<u>Disturbance condition no., $\gamma_d(G) =$</u>		
$d = F$ (feed rate)	11.8	4.3
$d = z_F$ (feed composition)	1.5	1.4
$d = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (setpoint in y_D)	110.7	54.9
$d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (setpoint in x_B)	88.5	44.6
<u>SV-decomposition, $G = U \Sigma V^H$</u>		
$U =$	$\begin{bmatrix} -0.625 & 0.781 \\ -0.781 & -0.625 \end{bmatrix}$	$\begin{bmatrix} -0.630 & 0.777 \\ -0.777 & -0.630 \end{bmatrix}$
$\Sigma =$	$\begin{bmatrix} 1.972 & 0 \\ 0 & 0.0139 \end{bmatrix}$	$\begin{bmatrix} 1.393 & 0 \\ 0 & 0.0197 \end{bmatrix}$
$V =$	$\begin{bmatrix} -0.707 & 0.708 \\ 0.708 & 0.707 \end{bmatrix}$	$\begin{bmatrix} 1.000 & -0.001 \\ 0.001 & 1.000 \end{bmatrix}$

Table 3. RGA , Condition numbers and SVD for distillation column.

and there is input uncertainty (Fig. 3).

- A diagonal controller cannot correct for the strong directionality of a plant with large RGA-elements (22). This results in responses which are strongly dependent on the disturbance (or setpoint) direction (Fig. 4). The response to a disturbance in F which has $\gamma_d(G) = 11.8$ is acceptable, but the response to the setpoint change y_{s1} is extremely sluggish. This system may be acceptable despite the large value of λ_{11} , provided setpoint changes are not important.
- An inverse-based controller may give very good response for an ill-conditioned plant even with diagonal input uncertainty, provided λ_{11} is small (Fig. 5).
- A diagonal controller may remove most of the directionality in the plant if $V \approx I$. However, "interactions" are still present because $U = \begin{bmatrix} -0.63 & 0.78 \\ -0.78 & -0.63 \end{bmatrix}$ is not diagonal (Fig. 6).

In practice, the improvement in response by using the DV-configuration instead of the LV-configuration, is probably less than indicated by these simulations. One reason is that the level loop for the condenser is not immediate as assumed in the simulations. This implies that a change in D (as opposed to L and V) does not affect compositions directly. The immediate effect of a change in D is a change in the condenser holdup (which does not effect composition). This change in holdup will eventually lead to a change in reflux, but this response may not be fast enough to counteract a large disturbance which may change the compositions considerably in a matter of minutes.

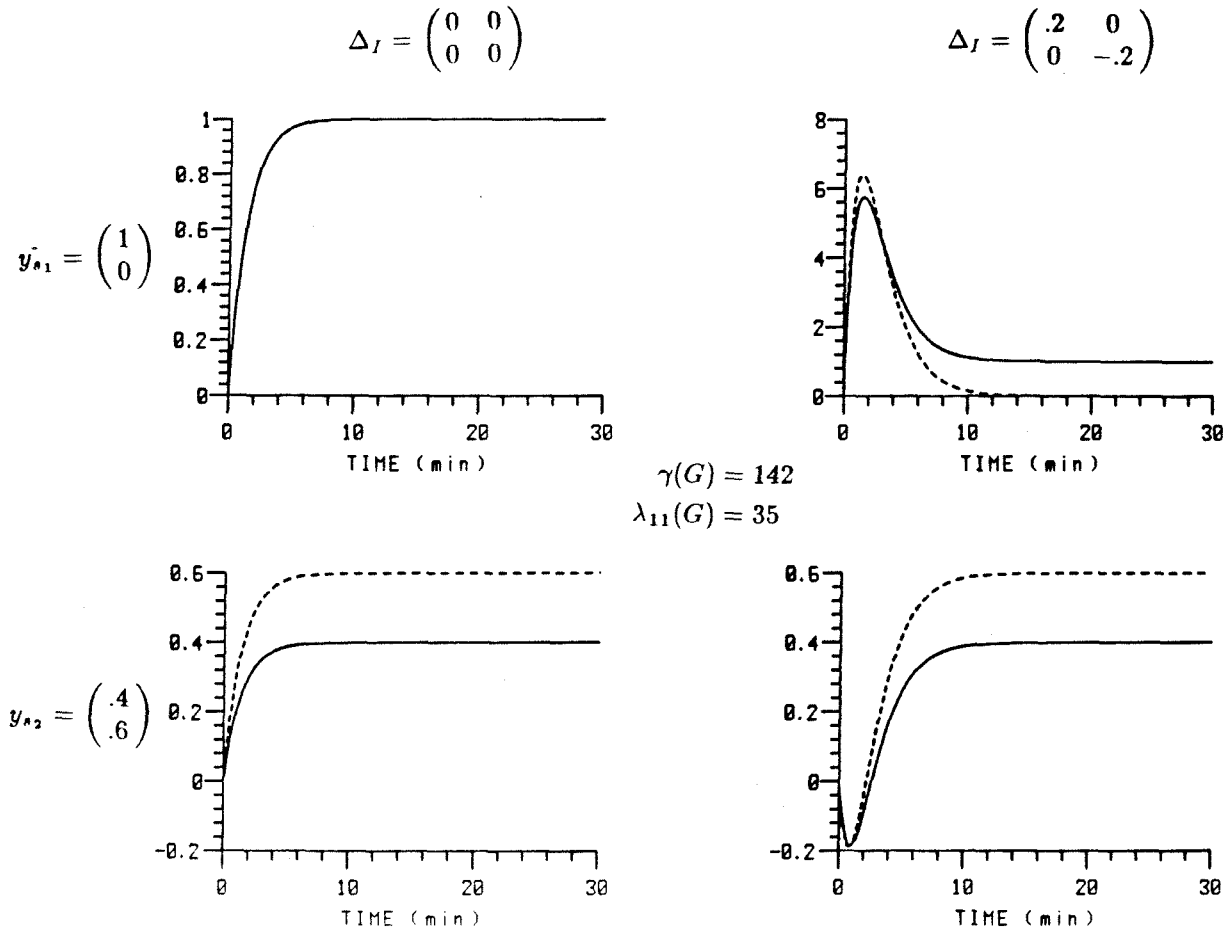


Figure 3. LV-configuration. Closed-loop responses y_1 and y_2 for inverse-based controller

$$C(s) = \frac{0.7}{s} G^{-1}_{LV} = \frac{0.7}{s} \begin{bmatrix} 39.94 & -31.49 \\ 39.43 & -32.00 \end{bmatrix}$$

An inverse-based controller gives poor response when λ_{11} is large and there is input uncertainty .

$$\Delta_I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Delta_I = \begin{pmatrix} .2 & 0 \\ 0 & -.2 \end{pmatrix}$$

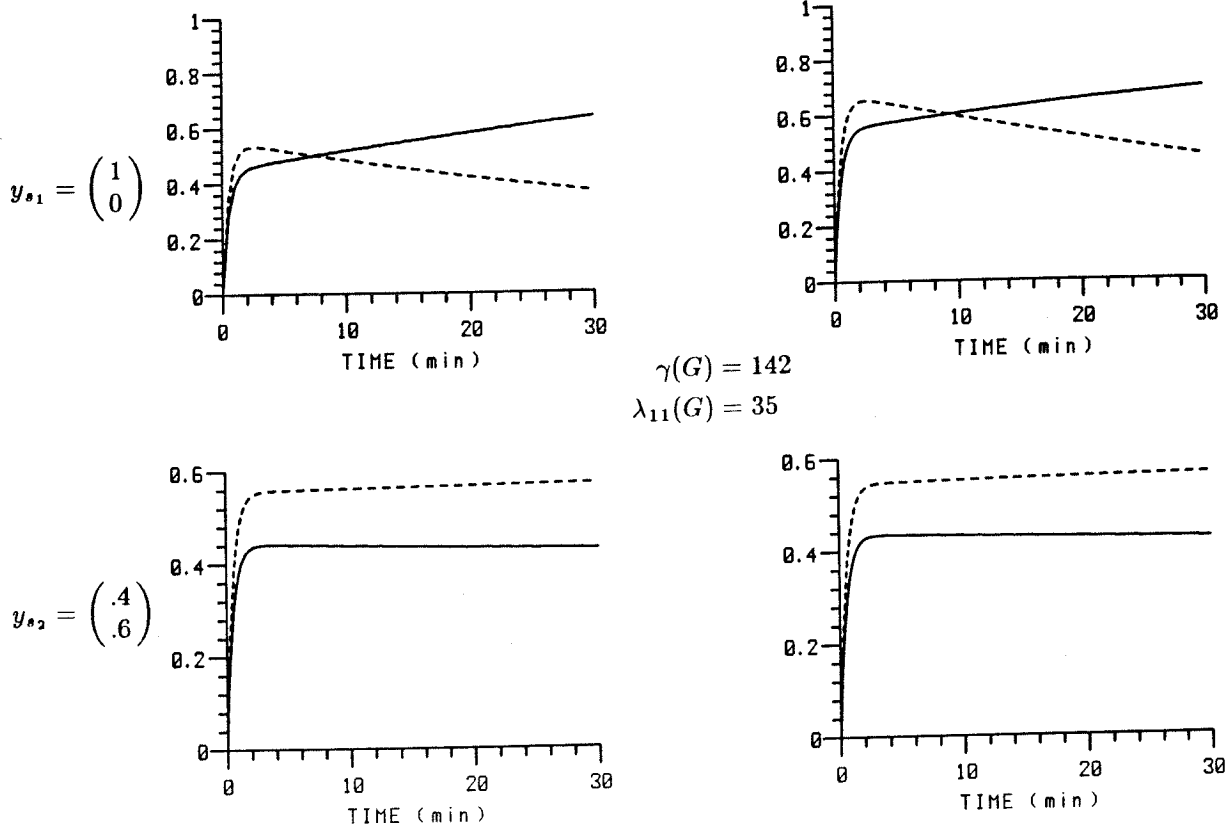


Figure 4. LV-configuration. Closed-loop responses y_1 and y_2 for diagonal controller

$$C(s) = \frac{1}{s} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

A diagonal controller which does not correct for the strong directionality of the plant gives responses which are strongly dependent on the disturbance (or setpoint) direction. The response to $y_{s2} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$ (disturbance in F, $\gamma_d(G) = 11.8$) is acceptable, but the response to the setpoint change y_{s1} ($\gamma_d(G) = 110.7$) is extremely sluggish. This system may be acceptable, despite the large value of λ_{11} , if setpoint changes are not important.

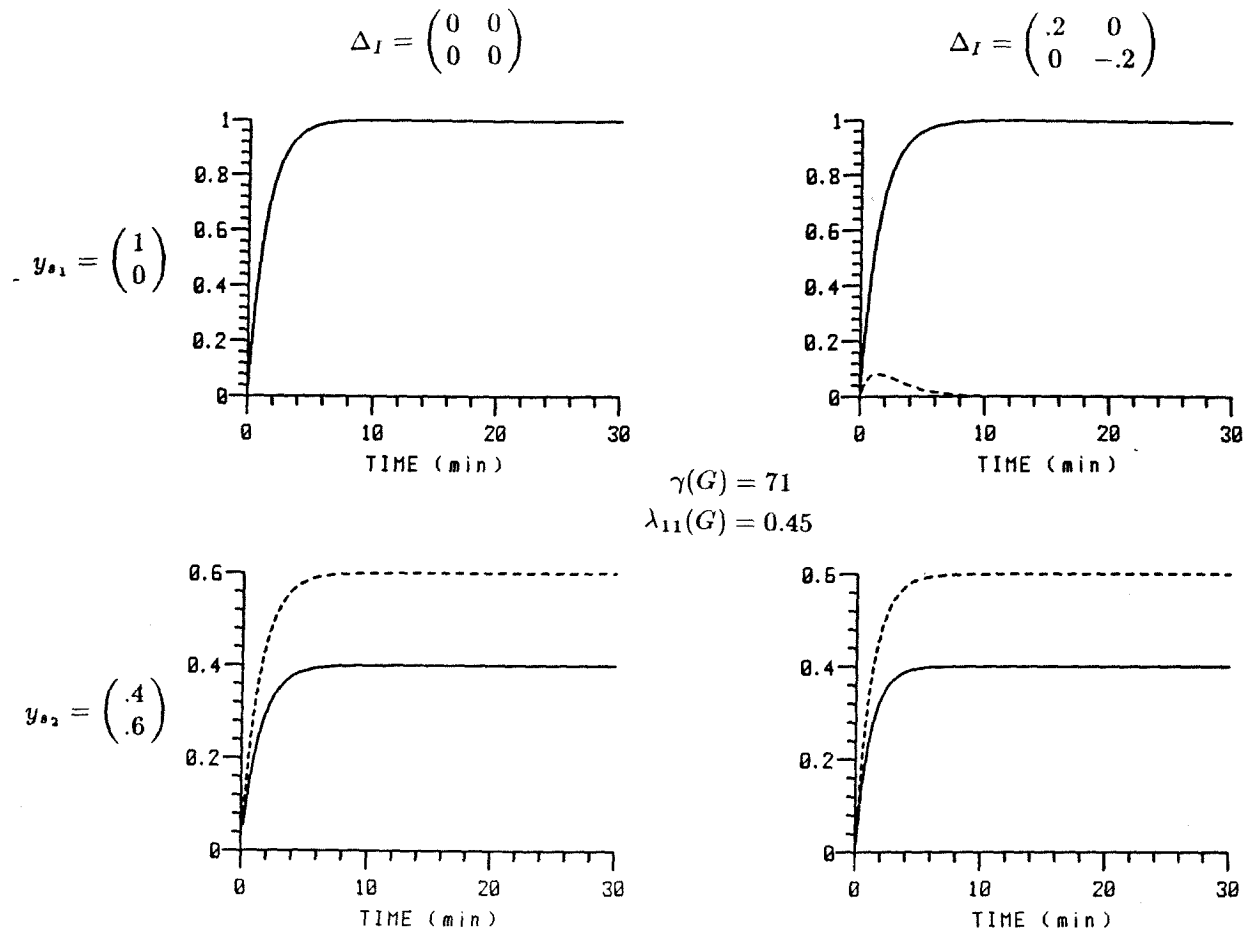


Figure 5. DV-configuration. Closed-loop responses y_1 and y_2 for inverse-based controller

$$C(s) = \frac{0.7}{s} G^{-1}_{DV} = \frac{0.7}{s} \begin{bmatrix} -0.5102 & -0.5102 \\ 39.43 & -32.00 \end{bmatrix}$$

An inverse-based controller may give very good response for an ill-conditioned plant, even with diagonal input uncertainty, provided λ_{11} is small .

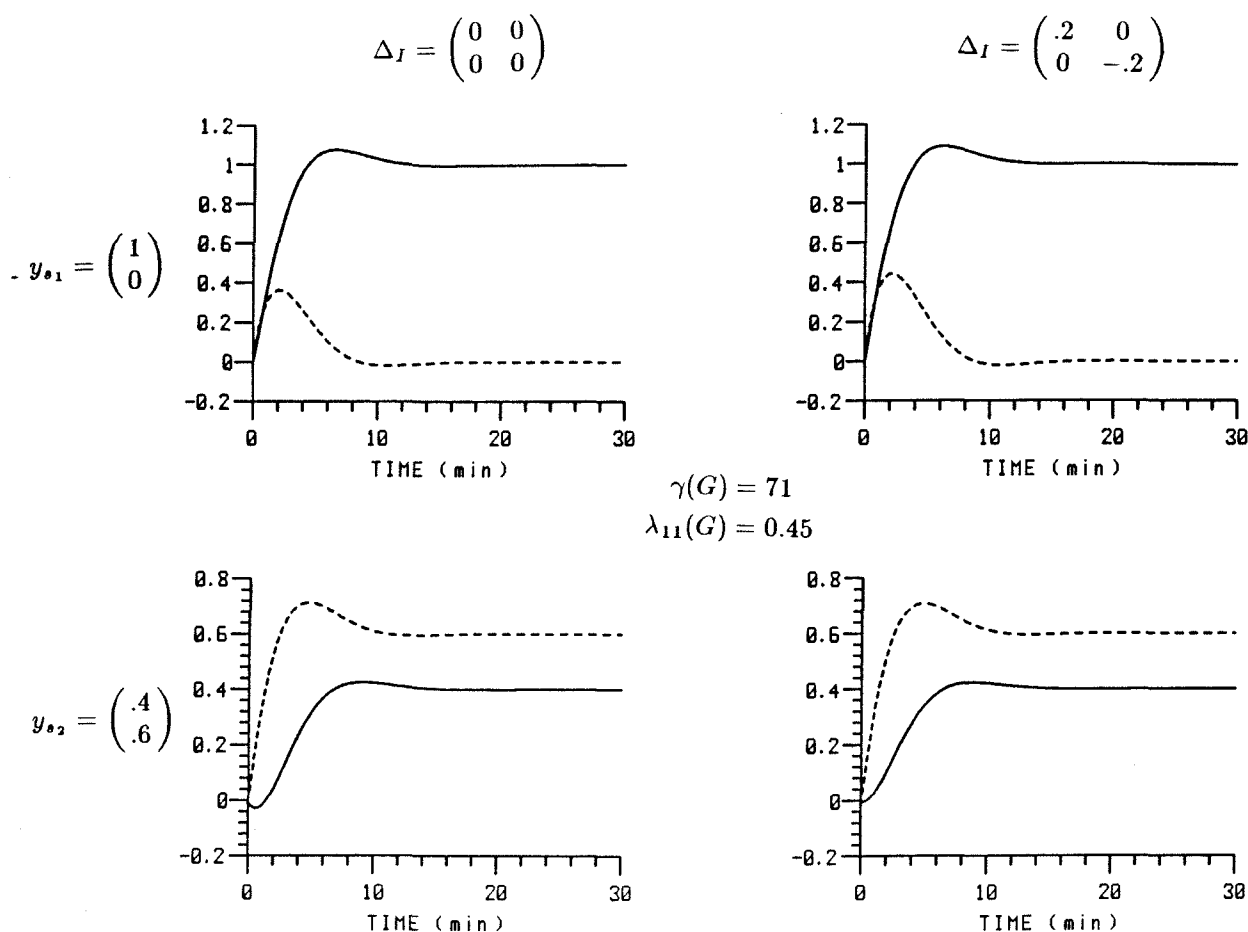


Figure 6. DV-configuration. Closed-loop responses y_1 and y_2 for diagonal controller

$$C(s) = \frac{-0.2}{s} \Lambda^{-1} = \frac{0.2}{s} \begin{bmatrix} -0.718 & 0 \\ 0 & -50.8 \end{bmatrix}$$

A diagonal controller may remove most of the directionality in the plant if $V \approx I$. However, "interactions" are still present

because $U = \begin{bmatrix} -0.63 & 0.78 \\ -0.78 & -0.63 \end{bmatrix}$ is not diagonal.

References

- Arkun, Y., B. Manousiouthakis and A. Palazoglu, Robustness Analysis of Process Control Systems. A Case Study of Decoupling Control in Distillation, Ind. & Eng. Chemistry Process Des. & Dev., 23, 93-101 (1984).
- Bristol, E. H., On a New Measure of Interactions for Multivariable Process Control, IEEE Trans. on Automatic Control., Vol. AC-11, 133-134 (1966).
- Grosdidier, P., M. Morari and B. R. Holt, Closed-Loop Properties from Steady-State Gain Information, Industrial and Eng. Chemistry Fund., 24, 221-235 (1985).
- Grosdidier, P. and M. Morari, Interaction Measures for Systems Under Decentralized Control, Automatica, 22, 3, 309-319 (1986).
- Luyben, W. L., Distillation Decoupling, AIChE Journal, 16, 198-203 (1970).
- McAvoy, T. J., Interaction Analysis, Instrumental Society of America, Research Triangle Park, NC (1983).
- Morari, M. Design of Resilient Processing Plants - III: A General Framework for the Assessment of Dynamic Resilience, Chemical Engineering Science, 38, 1881-1891 (1983).
- Nett, C. N. and V. Manousiouthakis, Euclidean Condition and Block Relative Gain: Connections, Conjectures and Clarifications, submitted to IEEE Trans. of Automatic Control (1986).
- Shinskey, F. G., Distillation Control, 2nd Edition, McGraw-Hill (1984).
- Skogestad, S. and M. Morari, Design of Resilient Processing Plants: Effect of Model Uncertainty on Dynamic Resilience, submitted to

Chemical Engineering Science (1986a).

Skogestad, S. and M. Morari, Understanding the Steady State Behavior of Distillation Columns, manuscript in preparation (1986b).

Skogestad, S. and M. Morari, Effect of Disturbance Directions on Closed Loop Performance, submitted to Industrial & Eng. Chemistry Process Des. & Dev. (1986c).

Toijala (Waller) K. and K. Fagervik, A Digital Simulation Study of Two-Point Feedback Control of Distillation Columns, Kemian Teollisuus, 29, 1-12 (1972).

Chapter VI

SOME NEW PROPERTIES OF THE STRUCTURED SINGULAR VALUE

Some New Properties of the Structured Singular Value

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Pasadena, California 91125

(818)356-4186

Submitted to IEEE Transactions on Automatic Control

January 1987

Abstract

Doyle et al. (1982) have shown that a necessary and sufficient condition for robust stability or robust performance in the H_∞ -framework may be formulated as a bound on the Structured Singular Value (μ) of a specific matrix M . M includes information on the system model, the controller, the model uncertainty and the performance specifications. Often it is desirable to express the robust stability and performance conditions as norm bounds on transfer matrices (T) which are of direct interest to the engineer, e.g., sensitivity or complementary sensitivity. This paper shows how to derive bounds on $\bar{\sigma}(T)$ from bounds on $\mu(M)$.

1. INTRODUCTION

The Structured Singular Value (μ) has proven to be a powerful tool for studying robustness of linear systems (Doyle et al., (1982)). Given a set of disturbances, a set of possible perturbations on the plant model (uncertainty) and a performance specification, μ is used to identify the "worst case" response for a given controller. Mathematically, a necessary and sufficient condition for achieving robust stability or robust performance is

$$\mu_{\Delta}(M) \leq k(\omega) \quad , \quad \forall \omega \quad (1)$$

In addition, the interconnection matrix M must be (internally) stable. The implications of (1) may not be easy to understand for the engineer. A simple bound on $\bar{\sigma}(T)$ may provide more insight. The goal of this paper is to derive such bounds. To this end assume that M can be written as a linear fractional transformation (LFT) of the transfer matrix T

$$M = N_{11} + N_{12}T(I - N_{22}T)^{-1}N_{21} \quad (2)$$

(A superscript on N (e.g., N^T) is sometimes used to denote a particular choice of T). The matrix N is used to derive the desired bound on $\bar{\sigma}(T)$. Since one objective is to use these bounds to assist the engineer in designing the controller (C), N should be independent of C . Typical choices of T include the loop transfer function PC and the closed-loop sensitivity (S) and complementary sensitivity (H) functions.

$$S = (I + PC)^{-1} \quad (3)$$

$$H = PC(I + PC)^{-1} \quad (4)$$

2. HOW TO FIND N

The first step is to find a LFT of M in terms of T (Eq. 2). In many cases this is easily done by inspection. In other cases the following procedure may be used:

- 1.) Write M as a LFT of C (Fig. 1):

$$M = G_{11} + G_{12}C(I - G_{22}C)^{-1}G_{21} \quad (5)$$

(The matrix G is easy to construct by inspection of the block diagram).

- 2.) Write the controller C as a LFT of the transfer matrix of interest (T).

$$C = J_{11} + J_{12}T(I - J_{22}T)^{-1}J_{21} \quad (6)$$

- 3.) Given G and J , N is easily derived (Fig. 1) because any interconnection of LFT's is again an LFT (Doyle, 1984).

$$\begin{aligned} N &= \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \\ &= \begin{bmatrix} G_{11} + G_{12}J_{11}(I - G_{22}J_{11})^{-1}G_{21} & G_{12}(I - J_{11}G_{22})^{-1}J_{12} \\ J_{21}(I - G_{22}J_{11})^{-1}G_{21} & J_{22} + J_{21}G_{22}(I - J_{11}G_{22})^{-1}J_{12} \end{bmatrix} \quad (7) \end{aligned}$$

For the special case $J_{11} = 0$ this reduces to

$$N = \begin{bmatrix} G_{11} & G_{12}J_{12} \\ J_{21}G_{21} & J_{22} + J_{21}G_{22}J_{12} \end{bmatrix} \quad (8)$$

Comments:

- If T is a *closed-loop* transfer function, (e.g., T is H or S), then $N_{22} = 0$ in (7) and (8). Readers familiar with the Q -parametrization of all stabilizing controllers (in which case $T = Q$) will not be surprised by this result (Doyle, 1984).
- Given N^H it is easy to derive N for other closed-loop transfer functions. For example, note that H is a LFT of S

$$H = I - S \quad (9)$$

and N^S is derived from N^H using (7). For $N_{22}^H = 0$ (which is generally the case) we find

$$N^S = \begin{bmatrix} N_{11}^H + N_{12}^H N_{21}^H & -N_{12}^H \\ N_{21}^H & 0 \end{bmatrix} \quad (10)$$

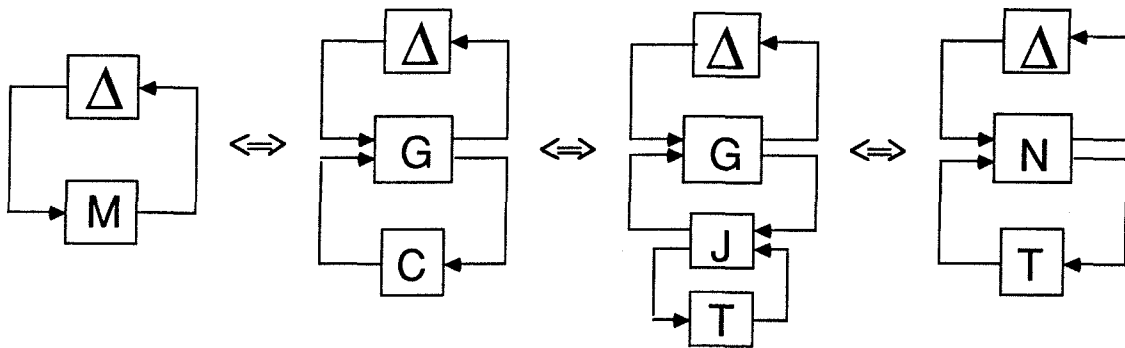


Figure 1. Equivalent representations of system M with perturbation Δ .

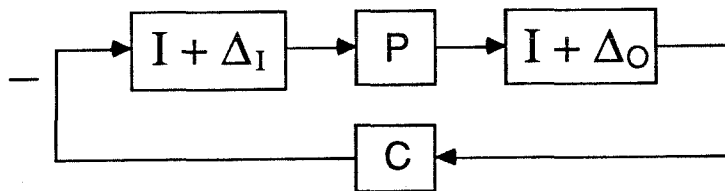


Figure 2. Plant with multiplicative input (Δ_I) and output (Δ_O) uncertainty .

Example 1. Input and Output Uncertainty (Fig. 2). For a system with input and output multiplicative uncertainty of magnitude $\bar{\sigma}(\Delta_I) < w(\omega)$ and $\bar{\sigma}(\Delta_O) < w(\omega)$ a necessary and sufficient condition for Robust Stability (R.S.) is (Doyle, 1982):

$$\mu(M) \leq w^{-1}(\omega) \quad , \quad \forall \omega \quad (11)$$

$$M = \begin{bmatrix} -(I + CP)^{-1}CP & -(I + CP)^{-1}C \\ (I + PC)^{-1}P & -(I + PC)^{-1}PC \end{bmatrix} \quad (12)$$

μ is computed with respect to the structure $\Delta = \text{diag}\{\Delta_I, \Delta_O\}$. Let us construct M as an LFT of H using the three-step procedure:

- 1.) Using Fig. 1B with $\Delta = \text{diag}\{\Delta_I, \Delta_O\}$ we find (constructing G directly is simpler than starting from (12):

$$G = \begin{bmatrix} O & O & -I \\ P & O & -P \\ P & I & -P \end{bmatrix} \quad (13)$$

Note that G_{11} is the upper left 2×2 block of G corresponding to Δ .

- 2.)

$$T = H : \quad J = \begin{bmatrix} O & P^{-1} \\ I & I \end{bmatrix} \quad (14)$$

- 3.) Substituting (13) and (14) into (7) yields

$$N_{11}^H = \begin{bmatrix} O & O \\ P & O \end{bmatrix}, \quad N_{12}^H = \begin{bmatrix} -P^{-1} \\ -I \end{bmatrix}, \quad N_{21}^H = [P \quad I], \quad N_{22}^H = 0 \quad (15)$$

To find M as a LFT of S , use N^H and Eq. (10). We get:

$$N_{11}^S = \begin{bmatrix} -I & -P^{-1} \\ 0 & -I \end{bmatrix}, \quad N_{12}^S = \begin{bmatrix} P^{-1} \\ I \end{bmatrix}, \quad N_{21}^S = [P \quad I], \quad N_{22}^S = 0 \quad (16)$$

3. THE STRUCTURED SINGULAR VALUE μ

Let M be a square complex matrix. $\mu(M)$ is defined such that $\mu^{-1}(M)$ is equal to the smallest $\bar{\sigma}(\Delta)$ needed to make $(I + \Delta M)$ singular, i.e.,

$$\mu^{-1}(M) = \min_{\delta} \{ \delta | \det(I + \Delta M) = 0 \text{ for some } \Delta, \bar{\sigma}(\Delta) < \delta \} \quad (17)$$

(If M is a transfer matrix this definition applies frequency-by-frequency). Δ is a block-diagonal perturbation matrix with a given *structure*

$$\Delta = \begin{bmatrix} \Delta_1 & & \\ & \ddots & \\ & & \Delta_n \end{bmatrix}, \quad \bar{\sigma}(\Delta_i) < \delta, \quad \forall i \quad (18)$$

Δ is allowed to be *any* complex matrix satisfying $\bar{\sigma}(\Delta) < \delta$. (It turns out that Δ may be restricted to being unitary without changing $\mu(M)$ (Doyle, 1982)). $\mu(M)$ depends on both the matrix M and the *structure* of the perturbations Δ . To write this more explicitly we will sometimes use the notation $\mu(M) = \mu_\Delta(M)$. An equivalent statement of (17) which is more useful for our purposes is the following:

$$\begin{aligned} \det(I + \Delta M) &\neq 0, \quad \forall \Delta (\bar{\sigma}(\Delta) < \delta) \\ \Leftrightarrow \rho(\Delta M) &\leq 1, \quad \forall \Delta (\bar{\sigma}(\Delta) < \delta) \\ \Leftrightarrow \mu_\Delta(M) &\leq 1/\delta \end{aligned} \quad (19)$$

The reader is referred to Doyle (1982) for further properties and computational aspects of μ .

4. NEW PROPERTIES OF μ

The results in this section apply to any complex matrices although in most cases these will be transfer matrices. Proofs of all results are given in Appendix.

Theorem 1. *Let M be written as a LFT of T :*

$$M = N_{11} + N_{12}T(I - N_{22}T)^{-1}N_{21} \quad (20)$$

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

$$\mu_\Delta(M) \leq k \quad (21)$$

if

$$\bar{\sigma}(T) \leq c_T \quad (22)$$

where c_T solves

$$\mu_{\tilde{\Delta}} \begin{bmatrix} N_{11} & N_{12} \\ kc_T N_{21} & kc_T N_{22} \end{bmatrix} = k \quad (23)$$

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

c_T is in general found numerically using the implicit expression (23). This search is straightforward since the value of μ increases monotonically with c_T . An explicit expression exists for the special case $N_{11} = N_{22} = 0$:

Lemma 1. *For $N_{11} = N_{22} = 0$ there exists an explicit formula for c_T :*

$$c_T = k \mu_{\tilde{\Delta}}^{-2} \begin{bmatrix} 0 & N_{12} \\ N_{21} & 0 \end{bmatrix} \quad (24)$$

Comments on Theorem 1:

1. If M is a transfer matrix, conditions (21)-(23) apply on a frequency-by-frequency basis (with $k(\omega)$ and $c_T(\omega)$).
2. The bound $\mu(M) \leq k$ may result from a robust stability or robust performance condition, and T may be a particular transfer function we want to bound (for example, T is H or S as in Example 1). In such cases M is a (internally) stable transfer matrix and the condition $\det(I - N_{22}T) \neq 0, \forall \omega$ is trivially satisfied since M has no $j\omega$ -axis poles. The condition that M be stable puts additional restrictions on the *allowable* T (although T does *not* necessarily have to be stable, for example, if $T = PC$). If $T = H$ or $T = S$ we must require T stable.
3. The condition $\mu_{\Delta}(N_{11}) < k$ is required for the existence of a solution $c_T > 0$ to (23). If $\mu(M) < k(\omega)$ is a robust stability (performance) condition, then the condition $\mu(N_{11}) < k(\omega)$ is equivalent to requiring that the robust stability (performance) condition be satisfied for $T = 0$ at this frequency.
4. Condition (22) is necessary and sufficient for (21) if we want (21) to be satisfied for *all* T 's satisfying $\bar{\sigma}(T) \leq c_T$. (This follows directly from the proof of the theorem). However, in most cases we are interested only in a specific M (and a specific T), and condition (22) is only sufficient for (21).

5. The previous comment implies that c_T gives the least conservative bound which may be derived on $\bar{\sigma}(T)$ given (21). (There will always exist a T with the *same* $\bar{\sigma}(T)$ for which condition (22) is “tight”).
6. μ in (23) is computed based on the structure of Δ *and* of T . The least restrictive bound on $\bar{\sigma}(T)$ (c_T large) is found if $T = tI$ is assumed, and the most restrictive bound (c_T small) is found if T is a full matrix. This may seem counter-intuitive; we expect to “pay” something for restricting $T = tI$. The reason it does not work out this way is that we are considering how large “perturbations” from $T = 0$ we may allow in T . By restricting $T = tI$, the class of perturbations is restricted, and the magnitude of the perturbations may be larger (we also arrive at the same conclusion by observing that tI will be included as a special case for any structure assumed for T).
7. Theorem 1 may be used to derive a bound on any transfer matrix T which is related to M through a linear fractional transformation (LFT). Note that these bounds (e.g., on $\bar{\sigma}(H)$ and $\bar{\sigma}(S)$) may be *combined* over different frequency-ranges since Theorem 1 applies on a frequency-by-frequency basis. This provides a more powerful method for deriving simple robustness bounds than the approach suggested by Postlethwaite and Foo (1985). Firstly, it is possible to take advantage of the structure of T using our approach. Secondly, the initial formulation of the uncertainty bounds is much more straightforward: Postlethwaite and Foo’s approach is to force the uncertainty of the plant to fit a specific single norm-bounded perturbations (e.g., use Δ_0 (Fig. 2) to derive a bound on $\bar{\sigma}(H)$). However, it is not clear how a tight bound on this perturbation may be derived if there are several sources of uncertainty. Our approach is to describe the uncertainty as it occurs physically using multiple perturbations (Δ ’s). This yields a necessary and sufficient μ -condition which subsequently is used to derive the tightest possible bound on $\bar{\sigma}(T)$.

The following property of μ is derived using Theorem 1 and Lemma 1.

Theorem 2. *Let $\tilde{\Delta} = \text{diag}\{\Delta, T\}$. Then*

$$\mu_{\Delta}(ATB) \leq \bar{\sigma}(T)\mu_{\tilde{\Delta}}^2 \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \quad (25)$$

((25) also applies on a frequency-by-frequency basis). Note that ATB and T are square matrices, while A and B may be non-square. (25) is a generalization of property (c) in Doyle (1982):

Assume Δ and T have the same structure. Then

$$\mu_{\Delta}(AT) \leq \bar{\sigma}(T)\mu_{\Delta}(A) \quad (26)$$

(The property is stated incorrectly in Doyle (1982) since the condition that Δ and T have the same structure is left out).

Special Cases of Theorem 2: In some cases Theorem 2 may be simplified by writing $\mu^2 \begin{bmatrix} O & A \\ B & O \end{bmatrix}$ in terms of other quantities:

Case	$\mu_{\tilde{\Delta}}^2 \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} =$
1. Δ and T are both full matrices:	$\bar{\sigma}(A)\bar{\sigma}(B)$
2. $T = tI$:	$\mu_{\Delta}(AB)$
3. $\Delta = \delta I$:	$\mu_T(BA)$
4. $\Delta = \delta I, T = tI$:	$\rho(AB) = \rho(BA)$
5. $B = I$:	$\mu_{\Delta T}(A)$

Case 1 can be interpreted as $\bar{\sigma}(ATB) \leq \bar{\sigma}(A)\bar{\sigma}(T)\bar{\sigma}(B)$. Case 5 yields

$$\mu_{\Delta}(AT) = \mu_{\Delta}(TA) \leq \bar{\sigma}(T)\mu_{\Delta T}(A) \quad (27)$$

where " ΔT " denotes the structure of the matrix ΔT . For ΔT to have a well-defined structure, the structures of Δ and T must "match up" in some consistent manner.

In particular, Δ and T must have the same size which implies that A must be square. Some examples are:

- (i) Δ and T have the same structure: " ΔT " = $\Delta = T$ (property (c) in Doyle (1982)).
- (ii) Δ diagonal: " ΔT " = T .
- (iii) $\Delta = \text{diag}\{\Delta_1, \Delta_2\}$, $T = \text{diag}\{T_{11}, T_{12}, T_2\}$ where Δ_2 and T_2 have the same size: " ΔT " = Δ .

5. EXAMPLES

The following two examples demonstrate the use of Theorems 1 and 2.

Example 2. Input Uncertainty. The robust stability condition for Fig. 2 when there is only input uncertainty of magnitude $\bar{\sigma}(\Delta_I) < w(\omega)$ is

$$\text{R.S.} \Leftrightarrow \mu(P^{-1}HP) \leq w^{-1}(\omega) \quad (28)$$

Here μ is computed with respect to the structure of Δ_I which may be a diagonal matrix. The least conservative bound on $\bar{\sigma}(H)$ which may be derived from (28) is found using Theorem 2:

$$\text{R.S.} \Leftarrow \bar{\sigma}(H) \leq 1/\mu^2 \begin{bmatrix} 0 & P^{-1} \\ P & O \end{bmatrix} w(\omega) \quad (29)$$

μ in (29) is computed with respect to the structure $\text{diag}\{\Delta_I, H\}$. Note the following special cases:

- (i) Δ_I and H are both "full" matrices: $\mu^2 \begin{bmatrix} O & P^{-1} \\ P & O \end{bmatrix} = \gamma(P)$ where $\gamma(P) = \bar{\sigma}(P)/\underline{\sigma}(P)$ is the condition number of the plant.
- (ii) $H = hI$: $\mu^2 \begin{bmatrix} O & P^{-1} \\ P & O \end{bmatrix} = 1$

Example 3. Robust Performance for SISO-plant. Robust Performance (R.P.) is achieved if the performance condition $\bar{\sigma}(S) = |S| < 1/|w_P|$ is satisfied for all possible plants. The set of possible plants is given in terms of multiplicative uncertainty of magnitude $w_O(s)$. Using the results of Doyle et al. (1982) we derive

$$\text{R.P.} \Leftrightarrow \mu(M) \leq 1 \quad \forall \omega \quad (30a)$$

where

$$M = \begin{bmatrix} w_O H & w_O H \\ w_P S & w_P S \end{bmatrix} \quad (30b)$$

μ in (30a) is computed with respect to the diagonal 2×2 -matrix $diag\{\Delta_O, \Delta_P\}$.

Bounds on $\bar{\sigma}(H) = |H|$ and $\bar{\sigma}(S) = |S|$ are easily derived using Theorem 1 with

$k = 1$. Write M as a LFT of H :

$$N_{11}^H = \begin{bmatrix} 0 & 0 \\ w_P & w_P \end{bmatrix} \quad N_{12}^H = \begin{bmatrix} w_O \\ -w_P \end{bmatrix} \quad N_{21}^H = [1 \quad 1] \quad N_{22}^H = 0 \quad (31)$$

Theorem 1 gives

$$\text{R.P.} \quad \Leftarrow \quad |H| \leq c_H \quad \forall \omega \quad (32a)$$

where c_H at each frequency solves

$$\mu \begin{bmatrix} 0 & 0 & w_O \\ w_P & w_P & -w_P \\ c_H & c_H & 0 \end{bmatrix} = 1 \quad (32b)$$

μ in (32b) is computed with respect to the structure $diag\{\Delta_O, \Delta_P, H\}$, i.e., a diagonal 3×3 matrix. Note that (32) is independent of the plant model (P). However, M (and therefore H) must be *stable*, and this implicitly makes the allowable H 's dependent on P . An analytic expression may be derived for c_H for this simple case.

We find:

$$\text{R.P.} \quad \Leftarrow \quad |H| \leq c_H = \frac{1 - |w_P|}{|w_O| + |w_P|} \quad (33)$$

Similarly, a condition in terms of S is derived

$$\text{R.P.} \quad \Leftarrow \quad |S| \leq c_S = \frac{1 - |w_O|}{|w_O| + |w_P|} \quad (34)$$

The expressions for c_H and c_S in (33) and (34) are most easily derived from the identity

$$\mu \begin{bmatrix} w_O H & w_O H \\ w_P S & w_P S \end{bmatrix} = |w_O H| + |w_P S| \quad (35)$$

combined with the triangle inequality (e.g., use $|S| = |1 - H| \leq 1 + |H|$ to derive (33)). Note that (33) is impossible to satisfy at low frequencies where tight

performance is desired and $|w_P|$ is larger than one (corresponds to $\mu(N_{11}) > k$ in Theorem 1). Similarly, (34) is impossible to satisfy at high frequencies where the uncertainty exceeds 100% and $|w_O|$ is larger than one. However, we can combine bounds: (30) is satisfied if (34) is satisfied at low frequencies and (33) at high frequencies. The bounds (33) and (34) (even when combined) tend to be conservative around cross-over where $|H|$ and $|S|$ have similar magnitude. This means that there will be systems which satisfy (30), but do not satisfy (33) and (34).

Conditions (33) and (34) are shown graphically in Fig.3A for the choice $w_O(s) = 0.2(\frac{s}{2} + 1)$ and $w_P(s) = 0.5(1 + \frac{1}{s})$. Assume that the plant is minimum phase such that $H = \frac{1}{s+1}$ is an allowable (stable) closed-loop transfer function. This corresponds to a nominal first-order response with time constant 1. This choice is seen to satisfy (33) for $\omega > 1.2$ and (34) for $\omega < 2$ (Fig. 3B). Consequently, (30) is satisfied at all frequencies and robust performance is guaranteed.

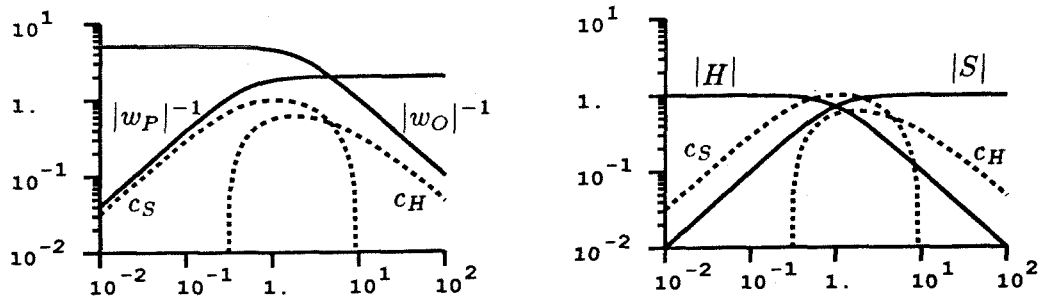


Figure 3. Graphical representation of conditions (33) and (34). R.P. is guaranteed since

$$|S| < c_S \text{ for } \omega < 2 \text{ and } |H| < c_H \text{ for } \omega > 1.4.$$

Acknowledgements. We are thankful to John C. Doyle for numerous useful discussions and remarks. The idea of treating T as a "perturbation" (which subsequently led to the derivation of Theorem 1) was first presented by Grosdidier and Morari (1986) in their derivation of the μ -interaction measure. Partial financial support from the National Science Foundation is gratefully acknowledged.

References

- Doyle, J. C., Analysis of Feedback Systems with Structured Uncertainties, Inst. Electrical Engrs. Proc., Pt. D., *129*, D(6), 242-247 (1982).
- Doyle, J. C., J. E. Wall and G. Stein, Performance and Robustness Analysis for Structured Uncertainty, Proc. IEEE Conference on Decision and Control, Orlando, FL (1982).
- Doyle, J. C., Lecture Notes for ONR/Honeywell Workshop, Oct. 8-10, 1984, Minneapolis, MN.
- Grosdidier, P. and M. Morari, Interaction Measures for Systems Under Decentralized Control. *Automatica*, *22*, 3, 309-319 (1986).
- Postlethwaite, I. and Y. K. Foo, Robustness with Simultaneous Pole and Zero Movements Across the $j\omega$ -Axis, *Automatica*, *21*, 4, 433-443 (1985).

Appendix. Proof of Theorems

Theorem 1. The theorem follows directly from the definition of μ after some algebra: Assume that T is defined such that $\bar{\sigma}(T) < c_T$. Then at each frequency the following holds

$$\begin{aligned} \mu_\Delta(M) &\leq k(\omega), \quad \forall T \ (\bar{\sigma}(T) < c_T) \\ \Leftrightarrow \det(I + \Delta M) &\neq 0, \quad \forall \Delta \ (\bar{\sigma}(\Delta) < 1/k), \quad \forall T \end{aligned} \quad (A1)$$

$$\Leftrightarrow \det \begin{bmatrix} I + \Delta N_{11} & \Delta N_{12} \\ -TN_{21} & I - TN_{22} \end{bmatrix} \neq 0 \quad \forall \Delta, \quad \forall T \quad (A2)$$

$$\Leftrightarrow \det(I + \begin{bmatrix} k\Delta & 0 \\ 0 & -\frac{1}{c_T}T \end{bmatrix} \begin{bmatrix} \frac{1}{k}N_{11} & \frac{1}{k}N_{12} \\ c_T N_{21} & c_T N_{22} \end{bmatrix}) \neq 0 \quad \forall \Delta, \quad \forall T$$

$$\Leftrightarrow \mu_{\tilde{\Delta}} \begin{bmatrix} \frac{1}{k}N_{11} & \frac{1}{k}N_{12} \\ c_T N_{21} & c_T N_{22} \end{bmatrix} \leq 1$$

$$\Leftrightarrow \mu_{\tilde{\Delta}} \begin{bmatrix} N_{11} & N_{12} \\ kc_T N_{21} & kc_T N_{22} \end{bmatrix} \leq k(\omega)$$

The step from (A1) to (A2) follows $M = N_{11} + N_{12}(I - TN_{22})^{-1}TN_{21}$ and Schurs formula

$$\det(A - BD^{-1}C) = \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} / \det D \quad (A3)$$

and the assumption $\det D = \det(I - TN_{22}) \neq 0$

Lemma 1. An equivalent statement of the lemma is: Let $\tilde{\Delta} = \text{diag}\{\Delta_1, \Delta_2\}$ where Δ_1 and Δ_2 have the same size as N_{11} and N_{22} , respectively. (Δ_1 and Δ_2 may have additional structure). Then:

$$\mu_{\tilde{\Delta}} \begin{bmatrix} 0 & N_{12} \\ cN_{21} & 0 \end{bmatrix} = \sqrt{c} \mu_{\tilde{\Delta}} \begin{bmatrix} 0 & N_{12} \\ N_{21} & 0 \end{bmatrix} \quad (A4)$$

Proof of (A4):

$$\mu_{\tilde{\Delta}} \begin{bmatrix} 0 & N_{12} \\ cN_{21} & 0 \end{bmatrix} \leq 1/k_1 \quad (A5)$$

$$\Leftrightarrow \det(I + k_1 \begin{bmatrix} \Delta_1 & \\ & \Delta_2 \end{bmatrix} \begin{bmatrix} 0 & N_{12} \\ cN_{21} & 0 \end{bmatrix}) \neq 0$$

$$\Leftrightarrow \det \begin{bmatrix} I & k_1 \Delta_1 N_{12} \\ k_1 c \Delta_2 N_{21} & I \end{bmatrix} \neq 0 \quad (A6)$$

$$\Leftrightarrow \det(I - k_1^2 c \Delta_1 N_{12} \Delta_2 N_{21}) \neq 0 \quad (A7)$$

$$\Leftrightarrow \det \begin{bmatrix} I & \sqrt{k_1^2 c} \Delta_1 N_{12} \\ \sqrt{k_1^2 c} \Delta_2 N_{21} & I \end{bmatrix} \neq 0 \quad (A8)$$

$$\Leftrightarrow \mu_{\tilde{\Delta}} \begin{bmatrix} 0 & N_{12} \\ N_{21} & 0 \end{bmatrix} \leq 1/\sqrt{k_1^2 c} \quad (\text{A9})$$

The conditions involving $\det(\cdot) \neq 0$ must hold for $\forall \Delta_1$ s.t. $\bar{\sigma}(\Delta_1) < 1$ and $\forall \Delta_2$ s.t. $\bar{\sigma}(\Delta_2) < 1$. The step from (A6) to (A7) and back to (A8) follows from (A3). Since (A5) and (A9) must hold for *any* value of k_1 , (A4) follows.

Theorem 2. From Theorem 1 and Lemma 1 for the case $N_{11} = N_{22} = 0$:

$$\mu_{\Delta}(N_{12}TN_{21}) \leq k \quad \text{if} \quad \bar{\sigma}(T)\mu_{\tilde{\Delta}}^2 \begin{bmatrix} 0 & N_{12} \\ N_{21} & 0 \end{bmatrix} \leq k \quad (\text{A10})$$

Since (A10) holds for *any* choice of k it is equivalent to

$$\mu_{\Delta}(N_{12}TN_{21}) \leq \bar{\sigma}(T)\mu_{\tilde{\Delta}}^2 \begin{bmatrix} 0 & N_{12} \\ N_{21} & 0 \end{bmatrix}$$

Theorem 2 follows by choosing $N_{12} = A, N_{21} = B$.

Special Cases of Theorem 2. Let Δ_1 and Δ_2 have the same *structure* as Δ and T in Theorem 2. Define $\tilde{\Delta} = \text{diag}\{\Delta_1, \Delta_2\}$. Then

$$\mu_{\tilde{\Delta}}^2 \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix} \leq 1/k \quad (\text{A11})$$

$$\Leftrightarrow \mu_{\tilde{\Delta}} \begin{bmatrix} 0 & kA \\ B & 0 \end{bmatrix} \leq 1 \quad (\text{A12})$$

$$\Leftrightarrow \det\left(I + \begin{bmatrix} \Delta_1 & \\ & \Delta_2 \end{bmatrix} \begin{bmatrix} 0 & kA \\ B & 0 \end{bmatrix}\right) \neq 0 \quad \forall \Delta_1, \Delta_2$$

$$\Leftrightarrow \det(I - k\Delta_2 B \Delta_1 A) = \det(I - k\Delta_1 A \Delta_2 B) \neq 0 \quad \forall \Delta_1, \Delta_2$$

$$\Leftrightarrow \mu_{\Delta_2}(B \Delta_1 A) \leq 1/k \quad \forall \Delta_1 \quad (\text{A13})$$

$$\Leftrightarrow \mu_{\Delta_1}(A \Delta_2 B) \leq 1/k \quad \forall \Delta_2 \quad (\text{A14})$$

$$\Leftrightarrow \rho(\Delta_1 A \Delta_2 B) = \rho(\Delta_2 B \Delta_1 A) \leq 1/k \quad \forall \Delta_1, \Delta_2 \quad (\text{A15})$$

By $\forall \Delta_i$ is understood all Δ_i s.t. $\bar{\sigma}(\Delta_i) < 1$. The step from (A11) to (A12) follows from (A4).

Case(1): Follows from (A15): Use the SVD of $A = U_A \Sigma_A V_A^H$ and $B = U_B \Sigma_B V_B^H$. Since Δ_1 and Δ_2 are “full”, Δ_1 may be chosen such that $\Delta_1 U_A = V_B$ and Δ_2 such that $V_A^H \Delta_2 = U_B^H$. Then $\rho(\Delta_1 A \Delta_2 B) = \rho(V_B \Sigma_1 \Sigma_2 V_B^H) = \rho(\Sigma_1 \Sigma_2) = \bar{\sigma}(A)\bar{\sigma}(B)$. (The generalization to the case then A and B are non-square is straightforward and involves “lining up” the directions corresponding to $\bar{\sigma}(A)$ and $\bar{\sigma}(B)$).

Case(2): Follows from (A14).

Case(3): Follows from (A13).

Case(4),(5): Follow from (A15).

Chapter VII

ROBUST PERFORMANCE OF DECENTRALIZED CONTROL SYSTEMS BY INDEPENDENT DESIGNS

**ROBUST PERFORMANCE
OF DECENTRALIZED CONTROL SYSTEMS
BY INDEPENDENT DESIGNS**

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January, 1987

Submitted to Automatica

Abstract

Decentralized control systems have fewer tuning parameters, are easier to understand and tune, and are more easily made failure tolerant than general multivariable control systems. In this paper the decentralized control problem is formulated as a series of independent designs. Simple bounds on these individual designs are derived, which when satisfied, guarantee robust performance of the overall system. The results provide a generalization of the μ -Interaction Measure introduced by Grosdidier and Morari (1986).

1. INTRODUCTION

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 (G), that is, we require nominal stability (NS) and nominal performance (NP). In addition, when a decentralized controller is used, it is desirable that the system be failure tolerant. This means that the system should remain stable as individual loops are opened or closed.

In practice the real (or "perturbed") 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 performance (RP) we mean that the performance requirements are satisfied for all $G_p \in \Pi$. Mainly for mathematical convenience, we choose to define performance using the H_∞ - norm. Define

$$NP \Leftrightarrow \bar{\sigma}(\Sigma) \leq 1, \quad \forall \omega \quad (1a)$$

$$RP \Leftrightarrow \bar{\sigma}(\Sigma_p) \leq 1, \quad \forall \omega, \quad \forall G_p \in \Pi \quad (1b)$$

In most cases Σ is the weighted sensitivity operator

$$\Sigma = W_1 S W_2, \quad S = (I + GC)^{-1} \quad (2a)$$

$$\Sigma_p = W_1 S_p W_2, \quad S_p = (I + G_p C)^{-1} \quad (2b)$$

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 weight each output according to its importance.

The definition of Robust Performance is of no value without simple methods to test if conditions like (1b) are satisfied for all G_p in the set Π of possible plants.

Doyle et al. (1982) have derived a computationally useful condition for (1b) involving the Structured Singular Value μ (μ is defined in Appendix). To use μ we must model the uncertainty (the set Π of possible plants G_p) as normbounded perturbations (Δ_i) on the nominal system. Through weights each perturbation is normalized to be of size one:

$$\bar{\sigma}(\Delta_i) \leq 1, \forall \omega \quad (3)$$

The perturbations, which may occur at different locations in the system, are collected in the diagonal matrix Δ_U (the subscript u denotes uncertainty)

$$\Delta_U = \text{diag}\{\Delta_1, \dots, \Delta_n\} \quad (4)$$

and the system is rearranged to match the structure in Fig. 1. The interconnection matrix M in Fig.1 is determined by the nominal model (G), the size and nature of the uncertainty, the performance specifications and the controller. For Fig.1 the robust performance condition (1b) becomes (Doyle et al., 1982)

$$RP \Leftrightarrow \mu(M) < 1, \quad \forall \omega \quad (5)$$

$\mu(M)$ depends on both the elements in the matrix M and the structure of the perturbation matrix $\Delta = \text{diag}\{\Delta_U, \Delta_P\}$. Sometimes this is shown explicitly by using the notation $\mu(M) = \mu_\Delta(M)$. Δ_P is a full square matrix with dimension equal to the number of outputs (the subscript P denotes performance). In addition to satisfying (5), the system must be nominally stable (i.e., M is stable). Also note that within this framework, the issue of robust stability (RS) is simply a special case of robust performance.

Decentralized Control

Decentralized control involves using a diagonal or block-diagonal controller (Fig. 2)

$$C = \text{diag}\{c_i\}$$

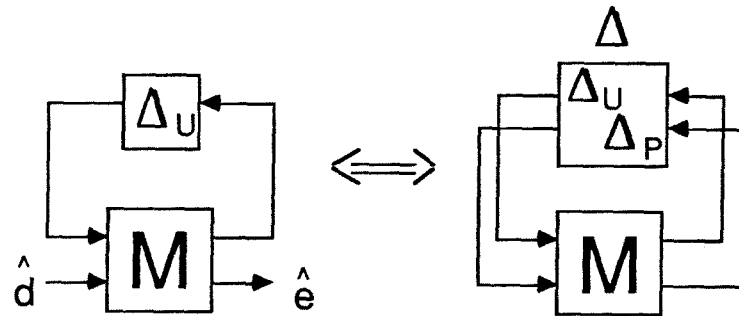


Fig. 1 General structure for studying effect of uncertainty (Δ_U) on performance. M is a function of the plant model (G) and the controller. \hat{d} : external inputs (disturbances, reference signals), \hat{e} : external outputs (weighted errors $y - r$), $\hat{e} = \Sigma_P \hat{d}$ (Eq. (1)). Δ_P is a full matrix corresponding to the dimensions of \hat{d} and \hat{e} . Robust performance is guaranteed if and only if $\mu_\Delta(M) < 1$.

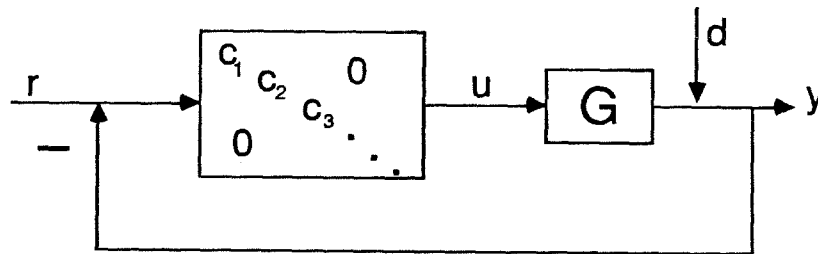


Fig. 2 Decentralized control structure.

Some reasons for using a decentralized controller are

- tuning and retuning is simple
- they are easy to understand
- they are easy to make failure tolerant

The design of a decentralized control system involves two steps

A) Choice of pairings (control structure)

B) Design of each SISO-controller c_i (or block).

The best way to proceed for each of these steps is still an active area of research. The RGA (Bristol, 1966) has proven to be an efficient tool for eliminating undesirable pairings in Step A. This paper deals with the Step B. Two design methods which may be applied for this step are 1) Sequential loop-closing and 2) Independent design of each loop.

1) Sequential loop-closing. This design approach (e.g., Mayne, 1973) involves designing each element (or block) in C sequentially. Usually the controller corresponding to a fast loop is designed first. This loop is then closed before the design proceeds with the next controller. This means that the information about the “lower-level” controllers is directly used as more loops are closed. The final step in the design procedure is to test if the overall system satisfies the RP-condition (5). The main disadvantages of this design method are

- Failure tolerance is not guaranteed when “lower-level” loops fail.
- The method depends strongly on which loop is designed first and how this controller is designed.
- There are no guidelines on how (and in which order) to design the controllers for each loop in order to guarantee robust performance of the overall system. Therefore the design proceeds by “trial-and-error”.

2) Independent design of each loop. This is the design approach used in this paper. In this case each controller element (or block) is designed independently of

the others. We present a procedure for these designs which guarantees robust performance of the overall system. The proposed method has the following advantage

- Failure tolerance: Nominal stability (of the remaining system) is guaranteed if any loop fails.
- Each controller is designed directly with no need for trial-and- error.

The main limitation of the approach is the assumption of independent designs, which means that we do not exploit information about the controllers used in the other loops. Therefore the derived bounds are only sufficient for robust performance.

Problem definition.

This paper addresses the following problem: Let \tilde{G} denote the diagonal (or block-diagonal) version of the plant corresponding to the chosen structure of C (i.e., \tilde{G} is found from G by deleting the off-diagonal elements). Assume that uncertainty and "interactions" are neglected when designing the controller C , that is, design each element (or block) of C independently based on the information contained in \tilde{G} only. What constraints have to be placed on the individual designs in order to guarantee robust performance of the overall system (which can be any plant G_p from the set Π)?

The constraints on the individual designs are chosen to be in terms of bounds on $|\tilde{h}_i|$ and $|\tilde{s}_i|$ where \tilde{h}_i and \tilde{s}_i are the closed-loop transfer functions for loop i :

$$\tilde{h}_i = g_{ii}c_i(1 + g_{ii}c_i)^{-1} \quad \tilde{H} = \text{diag}\{\tilde{h}_i\} \quad (6a)$$

$$\tilde{s}_i = (1 + g_{ii}c_i)^{-1} \quad \tilde{S} = \text{diag}\{\tilde{s}_i\} \quad (6b)$$

(In general, if C is block-diagonal, \tilde{h}_i, \tilde{s}_i and g_{ii} are matrices corresponding to the block-structure of C , and $|\tilde{s}_i|$ and $|\tilde{h}_i|$ are replaced by $\bar{\sigma}(\tilde{H}_i)$ and $\bar{\sigma}(\tilde{S}_i)$).

We solve the decentralized problem as defined above, by deriving the tightest possible bounds on

$$\bar{\sigma}(\tilde{H}) = \max_i |\tilde{h}_i| \quad \text{and} \quad \bar{\sigma}(\tilde{S}) = \max_i |\tilde{s}_i|$$

which guarantee robust performance:

$$RP \Leftrightarrow \bar{\sigma}(\tilde{H}) < \tilde{c}_H \text{ or } \bar{\sigma}(\tilde{S}) < \tilde{c}_S, \quad \forall \omega \quad (7)$$

In addition to satisfying (7) the system has to be nominally stable. The μ -interaction measure, introduced by Grosdidier and Morari (1986), gives a sufficient condition for nominal stability:

$$NS \Leftrightarrow \bar{\sigma}(\tilde{H}) \leq \mu_C(E_H), \quad \forall \omega, \quad E_H = (G - \tilde{G})\tilde{G}^{-1} \quad (8)$$

(μ is computed with respect to the structure of C which is equal to the structure of \tilde{G} , \tilde{H} and \tilde{S}). This paper provides a generalization of the μ -interaction measure from the case of nominal stability (NS) to the case of robust performance (RP). The results derived here also apply to robust stability (RS) or nominal performance (NP) if the μ -condition (5) is a RS- or NP-condition rather than a RP-condition.

Notation

The most important notation is summarized below.

G - model of the plant

$\tilde{G} = \text{diag}\{g_{ii}\}$ (corresponding to structure of C)

$G_p = f(G, \Delta_U)$, Δ_U : uncertainty, $G_p = G$ when $\Delta_U = 0$

$$\begin{aligned} \tilde{S} &= (I + \tilde{G}C)^{-1}, & \tilde{H} &= I - \tilde{S} \\ S &= (I + GC)^{-1}, & H &= I - S \\ S_p &= (I + G_p C)^{-1}, & H_p &= I - S_p \end{aligned} \quad (9)$$

Stability of individual loops $\Leftrightarrow \tilde{H}$ (and \tilde{S}) is stable

NS $\Leftrightarrow H$ (and S) is stable (overall system stable with no uncertainty)

RS $\Leftrightarrow H_p$ (and S_p) is stable (for all $G_p \in \Pi$).

NP $\Leftrightarrow S$ satisfies the performance specification

RP $\Leftrightarrow S_p$ satisfies the performance specification (for all $G_p \in \Pi$).

2. NOMINAL STABILITY (OF H AND S)

To apply the general robust performance condition $\mu(M) < 1$ (5) we must require that the system is nominally stable, that is, that the interconnection matrix

M is stable. Nominal stability is satisfied if H (and S) is stable. However, note that nominal stability (i.e., stability of H and S) is not necessarily implied by the stability of the individual loops (i.e., stability of \tilde{H} and \tilde{S}). The “interactions” (difference between G and \tilde{G}) may cause stability problems as discussed by Grosdidier and Morari (1986). If either one of the following conditions on $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$ is satisfied, then the stability of \tilde{H} (or \tilde{S}) implies nominal stability.

Condition 1 for NS (Grosdidier and Morari, 1986). *Assume \tilde{H} is stable (each loop is stable by itself), and that G and \tilde{G} have the same number of RHP (unstable) poles. Then H is stable (the system is stable when all loops are closed) if*

$$\bar{\sigma}(\tilde{H}) \leq \mu_C^{-1}(E_H) \quad \forall \omega \quad (10)$$

$$\text{where } E_H = (G - \tilde{G})\tilde{G}^{-1} \quad (11)$$

$\mu_C(E_H)$ is the μ -interaction measure and μ is computed with respect to the structure of the decentralized controller C . Note that the condition that G and \tilde{G} have the same number of RHP poles, is generally satisfied only when G and \tilde{G} are stable. In order to allow integral action ($\tilde{H}(0) = I$), we have to require that $\mu(E_H) < 1$ at $\omega = 0$, that is, we need diagonal dominance at low frequencies. If this is not the case the following alternative condition may be used:

Condition 2 for NS (Postlethwaite and Foo, 1985, Grosdidier, 1985). *Assume \tilde{S} is stable, and that G and \tilde{G} have the same number of RHP zeros. Then S (and H) is stable if*

$$\bar{\sigma}(\tilde{S}) \leq \mu_C^{-1}(E_S) \quad \forall \omega \quad (12)$$

$$\text{where } E_S = (G - \tilde{G})G^{-1} \quad (13)$$

Since we have to require $\tilde{S} = I$ as $\omega \rightarrow \infty$ for any real system, we have to require $\mu(E_S) < 1$ as $\omega \rightarrow \infty$, in order to be able to satisfy (12), that is, we must have diagonal dominance at high frequencies.

Conditions 1 and 2 are conditions for nominal stability (i.e., stability of H and S). These conditions cannot be combined over different frequency ranges. The reason is that the “uncertainty” $G - \tilde{G}$ is not a norm-bounded set and therefore is not “connected in the graph topology” (Postlethwaite, et al., 1985).

What to do when both conditions fail

In some cases it may be impossible to satisfy either (10) or (12). For example, in order to satisfy (10) and to have integral action ($\tilde{H}(0) = I$) we must require at least

$$\rho(E_H(0)) < 1 \quad (14)$$

(ρ is the spectral radius of E_H). (14) is derived from (10) by assuming $\tilde{H} = \tilde{h}I$ (all loops identical) which yields the least restrictive bound $\bar{\sigma}(\tilde{H}) \leq \rho^{-1}(E_H)$ in (10). In general (14) is conservative. For example, it is easily shown (Skogestad and Morari, 1987b), that it is always possible to find a diagonal controller which yields NS if the less restrictive condition

$$Re\{\lambda_i(E_H(0))\} \geq -1, \quad \forall i \quad (15)$$

is satisfied. One example for which (15) is satisfied, but not (14) is the following 2×2 plant

$$G(0) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \quad \tilde{G}(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\lambda_i(E_H(0)) = \pm i2, \quad \rho(E_H(0)) = 2$$

For 2×2 plants, (15) is always satisfied when $RGA_{11} > 0$ (RGA_{11} is the 1,1-element of the RGA (Bristol, 1966)), while (14) is only satisfied when $RGA_{11} > 0.5$.

Similarly, condition (12) may be impossible to satisfy because i) G and \tilde{G} do not have the same number of RHP-zeros, or ii) $\mu(E_S(j\infty)) \geq 1$.

In cases when neither Conditions (10) or (12) can be satisfied we may try to redefine the nominal model (G and \tilde{G}) such that either condition 1 or 2 is satisfied.

However, since the set Π of possible plants (G_p) still has to be the same, this generally means that we have to increase the magnitude of the model uncertainty. The three following “tricks” may be used (the last two of these are probably easiest to apply since uncertainty always dominates at high frequency):

- To satisfy (10): The plant is made diagonal dominant at low frequencies ($\mu(E_H)(0) < 1$), by reducing the magnitude of the nominal off-diagonal elements and replacing it by element uncertainty (at low frequency) (see Skogestad and Morari (1987c) on how to treat element uncertainty within the μ -framework).
- To satisfy (12): The plant is made diagonal dominant at high frequencies ($\mu(E_S(j\infty)) < 1$), by reducing the magnitude of the nominal off-diagonal elements and replacing it by element uncertainty (at high frequency).
- To have the same number of RHP-zeros in G and \tilde{G} : RHP-zeros (or time delays) are “removed” by treating them as uncertainty.

One extreme is obviously to treat the off-diagonal elements entirely as additive element uncertainty. In this case $\mu(E_H) = 0$ at all frequencies, and nominal stability (stability of H) is obviously satisfied if each loop \tilde{h} is stable (since $G = \tilde{G}$ and $H = \tilde{H}$ in this case). This approach is generally more conservative, however, since the offdiagonal elements in G (which nominally are equal to g_{ij}) for the case of element uncertainty are allowed to be any transfer function of magnitude $|g_{ij}|$ (in particular, both g_{ij} and $-g_{ij}$ are allowed). This additional uncertainty makes it more difficult to satisfy the robust stability and performance conditions.

3. ROBUST PERFORMANCE

Having derived conditions for nominal stability, we can now proceed to the case of robust performance. The objective of this section is to derive bounds on the individual designs (\tilde{H} and \tilde{S}), which when satisfied guarantee robust performance of the overall system (that is, $\mu(M) < 1$). This is accomplished in two steps:

1. Sufficient conditions for RP in terms of bounds on $\bar{\sigma}(H)$ and $\bar{\sigma}(S)$ are derived by writing M as a linear fractional transformation (LFT) of H and S .

2. These bounds are used to derive sufficient conditions for RP in terms of bounds on $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$.

3.1 Robust Performance Condition in Terms of H and S

The robust performance condition

$$RP \Leftrightarrow \mu_{\Delta}(M) \leq 1, \quad \forall \omega \quad (5)$$

may be used to derive sufficient conditions for RP in terms of bounds on $\bar{\sigma}(H)$ and $\bar{\sigma}(S)$ (Skogestad and Morari, 1987a). To this end write M as a LFT of H (Fig. 3)

$$M = N_{11}^H + N_{12}^H H (I - N_{22}^H H)^{-1} N_{21}^H \quad (16)$$

The matrix N^H , which is independent of C , can be obtained from M by inspection in many cases. Otherwise, the procedure given by Skogestad and Morari (1987a) can be used. They also point out that in general M is affine in H , that is, $N_{22}^H = 0$. Applying Theorem 1 of Skogestad and Morari (1987a) (the theorem is reproduced in the Appendix) the following sufficient condition for (5) is derived:

RP-condition in terms of H . *Assume M is given as a LFT of H (Eq. 16). Then at any given frequency*

$$\mu_{\Delta}(M) \leq 1 \quad \text{if} \quad \bar{\sigma}(H) \leq c_H \quad (17a)$$

where at this frequency c_H solves

$$\mu_{\hat{\Delta}} \left(\begin{array}{cc} N_{11}^H & N_{12}^H \\ c_H N_{21}^H & c_H N_{22}^H \end{array} \right) = 1 \quad (17b)$$

and μ is computed with respect to the structure $\hat{\Delta} = \text{diag}\{\Delta, H\}$.

Note that H is generally a "full" matrix if the controller is diagonal. A similar bound in terms of S is derived by replacing H by S in Eq. (16) and (17). (17) applies on a frequency-by-frequency basis. This implies that $\mu(M) \leq 1$ at a given frequency

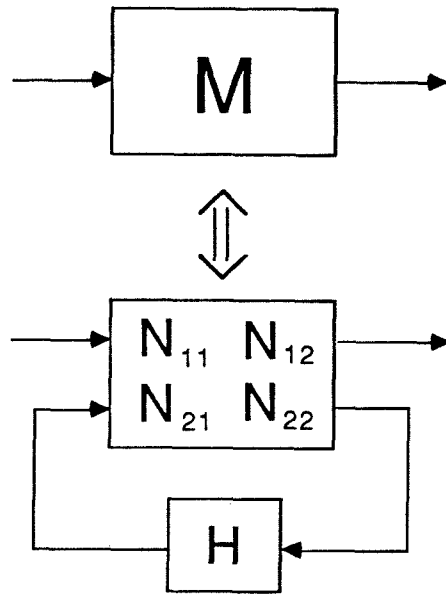


Fig. 3 M written as a LFT of H . N is independent of the controller.

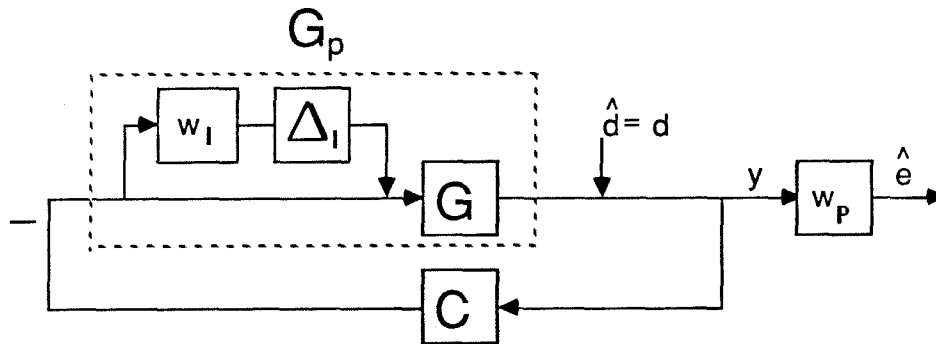


Fig. 4 Plant with input uncertainty (Δ_I) of magnitude $w_I(s)$. Robust performance is satisfied if $\bar{\sigma}(w_P(I + G_p C)^{-1}) \leq 1$, for all $\Delta_I(\bar{\sigma}(\Delta_I) \leq 1)$.

is guaranteed if either $\bar{\sigma}(H) < c_H$ or $\bar{\sigma}(S) < c_S$ at this frequency. Consequently, the bounds on $\bar{\sigma}(H)$ and $\bar{\sigma}(S)$ can be combined over different frequency ranges. In particular, the following holds

$$RP \Leftrightarrow \bar{\sigma}(H) \leq c_H \text{ or } \bar{\sigma}(S) \leq c_S, \quad \forall \omega \quad (18)$$

Example. Robust performance with input uncertainty (Fig. 4).

Let the set Π of possible plants be given by

$$G_p = G(I + w_I \Delta_I), \quad \bar{\sigma}(\Delta_I) \leq 1, \quad \forall \omega \quad (19)$$

Here w_I is the magnitude of the relative (multiplicative) uncertainty at the plant inputs. For robust performance we require that the magnitude of the sensitivity operator is bounded by $|w_p|^{-1}$:

$$RP \Leftrightarrow \bar{\sigma}(w_P S_p) = \bar{\sigma}(w_P (I + G_p C)^{-1}) \leq 1 \quad \forall \omega, \quad \forall G_p \in \Pi \quad (20)$$

This condition is most easily checked using μ (Eq. (5)):

$$RP \Leftrightarrow \mu_\Delta(M) \leq 1 \quad \forall \omega \quad (21a)$$

where the interconnection matrix M is (Skogestad and Morari, 1986):

$$M = \begin{pmatrix} -w_I C S G & -w_I C S \\ w_P S G & w_P S \end{pmatrix} \quad (21b)$$

and $\mu(M)$ is computed with respect to the structure $\Delta = \text{diag}\{\Delta_I, \Delta_P\}$. Δ_P is always a "full" matrix of the same dimension as S . Δ_I is often a diagonal matrix (if the inputs do not affect each other). Rewrite M in terms of S and H such that C does not appear

$$M = \begin{pmatrix} -w_I G^{-1} H G & -w_I G^{-1} H \\ w_P S G & w_P S \end{pmatrix} \quad (22)$$

By inspection M may be written as a LFT (16) of H (recall $S = I - H$)

$$M = N_{11}^H + N_{12}^H H N_{21}^H = \begin{pmatrix} 0 & 0 \\ w_P G & w_P I \end{pmatrix} + \begin{pmatrix} -w_I G^{-1} \\ -w_P I \end{pmatrix} H (G \quad I) \quad (23)$$

We derive from (23) and (17)

$$RP \quad \text{if} \quad \bar{\sigma}(H) \leq c_H \quad \forall \omega \quad (24a)$$

where at each frequency c_H solves

$$\mu \begin{pmatrix} 0 & 0 & -w_I G^{-1} \\ w_P G & w_P I & -w_P I \\ c_H G & c_H I & 0 \end{pmatrix} = 1 \quad (24b)$$

Similarly, a bound on $\bar{\sigma}(S)$ is derived by writing M as a LFT of S .

$$M = N_{11}^S + N_{12}^S S N_{21}^S = \begin{pmatrix} -w_I I & -w_I G^{-1} \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} w_I G^{-1} \\ w_P I \end{pmatrix} S (G \quad I) \quad (25)$$

(25) and (17) (with H replaced by S) yield

$$RP \quad \text{if} \quad \bar{\sigma}(S) \leq c_S \quad \forall \omega \quad (26a)$$

where at each frequency c_S solves

$$\mu \begin{pmatrix} -w_I I & -w_I G^{-1} & w_I G^{-1} \\ 0 & 0 & w_P I \\ c_S G & c_S I & 0 \end{pmatrix} = 1 \quad (26b)$$

In both (24b) and (26b) μ is computed with respect to the structure $\text{diag}\{\Delta_I, \Delta_P, H\}$. The bounds (24) and (26) may be combined over different frequency ranges, and RP is guaranteed if either one is satisfied at any frequency (Eq. (18)). In practice, (24) is most easily satisfied at high frequencies and (26) at low frequencies.

3.2 Robust Performance Condition in Terms of \tilde{H} and \tilde{S}

Sufficient conditions for Robust Performance in terms of $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$ may now be derived using the identities (Grosdidier, 1985)

$$H = G \tilde{G}^{-1} \tilde{H} (I + E_H \tilde{H})^{-1} \quad (27)$$

$$S = \tilde{S} (I - E_S \tilde{S})^{-1} \tilde{G} G^{-1} \quad (28)$$

Note that (27) and (28) both are LFT's of H (and S) in terms of \tilde{H} (and \tilde{S}). In Section 3.1 we pointed out that in general M can be written as a LFT of H with $N_{22}^H = 0$:

$$M = N_{11}^H + N_{12}^H H N_{21}^H \quad (29)$$

Substituting (27) into (29) yields

$$M = N_{11}^H + N_{12}^H G \tilde{G}^{-1} \tilde{H} (I + E_H \tilde{H})^{-1} N_{21}^H \quad (30)$$

which is a LFT of M in terms of \tilde{H} . Using Theorem 1 (Appendix 1) and (30) we derive:

RP-condition in terms of \tilde{H} . *Let $M = N_{11}^H + N_{12}^H H N_{21}^H$. Then at any frequency*

$$\mu_{\Delta}(M) \leq 1 \quad \text{if} \quad \bar{\sigma}(\tilde{H}) \leq \tilde{c}_H \quad (31a)$$

where at this frequency \tilde{c}_H solves

$$\mu_{\hat{\Delta}} \left(\begin{array}{cc} N_{11}^H & N_{12}^H G \tilde{G}^{-1} \\ \tilde{c}_H N_{21}^H & -\tilde{c}_H E_H \end{array} \right) = 1 \quad (31b)$$

and μ is computed with respect to the structure $\hat{\Delta} = \text{diag}\{\Delta, C\}$.

Note that the structure of C is block-diagonal and equal to that of \tilde{H} . An entirely equivalent condition may be derived in terms of $\bar{\sigma}(\tilde{S})$:

RP-condition in terms of \tilde{S} . *Let $M = N_{11}^S + N_{12}^S S N_{21}^S$. Then at any frequency*

$$\mu_{\Delta}(M) \leq 1 \quad \text{if} \quad \bar{\sigma}(\tilde{S}) \leq \tilde{c}_S \quad (32a)$$

where \tilde{c}_S solves

$$\mu_{\hat{\Delta}} \left(\begin{array}{cc} N_{11}^S & N_{12}^S \\ \tilde{c}_S \tilde{G} G^{-1} N_{21}^S & \tilde{c}_S E_S \end{array} \right) = 1 \quad (32b)$$

and μ is computed with respect to the structure $\hat{\Delta} = \text{diag}\{\Delta, C\}$.

Again, the bounds (31) and (32) may be combined over different frequency ranges:

Combined RP-condition.

$$RP \quad \text{if} \quad \bar{\sigma}(\tilde{H}) \leq \tilde{c}_H \quad \text{or} \quad \bar{\sigma}(\tilde{S}) < \tilde{c}_S \quad \forall \omega \quad (33)$$

Example. Robust Performance with Input Uncertainty (Continued)

Consider the same example as above (Fig. 4). However, in this case we will derive bounds in terms of $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$. A RP-condition in terms of $\bar{\sigma}(\tilde{H}) = |\tilde{h}_i|$ is derived by combining (31) and (23):

$$RP \quad \text{if} \quad \bar{\sigma}(\tilde{H}) \leq \tilde{c}_H \quad \forall \omega \quad (34a)$$

where at each frequency \tilde{c}_H solves

$$\mu_{\hat{\Delta}} \begin{pmatrix} 0 & 0 & -w_I \tilde{G}^{-1} \\ w_P G & w_P I & -w_P G \tilde{G}^{-1} \\ \tilde{c}_H G & \tilde{c}_H I & -\tilde{c}_H E_H \end{pmatrix} = 1 \quad (34b)$$

Similarly, the RP-condition in terms of $\bar{\sigma}(\tilde{S}) = |\tilde{s}_i|$ derived from (32) and (25) is

$$RP \quad \text{if} \quad \bar{\sigma}(\tilde{S}) \leq \tilde{c}_S \quad \forall \omega \quad (35a)$$

where at each frequency \tilde{c}_S solves

$$\mu_{\hat{\Delta}} \begin{pmatrix} -w_I I & -w_I G^{-1} & w_I G^{-1} \\ 0 & 0 & w_P I \\ \tilde{c}_S \tilde{G} & \tilde{c}_S G \tilde{G}^{-1} & \tilde{c}_S E_S \end{pmatrix} = 1 \quad (35b)$$

In both (34b) and (35b) μ is computed with respect to the structure $\hat{\Delta} = \text{diag}\{\Delta_I, \Delta_P, C\}$. Conditions (34) and (35) can be combined as shown in (33).

4. DESIGN PROCEDURE

The following design procedure for decentralized control systems based on the “independent designs”-assumption is proposed: Find a decentralized controller which yields individual loops (\tilde{H} and \tilde{S}) which are stable and in addition satisfy

- 1) Nominal Stability: Satisfy $\bar{\sigma}(\tilde{H}) \leq \mu^{-1}(E_H)$ (10) at all frequencies or satisfy $\bar{\sigma}(\tilde{S}) \leq \mu^{-1}(E_S)$ (12) at all frequencies. It is not allowed to combine (10) and (12).
- 2) Robust performance: At each frequency satisfy either $\bar{\sigma}(\tilde{H}) \leq \tilde{c}_H$ (31) or $\bar{\sigma}(\tilde{S}) \leq \tilde{c}_S$ (32). Combining (31) and (32) over different frequency ranges is allowed.

Consequently, two separate conditions must be satisfied by the individual designs: One for nominal stability and one for robust performance.

Remarks

1. The nominal stability condition $\bar{\sigma}(\tilde{H}) \leq \mu^{-1}(E_H)$ (10) is automatically satisfied at any frequency where the robust performance condition (31) is satisfied. This follows from the inequality

$$\mu_{diag\{\Delta_1, \Delta_2\}} \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} \geq \max\{\mu_{\Delta_1}(N_{11}), \mu_{\Delta_2}(N_{22})\} \quad (36)$$

applied to (31b). We find $\tilde{c}_H \leq \mu^{-1}(E_H)$ and therefore the RP-condition puts a tighter bound on $\bar{\sigma}(\tilde{H})$ than the NS- condition (10). A similar relationship exists between the RP-condition (32) on $\bar{\sigma}(\tilde{S})$ and the NS-condition (12).

2. This may seem to imply that NS is automatically guaranteed if RP is satisfied. This is not the case, however, since the NS-condition (10) (or (12)) must be satisfied at all frequencies. This is not necessarily implied by the combined RP-condition (33) since neither $\bar{\sigma}(\tilde{H}) \leq \tilde{c}_H$ or $\bar{\sigma}(\tilde{S}) \leq c_S$ have to be satisfied at all frequencies to satisfy (33). In the following two cases RP does imply NS:
3. If it happens that the RP-bound (31) on $\bar{\sigma}(\tilde{H})$ is satisfied at all frequencies, and if \tilde{H} is stable, then RP and NS are both guaranteed using a single condition. However, to be able to satisfy (31) at all frequencies we must require that there exists a $c_H > 0$ which solves (31b). This is equivalent to requiring $\mu(N_{11}^H) \leq 1$, which from (30) is equivalent to $\mu(M(\tilde{H} = 0)) \leq 1$. Consequently, to be able to satisfy (31) we must require that at each frequency the performance requirements are such that $\tilde{H} = 0$ is a possible solution. This may be the case, for example, if we are interested in robust stability only.
4. If it happens that the RP-bound (32) on $\bar{\sigma}(\tilde{S})$ is satisfied at all frequencies, and if \tilde{S} is stable, then RP and NS are both guaranteed. However, to be able to satisfy (32) at all frequencies we must require that at each frequency $\tilde{S} = 0$

is a possible solution. This may be the case if there is no uncertainty, that is, if we are interested in nominal performance only.

5. NUMERICAL EXAMPLE

In this section we continue the previous example of RP with diagonal input uncertainty (Fig.4). Consider the following plant (time is in minutes)

$$\hat{G} = \frac{1}{1 + 75s} \begin{bmatrix} -0.878 \frac{1-0.2s}{1+0.2s} & 0.014 \\ -1.082 \frac{1-0.2s}{1+0.2s} & -0.014 \frac{1-0.2s}{1+0.2s} \end{bmatrix} \quad (37)$$

Physically, this may correspond to a high-purity distillation column using distillate (D) and boilup (V) as manipulated inputs to control top and bottom composition (Skogestad and Morari, 1986). We want to design a decentralized (diagonal) controller for this plant such that robust performance is guaranteed when there is 10% uncertainty on each manipulated input. The uncertainty and performance weights are

$$\hat{w}_I(s) = 0.1 \quad (38a)$$

$$w_P(s) = 0.25 \frac{7s + 1}{7s} \quad (38b)$$

The robust performance condition is

$$\bar{\sigma}(S_p) < 1/|w_P| \quad \forall G_p \in \Pi \quad (39)$$

(38b) implies that we require integral action ($w_P(0) = \infty$) and allow an amplification of disturbances at high frequencies of at most a factor of four ($w_P(j\infty) = 0.25$). A particular sensitivity function which matches the performance bound (39) exactly at low frequencies and satisfies it easily at high frequency is $S = \frac{28s}{28s+1}I$. This corresponds to a first order response with time constant 28 min.

Nominal Stability (NS)

The nominal model has $RGA_{11} = 0.45$ and we find $\mu(E_H(0)) = 1.11$. Consequently, it is impossible to satisfy the NS-condition (10).

The NS-condition (12) for $\bar{\sigma}(\tilde{S})$ cannot be satisfied either. Firstly, \hat{G} has one RHP-zero, while the diagonal plant has two. Secondly, the plant is clearly not diagonal dominant at high frequencies, and $\mu(E_S(j\omega))$ is larger than one for $\omega > 4 \text{ min}^{-1}$. The simplest way to get around this problem is to treat the RHP-zeros as uncertainty. (This is actually not very conservative, since RHP-zeros limit the achievable performance anyway.) To this end define the following “new” nominal model

$$G = \frac{1}{1 + 75s} \begin{bmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{bmatrix} \quad (40)$$

and include the RHP-zeros in the input uncertainty by using the following new uncertainty weight

$$w_I(s) = 0.1 \frac{5s + 1}{0.25s + 1} \quad (41)$$

$|w_I(j\omega)|$ reaches a value at one at about $\omega = 2 \text{ min}^{-1}$. This includes the neglected RHP-zeros since the relative uncertainty introduced by replacing $\frac{1-0.2s}{1+0.2s}$ by 1 is $|1 - \frac{1-0.2s}{1+0.2s}|$, which reaches a value of one at about $\omega = 3 \text{ min}^{-1}$.

With the new model (40) we still cannot satisfy the NS-condition (10) for $\bar{\sigma}(\tilde{H})$. However, the NS-condition (12) on $\bar{\sigma}(\tilde{S})$ is easily satisfied since G and \tilde{G} have the same number of RHP-zeros (none), and $\mu(E_S) = 0.743$ at all frequencies. The only restriction this imposes on \tilde{S} is that the maximum peaks of $|\tilde{s}_1|$ and $|\tilde{s}_2|$ must be less than $1/0.743 = 1.35$. This is easily satisfied since both $\tilde{g}_{11} = \frac{-0.878}{1+75s}$ and $\tilde{g}_{22} = \frac{-0.014}{1+75s}$ are minimum phase.

In the remainder of this Section the model of the plant (G) is assumed to be given by (40) and the uncertainty weight (w_I) by (41).

Nominal Performance (NP)

The NP-requirement is

$$NP \Leftrightarrow \bar{\sigma}(S) \leq |w_P|^{-1} \quad \forall \omega \quad (42)$$

How should the individual loops ($\tilde{S} = \text{diag}\{\tilde{s}_1, \tilde{s}_2\}$) be designed in order to satisfy this requirement? Intuitively, we might expect that we have to require at least that the individual loops satisfy (42), that is, $\bar{\sigma}(\tilde{S}) \leq |w_P|^{-1}$. However, this is not necessarily the case, as illustrated by the example: (42) is equivalent to $\mu_{\Delta}(M) \leq 1$ with $M = w_P S$ and $\Delta = \Delta_P$ (Δ_P is a full matrix). (32) then yields the following sufficient condition for NP in terms of \tilde{S} :

$$NP \iff \bar{\sigma}(\tilde{S}) \leq c_{NP} \quad \forall \omega \quad (43a)$$

where c_{NP} at each frequency solves

$$\mu_{\hat{\Delta}} \begin{pmatrix} 0 & w_P I \\ c_{NP} \tilde{G} G^{-1} & c_{NP} E_S \end{pmatrix} = I \quad (43b)$$

and $\hat{\Delta} = \text{diag}\{\Delta_P, C\}$. In our example Δ_P is a “full” 2×2 matrix, and C is a diagonal 2×2 matrix. c_{NP} is shown graphically in Fig. 5 and it is seen to be larger than $|w_P|^{-1}$ at low frequency. Consequently, the performance of the overall system (S) may be better than that of the individual loops (\tilde{s}_1 and \tilde{s}_2), that is, the interactions may improve the performance.

Robust Performance (RP)

Bound on $\bar{\sigma}(\tilde{H})$. The bound \tilde{c}_H on $\bar{\sigma}(\tilde{H})$ is given by Eq. (34) and is shown graphically in Fig. 6. (μ of the matrix in (34b) is computed with respect to the structure $\hat{\Delta} = \text{diag}\{\Delta_I, \Delta_P, C\}$, where Δ_I is a diagonal 2×2 matrix, Δ_P is a full 2×2 matrix and C is a diagonal 2×2 matrix). It is clearly not possible to satisfy the bound $\bar{\sigma}(\tilde{H}) < \tilde{c}_H$ at all frequencies. In particular, we find $\tilde{c}_H \leq 0$ for $\omega < 0.03 \text{ min}^{-1}$. The reason is that the performance-weight $|w_P| > 1$ in this frequency range, which means that feedback is required (i.e., $\tilde{H} = 0$ is not possible, see Remark 3 in Section 4).

Bound on $\bar{\sigma}(\tilde{S})$. The bound \tilde{c}_S on $\bar{\sigma}(\tilde{S})$ is given by Eq. (35) and is shown graphically in Fig. 7 (μ is computed with respect to the same structure as above). Again it is

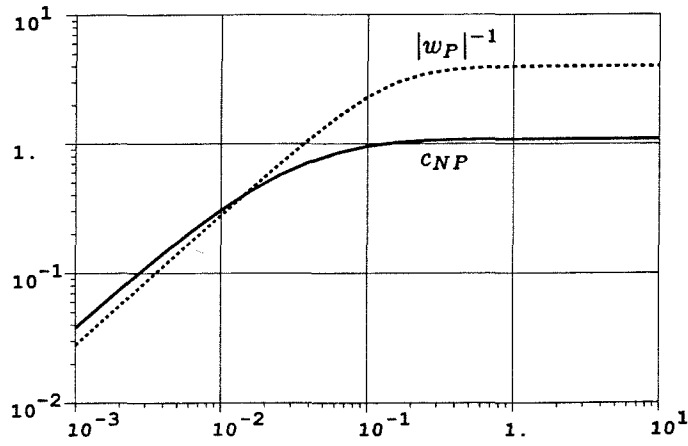


Fig. 5 NP is satisfied if and only if $\bar{\sigma}(S) \leq |w_P|^{-1}$ which is satisfied if $\bar{\sigma}(\tilde{S}) \leq c_{NP}$.

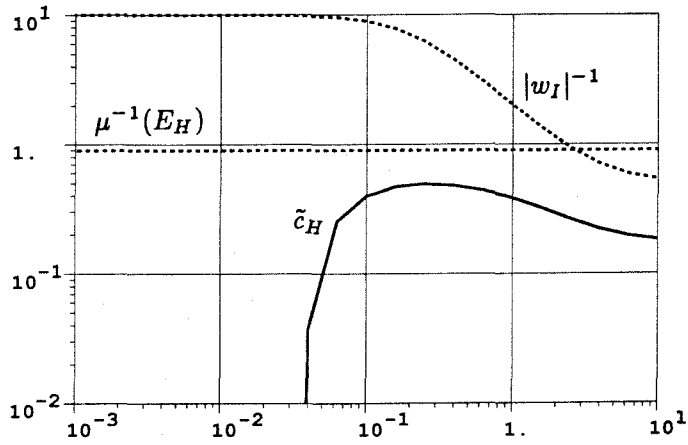


Fig. 6 Bounds $\mu^{-1}(E_H)$ and \tilde{c}_H on $\bar{\sigma}(\tilde{H})$.

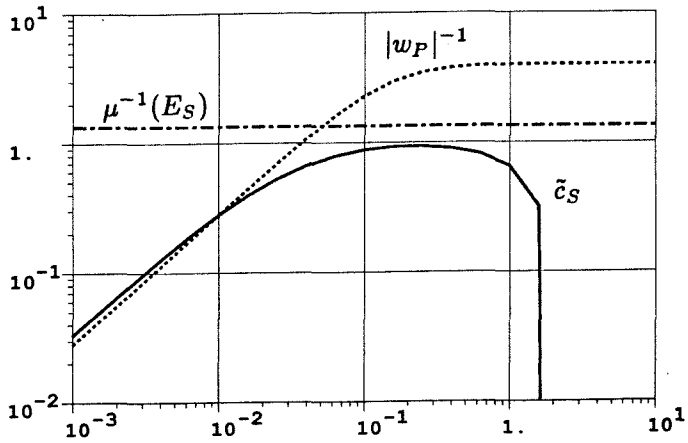


Fig. 7 Bounds $\mu^{-1}(E_S)$ and \tilde{c}_S on $\bar{\sigma}(\tilde{S})$.

not possible to satisfy this bound at all frequencies. In particular, we find $\tilde{c}_S \leq 0$ for $\omega > 2 \text{ min}^{-1}$. The reason is that the uncertainty weight $|w_I| > 1$ in this frequency range, which means that perfect control ($\tilde{S} = 0$) is not allowed.

Combining bounds on $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$. The bound on $\bar{\sigma}(\tilde{S})$ is easily satisfied at low frequencies, and the bound on $\bar{\sigma}(\tilde{H})$ is easily satisfied at high frequencies. The difficulty is to find a $\tilde{S} = I - \tilde{H}$ which satisfy either one of the conditions in the frequency range from 0.1 to 1 min^{-1} . The following design is seen to do the job (Fig. 8).

$$\tilde{h}_1 = \tilde{h}_2 = \frac{1}{7.5s + 1}, \quad \tilde{s}_1 = \tilde{s}_2 = \frac{7.5s}{7.5s + 1} \quad (44)$$

The bound on $|\tilde{s}_1|$ is satisfied for $\omega < 0.3 \text{ min}^{-1}$, and the bound on $|\tilde{h}_i|$ is satisfied for $\omega > 0.23 \text{ min}^{-1}$. (44) corresponds to the following controller

$$C = k \frac{(1 + 75s)}{s} \begin{pmatrix} \frac{-1}{0.878} & 0 \\ 0 & \frac{-1}{0.014} \end{pmatrix}, \quad k = 0.133 \quad (45)$$

Because the bounds \tilde{c}_H and \tilde{c}_S are almost flat in the cross-over region, the result is fairly insensitive to the particular choice of controller gain; it turns out that $0.06 < k < 0.25$ yields a design which satisfies at each frequency $\bar{\sigma}(\tilde{S}) < \tilde{c}_S$ or $\bar{\sigma}(\tilde{H}) < \tilde{c}_H$ and thus has RP. The controller (45) obviously yields an overall system which satisfies the robust performance condition, that is, $\mu(M)$ is less than one. This is also seen from Fig. 9 which shows $\mu(M)$ (M is given by (21b)) as a function of frequency. We find $\mu_{RP} = \sup_{\omega} \mu(M) = 0.63 < 1$ and RP is guaranteed. The fact that μ_{RP} is so much smaller than one, demonstrates some of the conservativeness of conditions (34) and (35) (which are only sufficient for RP).

6. CONCLUSION

This paper solves the problem of robust performance using independent designs as introduced in the Introduction. The example illustrates that this design approach may be useful for designing decentralized controllers.

The main limitation of the approach stems from the initial assumption regarding independent designs: Since each loop is designed separately, we cannot make

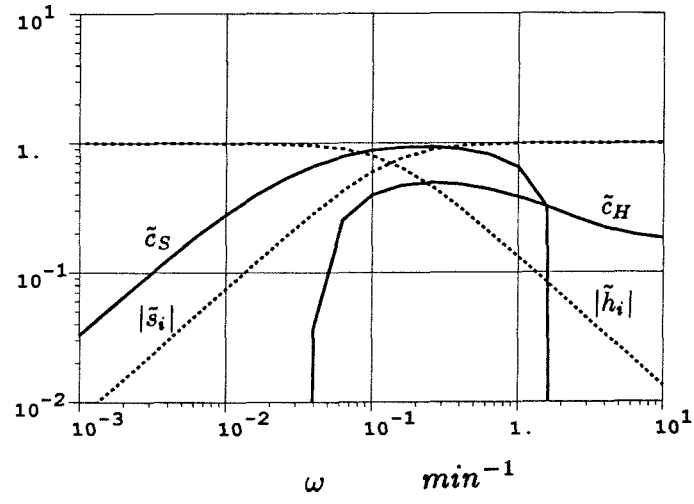


Fig. 8 RP is guaranteed since $|\tilde{s}_i| < \tilde{c}_S$ for $\omega < 0.3 \text{ min}^{-1}$ and $|\tilde{h}_i| < \tilde{c}_H$ for $\omega > 0.23 \text{ min}^{-1}$. $\tilde{h}_i = 1/1 + 7.5s$.

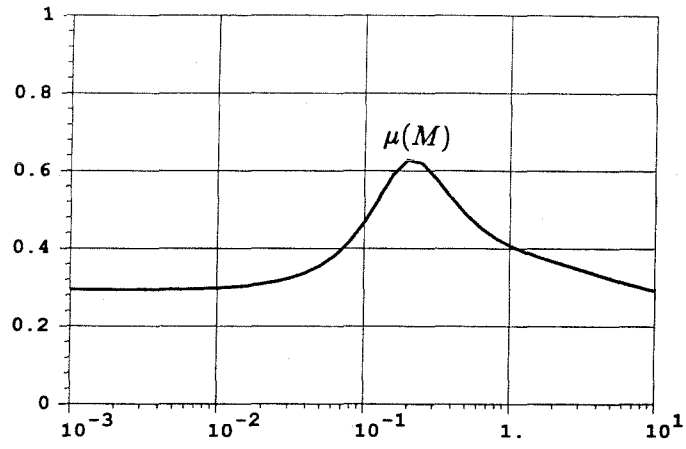


Fig. 9 $\mu(M)$ as a function of frequency. RP is guaranteed since $\mu(M) < 1$ at all frequencies.

use of information about the controllers used in the other loops. The consequence is that the bounds on $\bar{\sigma}(\tilde{S})$ and $\bar{\sigma}(\tilde{H})$ are only sufficient for robust performance; there will exist decentralized controllers which violate the bounds on $\bar{\sigma}(\tilde{S})$ and $\bar{\sigma}(\tilde{H})$, but which satisfy the robust performance condition. However, the derived bounds on $\bar{\sigma}(\tilde{S})$ and $\bar{\sigma}(\tilde{H})$ are the tightest norm bounds possible, in the sense that in such cases there will exist another controller with the same values of $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$ which does not yield robust performance.

The bounds on $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$ tend to be most conservative in the frequency range around crossover where $\bar{\sigma}(\tilde{H})$ and $\bar{\sigma}(\tilde{S})$ are both close to one. If, for a particular case, it is not possible to satisfy either $\bar{\sigma}(\tilde{H}) < \tilde{c}_H$ or $\bar{\sigma}(\tilde{S}) < \tilde{c}_S$ in this frequency range, then try the following: Design a controller for which the frequency range where both bounds are violated is as small as possible. Since the bounds are only sufficient for RP, this may still yield an acceptable design with robust performance. This may be checked using the tight RP-condition $\mu(M) < 1$ (5).

Another potential source of conservativeness is the inherent assumption of similar or equal bandwidths in all loops which is made when the same bounds on $|\tilde{h}_i|$ and $|\tilde{s}_i|$ are used for all loops. This limitation may be partially eliminated by including matrix valued weights on \tilde{H} and \tilde{S} (see Grosdidier and Morari, 1986). However, it is not obvious how these weights should be chosen apriori.

Acknowledgements. Partial support from the National Science Foundation and Norsk Hydro is gratefully acknowledged. We are thankful to Prof. John C. Doyle and Dr. Pierre Grosdidier for numerous useful discussions and remarks.

References

- Bristol, E. H., On a New Measure of Interaction for Multivariable Process Control, IEEE Trans. on Automatic Control, **AC-11**, 133-134 (1966).
- Doyle, J. C., Analysis of Feedback Systems with Structured Uncertainties, Inst. Electrical Engrs. Proc., **129**, Pt. D, 6, 242-250 (1982).
- Doyle, J. C., J. E. Wall and G. Stein, Performance and Robustness Analysis for Structured Uncertainty, Proc. IEEE Conf. on Decision and Control, Orlando, FL (1982).
- Grosdidier, P., Personal communication (1985).
- Grosdidier, P. and M. Morari, Interaction Measures for Systems Under Decentralized Control, Automatica, **22**, 3, 309-319 (1986).
- Mayne, D. Q., The Design of Linear Multivariable Systems, Automatica, **9**, 201-227 (1973).
- Postlethwaite, I. and V. K. Foo, Robustness with Simultaneous Pole and Zero Movement Across the $j\omega$ -Axis, Automatica, **21**, 433-443 (1985).
- Skogestad, S. and M. Morari, Control of Ill-Conditioned Plants: High Purity Distillation, AIChE Annual Meeting, Miami Beach, FL (1986).
- Skogestad, S. and M. Morari, Some New Properties of the Structured Singular Value, submitted to IEEE Trans. on Automatic Control (1987a).
- Skogestad, S. and M. Morari, Letter to the Editor concerning paper by Mijares et al., accepted for publication in AIChE Journal (1987b).
- Skogestad, S. and M. Morari, Effect of Model Uncertainty on Dynamic Resilience, Chem. Eng. Sci., to appear (1987c).

APPENDIX.

Definition of the Structured Singular Value μ . Let M be a square complex matrix. $\mu(M)$ is defined such that $\mu^{-1}(M)$ is equal to the smallest $\bar{\sigma}(\Delta)$ needed to make $(I + \Delta M)$ singular, i.e.,

$$\mu^{-1}(M) = \min_{\delta} \{ \delta | \det(I + \Delta M) = 0 \text{ for some } \Delta, \bar{\sigma}(\Delta) < \delta \} \quad (A1)$$

(If M is a transfer matrix this definition applies frequency-by-frequency). Δ is a block-diagonal perturbation matrix with a given structure

$$\Delta = \begin{bmatrix} \Delta_1 & & \\ & \ddots & \\ & & \Delta_n \end{bmatrix}, \quad \bar{\sigma}(\Delta_i) < \delta, \quad \forall i$$

Δ is allowed to be any complex matrix satisfying $\bar{\sigma}(\Delta) < \delta$. (It turns out that Δ may be restricted to being unitary without changing $\mu(M)$ (Doyle, 1982)). $\mu(M)$ depends on both the matrix M and the structure of the perturbations Δ . This is sometimes shown explicitly by using the notation $\mu(M) = \mu_{\Delta}(M)$. An equivalent statement of (A1) which is more useful for our purposes is the following:

$$\begin{aligned} \det(I + \Delta M) &\neq 0, \quad \forall \Delta (\bar{\sigma}(\Delta) < \delta) \\ \Leftrightarrow \rho(\Delta M) &\leq 1, \quad \forall \Delta (\bar{\sigma}(\Delta) < \delta) \\ \Leftrightarrow \mu_{\Delta}(M) &\leq 1/\delta \end{aligned} \quad (A2)$$

The reader is referred to Doyle (1982) for further properties and computational aspects of μ .

Theorem 1. *Let M be written as a LFT of T :*

$$M = N_{11} + N_{12}T(I - N_{22}T)^{-1}N_{21} \quad (A3)$$

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

$$\mu_{\Delta}(M) \leq k \quad (A4)$$

if

$$\bar{\sigma}(T) \leq c_T \quad (A5)$$

where c_T solves

$$\mu_{\Delta} \begin{bmatrix} N_{11} & N_{12} \\ kc_T N_{21} & kc_T N_{22} \end{bmatrix} = k \quad (A6)$$

and $\hat{\Delta} = \text{diag}\{\Delta, T\}$.

Proof. The theorem follows directly from the definition of μ (A2) after some algebra: Assume that T is defined such that $\bar{\sigma}(T) < c_T$. Then at each frequency the following holds

$$\begin{aligned} \mu_{\Delta}(M) &\leq k(\omega), \quad \forall T \ (\bar{\sigma}(T) < c_T) \\ \Leftrightarrow \det(I + \Delta M) &\neq 0, \quad \forall \Delta \ (\bar{\sigma}(\Delta) < 1/k), \quad \forall T \end{aligned} \quad (A7)$$

$$\Leftrightarrow \det \begin{bmatrix} I + \Delta N_{11} & \Delta N_{12} \\ -TN_{21} & I - TN_{22} \end{bmatrix} \neq 0 \quad \forall \Delta, \quad \forall T \quad (A8)$$

$$\Leftrightarrow \det \left(I + \begin{bmatrix} k\Delta & 0 \\ 0 & -\frac{1}{c_T}T \end{bmatrix} \begin{bmatrix} \frac{1}{k}N_{11} & \frac{1}{k}N_{12} \\ c_T N_{21} & c_T N_{22} \end{bmatrix} \right) \neq 0 \quad \forall \Delta, \quad \forall T$$

$$\Leftrightarrow \mu_{\hat{\Delta}} \begin{bmatrix} \frac{1}{k}N_{11} & \frac{1}{k}N_{12} \\ c_T N_{21} & c_T N_{22} \end{bmatrix} \leq 1$$

$$\Leftrightarrow \mu_{\hat{\Delta}} \begin{bmatrix} N_{11} & N_{12} \\ kc_T N_{21} & kc_T N_{22} \end{bmatrix} \leq k(\omega)$$

The step from (A7) to (A8) follows $M = N_{11} + N_{12}(I - TN_{22})^{-1}TN_{21}$ and Schurs formula

$$\det(A - BD^{-1}C) = \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} / \det D \quad (A9)$$

and the assumption $\det D = \det(I - TN_{22}) \neq 0$

Chapter VIII

A SYSTEMATIC APPROACH TO DISTILLATION COLUMN CONTROL

A SYSTEMATIC APPROACH TO DISTILLATION COLUMN CONTROL

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January 1987

Abstract

This paper presents a systematic approach to distillation column control. The main emphasis is on the steps which precede the actual controller design, namely the modelling of the column and the selection of the control configuration. By control configuration in this context we mean the two independent variable used for composition control (for example, L and V , D and V , or $\frac{L}{D}$ and $\frac{V}{B}$). The steps preceding the controller design are generally the most important, yet, they are not usually discussed in any detail in the academic literature. The goal of this paper is to fill in this apparent gap.

1. INTRODUCTION

Distillation columns constitute a major part of most chemical processing plants. The objective of a distillation column is to split the feed into two (or more) products with compositions different from that of the feed. The desired composition of the products may be fixed by product requirements or may result from some plantwide optimization. An important objective of the control system should be to keep these product compositions at their desired level. In practice, very few industrial columns maintain "dual" composition control, and it is still common to find that both compositions are controlled manually. Reports from industry indicate energy savings of 10-30% (Stanley and McAvoy, 1985) if dual composition control is used instead of applying manual control which usually results in overpurification or loss of valuable product. Also, a recent survey among plant managers (Dartt, 1985) cites distillation as the unit operation which could benefit most significantly from improved control.

A main reason for why dual composition control is not widely applied in industry is the stability problem often encountered when the controllers are tuned in order to get a reasonably fast response. In particular, high-purity columns tend to be difficult to control. Ironically, columns with low-purity products, which are simple to control, are the ones usually studied experimentally in university laboratories. Another reason for the infrequent use of dual composition control is the lack of systematic guidelines in the literature on how to design control systems for distillation columns.

In distillation control the gap between the industrial and the academic camp is clearly evident from the literature. The *industrial camp* argues from experience. On one side hardly anybody is willing to challenge the recommendations and conclusions. On the other side, there are few who claim to understand the reasoning and the explanations. Take for example the loop pairing recommendations based

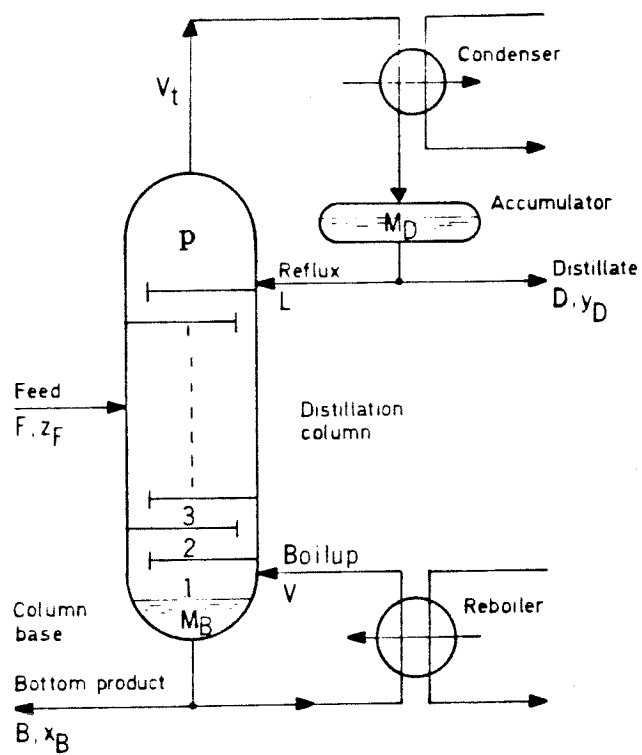


Fig. 1. Distillation column with 5 manipulated inputs (L, V, D, B and V_T) and 5 controlled outputs (y_D, x_B, M_D, M_B and p).

on the Relative Gain Array (RGA). From the industrial literature (Shinskey, 1984) one is led to believe that the RGA is just an interaction measure indicating the difficulties one can expect when tuning single loops for a multivariable system. But if this is true then the RGA recommendations regarding variable selection for distillation control would be arcane today, when the implementation of multivariable control systems is quite straightforward. Practical evidence suggests, however, that the RGA evaluation is very useful even for the design of multivariable control systems. Thus, while the RGA has proven to be a very useful tool for categorizing experience, it has not helped to explain the observed phenomena.

This paper is aimed at filling the apparent gap and to present a systematic approach to distillation column control. How should we go about designing a control system for a distillation column? Three separate steps are involved:

1. Modelling the column.
2. Choice of the control configuration.
3. Controller design/implementation.

This paper is aimed at discussing Step 1 and 2 which should precede the controller design. In fact, the controller design itself is probably the least important step in spite of the attention it is given in the literature. Let us consider Step 2. Most two-product distillation columns may be described as a 5×5 plant (Fig. 1). The five controlled variables are

- Vapor Holdup (M_V) (expressed by the pressure p).
- Liquid holdup in accumulator (condenser) (M_D).
- Liquid holdup in column base (reboiler) (M_B).
- Distillate product composition (y_D).
- Bottom product composition (x_B).

The five manipulated inputs are (corresponding to the available valves)

- Distillate flow (D)

- Bottom flow (B)
- Reflux (L)
- Boilup (V) (manipulated indirectly through the reboiler duty)
- Overhead vapor flow (V_T) (manipulated indirectly through the condenser duty)

In theory we could imagine designing the “optimal” 5×5 controller for the column. While this would certainly be of theoretical interest, it is very unlikely that such a controller would ever be implemented in practice. In order to make the control system failure tolerant and easier to understand and tune, simpler control structures are used. More specifically, we will not use all the five flows L , V , D , B and V_T for composition control, but only two independent combinations. The overall control system will then consist of a 2×2 controller (denoted by K) for composition control plus a control system for level and pressure control. This simplification does not necessarily lead to a serious degradation in achievable performance since the pressure and level loops often are so fast that they effectively give three static relationships between the five manipulated inputs (L , V , V_T , D and B) which have to be satisfied at any given time. This implies that there are only two degrees of freedom left for composition control.

Step 2 therefore involves answering the following question: Which two independent variables should be “manipulated” for composition control? We have put “manipulated” in quotes, because we are going to define new “manipulated” variables different from the real ones (which are the valve positions). In fact, we have already implicitly redefined the manipulated variables by assuming that we actually can manipulate the flows L , V , V_T and B directly instead of their valve positions. In practice, for L , D and B this may be implemented by measuring the actual flow rate and using a fast inner loop to adjust this measured rate to match the desired flow. By this we also remove the nonlinear relationship between the valve position and the flow rate. It is usually not possible to measure V and V_T and these flows must

be estimated in some other way, for example, by enthalpy balance calculations.

There is clearly an infinite number of relationships between L, V, V_T, D and B which can be defined as new “manipulated” variables. Of these we will only consider the flows themselves and ratios between the flows. A further simplification results because the condenser duty (i.e., V_T) is almost never used for controlling composition (Shinskey, 1984). The problem of dual composition control is then reduced to controlling the compositions y_D and x_B using two independent combinations of the inputs L, V, D and B . Each choice of manipulated inputs corresponds to a specific control configuration. For example, the choice of L and V for composition control is referred to as the LV-configuration.

This paper summarizes work done by the authors for the last two years, and more detailed information are found in [1-7].

2. MODELLING THE COLUMN

A distillation column is strongly nonlinear, but for control design we will describe it by a linear model. This model should reflect the issues most important for control purposes:

- RHP-zeros (time delays, inverse responses) (Holt and Morari, 1985).
- Ill-conditioning (large RGA-elements) (Skogestad and Morari, 1986b).

The flow dynamics are the main source of RHP-zeros. High-purity columns are known to be ill-conditioned at steady-state (Skogestad and Morari, 1986a). However, for control purposes the steady-state is not as important as the behavior around cross-over (the frequency equal to the inverse of the closed-loop time constant). For most columns the condition number ($\gamma(G)$) is smaller at higher frequencies. Still, since the steady-state behavior generally reflects what happens at higher frequency, a good steady-state model is very useful.

It turns out that large RGA-elements, rather than a large value of $\gamma(G)$, implies control problems (Skogestad and Morari, 1986b). It also turns out that the sum of

the absolute values in the RGA ($\|RGA\|_1 = \sum \lambda_{ij}$) (which is independent of scaling) is very close in magnitude to the value of $\gamma(G)$ obtained when it is minimized over all input and output scalings. Therefore, $\|RGA\|_1$ will be used as a measure of ill-conditionedness rather than $\gamma(G)$. Shinskey (1984) uses the steady-state RGA-elements as the main source of information for choosing the control configuration.

2.1 Steady-State Behavior

The linearized-steady state behavior is compactly written in terms of the steady-state gain matrix (G)

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G_{u_1 u_2} \begin{pmatrix} du_1 \\ du_2 \end{pmatrix}, G_{u_1 u_2} = \begin{bmatrix} \left(\frac{\partial y_D}{\partial u_1}\right)_{u_2} & \left(\frac{\partial y_D}{\partial u_2}\right)_{u_1} \\ \left(\frac{\partial x_B}{\partial u_1}\right)_{u_2} & \left(\frac{\partial x_B}{\partial u_2}\right)_{u_1} \end{bmatrix} \quad (1)$$

Here u_1 and u_2 are the manipulated variables chosen for composition control. We will derive expression for estimating $G_{u_1 u_2}$ and show that the steady-state matrices for various configurations are closely related. For example, we can always write

$$G_{u_1 u_2} = G_{LV} M \quad (2)$$

The matrix M usually does not change much with operating conditions (while G_{LV} and $G_{u_1 u_2}$ may change a lot). Two quantities which are easy to compute from the gain matrix are the condition number

$$\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G) \quad (3)$$

and the RGA (\times denotes element-by-element multiplication)

$$RGA = \{\lambda_{ij}\} = G \times (G^{-1})^T, \quad (2 \times 2) : \lambda_{11} = \frac{1}{1 - \frac{g_{12}g_{21}}{g_{11}g_{22}}} \quad (4)$$

2.1.1 Obtaining the Gain Matrix

Three alternative ways to find $G_{u_1 u_2}$ are

1. Experimentally (not recommended) or numerically.
2. Using analytic expressions.

3. From the gain matrix for another configuration (Eq. (2)).

Plant data may be used to adjust the number of theoretical trays in the simulation model, but we do not recommend using them for obtaining the gains. Simulation is generally the preferred approach. However, in this paper the goal is to obtain insight into how the gains vary by deriving simple analytical expressions. In particular, it is easy to derive simple consistency-relationships which may be used to correct data obtained by Method 1.

2.1.2 Internal and External Flows

The key to understanding the steady-state behavior of distillation columns is to understand the difference between external flows (products B and D) and internal flows ("reflux", L and V) and how they affect composition (Rosenbrock, 1962). Changing the external flows will usually have a large effect on the product compositions and both compositions will change in the same "direction". For example, an increase in B will make both products richer in the light component. This corresponds to making one product purer and the other product less pure. On the other hand, if the external flows are fixed, and only the internal flows L and V are changed, the effect on the compositions is much smaller and the compositions will change in opposite directions.

This results in an inherently ill-conditioned plant. By ill- conditioned we mean that the gain of the plant is strongly dependent on the input direction: A change in external flows gives a large change in product compositions; a similar change in internal flows usually has little effect on compositions. In particular, this is the case for columns with both products of high purity. As an illustration consider the column in Table 1 with $z_F = 0.5$, $y_D = 0.99$, $x_B = 0.01$ and $D = B = 0.5$ kmol/min. Assume the distillate flow D is increased by 5% to 0.525 kmol/min. Since there is only 0.5 kmol/min of light component in the feed at least 0.025 kmol/min of this has to be heavy component. The best attainable value for the top composition,

Binary separation, constant molar flows, feed liquid

Relative volatility	$\alpha = 1.5$
No. of theoretical trays	$N = 50$
Feed tray location	$N_F = 21$
Feed rate and composition	$F = 1 \text{ kmol/min}, z_F = 0.5$
Product compositions	$y_D = 0.99, x_B = 0.01$
Product rates	$D = B = 0.5 \text{ kmol/min}$
Reflux rate	$L = 2.71 \text{ kmol/min} (1.39L_{min})$

Table 1. Data for distillation column A.

even with total reflux, is then $y_D = 0.5/0.525 = 0.952$. This is far from the desired $y_D = 0.99$.

More generally, the effect of the external flows on the product compositions is found using

$$\frac{D}{B} = \frac{z_F - x_B}{y_D - z_F} \quad (5)$$

This exact expression can be derived from an overall material balance for the light component. For high-purity columns we find by differentiating (5) that the relative changes in y_D and x_B are extremely sensitive to changes in D/B .

High-purity distillation columns are always sensitive to changes in the external flows, and we can only to a limited degree be made less ill-conditioned by changing the manipulated variables for composition control. However, ill-conditioned plants are not necessarily difficult to control. For example, it can be shown that diagonal input uncertainty gives performance problems only when the elements in the RGA are large (Skogestad and Morari, 1986b). All plants with large RGA-elements are ill-conditioned, but there exist ill-conditioned plants for which the elements in the RGA are small (for example, the DV-configuration) and which are not sensitive to diagonal input uncertainty (see Section 3).

Example. SVD-Analysis of LV-Configuration

The purpose of this example is to show mathematically why high-purity distillation columns are ill-conditioned. Consider the distillation column in Table 1. For the LV-configuration we derive the steady-state gain matrix (Skogestad and Morari, 1986a)

$$G_{LV} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix}$$

The condition number is $\gamma(G_{LV}) = 141.7$ and the 1-1-element of the RGA is $\lambda_{11} = 35.1$. This shows a high degree of “directionality” in the plant. More specific

information is obtained from a SVD of G :

$$G = U\Sigma V^H, \quad \Sigma = \begin{pmatrix} \bar{\sigma}(G) & & \\ & \ddots & \\ & & \underline{\sigma}(G) \end{pmatrix} \quad (6)$$

Equivalently

$$G\bar{v} = \bar{\sigma}(G)\bar{u}, \quad G\underline{v} = \underline{\sigma}(G)\underline{u} \quad (7)$$

$\bar{v}(\underline{v})$: Most (least) effective input direction.

$\bar{u}(\underline{u})$: Most (least) easily affected output direction.

We find

$$\begin{pmatrix} \Delta y_D \\ \Delta x_B \end{pmatrix} = \begin{pmatrix} 0.625 & 0.781 \\ 0.781 & -0.625 \end{pmatrix} \begin{pmatrix} 1.972 & 0 \\ 0 & 0.0139 \end{pmatrix} \begin{pmatrix} 0.707 & 0.708 \\ -0.707 & 0.707 \end{pmatrix}^H \begin{pmatrix} \Delta L \\ \Delta V \end{pmatrix}$$

$\bar{u} \qquad \underline{u} \qquad \bar{\sigma} \qquad \underline{\sigma} \qquad \bar{v} \qquad \underline{v}$

The large plant gain, $\bar{\sigma}(G) = 1.972$, is obtained when the inputs are in the direction

$$\begin{pmatrix} \Delta L \\ \Delta V \end{pmatrix} = \bar{v} = \begin{pmatrix} 0.707 \\ -0.708 \end{pmatrix}. \text{ Since}$$

$$\Delta B = -\Delta D = \Delta L - \Delta V \quad (8)$$

this physically corresponds to the direction with the largest change in the external flows, D and B . From the direction of the output vector \bar{u} we see that the effect of the input \bar{v} is to increase both compositions.

The low plant gain, $\underline{\sigma}(G) = 0.0139$, is obtained for inputs $\begin{pmatrix} \Delta L \\ \Delta V \end{pmatrix} = \underline{v} = \begin{pmatrix} 0.708 \\ 0.707 \end{pmatrix}$. From (8) we see that this corresponds to increasing the internal flows only (while keeping B and D nearly constant). The effect (\underline{u}) is to increase y_D and decrease x_B (i.e., make both products purer). As we would expect such a change takes a large control action (large increase in internal flows).

2.1.3. The Separation Factor S

In order to gain insight we need simple analytical expressions for the gain matrix which capture the essential part of the behavior. Changes in the external

flows lead to large changes in the compositions, and it would be useful to find a function of the compositions which does not change very much in such cases. With such a function in hand we would be able to derive the main effect of the external flows on the compositions without having to know further details about the column. The separation factor (S) turns out to be such a function.

$$S = \frac{y_D(1 - x_B)}{(1 - y_D)x_B} \quad (9)$$

S usually does not change much with operating conditions. For S constant we derive

$$\frac{d(1 - y_D)}{y_D(1 - y_D)} + \frac{dx_B}{x_B(1 - x_B)} = 0 \quad (S \text{ constant}) \quad (10)$$

Since y_D and $1 - x_B$ are close to 1, this means that the relative change of impurities in the top ($1 - y_D$) and bottom (x_B) of the column will be nearly equal, but with different signs. Another advantage of using S is that simple shortcut models may often be expressed in terms of S . For constant relative volatility and infinite reflux Fenske's exact relationship applies

$$S = \alpha^N \quad (11)$$

For finite reflux the following approximation is useful [4]

$$S = \alpha^N \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \quad (12)$$

where subscripts T and B denote the top and bottom of the column. This model gives a good description of how S changes with internal flows, but describes poorly the effect of changes in the external material balance [4]. Fortunately, it turns out that this is of less importance if the model for S is used to obtain estimates for the steady-state gains.

2.1.4. Simple Gain Expressions

The total material balance for light component is

$$Fz_F = Dy_D + Bx_B \quad (13)$$

Let ζ represent any manipulated input or disturbance. Differentiating (13) with respect to ζ gives

$$D \frac{\partial y_D}{\partial \zeta} + B \frac{\partial x_B}{\partial \zeta} = e_\zeta \quad (14)$$

where

$$e_\zeta = -(y_D - x_B) \frac{\partial D}{\partial \zeta} + F \frac{z_F}{\partial \zeta} + (z_F - x_B) \frac{\partial F}{\partial \zeta} \quad (15)$$

Note that the value of $\frac{\partial y_D}{\partial \zeta}$ depends heavily on the particular choice of manipulated variables (u_1 and u_2). Sometimes this is shown explicitly by writing, for example, $(\frac{\partial y_D}{\partial u_1})_{u_2}$. Analytic expressions for the gains are found by combining (14) and (9).

We find for binary mixtures

$$\frac{1}{(1 - y_D)y_D} \frac{\partial y_D}{\partial \zeta} = \frac{1}{I_s} (e_\zeta + Bx_B(1 - x_B) \frac{\partial \ln S}{\partial \zeta}) \quad (16a)$$

$$\frac{1}{(1 - x_B)x_B} \frac{\partial x_B}{\partial \zeta} = \frac{1}{I_s} (e_\zeta - Dy_D(1 - y_D) \frac{\partial \ln S}{\partial \zeta}) \quad (16b)$$

where the "impurity sum" is defined as

$$I_s = Bx_B(1 - x_B) + Dy_D(1 - y_D) \quad (17)$$

For high-purity separations $I_s \approx Bx_B + D(1 - y_D)$. Except for the assumptions about binary mixture and two-product column, these expressions are exact and express in a compact form the major contributions to the steady-state gains. The important point is that the first term in (16a) and (16b) involving e_ζ is often dominating (unless e_ζ is identically zero). Furthermore, this "material balance" term can be evaluated exactly from the product compositions if constant molar flows are assumed. Other interesting insight may also be obtained from these expressions [2].

2.1.5. Consistency Relationships

One important consequence of (14) is that it provides an exact consistency relationship between the column elements for any steady-state gain matrix.

$$D \left(\frac{\partial y_D}{\partial u_1} \right)_{u_2} + B \left(\frac{\partial x_B}{\partial u_1} \right)_{u_2} = -(y_D - x_B) \left(\frac{\partial D}{\partial u_1} \right)_{u_2} \quad (18)$$

(F and z_F are not influenced by changes in u_1 or u_2). The only assumption used when deriving (18) is that the column has two products such that $dD = -dB$. For the case of constant molar flows we derive

$$\text{LV - configuration : } \begin{cases} Dg_{11} + Bg_{21} = (y_D - x_B) \\ Dg_{12} + Bg_{22} = -(y_D - x_B) \end{cases}$$

$$\text{DV - configuration : } \begin{cases} Dg_{11} + Bg_{21} = 0 \\ Dg_{12} + Bg_{22} = -(y_D - x_B) \end{cases}$$

Here $G_{u_1 u_2} = \{g_{ij}\}$. These relationships should be used to check the gain matrix.

2.1.6. The Matrix M

At steady state the following relationships must be satisfied

$$dD = dV - dL \quad (19)$$

$$dB = -dD \quad (20)$$

(19) applies for constant molar flows, but more general expressions are easily derived [2]. (20) implies that D and $-B$ as a manipulated variable always gives the same steady-state gains. Given the gain matrix G_{LV} for the LV-configuration, we find $G_{u_1 u_2}$ (Eq. (1)) by using the following coordinate transformation

$$\begin{pmatrix} dL \\ dV \end{pmatrix} = M \begin{pmatrix} du_1 \\ du_2 \end{pmatrix}, M = \begin{bmatrix} \left(\frac{\partial L}{\partial u_1}\right)_{u_2} & \left(\frac{\partial L}{\partial u_2}\right)_{u_1} \\ \left(\frac{\partial V}{\partial u_1}\right)_{u_2} & \left(\frac{\partial V}{\partial u_2}\right)_{u_1} \end{bmatrix} \quad (21)$$

For any particular choice of u_1 and u_2 , M is easily derived using (19) and (20). Some examples are given in Table 3.

2.1.7. Estimates of the RGA and Condition Number

Using the gains derived in 2.1.4 the following approximation is derived for the condition number for the LD- and DV-configuration for high- purity separations:

$$\gamma(G_{DV}^s) \simeq \gamma(G_{LD}^s) \approx \frac{2}{Bx_B + D(1 - y_D)} \frac{1}{(\partial \ln S / \partial L)_D} \quad (22)$$

G^s is the scaled gain matrix

$$G^s = \begin{pmatrix} 1/1 - y_D^0 & 0 \\ 0 & 1/x_B^0 \end{pmatrix} G \quad (23)$$

u_1	u_2	M
L	V	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
L	$D(-B)$	$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$
$D(-B)$	V	$\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$
D	V/B	$\begin{bmatrix} -1 + \frac{V}{B} & B \\ -\frac{V}{B} & B \end{bmatrix}$
L/D	V	$\begin{bmatrix} \frac{D}{1+L/D} & \frac{L/D}{1+L/D} \\ 0 & 1 \end{bmatrix}$
L/D	V/B	$\frac{1}{r} \begin{bmatrix} D(1 + \frac{V}{B}) & B\frac{L}{D} \\ D\frac{V}{B} & B(1 + \frac{L}{D}) \end{bmatrix}$

$$r = 1 + \frac{L}{D} + \frac{V}{B}$$

Table 3. Matrix M for some choices of manipulated inputs (u_1 and u_2). $G_{u_1, u_2} = G_{LV}M$. Constant molar flows assumed.

This corresponds to using $dy_D/1 - y_D^0$ and dx_B/x_B^0 as outputs. This scaling gives outputs of comparable magnitude. In general, the gain matrix should always be scaled. Reasonably good estimates of $(\partial \ln S / \partial L)_D$ may be obtained from the short-cut model (12). For other configurations we find

$$\gamma(G_{LV}^s) \approx 2\gamma(G_{LD}^s) \quad (24)$$

$$\gamma(G_{\frac{L}{D}\frac{V}{B}}^s) \approx \frac{1}{1 + \frac{L}{D} + \frac{V}{B}}(G_{LV}^s) \quad (25)$$

Note from (22)-(25) that the condition number is always large whenever both products are pure such that $Bx_B + D(1 - y_D)$ is small. However, note that the value can be significantly reduced for the $\frac{L}{D}\frac{V}{B}$ -configuration. Similarly, we can derive expressions for the 1-1-element in the RGA (which is scaling invariant).

$$\lambda_{11}(G_{LD}) \approx 1 - \lambda_{11}(G_{DV}) \approx \frac{1}{1 + \frac{D(1-y_D)}{Bx_B}} \quad (26)$$

$$\lambda_{11}(G_{LV}) \approx \frac{1}{Bx_B + D(1 - y_D)} \frac{1}{(\partial \ln S / \partial L)_D} \approx \gamma(G_{LV}^s) \quad (27)$$

$$\lambda_{11}(G_{\frac{L}{D}\frac{V}{B}}) \approx \frac{\lambda_{11}(G_{LV})}{1 + \frac{L}{D} + \frac{V}{B}} \approx \gamma(G_{\frac{L}{D}\frac{V}{B}}^s) \quad (28)$$

λ_{11} for the LD- and DV-configuration is always less than one. λ_{11} for the LV-configuration is large if both products are of high purity. λ_{11} for the $\frac{L}{V}\frac{D}{B}$ -configuration may be significantly smaller if the reflux is large.

2.2. Dynamic Behavior

The external and internal flows often have drastically different effects on the steady-state behavior, and it is not surprising that they also in such cases give very different dynamic responses. The time constant for changes in the external flows (the dominant time constant, τ_1) is generally significantly larger than the time constant for changes in the internal flows (τ_2). In particular, this is the case for columns with both products of about equal purity. The response to changes in internal flows is therefore faster than the response to changes in the external

flows. The consequence is that the condition number of the plant ($\gamma(G)$) is often much lower at high frequencies than at steady-state, and the column may not be as difficult to control as we may expect from studying the steady-state alone. It also turns out, that while the steady-state gains and steady-state condition number may change considerably with operating conditions, the behavior at higher frequencies ("initial response") does not change very much - at least if scaled gains (Eq. (23)) are used [6]. This means that a linear controller may give satisfactory control.

Below, we will first discuss simple methods for estimating the dominant time constant (τ_1) and then present a simple model which explicitly takes the difference between internal and external flows into account.

2.2.1. The Dominant Time Constant

The dynamics of most distillation columns are dominated by one large time constant, which is nearly the same, regardless of where a disturbance is introduced or where composition is measured. This is well known both from plant measurements (McNeill and Sachs, 1969) and from theoretical studies (Moczek et al., 1963). Because of the nonlinearity the numerical value of this time constant often varies drastically with operating conditions (Fig. 2). Viewed on the basis of the large number of nonlinear differential equations, the simple low order responses observed for most columns is somewhat surprising. Levy et al. (1969) showed that the dominant slowest model primarily involves composition effects and is nearly unaffected by flow dynamics. This leads to the conclusion that the dominating dynamics can be captured by considering the composition dynamics only.

The simplest approach is to consider the total holdup of each component in the column. By assuming that all trays have the same response, this directly leads to a first order model, and the dominating time constant can be estimated. Moczek et al. (1963) used this idea to introduce the "inventory time constant" for a column

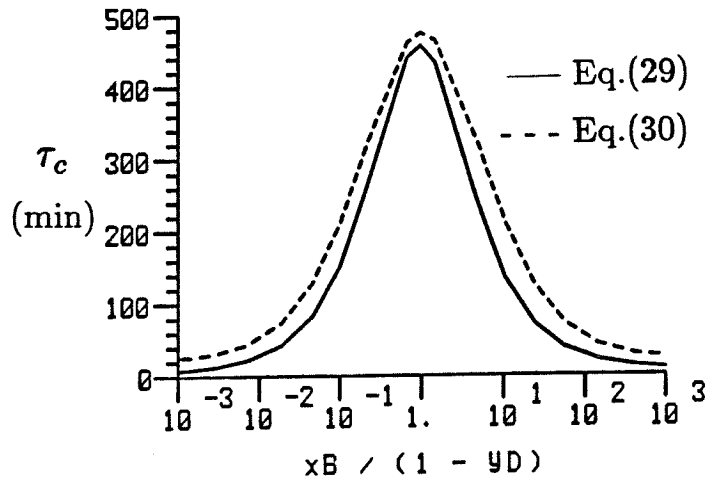


Fig. 2. Typical plot of linearized time constant as a function of product composition
(Plot derived by varying feed composition (z_F) with all flows fixed). See [5].

going from one steady state (subscript 0) to another (subscript f)

$$\tau_c = \frac{\Delta(\sum M_i x_i)}{\Delta(Fz_F) - y_{D0}\Delta D - x_{B0}\Delta B} = \frac{\Delta(\sum M_i x_i)}{D_f \Delta y_D + B_f \Delta x_B} \quad (29)$$

Here Δ represents the change, for example, $\Delta D = D_f - D_0$. A simple interpretation of (29) is

$$\tau_c = \frac{\text{"change in holdup of one component (mol)"}}{\text{"imbalance in supply of this component (mol/s)"}}$$

In spite of the excellent agreement often found between τ_c (29) and τ_1 from the actual response (Moczek et al., 1963) this expression does not seem to have found widespread use.

Very large time constants are found for small perturbations to columns with both products of high purity. The reason is that the compositions inside the column may change significantly (large change in component holdup, $\sum M_i x_i$), while the product compositions may almost be unchanged (resulting in a small value of $D\Delta y_D + B\Delta x_B$). Note that (29) does not apply to changes in the internal flows

because the denominator is zero in this case.

τ_c may be evaluated easily (and accurately) using a steady-state simulation program if the total holdup (M_i) on each stage is known. (Two simulations are needed). However, to gain insight, the following analytical expression, which is based on (29), S constant, and a linearized plant is useful:

$$\tau_c = \tau_{cI} + \tau_{cD} + \tau_{cB} \quad (30a)$$

$$\tau_{cI} \approx \frac{M_I}{I_s \ln S}, \quad \tau_{cD} \approx \frac{M_D(1 - y_D)}{I_s}, \quad \tau_{cB} \approx \frac{M_B x_B}{I_s} \quad (30b)$$

Here $M_I = \sum M_i$ is the total holdup *inside* the column and M_D and M_B are the condenser and reboiler holdups (in kmol). I_s is defined in (17). The time constant given by (30) is for the linearized plant and may therefore be misleading for large perturbations. For the special case of equal purities ($x_B = 1 - y_D$) we find

$$\tau_c \approx \frac{M_I/F}{x_B y_D \ln S} + \frac{M_D}{F} + \frac{M_B}{F} \quad (31)$$

(32) clearly shows that the contribution τ_{cI} from the holdup *inside* the column dominates for columns with both products of high purity.

The estimate (30) is excellent for columns with high reflux and large values of $\lambda_{11}(G_{LV})$. However, the main value of (30) is the insight it gives into the dynamic behavior: Since $\ln S$ usually does not change much with operating conditions, the value of τ_{cI} is mainly determined by I_s , which again is determined by the composition of the least pure product:

$$I_s \approx \begin{cases} D(1 - y_D) & \text{if distillate least pure} \\ Bx_B & \text{if bottoms least pure} \end{cases} \quad (31)$$

τ_{cI} reaches its maximum value approximately when both products have equal purity (Fig. 2). Case studies using linearized models and assuming equal purities are often presented in the academic literature, and the reported values for the time constants are therefore often misleading. For example, this assumption is used by Wahl and

Harriot (1971); the figures they present for estimating time constants are therefore of very limited value. In practice, for such columns, any disturbance or input to the column will take the column to a new steady state where one of the products is less pure, and the actual time constant will be smaller than found by linear analysis.

Example. The column in Table 1 has $x_B = 1 - y_D = 0.01$, $\ln S = 9.19$ and $I_s = x_B y_D = 0.0099$. With $M_I/F = M_D/F = M_B/F = 0.5$ min we get from (30)

$$\tau_c = \frac{39 \cdot 0.5}{0.0099 \cdot 9.19} + 0.5 + 0.5 = 215 \text{ min}$$

Assume there is a change in operating conditions (D/F changes from 0.5 to 0.555) such that $x_B = 0.1$ and $1 - y_D = 0.002$ (column C in Table 2). Then $\ln S = 8.51$ is almost unchanged but $I_s = 0.0510$ increases about five times. We find:

$$\tau_c = \frac{39 \cdot 0.5}{0.051 \cdot 8.51} + 0.98 + 0.02 = 46 \text{ min}$$

Thus, the nominal value of the time constant (215 min) holds only for very small perturbations from the nominal steady-state.

Column :	D/F	y_D	x_B	L/F	$\lambda_{11}(0)$	$\tau_c(\text{Eq. (30)})$
A	0.500	0.99	0.01	2.706	35.1	215min
C	0.555	0.90	0.02	2.737	7.53	46min

Table 2. Effect of change in operating conditions. Other column data as in Table 1.

2.2.2. Simplified Dynamic Model

The reflux (L) and boilup (V) are the only flows which immediately affect the product compositions. The effect of the product flows (D and B) depends heavily on how the level loops are tuned. In order to avoid this dependency on the level loops, we prefer to write the dynamic distillation model in terms of L and V . However, to model explicitly the difference in dynamic behavior between internal and external flows, we will first neglect the effect of flow dynamics and level loops.

Let $\{g_{ij}\}$ denote the elements of G_{LV} . Then at steady state we have

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (33)$$

Written in terms of D and V as manipulated variables (assuming constant molar flows):

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} dV - dD \\ dV \end{pmatrix} = \begin{pmatrix} g_{11} + g_{12} & -g_{11} \\ g_{21} + g_{22} & -g_{21} \end{pmatrix} \begin{pmatrix} dV \\ dD \end{pmatrix}$$

Next, assume that the effect of changes in external flows (D) is given by a first-order response with time constant τ_1 , and the effect of changes in internal flows (change V keeping D constant) has time constant τ_2 . Then

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} + g_{12} & -g_{11} \\ g_{21} + g_{22} & -g_{21} \end{pmatrix} \begin{pmatrix} dV/(1 + \tau_2 s) \\ dD/(1 + \tau_1 s) \end{pmatrix} \quad (34)$$

Switching back to L and V as manipulated input (still assuming immediate flow responses such that $dD = dV - dL$):

$$dy_D = \frac{g_{11}}{1 + \tau_1 s} dL_T + \left(\frac{g_{11} + g_{12}}{1 + \tau_2 s} - \frac{g_{11}}{1 + \tau_1 s} \right) dV_T \quad (35a)$$

$$dx_B = \frac{g_{21}}{1 + \tau_1 s} dL_B + \left(\frac{g_{21} + g_{22}}{1 + \tau_2 s} - \frac{g_{21}}{1 + \tau_1 s} \right) dV_B \quad (35b)$$

Here we have also introduced the “local” flows of L and V in order to take into account the effect of the flow dynamics. $dL_T = dL$ is the actual manipulated reflux and $dV_B = dV$ is the actual manipulated boilup. We have (Rademaker et al., 1975)

$$dL_T \approx dL \quad (36a)$$

$$dV_T \approx dV_B \approx dV \text{ (assuming immediate pressure control)} \quad (36b)$$

$$dL_B \approx e^{-\theta_L s} dL + (1 - \lambda + \lambda e^{-\theta_L s}) dV \quad (36c)$$

Here

$$\lambda = (\partial L_i / \partial V)_{M_i} \quad (37a)$$

$$\theta_L = N_{trays} \tau_L, \quad \tau_L = (\partial M_i / \partial L_i)_V \quad (37b)$$

A typical value for θ_L is $\frac{1}{3} \frac{M_L}{L}$. λ is difficult to estimate. If $\lambda > 0.5$, then x_B will show an inverse response for an increase in the boilup (V).

τ_1 is generally close to the time constant corresponding to the smallest eigenvalue, or may be approximated by τ_c as outlined in 2.2.1. τ_2 may be estimated by matching the RGA-values at high frequency [6]. We find

$$\frac{\tau_2}{\tau_1} \approx \frac{L}{F} \left(\frac{g_{12}}{g_{11}} - \frac{g_{22}}{g_{21}} \right) + \left(1 + \frac{g_{12}}{g_{11}} \right) \quad (38)$$

τ_2 may also be obtained from simulations (without flow dynamics) of changes in the internal flows. In most cases it will be very difficult to obtain τ_2 using plant data, since it is almost impossible, in practice, to avoid changes in the external flows (because of disturbance in feed rate, boilup, etc.).

Example. Consider again the column in Table 1. Neglecting flow dynamics (i.e., assume constant holdup (M_i) on all stages) results in a dynamic model with 41 states. The effect of a small change in external and internal flows is shown in Fig. 3 for the case with $M_i/F = 0.5$ min, on all stages (including reboiler and condenser). The responses are almost identical to those found using the simple linear model (35) with $\tau_1 = 194$ min and $\tau_2 = 15$ min (The estimates found using (30) and (38) are $\tau_1 = 215$ min and $\tau_2 = 18$ min):

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \frac{1}{1 + 194s} \begin{pmatrix} 0.878 & -0.864 \frac{1+12.1s}{1+15s} \\ 1.082 & -1.096 \frac{1+17.3s}{1+15s} \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (39)$$

In Fig. 4 $\|RGA\|_1$ is plotted as a function of frequency. The condition number ($\gamma(G)$) is not shown, but its value is almost identical to $\|RGA\|_1$ [7]. The RGA-elements are large at low frequency ($\lambda_{11} \approx 35$ at steady-state), but much smaller

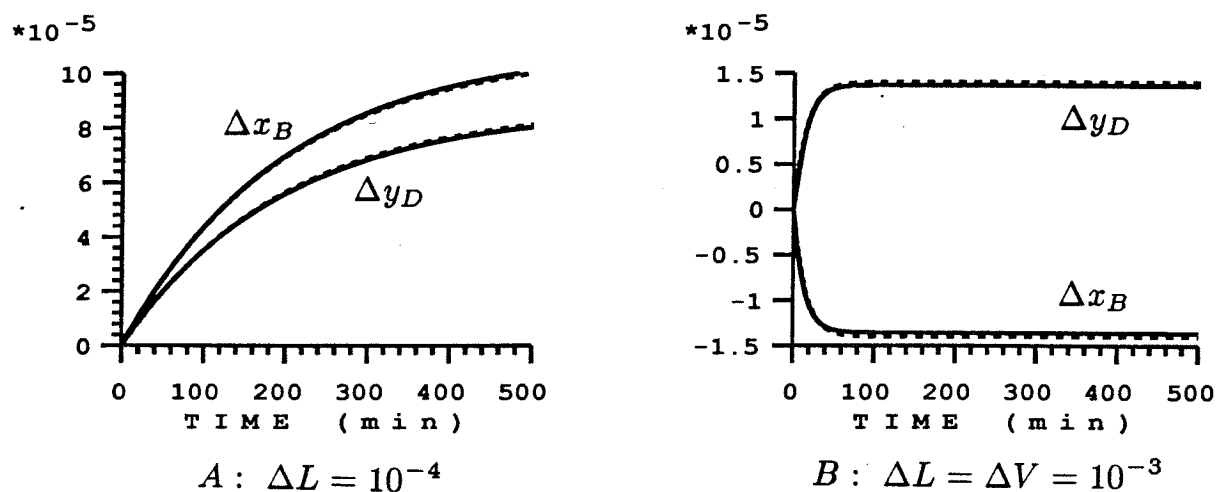


Fig. 3. Column A. Responses to small change in external (A) and internal (B) flows.

Dotted lines: Approximation with simplified model (39).

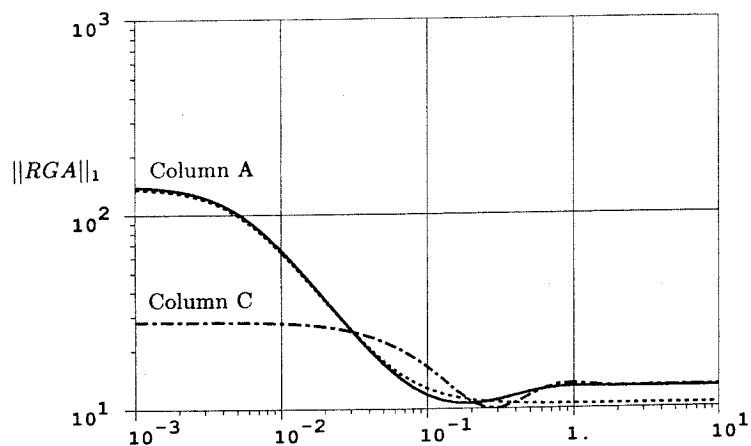


Fig. 4. $\|RGA\|_1$ as a function of frequency. Dotted line: Approximation for column A with model (39). ($M_i/F = 0.5$ min)

at high frequency ($|\lambda_{11}| \approx 3.5$ at $\omega = 0.1$). The simple model (39) captures this behavior very well. The column is therefore not as difficult to control as might be expected from the steady-state data.

Note that without the “correction terms” $\frac{1+12.1s}{1+15s}$ and $\frac{1+17.3s}{1+15s}$ the plant (39) would have the same condition number (and RGA) at all frequencies. Each transfer matrix element is often appreciated by a first-order lag with time delay ($ke^{-\theta s}/1 + \tau s$). It is clear that unless special care is taken, it is extremely unlikely that this model will capture the change in condition number with frequency.

We know that the time constant of 194 min used in (39) only holds for small perturbations from the nominal steady-state. Similarly, the steady-state gains vary a lot with operating conditions. However, it turns out that the column behavior at higher frequencies (“initial response”) changes much less. This is illustrated by the plots of $\|RGA\|_1$ for columns A and C in Fig. 4. Since the model (39) correctly predicts the behavior at high frequency it is useful also for large perturbations from the nominal steady state [6,7].

3. CHOICE OF CONTROL CONFIGURATION

The most important step when designing a composition control system is probably to choose which two independent combinations of L, V, D and B to use for composition control.

3.1. L, V, D and B as Manipulated Inputs

Let us first consider only the flows L, V, D and B themselves. There are $\binom{4}{2} = 6$ independent pair-combinations. However, only five of these are possible since D and B cannot be used together for composition control, because of the steady state material balance constraint $D + B = F$. Having chosen one of the remaining five pairs (LV, LD, LB, DV, VB), the control structure for the level loops usually follows easily.

Example. LV-Configuration. Assume L and V have been chosen for composition

control.* The condenser level may be controlled by D , pressure by V_T and the reboiler level by B resulting in the following control structure:

$$\begin{bmatrix} dL \\ dV \\ dD \\ dB \\ dV_T \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ K & 0 & 0 \\ 0 & 0 & c_D(s) & 0 & 0 \\ 0 & 0 & 0 & c_B(s) & 0 \\ 0 & 0 & 0 & 0 & c_V(s) \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \\ dM_D \\ dM_B \\ dM_V \end{bmatrix}$$

3.2. Ratios between L, V, D and B as Manipulated Inputs

Of the possible nonlinear relationships possible between L, V, D and B we will only consider ratios. These seem to be the only nonlinear combinations used in practice (Shinskey, 1984). The total number of independent ratios is six. They are

$$\frac{L}{V}, \frac{L}{D}, \frac{L}{B}, \frac{V}{D}, \frac{V}{B} \text{ and } \frac{D}{B}$$

Including the four flows themselves this results in $\binom{10}{2} = 45$ independent pairs of “manipulated” variables. Again, combinations of D, B , and $\frac{D}{B}$ cannot be used for composition control. This eliminates three of these options, but still leave us with 42 possible combinations.

If we look at the actual implementation there are even more than 42 options. Since the true manipulated variables are always L, V, D and B , we have to determine how L/V , for example, is implemented as a “manipulated” variable. To increase L/V we may either increase L , decrease V or change both at the same time. If the flow dynamics and level controls were immediate, these different implementations would not affect the composition response, but, because they are not it does make a difference. We adopt the following convention: Writing the ratio between L and V as $\frac{L}{V}$ means that L is manipulated to change $\frac{L}{V}$, and writing $\frac{V}{L}$ means that V is manipulated to change the ratio.

Ratio control systems have been used in the industry for at least forty years (Rademaker et al. (1975), p. 445). Yet, almost no discussion is found in the literature on why such schemes may be beneficial. The simplest justification for using

* This is the configuration most commonly used (Rademaker et al., 1975)

ratios as inputs follows from steady-state considerations: To keep the compositions constant, the ratio $\frac{L}{V}$ inside the column (slope of the operating line on the McCabe-Thiele diagram) should be constant. Intuitively, it seems that some disturbances may be counteracted by keeping this ratio constant. However, these arguments do not explain what happens when ratios are used for closed-loop control of compositions. Furthermore, the effect of using a given ratio depends entirely on which second manipulated variable is chosen for composition control.

Nonlinear Implementation. Clearly, using ratios as “manipulated” variables is a way of introducing a simple nonlinear control scheme. For example, the nonlinear implementation of L/D as a manipulated variable is (using the convention introduced above)

$$L = \left[\frac{L}{D} \right] D \quad (40)$$

The corresponding linear implementation is

$$dL = D d \left[\frac{L}{D} \right] + \frac{L}{D} dD \quad (41)$$

or equivalently

$$L = L_0 + D_0 \Delta \left[\frac{L}{D} \right] + \frac{L_0}{D_0} \Delta D \quad (42)$$

where we have used subscript 0 to denote the nominal values explicitly. The difference between (40) and (42) is important only if D and L/D change significantly with operating conditions. Because this is usually not the case there are only minor differences between the linear and nonlinear implementation. We would like to understand what kind of linear control system the ratio schemes correspond to. (Surprisingly, the results presented here on the linear interpretation of the ratio schemes seem to be new). To this end consider the following example.

Exampe. $\frac{L}{D} \frac{V}{B}$ -Configuration. The $\frac{L}{D} \frac{V}{B}$ -configuration is claimed by Shinskey (1984) to be applicable over the broadest range of cases and also Rademaker et al. (1985, p. 463) recommend this scheme. According to the convention introduced above dL

and dV are manipulated to change the ratios L/D and V/B . Linear analysis gives

$$dL = Dd\left[\frac{L}{D}\right] + \frac{L}{D}dD \quad (43)$$

$$dV = Bd\left[\frac{V}{B}\right] + \frac{V}{B}dB$$

Consequently, dL and dV depend on $d[\frac{L}{D}]$ and $d[\frac{V}{B}]$ (which are “manipulated” based on the compositions y_D and x_B) and on the flow rate changes dD and dB . Because L and V are manipulated for composition control, the condenser and reboiler levels are controlled using D and B , i.e.,

$$dD = c_D(s)dM_D \quad (44)$$

$$dB = c_B(s)dM_B$$

(The SISO controllers $c_D(s)$ and $c_B(s)$ are in many cases simple proportional controllers). Let the composition controller (possibly multivariable) be

$$\begin{pmatrix} d\frac{L}{D} \\ d\frac{V}{B} \end{pmatrix} = \hat{K} \begin{pmatrix} dy_D \\ dx_B \end{pmatrix} \quad (45)$$

and define

$$K = \begin{pmatrix} D & 0 \\ 0 & B \end{pmatrix} \hat{K} \quad (46)$$

The overall controller found by combining (43)-(46) is

$$\begin{bmatrix} dL \\ dV \\ dD \\ dB \end{bmatrix} \begin{bmatrix} K & \frac{L}{D}c_D & 0 \\ 0 & 0 & \frac{V}{B}c_B \\ 0 & 0 & c_D \\ 0 & 0 & c_B \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \\ dM_D \\ dM_B \end{bmatrix} \quad (47)$$

We see from (47) that the flow rates L and V are manipulated based both on the product compositions (y_D and x_B), and on the levels (M_D or M_B). Furthermore, the two SISO level controllers ($c_D(s)$ and $c_B(s)$) each manipulate two flow rates, and therefore appear at two places in the transfer matrix for the overall controller.

Thus, the use of ratios as manipulated variables introduce in an ad-hoc manner a multivariable control system.

In most cases the major effect of using ratios for composition control is captured by the linear analysis. Ratios do not tend to correct the nonlinear behavior of distillation columns because the manipulated inputs vary only moderately with operating conditions (neglecting startup). On the other hand, the product compositions do often vary significantly with operating conditions, and a significant “linearization” effect may be obtained, for example, by using $\ln(1 - y_D)$ and $\ln x_B$ as “redefined” controlled outputs [6,7].

3.3 Differences Between Control Configurations

Assuming immediate flow responses, perfect level control and constant molar flows we have in the absence of feed disturbances

$$dD = dV - dL \quad (19)$$

$$dB = -dD \quad (20)$$

These two equations suggest that any pair of input variables has the same effect: Changing L and V , for example, is equivalent to changing V and D or V and B . Consequently, we might expect to get good and almost identical control performance for any choice of control configuration. However, there are at least six reasons for why the choice of control configuration can make a significant difference:

1. “Uncertainty”
2. Dynamic considerations
3. Rejection of flow disturbances
4. One-point (“manual”) composition control
5. Changes between “manual” and “automatic”
6. Constraints

In many cases conflicting conclusions arise from these considerations, and the engineer has to perform a more detailed analysis or use his judgment in making

the final choice. In this paper we will discuss these issues only briefly. Let us first review the RGA which is used extensively by Shinskey (1984) to compare control configurations.

The RGA. From Shinskey's book the reader is led to believe that the RGA is useful because it provides a measure of interactions when using a decentralized controller. His rule (though he does not express it explicitly) is to choose a configuration with λ_{11} in the range of about 0.9 to 4 (Shinskey, 1984, Table 5.2). If λ_{11} were used only as an interaction measure this recommendation would not make any sense; in this case λ_{11} should be chosen to be as close to one as possible and $\lambda_{11} = 0.67$ would be almost equivalent to $\lambda_{11} = 2$ (both have $|\frac{g_{12}g_{21}}{g_{11}g_{22}}| = 0.5$). Consequently, Shinskey's use of the RGA is a way of categorizing his experience on distillation columns, rather than expressing the effect of interactions. In fact, his rules also apply when a multivariable controller is used. His recommendations regarding the RGA should therefore only be used for distillation columns. One objective of this section is to provide some justification for Shinskey's rules.

3.3.1. Uncertainty

Since we are considering different choices of manipulated inputs, the uncertainty associated with these manipulated inputs may cause different control behavior. The presence of input uncertainty favors using configurations with small elements in the RGA (Skogestad and Morari, 1986b). In general, all configurations involving D or B have $|\lambda_{11}| < 1$, while all others have $|\lambda_{11}| > 1$ (Shinskey, 1984, p. 146). The LV-configuration generally has the largest RGA-elements. Any configuration which uses D or B is therefore insensitive to input uncertainty, but the ratios $\frac{L}{D}$, $\frac{V}{B}$, $\frac{L}{B}$ or $\frac{V}{D}$ (or their inverses) may also be a good choice for columns with high reflux.

A simple physical reason for why input uncertainty may cause control problems is the following: As discussed in Section 2 the plant is very sensitive to changes in

the external flows, but rather insensitive to changes in the internal flows. To have tight control, the controller should counteract this, by making large changes in the internal flows (ΔL and ΔV), while keeping the changes in the external flows ($\Delta B = -\Delta D$) small. However, if the LV-configuration is used, uncertainty with respect to the values of ΔL and ΔV makes it impossible to keep $\Delta B = \Delta L - \Delta V$ small. The result is large (and undesired) changes in ΔB and ΔD , which result in large (and undesired) changes in the compositions because of the high plant gain in this direction. This is consistent with the large RGA-elements found for the LV-configuration. On the other hand, it is much easier to avoid changes in D or B if one of these flows is manipulated directly, and this is consistent with the small RGA-values found in this case.

3.3.2. Dynamic Considerations

These issues are addressed in detail in the literature (Rademaker et al., 1985, Shinskey, 1984), and only a short summary is given here: L and V should be manipulated directly for composition control to get a fast initial response. This is probably one of the main reasons for the popularity of the LV-configuration. The $\frac{L}{D} \frac{V}{B}$ -configuration also has this feature. Use of D or B for composition control is generally not recommended if a fast initial response is desired.

3.3.3. Rejection of Flow Disturbances

The major flow disturbances are in the feed rate (F), feed enthalpy (q_F), boilup (V), condenser vapor rate (V_T) and reflux temperature. There will also be disturbances in L , D and B (e.g., due to measurement noise), but those are usually of less importance. Three ways of handling flow disturbances are

1. feedforward control
2. through their effect on composition
3. through their effect on levels and pressure

The first option is possible only if the disturbance can be measured. The level and

pressure loops are usually much faster than the composition loops, and intuitively it seems preferable to try to reject the flow disturbances with the level loops (Option 3). However, since any flow disturbance which is not rejected by the level loops will result in a upset in composition, one may argue that the composition control system may as well take care of all disturbances (Option 2). The problem is that it may not be possible to tune the composition loops sufficiently fast to get acceptable response for large disturbances. This is in particular the case if L and V are not manipulated directly for composition control. Furthermore, by using Option 3 we retain some disturbance rejection capability in the case the composition loops are in "manual". As illustrated by the following example an important feature of some of the ratio control schemes is that they have a good "built-in" rejection of flow disturbances.

Example. Assume the feed is liquid and consider a feed flow disturbance. If the LV- or DV-configuration is used, this disturbance will immediately give an increase in bottoms flow rate (B), leading to a large upset in x_B and y_D . However, if the $\frac{L}{D} \frac{V}{B}$ -configuration is used, all flows are adjusted proportionally, and the effect on compositions is very small: The increased feed flow rate initially brings light components down the column which would increase x_B . However, it also leads to an increase in reboiler level. From (47) we see that this leads to a simultaneous increase in B and V (while the LV- and DV- configurations keeps V constant). The increased boilup (V) returns light components to the column, and counteracts the initial effect the increased feed flow had on compositions. Furthermore, the increase in V leads to an increase in distillate flow (D). The feed flow disturbance is therefore distributed to both products, and $\frac{D}{B}$ is kept unchanged.

Summary [5]. It is preferable to use the level control system to reject flow disturbances. V_T is usually used for pressure control, and disturbances in condenser duty are rejected perfectly (at least at steady state). However, no configuration

can reject all flow disturbances using the level control system: The commonly used LV-configuration does not reject disturbances in F, V, L and q_F . Configurations using D or B as one of the manipulated variables for composition control are insensitive to disturbances in V, L and q_F , but do not reject disturbances in F . (However, F is often measured and a feedforward control scheme may be used). The $\frac{L}{D} \frac{V}{B}$ -configuration is insensitive to disturbances in F , and also rejects other flow disturbances also well, provided the reflux is large.

3.3.4. One-Point ("Manual") Composition Control

Very few distillation columns are actually operated with a "two- point" control system. In most cases one of the compositions is controlled manually - at least part of the time. Since the operators do not monitor the compositions continually and manipulate the inputs accordingly, it is important that the effect of expected disturbances on the manually controlled ("uncontrolled") composition is as small as possible.

Both composition loops open. This issue was just discussed for the case of flow disturbances and the $\frac{L}{D} \frac{V}{B}$ -configuration was found to give good disturbance rejection. However, a feed composition (z_F) disturbance has no direct effect on the flows. Consequently, if both composition loops are "open", the effect of a feed composition disturbance will be the same for all configurations. Furthermore, the effect will usually be large because a change in feed composition requires a change in $\frac{D}{B}$ (Eq. (5)), and if this correction is not made, large changes in y_D and x_B will result for high-purity separations. Therefore, at least one of the compositions has to be controlled carefully, either by a feedback controller or by the operator.

One-point composition control (one composition loop open). Assume we have closed one loop, and are using u_2 to control y_2 . The output y_1 is not controlled and the manipulated input u_1 is constant. What is the effect of a disturbance d on the uncontrolled output y_1 ? First consider the steady-state where we have perfect

control of y_2 . The disturbance d has the effect $\begin{bmatrix} g_{1d} \\ g_{2d} \end{bmatrix}$ on the outputs when the inputs u_1 and u_2 are constant. Using deviation variables we have

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = G \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} g_{1d} \\ g_{2d} \end{bmatrix} d \quad (48)$$

Solving for $y_2 = 0$ and $u_1 = 0$ gives

$$\frac{y_1}{d} = -\frac{g_{12}}{g_{22}}g_{2d} + g_{1d} \quad (49)$$

Consequently, the disturbance will not affect the uncontrolled output y_1 if

$$\frac{g_{12}}{g_{22}} = \frac{g_{1d}}{g_{2d}} \quad (50)$$

This result should be obvious: If the disturbance has the same relative effect ($\frac{g_{1d}}{g_{2d}}$) on the outputs as the input u_2 ($\frac{g_{12}}{g_{22}}$), then we can get perfect disturbance rejection by using only this input.

Example. Consider again the column in Table 1. For a feed composition (z_F) disturbance all configurations have $\frac{g_{1d}}{g_{2d}} = 0.787$. The ratio $\frac{g_{1i}}{g_{2i}}$ to $\frac{g_{1d}}{g_{2d}}$ (denoted r_i) is given in below for various configurations ($u_1 u_2$). If this ratio is close to one then perfect disturbance rejection is achieved with u_i alone (the other input being constant).

	LV	$\frac{L}{D} \frac{V}{B}$	$\frac{L}{D} V$	$\frac{L}{D} D$	DV	LD
$r_1 =$	1.03	1.24	1.03	-1.27	1.03	-1.27
$r_2 =$	1.00	0.85	0.85	0.85	-1.27	1.00

D (or B) should obviously never be held constant. Configurations which keep L or V constant come out favorably. The same conclusion holds also for a feed flow disturbance. The reason is that in both cases the major effect of the disturbances may be counteracted by changing the product flow rates (adjusting $\frac{D}{B}$ to satisfy (5)).

Summary. Operating both composition loops open is not acceptable because no correction can be made for feed composition disturbances. When one-point composition control is used, reasonably good control of the "uncontrolled" composition

is maintained with most configurations, provided D or B are not kept constant. The LV -configuration (keeping L or V constant) comes out favorably when only steady-state considerations are taken into account, but it may be preferable to use one of the ratio control schemes (e.g., $\frac{L}{V} \frac{V}{B}$) in order to obtain better dynamic rejection of flow disturbances. One advantage of controlling only one composition is that tuning is simple and very tight control can be maintained for this composition.

3.3.5. Changes between “manual” and “automatic” control.

Changing one of the composition loops between “manual” and “automatic” control is frequently done when controlling distillation columns, for example, due to stability problems, constraints or failures in measurements or actuators. It is clearly desirable to be able to do this without upsetting the rest of the system or having to retune the controllers.

Configurations which use D or B may give very poor response for the uncontrolled composition when the loop involving D or B is put in manual. (This is the opposite of what one might expect from the RGA, since one can always choose pairings such that $0.5 < \lambda_{11} < 1$ in this case). The LV - and $\frac{L}{D} \frac{V}{B}$ -configurations which are preferable for one-point composition control (Section 4.5), are also most easily changed between manual and automatic (though the response for the controlled composition may deteriorate when the other loop is closed).

3.3.6. Constraints

Avoiding constraints. Constraints on flow rates or on holdups (level and pressure) may also be important when choosing the best configuration. Whenever a manipulated input hits a constraint, it is no longer useful for control purposes. Since level and pressure control always has to be maintained, this means that one of the product compositions can no longer be controlled. If a constraint on a flow used for composition control is reached and two-point composition control is still maintained, then the constraint is akin to input uncertainty. Therefore, constraints

are an additional reason for not using controllers with large RGA-elements (for example, a decoupler for the LV-configuration).

Flows used for level control will usually have the largest variations in magnitude, and are most likely to hit constraint. This leads to the following conclusions:

- A very small flow should not be used to control level. One example documented in the literature (McNeill and Sacks, 1969) is the use of distillate D to control M_D in a high reflux column with $L/D = 70$. This is clearly next to impossible.

Operating at Constraints. Many industrial columns are operated at their capacity limit, usually with respect to the boilup V , the reflux L , or the condensation rate V_T . This is another reason for why many columns are operated with only one composition being controlled. Fortunately, as pointed out in above, keeping L , V or V_T constant will also result in reasonably small variations in the uncontrolled product - at least at steady state.

3.3.7. Choice of Control Configuration. Conclusion

Two-Point Composition Control. The RGA is a useful tool for addressing the issue of input uncertainty. Configurations with large values of λ_{11} should be avoided. For distillation columns all material- balance configurations (using D or B) have $\lambda_{11} < 1$. However, these configurations often result in a poor dynamic response and give very poor disturbance rejection if the loop involving D or B is taken out of service. This is probably the reason for why Shinskey (1984) recommends avoiding configurations with $\lambda_{11} < 1$ (provided λ_{11} is not too large). (These considerations only hold for distillation column control, and for other processes there is no reason to try to avoid $\lambda_{11} < 1$).

One-point composition control (one loop in manual). Most industrial columns have closed-loop control of only one composition. This may seem suboptimal, but is in many cases reasonable, since one product is usually much more important than the other. Furthermore, if the column is operating at its capacity limit (which is often

the case), it is impossible to control more than one composition. Uncertainty does not pose any particular problem when only one composition is controlled. Reasonably good control of the uncontrolled composition is maintained provided D or B is not kept constant. The LV- and $\frac{L}{D}\frac{V}{B}$ -configurations will generally both perform satisfactory. The $\frac{L}{D}\frac{V}{B}$ - configuration is preferable because it has a better "build-in" rejection of flow-disturbances which leads to less variations in the uncontrolled composition. The only case when it may be worthwhile to use D or B as the manipulated input for one-point composition control, is for columns with very large reflux ($\frac{L}{D} \gg 1$ or $\frac{V}{B} \gg 1$) where level control using D or B may be almost impossible.

4. CONTROLLER DESIGN / IMPLEMENTATION

Given a good model of the plant, and with an appropriate choice of control configuration, the controller design itself should not be too difficult. Firstly, the level control system has to be designed ("closing the material balance"). This is usually reasonably straightforward once the choice of configuration has been made. Secondly, the 2×2 controller (K) for composition control must be designed. The academic literature has discussed for years whether to use "decouplers" as part of the controller. Recent work (Skogestad and Morari, 1986b) shows clearly that decouplers should only be used for plants with small RGA-elements. For distillation column control, this implies that decouplers should not be used for the LV-configuration, but may be helpful in improving the response if, for example, the DV-configuration is used. The presence of input uncertainty is important for these conclusions.

An important issue with regard to implementation is composition measurements. Often such measurements are not available for on-line control, or the measurements are delayed. In such cases temperature and pressure measurements are used to estimate the product compositions. A fairly sophisticated non-linear model may be needed to obtain acceptable estimates. This may be a tray-by-tray model of the column, or correlations between temperature profiles, pressure and compo-

sition obtained from simulations and plant measurements. The temperatures may also be used directly by the controller. However, this results in a more complicated controller, and may not give as good results as using a separate "estimator", since the controller is usually restricted to being linear.

References

- Bristol, E. H., "On a New Measure of Interactions for Multivariable Process Control", *IEEE Trans. Automatic Control*, **AC-11**, 133-134 (1966).
- Dartt, S. R., "A Survey on Process Control Applications Needs", *Chemical Engineering Progress*, 11-14 (December 1985).
- Holt, B. R. and M. Morari, "Design of Resilient Processing Plants V - The Effect of Deadtime on Dynamic Resilience", *Chem. Eng. Sci.*, **40**, 1229-1237 (1985).
- Levy, R. E. , A. S. Foss and E. A. Greens II, "Response Modes of a Binary Distillation Column", *Ind. Eng. Chem. Fundam.*, **8**, 765-776 (1969).
- McNeill, G. A. and J. D. Sachs, "High Performance Column Control", *Chemical Engineering Progress*, **65**, 3, 33-39 (1969).
- Moczek, J. S., R. E. Otto and T. J. Williams, "Approximation Model for the Dynamic Response of Large Distillation Columns", *Proc. 2nd IFAC Congress*, Basel (1963). Also published in: *Chem. Eng. Prog. Symp. Ser.*, **61**, 136-146 (1965).
- Rademaker, O., J. E. Rijnsdorp and A. Maarleveld, "*Dynamics and Control of Continuous Distillation Units*", Elsevier, Amsterdam (1975).
- Rosenbrock, H. H., "The Control of Distillation Columns", *Trans. Inst. Chem. Engrs.*, **40**, 35-53 (1962).
- Shinskey, F. G., *Distillation Control*, 2nd Ed., McGraw-Hill, New York (1984).
- Skogestad, S. and M. Morari, "Control of Ill-Conditioned Plants: High-Purity Distillation", paper 74a, AIChE Annual Mtg., Miami Beach (1986a).
- Skogestad, S. and M. Morari. "Implication of Large RGA-Elements on Control Performance", paper 6d, AIChE Annual Mtg., Miami Beach (1986b).
- Stanley, G. T. and T. J. McAvoy, "Dynamic Energy Conservation Aspects of Distillation Column Control", *Ind. Eng. Chem. Fundam.*, **24**, 439-443 (1985).
- Wahl, E. F. and P. Harriot, "Understanding and Prediction of the Dynamic Behavior of Distillation Columns", *Ind. Eng. Chem. Process Des. Develop.*, **9**, 396-407 (1970).

Papers by Skogestad and Morari in preparation :

- [1] Control Configuration Selection for Distillation Columns
- [2] Understanding the Steady-State Behavior I.
- [3] Understanding the Steady-State Behavior II.
- [4] Shortcut models - I. Steady-State Behavior.
- [5] Shortcut models - II. Dynamic Behavior.
- [6] Understanding the Dynamic Behavior.
- [7] LV-control of a High-Purity Distillation Column.

Chapter IX

SHORTCUT MODELS FOR DISTILLATION COLUMNS -

I. STEADY-STATE BEHAVIOR

SHORTCUT MODELS FOR DISTILLATION COLUMNS -

I. STEADY-STATE BEHAVIOR

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January 1987

Submitted to Computers and Chemical Engineering

Abstract

Jafarey et al. (1979) derived a simple analytical expression for the separation factor S . This paper provides a simpler and more instructive derivation of this expression, and also evaluates its validity for estimating steady-state gains and RGA-values. The results show that the expression gives a good estimate of how S changes with internal flows, but describes poorly its variation with external flows. Fortunately, when computing the steady-state gains, this error is often not important.

1. INTRODUCTION

Shortcut models. Simple "shortcut" models are useful for analyzing and understanding the strongly nonlinear behavior of distillation columns. For robust distillation column control, for example, it is important to know the correlation between the transfer matrix elements as they vary with operating conditions. The essential part of this behavior may be predicted from simple analytical models.

This paper presents a derivation and evaluation of analytical expressions for the separation factor S which is useful for steady-state calculations. In Part II (Skogestad and Morari, 1987a) we discuss the dynamic behavior and present an analytical expression for the dominant time constant. Though our main goal is to develop shortcut models for use in control system design, the results presented in these two papers are useful for obtaining general insight into distillation column behavior. For example, simple analytical models are invaluable as a teaching tool; a basic understanding of distillation columns is much more easily acquired by analyzing simple analytical models than by running tray-by-tray simulation programs.

We stress that the shortcut models presented are not intended to replace tray-by-tray simulations; for design purposes we recommend that more accurate models are obtained using simulation programs.

Steady-state behavior. A simple nonlinear steady-state model is useful for deriving analytical expressions for the gains g_{ij} between the process inputs (e.g., reflux L and boilup V for the LV-configuration) and the controlled outputs (top and bottom compositions, y_D and x_B).

$$dy_D = g_{11}dL + g_{12}dV$$

$$dx_B = g_{21}dL + g_{22}dV \tag{1}$$

To derive these gains analytically, two equations relating y_D and x_B to L and V and other operating variables for the column are needed. One is given by the overall

material balance for the light component

$$Fz_F = Dy_D + Bx_B \quad (2)$$

and the other is conveniently written in terms of the separation factor S (Shinskey, 1984)

$$S = \frac{y_D/1 - y_D}{x_B/1 - x_B} \approx \frac{1}{x_B(1 - y_D)} \quad (3)$$

(the approximation holds for high-purity separations with $y_D \approx 1$ and $(1 - x_B) \approx 1$).

S is in general a complex function of the operating and design variables for the column (flow rates, feed composition, VLE-data for system, number of trays, feed location, etc.). There are three reasons for choosing to express the column behavior in terms of S :

- (1) Reasonably simple and reliable expressions may be derived for S . This is motivated by Fenske's total reflux equation $S = \alpha^N$ which is exact for mixtures with constant relative volatility α .
- (2) Usually S does not change much with operating conditions. The essential part of the variation of the gains g_{ij} with operating conditions may therefore be captured by assuming S is constant.
- (3) S as defined by (3) is a function of y_D and x_B only. This makes it simple to derive analytical expressions for the steady-state gains.

There are other ways of expressing shortcut models for distillation columns, besides using S . For example, Gilliland's empirical correlation relates $\frac{R-R_{min}}{R+1}$ to $\frac{N-N_{min}}{N+1}$, and analytical equations have been derived to fit the correlation. However, this form is not convenient for obtaining analytical expressions for the gains, since it does not satisfy (2) and (3) above.

The separation factor S has been used extensively by Shinskey (1984) for predicting the steady-state value of the RGA. The 1-1 element of the RGA is

$$\lambda_{11} = \left(1 - \frac{g_{12}g_{21}}{g_{11}g_{22}}\right)^{-1} \quad (4)$$

For columns with both products of high purity Skogestad and Morari (1987b) derived the following approximations for the LV - and DV - configurations

$$\lambda_{11}(G_{LV}) \approx \frac{1}{D(1-y_D) + Bx_B} \frac{1}{(\partial \ln S / \partial L)_D} \quad (5)$$

$$\lambda_{11}(G_{DV}) \approx \frac{1}{1 + \frac{Bx_B}{D(1-y_D)}} \quad (6)$$

Consequently, a good model for S is needed to evaluate the RGA for the LV-configuration, but is less important for the DV-configurations. Also note that λ_{11} for the LV-configuration may be very large if both products are of high purity, while the value for the DV-configuration is always between 0 and 1.

Shinskey uses the simple analytical model for S developed by Jafarey et al. (1979) to estimate the RGA. In their paper, Jafarey et al. only checked the validity of the model for predicting the total number of theoretical trays. However, to estimate steady-state gains (g_{ij}) and RGA-values, *derivatives* of the model are needed. The objective of our paper is to check the models validity for estimating gains and the RGA.

We also feel that the derivation by Jafarey et al. does not provide much insight. They start from Smokers exact analytical solution, derive a simplified model which is not very good, delete some terms in this model, and finally arrive at a better model. Another goal of this paper is therefore to rederive Jafarey's model in a more direct manner which yields more insight. Our derivation yields a slightly different form of the model, but it reduces to Jafarey's expression if the assumption $D/F \approx z_F$ is made.

Assumptions. All the results in this paper are for a two-product column, binary mixture with constant relative volatility (α) and constant molar flows. The extension to multicomponent mixtures is discussed at the end of the paper.

2. CASE STUDY EXAMPLES

In this section we present steady-state data for seven high-purity distillation columns which may be used as case study examples in this and future work. The columns were selected to include a wide range of operating conditions. Data for the columns are given in Table 1, together with computed values for R/R_{min} , L/F and D/F .

Steady-state gains. The exact steady-state gain matrices for the seven columns were obtained by linearizing the material balance for each tray. For tray i with no feed

$$L(x_{i+1} - x_i) + V(y_{i-1} - y_i) = 0, \quad y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i} \quad (7)$$

Linearizing (7) for small perturbations

$$(x_{i+1} - x_i)dL + (y_{i-1} - y_i)dV + Ldx_{i+1} + (-L - K_i V)dx_i + K_{i-1}Vdx_{i-1} = 0 \quad (8)$$

where K_i is the linearized VLE-constant on each tray

$$K_i = \frac{dy_i}{dx_i} = \frac{\alpha}{(1 + (\alpha - 1)x_i)^2} \quad (9)$$

In matrix form (8) becomes

$$Ax + Bu + Ed = 0, \quad y = Cx \quad (10)$$

Here $x = (dx_1 \dots dx_N)^T$ are the tray compositions, $u = (dL \ dV)^T$ are the inputs, $d = (dF \ dz_F)^T$ are the disturbances, and $y = (dy_D \ dx_B)^T$ are the outputs. Solving these equations give

$$y = -CA^{-1}Bu - CA^{-1}Ed$$

and the steady-state gain matrix and disturbance matrix are

$$G = -CA^{-1}B \quad \text{and} \quad G_d = -CA^{-1}E \quad (11)$$

These matrices are given for the seven columns in Table 2. They are given both for the LV - and the DV -configurations:

Column	z_F	α	N	N_F	y_D	x_B	$\frac{D}{F}$	$\frac{L}{F}$	$\frac{N}{N_{min}}$	$\frac{R}{R_{min}}$
<i>A</i>	0.5	1.5	40	21	0.99	0.01	0.500	2.706	1.76	1.388
<i>B</i>	0.1	"	"	"	"	"	0.092	2.329	"	1.301
<i>C</i>	0.5	"	"	"	0.90	0.002	0.555	2.737	1.93	1.645
<i>D</i>	0.65	1.12	110	39	0.995	0.10	0.614	11.862	1.66	1.529
<i>E</i>	0.2	5	15	5	0.9999	0.05	0.158	0.226	1.99	1.144
<i>F</i>	0.5	15	10	5	0.9999	0.0001	0.500	0.227	1.47	3.183
<i>G</i>	0.5	1.5	80	40	0.9999	0.0001	0.500	2.635	1.76	1.318

Table 1. Steady-state data for distillation column examples. All columns have liquid feed ($q_F = 1$).

Column	G_{LV}	G_{DV}	$G_d (F, z_F)$
<i>A</i>	$\begin{pmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{pmatrix}$	$\begin{pmatrix} -0.878 & 0.014 \\ -1.082 & -0.014 \end{pmatrix}$	$\begin{pmatrix} 0.394 & 0.881 \\ 0.586 & 1.119 \end{pmatrix}$
<i>B</i>	$\begin{pmatrix} 1.748 & -1.717 \\ .9023 & -.9054 \end{pmatrix}$	$\begin{pmatrix} -1.748 & .03088 \\ -0.9023 & -.00312 \end{pmatrix}$	$\begin{pmatrix} 0.0858 & 1.636 \\ 0.0904 & 0.936 \end{pmatrix}$
<i>C</i>	$\begin{pmatrix} 1.604 & -1.602 \\ .01865 & -.02148 \end{pmatrix}$	$\begin{pmatrix} -1.604 & .00227 \\ -.01865 & -.00283 \end{pmatrix}$	$\begin{pmatrix} 0.8822 & 1.790 \\ 0.0197 & 0.0170 \end{pmatrix}$
<i>D</i>	$\begin{pmatrix} 0.1231 & -0.1211 \\ 2.126 & -2.129 \end{pmatrix}$	$\begin{pmatrix} -0.1231 & .00192 \\ -2.126 & -.00307 \end{pmatrix}$	$\begin{pmatrix} 0.0516 & 0.1252 \\ 1.344 & 2.395 \end{pmatrix}$
<i>E</i>	$\begin{pmatrix} .02033 & -.01315 \\ 1.124 & -1.126 \end{pmatrix}$	$\begin{pmatrix} -.02033 & .00719 \\ -1.124 & -.00135 \end{pmatrix}$	$\begin{pmatrix} .00045 & .00949 \\ 0.1780 & 1.186 \end{pmatrix}$
<i>F</i>	$\begin{pmatrix} 1.074 & -1.073 \\ 0.9257 & -0.9267 \end{pmatrix}$	$\begin{pmatrix} -1.074 & .000998 \\ -0.9257 & -.000998 \end{pmatrix}$	$\begin{pmatrix} 0.5362 & 1.073 \\ 0.4636 & 0.9269 \end{pmatrix}$
<i>G</i>	$\begin{pmatrix} 0.8649 & -0.8646 \\ 1.135 & -1.135 \end{pmatrix}$	$\begin{pmatrix} -0.8649 & .000294 \\ -1.135 & -.000294 \end{pmatrix}$	$\begin{pmatrix} 0.4315 & 0.8647 \\ 0.5683 & 1.135 \end{pmatrix}$

Table 2. Steady-state gain matrices for examples obtained using (11). $G_{DV} = G_{LV}M$ (13). Unscaled product compositions (y_D and x_B) are used as defined in (12).

$$\text{LV - configuration : } \begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G_{LV} \begin{pmatrix} dL \\ dV \end{pmatrix} + G_{dLV} \begin{pmatrix} dF \\ dz_F \end{pmatrix} \quad (12a)$$

$$\text{DV - configuration : } \begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G_{DV} \begin{pmatrix} dD \\ dV \end{pmatrix} + G_{dDV} \begin{pmatrix} dF \\ dz_F \end{pmatrix} \quad (12b)$$

Note that G_{LV} and G_{DV} are not independent. Because of the assumption of constant molar flows we have $dL = dV - dD$ which yields

$$\begin{pmatrix} dL \\ dV \end{pmatrix} = M \begin{pmatrix} dD \\ dV \end{pmatrix} \text{ where } M = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$$

and we derive

$$G_{DV} = G_{LV} M \quad (13)$$

Similar relationships are easily derived for other choices of manipulated inputs (Skogestad and Morari, 1987b). The matrix G_{LV} is very sensitive to small relative errors in the elements (as is seen from the large value of λ_{11} (Skogestad and Morari, 1987c)), while G_{DV} is insensitive to such errors. The reader should therefore find G_{LV} from $G_{LV} = G_{DV} M^{-1}$ rather than using the G_{LV} -matrix given in Table 2. The disturbance matrices G_{dLV} and G_{dDV} are in general different, but happen to be equal in this case since the feed is liquid.

3. A SIMPLE EXPRESSION FOR S

Previous Work

For mixtures with constant relative volatility (α) the following exact expressions for N_{min} and R_{min} hold (e.g., Henley and Seader, 1981).

$$R \rightarrow \infty : S = \alpha^{N_{min}} \quad (\text{Fenske}) \quad (14)$$

$$N \rightarrow \infty : R_{min} = \frac{1}{\alpha - 1} \left(\frac{y_D}{x_F} - \alpha \frac{1 - y_D}{1 - x_F} \right) \approx \frac{1}{\alpha - 1} \frac{1}{x_F} (y_D \approx 1) \quad (15)$$

x_F is the composition of the flashed feed. Simple extensions to multicomponent mixtures exist (e.g., Henley and Seader). Jafarey, Douglas and McAvoy (1979)

derived a simple expression for S which has also been adopted by Shinskey (1984). For liquid feed we have $x_F = z_F$ and Jafarey's expression becomes

$$q_F = 1 : \quad S = \alpha^N \left(\frac{1}{1 + 1/(Rz_F)} \right)^{N/2} \quad (16)$$

The expression (16) gives the correct limiting value (14) as $R \rightarrow \infty$, but its value of R_{min}

$$N \rightarrow \infty : \quad R_{min} = \frac{1}{(\alpha^2 - 1)z_F} \quad (17)$$

is lower than the correct one (15). Jafarey et al. used (16) for design purposes to find the number of theoretical stages N needed to accomplish a given separation. For 24 cases studied the average error in N was 5.3% compared to 2.5% for the Gilliland correlation.

Derivation of Expression for S

The main reason for including this section is that the derivation given here is much simpler and more instructive than the one given by Jafarey et al. Furthermore, extension to multicomponent mixtures is straightforward. We will initially linearize the vapor-liquid equilibrium (VLE) curve in the top and bottom parts of the column.

$$\text{Above feed :} \quad (1 - y_i) = \frac{1}{\alpha_T} (1 - x_i) \quad (18a)$$

$$\text{Below feed :} \quad y_i = \alpha_B x_i \quad (18b)$$

Here α_T and α_B denote the relative volatility in the top and bottom part of the column. The assumption of linear VLE will clearly make the separation in the middle of the column simpler as seen from Fig. 1. Note that for high-purity columns, most of the stages will be at the column ends in the regions of high purity where the linear approximation is good. The operating lines (material balance of light component for each tray) are

$$\text{Above feed :} \quad V_T y_{i-1} = L_T x_i + D y_D \quad (19a)$$

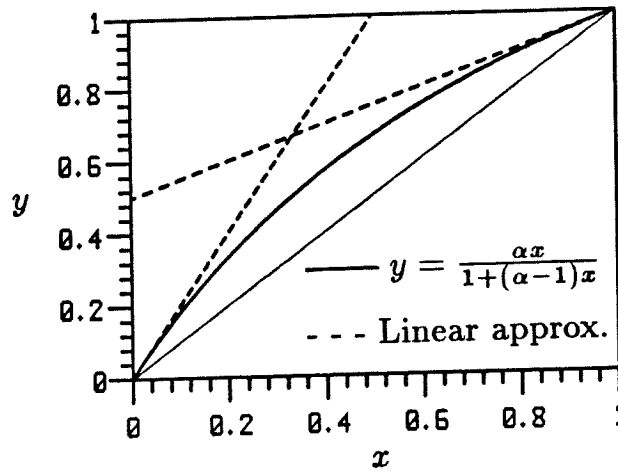


Fig. 1. VLE - curve for relative volatility $\alpha = 2.0$ and linear approximation with $y = \alpha x$ and $1 - y = \frac{1-x}{\alpha}$.

$$\text{Below feed : } L_B x_{i+1} = B_B y_i + B x_B \quad (19b)$$

Since the equilibrium and operating lines are assumed linear it is simple to derive exact relationships between the feed tray composition and y_D and x_B . Consider the top part of the column. The vapor composition of the heavy component ($1 - y_j$) on the j th tray from the top is found by repeated use of (18a) and (19a) (Kremser equations, see McCabe and Smith, 1976).

$$(1 - y_j) = (1 - y_D) \left(A_T^j + \frac{(A_T^j - 1)(1 - L_T/V_T)}{A_T - 1} \right) \quad (20)$$

where A_T is the absorption factor

$$A_T = \frac{L_T/V_T}{1/\alpha_T} \quad (21)$$

Since $A_T > 1$ we have $A_T^j \gg 1$ for j large. Using this assumption for the feed tray we derive

$$(1 - y_{NF}) \approx (1 - y_D) A_T^{N_T} \left(\frac{A_T - L_T/V_T}{A_T - 1} \right) \quad (22)$$

where N_T is the number of theoretical stages above the feed. Similarly, we derive for the bottom section of the column an expression for the liquid composition on the feed tray

$$x_{NF} = x_B \left(A_B^{-N_B} + \frac{(A_B^{-N_B} - 1)(1 - V_B/L_B)}{A_B^{-1} - 1} \right) \approx x_B A_B^{-N_B} \left(\frac{A_B^{-1} - V_B/L_B}{A_B^{-1} - 1} \right) \quad (23)$$

Here N_B is the number of theoretical trays below the feed and $A_B < 1$ is the absorption factor in the bottom part of the column

$$A_B = \frac{L_B/V_B}{\alpha_B} \quad (24)$$

Multiplying equations (22) and (23) gives an approximate expression for the separation factor S :

$$S \approx \frac{1}{x_B(1 - y_D)} = \frac{\alpha_B^{N_B} \alpha_T^{N_T}}{(1 - y_{NF})x_{NF}} \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \cdot c \quad (25)$$

$$c = \frac{(\alpha_T - 1)(\alpha_B - 1)}{(\alpha_T - V_T/L_T)(\alpha_B - L_B/V_B)}$$

In most cases c is close to one; in particular, this is the case if the reflux is high. Assuming constant relative volatility $\alpha = \alpha_B = \alpha_T$ and using $\alpha = \frac{y_{NF}/1 - y_{NF}}{x_{NF}/1 - x_{NF}}$ and $N = N_B + N_T + 1$ we derive from (25)

$$S = \frac{\alpha^N}{y_{NF}(1 - x_{NF})} \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \cdot c \quad (26)$$

We know that S predicted by (26) is too large because of the linearized VLE. However, (26) may be corrected to satisfy the exact relationship $S = \alpha^N$ at infinite reflux simply by dropping $1/y_{NF}(1 - x_{NF})$ from (26) . By assuming in addition $c = 1$ we get

$$S = \alpha^N \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \quad (27)$$

This expression is somewhat misleading since it suggests that the separation may always be improved by transferring stages from the bottom to the top section if

$(L/V)_T > (V/L)_B$. This is clearly not generally true, and to avoid this problem we follow Jafarey et al. (1979) and choose $N_T \approx N_B \approx N/2$ and derive the alternative expression

$$S = \alpha^N \left(\frac{(L/V)_T}{(L/V)_B} \right)^{N/2} \quad (28)$$

If the feed is liquid then $V_B = V_T$ and

$$q_F = 1 : \quad S = \alpha^N \left(\frac{L_T}{L_B} \right)^{N/2} = \alpha^N \left(\frac{L}{L + F} \right)^{N/2} \quad (28a)$$

Since $L/F = \frac{L}{D} \frac{D}{F}$, this reduces to Jafarey's expression (16) if the additional assumption $\frac{D}{F} \approx z_F$ is made. Jafarey et al. derived (16) as a design equation and the assumption $\frac{D}{F} \approx z_F$ is reasonable in this case. However, the assumption is not so easy to justify for control purposes where D is a variable which may take on values different from $D = z_F F$. The assumption seems particularly misleading if the model is used to compute gains for disturbances in z_F .

Column	Actual N	Estimated N	
		This work (28)	Jafarey (16)
A	40	37.02	37.02
B	40	40.51	38.41
C	40	33.67	35.73
D	110	102.81	99.89
E	15	15.92	14.17
F	10	9.88	9.88
G	80	75.32	75.32
Average Error		6.3%	6.3%

Table 3. Estimated number of theoretical trays in columns.

The main shortcoming of (28) is obvious from the derivation: The model is poor if a pinch zone appears around the feed plate such that the assumption that most of the trays are in the region of high or low purity which led to (26) is no longer valid. A pinch zone occurs around the feed plate if the reflux is near minimum; this explains why (16) gives a value of R_{min} (17) which is too low. More importantly, pinch zones appear above or below the feed plate if the feed location is not optimal. This is likely to happen during operation of the column. The use of (28) implicitly assumes that the feed location is optimal. Finally, for columns with large differences between N_T and N_B , (27) should be used instead of (28). We will study the validity of the models (16) and (28) for estimating steady-state gains in Section 5.

4. WHY ARE DISTILLATION COLUMNS NONLINEAR?

For the simple case of constant molar flows and constant relative volatility considered in this paper, there are two possible sources of nonlinearity

(A) Nonlinear VLE

$$y_i = \frac{\alpha x_i}{1 + (\alpha + 1)x_i} \quad (29)$$

(B) Bilinear terms (Lx_i, Vy_i , etc). in the material balance for each tray

$$Lx_{i+1} + Vy_{i-1} = Lx_i + Vy_i \quad (30)$$

None of these nonlinearities seem very strong, and the strongly nonlinear behavior observed for (high-purity) distillation columns is therefore somewhat surprising. The main reason for the nonlinear behavior is, as we will show, the nonlinear VLE.

Consider changes in product compositions (y_D and x_B) caused by a change in feed composition (z_F), when all flows, including reflux L and boilup V , are constant. In this case the material balance (30) is linear, and the only possible source of nonlinearity is the VLE.

Linear VLE. If the VLE were assumed to be linear (e.g., by linearizing (29) on each tray; $y_i = a_i x_i + b_i$), all equations describing the column behavior are linear,

and a linear relationship between the product compositions and z_F results

$$y_D = k_1 z_F, \quad x_B = k_2 z_F$$

Here the gains $(\partial y_D / \partial z_F)_{L,V} = k_1$ and $(\partial y_D / \partial z_F)_{L,V} = k_2$ are constant and independent of the feed composition.

Nonlinear VLE. A completely different result is found, however, when the VLE is not linearized. The simplified model (27) for S was derived using the nonlinear VLE model (29). (Actually, the VLE was first linearized, but the final expression was corrected to match $S = \alpha^N$ at total reflux.) This model predicts that S is constant for changes in feed composition provided L and V are unchanged. Consider a column which nominally has $z_F^o = 0.5$ and $1 - y_D^o = x_B^o = 0.01$. For S constant the following two equations give y_D and x_B as a function of z_F

$$\frac{y_D(1 - x_B)}{(1 - y_D)x_B} = \frac{y_D^o(1 - x_B^o)}{(1 - y_D^o)x_B^o} = 9801 \quad (31)$$

$$z_F = \frac{D}{F} y_D + \frac{B}{F} x_B \quad (32)$$

(The second equation is the overall material balance for light component.) Note that all flows are constant, so $D/F = B/F = 0.5$ is constant. (31) and (32) clearly result in a nonlinear relationship between y_D and x_B , and z_F . The steady-state gains found by combining these equations are strongly dependent on the operating point. For example,

$$S \text{ constant : } \left(\frac{dy_D}{dz_F} \right)_{L,V} = \frac{1}{\frac{D}{F} + \frac{B}{F} \frac{x_B(1 - x_B)}{y_D(1 - y_D)}} \quad (33)$$

At the nominal operating point $(\partial y_D / \partial z_F)_{L,V} = 1$. However, for $z_F = 0.55$, (31) and (32) give $x_B = 0.10$ and $1 - y_D = 0.0092$, and the linearized gain at this operating point is $(\partial y_D / \partial z_F)_{L,V} = 0.0202$. This is 50 times smaller than the nominal value.

The conclusion is that the nonlinear VLE results in a strongly nonlinear behavior for high-purity columns. The bilinear terms in (30) are of much less importance for the observed nonlinearity of distillation columns.

5. STEADY-STATE GAINS FROM EXPRESSION FOR S

Consider any input or disturbance ζ (for example $\zeta = L, V, B, F, z_F$, etc.). The steady-state gains $\partial y_D / \partial \zeta$ and $\partial x_B / \partial \zeta$ may be found by differentiating the material balance (2) and using the defining expression for the separation factor (3)

$$D \frac{\partial y_D}{\partial \zeta} + B \frac{\partial x_B}{\partial \zeta} = e_\zeta \quad (34)$$

$$\frac{\partial \ln S}{\partial \zeta} = \frac{1}{y_D(1-y_D)} \frac{\partial y_D}{\partial \zeta} - \frac{1}{x_B(1-x_B)} \frac{\partial x_B}{\partial \zeta} \quad (35)$$

Here e_ζ is defined as

$$e_\zeta = -(y_D - x_B) \frac{\partial D}{\partial \zeta} + F \frac{\partial z_F}{\partial \zeta} + (z_F - x_B) \frac{\partial F}{\partial \zeta} \quad (36)$$

Solving for the gains give

$$\frac{1}{(1-y_D)y_D} \frac{\partial y_D}{\partial \zeta} = \frac{1}{I_s} \left(e_\zeta + Bx_B(1-x_B) \frac{\partial \ln S}{\partial \zeta} \right) \quad (37a)$$

$$\frac{1}{(1-x_B)x_B} \frac{\partial x_B}{\partial \zeta} = \frac{1}{I_s} \left(e_\zeta - Dy_D(1-y_D) \frac{\partial \ln S}{\partial \zeta} \right) \quad (37b)$$

where I_s is the "sum" of impurities leaving the column

$$I_s = Bx_B(1-x_B) + Dy_D(1-y_D) \quad (38)$$

For high-purity separations $I_s \approx Bx_B + D(1-y_D)$. There are two contributions to the gains in (37): The e_ζ -term and the contribution from changes in the separation factor S . The e_ζ -term physically represents the effect on the gains of changing the external material balance. This is clear since $\partial D / \partial \zeta = 0$ and $e_\zeta = 0$ is obtained when we change the internal flows in the column only (change L and V keeping D and B constant).

For changes in the external flows (which represent most disturbances and inputs) the e_ζ -term in (37) is usually dominating - at least when both products are high-purity. This means that the major contribution to the gains in this case may be obtained by assuming S constant and the exact value of $\partial \ln S / \partial \zeta$ is of minor importance in this case.

On the other hand, a good model for S is important for obtaining the gains for changes in the internal flows : In this case the gains, for example $(\partial y_D / \partial L)_D$, are directly proportional to $(\partial \ln S / \partial L)_D = (\partial \ln S / \partial V)_D$ and a good estimate of this quantity is required for estimating the correct value of the gain.

Estimating $\partial \ln S / \partial \zeta$ from shortcut models

Equations (37) are exact. I_s is a given constant and e_ζ is trivial to find for the case of constant molar flows. The only "unknown" in (37) is $\partial \ln S / \partial \zeta$, which may be estimated from the shortcut models (28) or (16). Analytic expressions for $\partial \ln S / \partial \zeta$ obtained with these models for different choices of ζ are given in Table 4. Estimated numerical values of $\partial \ln S / \partial \zeta$, and of the gains and the RGA for the LV-configuration are compared with exact values for the seven columns in Table 5-7 and Fig.2. The results are discussed below for each shortcut model.

Using the model (28) for $\ln S$. From Table 5 we see that, with the exception of $(\partial \ln S / \partial V)_D$, the estimates of $\partial \ln S / \partial \zeta$ are extremely poor. Note that because of the assumption of constant molar flows we have

$$\left(\frac{\partial \ln S}{\partial V} \right)_D = \left(\frac{\partial \ln S}{\partial V} \right)_L + \left(\frac{\partial \ln S}{\partial L} \right)_V \quad (39)$$

and even though the two terms $(\partial \ln S / \partial V)_L$ and $(\partial \ln S / \partial L)_V$ individually are estimated very poorly with (28), the estimate of the sum $(\partial \ln S / \partial V)_D$ is in reasonable agreement (Fig.2). Also note from Fig. 2 that $(\partial \ln S / \partial V)_L$ and $(\partial \ln S / \partial L)_V$ are very sensitive to the feed point location, while their sum $(\partial \ln S / \partial V)_D$ is nearly constant. The conclusion is that the model (28) is only useful for estimating the

	This work (28)	Jafarey (16) (Feed liquid, $q_F = 1$)
$\ln S$	$N \ln \alpha + \frac{N}{2} \frac{(L/V)_T}{(L/V)_B}$	$N \ln \alpha - \frac{N}{2} \ln \left(1 + \frac{1}{Rz_F} \right)$
$\left(\frac{\partial \ln S}{\partial L} \right)_V$	$\frac{N}{2} \frac{1}{L} \frac{q_F F}{L + q_F F}$	$\frac{N}{2} \frac{1}{L} \frac{R+1}{Rz_F+1}$
$\left(\frac{\partial \ln S}{\partial V} \right)_L$	$\frac{N}{2} \frac{1}{V} \frac{(1-q_F)F}{L+D}$	$-\frac{N}{2} \frac{1}{D} \frac{1}{Rz_F+1}$
$\left(\frac{\partial \ln S}{\partial V} \right)_D$	$(q_F = 1) : \frac{N}{2} \frac{1}{L} \frac{1}{L/F+1}$	$\frac{N}{2} \frac{1}{L} \frac{1}{Rz_F+1}$
$\left(\frac{\partial \ln S}{\partial z_F} \right)_{L,V}$	0	$\frac{N}{2} \frac{1}{z_F} \frac{1}{Rz_F+1}$
$\left(\frac{\partial \ln S}{\partial F} \right)_{L,V}$	$-\frac{N}{2} \left(\frac{q_F}{L+q_F F} + \frac{1-q_F}{L+D} \right)$	0
$\left(\frac{\partial \ln S}{\partial q_F} \right)_{L,V}$	$-\frac{N}{2} \left(\frac{1}{L/F+q_F} - \frac{1}{L/F+D/F} \right)$	$-\frac{N}{2} \left(\frac{1}{Rz_F+1} - \frac{1}{Rz_F+z_F} \right)$

Table 4. Analytical expressions for $\partial \ln S / \partial \zeta$ obtained from shortcut models (28) and (16). $(\partial \ln S / \partial V)_D$ for $q_F \neq 1$ is equal to $(\partial \ln S / \partial L)_V + (\partial \ln S / \partial V)_L$ (39).

	Column A			Column B		
	Exact	This work (28)	Jafarey (16)	Exact	This work (28)	Jafarey (16)
$(\partial \ln S / \partial L)_V$	-21.1	1.99	12.8	85.3	2.58	64.0
$(\partial \ln S / \partial V)_L$	23.9	0	-10.8	-81.9	0	-61.6
$(\partial \ln S / \partial V)_D$	2.76	1.99	1.99	3.44	2.58	2.42
$(\partial \ln S / \partial z_F)_{L,V}$	-24.0	0	10.8	70.7	0	56.5
$(\partial \ln S / \partial F)_{L,V}$	-19.3	-5.4	0	-0.5	-6.0	0
$(\partial \ln S / \partial q_F)_{L,V}$	-22.7	0.8	0.8	84.8	2.3	1.9

Table 5. Estimated values for $\partial \ln S / \partial \zeta$ for columns A and B.

Column	Exact	Estimates	
		This work (28)	Jafarey (16)
A	2.76	1.99	1.99
B	3.44	2.58	2.43
C	1.44	1.96	2.11
D	0.42	0.36	0.34
E	71.9	27.1	25.8
F	20.0	18.0	18.0
G	5.87	4.18	4.18
Average error		29.2%	32.2%

Table 6. Estimates of $(\partial \ln S / \partial V)_D = (\partial \ln S / \partial L)_D$. This represents the effect of changing the internal flows on S .

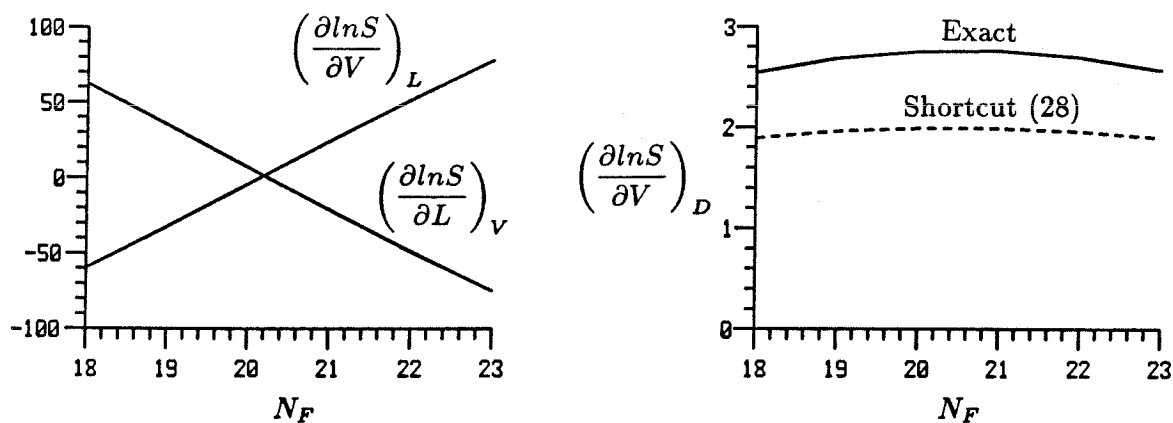


Fig. 2. Column A. Effect of feed point location N_F on derivatives of $\ln S$ with all other column data fixed (including $1 - y_D = x_B = 0.01$). Exact values are compared with shortcut model (28). Note that $(\partial \ln S / \partial L)_V + (\partial \ln S / \partial V)_L = (\partial \ln S / \partial V)_D$ and that $(\partial \ln S / \partial V)_L = 0$ for the shortcut model since $q_F = 1$ (see Table 4.).

Column	Exact	Estimated	
		This work (28)	Jafarey (16)
<i>A</i>	$\begin{pmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{pmatrix}$	$\begin{pmatrix} 0.990 & -0.980 \\ 0.970 & -0.980 \end{pmatrix}$	$\begin{pmatrix} 1.0434 & -1.0335 \\ 0.9166 & -0.9265 \end{pmatrix}$
<i>B</i>	$\begin{pmatrix} 1.748 & -1.717 \\ .9023 & -.9054 \end{pmatrix}$	$\begin{pmatrix} 1.0032 & -0.9800 \\ 0.9777 & -0.9800 \end{pmatrix}$	$\begin{pmatrix} 1.5554 & -1.5337 \\ 0.9218 & -0.9240 \end{pmatrix}$
<i>C</i>	$\begin{pmatrix} 1.604 & -1.602 \\ .01865 & -.02148 \end{pmatrix}$	$\begin{pmatrix} 1.5940 & -1.5909 \\ .03145 & -.03528 \end{pmatrix}$	$\begin{pmatrix} 1.6094 & -1.6061 \\ .01074 & -.01488 \end{pmatrix}$

Table 7. Estimates of G_{LV} for columns A, B and C obtained using (37) with shortcut models (28) and (16).

Column	λ_{11} , Exact	λ_{11} , Estimated	
		This work (28)	Jafarey (16)
A	35.1	50.3	50.0
B	47.5	39.3	61.0
C	7.53	9.06	3.58
D	58.7	66.7	85.9
E	2.82	1.87	3.35
F	499	558	558
G	1673	2394	2394
Average error		26.2%	36.5%

Table 8. Estimated and exact RGA-values for the LV-configuration. Note that the values of Jafarey are used in the book of Shinskey (1984).

effect of changing the internal flows on S

$$\left(\frac{\partial \ln S}{\partial V}\right)_D = \left(\frac{\partial \ln S}{\partial L}\right)_D \quad (40)$$

Fortunately, as pointed out above, this corresponds exactly to the case when the term $\partial \ln S / \partial \zeta$ in (37) is most important. Therefore, in spite of the generally poor estimates of $\partial \ln S / \partial \zeta$, the model (28) can be useful for estimating steady-state gains and the RGA (recall Eq. (5)). This will in particular be the case for columns with both products of high purity. In other cases the errors may be significant. This is evident by comparing the estimated steady-state gain matrices in Table 7 with the exact values in Table 2.

Using Jafarey's model (16) for $\ln S$. As seen from Table 5 and 6, the estimated values for $(\partial \ln S / \partial V)_D$ are very similar to those found using (28). The estimated values for $\partial \ln S / \partial \zeta$ with respect to changes in the external material balance are again poor. Surprisingly, the values are also very different from what is obtained using (28). The estimated steady-state gain matrices are also very different for the two models in some cases as seen from Table 7. Shinskey (1984) uses the model (16) to estimate the RGA. This estimate can be quite poor as seen from Table 8. The model (28) seems to be somewhat better than (16) for estimating the RGA.

Choice of Model for $\ln S$

The two seemingly very similar shortcut models (16) and (28) for $\ln S$ may result in quite different estimates for the steady-state gain matrices (Table 7). None of the models give very accurate results, but they may still be useful for obtaining a first estimate. Based on the numerical results presented above, there is no reason to give preference to one of the models (16) or (28). However, the assumption $D \approx z_F F$ which led to (16) does not seem to be justified from a theoretical point of view, and we therefore recommend using (28), which also gives simpler analytical expressions for $\partial \ln S / \partial \zeta$ (Table 4).

6. A NEW FORMULA FOR THE OPTIMAL FEED LOCATION

Equations (22) and (23) are reasonably accurate if most of the trays are located in the region of high and low purity. Divide equation (22) by (23) and assume the feed plate is optimally located such that $x_{N_F} = x_F$ and $y_{N_F} = y_F$

$$\frac{1 - y_F}{x_F} = \frac{1 - y_D}{x_B} \alpha^{N_T - N_B} \frac{(L/V)_T^{N_T}}{(V/L)_B^{N_B}} \frac{\alpha - (L/V)_B}{\alpha - (V/L)_T} \quad (41)$$

As a crude approximation neglect the last two terms. This approximation is reasonable if N_T and N_B are not too different since $(L/V)_T$ and $(V/L)_B$ are often reasonably close in magnitude. (41) then gives

$$N_T - N_B = \frac{\ln \frac{1 - y_F}{x_F} \frac{x_B}{1 - y_D}}{\ln \alpha} \quad (42)$$

x_F and y_F are functions of z_F, q_F and α and are obtained by flashing the feed. The optimal feed tray location is then given by

$$N_F = N_B + 1 = \frac{N + 1 - (N_T - N_B)}{2} \quad (43)$$

(N is the total number of theoretical stages.) Estimated and exact values of the optimal feed stage locations are shown for the seven columns in Table 9. The exact value is found using Stoppel (1946) (for the feed as liquid $N_F = n_2 + 1$, where n_2 is found from Stoppel's paper). The average error $|\Delta N_F|$ in percent of $N/2$ for the seven columns using (42) is 7.0%. This compares to 12.2% when using the Fenske ratio (Henley and Seader, 1981)

$$\frac{N_T}{N_B} = \frac{\ln \left(\frac{y_D}{1 - y_D} \frac{1 - z_F}{z_F} \right)}{\ln \left(\frac{1 - x_B}{x_B} \frac{z_F}{1 - z_F} \right)} \quad (44)$$

and 18.0% when using the empirical Kirkbride formula (Henley and Seader, 1981)

$$\frac{N_T}{N_B} = \left[\frac{1 - z_F}{z_F} \frac{B}{D} \left(\frac{x_B}{1 - y_D} \right)^2 \right]^{0.206} \quad (45)$$

Column	N	N_F , Exact Stoppel	N_F , Estimated		
			This work (42)	Fenske ratio (44)	Kirkbride (45)
A	40	21.0	20.8	20.5	20.5
B	40	18.5	17.9	11.2	12.1
C	40	30.1	25.6	29.8	33.8
D	110	39.1	45.3	42.0	30.1
E	15	5.16	5.82	2.79	1.55
F	10	5.84	5.88	5.50	5.50
G	80	40.0	40.8	40.5	40.5
Average error, $\frac{ \Delta N_F }{N/2}$:			7.0%	12.2%	18.0%

Table 9. Optimal feed point locations.

(The crude estimate $N_F = N/2$ gives an average error of 20.0%). We also computed the optimal feed point location for the 24 columns given by Jafarey et al. (1979). (There is a misprint in this paper, and R/R_{min} should be 1.75 for case b). The following average errors were found

	$\frac{ \Delta N_F }{N/2}$
This work (42)	4.4%
Fenske ratio (44)	15.7%
Kirkbride (45)	12.8%

7. EXTENSIONS TO MULTICOMPONENT MIXTURES

Define separation in terms of two key components (L and H), and define the pseudo-binary values of composition and flows (superscript ') by considering only these key components. For example

$$y' = y'_L = \frac{y_L}{y_L + y_H}, \quad V' = V(y_L + y_H) \quad (46)$$

(as usual the subscript L has been dropped for the light component). The nonkey components are assumed to be non-distributing; the heavy nonkey (HN) is assumed to have much lower relative volatility than the heavy key (H). Also, the light nonkey (LN) is assumed to have much higher relative volatility than the light key (L). Under these assumptions (27) still holds if a pseudo-binary basis is used (Appendix):

$$S = \frac{y'_D/1 - y'_D}{x'_B/1 - x'_B} = \alpha^N \frac{(L'/V')_T^{N_T}}{(L'/V')_B^{N_B}} \quad (47)$$

where as before

$$\alpha = \frac{y_L/y_H}{x_L/x_H} = \frac{y'/1 - y'}{x'/1 - x'} \quad (48)$$

The material balance may also be written on a pseudo-binary basis

$$F'z'_F = B'x'_B + D'y'_D \quad (49)$$

Consequently, all the gain equations (34)-(38) derived above still apply when a pseudo-binary basis is used.

Similarly, Eq. (42) for the optimal feed point location applies if x_B, y_D, y_F and x_F are replaced by the pseudo-binary values x'_B, y'_D, y'_F and x'_F .

8. CONCLUSION

It is convenient to express the column behavior in terms of the separation factor

$$S = \frac{y_D(1 - x_B)}{(1 - y_D)x_B} \quad (3)$$

S is often nearly constant for varying operating conditions. For columns with constant relative volatility and constant molar flows, Fenske's exact equation applies at total reflux

$$S = \alpha^N \quad (4)$$

For finite reflux the following generalization is useful

$$S = \alpha^N \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \quad (27)$$

This model gives a good description of how S changes with the internal flows, but describes poorly the effect of changes in the external material balance. Fortunately, it turns out that this is of less importance if the model for S is used to obtain estimates for the steady state gains.

Finally, a new formula for estimating the optimal feed stage location is proposed

$$N_T - N_B = \frac{\ln \frac{1-y_F}{x_F} \frac{x_B}{1-y_D}}{\ln \alpha} \quad (42)$$

From its derivation we know that this formula may give poor results if N_T and N_B are very different, but it gave better results than other proposed methods for the seven columns in Table 1 and for the 24 columns studied by Jafarey et al. (1979).

NOMENCLATURE

B - bottom product rate

D - distillate (top product) rate

F - feed rate

$L_T = L$ - liquid flow in top section of column

$L_B = L + q_F F$ - liquid flow in bottom section of column

N_B - number of theoretical trays below feed (incl. reboiler)

$N_F = N_B + 1$ - feed tray location

N_T - number of theoretical trays above feed

$N = N_B + N_T + 1$ - total number of theoretical trays in column

q_F - fraction of liquid in feed

$R = L/D$ - reflux ratio

$S = \frac{y_D(1-x_B)}{(1-y_D)x_B}$ - separation factor

$V_B = V$ - vapor flow in bottom section of column

$V_T = V + (1 - q_F)F$ - vapor flow in top part of column

x_B - mole fraction of light component in bottom product

x_i - liquid mole fraction of light component on stage i

x_F, y_F - mole fraction in feed at feed stage pressure

x_{N_F}, y_{N_F} - mole fractions on feed tray

$y_D = x_D$ - mole fraction of light component in distillate (top product)

y_i - vapor mole fraction of light component on stage i

Subscripts

B - bottom part of column, bottom product

D - distillate product

T - top part of column

i - tray no. numbered from bottom ($i=1$ for reboiler, $i=2$ for first tray, $i=N$ for top tray, $i=N+1$ for condenser)

Superscript

' - pseudo-binary basis is used

Acknowledgements. Partial support from the National Science Foundation and Norsk Hydro is gratefully acknowledged.

REFERENCES

- Gilliland, E. R., *Ind. and Eng. Chem.*, **32**, 1220 (1940).
- Henley, E. J. and J. D. Seader, *Equilibrium-State Separation Operations in Chemical Engineering*, John Wiley and Sons (1981).
- Jafarey, A., J. M. Douglas and T. J. McAvoy, Short-Cut Techniques for Distillation Column Design and Control. 1. Column Design, *Ind. & Eng. Chemistry Process Des. & Dev.*, **18**, 197-202 (1979).
- McCabe, W. L. and J. C. Smith, *Unit Operations of Chemical Engineering*, 3rd Edition, McGraw-Hill (1976).
- Shinskey, F. G., *Distillation Control*, 2nd Edition, McGraw-Hill (1984).
- Skogestad, S. and M. Morari, Shortcut Models for Distillation Columns - II. Dynamic Composition Response, This issue of *Comp. & Chem. Eng.* (1987a)
- Skogestad, S. and M. Morari, Understanding the Steady State Behavior of Distillation Columns, in preparation (1987b).
- Skogestad, S. and M. Morari, Implications of Large RGA-elements on Control Performance, submitted to *Ind. & Eng. Chemistry Process Des. & Dev.* (1987c)
- Stoppel, A. E., Calculation of Number of Theoretical Plates for Rectifying Column, *Ind. & Eng. Chemistry*, **38**, 12, 1271-1272 (1946).

APPENDIX

Extensions to multicomponent mixtures

The relative volatility between the key components is defined as (subscript i for the tray is dropped)

$$\alpha = \frac{y_L/x_L}{y_H/x_H} = \frac{y'(1-x')}{x'(1-y')} \quad (48)$$

Bottom section. From (48) we get

$$y_L = \left(\alpha_B \frac{y_H}{x_H}\right) x_L = \hat{\alpha}_B x_L \quad (A1)$$

In Section 3 we linearized the VLE by assuming $y_H \approx 1$ and $x_H \approx 1$, but the last approximation does not hold in presence of heavy nonkey-components (HN). Assuming $y_{HN} \approx 0$ (i.e., the heavy nonkey has very low relative volatility) we get

$$x_{HN} \approx \frac{z_{HN}F}{L_B}$$

and we find $y_H \approx 1$ and $x_H \approx 1 - x_{HN} = 1 - \frac{z_{HN}F}{L_B}$, i.e.,

$$\hat{\alpha}_B = \alpha_B \left(1 - \frac{x_{HN}F}{L_B}\right) \quad (A2)$$

and we have a linear VLE relationship for the bottom section. This leads to an expression for x_{LN_F} similar to (23) but with α_B replaced by $\hat{\alpha}_B$.

Top section. From (48) we get

$$y_H = \left(\frac{1}{\alpha_T} \frac{y_L}{x_L}\right) x_H = \frac{1}{\hat{\alpha}_T} x_H \quad (A3)$$

A derivation similar to the one for the bottom section yields

$$\hat{\alpha}_T = \frac{\alpha_T}{1 - z_{LN}F/V_T} \quad (A4)$$

This leads to an expression for y_{HN_F} as a function of y_{HD} similar to (22), but with α_T replaced by $\hat{\alpha}_T$.

Note that Fenske's expression for S at infinite influx is exact also when there are nonkey components.

$$S = \frac{y_{LD}/x_{LD}}{y_{HD}/x_{HD}} = \frac{y'_D(1-x'_B)}{x'_B(1-y'_D)} = \alpha^N \quad (A5)$$

Multiplying (22) with (24) as before, and correcting with the exact expression (A5) we derive

$$S = \alpha^N \left(1 - \frac{z_{HN}F}{L_B}\right)^{N_B} \left(1 - \frac{z_{LN}F}{V_T}\right)^{-N_T} \frac{(L/V)_T^{N_T}}{(L/V)_B^{N_B}} \quad (A6)$$

Since the nonkey components are non-distributing we have

$$L'_T = L_T, \quad V'_T = V_T - z_{LN}F$$

$$L'_B = L_B - x_{HN}F, \quad V'_B = V_B$$

Substituting this into (A6) we finally derive

$$S = \alpha^N \frac{(L'/V')_T^{N_T}}{(L'/V')_B^{N_B}} \quad (47)$$

Chapter X

SHORTCUT MODELS FOR DISTILLATION COLUMNS -

II. DYNAMIC COMPOSITION RESPONSE

SHORTCUT MODELS FOR DISTILLATION COLUMNS -

II. DYNAMIC COMPOSITION RESPONSE

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January 1987

Submitted to Computers and Chemical Engineering

Abstract

It is shown that the dynamic composition response can be approximated by a linear first-order response. This applies also for large perturbations to the column. The numerical value of this dominant time constant (τ_c) can be derived from steady-state simulations. A simple analytical expression for small perturbations is derived which provides insight into the variation of τ_c with operating conditions. The time constant τ_c does not apply when there are changes in the internal flows only.

1. INTRODUCTION

The dynamics of most distillation columns are dominated by one large time constant, which is nearly the same, regardless of where a disturbance is introduced or where composition is measured. This is well known both from plants measurements (McNeill and Sachs, 1969) and from theoretical studies (Moczek et al., 1963, Wahl and Harriot, 1970, Kim and Friedly, 1974).

A distillation column is described by a large number of differential equations. On each tray a differential equation may be formulated for , (i) the material balance of each component (composition dynamics)

$$\frac{d}{dt}(M_i x_i) = L_{i+1} x_{i+1} + V_{i-1} y_{i-1} - L_i x_i - V_i y_i$$

$$\text{where from the VLE : } y_i = K_i(x_i, T)$$

(ii) the overall material balance (flow dynamics)

$$\frac{d}{dt} M_i = L_{i+1} + V_{i-1} - L_i - V_i$$

$$\text{where (tray hydraulics) } L_i = f_1(V_i, M_i, \Delta p_i)$$

$$\text{(pressure drop) } V_i = f_2(M_i, p_i)$$

and (iii) the enthalpy balance

$$\frac{d}{dt}(M_i H_i) = L_{i+1} H_{i+1}^L + V_{i-1} H_{i-1}^V - L_i H_i^L - V_i H_i^V$$

$$\text{where } H_i^L = f_3(x_i, p), \quad H_i^V = f_4(y_i, p)$$

We only want to outline the structure of the equations and the reader should not be concerned about the details. Viewed against the background of the large number of nonlinear differential equations, the simple low-order responses (often first order) observed for most distillation columns are somewhat surprising. Levy et al. (1969) showed through modal analysis that for a specific example the slowest mode involved

primarily the composition effects (i). This leads to the conjecture that the essential part of the dynamics can be captured by modelling only the composition dynamics. This is the approach taken in this paper.

The simplest approach is to consider the total holdup of each component in the column. By assuming that all trays have the same response, this directly leads to a first order model, and the dominant time constant can be estimated. According to Rademaker et al. (1975, p.280) this idea dates back to the beginning of the century (Lord Raleigh) and seems to get rediscovered every few years. Moczek et al. (1963) used it to introduce the "inventory time constant" for a column going from one steady state (subscript 0) to another (subscript f):

$$T_{inv} = \frac{\Delta(\sum M_i x_i)}{\Delta(Fz_F) - y_{D0}\Delta D - x_{B0}\Delta B} \quad (1)$$

Here $\Delta D = D_f - D_0$ and $\Delta B = B_f - B_0$ are the changes in distillate and bottoms flow rate, $\Delta(\sum M_i x_i)$ is the change in holdup in the column of any component, $\Delta(Fz_F)$ is the change in feed rate of this component and y_{D0} and x_{B0} are the initial product mole fractions of this component.

Later, Wahl and Harriot (1970) and Waller et al. (Toijala, 1969) introduced similar concepts (T_s and T_{eq}), but they considered only the time constant of the linearized system (called the "linearized time constant" in the following). Moczek et al. introduce (1) through somewhat intuitive arguments, and we will return with a complete derivation and interpretation below.

In spite of the excellent agreement found in several studies (Moczek et al., 1963, Weigand et al., 1972) between (1) and the actual time response of the column, the usefulness of (1) does not seem to be appreciated in the chemical engineering community. For example, Shinskey(1984, p.157) claims that the only general relationship that seems to hold for the dominant time constant is that it is proportional to $\sum M_i/F$. In fact, this follows directly from (1), but (1) certainly contains

much more information than this. The limited use of (1) is probably caused by the following two misconceptions:

- (i) (1) gives the value of the linearized time constant, that is, it is only useful for small perturbations from steady state.
- (ii) The linearized time constants are substantially in error (much larger) when compared to either the actual or simulated response.

Both these claims are incorrect. Misconception (i) is probably based on the work by Wahl and Harriot (1970) and Toijala (1969) who derived (1) for small perturbations. In fact, (1) can be used to estimate the "average" time constant between any two steady states. Misconception (ii) is due to the unfortunate assumption of equal product purities ($1 - y_D = x_B$) used in the majority of the academic case studies. It turns out, as we will show, that the time constant has its peak value approximately for $1 - y_D = x_B$. However, in general, the linearized time constant is not necessarily larger than the actual time constant (e.g., see Fig. 2). In fairness we should add that the claim (ii) is correct if both products are of high purity. The reason is that in this case the perturbations to the column will bring the column to a new steady state where one of the products is less pure. Since the time constant is determined by the least pure product this implies that the actual time constant will be smaller than the linearized one (see Section 3.).

A major source of misconception (ii) is probably the work of Wahl and Harriot (1970). They present a figure (Fig. 7 in their paper) for estimating the linearized time constant as a function of operating variables for the column. However, although not stated in the paper, these values apply only to the special case $1 - y_D = x_B$. The figure is therefore of very limited practical value and will generally yield too large values for the time constant. Similar figures, which are misleading for the same reason, are presented by Tyreus et al. (1975). Wahl and Harriot also claim that the time constant is relatively constant for a large perturbation in the loads;

this is not correct for high-purity columns as seen from Fig.1 .

Misconception (ii) is present in a recent work by Kapoor et al. (1986). They claim that "published tower time constants based on linear analysis have been substantially in error when compared to actual responses for many cases. The reason for this error is a ... positive feedback loop produced by the recycle tower structure ... (for which the) gain drops sharply for small perturbations from steady state". Firstly, the results based on linear analysis are not "in error" as they claim. (On the other hand, the range of validity may be very limited if both products are of high purity.) Secondly, the simple mixing-tank model (1) explains in at least as simple a fashion as the "positive feedback loops" when and why the linearized time constant is "in error". This is discussed in detail below.

2. DERIVATION OF EXPRESSION FOR τ_c

Consider a column which initially ($t = 0$) is at steady state (subscript 0). At $t = 0$ a step change is introduced to the column which eventually ($t \rightarrow \infty$) moves the column to a new steady state (subscript f). The nature of this step change is not important as long as i) the new steady state is known and ii) it leads to a change in the total holdup in the column of one or more component. This includes most disturbances and inputs except changes in the internal flows (changes in L and V keeping product rates constant).

Assumption 1. *The flow dynamics are immediate, i.e., for $t > 0$: $M_i(t) = M_{if}$, $D(t) = D_f$, $B(t) = B_f$.*

This assumption is also used in all simulations. The assumption is reasonable when considering the composition dynamics, provided the flow response is much faster than the composition response. Using Assumption 1 the overall material balance for any component for $t > 0$ becomes:

$$\frac{d}{dt} \left[\sum_{i=1}^{N+1} M_{if} x_i(t) \right] = F_f z_{Ff} - D_f y_D(t) - B_f x_B(t) \quad (2)$$

Subtracting the final steady state ($0 = F_f z_{Ff} - D y_{Df} - B_f x_{Bf}$) yields

$$\sum_{i=1}^{N+1} M_{if} \Delta \dot{x}_i(t) = -D_f \Delta y_D(t) - B_f \Delta x_B(t) \quad (3)$$

Assumption 2. *All trays have the same dynamic responses, i.e. : $\Delta x_i(t) = \Delta x_i k(t)$, $\Delta y_D(t) = \Delta y_D k(t)$, $\Delta x_B(t) = \Delta x_B k(t)$. (Here $k(0) = 1$ and $\Delta x_i = \Delta x_i(0)$, $\Delta y_D = \Delta y_D(0)$ and $\Delta x_B = \Delta x_B(0)$ denote the difference between the initial and final steady state.)*

Assumption 2 corresponds to treating the column as a large mixing tank. This assumption is reasonable if the time constant for the internal mixing in the column, $\tau_M = M_I/L$, is much shorter than the dominant time constant. τ_M is approximately the time it takes for a composition change at the top to travel to the bottom; for a composition change starting from the top each tray acts as a first order lag with time constant M_i/L . The overall transfer function is the product of these lags which may be approximated by a time delay with time constant $\tau_M = M_I/L$. τ_M was introduced by Harriot and Wahl (1970) who called it the circulation time. Assumption 2 and Eq.(3) yield

$$\left(\sum_{i=1}^{N+1} M_{if} \Delta x_i \right) \dot{k}(t) = (-D_f \Delta y_D - B_f \Delta x_B) k(t) \quad (4)$$

Solving (4) gives a linear first-order response

$$k(t) = e^{-t/\tau_c}$$

where the time constant τ_c is defined as (subscript c denotes *change* in component holdup):

$$\tau_c = \frac{\sum_{i=1}^{N+1} M_{if} \Delta x_i}{\Delta S_i}, \quad \Delta S_i \stackrel{\text{def}}{=} D_f \Delta y_D + B_f \Delta x_B \quad (5)$$

(ΔS_i is the supply imbalance). A simple interpretation of (5) is

$$\tau_c = \frac{\text{"change in holdup of one component" (kmol)}}{\text{"imbalance in supply of this component" (kmol/min)}}$$

Note that

$$\Delta S_i = D_f \Delta y_D + B_f \Delta x_B = \Delta(Fz_f) - y_{D0} \Delta D - x_{B0} \Delta B \quad (6)$$

and τ_c defined by (5) is therefore equal to T_{inv} defined by (1) if we *assume that the holdup on each tray is constant, i.e., $M_{if} = M_i$* . This assumption is used in the rest of the paper; it clearly has almost no significance on the value of τ_c .

Comments on (5):

1. The column model was not linearized, and (5) applies to any finite change provided Assumptions 1 and 2 hold.
2. The time constant depends on the magnitude and "direction" (negative or positive change) of the step change introduced.
3. The expression for τ_c applies to any component in a multicomponent mixture.
4. Eq. (5) applies to any change which changes the external material balance, i.e., which has $\Delta S_i \neq 0$. Eq. (5) does not apply for changes in the internal flows (changing L and V while keeping D and B constant) because the denominator $\Delta S_i = 0$ in this case (see (6)). Furthermore, in this case there is very little change in component holdup, and the entire holdup approach is not appropriate. Methods for estimating the time constant for changes in the internal flows are discussed in another paper (Skogestad and Morari, 1987b).
5. To compute τ_c according to (5) a steady-state model of the column is needed. For obtaining accurate numerical values a nonlinear simulation program should be used. Such programs are usually readily available to the engineer. For any given step change two simulations is all what is needed to compute τ_c . To simplify the computations the program should be modified to print out $\sum_{i=1}^{N+1} M_i x_i$.
6. Very large time constants are found for small perturbations to columns with both products of high purity. This agrees with the observations of Wahl and

Harriot(1970), Tyreus et al. (1975) and Fuentes and Luyben (1983). The reason is that the compositions inside the column may change significantly (the entire column profile may shift resulting in a large change in component holdup), while the change in product compositions may be very small (resulting in a small imbalance ΔS_i to cause the change in component holdup).

7. The expression for τ_c can be split into three contributions

$$\tau_c = \tau_{cI} + \tau_{cD} + \tau_{cB} \quad (7a)$$

$$\tau_{cI} = \frac{M_I \Delta \bar{x}_I}{\Delta S_i}, \tau_{cD} = \frac{M_D \Delta y_D}{\Delta S_i} \leq \frac{M_D}{D_f}, \tau_{cB} = \frac{M_B \Delta x_B}{\Delta S_i} \leq \frac{M_B}{B_f} \quad (7b)$$

where $M_I = \sum_{i=2}^N M_i$ is the total holdup and $\bar{x}_I = \sum_{i=2}^N x_i M_i / M_I$ is the average composition inside the column. The contribution to τ_c from the change in holdup inside the column (τ_{cI}) is often dominating. Furthermore, the reboiler and condenser are to some degree "decoupled" from the rest of the column, and their contribution to τ_c may be less than what is indicated by (7) (see Example 4 below).

8. One disadvantage of (5) is that the compositions on all trays are needed to compute τ_c . We will therefore proceed to derive an analytical expression for τ_c , based on a very simple model, which involves only the product compositions (y_D and x_B). This expression is useful for gaining insight into the nonlinear behavior of distillation columns.

Comparison with nonlinear simulations

Moczek et al. (1963) reported excellent agreement between (5) and the observed nonlinear response for a high-purity BTX-column. Weigand et al. (1972) studied six different columns and found very good agreement for the high-purity column (Column V) and reasonably good agreement for the five low-purity columns. The agreement was found to be best for small perturbations to high-purity columns, which is expected, since this gives large time constants and Assumption 1 and 2 are

likely to hold. For columns with only one product of high purity, τ_c was found to apply to the low-purity end, while the time constant for the high-purity end was often significantly smaller.

In this paper the seven columns (A-G) introduced in Part I are used as examples (Table 1.). In all examples we assume constant molar flows, constant liquid and vapor holdup (i.e., instantaneous flow responses), constant relative volatility, liquid feed, 100% tray efficiency, total condenser and equal holdup on all trays inside the column.

Example 1. Column A. ($M_i/F=0.5$ min, $M_D/F=32.1$ min, $M_B/F=11.1$ min) Table 2 compares τ_c with the actual time constant τ_1 observed for nonlinear responses to small and large changes in F , V and L for column A. Each nonlinear response was fitted by eye to a second order linear response $1/(1 + \tau_1 s)(1 + \tau_2 s)$ as shown in Fig. 1. (We do not propose this as a good column model, but it gives the reader an idea of the shape of the nonlinear response.) The agreement between τ_c and τ_1 is good for small perturbations. For the larger perturbations τ_c agrees better with $\tau_1 + \tau_2$. The only large deviation is found for the response in x_B to a large increase in $\Delta V/F_0 = 0.20$ (L constant); this change takes the column from the initial steady state with $y_D = 0.99$ and $x_B = 0.01$ to a new steady state with $y_D = 0.714$ and $x_B = 0.0006$. For the pure bottom product we find $\tau_1 + \tau_2 = 14.5$ min which is much smaller than $\tau_c = 75.3$ min. However, as observed by Weigand et al. (1972) the agreement is very good for the less pure top product ($\tau_1 + \tau_2 = 75$ min).

Example 2. Column D. ($M_i/F=M_D/F=M_B/F=1$ min) To show that the linearized time constant is not necessarily larger than the actual "nonlinear" time constant, we considered a feed composition disturbance to column D. The linearized time constant for a small disturbances in z_F found using (5) is $\tau_c = 319$ min. This compares nicely with the actual response which has a time constant of about 388 min for y_D and 341 min for x_B (Fig. 2A). Next, we studied a 7.7%

Column	z_F	α	N	N_F	y_D	x_B	D/F	L/F	$\lambda_{11}(0)$
A	0.5	1.5	40	21	0.99	0.01	0.500	2.706	35.1
B	0.1	"	"	"	"	"	0.092	2.329	47.5
C	0.5	"	"	"	0.90	0.002	0.555	2.737	7.53
D	0.65	1.12	110	39	0.995	0.10	0.614	11.862	58.7
E	0.2	5	15	5	0.9999	0.05	0.158	0.226	2.82
F	0.5	15	10	5	0.9999	0.0001	0.500	0.227	499
G	0.5	1.5	80	40	0.9999	0.0001	0.500	2.635	1673
WH	0.5	2.0	26	13	0.995	0.005	0.500	1.477	41.6

Table 1. Steady-state data for distillation column examples. $\lambda_{11}(0)$ denotes the 1-1-element in the RGA for the LV-configuration .

Input	Holdup model (7)				Fit of nonlinear response to $\frac{1}{(\tau_1 s + 1)(\tau_2 s + 1)}$			
	τ_{cI}	τ_{cD}	τ_{cB}	τ_c = Sum	τ_1 (y_D)	τ_2 (y_D)	τ_1 (x_B)	τ_2 (x_B)
$\Delta F/F_0 = 0.001$	190	24.6	13.7	228	207	50	207	0
$\Delta V/F_0 = 0.001$	194	30.6	11.6	236	210	50	210	0
$\Delta L/F_0 = 0.001$	190	26.4	13.1	229	228	20	228	0
$\Delta L/F_0 \rightarrow 0$	191	28.7	12.3	232				
$\Delta F/F_0 = 0.150$	41.1	1.9	16.6	59.6	50	25	50	15
$\Delta V/F_0 = 0.200$	29.6	45.2	0.5	75.3	50	25	14.5	0
$\Delta V/F_0 = -0.20$	25.7	1.5	15.6	42.8	43	0	25	20
$\Delta L/F_0 = 0.200$	26.5	1.5	15.6	43.7	40	0	25	20

Table 2. Column A. Time constants for inputs of various magnitude. Note that (7) often overestimates the magnitude of τ_{cD} and τ_{cB} (see Example 4). ($M_I/F = 0.5$ min, $M_D/F = 32.1$ min, $M_B/F = 11.1$ min)

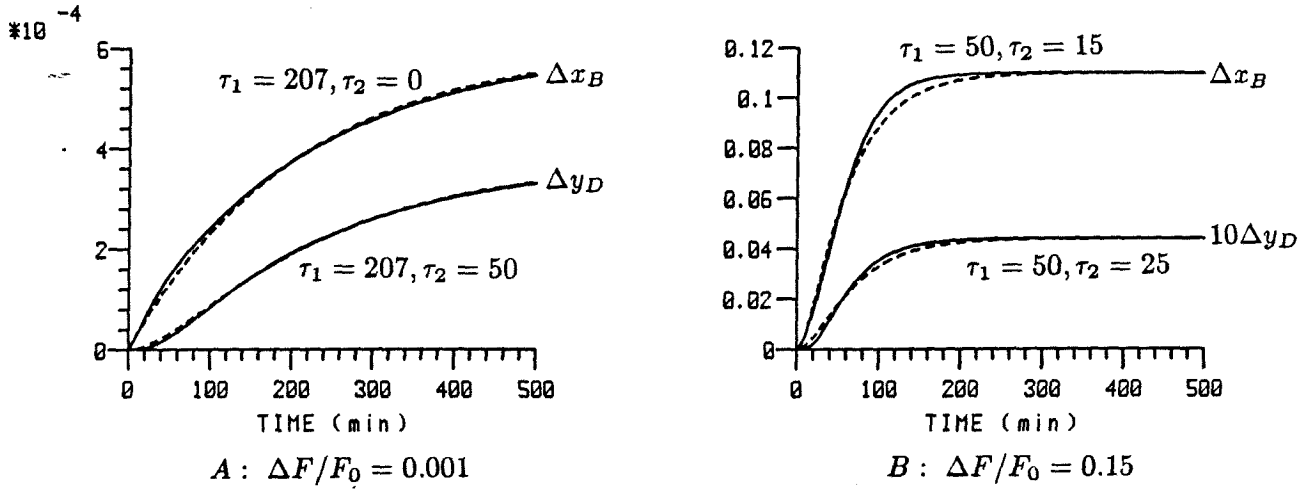


Fig. 1. Column A. Nonlinear open-loop response to small (A) and large (B) change in feed rate F with L and V constant. Dotted line : Approximation with second-order response $k/(\tau_1 s + 1)(\tau_2 s + 2)$. ($M_i/F = 0.5$ min, $M_D/F = 32.1$ min, $M_B/F = 11.1$ min)

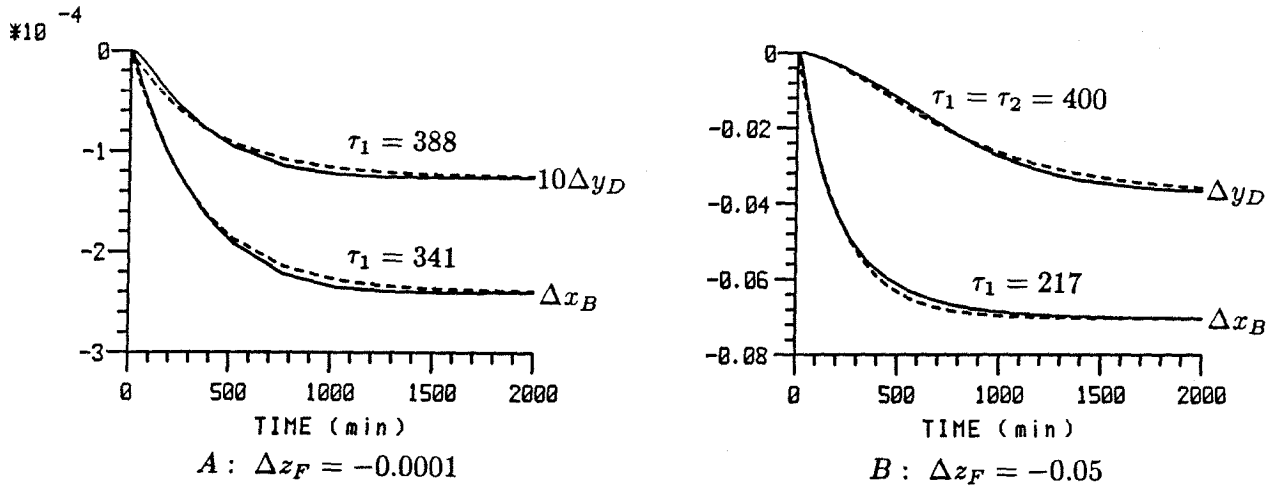


Fig. 2. Column D. Nonlinear open-loop response to small (A) and large (B) change in feed composition z_F . Dotted line : Approximation with second-order response $k/(\tau_1 s + 1)(\tau_2 s + 2)$, $\tau_2 = 0$ if not specified. ($M_i/F = M_D/F = M_B/F = 1$ min)

decrease in z_F (Fig. 2B):

Initial steady state: $z_F = 0.65$, $y_D = 0.9950$, $x_B = 0.1000$, $\bar{x}_I = 0.714$

Final steady state: $z_F = 0.60$, $y_D = 0.9578$, $x_B = 0.0296$, $\bar{x}_I = 0.495$

We find for this change

$$\tau_{cI} = \frac{(N-1)M_i\Delta\bar{x}_I}{\Delta(Fz_F)} = \frac{109 \ 1 \ 0.219}{0.05} = 477 \text{ min}$$

$$\tau_{cB} = \frac{M_B\Delta x_B}{\Delta(Fz_F)} = 1.4 \text{ min}, \quad \tau_{cD} = \frac{M_D\Delta y_D}{\Delta(Fz_F)} = 0.7 \text{ min}$$

Consequently, for this disturbance in z_F , (5) gives $\tau_c = 479$ min, which is higher than the linearized value of 319 min. This is also confirmed by the simulations in Fig.2B; the response of y_D to the large disturbance in z_F is clearly more sluggish ($\tau_1 \approx \tau_2 \approx 400$ min) than for the small disturbance.

Linearized time constant

The linearized time constant is derived from (7) by replacing the Δ 's by differentials:

$$\tau_{cI} = \frac{M_I d\bar{x}_I/dx_B}{B + Ddy_D/dx_B}, \quad \tau_{cD} = \frac{M_D}{D + Bdx_B/dy_D}, \quad \tau_{cB} = \frac{M_B}{B + Ddy_D/dx_B} \quad (8)$$

These values are in general different depending on the disturbance or input because the linearized gains (dy_D/dx_B) are different. The values of τ_{cD} and τ_{cB} are easily obtained from the steady-state gain matrix.

Example 3. Column A. ($M_i/F=0.5$ min, $M_D/F=32.1$ min, $M_B/F=11.1$ min)

The steady-state gain matrix from Part I (Skogestad and Morari, 1987b) is

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix}$$

For a small change in L with V constant this gives $(dy_D/dx_B)_V = 0.878/1.082 = 0.811$ which yields $\tau_{cD} = 28.7$ min and $\tau_{cB} = 12.3$ min. The value of $(d\bar{x}_I/dx_B)_V = 8.85$ for a small perturbation in L was obtained numerically. This gives

$$\tau_{cI} = \frac{39 \ 0.58.85}{0.5 + 0.5 \ 0.811} = 191 \text{ min}, \quad \tau_c = \tau_{cI} + \tau_{cD} + \tau_{cB} = 232 \text{ min}$$

Very similar values are obtained for a small change in V (L constant) since $(dy_D/dx_B)_L = 0.864/1.096 = 0.799$ and $(d\bar{x}_I/dx_B)_L = 8.74$ are almost unchanged. Note that the 1-1 element in the RGA for the LV-configuration is given by (Shinsky,1984)

$$\lambda_{11} = \left(1 - \frac{(dy_D/dx_B)_L}{(dy_D/dx_B)_V}\right)^{-1} \quad (9)$$

and columns with a large RGA-elements are therefore expected to yield similar values for τ_c for small perturbations in L or V. This will correspond to columns with both products of high purity.

Comparison with eigenvalues

To find the linearized dynamics more accurately, linearize the material balance and VLE on each tray

$$M_i \dot{x}_i = Ldx_{i+1} - (L + K_i V)dx_i + K_{i-1} V dx_{i-1} + (x_{i+1} - x_i)dL - (y_i - y_{i-1})dV$$

Here K_i is the linearized VLE-constant

$$K_i = \frac{dy_i}{dx_i} = \frac{\alpha}{(1 + (\alpha - 1)x_i)^2}$$

Written in the standard state-variable form in terms of deviation variables

$$\dot{x} = Ax + Bu + Ed, \quad y = Cx \quad (10)$$

Here $x = (dx_1, \dots, dx_{N+1})^T$ are the tray compositions, $u = (dL, dV)^T$, $d = (dF, dz_F)^T$ and $y = (dy_D, dx_B)^T$. Written on transfer matrix form

$$y = C(sI - A)^{-1}(Bu + Ed) \quad (11)$$

The poles (eigenvalues of A) are the same for any input or disturbance and independent of where composition is measured (y_D, x_B, x_i). However, the zero locations are different which may yield entirely different responses.

Wahl and Harriot (1970) found very good agreement between the linearized value of τ_c and the time constant τ_{1e} corresponding to the smallest eigenvalue of A, and this is confirmed by our results. Wahl and Harriot found that the agreement was acceptable (errors less than 20%) even for some cases with $\tau_M/\tau_c < 1$.

Example 4. Column A. The value of the time constants corresponding to the three smallest eigenvalues of A are (all numbers in minutes, τ_c from Example 1)

M_i/F	M_D/F	M_B/F	τ_{1e}	τ_{2e}	τ_{3e}	τ_c
0.5	0	0	193	12	3.4	191
0.5	32.1	11.1	220	32	11.6	232

Note that τ_c and τ_{1e} are almost identical for the case $M_D = M_B = 0$. This indicates that Assumption 2 holds very well for what happens inside the column. However, with $M_D/F=32.1$ min and $M_B/F= 11.1$ min the increase in τ_{1e} is only 27 minutes, while τ_c increases by 43 minutes. This indicates that Assumption 2 does not hold for the reboiler and condenser, and that these are partially decoupled from the rest of the column. This is not surprising since the larger holdups in the reboiler and condenser make these less sensitive to "interactions" with the other trays. In addition, there is only one stream entering the reboiler and condenser, while the trays inside the column have two.

τ_{1e} and linearized values of τ_c (5) are compared for some other columns with small reboiler and condenser holdup in Table 3. The agreement is amazing for columns A,B,F,G and WH all of which have $\tau_M \ll \tau_c$, and also for the other cases the agreement is very good. The only exception is column E for which τ_c for a change in V is 42% smaller than τ_{1e} . This is not surprising because $\tau_c/\tau_M = 1.34$ in this case and Approximation 2 is not likely to be valid. The large difference in the linearized time constant τ_c for changes in L and V which is observed in this case is also expected because of the low values for the RGA for this column ($\lambda_{11} = 2.82$) which implies that $(dy_D/dx_B)_L$ and $(dy_D/dx_B)_V$ are quite different.

<i>Column</i>			Eigenvalues		
	$\tau_c(5)$	$\tau_{sc}(17)$	τ_{1e}	τ_{2e}	τ_M
<i>A</i>	387	427	388	24	15
<i>B</i>	504(496)	428	500	23	18
<i>C</i>	58(60)	91	49	16	15
<i>D</i>	300(299)	385	308	47	9
<i>E</i>	142(95)	29	165	16	71
<i>F</i>	5992	4886	5992	9	48
<i>G</i>	40664	42891	40667	40	31
<i>WH</i>	459	477	459	15	18

Table 3. Linearized time constants (min). Values for τ_c (5) are given for small perturbations in L with V constant (and, if different, value for small change in V with L constant is in parantheses). Note that $\tau_M = M_I/L \ll \tau_c$ for Assumption 2 to hold. All columns: $M_i/F = M_D/F = M_B/F = 1$ min.

The simple holdup model (5) explains exactly why large linearized time constants are observed for high-purity columns: There is a "large" change in component holdup inside the column, but only a small imbalance (change in product compositions) to bring about the change. Other explanations for the large time constants are given in the literature: Kapoor et al. (1986) claim that they are caused by "positive feedback loops produced by the recycle structure". Fuentes and Luyben (1985) claim that the cause is "small concentration changes from tray to tray (which) make some of the coefficients in the linearized equation very small, giving small eigenvalues". Both of these explanations seem unnecessary complicated.

Note from Table 3 that very large linearized time constant may be encountered even for easy separations with few trays: Column F has 10 theoretical trays and $L/D=0.45$, yet $\tau_c = \tau_{1e} = 5992$ min. The reason for the large value of the linearized time constant is the high purity of the products ($1 - y_D = x_B = 10^{-4}$). However, the time constant will be drastically smaller for any realistic perturbation to the column. This is discussed below.

3. A SIMPLE FORMULA FOR τ_c

We will now use (5) to derive a simple formula for the linearized time constant for the case of a binary separation. The following additional assumptions are made:

Assumption 3. All trays ($i = 2, N$) have equal and constant holdup (M_i).

Assumption 4. The average composition inside the column is (Appendix 1):

$$\bar{x}_I = \left(1 + \frac{\ln \frac{1-x_B}{x_B}}{\ln \frac{y_D}{1-y_D}} \right)^{-1} \quad (12)$$

Assumption 5: The separation factor $S = \frac{y_D(1-x_B)}{(1-y_D)x_B}$ is constant for any given change considered.

(12) was derived by assuming the composition profile is the same as that of a column with total reflux, and Assumption 4 is therefore most likely to hold for columns with

large reflux. Assumption 5 is most likely to hold for high-purity columns because then the term e_ζ in (15) will dominate.

In this section let y_D , x_B and x_i denote the mole fraction of the light component, and let ζ represent any input or disturbance to the column which changes the external material balance such that (5) holds. For a small perturbation, $d\zeta$, (5) yields

$$\tau_c = \frac{M_I \frac{\partial \bar{x}_I}{\partial \zeta} + M_D \frac{\partial y_D}{\partial \zeta} + M_B \frac{\partial x_B}{\partial \zeta}}{D \frac{\partial y_D}{\partial \zeta} + B \frac{\partial x_B}{\partial \zeta}} \quad (13)$$

Differentiating the component material balance, $Fz_F = Dy_D + Bx_B$, yields the following expression for the denominator of (13)

$$D \frac{\partial y_D}{\partial \zeta} + B \frac{\partial x_B}{\partial \zeta} = e_\zeta \quad (14a)$$

where e_ζ is defined as

$$e_\zeta = -(y_D - x_B) \frac{\partial D}{\partial \zeta} + F \frac{z_F}{\partial \zeta} + (z_F - x_B) \frac{\partial F}{\partial \zeta} \quad (14b)$$

To evaluate the numerator, we need to find the linearized gains $\partial y_D / \partial \zeta$ and $\partial x_B / \partial \zeta$. (This will also yield $\partial \bar{x}_I / \partial \zeta$ because of Assumption 4.) For binary mixtures exact expressions for these gains are derived by combining (14a) with the definition of the separation factor S (see Part I (Skogestad and Morari, 1987a))

$$\frac{1}{(1 - y_D)y_D} \frac{\partial y_D}{\partial \zeta} = \frac{1}{I_s} (e_\zeta + Bx_B(1 - x_B) \frac{\partial \ln S}{\partial \zeta}) \quad (15a)$$

$$\frac{1}{(1 - x_B)x_B} \frac{\partial x_B}{\partial \zeta} = \frac{1}{I_s} (e_\zeta - Dy_D(1 - y_D) \frac{\partial \ln S}{\partial \zeta}) \quad (15b)$$

Here I_s is the "sum" of impurities leaving the column

$$I_s = Bx_B(1 - x_B) + Dy_D(1 - y_D) \quad (16)$$

For upsets $d\zeta$ which change the external material balance ($e_\zeta \neq 0$), the term involving the change in separation factor S is usually of minor importance (Part I), and

a reasonable approximation is found by assuming S constant (Assumption 5). (15) yields

$$\begin{aligned} \left(\frac{\partial y_D}{\partial \zeta} \right)_S &= \frac{(1 - y_D)y_D e_\zeta}{I_S} & \left(\frac{\partial x_B}{\partial \zeta} \right)_S &= \frac{(1 - x_B)x_B e_\zeta}{I_s} \\ \left(\frac{\partial \bar{x}_I}{\partial \zeta} \right)_S &= \frac{\partial}{\partial \zeta} \left(\frac{\ln \frac{y_D}{1 - y_D}}{\ln S} \right)_S = \frac{1}{\ln S} \frac{\partial}{\partial \zeta} \left(\ln \frac{y_D}{1 - y_D} \right) = \frac{e_\zeta}{I_s \ln S} \end{aligned}$$

Substituting this into (13) yields a short-cut formula for the linearized value of τ_c (note that e_ζ drops out)

$$\tau_{sc} = \tau_{scI} + \tau_{scD} + \tau_{scB} = \frac{M_I}{I_s \ln S} + \frac{M_D(1 - y_D)y_D}{I_s} + \frac{M_B(1 - x_B)x_B}{I_s} \quad (17)$$

For the special case of equal purities ($x_B = 1 - y_D$) we get $I_s = F x_B y_D$ and (17) becomes

$$\tau_{sc} = \frac{M_I/F}{x_B y_D \ln S} + M_D/F + M_B/F \quad (18)$$

Here the term $1/x_B y_D \ln S$ which multiplies the holdup inside the column is

$x_B = 1 - y_D :$	0.3	0.1	0.01	0.001	10^{-4}	10^{-6}
$1/x_B y_D \ln S :$	2.81	2.53	11.0	72.5	543	36191

This clearly shows that usually the contribution to τ_c from the holdup inside the column dominates, especially for separations with both products of high purity. Also recall that the contributions to τ_c from the condenser and reboiler holdup are generally overestimated by (17) and (18) because Assumption 2 does not hold .

Values for τ_{sc} found using (17) are compared with τ_c (8) in Table 3. The agreement is of course best for the cases when Assumptions 4 and 5 are likely to hold, i.e., for columns with high reflux (Assumption 4 is valid) and for high-purity columns (Assumption 5 is valid). However, the main value of the analytical formula (17) is the insight it yields into the nonlinear dynamic behavior of distillation columns. Consider the contribution from the holdup inside the column

$$\tau_{scI} = \frac{M_I}{I_s \ln S} \quad (19)$$

In S usually does not change very much with operating conditions. The time constant is therefore determined mainly by I_s which again is determined by the composition of the least pure product:

$$I_s \approx \begin{cases} D(1 - y_D) & \text{if distillate least pure (} (1 - y_D) \gg x_B \text{)} \\ Bx_B & \text{if bottoms least pure (} x_B \gg (1 - y_D) \text{)} \end{cases} \quad (20)$$

For columns with both products of high purity any disturbance or input to the column is likely to take the column to a new steady state where one of the products is less pure, and the actual time constant for this change will be smaller than found by linear analysis (which (19) is based on).

It also follows from (19) that τ_{cI} reaches its maximum value approximately when both products have equal purity; by differentiating (19) we find that for $z_F=0.5$ the maximum value of τ_{cI} when D/B is varied is obtained for $D/B = 1$ corresponding to $1 - y_D = x_B$. Also for other values of z_F the maximum is obtained when $(1 - y_D)/x_B$ is of the order one.

Example 5. Column WH. Consider the high-purity column ($1 - y_D = x_B = 0.005$) presented by Wahl and Harriot (1970). In Fig. 3 the linearized time constant τ_{cI} is shown as a function of z_F with all flows fixed. Each value of z_F corresponds to a specific steady state with a given ratio $x_B/(1 - y_D)$ and we have also plotted τ_{cI} against this ratio. We see that τ_{cI} has a very high peak at the nominal operating point $z_F = 0.5$ which corresponds to $x_B/(1 - y_D) = 1$. This illustrates that Fig. 7 in Wahl and Harriot (1970) for estimating the linearized time constant is highly misleading; it is based on the assumption $1 - y_D = x_B$. Also note from Fig.3 that very similar values for τ_c are obtained from (8) and the short-cut formula (18).

Example 6. Column B. A similar example for column B is shown in Fig. 4. In this case $z_F = 0.1$ and $L/F = 2.329$ are fixed, and different steady states are obtained by varying D/F . The nominal operating points has $D/F = 0.0918$ corresponding to $(1 - y_D)/x_B = 1$ which yields $\tau_{cI} = 501$ min. The peak value of the linearized time constant τ_{cI} is 656 min which is obtained for $D/F = 0.0952$ corresponding to

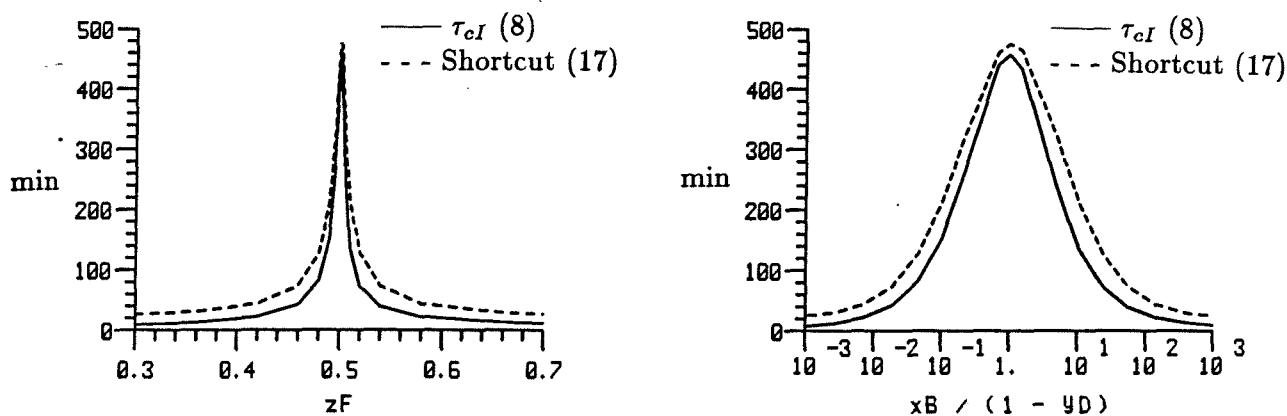


Fig. 3. Column WH. Linearized time constant τ_{cI} as a function of z_F with all flows fixed. Each value of z_F corresponds to a new steady state with a particular y_D and x_B , and the right figure is included to show clearer the variation in τ_{cI} with $\frac{x_B}{1-y_D}$. ($M_i/F = M_D/F = M_B/F = 1$ min)

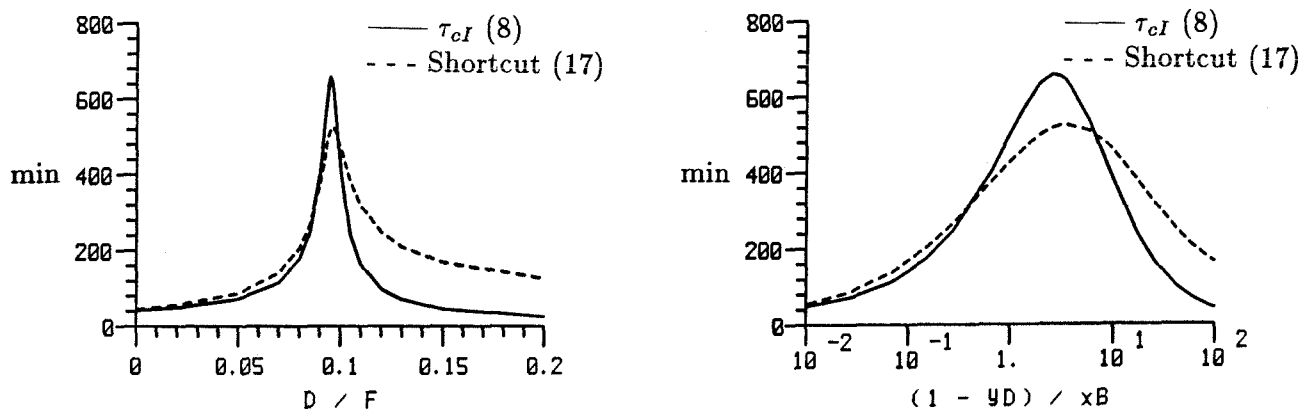


Fig. 4. Column B. Linearized time constant τ_{cI} as a function of D/F (with $z_F = 0.1$ and $L/F = 2.329$ fixed). Each value of D/F corresponds to a particular $\frac{1-y_D}{x_B}$. ($M_i/F = M_D/F = M_B/F = 1$ min)

$$(1 - y_D)/x_B = 2.8.$$

4. DISCUSSION AND CONCLUSION

The dynamics of most distillation columns are dominated by one large time constant (τ_c), which is nearly the same, regardless of where a disturbance is introduced or where composition is measured. Physically, this dominant time constant reflects the change in component holdup inside the column. Based on the inventory time concept introduced by Moczek et al. (1963) we have derived a simple formula for estimating its linearized value for binary separations

$$\tau_{cI} \approx \tau_{scI} = \frac{M_I}{I_s \ln S} \quad (19)$$

(In addition there are contributions to τ_c from the reboiler and condenser holdup; these are usually of less importance.) The formula (19) gives reasonable agreement with observed values, but its main value is the simple analytical form which provides insight into the nonlinear behavior of distillation columns. In particular, the value of τ_{cI} is determined mainly by the purity of the least pure product (Eq. (20)) and it will be large when both products have high purity (I_s is small).

From the derivation of τ_c it is clear that this dominant time constant only applies if there is a change in the external material balance. The time constant τ_2 for changes in the internal flows is often much smaller. As an example consider Fig.5 which shows the condition number as a function of frequency for column A. If a single time constant τ_c were used for all transfer function elements, corresponding to the model $G(s) = G(0)/(1 + \tau_c s)$, then the condition number would be the same at all frequencies. This is clearly not the case as seen from Fig.5, and the lower value at high frequencies is caused by τ_2 being significantly smaller than τ_c . This is discussed in detail in another paper (Skogestad and Morari, 1987b) where we propose a model in terms of τ_c and τ_2 which is consistent with the behavior in Fig.5 .

From the derivation of (5) we know that if Assumptions 1 and 2 hold then the composition response should be first order, even for large deviations from steady

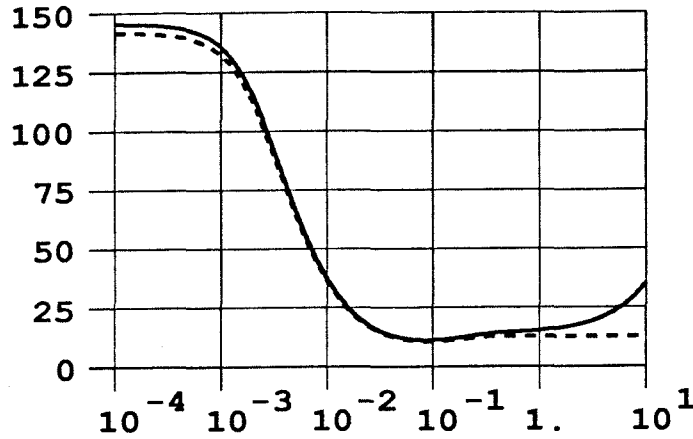


Fig. 5. Column A. Condition number $\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G)$ and $\|RGA\|_1 = \sum |\lambda_{ij}|$ (dotted line) as a function of frequency. ($M_i/F = M_D/F = M_B/F = 1$ min)

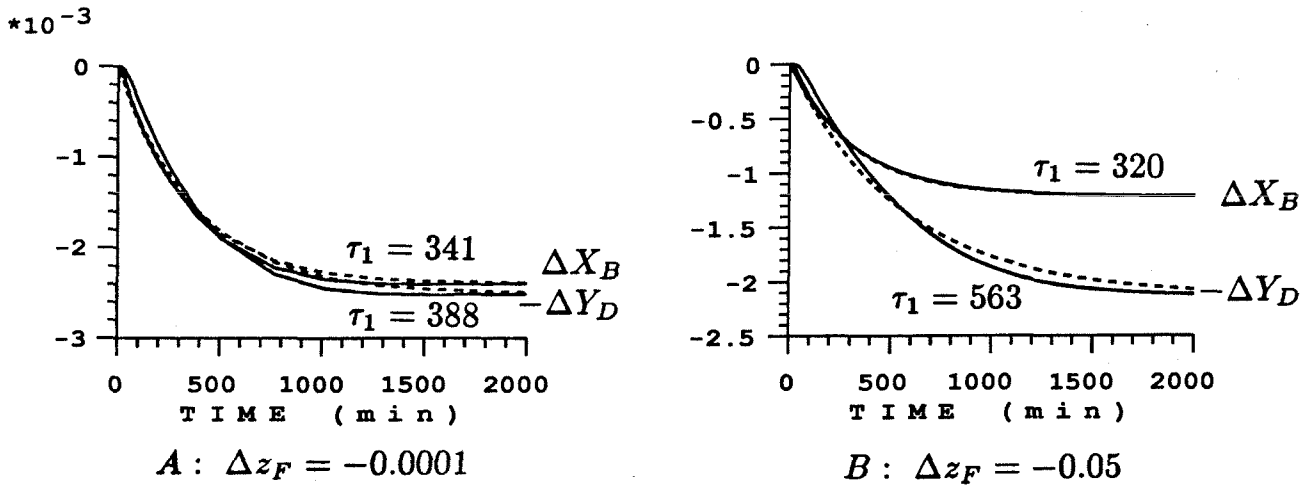


Fig. 6. Column D. Nonlinear open-loop responses to small (A) and large (B) change in z_F . logarithmic compositions $Y_D = \ln(1 - y_D)$ and $X_B = \ln x_B$ are used. Dotted line: Approximation with first-order response $k/(\tau_1 s + 1)$. The responses should be compared to the ones in terms of y_D and x_B in Fig.2.

state. The simulations indicate that the first order approximation is indeed valid for small perturbations. However, recall Fig.2 which shows that the response in y_D to a small decrease in z_F is first order (Fig.2A), but that it is approximately second order for a large decrease (Fig.2B). In the last case y_D decreases from 0.995 to 0.958, that is, the amount of heavy impurity increases from 0.005 to 0.042. Although the initial response in terms of absolute compositions is sluggish (second order), there is a large relative change in the amount of impurity. This suggests that a lower-order response may be obtained if logarithmic compositions are used

$$Y_D = \ln(1 - y_D) \quad \text{and} \quad X_B = \ln x_B \quad (21)$$

Fig.6 shows the same responses as in Fig.2, but using logarithmic compositions. The time constants of the first- and second-order approximations of the observed responses are summarized below:

	$\Delta z_F = -0.0001 :$	$\Delta z_F = -0.05 :$
y_D	$\tau_1 = 388$	$\tau_1 = \tau_2 = 400$
Y_D	$\tau_1 = 388$	$\tau_1 = 563$
x_B	$\tau_1 = 341$	$\tau_1 = 217$
X_B	$\tau_1 = 341$	$\tau_1 = 320$

The responses are much less dependent on the magnitude of the disturbance when logarithmic compositions are used, and the observed second order response for y_D almost "disappears" (becomes first order) when Y_D is used. This confirms the results by Skogestad and Morari (1987b) who suggest using logarithmic compositions as a means of reducing the effect of nonlinearity. Note that for small deviations from steady state the use of logarithmic compositions merely corresponds to a rescaling of the outputs and the time constants in terms Y_D and X_B are the same as for y_D and x_B . This of course means that the linearized time constants derived in this paper (Eq. 8 and 17) also apply to Y_D and X_B .

However, as shown above, for large deviations from steady state the responses are different, and it seems that the response in terms of logarithmic compositions is

closer to first order than when using y_D and x_B . We can not explain this from (5), which was derived in terms of absolute compositions, and our only justification is that the linear response at each operating point is approximately first order (from (5)) and that the use of logarithmic compositions makes the linear response only weakly dependent on operating conditions (Skogestad and Morari, 1987b).

Acknowledgements. Partial support from the National Science Foundation and Norsk Hydro is gratefully acknowledged.

NOMENCLATURE

B - bottom product rate (kmol/min)

D - distillate (top product) rate (kmol/min)

F - feed rate (kmol/min)

$I_s = Dy_D(1 - y_D) + Bx_B(1 - x_B)$ - "impurity sum"

L - reflux flow rate (kmol/min)

M_B - holdup in reboiler (kmol)

M_D - holdup in condenser (kmol)

M_i - holdup on tray i (kmol)

$M_I = \sum_{i=2}^N M_i$ - total holdup of liquid inside column (kmol)

N - total number of theoretical trays (incl. reboiler)

N_F - feed tray location from bottom (feed enters above this tray)

$S = \frac{y_D(1-x_B)}{(1-y_D)x_B}$ - separation factor

V - boilup from reboiler (kmol/min)

x_B - mole fraction of light component in bottom product

x_i - liquid mole fraction of light component on stage i

$\bar{x}_I = \sum_{i=2}^N x_i M_i / M_I$ - average liquid mole fraction inside column

$y_D = x_D$ - mole fraction of light component in distillate (top product)

y_i - vapor mole fraction of light component on stage i

z_F - mole fraction of light component in feed

$\alpha = \frac{y_i/x_i}{(1-y_i)/(1-x_i)}$ - relative volatility

τ_c - dominant time constant for change in holdup (min)

τ_{sc} - shortcut estimate of τ_c

$\tau_{je} = 1/\lambda_j(A)$ - time constant corresponding to the j -th smallest eigenvalue of A in Eq.(10) (min)

$\tau_M = M_I/L$ - mixing time for column (min)

Subscripts

B - bottom product

D - distillate product

I - inside column

f - final steady state

0 - initial steady state

i - tray no. numbered from bottom ($i=1$ for reboiler, $i=2$ for first tray, $i=N$ for top tray, $i=N+1$ for condenser)

REFERENCES

- Fuentes, C. and W. L. Luyben, Control of High-Purity Distillation Columns, *Ind. Eng. Chem. Process Des. Dev.*, **22**, 361-366 (1983)
- Kapoor, N., T. J. Mcavoy and T. E. Marlin, Effect of Recycle Structure on Distillation Tower Time Constants, *AIChE Journal*, **32**, 3, 411-418 (1986)
- Kim, C. and C. F. Friedly, Dynamic Modeling of Large Staged Systems, *Ind. Eng. Chem. Process Des. Dev.*, **13**, 2, 177-181 (1974)
- Levy, R. E. , A. S. Foss and E. A. Greens II, Response Modes of a Binary Distillation Column, *Ind. Eng. Chem. Fund.*, **8**, 4, 765-776 (1969)
- Moczek, J. S. , R. E. Otto and T. J. Williams, Approximation Models for the Dynamic Response of Large Distillation Columns, *Proc, 2nd IFAC Congress, Basel* (1963). Also published in: *Chem. Eng. Progress Symp. Series*, **61**, 136-146 (1965)
- McNeill, G. A. and J. D. Sachs, High Performance Column Control, *Chem. Eng. Progress*, **65**, 3, 33-39 (1969)
- Rademaker, O., J. E. Rijnsdorp and A. Maarleveld, *Dynamics and Control of Distillation Columns*, Elsevier, Amsterdam (1975)
- Shinskey, F. G., *Distillation Control*, 2nd Edition, McGraw-Hill, New York (1984)
- Skogestad, S. and M. Morari, Shortcut models for distillation columns - I. Steady-State Behavior, *This issue of Comp. & Chem. Eng.* (1987a)
- Skogestad, S. and M. Morari, Understanding the Dynamic Behavior of Distillation Columns, in preparation (1987b)
- Toijala (Waller), K. , *Acta Acad. Aboensis Math Phys.*, **29**, 10 (1969) (as quoted in K. Toijala and S. Gustafsson, On The General Characteristics of Multicomponent Distillation, *Kemian Teollisuus*, **29**, 3, 173-184 (1972))
- Tyreus, B. D., W. L. Luyben and W. E. Schlosser, Stiffness in Distillation Models and the Use of Implicit Integration Methods to Reduce Computation Times, *Ind. Eng. Chem. Process Des. Dev.*, **14**, 427-433 (1975)
- Wahl, E. F. and P. Harriot, Understanding and Prediction of the Dynamic Behavior of Distillation Columns, *Ind. Eng. Chem. Process Des. Dev.*, **9**, 3, 396-407 (1970)
- Weigand, W. A., A. K. Jhavar and T. J. Williams, Calculation Method for the Response Time to Step Inputs for Approximate Dynamic Models of Distillation Columns, *AIChE Journal*, **18**, 6, 1243-1252 (1972)

APPENDIX. Formula for \bar{x}_I .

The expression (12) for \bar{x}_I is derived by assuming

- (i) All trays have the same holdup (Assumption 3).
- (ii) Constant relative volatility α .
- (iii) The shape of the composition profile is the same as that derived from Fenske's exact equation for total reflux

$$\frac{y_i/1-y_i}{x_B/1-x_B} = \alpha^i$$

- (iv) This composition profile may be approximated by straight lines:

$$\begin{aligned} x_B \leq x_i \leq w : \quad \bar{x} &\approx 0 & (i \leq N_w) \\ w < x_i < 1-w : \quad \bar{x} &\approx 0.5 \\ 1-w \leq x_i \leq y_D : \quad \bar{x} &\approx 1 & (i \geq N_{1-w}) \end{aligned}$$

(N_w and N_{1-w} are defined by the above equations, w is a fixed number.)

The average composition in the column from (iii)

$$\bar{x}_I = \frac{1}{N_{min} - 1} \int_{i=1}^{N_{min}} x_i di$$

Using (iv)

$$\bar{x}_I \approx \frac{1}{N_{min} - 1} [0.5(N_{1-w} - N_w) + 1.0(N_{min} - N_{1-w})] \quad (A1)$$

Here N_{min} , N_w and N_{1-w} are found using (ii)

$$N_{min} = \frac{\ln S}{\ln \alpha}, \quad N_w = \frac{\ln \frac{w}{1-w} \frac{1-x_B}{x_B}}{\ln \alpha} + 1, \quad N_{1-w} = \frac{\ln \frac{1-w}{w} \frac{1-x_B}{x_B}}{\ln \alpha} + 1$$

This gives

$$\begin{aligned} 0.5(N_{1-w} - N_w) &= \frac{\ln \frac{1-w}{w}}{\ln \alpha} \\ (N_{min} - N_{1-w}) &= \frac{\ln \frac{y_D}{1-y_D} - \ln \frac{1-w}{w} - 1}{\ln \alpha} \end{aligned}$$

Note that w drops out when this is substituted into (A1) and we get

$$\bar{x}_I = \frac{\ln \frac{y_D}{1-y_D} - \ln \alpha}{\ln S - \ln \alpha} \approx \frac{\ln \frac{y_D}{1-y_D}}{\ln S} = \left(1 + \frac{\ln \frac{1-x_B}{B}}{\ln \frac{y_D}{1-y_D}} \right)^{-1}$$

Chapter XI

UNDERSTANDING THE DYNAMIC BEHAVIOR OF DISTILLATION COLUMNS

UNDERSTANDING THE DYNAMIC BEHAVIOR OF DISTILLATION COLUMNS

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January, 1987

Submitted to Ind. & Eng. Chemistry Research

Abstract

The dynamic behavior of a distillation column is approximated with a two time constant model. The response to changes in the external flows is approximately first order with time constant τ_1 . This dominant time constant can be estimated using a simple mixing tank model for the column. The response to changes in the internal flows is also first order, but its time constant τ_2 is generally significantly smaller than τ_1 .

The condition number and the RGA are smaller at high frequency than at steady state. Most models presented in the literature do not take this into account. The two time constant model does predict this behavior, and τ_2 can be estimated by matching the RGA at high frequency. Finally, it is shown that the effect of nonlinearity is almost eliminated if logarithmic compositions $\ln x_B$ and $\ln(1 - y_D)$ are used. In particular, this applies to the initial response which is of primary importance for feedback control.

1. INTRODUCTION

The dynamic response of most distillation columns (Fig. 1) is dominated by one large time constant, which is nearly the same, regardless of where a disturbance is introduced or where composition is measured. This is well known both from plant measurements (McNeill and Sachs, 1969) and from theoretical studies (Moczek et al., 1963). Furthermore, the value of this time constant is largely unaffected by the flow dynamics.

The composition dynamics are described by one differential equation for each tray in the column

$$\frac{d}{dt}(M_i x_i) = L_{i+1} x_{i+1} + V_{i-1} y_{i-1} - L_i x_i - V_i y_i \quad (1)$$

It is somewhat surprising that the response of a distillation column with, for example, 100 trays, corresponding to at least a 100th order model, may be adequately described by a simple first-order model. Skogestad and Morari (1987a) and others (Moczek et al., 1963, Wahl and Harriot, 1970) have studied this in more detail. They found that the main reason for the low-order behavior is that all the trays have essentially the same composition response. This leads to the conclusion that the distillation column can be approximated by one large mixing tank, for which the time constant τ_{1c} (subscript c denotes change) is given by

$$\begin{aligned} \tau_{1c} &\approx \frac{\Delta(\sum_{i=1}^N M_i x_i)}{D_f \Delta y_D + B_f \Delta x_B} \\ &= \frac{\text{"change in holdup of one component inside column (mol)"}}{\text{"imbalance in supply of this component (mol/s)"}} \end{aligned} \quad (2)$$

Here Δ represents the difference between the final (subscript f) and initial (subscript 0) steady state. For example, $\Delta y_D = y_{Df} - y_{D0}$. The agreement between (2) and observed responses is very good in many cases. This is illustrated by Fig. 2A and 2B which show the response to small increases in reflux L (V constant) and boilup V (L constant) for column A. This column has 40 theoretical trays plus a condenser

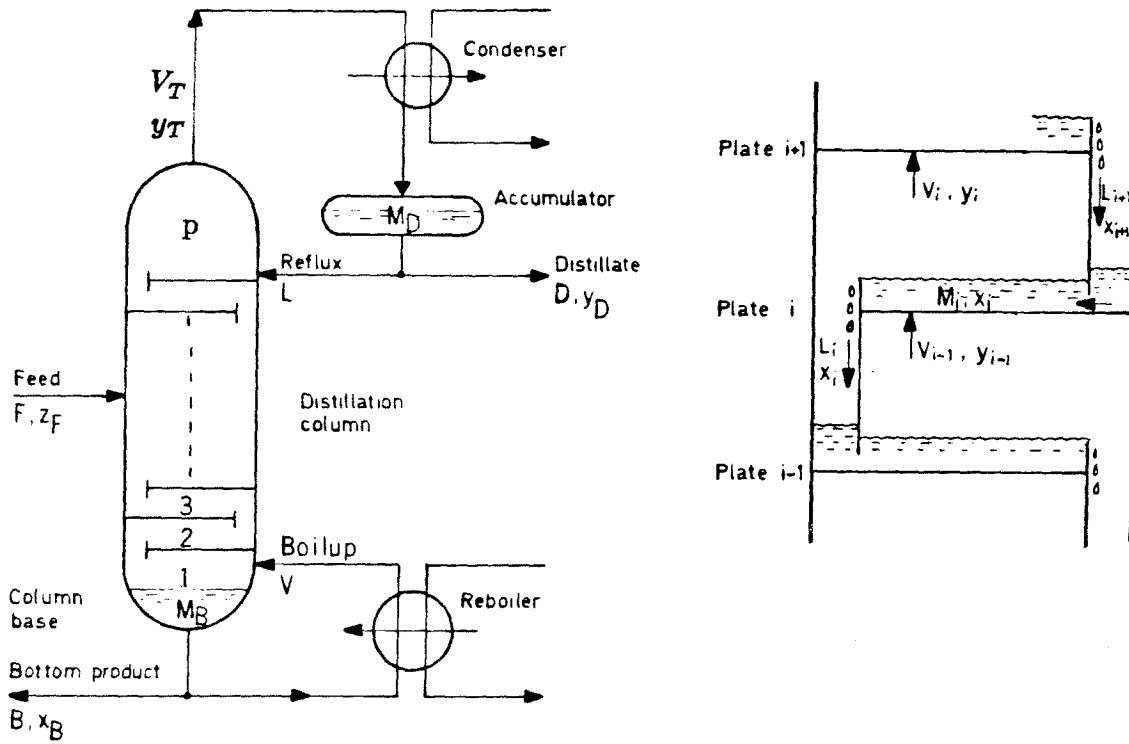


Figure 1. Two product distillation column with single feed and total condenser. Details are shown of the flows and holdups on a plate.

and the exact model is 41st order. This response is compared with a first order response with time constant 194 minutes corresponding to the linear model

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \frac{1}{1 + 194s} \begin{pmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (3)$$

The agreement is so good that the dotted line corresponding to this approximation is hardly visible. The value of the time constant (194 min) was found using (2) and in this case it is almost identical to the inverse of the smallest eigenvalue of the linearized model (see below).

For high-purity binary separations and small perturbations to the column (linear model valid) Skogestad and Morari (1987a) have derived an analytical expression for (2)

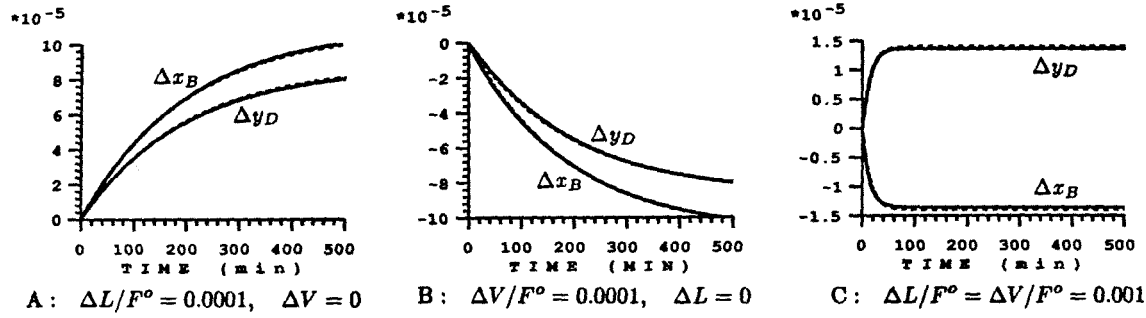


Figure 2. Column A. Responses to small changes in external (A & B) and internal (C) flows. Dotted lines for A & B: First order model (3) with time constant $\tau_1 = 194$ min. Dotted line for C: First order model (8) with time constant $\tau_2 = 15$ min. ($M_i/F = 0.5$ min)

$$\tau_{1c} \approx \frac{M_I}{\ln S \cdot I_s} + \frac{M_D y_D (1 - y_D)}{I_s} + \frac{M_B x_B (1 - x_B)}{I_s} \quad (4)$$

Here M_I is the total holdup inside the column, and M_D and M_B are the condenser and reboiler holdups. S is the separation factor

$$S = \frac{y_D (1 - x_B)}{(1 - y_D) x_B} \quad (5)$$

and I_s is the "impurity sum" defined by

$$I_s = D y_D (1 - y_D) + B x_B (1 - x_B) \quad (6)$$

The first term in (4), which represents the contribution from changing the holdup inside the column, dominates for columns with both products of high purity ($1 - y_D$ and x_B both small). Note that I_s may be extremely small in such cases resulting in very large values of τ_{1c} . This agrees with the observations of Fuentes and Luyben (1983). The reader is encouraged to study the paper by Skogestad and Morari (1987a) which discusses the use of (2) and (4) in detail.

It is clear from the derivation of (2) and (4) (Skogestad and Morari, 1987a) that τ_{1c} applies only to cases when there is a change in the total holdup $\Sigma M_i x_i$ of

some component in the column. Furthermore, from the total component material balance ($Fz_F = Dy_D + Bx_B$) we derive

$$D_f \Delta y_D + B_f \Delta x_B = \Delta(Fz_F) - y_{D0} \Delta D - x_{D0} \Delta B \quad (7)$$

From (7) we see that the denominator of (2) is non-zero only if there is a change in the $\Delta(Fz_F)$, ΔD or ΔB , that is, if there is a change in the external material balance. If we change the internal flows only (for example, increase the reflux L and the boilup V keeping the product flows B and D constant), then the numerator of (2) will be small, and the denominator will be identically zero. Consequently, (2) does not apply in such cases.

This is indeed confirmed by simulations. Fig. 2C shows the response to a simultaneous increase in L and V (D and B constant). The response is much faster than expected from the value $\tau_{1c} = 194$ min. In fact, an excellent fit is obtained using a time constant $\tau_2 = 15$ min, corresponding to the linear model

$$dL = dV : \quad \begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \frac{1}{1 + 15s} \begin{pmatrix} 0.014 \\ -0.014 \end{pmatrix} dV \quad (8)$$

(The gains 0.014 and -0.014 are derived from (3) using $dL=dV$.) Consequently, similar to what is known for the steady state (Skogestad and Morari, 1987b), there is a fundamental difference in column behavior for changes in external and internal flows. The objective of this paper is to study this in more detail, and to develop simple column models which display this behavior.

All results in this paper (gains, RGA-values, etc.) are for reflux L and boilup V as manipulated inputs. Distillate (D) and bottom flow (B) are manipulated to keep constant holdups in the accumulator/condenser (M_D) and the column base/reboiler (M_B). This does not imply that the LV-configuration is the preferred choice for control purposes. The choice is made because the column model is most naturally written in terms of L and V as manipulated inputs (see comments below).

Column	z_F	α	N	N_F	$1 - y_D$	x_B	D/F	L/F
<i>A</i>	0.5	1.5	40	21	0.01	0.01	0.500	2.706
<i>B</i>	0.1	1.5	40	21	0.01	0.01	0.092	2.329
<i>C</i>	0.5	1.5	40	21	0.10	0.002	0.555	2.737
<i>D</i>	0.65	1.12	110	39	0.005	0.10	0.614	11.862
<i>E</i>	0.2	5	15	5	0.0001	0.05	0.158	0.226
<i>F</i>	0.5	15	10	5	0.0001	0.0001	0.500	0.227
<i>G</i>	0.5	1.5	80	40	0.0001	0.0001	0.500	2.635

Table 1. Steady-state data for distillation column examples.

Column	G_{LV}^S	$\gamma(G_{LV}^S)$	$\lambda_{11}(G_{LV})$
<i>A</i>	$\begin{pmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{pmatrix}$	141.7	35.1
<i>B</i>	$\begin{pmatrix} 174.79 & -171.7 \\ 90.191 & -90.5 \end{pmatrix}$	229.2	47.5
<i>C</i>	$\begin{pmatrix} 16.023 & -16.0 \\ 9.29 & -10.7 \end{pmatrix}$	31.3	7.53
<i>D</i>	$\begin{pmatrix} 24.585 & -24.2 \\ 21.270 & -21.3 \end{pmatrix}$	234.9	58.7
<i>E</i>	$\begin{pmatrix} 203.4 & -131.5 \\ 22.47 & -22.5 \end{pmatrix}$	36.7	2.82
<i>F</i>	$\begin{pmatrix} 10740 & -10730 \\ 9257 & -9267 \end{pmatrix}$	2014	499
<i>G</i>	$\begin{pmatrix} 8648.94 & -8646 \\ 11347.06 & -11350 \end{pmatrix}$	6939	1673

Table 2. Steady-state values of the scaled gain matrix, the condition number and the 1,1-element of the RGA.

Assumptions. The seven columns studied previously by Skogestad and Morari (1987a,f) are used as examples. Steady-state data for these columns are given in Table 1. For all examples we assume constant molar flows, no flow dynamics (constant holdup), and binary mixtures with constant relative volatility. A nonlinear model based on these assumptions is given in Appendix. A linear (N+1)th order model (N theoretical trays plus the condenser) is obtained by linearizing the nonlinear model at the nominal operating point as shown in Appendix. This results in the following linear model

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G_{LV}(s) \begin{pmatrix} dL \\ dV \end{pmatrix}, \quad G_{LV}(s) = C(sI - A)^{-1}B \quad (9)$$

Based on this linear model we derive the steady-state gain matrices $G_{LV}(0) = -CA^{-1}B$ as given in Table 2. The poles of $G_{LV}(s)$ are equal to the eigenvalues of A. The time constants τ_{1e} and τ_{2e} corresponding to the two smallest eigenvalues of A are given in Table 3. The assumed holdup on the trays is $M_i/F = 0.5$ min in all examples (except for Table 3 and 4 and Fig. 4 which have $M_i/F = 1$ min). The holdups in the reboiler and condenser are also $M_i/F = 0.5$ min, unless stated otherwise.

2. A SIMPLIFIED DYNAMIC MODEL BASED ON INTERNAL AND EXTERNAL FLOWS

In this section we neglect the flow dynamics by assuming constant holdups. However, in order to allow us to easily extend our models to include flow dynamics it is preferable to write the dynamic model in terms of reflux (L) and boilup (V) as inputs. The reason is that L and V are the only flows which affect the product compositions immediately. The immediate effect of a change in one of the product flows (D and B) is to change the condenser or reboiler holdup which has no direct effect on composition. Therefore the response to changing D and B depends strongly on the tuning of the level loops.

The steady-state model using L and V as manipulated inputs is

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix}$$

In order to model explicitly the difference in dynamic behavior between internal and external flow changes we will consider V and D as manipulated inputs for the moment. Assuming constant molar flows we find

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} dV - dD \\ dV \end{pmatrix} = \begin{pmatrix} g_{11} + g_{12} & -g_{11} \\ g_{21} + g_{22} & -g_{21} \end{pmatrix} \begin{pmatrix} dV \\ dD \end{pmatrix} \quad (10a)$$

To get a dynamic model we make the following assumption which will be justified by examples and theoretical arguments later:

Modelling assumption. *The response to changes in the external flows (D in (10a)) is first-order with time constant τ_1 . The response to changes in internal flows (V in (10a)) is first-order with time constant τ_2 .*

With this assumption we derive the following dynamic model from (10a):

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} g_{11} + g_{12} & -g_{11} \\ g_{21} + g_{22} & -g_{21} \end{pmatrix} \begin{pmatrix} dV/(1 + \tau_2 s) \\ dD/(1 + \tau_1 s) \end{pmatrix} \quad (10b)$$

Switching back to L and V as manipulated inputs, and assuming constant holdups (perfect level control) and constant molar flows such that $dD(s) = dV(s) - dL(s)$, we finally derive

$$\begin{aligned} dy_D &= \frac{g_{11}}{1 + \tau_1 s} dL + \left(\frac{g_{11} + g_{12}}{1 + \tau_2 s} - \frac{g_{11}}{1 + \tau_1 s} \right) dV \\ dx_B &= \frac{g_{21}}{1 + \tau_1 s} dL + \left(\frac{g_{21} + g_{22}}{1 + \tau_2 s} - \frac{g_{21}}{1 + \tau_1 s} \right) dV \end{aligned} \quad (11)$$

This simple model is obviously not an accurate description of all distillation columns, but it is adequate for controller design. The model is best when the reboiler and condenser holdups are small as discussed in Section 9. The model's main advantage is its simplicity and that it gives a reasonable description of both the low- and high-frequency behavior.

τ_1 is generally close to the time constant corresponding to the smallest eigenvalue of A, and may be estimated as shown above (Eq. (2)). τ_2/τ_1 can be estimated by matching the high-frequency behavior as shown in Section 7. τ_1 is also simple to obtain from plant data or simulations. τ_2 may also be obtained from simulations (without flow dynamics) of changes in the internal flows (Fig. 2C). In most cases it will be very difficult to obtain τ_2 from plant data, since it is almost impossible, in practice, to carry out test runs for changes in the internal flows without changing the external flows (because of uncertainty and disturbances in feed rate, boilup, etc.). Also note from Fig. 2A and B that the small time constant (τ_2) is not detectable from the response to changes in reflux (L) and boilup (V). Both these responses are almost perfectly fitted by a first order response with time constant τ_1 .

Example. Column A. With the values $\tau_1 = 194$ min and $\tau_2 = 15$ min proposed in Section 1 (11) becomes

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \frac{1}{1 + 194s} \begin{pmatrix} 0.878 & -0.864 \frac{1+12.1s}{1+15s} \\ 1.082 & -1.096 \frac{1+17.3s}{1+15s} \end{pmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (12)$$

The agreement between this model and the exact 41st order model is excellent for small perturbations as seen from Fig. 2. The relative error $\bar{\sigma}((G - \tilde{G})\tilde{G}^{-1})$ (here $\bar{\sigma}$ is the maximum singular value) between the two time constant model (12) (denoted by \tilde{G}) and the full linear 41st order model (denoted by G) is shown as a function of frequency in Fig. 3. It is clear that \tilde{G} (12) is an excellent approximation of G up to about a frequency of 1 min^{-1} . On the other hand, the one time constant model (3) which has $\tau_1 = \tau_2 = 194$ min, gives a very poor approximation as seen from the dotted line in Fig. 3.

Note that without the seemingly negligible "correction terms" $\frac{1+12.1s}{1+15s}$ and $\frac{1+17.3s}{1+15s}$ the responses to changes in the internal flows would have a time constant of 194 min instead of the observed 15 min (Fig. 2C). In the literature each transfer matrix element in (11) is often approximated by a first-order lag with time delay

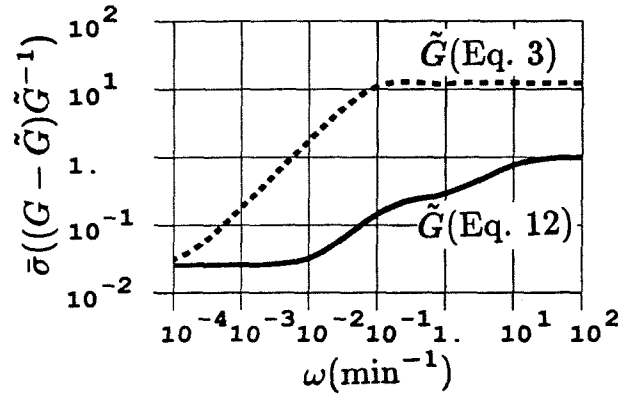


Figure 3. Column A. Relative difference between low order model \tilde{G} and 41st order plant G (9). The two time constant model (12) provides an excellent approximation, while the one time constant model (3) is poor at high frequency. ($M_i/F = 0.5$ min)

$(ge^{-\theta s}/(1+\tau s))$ where g is obtained by matching steady-state data. It is clear that, unless special care is taken, it is very unlikely that such a model will be able to capture the difference in time constants between external and internal flows.

It is important to note that if we consider each transfer function element independently, then an excellent fit is obtained with the one time constant model (3) (Fig. 2A and B). Consequently, each response by itself is essentially first order with time constant $\tau_1 = 194$ min. However, in order to capture the multivariable aspects of the model, for example, a simultaneous increase in L and V (change in internal flows), the two time constant model (12) is needed.

3. SINGULAR VALUES, THE CONDITION NUMBER AND THE RGA

In this section we introduce some tools which are commonly used when evaluating multivariable systems.

Singular Values. The maximum and minimum singular values are the maximum and minimum gain of the plant

$$\bar{\sigma}(G) = \max_{u \neq 0} \frac{\|Gu\|_2}{\|u\|_2} \equiv \|G\bar{v}\|_2, \quad \|\bar{v}\|_2 = 1 \quad (13a)$$

$$\underline{\sigma}(G) = \min_{u \neq 0} \frac{\|Gu\|_2}{\|u\|_2} \equiv \|G\underline{v}\|_2, \quad \|\underline{v}\|_2 = 1 \quad (13b)$$

(Both of these are functions of frequency, but this is not shown explicitly to simplify notation). $\|\cdot\|_2$ denotes the usual Euclidean norm. For distillation columns Skogestad and Morari (1986b) have shown that at steady state the most sensitive input direction \bar{v} (corresponding to the maximum gain) is obtained by changing the external flows, and the least sensitive input direction \underline{v} is obtained by changing the internal flows. Consequently, with L and V as manipulated inputs the most sensitive direction is found for $dL = -dV$ and the least sensitive direction is obtained for $dL = dV$. This also holds at higher frequencies as we will show below.

Condition Number. The condition number is defined as

$$\gamma(G(j\omega)) = \bar{\sigma}(G(j\omega)) / \underline{\sigma}(G(j\omega)) \quad (14)$$

It is possible to define $\gamma(G)$ in terms of other norms, but (14) is most common. A plant with a large value of $\gamma(G)$ is called ill-conditioned. Physically this means that the gain of the plant is strongly dependent on the input direction.

Scaling. The singular values and condition number are scaling dependent. For correct interpretation of these quantities the plant should be scaled such that outputs are of comparable magnitude and the inputs are of comparable magnitude. Let y_D^o and x_B^o denote the nominal product compositions. To get outputs of comparable magnitude scale each output with respect to the amount of impurity, that is,

$$y_D^S = \frac{y_D}{1 - y_D^o}, \quad x_B^S = \frac{x_B}{x_B^o} \quad (15)$$

The inputs are scaled with respect to feed flow rate (or equivalently set $F^o = 1$). In terms of scaled variables the model (9) becomes

$$\begin{pmatrix} dy_D^S \\ dx_B^S \end{pmatrix} = \begin{pmatrix} dy_D/1 - y_D^o \\ dx_B/x_B^o \end{pmatrix} = G_{LV}^S(s) \begin{pmatrix} dL/F^o \\ dV/F^o \end{pmatrix} \quad (16a)$$

where the scaled gain matrix is

$$G_{LV}^S(s) = \begin{pmatrix} \frac{F^o}{1-y_D^o} & 0 \\ 0 & \frac{F^o}{x_B^o} \end{pmatrix} G_{LV}(s) \quad (16b)$$

Note that the output scaling (15) is automatically obtained if we use logarithmic (relative) compositions since

$$d \ln(1 - y_D) = -\frac{dy_D}{1 - y_D} \quad \text{and} \quad d \ln x_B = \frac{dx_B}{x_B} \quad (17)$$

There are also other significant advantages of using logarithmic composition as we will discuss in detail later.

Minimum Condition Number. The minimum condition number is the value of $\gamma(G)$ obtained when minimizing over all possible input and output scalings

$$\gamma^*(G) = \min_{S_1, S_2} \gamma(S_1 G S_2) \quad (18)$$

Here S_1 and S_2 are diagonal matrices with real, positive entries. A large value of $\gamma^*(G)$ has been suggested to imply a plant which is fundamentally difficult to control (Grosdidier et al., 1985). However, obtaining $\gamma^*(G)$ numerically is not simple. Fortunately, there is a very close relationship between $\gamma^*(G)$ and the magnitude of the elements in the RGA. The RGA is defined below and is easy to compute. Thus, for practical computations, we recommend using the RGA rather than $\gamma^*(G)$.

RGA (Bristol, 1966). For 2×2 plants the RGA is defined as

$$RGA = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 1 - \lambda_{11} \\ 1 - \lambda_{11} & \lambda_{11} \end{bmatrix} \quad (19a)$$

where

$$\lambda_{11} = \frac{1}{1 - \kappa}, \quad \kappa(j\omega) = \frac{g_{12}g_{21}}{g_{11}g_{22}}(j\omega) \quad (19b)$$

Define the 1-norm of the RGA as the sum of the absolute values of the elements

$$\|RGA\|_1 = \sum_{i,j} |\lambda_{ij}| = 2|\lambda_{11}| + 2|1 - \lambda_{11}| \quad (20)$$

$\|RGA\|_1$ and $\gamma^*(G)$ are always close to each other in magnitude for 2×2 plants as seen from the following inequalities (Nett et al., 1986, Grosdidier et al., 1985)

$$\|RGA\|_1 - \frac{1}{\gamma^*(G)} \leq \gamma^*(G) \leq \|RGA\|_1 \quad (21)$$

The difference between $\|RGA\|_1$ and $\gamma^*(G)$ is at most one (since $\gamma^*(G) \geq 1$) and goes to zero as $\gamma^*(G) \rightarrow \infty$.

A plant with large RGA-elements is fundamentally difficult to control (Skogestad and Morari, 1986b): To have tight control of a multivariable plant it is desirable to use a controller which inverts the plant. However, such a controller is extremely sensitive to uncertainty with respect to the magnitude of the inputs if the RGA-elements are large. In particular, control problems are expected if the RGA has large elements around the crossover frequency (the frequency corresponding to the closed loop time constant). Consequently, the high-frequency behavior of the RGA is as important for evaluating potential control problems as the steady state value. We will show that the simple model (12) predicts correctly the value of the RGA at both low and high frequencies.

Importance of High-Frequency Behavior. For control purposes the behavior around crossover (the frequency corresponding to the inverse of the closed-loop time constant) is more important than the steady-state behavior. Consider the sensitivity functions S which is the closed-loop transfer function between disturbances (d) and reference signals (r) and the errors ($y - r$):

$$e = (y - r) = S(d - r), \quad S = (I + GC)^{-1} \quad (22)$$

If the controller has integral action ($C(0) \rightarrow \infty$), then perfect disturbance rejection is obtained at steady-state ($S(0) = 0$). However, the speed of the response is determined by the closed-loop bandwidth, which is approximately the frequency at which $\bar{\sigma}(S)$ reaches one. To improve the speed of the response (and obtain better performance) it is therefore important to have a good model of the plant in this frequency region. The steady-state behavior is not too important by itself, unless it reflects the plant behavior around crossover (which it usually does).

4. DISTILLATION COLUMN EXAMPLES

In this section we study the example columns in Table 1 and we will make some observations which will be explained in subsequent sections. Results obtained with the “full” linear model (9) will be compared with the simple two time constant model (11).

4.1 Singular Values and RGA as a Function of Frequency

It is well known that high-purity columns are ill-conditioned, at least at steady state (Skogestad and Morari, 1987a). The reason for the ill-conditioning is that the external flows have a much larger effect (“gain”) on the product compositions than the internal flows. This was illustrated in Section 1 for column A which is a high-purity column with $1 - y_D = x_B = 0.01$: At steady state the effect of changes in the external flows on y_D is $(\partial y_D / \partial D)_L = -0.864$, while the effect of changes in the internal flows is only $(\partial y_D / \partial L)_D = 0.014$.

What happens at higher frequencies ? In the example in Section 2 we found that the time constant for changes in the external flows (τ_1) was much larger than that for changes in the internal flows (τ_2). Furthermore, since the external and internal flows correspond to the most (\bar{v}) and least (\underline{v}) sensitive directions, respectively, we expect the condition number to be smaller at higher frequencies. This is indeed confirmed by the results presented below.

The singular values and $\|RGA\|_1$ for the seven columns are shown as a function of frequency in Fig. 4. The condition number $\gamma(G^S)$ is not shown, but its value is equal to the difference between the curves for $\bar{\sigma}(G^S)$ and $\underline{\sigma}(G^S)$ and it is close to $\|RGA\|_1$ for our examples. The most interesting fact to note from Fig. 4 is that the value of the condition number and the RGA is generally much lower at high frequencies. In particular, this is the case for columns A, B, F and G which have both products of high purity. We also note that the value of $\|RGA\|_1$ around $\omega = 0.1 \text{ min}^{-1}$ (which is probably close to the closed-loop bandwidth) does not depend very much on the purity of the products. Columns A, B, C and G all have $\|RGA\|_1 \approx 12$ at $\omega = 0.1$. (The reason, as shown in Section 7, is that L/F is similar for these columns.) On the other hand, the value of $\|RGA\|_1$ at steady-state is determined mainly by the purity of the least pure product: The steady-state value of $\|RGA\|_1$ for columns A, B, C and G vary in magnitude from 28 (column C) to 6690 (column G) depending on the purity of the least pure product. This agrees with the results of Skogestad and Morari (1987b) (also see (52) below).

The implication for control purposes is that an analysis based on steady-state data only can be misleading. For example, for column A $\|RGA\|_1$ is about 140 at steady state, which indicates a system which is fundamentally difficult to control. However, at high frequency it is only about 12, which seems acceptable. Thus, control of this column using L and V as manipulated inputs is simpler than might be expected from the steady-state data.

Steady-state analysis can be particularly misleading for columns with both products of high purity. The reason is that the steady-state value of the condition number (and $\|RGA\|_1$) is very large for such columns (Skogestad and Morari, 1987c). At the same time, the dominant time constant (τ_1) is also very large (recall comments following (6)). Since generally τ_2 does not depend very much on the product purities (see (58) below), it will be much smaller than τ_1 . This implies that

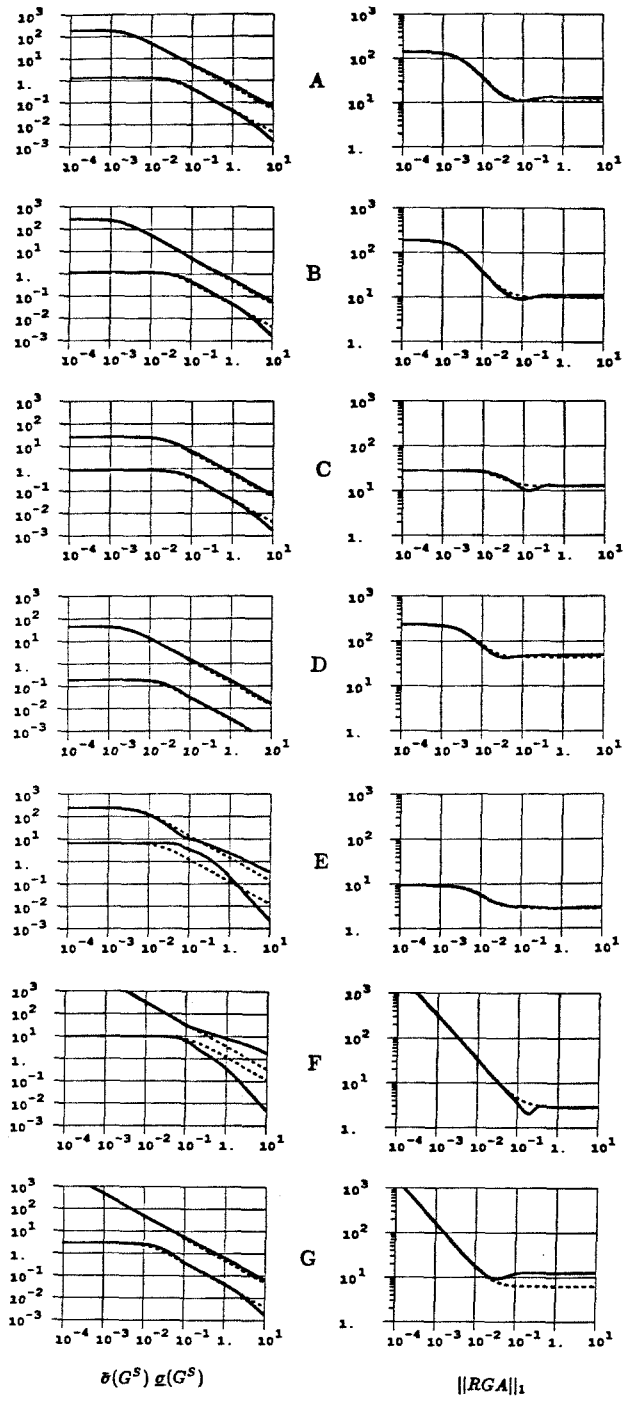


Figure 4. Singular values and $\|RGA\|_1$ as a function of frequency for column A-G. Dotted lines: Approximation using two time constant model (11) with data from Table 2 and 3. ($M_i/F = 1$ min)

$\gamma(G)$ and $\|RGA\|_1$ will be much higher at steady state than at high frequency for columns with both products of high purity.

This is confirmed by the results in Fig. 4 for columns A, B, F and G. On the other hand, note that the variation in $\|RGA\|_1$ with frequency is much less for columns C, D and E which do not have both products of high purity. From the RGA-values in Fig. 4, the only column which seems to be fundamentally difficult to control at all frequencies using the LV-configuration is column D, which has $\|RGA\|_1 \approx 50$ at high frequencies. Column D is similar to a C_3 -splitter and has 110 trays, $L/D = 19.3$, $1 - y_D = 0.005$ and $x_B = 0.10$. The large value of $\|RGA\|_1$ at steady state may lead to a sluggish return to steady-state also for columns A, B, F and G (Skogestad and Morari, 1987d).

4.2 Comparison with Simplified Model

The dotted lines in Fig. 4 are the values obtained with the simple two time constant model (11). The steady-state gain matrices needed for (11) are given in

Column	$\tau_1 = \tau_{1e}$	τ_{1c}	τ_2	τ_{2e}
A*	388	387	30	24
B	500	504	30	23
C	49	58	20	16
D	308	300	60	47
E	165	142	60	16
F	5992	5992	8	9
G	40667	40664	60	40

Table 3. Time constants τ_1 and τ_2 (in min) used in model (11) when computing dotted lines in Fig.4. For comparison are shown τ_{1c} estimated from (2) for a small perturbation in L (Skogestad and Morari, 1987a) and τ_{2e} (the inverse of the second smallest eigenvalue of A). All time constants are for $M_i/F = 1$ min.

*) Note that $\tau_1 = 194$ min and $\tau_2 = 15$ min are used for examples with $M_i/F = 0.5$ min.

Table 2 and the time constants τ_1 and τ_2 are given in Table 3. τ_1 was chosen to be equal to τ_{1e} (the inverse of the smallest eigenvalue). The value of τ_2 was adjusted by trial and error to get a good fit to the singular value and $\|RGA\|_1$ plots in Fig. 4. In many cases τ_2 is close to inverse of the second smallest eigenvalue (τ_{2e}).

We see from Fig.4 that the agreement between the full-order model and the simple two time constant model (11) is very good.

4.3 The Effect of Nonlinearity

For column A we know that the linear model (12) provides a very good representation of the full-order linear model (9). However, distillation columns are known to be strongly nonlinear (e.g., Moczek et al., 1963, Fuentes and Luyben, 1983). Is (12) still useful? We will return with a more complete discussion of these issues in Section 6. Here we show some examples which indicate that distillation columns are not as nonlinear as one would tend to believe – at least if we consider the initial response – and a linear model may be adequate. This is illustrated by Fig. 5 and 6 which show the response to small and large changes in the external (Fig. 5) and internal flows (Fig. 6): The initial responses $\frac{\Delta y_D}{\Delta L}$ and $\frac{\Delta x_B}{\Delta L}$ are almost independent of the magnitude of ΔL , although the steady-state behavior is entirely different. $\frac{\Delta y_D}{\Delta L}$ and $\frac{\Delta x_B}{\Delta L}$ are the responses to a unit change in ΔL , and we will call them the unit responses. Fig. 6 indicates that the initial unit response is independent of the magnitude of ΔL and ΔV .

However, are these initial unit responses also independent of operating conditions? Within a linear framework, one way of studying the effect of changing operating conditions, is to study how the linearized model changes with operating conditions. To this end consider column A and C. These actually represent the same column, but at two entirely different operating conditions. The product composition for column A are $1 - y_D = x_B = 0.01$. Column C is obtained by changing D/F from 0.500 to 0.555, which yields $1 - y_D = 0.10$ and $x_B = 0.002$. The steady-

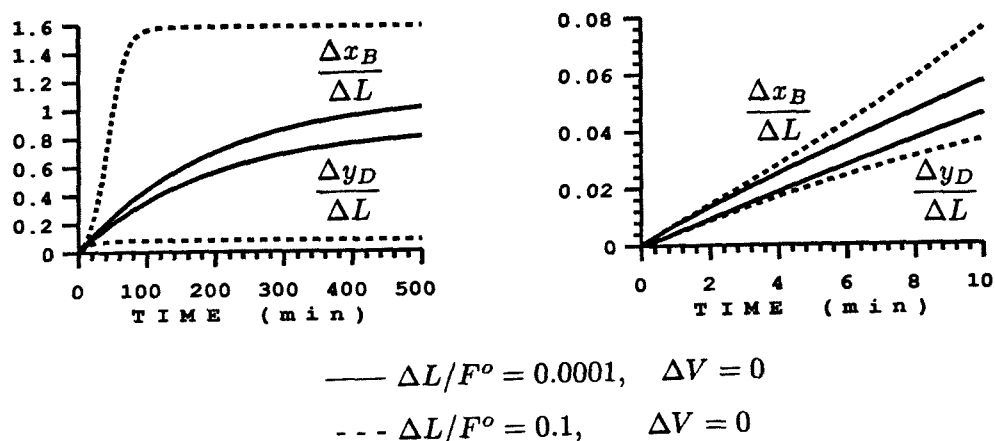


Figure 5. Column A. Unit responses to a small and large increase in reflux L . The initial unit response is almost independent of the magnitude of ΔL , but the steady-state behavior is entirely different. ($M_i/F = 0.5$ min)

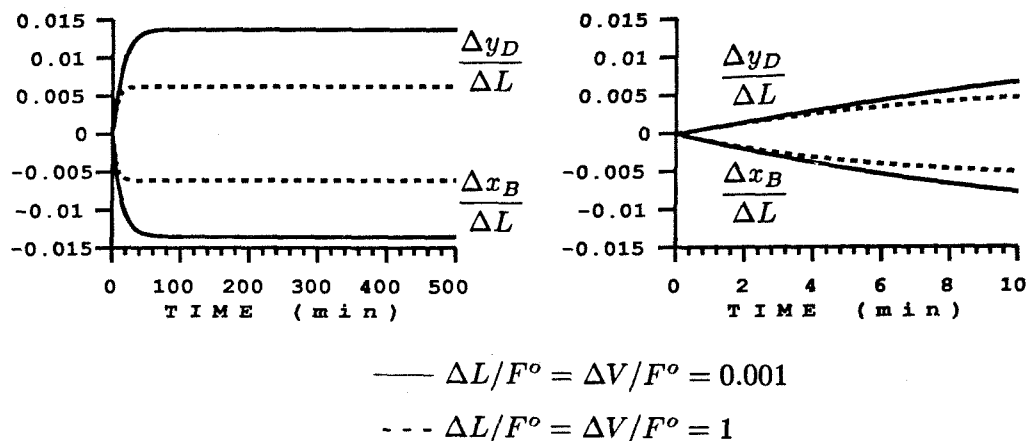


Figure 6. Column A. Unit responses to a small and large increase in internal flows. ($M_i/F = 0.5$ min)

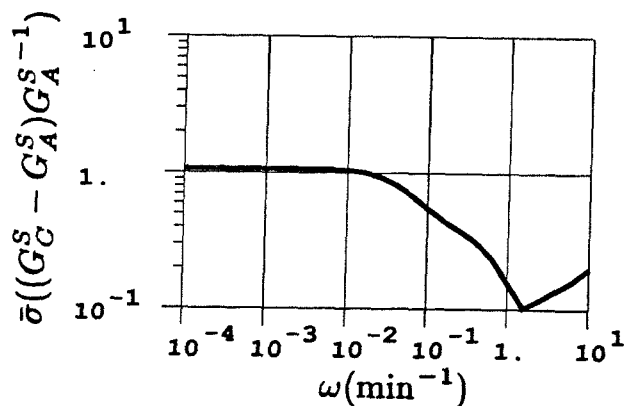


Figure 7. The scaled linear models for column A and C are entirely different at low frequency, but almost identical at higher frequencies. ($M_i/F = 0.5$ min)

state values of the scaled gains are drastically different for columns A and C. (The unscaled gains are even more different). The steady-state values of the condition number and the RGA (Table 2) and the time constants τ_{1e} and τ_{2e} (Table 3) are also very different. However, note from Fig. 4 that values of $\bar{\sigma}(G^S)$, $\underline{\sigma}(G^S)$ and $\|RGA\|_1$ are quite similar at high frequencies. This is also clear from Fig. 7 which shows the relative difference between the linearized scaled models for column A and C; the models are almost identical at higher frequencies. This implies that, even though the steady-state behavior is quite different, the initial response in terms of scaled (logarithmic) compositions is similar. In a separate study (Skogestad and Morari, 1987d) we show that a single linear controller, which is designed based on the linear model for column A, also yields good performance at the operating condition corresponding to column C. In these simulations $\ln(1 - y_D)$ and $\ln x_B$ were used as controlled outputs instead of y_D and x_B .

The objective of the remainder of the paper is to explain the observations made in this section.

5. HIGH-FREQUENCY BEHAVIOR (INITIAL RESPONSE)

The main approach to modelling the dynamic behavior of distillation columns in the literature (including our own work) has been to use a model which matches the steady-state behavior. In this case the model often represents the condition number and RGA well at low frequency, but poorly at high frequencies. The reasons for choosing this approach are i) that steady-state data are easily available from shortcut models or simulations and ii) that the steady-state behavior reflects the behavior at higher frequencies and is therefore useful for screening design alternatives for control purposes. However, for a more careful analysis the high-frequency behavior should also be considered. In fact, Rademaker et al.(1975, p. 137) suggest using a model which matches the high-frequency behavior (including the RGA), but which may be poor at steady-state. The preferred approach is clearly to use a

model which matches both the high- and low-frequency behavior. Deriving such a model was one of the objectives of this work and, as we will show, the simple two time constant model (11) does indeed satisfy this objective.

5.1 Slope of Initial Response

Assume constant molar flows and constant holdup. The component material balance for tray i at steady-state is

$$M_i \dot{x}_i = 0 = L_i(x_{i+1} - x_i) + V_i(y_{i-1} - y_i) \quad (23)$$

Assume a step change is made in L_i and V_i such that the flows for $t > 0$ are $L_i + \Delta L_i$ and $V_i + \Delta V_i$. Immediately following this change the values of the product compositions are unchanged. Thus we have for $t = 0^+$:

$$M_i \dot{x}_i = (L_i + \Delta L_i)(x_{i+1} - x_i) + (V_i + \Delta V_i)(y_{i-1} - y_i) \quad (24)$$

Subtracting the steady-state (23) yields (Rademaker et al., 1975, p.129)

$$(t \approx 0) : \quad M_i \dot{x}_i = (x_{i+1} - x_i)\Delta L_i + (y_{i-1} - y_i)\Delta V_i \quad (25)$$

\dot{x}_i given by this equation is equal to the initial slope of the response for Δx_i . Note that (25) is linear in ΔL_i and ΔV_i . This explains why the initial unit responses are independent of the magnitude of ΔL and ΔV as was observed in Fig. 5 and 6. (25) shows that the pole to zero excess is one at high frequency (neglecting dynamics involved in changing L and V). We know that for a column with N trays there are N poles if only the composition dynamics are considered. (25) shows that each transfer function has exactly $N - 1$ zeros (Kim and Friedly, 1974). In Section 6 we will justify that this N th order transfer function may be well approximated by a first order model.

5.2 The Effect of Operating Point on Initial Response

We used (25) to explain why

- 1) The initial unit response of Δx_i is independent of the magnitude of ΔL and ΔV . We now want to prove the claim made in Section 4.3:
- 2) The initial unit response of Δx_i (Δx_i may be Δy_D or Δx_B) is independent of the operating point if relative (logarithmic) compositions are used.

The steady-state relationship (23) may be used to rewrite the expression (25) for the initial slope of the response.

$$\dot{x}_i = \frac{x_{i+1} - x_i}{M_i} \left(\Delta L_i - \frac{L_i}{V_i} \Delta V_i \right) \quad (26)$$

Rewriting (26) gives

$$\frac{\dot{x}_i}{x_i} = \frac{d \ln x_i}{dt} = \frac{1}{M_i} \left(\frac{x_{i+1}}{x_i} - 1 \right) \left(\Delta L_i - \frac{L_i}{V_i} \Delta V_i \right) \quad (27a)$$

Alternatively

$$\frac{\dot{x}_i}{1 - x_i} = - \frac{d \ln(1 - x_i)}{dt} = \frac{1}{M_i} \left(1 - \frac{1 - x_{i+1}}{1 - x_i} \right) \left(\Delta L_i - \frac{L_i}{V_i} \Delta V_i \right) \quad (27b)$$

Consider a binary mixture and let x_i denote the mole fraction of light component. It is easily shown that the near the bottom of the column the ratio $\frac{x_{i+1}}{x_i}$ is 1) almost the same for any tray i and 2) only weakly dependent on operating conditions. Similarly, the ratio $\frac{1 - x_{i+1}}{1 - x_i}$ is nearly constant near the top of the column. Consequently, if the logarithm of the composition is used then the initial unit responses are nearly independent of operating conditions and the entire top or bottom of the column has almost the same response. On the other hand, the slope of the initial unit response is not independent of the operating point if composition are measured in terms of mole fractions.

To show that the ratio $\frac{x_{i+1}}{x_i}$ is nearly constant near the bottom of the column, assume that the equilibrium line operating lines are linear.

$$y_i = K_B x_i \quad (28)$$

$$x_{i+1} = \left(\frac{V}{L} \right)_B y_i + \frac{B}{L_B} x_B \quad (29)$$

These assumptions are reasonable for high-purity columns. V_B and L_B denote the vapor and liquid flows in the bottom of the column. Combining (28) and (29) yields

$$\frac{x_{i+1}}{x_i} = \left(\frac{V}{L}\right)_B K_B + \frac{B}{L_B} \frac{x_B}{x_i}$$

The second term is negligible as we go up the column and it is also small near the bottom for columns with $V/B > 1$. We get

$$\frac{x_{i+1}}{x_i} \approx \left(\frac{V}{L}\right)_B K_B \quad (30)$$

Thus $\frac{x_{i+1}}{x_i}$ is 1) independent of the tray location and 2) only weakly dependent on the operating point (since K_B and $(\frac{V}{L})_B$ are only weakly dependent on the operating point). Substituting (30) into (27a) yields

$$\text{Bottom part : } \quad \frac{\dot{x}_i}{x_i} \approx \frac{1}{M_i} \left(\left(\frac{V}{L}\right)_B K_B - 1 \right) \left(\Delta L_i - \left(\frac{L}{V}\right)_B \Delta V_i \right) \quad (31a)$$

A similar expression is derived for the top part where $(1 - y_i) \approx (1 - x_i)/K_T$

$$\text{Top part : } \quad \frac{\dot{x}_i}{1 - x_i} \approx \frac{1}{M_i} \left(1 - \frac{(V/L)_T}{K_T} \right) \left(\Delta L_i - \left(\frac{L}{V}\right)_T \Delta V_i \right) \quad (31b)$$

Multicomponent Mixtures. (31) was derived under the assumption of binary mixtures. It also holds for multicomponent mixtures if i) x_i in (31a) denotes the mole fraction of any light component (such that $x_B/x_i \ll 1$) and ii) \dot{x}_i and $1 - x_i$ in (31b) are replaced by \dot{x}_h and x_h where x_h denotes the mole fraction of any heavy component.

Range of Validity. From the derivation of (31) we see that the approximation is most likely to hold for high-purity column with large reflux. Note that for the case of constant relative volatility α we have $K_B = K_T = \alpha$. This will be used in the following.

Reboiler and Condenser. (31) does not apply to the reboiler and condenser. Assume perfect level control, total condenser and let y_T denote the vapor composition on

the top tray. (Note that $y_T = y_D$ at steady state. Then a derivation similar to that which led to (25) and (27) yields ($t = 0^+$)

$$\text{Reboiler : } \frac{\dot{x}_B}{x_B} = \frac{\alpha - 1}{M_B} \left(\left(\frac{V}{L} \right)_B \Delta L_B - \Delta V_B \right) \quad (32a)$$

$$\text{Condenser : } \dot{y}_D = \frac{y_T - y_D}{M_D} \Delta V_T = 0 \quad (32b)$$

For the case $M_B = M_i$ we find that $\frac{\dot{x}_B}{x_B}$ from (32a) is larger than for the other trays in the bottom (31a). On the other hand, from (32b) we find that \dot{y}_D is initially zero because a change in L and V has no direct effect on the composition in the condenser. However, note from Fig. 5B that the observed initial slope for y_D is not zero (at least not for $t > 0.1$ min). The reason for the discrepancy are the "second order effects" (changes in y_T) which were neglected when deriving (32b). For columns with small reboiler and condenser holdup, the following example shows that (31) may in fact also give a good approximation of the initial response in the reboiler and condenser. (However, if the holdup is large, then the response for y_D is second order and the initial slope is indeed zero as discussed in Section 9 (Fig. 8)).

Example. Column A. The slopes of the initial unit response obtained using (31a) and (31b) with $M_i/F = 0.5$ min are

$$\frac{\dot{x}_B}{\Delta L} = \frac{x_B}{M_i} \left(\left(\frac{V}{L} \right)_B \alpha - 1 \right) = \frac{0.01}{0.5} \left(\frac{3.21}{3.71} 1.5 - 1 \right) = 0.0060 \quad (33a)$$

$$\frac{\dot{y}_D}{\Delta L} = \frac{1 - y_D}{M_i} \left(1 - \frac{(V/L)_T}{\alpha} \right) = \frac{0.01}{0.5} \left(1 - \frac{3.21/2.71}{1.5} \right) = 0.0042 \quad (33b)$$

These are very close to the observed values in Fig. 5B.

Implications for control purposes. Equations (31) show that the initial response in terms of logarithmic (relative) compositions is independent of operating point. Skogestad and Morari (1987c) have shown that the steady-state gains also are much less affected by changes in operating points if logarithmic (rather than absolute) compositions are used. Therefore, the effect of nonlinearity is effectively

counteracted by using logarithmic compositions. The implication of these findings for control purposes is obviously that

$$Y_D = \ln(1 - y_D) \quad \text{and} \quad X_B = \ln x_B \quad (34)$$

should be used as controlled outputs if significant variations in product compositions are expected. This has also been suggested previously by Ryskamp (1981). Furthermore, this implies that the simple model (11) will also give a good description of the nonlinear column behavior if logarithmic (relative) compositions are used. To this end use (17) and rewrite (11) as follows:

$$\begin{aligned} -dY_D &= \frac{g_{11}^S}{1 + \tau_1 s} dL + \left(\frac{g_{11}^S + g_{12}^S}{1 + \tau_2 s} - \frac{g_{11}^S}{1 + \tau_1 s} \right) dV \\ dX_B &= \frac{g_{21}^S}{1 + \tau_1 s} dL + \left(\frac{g_{21}^S + g_{22}^S}{1 + \tau_2 s} - \frac{g_{21}^S}{1 + \tau_1 s} \right) dV \end{aligned} \quad (35)$$

Here $g_{11}^S = g_{11}/1 - y_D^o$, $g_{12}^S = g_{12}/1 - y_D^o$, $g_{21}^S = g_{21}/1 - x_B^o$, $g_{22}^S = g_{22}/1 - x_B^o$ are the scaled gains and g_{ij} , τ_1 and τ_2 are evaluated at the nominal operating point (y_D^o, x_B^o) .

6. VALIDITY OF FIRST-ORDER MODEL AT HIGH FREQUENCY

In the previous section we derived expressions (31) for the initial response. Is the simple two time constant model (11) consistent with these predictions? It is not obvious that this is the case since (11) was derived based on matching steady-state gains and low-frequency behavior (τ_1 and τ_2). Let us first consider the individual elements of the plant. Recall that each element in (11) is approximately first-order with time constant τ_1 (the effect of τ_2 on each individual element is usually negligible). If the initial response for each transfer matrix element predicted by (11) is indeed consistent with (31), then we may conclude that the time constant τ_1 also applies to the initial response. The initial slope derived from (11) is (let $s \rightarrow \infty$ and use $\mathcal{L}\{sdy_D\} = \dot{y}_D$):

$$\begin{pmatrix} \dot{y}_D \\ \dot{x}_B \end{pmatrix} = \begin{bmatrix} \frac{g_{11}}{\tau_1} & \frac{g_{11} + g_{12}}{\tau_2} - \frac{g_{11}}{\tau_1} \\ \frac{g_{21}}{\tau_1} & \frac{g_{21} + g_{22}}{\tau_2} - \frac{g_{21}}{\tau_1} \end{bmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (36)$$

As we will discuss in Section 9, the simple model (11) (and (36)) applies to the case with small reboiler and condenser holdups. In this case the predicted initial response obtained from (31) (recall comments following (32)) is

$$\begin{pmatrix} \dot{y}_D \\ \dot{x}_B \end{pmatrix} = \frac{1}{M_i} \begin{bmatrix} (1-y_D)(1-\frac{(V/L)\tau}{\alpha}) & -(1-y_D)(1-\frac{(V/L)\tau}{\alpha})(\frac{L}{V})_T \\ x_B((\frac{V}{L})_B\alpha-1) & -x_B((\frac{V}{L})_B\alpha-1)(\frac{L}{V})_B \end{bmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (37)$$

Example. Column A. Using (36) and (12) we derive

$$\begin{pmatrix} \dot{y}_D \\ \dot{x}_B \end{pmatrix} = \begin{pmatrix} 0.0045 & -0.0036 \\ 0.0056 & -0.0065 \end{pmatrix} \quad (38)$$

From (37) with $M_i/F = 0.5$ min

$$\begin{pmatrix} \dot{y}_D \\ \dot{x}_B \end{pmatrix} = \begin{pmatrix} 0.0042 & -0.0036 \\ 0.0060 & -0.0069 \end{pmatrix} \quad (39)$$

The consistency between (36) and (37) is excellent for this particular example. The objective of the following derivation is to show that is the case also in general – at least for high-purity columns with L/V close to one.

General derivation of consistency between (36) and (37)

Consider the 1,1-element $\frac{g_{11}}{\tau_1}$ in (36). The objective is to show that it is close in magnitude to $\frac{1-y_D}{M_i}(1-\frac{(V/L)\tau}{\alpha})$ in (37). (Similar derivations hold for other elements). Assume (i) constant molar flows, (ii) constant relative volatility, (iii) negligible condenser and reboiler holdup, (iv) equal holdups M_i on all trays, (v) large reflux (L/V close to one) and (vi) that the effect of changes in the separation factor $S = \frac{y_D(1-x_B)}{(1-y_D)x_B}$ can be neglected when evaluating the steady-state gains. Note that assumption (iii) is needed anyway in order for (36) and (37) to be valid. In practice, assumption (v) is the same as assuming α close to one (see (55) below). Assumption (vi) is most likely to hold for high-purity columns. Using assumptions (i) and (vi) we derive (Skogestad and Morari, 1987b)

$$g_{11} = \left(\frac{\partial y_D}{\partial L} \right)_V = \frac{(1-y_D)y_D(y_D-x_B)}{I_s} \approx \frac{1-y_d}{I_s} \quad (40)$$

where I_s is defined by (6). Similarly, assumptions (i)-(iv) yield (see (4) above) (Skogestad and Morari, 1987a)

$$\tau_1 = \frac{NM_i}{I_s \ln S} \quad (41)$$

Combining (40) and (41)

$$\frac{g_{11}}{\tau_1} = \frac{(1 - y_D) \ln S}{NM_i} \quad (42)$$

A good approximation for $\ln S$ when assumptions (i) and (ii) hold is (Skogestad and Morari, 1987f)

$$\ln S = N \ln \alpha + \frac{N}{2} \ln \frac{(L/V)_T}{(L/V)_B} \quad (43)$$

Assume that $(L/V)_T \approx (V/L)_B$ are close to one (Assumption v). Then

$$\frac{\ln S}{N} = \ln \alpha - \ln(V/L)_T \approx 1 + \alpha - 1 - (V/L)_T = \alpha - (V/L)_T$$

Substituting this into (40) yields

$$\frac{g_{11}}{\tau_1} = \frac{(1 - y_D)}{M_i} \left(\alpha - \left(\frac{V}{L} \right)_T \right) = \alpha \frac{1 - y_D}{M_i} \left(1 - \frac{(V/L)_T}{\alpha} \right) \quad (44)$$

Except for the factor α , which is close to one, this is identical to the 1,1-element of (37).

We should also note from this derivation the strong correlation between the scaled steady-state gain and the dominant time constant (see (42) above)

$$\frac{g^S}{\tau_1} \approx \frac{\ln S}{\sum_i M_i} \quad (45)$$

It is this correlation which makes the initial response in terms of scaled (relative) compositions nearly independent of operating conditions.

An alternative way of estimating τ_1 . Finally, note that if we assume that the entire response is approximated well by a first order model, then the expression (37) for the initial slope can be used to estimate τ_1 (rather than using (2)). This approach

is suggested by Rademaker et al. (1975, p. 137). For example, matching \dot{y}_D/dL in (36) and (37) we derive

$$\frac{g_{11}}{\tau_1} = \frac{1 - y_D}{M_i} \left(1 - \frac{(V/L)\tau}{\alpha} \right)$$

In terms of scaled gains ($g_{11}^S = g_{11}/1 - y_D$)

$$\tau_{1i} = \frac{g_{11}^S M_i}{1 - \frac{(V/L)\tau}{\alpha}} \quad (46)$$

(The subscript i on τ_1 denotes that it is derived by matching the initial responses). Other relationships result if we try to match other elements, but the resulting τ_{1i} is about the same. Because condenser holdup was neglected when deriving (37), M_i in (46) represents the holdup on a plate in the column. τ_{1i} computed from (46) is compared to τ_{1e} (the dominant time constant corresponding to the largest eigenvalue) in Table 4. As expected τ_{1i} and τ_{1e} are very similar for columns with high-purity and/or large reflux. For these columns (A, B, D and G) we conclude that the composition response is indeed well approximated by a first order model. For cases with $\tau_{1e} > \tau_{1i}$ the initial response is somewhat slower than predicted by τ_{1e} , indicating a higher-order initial response. However, the difference between $\tau_{1i} > \tau_{1e}$ is at the most a factor of two for the seven columns, and the assumption regarding first-order model is justified.

In this section we have shown that the magnitude of each individual element in (36) is similar to that derived from the initial response (37). We used this to conclude that each transfer function element in the plant is approximately first order with time constant τ_1 . However, for multivariable control the relative magnitude of the elements, for example, as expressed by the RGA, is important. In Section 7 we use (37) to estimate the high-frequency RGA-values. These values should be consistent with those obtained from the two time constant model (36), and this may be achieved by adjusting the value of τ_2 in (11). This is discussed in Section 8.

Column	τ_{1e}	τ_{1i} <i>Eq.(46)</i>
<i>A</i>	388	412
<i>B</i>	500	569
<i>C</i>	49	81
<i>D</i>	308	405
<i>E</i>	165	308
<i>F</i>	5992	13655
<i>G</i>	40667	41715

Table 4. Time constants (min) τ_{1e} corresponding to largest eigenvalue and τ_{1i} obtained by matching initial response (46). $\tau_{1i} > \tau_{1e}$ indicates that the initial response is higher than first order ($M_i/F = 1$ min).

7. ILL-CONDITIONING AT HIGH FREQUENCY

7.1 Most Sensitive Direction

The high-frequency behavior estimated from the initial responses (37) is

$$M_i \begin{pmatrix} \frac{dy_D}{1-y_D} \\ \frac{dx_B}{x_B} \end{pmatrix} = \frac{1}{s} \begin{bmatrix} k_T & -k_T \left(\frac{L}{V}\right)_T \\ k_B & -k_B \left(\frac{L}{V}\right)_B \end{bmatrix} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (47)$$

with $k_T = 1 - \frac{(V/L)_T}{\alpha}$ and $k_B = \left(\frac{V}{L}\right)_B \alpha - 1$. Recall that at steady state the maximum gain is found when $dL = -dV$ corresponding to changing the external flows, and the minimum gain is found when $dL = dV$ corresponding to changing the internal flows. The same also holds at high frequency for most columns as is shown next.

We will use (47) to show that at high frequency $dL \approx dV$ defines a direction corresponding closely to the minimum gain. (This will prove that $dL = -dV$ corresponds to the maximum gain since $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are orthogonal). From (47) we see that the minimum effect on y_D (zero) is obtained when $dL/dV = \left(\frac{L}{V}\right)_T$. Similarly, the minimum effect on x_B (zero) is found when $dL/dV = \left(\frac{L}{V}\right)_B$. Note that $\left(\frac{L}{V}\right)_T < 1$ and $\left(\frac{L}{V}\right)_B > 1$. Consequently, the minimum gain (effect on y_D and x_B) is obtained for $\left(\frac{L}{V}\right)_T < \frac{dL}{dV} < \left(\frac{L}{V}\right)_B$. For columns with large reflux this clearly gives $dL/dV \approx 1$. Also, note that if the feed is liquid, then $\left(\frac{L}{V}\right)_B = \left(\frac{L}{V}\right)_T + \frac{F}{V}$. Thus, with the possible exception for columns with small values of V/F , we will always find the minimum gain for $dL \approx dV$. Since the simple model (11) yields the maximum gain for $dL = dV$ at all frequencies, this provides a further justification for using (11).

7.2 Estimate of the RGA at High Frequency

We want to estimate $\lambda_{11} = (1 - \kappa)^{-1}$ at high frequency. Note that

$$\kappa(j\omega) = \frac{g_{21}/g_{22}}{g_{11}/g_{12}}(j\omega)$$

Column	$\ RGA(\infty)\ _1$ observed	$4\frac{L}{F} + 2$ Eq.(51)
<i>A</i>	12.83	12.82
<i>B</i>	11.32	11.32
<i>C</i>	12.95	12.94
<i>D</i>	49.10	49.44
<i>E</i>	2.90	2.90
<i>F</i>	2.91	2.91
<i>G</i>	12.54	12.54

Table 5. Observed (using “full” linear model (9)) and estimated values of $\|RGA\|_1$ at high frequency ($\omega = \infty$). The agreement is amazing.

where

$$\frac{g_{21}}{g_{22}} = \frac{(\partial x_B / \partial L)_V}{(\partial x_B / \partial V)_L} \quad \text{and} \quad \frac{g_{11}}{g_{12}} = \frac{(\partial y_D / \partial L)_V}{(\partial y_D / \partial V)_L} \quad (48)$$

At high frequency these ratios are given by the ratio between the slopes of the initial response of x_B (and y_D) to changes in L and V . From (31a) and (31b) we get (these apply to the entire bottom and top part of the column)

$$\frac{g_{21}}{g_{22}}(\infty) = - \left(\frac{V}{L} \right)_B \quad \text{and} \quad \frac{g_{11}}{g_{12}}(\infty) = - \left(\frac{V}{L} \right)_T$$

and we derive

$$\kappa(\infty) = \frac{(L/V)_T}{(L/V)_B} \quad (49)$$

This derivation does not depend on the amount of holdup, and therefore we expect $\lambda_{11}(\infty)$ to be almost independent of condenser and reboiler holdup. For the case of constant molar flows and feed as liquid ($L_B = L_T + F$, $V_T = V_B$) we find $\kappa(\infty) = L_T/L_B$ and the RGA becomes

$$\text{Feed liquid :} \quad \lambda_{11}(\infty) = \frac{1}{1 - \kappa(\infty)} = 1 + \frac{L}{F} \quad (50)$$

which gives

$$\|RGA(\infty)\|_1 = 4\frac{L}{F} + 2 \quad (51)$$

For the seven examples the agreement between the RGA-values estimated from (51) and those obtained from the full linearized model is amazing (Table 5).

7.3 Variation of RGA with Frequency

From (50) we conclude that large reflux ($L/F \gg 1$) is necessary for the column to have large RGA-values at high frequency. The RGA-values at low frequency are mainly determined by the purity of the last pure product. This follows from the following approximation for $\lambda_{11}(0)$ for the LV- configuration which applies to columns with both products of high purity (Skogestad and Morari, 1987b)

$$\lambda_{11}(0) = \frac{1}{I_s} \frac{1}{(\partial \ln S / \partial L)_D} \quad (52)$$

Here $I_s = Bx_B(1 - x_B) + Dy_D(1 - y_D)$ depends strongly on operating conditions and its value is determined by the purity of the least pure product. $I_s \rightarrow 0$ and $\lambda_{11}(0) \rightarrow \infty$ when both products are of high purity. $(\partial \ln S / \partial L)_D$ shows less variation with operating conditions as seen from the following approximation (Skogestad and Morari, 1987f)

$$\text{Feed liquid :} \quad \left(\frac{\partial \ln S}{\partial L} \right)_D \approx \frac{N}{2} \frac{1}{L} \frac{1}{L/F + 1} \quad (53)$$

Is it possible to have larger RGA-values at low than at high frequencies? This would require $\tau_2 > \tau_1$ in the model (11), and thus does not seem likely. A simplified analysis yields the same conclusion. Substituting (53) into (52) we derive

$$\lambda_{11}(0) \approx \frac{2L}{NI_s} \left(1 + \frac{L}{F} \right) \quad (54)$$

Recall that $\lambda_{11}(\infty) \approx 1 + \frac{L}{F}$ when the feed is liquid. The question is therefore whether it is possible to have the factor multiplying $(1 + \frac{L}{F})$ in (54) smaller than one. This does not seem likely. At least we must require $\frac{L}{F} > (\frac{L}{F})_{min}$. For columns with constant relatively volatility and feed liquid the minimum reflux ratio is approximately (e.g., Skogestad and Morari, 1987f) $(\frac{L}{D})_{min} \approx \frac{1}{\alpha - 1} \frac{1}{z_F}$. For most columns $D \approx Fz_F$ and we derive

$$\left(\frac{L}{F}\right)_{min} \approx \frac{1}{\alpha - 1} \quad (55)$$

Furthermore, for a column which is well designed $N/N_{min} \approx 2$ (Henley and Seider, p. 453) where $N_{min} = \ln S / \ln \alpha$. The factor multiplying $(1 + \frac{L}{F})$ in (54) is then bounded below as follows:

$$\frac{2L}{NI_s} > \frac{\ln \alpha}{\ln S(\alpha - 1)I_s/F} \approx \frac{1}{\ln S \cdot I_s/F}$$

The product $\ln S \cdot I_s/F$ is less than one for most columns, except for cases where the purities of the products are completely different. However, for such columns the approximation (52) is poor (gives too small values of $\lambda_{11}(0)$) and our analysis is not valid. In fact, we have not been able to find any examples of columns which exhibit $\lambda_{11}(0) < \lambda_{11}(\infty)$. †

7.4 Implications for Control Purposes

For control purposes, it is clearly bad if the RGA-values are large either at low or high frequencies. We conclude from (50) and (54) that the worst columns to control using the LV-configuration are those with large reflux and with both products of high purity. Most well-designed columns have a reflux L/F which is about 5 to 50% higher than $(L/F)_{min} \approx 1/\alpha - 1$ (55). This means that in industrial columns (which hopefully are reasonably optimal), large reflux flows (and large RGA-values at high frequency) are observed only for columns with relative volatility α close to one. Furthermore most well-designed columns have $N \approx 2N_{min}$ and

† The column we have found which have $\lambda_{11}(0)$ and $\lambda_{11}(\infty)$ closest is the following: $z_F = 0.65, y_D = 0.9, x_B = 0.002, \alpha = 1.12, N = 110, N_F = 39$. This yields $L/F = 49.6, \lambda_{11}(0) = 57.7$ and $\lambda_{11}(\infty) = 1 + \frac{L}{F} = 50.6$.

because $N_{min} = \ln S / \ln \alpha$ such separations usually require a large number of trays (e.g., column D).

To observe large RGA-values throughout the frequency range in a university column, which usually has few trays, one should use systems with high relative volatility (to get high-purity products) and operate at high reflux.

8. ESTIMATION OF τ_2

The ability to estimate the RGA at high frequency (Eq. (49)) suggests that τ_2 may be estimated by matching the RGA-value at high frequency. $\kappa(\infty)$ predicted by the two-time constant model (11) is

$$\kappa(\infty) = \frac{\frac{g_{21}}{\tau_1} \left(\frac{g_{11}+g_{12}}{\tau_2} - \frac{g_{11}}{\tau_1} \right)}{\frac{g_{11}}{\tau_1} \left(\frac{g_{21}+g_{22}}{\tau_2} - \frac{g_{21}}{\tau_1} \right)} = \frac{\left(1 + \frac{g_{12}}{g_{11}} \right) - \frac{\tau_2}{\tau_1}}{\left(1 + \frac{g_{22}}{g_{21}} \right) - \frac{\tau_2}{\tau_1}} \quad (56)$$

The ratio τ_2/τ_1 may be estimated by equating (56) and (49). For the case of constant molar flows and feed liquid (49) gives $\kappa(\infty) = L/(L+F)$ and we derive

$$\text{Feed liquid : } \left(\frac{\tau_2}{\tau_1} \right)_i = \frac{L}{F} \left(\frac{g_{12}}{g_{11}} - \frac{g_{22}}{g_{21}} \right) + \left(1 + \frac{g_{12}}{g_{11}} \right) \quad (57)$$

(subscript i denotes initial). This ratio is shown in Table 6 for our seven examples and is compared with the actual value used in Fig. 4. The agreement is obviously good since τ_2 for Fig. 4 was derived partially based on matching the RGA-values.

An analytical formula for τ_2 is obtained by using the following approximations which apply to columns with both products of high purity: (i) Gains (Skogestad and Morari, 1987f)

$$\frac{g_{12}}{g_{11}} \approx 1 - Bx_B(1-x_B) \left(\frac{\partial \ln S}{\partial L} \right)_D, \quad \frac{g_{22}}{g_{21}} \approx -1 + Dy_D(1-y_D) \left(\frac{\partial \ln S}{\partial L} \right)_D$$

Column	τ_1/τ_2 (Table 3)	$(\tau_1/\tau_2)_i$ (Eq.(57))
<i>A</i>	0.077	0.092
<i>B</i>	0.060	0.069
<i>C</i>	0.408	0.420
<i>D</i>	0.195	0.218
<i>E</i>	0.364	0.434
<i>F</i>	.00133	.00139
<i>G</i>	.00147	.00192

Table 6. Ratio τ_1/τ_2 used in examples (Fig.4, Table 3) and estimate (57) derived by matching the high-frequency RGA.

(ii) $(\frac{\partial \ln S}{\partial L})_D$ from (52) and (iii) $\tau_c \approx \frac{M_I}{\ln S I_s}$ from (4). We derive

$$\text{Feed liquid : } \tau_{2i} \approx \frac{N}{2 \ln S} \frac{M_I/F}{\frac{L}{F} + 1} \left(1 + \frac{F/L}{1 + \frac{D y_D (1-y_D)}{B x_B (1-x_B)}} \right) \quad (58)$$

Note that $\frac{N}{2 \ln S} \approx 1/\ln \alpha$ because most columns have $N/N_{min} \approx 2$. (58) demonstrates that τ_2 is only weakly dependent on operating conditions. The agreement between (58) and the values for τ_2 in Table 3 is good. However, for practical calculations we recommend using (57) to estimate τ_2 , rather than (58), in order to get the correct high-frequency behavior.

9. THE EFFECT OF REBOILER AND CONDENSER HOLDUP

All examples considered so far have had negligible reboiler and condenser holdup. However, the responses may depend strongly on the amount of holdup as seen from Fig. 8: Curve 1 shows the initial response in y_D to a small change in reflux L for a column with $M_i/F = 0.5$ min on all trays. This response is closely approximated by (3): $dy_D = \frac{0.878}{194s+1} dL$. Curve 2 shows the response with the same

holdup inside the column, but with M_D/F increased from 0.5 to 32.1 min. The initial response is significantly more sluggish than estimated from (3). To derive a better approximation we argue as follows: The composition in the condenser (y_D) is not directly affected by the reflux (L), but only by the composition y_T in the overhead vapor (V_T). (Note that $y_T = y_D$ at steady-state because of the assumption of total condenser). The effect of y_T on y_D is given by a first order response with time constant $\tau_D = M_D/V_T = 10$ min. The following approximation to curve 2 is then derived (shown as curve 3)

$$dy_D = \frac{0.878}{(10s + 1)(194s + 1)} dL \quad (59)$$

Fig. 8 shows that this is a much better approximation than (3).

The value of the RGA is hardly affected at all by adding the reboiler and condenser holdups (Fig. 9). This confirms the comments made following Eq. (49). On the other hand, the singular values do change significantly. Again, from the dotted line in Fig. 9 we see that (59) provides a reasonable way of taking into account the effect of condenser holdup.

Approximate model with reboiler and condenser holdup. To include the real case with reboiler and condenser holdup we propose to simply "add on" two lags for the condenser and reboiler to the model (11) (or (35) if logarithmic compositions are used).

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = \begin{pmatrix} \frac{1}{\tau_D s + 1} & 0 \\ 0 & \frac{1}{\tau_B s + 1} \end{pmatrix} G_{(Eq.(11))} \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (60)$$

where $\tau_D = M_D/V_T$ and $\tau_B = M_B/L_B$. In most cases τ_D is significantly larger than τ_B .

Measuring overhead composition. In practice, composition is often not measured in the condenser (y_D), but rather in the overhead vapor (y_T). Furthermore, temperatures are often used to infer compositions, and these temperatures will have a dynamic response similar to y_T , rather than y_D . Curve 4 on Fig. 8 shows

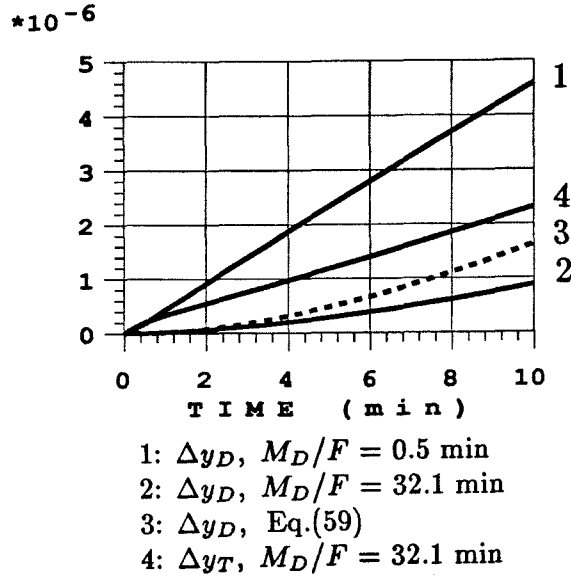


Figure 8. Column A. Effect of condenser holdup M_D on initial response of y_D to a small increase in reflux ($\Delta L/F^o = 0.0001$).

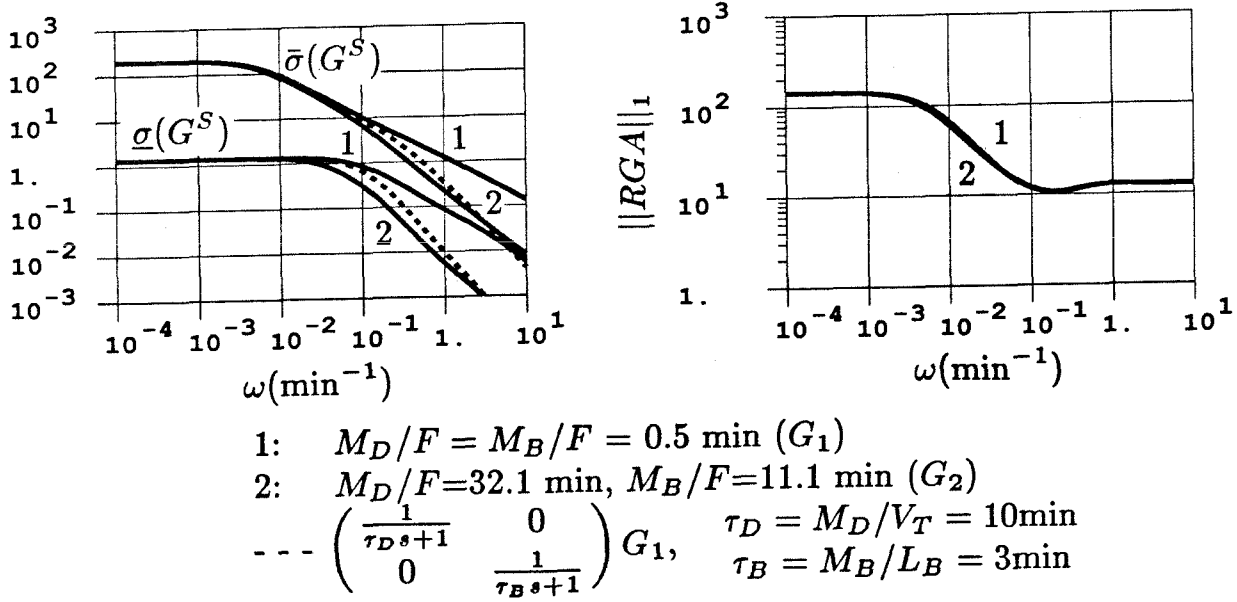


Figure 9. Column A. Effect of reboiler and condenser holdup on singular values and RGA.

that the initial response of y_T is much faster than that of y_D for the case with large condenser holdup. The same arguments also apply to the bottom section: Temperature measurements which are used to infer compositions have a dynamic response which in many cases may be approximated well using (11). Consequently, often in practice, the simplified model (11), which neglects the condenser holdup, may be satisfactory for controller design. This is indeed confirmed by a simulation study by Skogestad and Morari (1987d)

10. SIMPLIFIED MODEL INCLUDING FLOW DYNAMICS

So far we have neglected the flow dynamics. However, the flow responses are very important for the initial dynamic response. These issues are discussed by Rademaker et al. (1975). If we make the simplifying assumption that the flow dynamics are essentially decoupled from the composition dynamics, we derive from (35) the following simplified column model when flow dynamics are included

$$-dY_D = \frac{g_{11}^S}{1 + \tau_1 s} dL_T + \left(\frac{g_{11}^S + g_{12}^S}{1 + \tau_2 s} - \frac{g_{11}^S}{1 + \tau_1 s} \right) dV_T \quad (61a)$$

$$dX_B = \frac{g_{21}^S}{1 + \tau_1 s} dL_B + \left(\frac{g_{21}^S + g_{22}^S}{1 + \tau_2 s} - \frac{g_{21}^S}{1 + \tau_1 s} \right) dV_B \quad (61b)$$

Comments on (61):

- Logarithmic compositions $Y_D = \ln(1 - y_D)$ and $X_B = \ln x_B$ are used to reduce the effect of nonlinearity. Of course, if only small variations in the product compositions are expected we may as well use the unscaled composition y_D and x_B . For multicomponent mixtures $Y_D = \ln y_{DH}$ and $X_B = \ln x_{BL}$ should be used. Here H and L denote the heavy and light key components.
- g_{ij}^S in (61) represents the scaled steady-state gains for the LV- configuration. For example, $g_{11}^S = g_{11}/1 - y_D^o$, where y_D^o is the composition for the nominal operating point used to evaluate g_{11} .
- τ_1 and τ_2 should be evaluated at the same nominal operating points as the

steady-state gains. τ_1 may be estimated using (2) or (46); τ_2/τ_1 may be estimated using (57).

- Extra lags $1/(1 + \tau_D s)$ and $1/(1 + \tau_B s)$ should be added to the response in y_D and x_B if composition is measured in the condenser and reboiler (see Eq. (60). Here $\tau_D = M_D/V_T$ and $\tau_B = M_B/L_B$.
- Measurement dynamics are not included in (61).
- L_T, V_T, L_B and V_B are the "local" liquid and vapor flows at the top and bottom of the column. Let L and V denote the manipulated values of the reflux and the boilup. Then the following approximations, which apply to case of constant molar flows, are useful (Rademaker et al. (1975)).

$$dL_T = dL, \quad dV_B = dV \quad (62a)$$

$$dV_T = dV \quad (\text{assuming perfect pressure control}) \quad (62b)$$

$$dL_B \approx e^{-\theta_L s} dL + \lambda(1 - e^{-\theta_L s}) dV \quad (62c)$$

Here

$$\lambda = (\partial L_i / \partial V)_{M_i} \quad (63a)$$

$$\theta_L = N_{tr} \tau_L, \quad \tau_L = (\partial M_i / \partial L_i)_V \quad (63b)$$

and N_{tr} is the number of physical trays in the column.

- Additional dynamics (valve dynamics, etc.) may be needed in order to physically change L and V .

Eq. (62c) is derived by repeated combination of the following two equations for each tray (Rademaker et al., 1975)

$$dL_i = \left(\frac{\partial L_i}{\partial V} \right)_{M_i} dV + \left(\frac{\partial L_i}{\partial M_i} \right)_V dM_i = \lambda dV + \frac{1}{\tau_L} dM_i \quad (64a)$$

$$\frac{dM_i}{dt} = dL_{i+1} - dL_i \quad (64b)$$

and using the approximation

$$e^{-\theta_L s} \approx \frac{1}{(1 + \tau_L s)^{N+1}} \quad (64c)$$

The values for θ_L and λ should preferably be determined experimentally: θ_L is the time it takes for an increase in reflux to affect the reboiler level; λ is obtained by observing the response in reboiler level to a change in boilup.

For columns with trays θ_L can be estimated as follows: Let M_{oi} represent the amount of liquid on each tray over the weir. According to Francis weir formula $L_i = kM_{oi}^{3/2}$ and we derive

$$\tau_L = \left(\frac{\partial M_i}{\partial L_i} \right)_V = \left(\frac{\partial M_{oi}}{\partial L_i} \right)_V = \frac{2}{3} \frac{M_{oi}}{L} \quad (65)$$

Assuming $M_{oi}/M_i \approx 0.5$ we obtain the estimate

$$\theta_L = \frac{1}{3} \frac{M_I}{L} \quad (66)$$

where M_I is the total holdup inside the column.

Physically λ represents the initial change (M_i constant) in liquid flow on each tray caused by a change in boilup. This effect may be caused by vapor pushing liquid off the trays. λ is difficult to estimate, and is even found to be negative in some cases. For $\lambda > 0.5$ both x_B and the reboiler level will show an inverse response for an increase in boilup (Rademaker et al., 1975). Such behavior can be detrimental for control purposes.

Packed Columns. Eq. (61)-(62) apply also to packed columns. However, the values for θ_L and λ may be quite different in this case, and should be obtained experimentally. The holdup M_I inside the column is generally smaller for packed columns, which causes reboiler and condenser holdups to be of greater importance than for columns with trays.

Pressure Response. The pressure control was assumed perfect to derive $dV_T = dV$ (62b). This is not quite true in practice, but a first order response is probably

adequate in most cases

$$dV_T = \frac{1}{\tau_p s + 1} dV \quad (67)$$

The value of τ_p depends on the pressure control system. The open-loop pressure response is given by (Note that the pressure is approximately proportional to the vapor holdup, i.e., $p = kM_V$)

$$dM_V(s) = \frac{1}{s + k_p} (dV - dV_T) \quad (68)$$

The constant k_p takes care of the pressure's self-regulating effect (for example, an increase in V increases pressure which leads to condensation of vapor). Its value is usually small and may be neglected for control purposes, i.e., $k_p = 0$. The control system manipulates dV_T (in some cases dV) for pressure control

$$dV_T = c(s) dM_V(s) \quad (69)$$

If $c(s) = c$ is purely proportional then a first-order response (67) with $\tau_p = 1/c$ is derived. In practice $c(s)$ will also include integral action and there will also be dynamics included in order to actually change V_T (which is done indirectly by manipulating the cooling duty).

Decoupling at High-Frequency. The flow dynamics have the effect of decoupling the response of high frequency. The reflux (L) has a direct influence the compositions in the top of the column, but only a delayed influence x_B . Similarly, the boilup directly influences x_B (at least for $\lambda \approx 0$), but has a smaller initial effect on compositions in the top (because of (67)). This decoupling at high frequency yields $\lambda_{11}(\infty) = 1$ which may be beneficial for control purposes.

11. DISCUSSION/CONCLUSION

The main advantage of the simple models (11) and (61) is that they give a good description of both the low- and high-frequency behavior of distillation columns. Such models were not available in the literature. The traditional approach has been

to use a model which matches the steady-state gains, but which is not necessarily accurate for high-frequencies. The other extreme is to match the high-frequency gains (Rademaker et al., 1975, p. 137). (11) provides a link between the low- and high-frequency regions.

It is well known that the dynamic response for distillation columns is essentially first order, and we have presented results which justify this claim. (11) was derived by considering the fundamental difference between external and internal flows, both at steady state and dynamically. The parameters in (11) are the steady-state gains, the dominant first-order time constant τ_1 associated with the external flows, and the first-order time constant τ_2 associated with the internal flows. τ_1 can be estimated from the steady-state data using (2) or (4) or from the estimated initial slope using (46). τ_2/τ_1 is found using (57) which is based on matching the RGA at high frequency.

From the derivation and analysis of the model it is clear that (11) is most likely to hold for high-purity columns with large reflux. This is exactly the case for which control is expected to be most difficult.

Disturbances in feed rate (F) and feed composition (z_F) have not been discussed in this paper. Their low-frequency response is approximately first order with time constant τ_1 (Skogestad and Morari, 1987a), but the initial response of is higher order, especially for a disturbance in z_F . For feedback control the exact dynamics of the disturbances are not important. However, a good model is desirable if feedforward control is used.

An important conclusion of this paper is that the high-frequency behavior is generally much less affected by changing operating conditions than the steady-state. This partially explains why highly nonlinear distillation columns may be controlled satisfactory using linear controllers. In particular, we showed that the initial response is independent of operating conditions if relative (logarithmic) compositions

are used. This suggests that $\ln(1 - y_D)$ and $\ln x_B$ should be used as controlled outputs for columns where y_D or x_B may vary significantly.

It may be misleading to use steady-state data as an indicator of the expected control quality. In particular, this is the case for columns with both products of high purity. For such columns the RGA-values at high frequency are generally much smaller than at steady-state. From (50) and (54) we conclude that the columns which are going to be most difficult to control are high-purity columns with large reflux.

The traditional approach to modelling distillation columns is to approximate each transfer function by a first-order lag and a time delay ($ge^{-\theta s}/(1 + \tau s)$) where g is obtained by matching the steady-state gains. It is very difficult to obtain a good model for high-purity columns which captures the difference between external and internal flows using this approach. Furthermore, it is unlikely that the correct behavior at high-frequency (for example, the RGA) is obtained. Kapoor et al. (1986) have suggested to base the controller design on a model for the "perturbed" steady-state. This model is more likely to yield a reasonable high-frequency behavior. However, such "tricks" are unnecessary if one uses a model, for example (61), which accurately describes both the low and high frequency behavior.

Acknowledgements. Financial support from the National Science Foundation and Norsk Hydro is gratefully acknowledged.

References

- Bristol, E. H., "On a New Measure of Interactions for Multivariable Process Control", IEEE Trans. Automatic Control, **AC-11**, 133-134 (1966).
- Fuentes, C. and W. L. Luyben, "Control of High-Purity Distillation Columns", Ind. & Eng. Chem. Process Des. & Dev., **22**, 3, 361-366 (1983).
- Grosdidier, P. and M. Morari, "Interaction Measures for Systems Under Decentralized Control", Automatica, **22**, 3, 309-319 (1986).
- Henley, E. J. and J. D. Seader, Equilibrium-State Separation Operations in Chemical Engineering, John Wiley and Sons, New York (1981).

- Kapoor, N., T. J. McAvoy and T. E. Marlin, "Effect of Recycle Structure on Distillation Tower Time Constants," *AIChE Journal*, **32**, 3, 411-418 (1986).
- McNeill, G. A. and J. D. Sachs, "High Performance Column Control", *Chemical Engineering Progress*, **65**, 3, 33-39 (1969).
- Moczek, J. S., R. E. Otto and T. J. Williams, "Approximation Model for the Dynamic Response of Large Distillation Columns", *Proc. 2nd IFAC Congress*, Basel (1963).
- Nett, C. N. and V. Manousiouthakis, "Euclidean Condition Number and Block Relative Gain: Connections, Conjectures and Clarifications", submitted to *IEEE Trans. of Autom. Control* (1986).
- Rademaker, O., J. E. Rijnsdorp and A. Maarleveld, "Dynamics and Control of Continuous Distillation Units", Elsevier, Amsterdam (1975).
- Ryskamp, C. J., "Explicit Versus Implicit Decoupling in Distillation Control" *Proc. Chemical Process Control 2*, Sea Island, Georgia, Jan. 18-23 (1981) (T. F. Edgar and D. E. Seborg, eds., United Engineering Trustees (1982) (available from AIChE).
- Skogestad, S. and M. Morari, "Control of Ill-Conditioned Plants: High Purity Distillation", paper 74a, *AIChE Annual Mtg.*, Miami Beach (1986a).
- Skogestad, S. and M. Morari, "Implication of Large RGA-Elements on Control Performance", paper 6d, *AIChE Annual Mtg.*, Miami Beach (1986b).
- Skogestad, S. and M. Morari, "Shortcut Models - II. Dynamic Behavior", submitted to *Comp. and Chem. Eng.* (1987a).
- Skogestad, S. and M. Morari, "Understanding the Steady-State Behavior I." in preparation (1987b).
- Skogestad, S. and M. Morari, "Understanding the Steady-State Behavior. II." in preparation (1987c).
- Skogestad, S. and M. Morari, "Robust Control of a Distillation Column Using the LV-Configuration", submitted to *Chem. Eng. Sci.* (1987d).
- Skogestad, S. and M. Morari, "Control Configuration Selection for Distillation Columns", submitted to *AIChE J.* (1987e).
- Skogestad, S. and M. Morari, "Shortcut Models - I. Steady-State Behavior", submitted to *Comp. and Chem. Eng.* (1987f).
- Wahl, E. F. and P. Harriot, "Understanding and Prediction of the Dynamic Behavior of Distillation Columns," *Ind. & Eng. Chem. Process Des. & Dev.*, **9**, 396-407 (1970).

Appendix. Dynamic model of distillation column.

Assumptions:

- Constant molar flows
- No vapor holdup (immediate vapor response, $dV_T = dv_B$)
- Liquid holdup M_i on all trays constant (immediate liquid response, $dL_T = dL_B$)
- Vapor-Liquid Equilibrium (VLE) and perfect mixing on all stages

N - no. of equilibrium (theoretical) stages including the reboiler

$N + 1$ - total number of stages including total condenser

N_F - feed stage location

F - feed rate (kmol/min)

z_F - mole fraction of light component in feed

q_F - fraction liquid in feed

Nonlinear model

Material balances for change in holdup of light component on each tray:

$i = 2, N$ ($i \neq N_F, i \neq N_F + 1$):

$$M_i \dot{x}_i = L_{i+1}x_{i+1} + V_{i-1}y_{i-1} - L_i x_i - V_i y_i$$

Above feed location, $i = N_F + 1$:

$$M_i \dot{x}_i = L_{i+1}x_{i+1} + V_{i-1}y_{i-1} - L_i x_i - V_i y_i + F_V y_F$$

Below feed location, $i = N_F$:

$$M_i \dot{x}_i = L_{i+1}x_{i+1} + V_{i-1}y_{i-1} - L_i x_i - V_i y_i + F_L x_F$$

Reboiler, $i = 1$:

$$M_B \dot{x}_i = L_{i+1}x_{i+1} - V_i y_i - B x_i, \quad x_B = x_1$$

Total condenser, $i = N + 1$:

$$M_D \dot{x}_i = V_{i-1}y_{i-1} - L_i x_i - D x_i, \quad y_D = x_{N+1}$$

VLE on each tray ($i = 1, N$), constant relative volatility:

$$y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i}$$

Flow rates assuming constant molar flows (L - reflux, V - boilup):

$$i > N_F \text{ (above feed): } L_i = L, \quad V_i = V + F_V$$

$$i \leq N_F \text{ (below feed): } L_i = L + F_L, \quad V_i = V$$

$$F_L = q_F F, \quad F_V = F_{FL}$$

$$D = V_N - L = V + F_V - L \quad (\text{condenser holdup constant})$$

$$B = L_2 - V_1 = L + F_L - V \quad (\text{reboiler holdup constant})$$

Compositions x_F and y_F in the liquid and vapor phase of the feed are obtained by solving the flash equations:

$$F z_F = F_L x_F + F_V y_F$$

$$y_F = \frac{\alpha x_F}{1 + (\alpha - 1)x_F}$$

Linear model

Linearized material balance on each tray ($dL_i = dL$, $dV_i = dV$):

$$M_i \dot{x}_i = L_{i+1} dx_{i+1} - (L_i + K_i V_i) dx_i + K_{i-1} V_{i-1} dx_{i-1} + (x_{i+1} - x_i) dL - (y_i - y_{i-1}) dV$$

where K_i is the linearized VLE-constant:

$$K_i = \frac{dy_i}{dx_I} = \frac{\alpha}{(1 + (\alpha - 1)x_i)^2}$$

and y_i, x_i, L_i and V_i are the steady-state values at the nominal operating point. Written in the standard state variable form in terms of deviation variables:

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

Here $x = (dx_1, \dots, dx_{N+1})^T$ are the tray compositions, $u = (dL, dV)^T$ are the manipulated inputs and $y = (dy_D, dx_B)^T$ are the controlled outputs. The state matrix $A = \{a_{i,j}\}$ is tri-diagonal:

$$\begin{aligned} i \neq N+1: \quad a_{i,i+1} &= L_{i+1}/M_i \\ a_{i,i} &= -(L_i + K_i V_i)/M_i \\ i \neq 1: \quad a_{i,i-1} &= K_{i-1} V_{i-1}/M_i \end{aligned}$$

Input matrix $B = \{b_{i,j}\}$:

$$\begin{aligned} i \neq N+1: \quad b_{i,1} &= (x_{i+1} - x_i)/M_i, \quad b_{N+1,1} = 0 \\ i \neq 1, i \neq N+1: \quad b_{i,2} &= -(y_i - y_{i-1})/M_i, \quad b_{N+1,2} = 0, \quad b_{1,2} = (y_1 - x_1)/M_1 \end{aligned}$$

Output matrix C :

$$C = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Chapter XII

CONTROL CONFIGURATION SELECTION FOR DISTILLATION COLUMNS

CONTROL CONFIGURATION SELECTION FOR DISTILLATION COLUMNS

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

(818)356-4186

January 1987

Submitted to AIChE Journal

Abstract

Most two-product distillation columns can be described as a 5×5 plant, but the control system design is usually simplified by means of the following procedure:

1. Choose two manipulated inputs for composition control (corresponding to a specific control "configuration").
2. Design the level and pressure control system (usually three SISO controllers).
3. Design a 2×2 controller for composition control.

The goal of this paper is to provide guidelines for Step 1 which is considered the most important. Ratios (e.g. $\frac{L}{D}$ or $\frac{V}{B}$) are frequently chosen as "manipulated" inputs in Step 1. We show that the ratio configurations are effectively complex multivariable controllers which provide - among other features - improved flow disturbance rejection.

1. Introduction

Distillation columns constitute a major part of most chemical processing plants. The objective of a distillation column is to split the feed into two (or more) products with compositions different from that of the feed. The desired composition of the products may be fixed by product requirements or may result from some plantwide optimization. An important objective of the control system should be to keep these product compositions at their desired level. In practice, very few industrial columns maintain “dual” composition control, and it is still common to find found that both compositions are controlled manually. Reports from industry indicate energy saving of 10-30% (Stanley and McAvoy, 1985) if dual composition control is used instead of applying manual control which usually results in overpurification or loss of valuable product. Also, a recent survey among plant managers (Dartt, 1985) cites distillation as the unit operation which could benefit most significantly from improved control.

A main reason why dual composition control is not widely applied in industry is the stability problem often encountered when such a system is tuned in order to get a reasonably fast response. Some reasons usually cited in the literature for the problems with dual compositions control are

- strongly nonlinear behavior
- very sluggish response
- measurement problems, dead times for composition measurements
- difficult to choose appropriate manipulated variables for composition control
- strongly interactive system

These problems do not apply to all columns. Columns with low purity products tend to be simpler to control. Ironically, simple columns are the ones usually studied experimentally in university laboratories. Another reason for the infrequent use of dual composition control is the lack of systematic guidelines in the literature on how to design such control systems.

In distillation control the gap between the industrial and the academic camp is clearly evident from the literature. The *industrial camp* argues from experience. On one side hardly anybody is willing to challenge the recommendations and conclusions. On the other side, there are few who claim to understand the reasoning and the explanations. Take for example the loop pairing recommendations based on the Relative Gain Array (RGA). From the industrial literature one is led to believe that the RGA is just an interaction measure indicating the difficulties one can expect when tuning single loops for a multivariable system. But if this were true then the RGA recommendations regarding variable selection for distillation control would be archaic today, when the implementation of multivariable control systems is quite straightforward. Practical evidence suggests however, that the RGA evaluation is very useful even for the design of multivariable control systems. Thus, while the RGA has proven to be a very useful tool for categorizing experience, it has not helped to explain the observed phenomena.

The *academic camp* tends to tackle quite specific control problems with a heavy machinery (whatever multivariable control theory happens to be trendy), but efforts using an organized approach toward understanding distillation as a system have been rare. This paper is aimed at filling the apparent gap. Distillation is analyzed from a system point of view and practical experience and recommendations are interpreted in this context.

2. The Distillation Column from a System Point of View

A schematic picture of a two product distillation column is shown in Fig. 1. Conventional notation is used. A total condenser has been assumed, but this has little significance on the results which follow.

2.1 Input and Output Signals

Viewed from a systems point of view the distillation column is a “box” which takes some input functions and maps them into a set of output functions. The

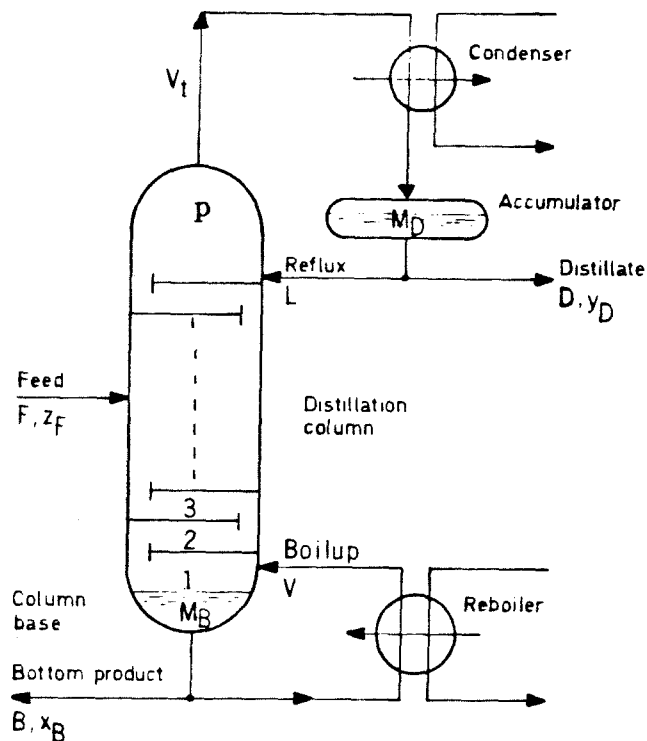


Fig. 1. Distillation column with 5 manipulated inputs (L, V, D, B and V_T) and 5 controlled outputs (y_D, x_B, M_D, M_B and p).

inputs are divided into those which can be adjusted (manipulated variables (u), usually corresponding to valves) and those which cannot be affected within the system (disturbances (d) and setpoints (y_s)). Similarly the outputs are divided into those of interest (controlled variables (y)) and the known or measured signals (y_m). Obviously, in many cases an output will be both a controlled variable and a measurement, but this is not necessarily the case. The distillation column in Fig. 1 has five manipulated inputs u and five controlled outputs y .

Controlled Outputs (y). The five controlled outputs in Fig. 1 are:

- Vapor Holdup (M_V) (expressed by the pressure p).
- Liquid holdup in accumulator (condenser) (M_D).
- Liquid holdup in column base (reboiler) (M_B).

- Distillate product composition (y_D).
- Bottom product composition (x_B).

The reason for choosing these five variables as controlled outputs is briefly discussed: Since vapor and liquid holdups must always be controlled to ensure stable operation, the pressure and the condenser and reboiler holdups (M_V, M_B, M_D) clearly have to be controlled. The liquid holdup inside the column is “self regulating” and does not have to be controlled unless the column is overloaded.

We have chosen to use the mole fractions, y_D and x_B , of the light component in the top and bottom product as our product specification. In general, other choices are possible, e.g., ratios between compositions, densities, boiling points, etc., but y_D and x_B are most common. Also note that for a multicomponent system, only one composition variable may be controlled independently for each product. In addition to the five controlled outputs mentioned above there will also be other signals which we may be concerned about. In particular we want to avoid excessive movements of the manipulated variables, mainly because of constraints. Therefore these signals should also be included as controlled outputs.

Manipulated inputs (u). The five manipulated variables in Fig. 1 are:

- Distillate flow (D)
- Bottom flow (B)
- Reflux (L)
- Boilup (V) (manipulated indirectly through the reboiler duty)
- Overhead vapor flow (V_T) (manipulated indirectly through the condenser duty)

Essentially, these correspond to the available valves. The flow rates V and V_T are controlled indirectly, usually with the flow rates of the heating and cooling medium. In some cases additional manipulated variables are available, for example, the feed rate (F) or the feed enthalpy (q_F), but this will not be considered here.

Disturbances (d) and Setpoints (y_s). The disturbances to the column are often related to the feed: the feed flow rate (F), the feed enthalpy expressed in terms of its fraction of liquid (q_F), and the feed composition (z_F). In addition, there are disturbances on the five manipulated inputs. Of these, the disturbances on V and V_T are most important. Typical sources of the disturbances in V and V_T are temperature or pressure changes of the heating or cooling medium.

Other “disturbances” are setpoint changes for y_D and x_B . Setpoint changes are not common, but will be encountered if there is a higher level optimization scheme which changes setpoints based on some overall economic objective. This kind of optimization is believed to become used increasingly in the future, and will probably constitute a major driving force towards implementing dual composition control schemes.

Measurements (y_m). The measurements typically include the pressure p (usually at several locations), the liquid holdup (“level”) in the reboiler and condenser, the top and bottom compositions (often delayed and/or sampled) and temperatures at several locations. Often some of the disturbances are measured; typically these include the flowrate and temperature of the feed and the flow rate, pressure and temperature of the heating and cooling medium.

2.2 Performance Specifications

An important factor to consider when designing a control system are the performance specifications. More precisely, these are specifications on how the controlled outputs are to behave in response to certain inputs.

Consider first liquid and vapor holdups (M_D, M_B, M_V) which have to be controlled to ensure stability. From a steady state point of view only M_V (i.e., the pressure p) has any bearing on the performance of the column. The setpoint for the pressure may be based on an optimization of the column performance. Since

separation is usually favored by low pressure, the optimal pressure is often the minimum attainable, that is, the pressure determined by the constraint of maximum cooling in the condenser ("Floating- Pressure Control" (Shinskey, 1984)). However, the pressure should always be kept slightly above the minimum attainable in order to maintain short term pressure control. Short terms pressure control is needed to avoid fluctuations in the pressure, for example due to changes in the cooling medium (Shinskey, 1984).

The control of the condenser and reboiler holdups is important not because the holdups themselves have any significance, but because changes in the holdups affect the flows controlled by them. Perfect level control is not desirable since this removes the "smoothing" effect of the holdups. This is the main reason why the holdups are there in the first place. We will not go into any detail about the performance specifications here, but only state that if reflux (L) or boilup (V) are used for level control, then these level loops should be considerably faster than the composition response.

The most important controlled outputs are the top and bottom compositions, y_D and x_B . Their setpoints may be given by strict product specifications or as a result of a column optimization. The optimization may involve, for example, a trade-off between the cost of heating medium and the money earned by recovering more of the valuable product. Obviously the error

$$e = y - y_s$$

which expresses the deviation between actual and desired product purity, should be "small". We have to define more precisely what we mean by "small", i.e., what kind of norm should be used for e . The choice of norm depends on the reasons for keeping e small. Assume there is a fixed product specification, (e.g., $x_B \leq 0.01$) which should never be violated. (The bottom stream may be a feed stream to another unit where $x_B > 0.01$ is not allowed). In this case we might choose the

setpoint to be $x_{B,s} = 0.008$ and use the performance specification:

$$\max_t |e(t)| \triangleq \|e(t)\|_\infty \leq 0.002$$

In other cases the bottom stream might go to a large storage tank, which will average out the composition such that only the average composition matters. In this case it would be desirable to have

$$\int_0^\infty e(t) dt$$

as small as possible. This may be achieved even if $e(t)$ is fluctuating wildly. This is not desirable, however, because the cost of separation increases (Shinskey, 1984); the energy saved when $x_B > 0.01$ is less than the extra energy needed when $x_B < 0.01$. Consequently, in this case a more appropriate performance specification may be to keep

$$ISE = \left(\int_0^\infty |e(t)|^2 dt \right)^{1/2} \triangleq \|e(t)\|_2 \quad \text{or} \quad IAE = \int_0^\infty |e(t)| dt \triangleq \|e(t)\|_1$$

as small as possible (but there may not be a specified upper bound on these norms). The ISE (2-norm) or IAE (1-norm) may be even more appropriate if the setpoint is determined by some optimization rather than by a product specification.

2.3 Linear Model

A distillation column is strongly nonlinear, but for control design we will describe it by a linear model. We will only outline the structure of this model.

One complication in obtaining a linear model is that without the pressure and level loops closed, the distillation column is unstable. It is then difficult to obtain open loop transfer functions for the composition responses using simulation because the reboiler and condenser overflow or run dry long before the composition response has settled. However, since the composition response is only very weakly dependent on the actual level in the condenser and reboiler (M_D and M_B), and since these levels are usually tightly controlled, a good approximation of the open

Controlled Output	Manipulated Input				
	L	V	D	B	V_T
y_D	$g_{11}(s)$	$g_{12}(s)$	0	0	0
x_B	$g_{21}(s)$	$g_{22}(s)$	0	0	0
M_D	$-\frac{1}{s}$	0	$-\frac{1}{s}$	0	$\frac{1}{s}$
M_B	$\frac{1}{s}e^{-\theta s}$	$-\frac{1-\lambda(1-e^{-\theta s})}{s}$	0	$-\frac{1}{s}$	0
$M_V(p)$	0	$\frac{1}{s+k_p}$	0	0	$-\frac{1}{s+k_p}$

Table 1. Approximate open-loop transfer matrix for distillation column (Skogestad and Morari, 1986c).

1. Constant molar flows assumed.
2. The transfer function for M_V is not a pure integrator because of condensation effects which are included in k_p .
3. $e^{-\theta s}$ with $\theta = \tau_L N$ is an approximation for $1/(1 + \tau_L s)^N$. $\tau_L = (\partial M_i / \partial L_i) V_i$ is the hydraulic time constant. N is the total number of trays.
4. $\lambda = (\partial L_i / \partial V_i) M_i$ is the initial change in liquid flow due to a change in vapor flow (V may "push" liquid off the tray and give $\lambda > 0$). An inverse response occurs if $\lambda \geq 0.5$.
5. In addition there will be dynamics involved in order to change L, V, D, B and V_T (valve dynamics, etc.)
6. Consistent units have been assumed for holdups and flows (e.g., M_D in kmol and D in kmol/min).

loop composition response is found by assuming these levels to be constant. The pressure (i.e., vapor holdup (M_V)) does have a significant effect on the composition, but since pressure is usually tightly controlled, this effect may be neglected as a first approximation. Approximate "open loop" responses are therefore obtained by varying L and V , and assuming V_T, D and B to be fixed by the requirement of perfect control of M_V, M_D and M_B . For consistency, we then also have to neglect the effect changes in V_T, D and B have indirectly on the compositions because of their effect on M_V, M_D and M_B . With these assumptions the structure of the open loop transfer matrix is as shown in Table 1. Similar transfer matrices may be

derived for the disturbances.

2.4 General Interconnection Structure

Schematically the distillation column may be written as a box as shown in Fig. 2. In general, P is a nonlinear operator giving the “nominal” relationship (model) between inputs and outputs. We will be using linear models in which case P is a transfer matrix. For linear systems P is conveniently divided into four subsystems. P_{11} and P_{21} represent the disturbance model between the disturbances d and the controlled variables y and measured outputs y_m . P_{11} will be denoted by

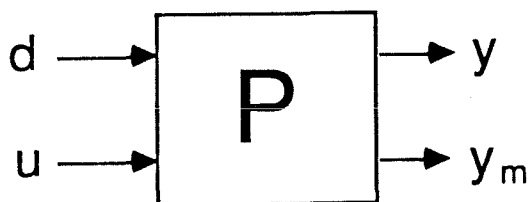


Fig. 2. Schematic representation of distillation column.

G_d later. P_{12} and P_{22} represent the model between the manipulated inputs u and the controlled outputs y and the measured outputs y_m . P_{12} is what usually is called the "process" and will be denoted G . Note that with the usual assumption $y_m = y$, Fig. 2 may be represented as in Fig. 3.

In Fig. 4 we have added two additional blocks to Fig. 2. One is the controller (C) which computes the appropriate inputs u based on the information about the process (y_m). The other block (Δ) represents the model uncertainty (Doyle et al., 1982). Here we will not dwell on this particular way of representing uncertainty, but just note that it clearly shows that \hat{P} and P are models only, and that the actual plant is different depending on Δ . Based on the measurements (y_m), the objective of the controller C is to generate inputs (u) which keep the outputs (y) as close as possible to their setpoints (y_s) in spite of disturbances (d) and model uncertainty (Δ). The controller C is often nonsquare as there are usually more measurements than manipulated variables. For the design of the controller C information about the expected model uncertainty should be taken into account. The case when $y_m \neq y$ is often called "inferential control". It is seen to be handled "automatically" in this framework.

Figure 4 was introduced by Doyle et al. (1982) and represents a unifying framework for studying linear control problems. The interconnection matrix \hat{P} includes all information needed in order to design the "optimal" C . In particular \hat{P} includes the matrix P in Fig. 2, i.e., the process G and the disturbance model G_d . Furthermore, performance "weights" are included in \hat{P} in order to be able to compare mathematically the controlled variables, which have different physical significance and in order to decide on the type of desired response. Finally \hat{P} contains information on how the uncertainty affects the overall system.

Above we have outlined a unifying framework for control problems (Fig. 4) and we have tried to give some indication on how distillation fits into it. Clearly, our

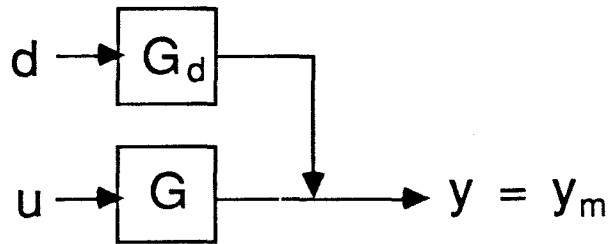


Fig. 3. Equivalent representation of Fig. 2 for a linear plant with $y = y_m$.

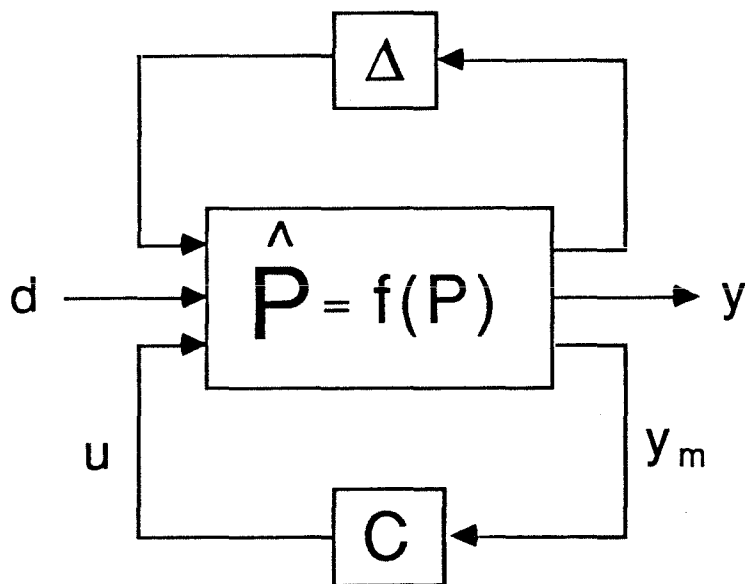


Fig. 4. General structure for studying any linear control problem.

treatment has been very brief, and for more details the reader should consult Doyle et al. (1982). The main objective is to show that the distillation control problem may be put into a systematic framework and to point out what information is needed about the process. The problem with this approach is that it results in a controller which is complicated and difficult to design and understand. Engineering judgment cannot be easily brought into this design process. Below, we will outline a stepwise procedure which leads to a much simpler design. The basic idea is to use only two independent manipulated variables for composition control. This is the approach used in practice (Shinskey, 1984). The first (and most important) step of this design approach is to choose the best control "configuration". i.e., to decide which two manipulated inputs to use for composition control.

3. A simplified approach

Assume as a simplification that all five controlled outputs (including y_D and x_B) are measured. Given the open-loop model for the distillation column, and information about disturbances, performance, uncertainty, etc., we can then imagine designing the "optimal" 5×5 controller for the column. While it is certainly of theoretical interest to find this optimal controller, it is very unlikely that such a controller would ever be implemented in practice. In order to make the control system failure tolerant and easier to understand and tune, simpler control structures are used. This will be the topic of the remainder of this paper.

More specifically, we will not use all five flows L, V, D, B and V_T for composition control, but only two independent combinations. The overall control system will then consist of a 2×2 controller (denoted by K) for composition control plus a control system for level and pressure control.

3.1 Inventory Control Leaves Only Two Degrees of Freedom for Composition

Control

The task of subdividing the problem is simplified by the observation that the

pressure and level controls are almost always much faster than the composition control (because the flow dynamics are usually significantly faster than the composition dynamics). As a first approximation assume that the pressure and level loops are so fast that they effectively give three static relationships between the five manipulated inputs (L, V, V_T, D and B) which have to be satisfied at any given time. This implies that there are only two degrees of freedom left for composition control.

In practice, the pressure and level loops are not immediate and, at least on a short time scale, more than two independent inputs could be used for composition control. However, as a first step it is reasonable to design a composition control system using only two manipulated inputs. This system may subsequently be modified to reduce the effect of the lags introduced by the level loops.

3.2 Design of a Simplified Control System

Step 1. Choose two manipulated inputs for composition control. Each choice of manipulated inputs corresponds to a specific control configuration. For example, the choice of L and V for composition control is referred to as the LV-configuration.

Step 2. Design the level and pressure control system. Shinskey (1984) calls this “closing the material and energy balance”. In most cases a simple control system using SISO-controllers is chosen, and the choice of “pairings” is usually obvious once the choice in Step 1 is made. Note, however, that the level control system can significantly affect the composition control. The main importance of this step in our context is to derive “new” “open-loop” composition responses (assuming the pressure and level loops are closed), which may be used to design the controller in

Step 3. In many cases Step 2 is simplified by assuming that the level and pressure controls are “perfect”.

Step 3. Design the 2×2 controller K for composition control. This is not a trivial step, but it is certainly much simpler than designing a 5×5 controller including

all inputs and outputs. In many cases K is restricted to be diagonal (decentralized control).

The most important step in the above procedure for designing a simplified dual composition control system is to decide on the control configuration (Step 1): Which variables should be “manipulated” in order to maintain composition control? We have put “manipulate” in quotes, because we are going to define new “manipulated” variables different from the real ones (which are the valve positions). In fact, we have already implicitly redefined the manipulated variables by assuming that we can actually manipulate the flows L, V, V_T, D and B directly instead of their valve positions. In practice, for L, D and B this may be implemented by measuring the actual flow rate and using a very fast inner loop to adjust this measured rate to match the desired flow. By this we also remove the nonlinear relationship between the valve position and the flow rate. It is usually not possible to measure V and V_T and these flows must be estimated in some other way, for example, by enthalpy balance calculations.

There is clearly an infinite number of relationships between L, V, V_T, D and B which can be defined as new “manipulated” variables. Of these we will only consider the flows themselves and ratios between the flows. A further simplification results because the condenser duty (i.e., V_T) is almost never used for controlling composition (Shinskey, 1984). The reason for this is probably that V and V_T have almost the same effect on composition, and cannot be used independently for composition control. Furthermore:

- V has a more direct effect on bottom composition (x_B), and is therefore preferable over V_T from a dynamic point of view.
- V_T is generally better for pressure control since the primary and secondary effects on pressure are always in the same direction. On the other hand, V may yield an inverse response: Initially, pressure increases in response to an

increase in V . However, if composition is uncontrolled the temperature in the column will start rising, thereby decreasing ΔT in the reboiler and increasing ΔT in the condenser, resulting in reduced pressure (Rademaker, et al., 1985).

The problem of dual composition control is then reduced to controlling the compositions y_D and x_B using two independent combinations of the inputs L , V , D and B (Shinskey, 1984).

3.3 L, V, D and B as manipulated inputs

Let us first consider the case when the flows L, V, D and B themselves are used as manipulated variables for composition control. There are $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 6$ independent pair combinations. However, only five of these are possible since D and B cannot be used together for composition control, because of the steady state material balance constraint $D + B = F$. Having chosen one of the remaining five pairs (LV, LD, LB, DV , or VB) for composition control, the control structure for the level loops usually follows easily.

Example. LV-configuration. Assume L and V have been chosen for composition control. (This is the configuration most commonly used (Rademaker et al., 1975)). The condenser level may be controlled by D , pressure by V_T and the reboiler level by B resulting in the following control structure:

$$\begin{bmatrix} dL \\ dV \\ dD \\ dB \\ dV_T \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ K & 0 & 0 & 0 \\ 0 & 0 & c_D(s) & 0 \\ 0 & 0 & 0 & c_B(s) \\ 0 & 0 & 0 & 0 & c_V(s) \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \\ dM_D \\ dM_B \\ dM_V \end{bmatrix} \quad (1)$$

3.4. Ratios between L, V, D and B as manipulated inputs

Of the possible nonlinear relationships possible between L, V, D and B we will only consider ratios. These seem to be the only nonlinear combinations used in practice (Shinskey, 1984). The total number of independent ratios is six. They are

$$\frac{L}{V}, \frac{L}{D}, \frac{L}{B}, \frac{V}{D}, \frac{V}{B} \text{ and } \frac{D}{B}$$

Including the four flows themselves this results in $\left[\begin{smallmatrix} 10 \\ 2 \end{smallmatrix} \right] = 45$ independent pairs of “manipulated” variables. Again, combinations of D , B , and $\frac{D}{B}$ cannot be used for composition control. This eliminates three of these options, but still leave us with 42 possible combinations.

Shinskey (1984) excludes the ratio $\frac{D}{B}$ because it “is not independent of D ”. However, even though configurations involving D , B , and $\frac{D}{B}$ have the same value of the RGA, the resulting control systems are generally different. Shinskey also groups $\frac{L}{V}$, $\frac{L}{D}$ and $\frac{V}{D}$ into a single manipulated variable, the separation factor S . He claims that this may be done because $\frac{L}{V}$ and $\frac{V}{D}$ uniquely determine $\frac{L}{D}$, and because $\frac{L}{D}$ determines S uniquely. However, the relationships between the flows hold only at steady state and when the feed is liquid, and the relationship between $\frac{L}{D}$ and S is only approximate. In practice, the three choices $(\frac{L}{V}, \frac{L}{D}, \frac{V}{D})$ can yield entirely different control systems.

If we look at the actual implementation there are even more than 42 options. Since the true manipulated variables are always L , V , D and B , we have to determine how $\frac{L}{V}$, for example, is implemented as a “manipulated” variable. To increase $\frac{L}{V}$ we may either increase L , decrease V or change both at the same time. If the flow dynamics and level controls were immediate, these different implementations would not affect the composition response, but, because they are not it does make a difference. We adopt the following convention: Writing the ratio between L and V as $\frac{L}{V}$ means that L is manipulated to change $\frac{L}{V}$, and writing $\frac{V}{L}$ means that V is manipulated to change the ratio.

Ratio control systems have been used in industry for at least forty years (Rade-maker et al. (1975), p. 445). Yet, almost no discussion is found in the literature on why such schemes may be beneficial. The simplest justification follows from steady-state considerations: To keep the compositions constant, the ratio L/V inside the column (slope of the operating line on the McCabe-Thiele diagram) should

be constant. Intuitively, it seems that some disturbances may be counteracted by keeping this ratio constant. However, these arguments do not explain what happens when ratios are used for closed-loop control of compositions. Furthermore, as will be shown, the effect of using a given ratio depends entirely on which second manipulated variable is chosen for composition control.

Clearly, using ratios as “manipulated” variables is a way of introducing a simple nonlinear control scheme. For example, the nonlinear implementation of $\frac{L}{D}$ as a manipulated variable is (using the convention introduced above)

$$L = \left[\frac{L}{D} \right] D \quad (2)$$

As usual, when a linear approach is taken, we consider deviations (dL) from the nominal steady state, (for example, $L = L_0 + dL$ where L_0 is the steady state value). The linear implementation corresponding to (2) then becomes

$$dL = D d \left[\frac{L}{D} \right] + \frac{L}{D} dD \quad (3)$$

or equivalently

$$L = L_0 + D_0 \Delta \left[\frac{L}{D} \right] + \frac{L_0}{D_0} \Delta D \quad (4)$$

The difference between (2) and (4) is important only if the flow rates D and $\frac{L}{D}$ change significantly with operating conditions. Because this is usually not the case there are only minor differences between the linear and nonlinear implementation.

Consequently, in most cases it does not make much difference if we use linear combinations of L, V, D and B as new manipulated variables instead of ratios. We would like to understand what kind of linear control system this corresponds to. (Surprisingly, the results presented here on the linear interpretation of the ratio schemes seem to be new). To this end consider the following examples.

Example. $\frac{L}{D} \frac{V}{B}$ -Configuration. The $\frac{L}{D} \frac{V}{B}$ -configuration is claimed by Shinskey (1984) to be applicable over the broadest range of cases and also Rademaker et al. (1975,

p. 463) recommend this scheme. A small change in $\frac{L}{D}$ and $\frac{V}{B}$ is written

$$d\left[\frac{L}{D}\right] = \frac{1}{D}dL - \frac{L}{D^2}dD \quad (5)$$

$$d\left[\frac{V}{B}\right] = \frac{1}{B}dV - \frac{V}{B^2}dB$$

Note that the constant coefficients multiplying dD , dL , dV and dB are determined at a chosen nominal steady state. The idea is to use $d\left[\frac{L}{D}\right]$ and $d\left[\frac{V}{B}\right]$ for composition control. According to the convention defined above dL and dV are manipulated to change the ratios $\frac{L}{D}$ and $\frac{V}{B}$. Rearranging (5) yields:

$$dL = Dd\left[\frac{L}{D}\right] + \frac{L}{D}dD \quad (6)$$

$$dV = Bd\left[\frac{V}{B}\right] + \frac{V}{B}dB$$

Consequently, dL and dV depend on $d\left[\frac{L}{D}\right]$ and $d\left[\frac{V}{B}\right]$, (which are “manipulated” based on the compositions y_D and x_B) and on the flow rate changes dD and dB . We could measure dD and dB and use this in (6), but note that the values of dD and dB are determined by the level control system. Therefore, consider Step 2 in the design procedure which is the design of the level control system: Because L and V are manipulated for composition control the most reasonable choices for the control of the condenser and reboiler levels are D and B , i.e.,

$$dD = c_D(s)dM_D \quad (7)$$

$$dB = c_B(s)dM_B$$

(The SISO controllers $c_D(s)$ and $c_B(s)$ are in many cases simple proportional controllers). Using (7) to eliminate dD and dB , (6) yields

$$dL = Dd\left[\frac{L}{D}\right] + \frac{L}{D}c_D(s)dM_D \quad (8)$$

$$dV = Bd \left[\frac{V}{B} \right] + \frac{V}{B} c_B dM_B$$

Let the composition controller (possibly multivariable) be

$$\begin{bmatrix} d\frac{L}{D} \\ d\frac{V}{B} \end{bmatrix} = \hat{K} \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} \quad (9)$$

and define

$$K = \begin{bmatrix} D & 0 \\ 0 & B \end{bmatrix} \hat{K} \quad (10)$$

Then the overall controller becomes

$$\begin{bmatrix} dL \\ dV \\ dD \\ dB \end{bmatrix} = \begin{bmatrix} & & \frac{L}{D} c_D & 0 \\ & K & & \\ & & 0 & \frac{V}{B} c_B \\ 0 & 0 & c_D & 0 \\ 0 & 0 & 0 & c_B \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \\ dM_D \\ dM_B \end{bmatrix} \quad (11)$$

We see from (11) that the flow rates L and V are manipulated based both on the product compositions (y_D and x_B), and on the levels (M_D or M_B). Furthermore, the two SISO level controllers ($c_D(s)$ and $c_B(s)$) each manipulate two flow rates, and therefore appear at two places in the transfer matrix for the overall controller.

If the $\frac{D}{L} \frac{B}{V}$ - configuration had been used instead, we would get a similar controller structure, but with K in the lower left corner. Also the $\frac{L}{D} \frac{B}{V}$ - or $\frac{D}{L} \frac{V}{B}$ - configurations would result in similar controller structures.

In summary, from a linear point of view, the main feature of this ratio control system is to let the level be controlled by more than one flow; the controller changes both L and D in response to a change in M_D and both V and B in response to a change in M_B . Thus, the use of ratios as manipulated variables introduces in an ad-hoc manner a multivariable control system. In other cases it leads to simplified MIMO controller for the composition control (but which is tuned as two SISO controllers). This is illustrated by the following example.

Example $D\frac{L}{D}$ -configuration. When L is manipulated in the $D\frac{L}{D}$ -configuration then a linear analysis shows

$$dL = Dd \left[\frac{L}{D} \right] + \frac{L}{D} dD \quad (12)$$

Assume that SISO controllers (decentralized control) are used for composition control: D is manipulated based on y_D and $\frac{L}{D}$ is manipulated based on x_B :

$$dD = k_1(s)dy_D \quad (13)$$

$$d\left(\frac{L}{D}\right) = k_2(s)dx_B$$

Combining (12) and (13) yields (in this case the level control system influences B and V only, but not L and D)

$$\begin{bmatrix} dD \\ dL \end{bmatrix} = \begin{bmatrix} k_1(s) & 0 \\ \frac{L}{D}k_1(s) & Dk_2(s) \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} \quad (14)$$

Effectively, a MIMO (in this case triangular) composition controller results which is tuned like two SISO controllers.

Note that in this case the effect of using $\frac{L}{D}$ as a manipulated variable, is entirely different from that found for the $\frac{L}{D}\frac{V}{B}$ -configuration. In fact, the $D\frac{L}{D}$ -configuration is not much different from the DL -configuration as seen from (14). On the other hand, the $\frac{L}{D}\frac{V}{B}$ - configuration may behave significantly different from the $L\frac{V}{B}$ - or (even more so) the LV -configuration.

Example. $\frac{D}{V}\frac{V}{B}$ -Configuration. This example combines the features found in the previous two examples. Linearizing yields

$$dD = Vd\left[\frac{D}{V}\right] + \frac{D}{V}dV \quad (15)$$

$$dV = Bd\left[\frac{V}{B}\right] + \frac{V}{B}dB$$

Let the levels be controlled as follows

$$dL = c_D(s)dM_D \quad (16)$$

$$dB = c_B(s)dM_B$$

Combining (15) and (16) yields

$$dD = Vd\left[\frac{D}{V}\right] + \frac{DB}{V}d\left[\frac{V}{B}\right] + \frac{D}{B}c_B dM_B \quad (17)$$

$$dV = Bd\left[\frac{V}{B}\right] + \frac{V}{B}c_B dM_B$$

corresponding to the control structure

$$\begin{bmatrix} dD \\ dV \\ dL \\ dB \end{bmatrix} = \begin{bmatrix} & & 0 & \frac{D}{B}c_B \\ & K & & \\ & & 0 & \frac{V}{B}c_B \\ 0 & & c_D & 0 \\ 0 & & 0 & c_B \end{bmatrix} \begin{bmatrix} dy_D \\ dx_B \\ dM_D \\ dM_B \end{bmatrix} \quad (18)$$

where

$$K = \begin{bmatrix} V & \frac{DB}{V} \\ 0 & B \end{bmatrix} \hat{K}, \quad \begin{bmatrix} d\left(\frac{D}{V}\right) \\ d\left(\frac{V}{B}\right) \end{bmatrix} = \hat{K} \begin{bmatrix} dy_D \\ dx_B \end{bmatrix} \quad (19)$$

If \hat{K} is diagonal, this results in a triangular K , but tuned as two SISO controllers. Also note that an increase in the reboiler level, M_B , will result in a simultaneous increase in D, V and B .

Summary. Based on the three examples above let us state the following generalization: Assume that one of the “manipulated” variables for composition control is $\frac{\ell_1}{\ell_2}$ and assume that ℓ_1 is the flow which is manipulated to adjust $\frac{\ell_1}{\ell_2}$. Then the linear control system corresponding to $\frac{\ell_1}{\ell_2}$ has the following features compared to using the flow ℓ_1 alone for composition control.

1. If ℓ_2 is used for level control then this level is controlled both by ℓ_1 and ℓ_2 (but tuned as a single controller).
2. If ℓ_2 is used for composition control then this composition is controlled both by ℓ_1 and ℓ_2 , i.e., we may get an effective MIMO controller using a SISO design.

In case 2 it makes little sense to use a MIMO controller to “manipulate” ℓ_1/ℓ_2 , since the same result may be obtained by using ℓ_1 alone: Consider Eq. (19) in the last example. If \hat{K} is a “full” 2×2 matrix, then the tuning is not simpler than when K is designed directly.

In most cases the major effect of using ratios for composition control is captured by the linear analysis summarized in 1) and 2) above. Ratios do not tend to correct the nonlinear behavior of distillation columns because the manipulated inputs vary only moderately with operating conditions (neglecting startup). On the other hand, the product compositions do often vary significantly with operating conditions, and a significant “linearization” effect may be obtained, for example, by using $\ln(1 - y_D)$ and $\ln x_B$ as “redefined” controlled outputs. This will be discussed in a future paper (Skogestad and Morari, 1986c).

4. Differences Between Control Configurations

Assuming immediate flow responses, perfect level control and constant molar flows we have in the absence of feed disturbances

$$dV = dL + dD \quad (20)$$

$$dL = dV + dB \quad (21)$$

These two equations suggest that any pair of input variables has the same effect: Changing L and V , for example, is equivalent to changing V and D or V and B .

Consequently, we might expect to get good and almost identical control performance for any choice of control configuration. However, there are at least seven reasons for why the choice of control configuration can make a significant difference:

1. “Uncertainty”
2. Disturbances vs. setpoints
3. Dynamic considerations
4. Rejection of flow disturbances
5. One-point (“manual”) composition control
6. Changes between “manual” and “automatic”
7. Constraints

In many cases conflicting conclusions arise from these considerations, and the engineer has to perform a more detailed analysis or use his judgment in making the

final choice. Before looking into these seven points, we will consider some general characteristics of distillation columns which are used in the subsequent discussions.

Model Characteristics of Distillation Columns. From a control point of view the most important characteristic of distillation columns appears to be that for high-purity separations (x_B and $(1 - y_D)$ are small) the 2×2 system considered for composition control is always ill-conditioned regardless of what control configuration is used (Skogestad and Morari, 1986c). By "ill-conditioned" we mean that the plant gain in certain directions is much larger than in others. Irrespective of the control configuration, the two operating variables corresponding to the high and low plant gain are the external flows (product flow rates, D and B) and the internal flows (which are changed by changing the reflux L and boilup V while keeping D and B constant) (Rosenbrock, 1962). As an illustration consider the column in Table 2 with $z_F = 0.5$, $y_D = 0.99$, $x_B = 0.01$ and $D = B = 0.5$ kmol/min. Assume the distillate flow D is increased by 5% to 0.525 kmol/min. Since there is only 0.5 kmol/min of light component in the feed at least 0.025 kmol/min of this has to be heavy component. The best attainable value for the top composition, even with total reflux, is then $y_D = 0.5/0.525 = 0.952$. This is far from the desired $y_D = 0.99$.

More generally, the effect of the external flows on the product compositions is found using

$$\frac{D}{B} = \frac{z_F - x_B}{y_D - z_F} \quad (22)$$

This exact expression can be derived from an overall material balance for the light component. It implies that the ratio D/B should be kept constant for any flow disturbance. Furthermore, for high-purity the columns, the relative changes in y_D and x_B are extremely sensitive to changes in D/B . For example, with y_D constant, differentiation of (22) yields

$$\frac{dx_B}{x_B} = -\frac{y_D - z_F}{x_B} d\left(\frac{D}{B}\right) \quad (23)$$

The factor multiplying $d[\frac{D}{B}]$ approaches infinity when $x_B \rightarrow 0$.

Binary Separation, Constant Molar Flows, Feed Liquid

Relative volatility	$\alpha = 1.5$
No. of theoretical trays	$N = 50$
Feed tray location	$N_F = 21$
Feed rate and composition	$F = 1 \text{ kmol/min}, z_F = 0.5$
Product compositions	$y_D = 0.99, x_B = 0.01$
Product rates	$D = B = 0.5 \text{ kmol/min}$
Reflux rate	$L = 2.71 \text{ kmol/min} (1.39L_{min})$

Linearized steady-state gains, LV -configuration:

$$\begin{bmatrix} dy_D \\ dx_B \end{bmatrix} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \begin{bmatrix} dL \\ dV \end{bmatrix} + \begin{bmatrix} 0.394 \\ 0.586 \end{bmatrix} dF + \begin{bmatrix} 0.881 \\ 1.119 \end{bmatrix} dz_F$$

1-1 element in the RGA for various configurations:

	LV	$\frac{L}{D} \frac{V}{B}$	$\frac{L}{D} V$	$\frac{L}{D} D$	DV	DV
λ_{11}	35.1	3.22	5.85	0.60	0.45	0.56

Table 2. Data for distillation column example (Skogestad and Morari, 1986a).

For a more quantitative analysis, a Singular Value Decomposition (SVD) can be performed on the 2×2 transfer function model for each configuration (Skogestad and Morari, 1986a). For high-purity columns, the singular values are always found to be very different in magnitude (ill-conditioned system). The singular vectors confirm that the large plant gain is associated with a change in the external flows while changes in the internal flows have a much smaller effect on the compositions. As we will show, the observed advantages of certain control configurations can be explained from these basic characteristics.

The RGA. The RGA is determined by the plant transfer matrix (G)

$$\text{RGA} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 1 - \lambda_{11} \\ 1 - \lambda_{11} & \lambda_{11} \end{bmatrix}, \lambda_{11} = \frac{1}{1 - \frac{g_{12}g_{21}}{g_{11}g_{22}}} \quad (24)$$

The RGA is used extensively by Shinskey (1984) to compare control configurations. From his book the reader is led to believe that the RGA is useful because it provides a measure of interactions when using a decentralized controller. His rule (though he does not express it explicitly) is to choose a configuration with λ_{11} in the range of about 0.9 to 4 (Shinskey, 1984, Table 5.2). If λ_{11} were used only as an interaction measure this recommendation would not make any sense; in this case λ_{11} should be chosen to be as close to one as possible and $\lambda_{11} = 0.67$ would be almost equivalent to $\lambda_{11} = 2$ (both have $|\frac{g_{12}g_{21}}{g_{11}g_{22}}| = 0.5$). Consequently, Shinskey's use of the RGA is a way of categorizing his experience on distillation columns, rather than expressing the effect of interactions. In fact, his rules also apply when a multivariable controller is used. His recommendations regarding the RGA should therefore only be used for distillation columns. One objective of this section is to provide some justification for Shinskey's rules.

4.1 Uncertainty

Since we are considering different choices of manipulated inputs, the uncertainty associated with these manipulated inputs may cause different control behav-

ior. These issues have been discussed in detail by Skogestad and Morari (1986 a,b) and the main result is summarized below.

The RGA and input uncertainty. The RGA is a good indicator of plant-sensitivity to input uncertainty (Skogestad and Morari, 1986b). In general, a plant with large elements in the RGA is difficult to control in the presence of input uncertainty and, in particular, inverse-based controllers should be avoided.

Let Δ_1 and Δ_2 represent the magnitude of the relative uncertainty on each manipulated input. Then the actual ("perturbed") plant can be written in terms of the model G and this uncertainty:

$$G_p = G(I + \Delta_I), \quad \Delta_I = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \quad (25)$$

For good closed-loop performance, an inverse-based controller is desirable, for example, $C(s) = c(s)G(s)^{-1}$ where $c(s)$ is a scalar. The loop transfer function in this case becomes

$$G_p C = GC(I + G\Delta_I G^{-1}) \quad (26)$$

If the error term $G\Delta_I G^{-1}$ is large, the actual loop transfer function $G_p C$ will be significantly different from the desired (nominal) loop transfer function GC , and the closed-loop response is expected to be poor or even unstable. The diagonal elements of $G\Delta_I G^{-1}$ are a function of the RGA only

$$G\Delta_I G^{-1} = \begin{bmatrix} \lambda_{11}\Delta_1 + \lambda_{12}\Delta_2 & -\lambda_{11}\frac{g_{12}}{g_{22}}(\Delta_1 - \Delta_2) \\ \lambda_{11}\frac{g_{21}}{g_{11}}(\Delta_1 - \Delta_2) & \lambda_{21}\Delta_1 + \lambda_{22}\Delta_2 \end{bmatrix} \quad (27)$$

(27) clearly shows that the closed-loop response for plants with large RGA-elements is extremely sensitive to input uncertainty if a tight (inverse-based) controller is chosen. Note that it is the value of the RGA around the crossover-frequency which is of main interest. Using the steady-state value may be misleading (yield too large values). This is generally the case for columns with both products of equal purity (for example, the column in Table 2).

The result (27) explains in a quantitative way why configurations with large RGA-elements should be avoided. However, through the following discussion we want to give the reader a more intuitive feeling for why some configurations are sensitive to input uncertainty and others are not.

A Physical Interpretation of the Effect of Flow (Input) Uncertainty. From Eq. (20) (which applies to the case with perfect level control) it seems that a change in distillate flow (dD) may be achieved in two equivalent ways

- i) manipulate D directly
- ii) manipulate L and V such that $dV - dL = dD$

Similar arguments apply to other flows. However, such arguments only hold in the absence of input uncertainty. In practice, the actual flows are not the same as those demanded by the controller (the controller may try to increase a particular flow by 1 kmol/min, but the actual increase may be only 0.9 kmol/min, corresponding to 10% uncertainty with respect to the change). This input uncertainty may result in an enormous difference between various configurations since for tight control it is often desirable to make dL and dV large, while keeping dD small. This is almost impossible if, for example, the LV-configuration is used since we cannot, in practice, control differences ($dL - dV$) between two large flows accurately.

Example. Consider the column in Table 2 with $D = B = 0.5 \text{ kmol/min}$, $L/D = 5.4$, $V/B = 6.4$ and assume that we want to increase the internal flows (desired: $dL = dV = 1 \text{ kmol/min}$) without changing the external flows (desired: $dD = dB = 0$). Let us look how three different configurations would perform under these assumptions in the presence of uncertainty.

LV-configuration. Assume there is 10% uncertainty about the flow rate change, i.e.

$$dL = 1 \pm 0.1 \text{ kmol/min}, \quad dV = 1 \pm 0.1 \text{ kmol/min}$$

(in practice the uncertainty in the boilup (V) is probably larger than that for the reflux (L)). With the LV-configuration the distillate (D) and bottom (B) flows will

feel the full effect of this uncertainty

$$dD = 0 \pm 0.2 \text{ kmol/min}, \quad dB = 0 \pm 0.2 \text{ kmol/min} \quad (28)$$

This is highly undesirable because of the strong sensitivity of the compositions to changes in the external flows. The high value of the RGA ($\lambda_{11} = 35.1$) for this configuration predicts the sensitivity.

DV-configuration. For the same flow uncertainty we get

$$dL = dV = 1 \pm 0.1 \text{ kmol/min}$$

However, since D is manipulated directly, these changes do not result in any change in D .

$$dB = dD = 0 \text{ kmol/min} \quad (29)$$

Not surprisingly, the RGA-elements are generally less than one for this configuration ($\lambda_{11} = 0.45$ for this example).

$\frac{L}{D} \frac{V}{B}$ -configuration. If initially $dL \neq dV$ (because of uncertainty) then changes in the top and bottom accumulator levels occur. As is apparent from (8), these changes lead to adjustments of dL and dV , which will counteract the initial imbalance.

Assume now that the 10% uncertainty on L and V initially (before the level loops take action) results in $dL_1 = 1.1$ and $dV_1 = 0.9$ kmol/min. Let the subsequent flow adjustments made by the level control system be denoted as dL_2 and dV_2 . Then the final steady-state flows are

$$dL = dL_1 + dL_2, dV = dV_1 + dV_2$$

Furthermore, we must have perfect level control at steady-state

$$dV = dL + dD, dL = dV + dB$$

and according to (11) the levels are adjusted such that

$$dL_2 = \frac{L}{D} dD, dV_2 = \frac{V}{B} dB$$

Solving these equations gives

$$dB = (dL_1 - dV_1)/(1 + \frac{L}{D} + \frac{V}{B}) = 0.2/12.8 = 0.015 \text{ kmol/min} \quad (30)$$

The resulting error in B and D due to uncertainty in L and V is therefore reduced by a factor of $(1 + \frac{L}{D} + \frac{V}{B})$ compared to the LV-configuration. Interestingly, Skogestad and Morari (1986c) have shown that the elements in the RGA are also reduced by a factor of about $(1 + \frac{L}{D} + \frac{V}{B})$ compared to the LV-configuration. (The exact value is $\lambda_{11} = 3.22$ for this example). However, we may still have control problems for very high-purity columns because of the extreme sensitivity to changes in D and B . In this case the RGA should be computed to get a reliable indication of whether input uncertainty will cause problems or not.

Summary. The presence of input uncertainty favors using configurations with small elements in the RGA (Skogestad and Morari, 1986b). In general, all configurations involving D or B have $|\lambda_{11}| < 1$, while all others have $|\lambda_{11}| > 1$ (Shinskey, 1984, p. 146). The LV-configuration generally has the largest RGA-elements. Any configuration which uses D or B is therefore insensitive to input uncertainty, but the ratios $\frac{L}{D}$, $\frac{V}{B}$, $\frac{L}{B}$ or $\frac{V}{D}$ (or their inverses) may also be a good choice for columns with high reflux.

4.2 Disturbances vs. Setpoints

Although we just concluded that plants with large RGA-elements should be avoided, it is really large RGA-elements in the controller which cause control problems (Skogestad and Morari, 1986b). However, in most cases (in particular, if good setpoint tracking is desired), it is desirable to use an inverse-based controller to get good performance, and in this case the controller has large RGA-elements whenever the plant does.

For distillation columns, if we do not care too much about setpoint tracking, it may not be necessary to use an inverse-based controller to achieve good control performance. A diagonal controller always has $\lambda_{11}(C) = 1$ and is therefore not

sensitive to input uncertainty, but it often does not yield adequate control performance. For distillation columns, however, the disturbances are often “aligned” with the plant and may be counteracted with a diagonal controller. An accurate measure of how a disturbance d (which has the effect g_d on the outputs (y)) is aligned with the plant (G) is provided by the disturbance condition number (Skogestad and Morari, 1986d).

$$\gamma_d(G) = \frac{\|G^{-1}g_d\|_2}{\|g_d\|_2} \bar{\sigma}(G)$$

($\bar{\sigma}(G)$ denotes the maximum singular value of G , $\|\cdot\|_2$ denotes the Euclidean norm). Depending on the direction of g_d , $\gamma_d(G)$ ranges in magnitude between 1 and $\gamma(G)$ (the condition number of G). For distillation columns the values of $\gamma_d(G)$ for the disturbances are usually significantly smaller than $\gamma(G)$. For example, consider the distillation column in Table 2 which has $\gamma(G) = 141.7$ for the LV -configuration. Disturbances in $d = z_F, F, q_F, L$ and V yield $\gamma_d(G) = 1.48, 11.75, 1.09, 1.41$ and 1.41 and a diagonal controller may give acceptable response (Skogestad and Morari, 1986b).

Summary. Configurations with large RGA-elements (e.g., the LV -configuration) are not sensitive to input uncertainty if a diagonal controller is used. A diagonal controller may be acceptable if the disturbance condition number is small for all expected disturbances (and tight setpoint tracking is not required). This means that the LV -configuration may be acceptable in some cases even when it yields large RGA-values.

4.3 Dynamic Considerations

These issues are addressed in detail in the literature (Rademaker et al., 1985, Shinskey, 1984), and only a short summary is given here.

The flow rates L and V (or V_T) are the only ones which influence compositions directly. The direct effect of changing B or D is to change M_B and M_D , which has no effect on compositions. The effect on composition is caused by the level loops

which change L , V or V_T in response to the change in B and D . However, even a 1- or 2-minute lag caused by the level loops may make it difficult to counteract a large disturbance, which may change the product composition considerably in a matter of minutes. (The speed of the level loops is limited by noise on the level measurements, but is otherwise independent of the amount of holdup). These considerations are even more important for packed columns where the holdup inside the column is smaller. A possible solution (Shinskey, 1984, p. 128) is to let the composition loop also influence the flow used for level control (i.e., L or V) (This effect is only temporary, but will improve the dynamic response).

Other issues which should be considered are:

- Even L has only a delayed effect on x_B .
- An increase in boilup (V), may in some cases initially push liquid off the trays and result in a temporary increase in liquid flow in the column ($\lambda \geq 0.5$ in Table 1). The effect is a possible inverse response for V 's effect on M_B and x_B .
- Large overshoots in the open-loop response are often encountered with the material-balance configurations (using D or B): For example, for the DV-configuration, an increase in V will first cause x_B to fall. However, since D is constant, the increase in V will eventually produce an equal increase in L , which brings more light component back to the bottom, and cause x_B to return almost to its original value. This large overshoot in the response corresponds to a LHP-zero close to the origin. Shinskey (1984, p.157) claims that this LHP-zero causes control problems. This may be the case if a PID-controller is used (which cannot easily counteract the effect of the zero), but should not cause problems in general.

Summary. L and V should be manipulated directly for composition control to get a fast initial response. This is probably one of the main reasons for the popularity

of the LV-configuration. The $\frac{L}{D} \frac{V}{B}$ -configuration also has this feature. Use of D or B for composition control is generally not recommended if a fast initial response is desired.

4.4 Rejection of Flow Disturbances

The major flow disturbances are in the

- feed rate (F)
- feed enthalpy (q_F)
- boilup (V)
- condenser vapor rate (V_T)
- reflux temperature

The fraction liquid in the feed, q_F , is used as a measure of feed enthalpy. The result of a decrease in reflux temperature (possibly caused by sub-cooling the reflux) is equivalent to a simultaneous increase in L and a decrease in V_T . There will also be disturbances in L , D and B (e.g., due to measurement noise), but those are usually of less importance. Three ways of handling flow disturbances are

1. feedforward control
2. through their effect on composition
3. through their effect on levels and pressure

The first option is possible only if the disturbance can be measured. The level and pressure loops are usually much faster than the composition loops, and intuitively it seems preferable to try to reject the flow disturbances with the level loops (Option 3). However, since any flow disturbance which is not rejected by the level loops, will result in a upset in composition, one may argue that the composition control system may as well take care of all disturbances (Option 2). The problem is that it may not be possible to tune the composition loops sufficiently fast to get acceptable response for large disturbances. This is in particular the case if L and V are not manipulated directly for composition control (see 4.3 above). Furthermore, by using Option 3

we retain some disturbance rejection capability in the case the composition loops are in "manual".

Rejecting flow disturbances with the level loops (Option 3). The effect of flow disturbances on compositions in this case is found by assuming that the inputs used for composition control (u_1 and u_2) are constant. The effect depends strongly on the chosen control configuration:

- If there is a disturbance directly on a flow which is manipulated for pressure or level control alone, it will be corrected almost immediately by the level loop.
- On the other hand, if there is a disturbance on a flow used for composition control alone then the flow itself is not corrected. However, corrections on other flows may counteract the effect of the disturbance on the compositions.

As an example, consider the DV -configuration and assume there are disturbances on the boilup (V). Then disturbances on V are not corrected and affect the operation. However, the effect on the compositions is small because the disturbance on V causes L to increase and the product flowrates (D and B) do not change. Therefore the steady-state values of the compositions are almost unaffected.

The Disturbance Gain Matrix (G_d). The effect of disturbances on the product compositions is expressed mathematically by the disturbance gain matrix. (The steady-state matrix may be used since the level and pressure loops are much faster than the composition loops). Assume all the gains (including $\left[\frac{\partial y_D}{\partial d}\right]_{L,V}$) are known for L and V as manipulated inputs. (This is the most "natural" choice as seen from Table 1). We can then express $(\partial y_D / \partial d)$ for any other set of manipulated inputs as follows

$$\left[\frac{\partial y_D}{\partial d}\right]_{u_1, u_2} = \left[\frac{\partial y_D}{\partial L}\right]_V \left[\frac{\partial L}{\partial d}\right]_{u_1, u_2} + \left[\frac{\partial y_D}{\partial V}\right]_L \left[\frac{\partial V}{\partial d}\right]_{u_1, u_2} + \left[\frac{\partial y_D}{\partial d}\right]_{L, V} \quad (31)$$

The terms $(\frac{\partial L}{\partial d})_{u_1, u_2}$ and $(\frac{\partial V}{\partial d})_{u_1, u_2}$ are easy to evaluate if constant molar flows are assumed. Clearly, it is advantageous to choose configurations which have small

values of $(\frac{\partial y_D}{\partial d})_{u_1, u_2}$ for all disturbances. We will return with a more detailed discussion on how to evaluate the disturbance gains in a future paper.

Effect of Flow Disturbances on $\frac{D}{B}$. The described procedure (31) is exact, but does not give much insight. Since the product compositions are most sensitive to changes in the external flows (or equivalently $\frac{D}{B}$ (22)), an alternative approach is to consider the effect of flow disturbances on $\frac{D}{B}$. Configurations for which the effect is large should be avoided. It can be shown that an important feature of some of the ratio control schemes is that they have a good "built-in" rejection of flow disturbances.

Example. Assume the feed is liquid and consider a feed flow disturbance. If the LV- or DV-configuration is used, this disturbance will immediately give an increase in bottoms flow rate (B), leading to a large upset in x_B and y_D . However, if the $\frac{L}{D} \frac{V}{B}$ -configuration is used, all flows are adjusted proportionally, and the effect on compositions is very small: The increased feed flow rate initially brings light components down the column which would increase x_B . However, it also leads to an increase in reboiler level. From (11) we see that this leads to a simultaneous increase in B and V (while the LV- and DV-configurations keeps V constant). The increased boilup (V) returns light components to the column, and counteracts the initial effect the increased feed flow had on compositions. Furthermore, the increase in V leads to an increase in distillate flow (D). The feed flow disturbance is therefore distributed to both products, and $\frac{D}{B}$ is kept unchanged.

Table 3 summarizes the effect of some flow disturbances on $\frac{D}{B}$. Note that disturbances in V, L and $-q_F$ all increase the net flow from the reboiler to the condenser, and have the same effect on $\frac{D}{B}$. The results in Table 3 seem to be new and provide a simple explanation for why, for example, the $\frac{L}{D} \frac{V}{B}$ -configuration is less sensitive to flow disturbances than the LV-configuration.

Summary. It is preferable to use the level control system to reject flow disturbances. V_T is usually used for pressure control, and disturbances in condenser

Configuration (u_1, u_2)	Disturbance d			
	dF	$dV_d - dL_d - F dq_F$	dD_d	dB_d
LV	$k(1 - q_F - D/F)$	k	0	0
$\frac{L}{D} \frac{V}{B}$	0	$\frac{k}{1 + L/D + V/B}$	$\frac{kL/D}{1 + L/D + V/B}$	$\frac{-kV/B}{1 + L/D + V/B}$
$\frac{L}{D} V$	$-k \frac{V/F}{1 + L/D}$	$\frac{k}{1 + L/D}$	$\frac{kL/D}{1 + L/D}$	0
$L \frac{V}{B}$	$k \frac{L/F}{1 + V/B}$	$\frac{k}{1 + V/B}$	0	$-\frac{kV/B}{1 + V/B}$
DX	$-kD/F$	0	k	0
BX	kB/F	0	0	$-k$

q_F - fraction of liquid in feed

X - denotes any other manipulated input ($L, V, \frac{L}{D}$, etc.) except D, B and $\frac{D}{B}$.

$k = (1 + D/B)/B = F/B^2$

subscript d denotes an additive disturbance on this flow.

Table 3. $(\frac{\partial D/B}{\partial d})_{u_1, u_2}$ = Linearized effect of flow disturbances on D/B when both composition loops are open. Applies to steady state and constant molar flows. For derivation of Table, see Appendix.

duty are rejected perfectly (at least at steady state). However, no configuration can reject all flow disturbances using the level control system: The commonly used LV-configuration does not reject disturbances in F, V, L and q_F . Configurations using D or B as one of the manipulated variables for composition control are insensitive to disturbances in V, L and q_F , but do not reject disturbances in F . (However, F is often measured and a feedforward control scheme may be used). The $\frac{L}{D} \frac{V}{B}$ -configuration is insensitive to disturbances in F , and also rejects other flow disturbances also well, provided the reflux is large.

4.5 One-Point ("Manual") Composition Control

Very few distillation columns are actually operated with a "two-point" control system. In most cases one of the compositions is controlled manually - at least part of the time. Since the operators do not monitor the compositions continually

and manipulate the inputs accordingly, it is important that the effect of expected disturbances on the manually controlled ("uncontrolled") composition is as small as possible.

Both composition loops open. This issue was discussed above (Section 4.4) for the case of flow disturbances and the $\frac{L}{D} \frac{V}{B}$ -configuration was found to give good disturbance rejection. However, a feed composition (z_F) disturbance has no direct effect on the flows. Consequently, if both composition loops are "open", the effect of a feed composition disturbance will be the same for all configurations. Furthermore, the effect will usually be large because a change in feed composition requires a change in $\frac{D}{B}$ (Eq. (22)), and if this correction is not made, large changes in y_D and x_B will result for high-purity separations. As an example assume that initially $z_F = 0.5$, $x_B = 1 - y_D = 0.01$ and $\frac{D}{B} = 0.5$. A feed composition disturbance results in $z_F = 0.6$, but $\frac{D}{B} = 0.5$ remains constant. Then, according to Eq. (22), x_B has to increase at least to $x_B = 0.20$ (corresponding to $y_D = 1.0$). This is clearly not acceptable. Therefore, at least one of the compositions has to be controlled carefully, either by a feedback controller or by the operator.

One-point composition control (one composition loop open). Assume we have closed one loop, and are using u_2 to control y_2 . The output y_1 is not controlled and the manipulated input u_1 is constant. What is the effect of a disturbance d on the uncontrolled output y_1 ? First consider the steady-state where we have perfect control of y_2 . The disturbance d has the effect $\begin{bmatrix} g_{1d} \\ g_{2d} \end{bmatrix}$ on the outputs when the inputs u_1 and u_2 are constant. Using deviation variables we have

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = G \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} g_{1d} \\ g_{2d} \end{bmatrix} d \quad (32)$$

Solving for $y_2 = 0$ and $u_1 = 0$ gives

$$\frac{y_1}{d} = -\frac{g_{12}}{g_{22}} g_{2d} + g_{1d} \quad (33)$$

Consequently, the disturbance will not affect the uncontrolled output y_1 if

$$\frac{g_{12}}{g_{22}} = \frac{g_{1d}}{g_{2d}} \quad (34)$$

This result should be obvious: If the disturbance has the same relative effect ($\frac{g_{1d}}{g_{2d}}$) on the outputs as input $u_2(\frac{g_{12}}{g_{22}})$, then we can get perfect disturbance rejection by using only this input.

Example. Consider again the column in Table 2. For a feed composition (z_F) disturbance all configurations have $\frac{g_{1d}}{g_{2d}} = 0.787$. The ratio $\frac{g_{1i}}{g_{2i}}$ to $\frac{g_{1d}}{g_{2d}}$ (denoted r_i) is given in Table 4 for various configurations. If this ratio is close to one then perfect disturbance rejection is achieved with manipulated input i alone (the other input being constant). D (or B) should obviously never be held constant. Configurations which keep L or V constant come out favorably. The same conclusion holds also for a feed flow disturbance. The reason is that in both cases the major effect of the disturbances may be counteracted by changing the product flow rates (adjusting $\frac{D}{B}$ to satisfy (22)) which is easily accomplished using the LV-configuration.

Configuration (u_1, u_2)	$r_1 = \frac{g_{11}/g_{21}}{g_{1d}/g_{2d}}$ (u_2 constant)	$r_2 = \frac{g_{12}/g_{22}}{g_{1d}/g_{2d}}$ (u_1 constant)
LV	1.03	1.00
$\frac{L}{V} \frac{D}{B}$	1.24	0.85
$\frac{L}{D} V$	1.03	0.85
$\frac{L}{D} D$	-1.27	0.85
DV	1.03	-1.27
LD	-1.27	1.00

Table 4. Effect of feed composition disturbance ($d = z_f$) when one composition loop is in manual for column in Table 1. If $r_1(r_2)$ is close to one, then good composition control is maintained over the "uncontrolled" composition $x_B(y_D)$ when $u_2(u_1)$ is constant.

However, we have not considered the dynamic effects. For the LV-configuration flow disturbances are very poorly rejected by the level loops, and large changes in the uncontrolled composition may occur: Assume the top composition y_D is controlled with L , and x_B is left uncontrolled (i.e., V is constant). If the feed is liquid, a feed flow disturbance will reach the reboiler very fast and lead to a large change in bottom composition in a matter of minutes. Because it will take time before the feed flow disturbance is noticed in the top composition, and because of the “time delay” between a change in liquid flow (L) at the top and its effect on liquid flow in the bottom, the bottom composition will experience a large deviation before returning to its desired value. The $\frac{L}{D} \frac{V}{B}$ -configurations may be preferable from a dynamic viewpoint since the level loops will counteract the feed flow disturbance directly (without having to wait for the compositions to change).

Summary. Operating both composition loops “open” is not acceptable because no correction can be made for feed composition disturbances. When one-point composition control is used, reasonably good control of the “uncontrolled” composition is maintained with most configurations, provided D or B are not kept constant. The LV-configuration (keeping L or V constant) comes out favorably when only steady-state considerations are taken into account, but it may be preferable to use one of the ratio control schemes (e.g., $\frac{L}{V} \frac{V}{B}$) in order to obtain better dynamic rejection of flow disturbances. One advantage of controlling only one composition is that tuning is simple and very tight control can be maintained for this composition.

4.6 Changes between “manual” and “automatic” control.

Changing one of the composition loops between “manual” and “automatic” control is frequently done when controlling distillation columns, for example, due to stability problems, constraints or failures in measurements or actuators. It is clearly desirable to be able to do this without upsetting the rest of the system or having to retune the controllers.

From its definition, we might expect the RGA to give a reliable measure of how the system is affected by changing loops from “manual” to “automatic”: Each element in the RGA is defined as the open-loop gain (all the other loops in manual) divided by the gain between the same two variables when all the other loops are under “perfect” control (in automatic) (Bristol, 1966). For example, for 2×2 plants

$$\lambda_{11} = \frac{(\partial y_1 / \partial u_1)_{u_2}}{(\partial y_1 / \partial u_1)_{y_2}} = \frac{\text{Gain all other loops open}}{\text{Gain all other loops closed}} \quad (35)$$

However, the RGA is actually of very limited usefulness, because it does not take into account the effect of disturbances as illustrated for the *DV*- and *LV*-configurations below.

DV-configuration. Assume a decentralized control system is used (*D* controls y_D , and *V* controls x_B). This control system will provide acceptable control of both compositions in many cases. However, if the loop involving *D* is put in manual (i.e., *D* is constant), the response of y_D will be very poor when there are disturbances in the feed conditions. This was discussed in Section 4.5 (Table 4), and is even more transparent from the following exact expression

$$\left[\frac{\partial y_{DH} / y_{DH}}{\partial F / F} \right]_{D, x_B} = - \frac{z_F - x_B}{y_{DH}} \frac{F}{D} \quad (36)$$

Here $y_{DH} = 1 - y_D$ represents the mole fraction of heavy component in *D*. y_{DH} is seen to be extremely sensitive to changes in *F* if the distillate is of high purity ($y_{DH} \rightarrow 0$).

LV-configuration. A decentralized control system may in some cases give reasonable control of both compositions when there are feed disturbances (Section 4.2). Furthermore, if the loop involving *L* or *V* is put in manual, we still get reasonably good control of the uncontrolled composition (Section 4.5).

In general, the *LV*-configuration yields large RGA-elements while the *DV*-configuration yields small (for the column in Table 1, λ_{11} is 35.1 and 0.45 for the two cases). Yet, when the loop involving y_D is put in manual, the response of this

uncontrolled composition is still acceptable for the LV-configuration, but poor for the DV-configuration. The RGA is therefore not a reliable indicator of changes in performance when changes from automatic to manual are made.

Summary. Configurations which use D or B may give very poor response for the uncontrolled composition when the loop involving D or B is put in manual. (This is the opposite of what one might expect from the RGA, since one can always choose pairings such that $0.5 < \lambda_{11} < 1$ in this case). The LV- and $\frac{L}{D}\frac{V}{B}$ -configurations which are preferable for one-point composition control (Section 4.5), are also most easily changed between manual and automatic (though the response for the controlled composition may deteriorate when the other loop is closed).

4.7 Constraints

Avoiding constraints. Constraints on flow rates or on holdups (level and pressure) may also be important when choosing the best configuration. Whenever a manipulated input hits a constraint, it is no longer useful for control purposes. Since level and pressure control always has to be maintained, this means that one of the product compositions can no longer be controlled. If a constraint on a flow used for composition control is reached and two-point composition control is still maintained, then the constraints is akin to an input uncertainty. Therefore, constraints is an additional reason for not using controllers with large RGA-elements (for example, a decoupler for the LV-configuration).

Flows used for level control will usually have the largest variations in magnitude, and are most likely to hit constraint. This leads to the following conclusions:

- A flow which may easily reach its constraint should not be used to control holdup. In particular this statement will generally imply the following:
- A very small flow should not be used to control level. One example documented in the literature (McNeill and Sacks, 1969) is the use of distillate D to control M_D in a high reflux column with $L/D = 70$. This is clearly next to impossible.

Any imbalance in the large flows L and V_T will result in wild variations in D , and because of constraints on D the reflux drum is likely to run empty or overflow.

The possibility of meeting constraints makes it necessary to have some “override” control system (e.g., the operator) which is able to identify constraints and change the control configuration if the constrained flow rate is used for inventory control.

Operating at Constraints. Many industrial columns are operated at their capacity limit, usually wrt. the boilup V , the reflux L , or the condensation rate V_T . This is another reason for why many columns are operated with only one composition being controlled. Fortunately, as pointed out in 4.5, keeping L, V or V_T constant will also result in reasonably small variations in the uncontrolled product - at least at steady state. Since the active constraint may vary with operating conditions, an “override” control system is needed also in this case.

4.8 Choice of Control Configuration. Conclusion

The $\frac{L}{D} \frac{V}{B}$ -configuration comes out very favorably when all the points mentioned above are considered as a whole. This is also in accordance with the recommendation given by Shinskey (1984), and our analysis provides added justification for his claim. The main exception is very high-purity columns or columns with low reflux (i.e., large relative volatility) which may result in large elements in the RGA and give a system which is sensitive to input uncertainty and flow disturbances. For these columns a configuration using D or B for composition control should be considered (e.g., the $D \frac{V}{B}$ -configuration). These configurations have all RGA-elements less than one and are always insensitive to input uncertainty.

5. Conclusions

The main goal of this paper has been to present in a systematic manner the main issues which have to be addressed when designing a composition control sys-

tem. In order to avoid excessive length, a number of important issues have been addressed only qualitatively. More quantitative results on specific issues will follow. These include relationships for computing steady state gains for various configurations, and simple dynamic models. Nevertheless, it is clear that a number of generally conflicting considerations have to be taken into account.

Two-Point Composition Control. The RGA is a useful tool for addressing the issue of input uncertainty. Configurations with large values of λ_{11} should be avoided. For distillation columns all material-balance configurations (using D or B) have $\lambda_{11} < 1$. However, these configurations often result in a poor dynamic response and give very poor disturbance rejection if the loop involving D or B is taken out of service. This is probably the reason for why Shinskey (1984) recommends avoiding configurations with $\lambda_{11} < 1$ (provided λ_{11} is not too large). (These considerations only hold for distillation column control, and for other processes there is no reason to try to avoid $\lambda_{11} < 1$).

One-point composition control (one loop in manual). Most industrial columns have closed-loop control of only one composition. This may seem suboptimal, but is in many cases reasonable, since one product is usually much more important than the other. Furthermore, if the column is operating at its capacity limit (which is often the case), it is impossible to control more than one composition. Uncertainty does not pose any particular problem when only one composition is controlled. Reasonably good control of the uncontrolled composition is maintained provided D or B is not kept constant. The LV- and $\frac{L}{D} \frac{V}{B}$ -configurations will generally both perform satisfactory. The $\frac{L}{D} \frac{V}{B}$ - configuration is preferable because it has a better "build-in" rejection of flow-disturbances which leads to less variations in the uncontrolled composition. The only case when it may be worthwhile to use D or B as the manipulated input for one-point composition control, is for columns with very large reflux ($\frac{L}{D} \gg 1$ or $\frac{V}{B} \gg 1$) where level control using D or B may be almost impossible.

References

- Bristol, E. H., "On A New Measure of Interactions for Multivariable Process Control", *IEEE Trans. Automatic Control*, **AC-11**, 133-134 (1966).
- Dartt, S. R., "A Survey on Process Control Applications Needs", *Chemical Engineering Progress*, 11-14, (December 1985).
- Doyle, J. C., J. E. Wall and G. Stein, "Performance and Robustness Analysis for Structured Uncertainty", *IEEE Conf. on Decision and Control*, Orlando, FL (1982).
- McNeill, G. A. and J. D. Sacks, "High Performance Column Control", *Chemical Engineering Progress*, **65**, 3, 33-39 (1969).
- Rademaker, O., J. E. Rijnsdorp and A. Maarleveld, "*Dynamics and Control of Continuous Distillation Units*", Elsevier, Amsterdam (1975).
- Rosenbrock, H. H., "The Control of Distillation Columns", *Trans. Inst. Chem. Engrs.*, **40**, 35-53 (1962).
- Shinskey, F. G., *Distillation Control*, 2nd Edition, McGraw-Hill, New York (1984).
- Skogestad, S. and M. Morari, "Control of Ill-Conditioned Plants: High Purity Distillation", paper 74a, AIChE Annual Mtg., Miami Beach (1986a).
- Skogestad, S. and M. Morari, "Implication of Large RGA-Elements on Control Performance", paper 6d, AIChE Annual Mtg., Miami Beach (1986b).
- Skogestad, S. and M. Morari, "Understanding the Steady-State and Dynamic Behavior of Distillation Columns", in preparation (1986c).
- Skogestad, S. and M. Morari, "Effect of Disturbance Directions on Closed-Loop Performance", submitted to *Ind. Eng. Chem. Res.* (1986d).
- Stanley, G. T. and T. J. McAvoy, "Dynamic Energy Conservation Aspects of Distillation Control", *Ind. Eng. Chem. Fundam.*, **24**, 4, 439-443 (1985).

Appendix

Derivation of Table 3

Assuming constant molar flows the following exact steady-state relationships apply

$$dD = (1 - q_F)dF - Fdq_F + dV - dL \quad (A1)$$

$$dD = dF - dB \quad (A2)$$

Furthermore

$$d(D/B) = \frac{1}{B}dD - \frac{D}{B^2}dB \quad (A3)$$

combining (A2) and (A3)

$$d(D/B) = kdD - kD/FdF, k = F/B^2 \quad (A4)$$

We consider disturbances in F and q_F . In addition, each manipulated flow may have an additive disturbance. For example, the distillate flow D can be expressed

$$D = D_d + D_m \quad (A5)$$

Here D_m represents the “manipulated” part of the distillate (which is what D is “believed” to be), while D_d represents the disturbance. We want to find the effect of the disturbances on D/B when the composition loops are open, i.e.

$$du_{1m} = du_{2m} = 0 \quad (A6)$$

To derive Table 3 these equations are combined to express $d(D/B)$ in (A4) as a function of the disturbances only (i.e., express dD as a function of dD_d, dV_d, dL_d, dF , etc.).

Example. LV-Configuration

With

$$dL_m = dV_m = 0$$

(A1) becomes

$$dD = (1 - q_F)dF - Fdq_F + dV_d - dL_d$$

which upon inserting in (A4) yields

$$d(D/B) = k(1 - q_F F - D/F)dF + k(dV_d - dL_d - Fdq_F) \quad (A7)$$

Example. DX-Configuration

With

$$dD_m = 0$$

(A4) becomes

$$d(D/B) = kdD_d - kD/FdF \quad (A8)$$

Example. L/D V/B-Configuration

With

$$dL_m = L/D dD_m, dV_m = V/B dB_m$$

(A1) becomes

$$dD = (1 - q_F)dF - Fdq_F + dV_d - dL_d + V/B dB_m - L/D dD_m \quad (A9)$$

Here

$$dD_m = dD - dD_d$$

$$dB_m = dB - dB_d = dF - dD - dB_d$$

which upon inserting in (A9) yields

$$dD(1 + \frac{L}{D} + \frac{V}{B}) = (1 - q_F)dF - Fdq_F + dV_d - dL_d + \frac{V}{B}dF - \frac{V}{B}dB_d + \frac{L}{D}dD_d$$

The term involving dF drops out when this is substituted into (A4) (Use $(1 - q_F)F + V = D - L$) and we derive the expression given in Table 3:

$$d(\frac{D}{B})(1 + \frac{L}{D} + \frac{V}{B}) = k(dV_d - dL_d - Fdq_F + \frac{L}{D}dD_d - \frac{V}{B}dB_d) \quad (A10)$$

The fact that a change in F does not affect D/B when L/D and V/B are constant is expected, since a feed flow change is counteracted by keeping all flow ratios constant.

Chapter XIII

IV-CONTROL OF A HIGH-PURITY DISTILLATION COLUMN

LV-CONTROL OF A HIGH-PURITY DISTILLATION COLUMN

Sigurd Skogestad

Manfred Morari

California Institute of Technology

Chemical Engineering, 206-41

Pasadena, CA 91125

January 1987

submitted to Chem. Eng. Sci.

Abstract

A realistic study of the LV-control of a high-purity distillation column is presented. Linear controllers designed based on a linearized model of the plant are found to yield acceptable performance also when there is model-plant mismatch. The mismatch can be caused by uncertainty on the manipulated inputs, nonlinearity and variations in reboiler and condenser holdup. The presence of input uncertainty makes the use of a steady-state decoupler unacceptable. The effect of nonlinearity is strongly reduced by using the logarithm of the compositions. A simple diagonal PI-controller is not sensitive to model-plant mismatch, but yields a response with a sluggish return to steady-state.

1. INTRODUCTION

In this paper we study the high-purity distillation column in Table 1 using reflux (L) and boilup (V) as manipulated inputs to control the top (y_D) and bottom (x_B) compositions. This column was analyzed previously by the authors (Skogestad and Morari, 1986a), but the objective of that paper was to study general properties of ill-conditioned plants rather than distillation column control. The LV-configuration is chosen because this is the choice of manipulated inputs most commonly used in industrial practice. This does not necessarily mean that this is the best configuration, and, for example, the $\frac{L}{D} \frac{V}{B}$ -configuration may be preferable (Shinskey, 1984, Skogestad and Morari, 1987c).

The distillation column used in this paper was chosen to be representative of a large class of moderately high-purity distillation columns. The goal of this paper is to provide a realistic control design and simulation study for the column. To be realistic at least the issues of 1) uncertainty and 2) nonlinearity must be addressed.

1.1 Uncertainty

Skogestad and Morari (1986a) showed that the closed-loop system may be extremely sensitive to input uncertainty when the LV-configuration is used. In particular, inverse-based controllers were found to display severe robustness problems. In this paper the uncertainty is explicitly taken into account when designing and analyzing the controllers by using the Structured Singular Value (μ) introduced by Doyle (1982). We also find that μ provides a much easier way of comparing and analyzing the effect of various combinations of controllers, uncertainty and disturbances than the traditional simulation approach.

1.2 Nonlinearity

High-purity distillation columns are known to be strongly nonlinear (e.g. Moczek et al., 1963, Fuentes and Luyben, 1983), and any realistic study should take this into account. Our approach is to base the controller design on a linear

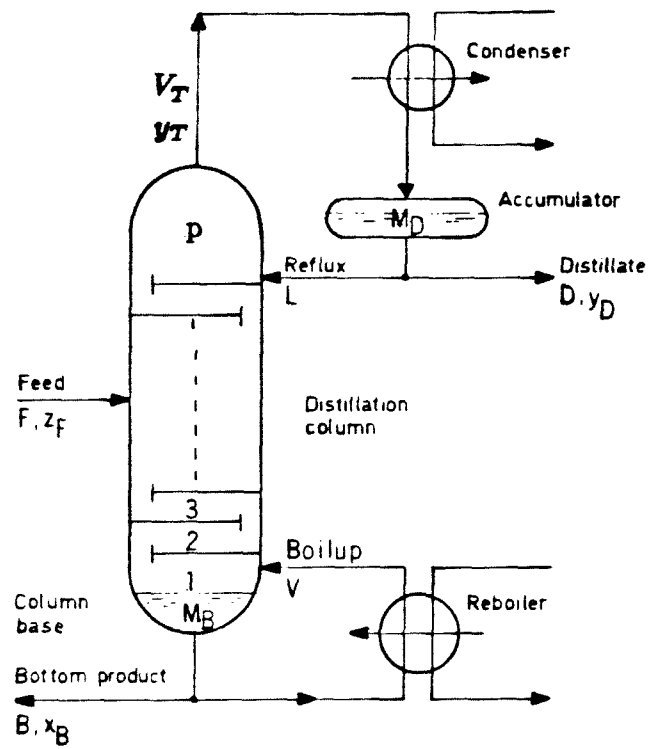


Figure 1. Two product distillation column with single feed and total condenser.

model. The effect of nonlinearity is taken care of by analyzing this controller for linearized models at different operating points. Furthermore, all simulations are based on the full nonlinear model.

1.3 Logarithmic Compositions

In another paper (Skogestad and Morari, 1987a) we study the dynamic behavior of distillation columns in general. One conclusion from that paper is that the high-frequency behavior is only weakly affected by operating conditions when the scaled transfer matrix is considered

$$\begin{pmatrix} dy_D^S \\ dx_B^S \end{pmatrix} = G^S \begin{pmatrix} dL \\ dV \end{pmatrix}, \quad G^S = \begin{pmatrix} \frac{1}{1-y_D^o} & 0 \\ 0 & \frac{1}{1-x_B^o} \end{pmatrix} G \quad (1)$$

All plant models and controllers in this paper are for the scaled plant. G^S is obtained by scaling the outputs with respect to the amount of impurity in each product

$$y_D^S = \frac{y_D}{1-y_D^o}, \quad x_B^S = \frac{x_B}{x_B^o} \quad (2)$$

Here x_B^o and y_D^o are the compositions at the nominal operating point. This relative scaling is automatically obtained by using logarithmic compositions

$$Y_D = \ln(1 - y_D) \quad (3)$$

$$X_B = \ln x_B$$

because

$$dY_D = -\frac{dy_D}{1-y_D}, \quad dX_B = \frac{dx_B}{x_B} \quad (4)$$

Furthermore, the use of logarithmic compositions (Y_D and X_B) effectively eliminates the effect of nonlinearity at high frequency (Skogestad and Morari, 1987a) and also reduces its effect at steady-state (Skogestad and Morari, 1987b). For control purposes the high frequency behavior (initial response) is of principal importance. Consequently, if logarithmic compositions are used we expect a linear controller to

perform satisfactorily also when we are far removed from the nominal operating point for which the controller was designed. Another objective of this paper is to confirm that this is indeed true.

In most cases the column is operated close to its nominal operating point and there is hardly any advantage in using logarithmic compositions which in this case merely corresponds to a rescaling of the outputs. However, if, for some reason, the column is taken far from this nominal operating point, for example, during startup or due to a temporary loss of control, the use of logarithmic compositions may bring the column safely back to its nominal operating point, whereas a controller based on unscaled compositions (y_D and x_B) may easily yield an unstable response.

1.4 Choice of Nominal Operating Point

The design approach suggested by the above discussion is to design a linear controller based on a linearized model for some nominal operating point. What operating point should be used? If an operating point corresponding to both products of high and equal purities is chosen (i.e., $1 - y_D = x_B$ is small), it is easily shown (Skogestad and Morari, 1987a,b, Kapoor et al., 1986) that the values of the steady-state gains and the linearized time constant will change drastically for small perturbations from this operating point. We may therefore question if acceptable closed-loop control can be obtained by basing the controller design on a linearized model at such an operating point. Kapoor et al. (1986) indicate that this is not advisable, and that a model based on a perturbed operating point should be used. However, as we just discussed, the high-frequency behavior, which is of primary importance for feedback control, shows much less variation with operating conditions. Therefore, provided the model gives a good description of the high-frequency behavior, we expect to be able to design an acceptable controller also when the nominal point has both products of high purity. This is also confirmed by the results in this paper.

A main conclusion of this paper is therefore that acceptable closed-loop performance may be obtained by designing a linear controller based on a linear model at any nominal operating point. If large perturbations from steady state are expected then logarithmic compositions should be used to reduce the effect of nonlinearity.

2. THE DISTILLATION COLUMN

Steady-state data for the distillation column are given in Table 1.

The following simplifying assumptions are made: a1) binary separation, a2) constant relative volatility, a3) constant molar flows and a4) constant holdups on all trays and perfect level control. The last assumption results in immediate flow response, that is, we are neglecting flow dynamics. This is somewhat unrealistic, and in order to avoid unrealistic controllers, we will add "uncertainty" at high frequency to include the effect of neglected flow dynamics when designing and analyzing the controllers (see Section 3).

We investigate the column at two different operating points. At the nominal operating point, A, both products are high-purity and $1 - y_D^o = x_B^o = 0.01$. Operating point C is obtained by increasing D/F from 0.500 to 0.555 which yields a less pure top product and a purer bottom product; $1 - y_{DC}^o = 0.10$ and $x_{BC}^o = 0.002$ (subscript C denotes operating point C while no subscript denotes operating point A). We will study the column for the following three assumptions regarding reboiler and condenser holdup

Case 1: Almost negligible condenser and reboiler holdup ($M_D/F = M_B/F = 0.5$ min).

Case 2: Large condenser and reboiler holdup ($M_D/F = 32.1$ min, $M_B/F = 11$ min).

Case 3: Same holdup as in Case 2, but the composition of the overhead vapor (y_T) is used as a controlled output instead of the composition in the condenser (y_D).

Binary separation, constant molar flows, feed liquid.

Column Data:

Relative Volatility	$\alpha = 1.5$
No. of theoretical trays	$N = 40$
Feed tray (1=reboiler)	$N_F = 21$
Feed composition	$z_F = 0.5$

Operating variables:

	A	C
$y_D =$	0.99	0.90
$x_B =$	0.01	0.002
$D/F =$	0.500	0.555
$L/F =$	2.706	2.737

Steady-state gains (unscaled compositions):

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G(0) \begin{pmatrix} dL/F \\ dV/F \end{pmatrix}$$

$$G(0) = \begin{pmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{pmatrix} \begin{matrix} \text{A} & \text{C} \\ \begin{pmatrix} 1.604 & -1.602 \\ .01865 & -.02148 \end{pmatrix} \end{matrix}$$

Table 1. Steady-state data for distillation column at operating points A and C.

These three cases will be denoted by subscripts 1, 2 and 3, respectively. The holdup on each tray inside the column is $M_i/F = 0.5$ min in all three cases.

2.1 Modelling

Nominal operating point (A). A 41st order linear model for the columns is easily derived based on the data given in Table 1 (see Skogestad and Morari, 1987a)

$$\begin{pmatrix} dy_D \\ dx_B \end{pmatrix} = G(s) \begin{pmatrix} dL \\ dV \end{pmatrix} \quad (5)$$

The scaled steady-state gain matrix is

$$G^S(0) = \begin{bmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{bmatrix} \quad (6)$$

which yields the following values for the condition number and the 1,1-element in the RGA

$$\gamma(G^S(0)) = \bar{\sigma}(G^S(0))/\underline{\sigma}(G^S(0)) = 141.7 \quad \lambda_{11}(G^S(0)) = 35.1$$

However, $\gamma(G^S)$ and $\lambda_{11}(G^S)$ are much smaller at high frequencies as seen from Fig. 2. A very crude model of the column was presented by Skogestad and Morari (1986a) (time in minutes)

$$\text{Model 0: } G(s) = \frac{1}{1 + 75s} G(0) \quad (7)$$

This model gives the same values of $\gamma(G)$ and $\lambda_{11}(G)$ at all frequencies, and is therefore a poor description of the actual plant at high frequency. In our previous study (Skogestad and Morari, 1986a) the controller design was based on this simplified model, and one objective of this paper is to study how these controllers perform when a more realistic model is used.

Case 1. For the case of negligible reboiler and condenser holdup the following simple two time-constant model yields an excellent approximation of the 41st order linear model (Skogestad and Morari, 1987a).

$$\text{Model 1: } G_1^S(s) = \begin{pmatrix} \frac{87.8}{1+\tau_1 s} & -\frac{87.8}{1+\tau_1 s} + \frac{1.4}{1+\tau_2 s} \\ \frac{108.2}{1+\tau_1 s} & -\frac{108.2}{1+\tau_1 s} - \frac{1.4}{1+\tau_2 s} \end{pmatrix} \quad \begin{matrix} \tau_1 = 194 \text{ min} \\ \tau_2 = 15 \text{ min} \end{matrix} \quad (8)$$

This model has only two states as seen from the minimal realization in the Appendix. $G_1(s)$ uses two time constants: τ_1 is the time constant for changes in the external flows. It corresponds to the dominant time constant and may be estimated, for example, by using the inventory time constant of Moczek et al. (1963). τ_2 is the time constant for changes in internal flows (simultaneous change in L and V with constant product rates, D and B) and can be estimated by matching the high-frequency behavior as shown by Skogestad and Morari (1987a). The simple model (8) matches the observed variation in condition number with frequency (Fig.2).

The effect of the reboiler and condenser holdups (Case 2) can be partially accounted for with Model 1 by multiplying $G_1(s)$ by $diag\{(1+\tau_D s)^{-1}, (1+\tau_B s)^{-1}\}$, where in our case $\tau_D = M_D/V_T = 10$ min and $\tau_B = M_B/L_B = 3$ min. However, in practice the top composition is often measured in the overhead vapor line (Case 3), rather than in the condenser. $G_1(s)$ provides a good approximation of the plant in such cases.

Cases 2 and 3. In order to obtain a low-order model for Case 2 and 3, we performed a model reduction (Balanced Realization, Moore (1981)) on the full 41st order model. A good approximation was obtained with a 5th order model as illustrated in Fig. 3. The state-space realizations of these models ($G_2^S(s)$ and $G_3^S(s)$) are given in Appendix.

Operating point C. We will return with a discussion of the model for this case in Section 6 when we also discuss the control of the plant.

2.2 Simulations

The design and analysis of the controller are based on the linear models $G_1(s)$, $G_2(s)$ and $G_3(s)$. However, except for the four simplifying assumptions a1-a4 stated above, all simulations are carried out with the full nonlinear model. (In some cases the changes are so small, however, that the results are equivalent to linear simulations.) To get a realistic evaluation of the controllers input uncertainty must

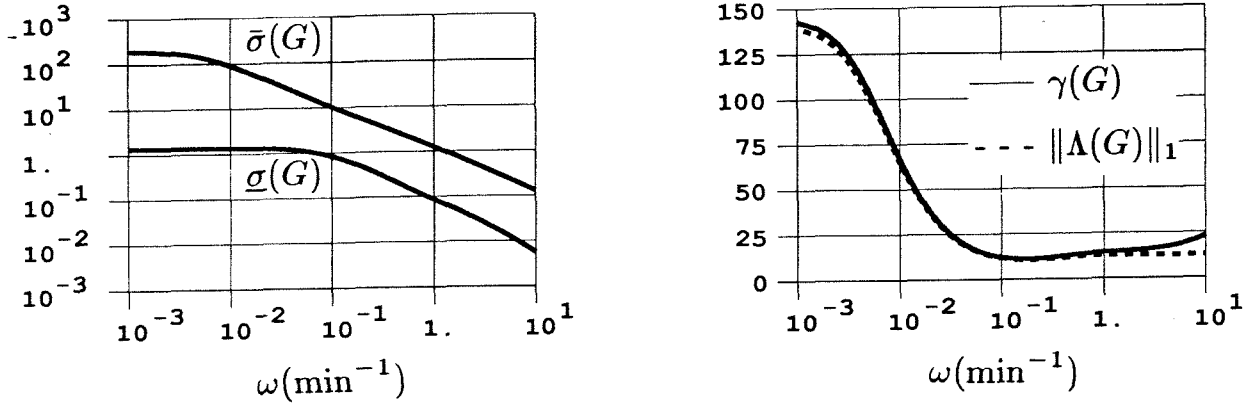


Figure 2. Column A, Case 1 ($G = G_1^S$). The condition number of the plant is about 10 times lower at high frequencies than at steady state.

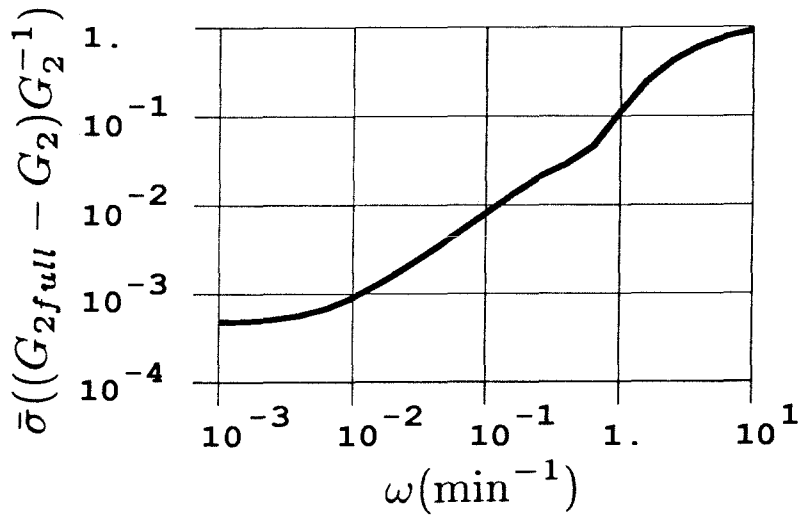


Figure 3. Column A, Case 2. Relative difference between the 5th order model $G_2(s)$ (Appendix) and the 41st order plant $G_{2full}(s)$. The 5th order model provides an excellent approximation within the frequency range of interest ($\omega \leq 1 \text{ min}^{-1}$).

be included (Skogestad and Morari, 1986a,b). Simulations are therefore shown both with and without 20% uncertainty with respect to the change of the two inputs. The following uncertainties are used

$$\begin{aligned}\Delta L &= (1 + \Delta_1)\Delta L_c, \quad \Delta_1 = 0.2 \\ \Delta V &= (1 + \Delta_2)\Delta V_c, \quad \Delta_2 = -0.2\end{aligned}\tag{9}$$

Here ΔL and ΔV are the actual changes in manipulated flow rates, while ΔL_c and ΔV_c are the desired values as computed by the controller. $\Delta_1 = -\Delta_2$ was chosen to represent the worst combination of the uncertainties (Skogestad and Morari, 1986b).

3. CONTROL THEORY

3.1 Robust performance and robust stability

The objective of using feedback control is to keep the controlled outputs (in our case y_D and x_B) "close" to their desired setpoints. What is meant by "close" is more precisely defined by the performance specifications. These performance requirements should be satisfied in spite of unmeasured disturbances and model-plant mismatch (uncertainty). Consequently, the ultimate goal of the controller design is to achieve Robust Performance (RP): The performance specification should be satisfied for the worst case combination of disturbances and model-plant mismatch.

To check for RP we will use the Structured Singular Value μ (Doyle, 1982). μ of a matrix N (denoted $\mu(N)$ or $\mu_\Delta(N)$) is equal to $1/\bar{\sigma}(\Delta)$ where $\bar{\sigma}(\Delta)$ is the magnitude of the smallest perturbation needed to make the matrix $(I + \Delta N)$ singular. $\mu(N)$ depends both on the matrix N and of the structure (e.g., diagonal or full matrix) of the perturbation Δ .

As stated, achieving robust performance is the overall goal. The implications of this requirement are easier to understand if we consider some subobjectives which have to be satisfied in order to achieve this goal:

Nominal Stability (NS): The model is assumed to be a reasonable approximation of the true plant. Therefore the closed loop system with the controller applied to the (nominal) plant model has to be stable.

Nominal Performance (NP): In addition to stability, the quality of the response should satisfy some minimum requirements – at least when the controller is applied to the plant model. We will define performance in terms of the weighted H^∞ -norm of the closed-loop transfer function S from the disturbances (d) and setpoints (y_s) to the errors ($e = y - y_s$, i.e., $y_D - y_{D,s}$, $x_B - x_{B,s}$). The performance specification is

$$NP \Leftrightarrow \bar{\sigma}(w_P S) \leq 1 \quad \forall \omega, \quad S = (I + GC)^{-1} \quad (10)$$

The weight w_P is used to specify the frequency range over which the errors are to be small. To get consistency with the notation used below define $\bar{\sigma}(w_P S) = \mu(N_{NP})$ such that (10) becomes

$$NP \Leftrightarrow \mu(N_{NP}) \leq 1 \quad \forall \omega \quad (11)$$

where $N_{NP} = w_P S$, and μ is computed with respect to the structure of a “full” matrix Δ_P .

Robust Stability (RS). The closed loop system must remain stable for all possible plants as defined by the uncertainty description. For example, assume there is uncertainty with respect to the actual magnitude of the manipulated inputs (which is always the case!). The possible plants, G_p , are then given by

$$G_p = G(I + \Delta_I), \quad \Delta_I = \begin{pmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{pmatrix} \quad (12)$$

where $\Delta_i(s)$ is the uncertainty for input i . We will consider the case when the magnitude of uncertainty is equal for both inputs

$$|\Delta_i| \leq |w_I(j\omega)|, \quad i = 1, 2 \quad (13)$$

The robust stability requirement can be checked using μ . In this particular case (Skogestad and Morari, 1986a)

$$RS \Leftrightarrow \mu(N_{RS}) \leq 1, \quad \forall \omega \quad (14)$$

where $N_{RS} = w_I C G S$ and μ is computed with respect to the diagonal 2×2 matrix Δ_I .

Robust Performance (RP): The closed loop system must satisfy the performance requirements for all possible plants as defined by the uncertainty description. As an example we may require (10) to be satisfied when G is replaced by any of the possible perturbed plants G_p as defined by the uncertainty description (12).

$$RP \Leftrightarrow \bar{\sigma}(w_P(I + G_p C)^{-1}) \leq 1 \quad \forall \omega, \quad \forall G_p \quad (15)$$

This definition of Robust Performance is of no value without a simple method to test if condition (15) is satisfied for all possible perturbed plants G_p generated by (12) and (13). Again it turns out that the structured singular value μ gives a condition which is relatively easy to check:

$$RP \Leftrightarrow \mu(N_{RP}) \leq 1, \quad \forall \omega \quad (16a)$$

where

$$N_{RP} = \begin{pmatrix} w_I C S G & w_I C S \\ w_P S G & w_P S \end{pmatrix} \quad (16b)$$

and μ is computed with respect to the structure $diag\{\Delta_I, \Delta_P\}$ where Δ_I is 2×2 diagonal matrix and Δ_P is a full 2×2 matrix.

3.2 The RGA

Let \times denote element-by-element multiplication. The RGA of the matrix G (Bristol, 1966) is defined as

$$\Lambda(G) = G \times (G^{-1})^T \quad (17)$$

The RGA is independent of input and output scaling. The RGA of the plant is commonly used as a tool for selecting control configurations for distillation columns (Shinskey, 1984). However, in this paper we will make use of the RGA of the controller as a measure of a system's sensitivity to input uncertainty (Skogestad and Morari, 1986b). Before stating this result, we will point out the close relationship between large plant RGA-elements and a high condition number. The condition number of the plant is $\gamma(G) = \bar{\sigma}(G)/\underline{\sigma}(G)$. This quantity is strongly dependent on how the inputs and outputs are scaled. The minimized scaled condition number $\gamma^*(G)$ is obtained by minimizing $\gamma(S_1 G S_2)$ over all possible input and output scalings, S_1 and S_2 . There is a very close relationship between γ^* and $\|\Lambda\|_1 = \sum_{i,j} |\lambda_{ij}|$. For 2×2 plants (Nett et al., 1986, Grosdidier et al., 1985)

$$\|\Lambda\|_1 - \frac{1}{\gamma^*(G)} \leq \gamma^*(G) \leq \|\Lambda\|_1 \quad (18)$$

Consequently, for 2×2 plants the difference between these quantities is at most one and $\|\Lambda\|_1$ approaches $\gamma^*(G)$ as $\gamma^*(G) \rightarrow \infty$. Since $\|\Lambda\|_1$ is much easier to compute than $\gamma^*(G)$, it is the preferred quantity to use.

The RGA and input uncertainty (Skogestad and Morari, 1986b). Again, consider uncertainty on the plant inputs as given by (12). The loop transfer matrix, $G_p C$, for the perturbed plant may be written in terms of its nominal value, GC :

$$G_p C = GC(I + C^{-1} \Delta_I C) \quad (19)$$

$G_p C$ is closely related to performance because of (15). For 2×2 plants the error term $C^{-1} \Delta_I C$ in (19) may be expressed in terms of the RGA of the controller

$$G^{-1} \Delta_I C = \begin{bmatrix} \lambda_{11}(C) \Delta_1 + \lambda_{21}(C) \Delta_2 & \lambda_{11}(C) \frac{c_{12}}{c_{11}} (\Delta_1 - \Delta_2) \\ -\lambda_{11}(C) \frac{c_{21}}{c_{22}} (\Delta_1 - \Delta_2) & \lambda_{12}(C) \Delta_1 + \lambda_{22}(C) \Delta_2 \end{bmatrix} \quad (20)$$

If any element in $C^{-1} \Delta_I C$ is large compared to 1, the loop transfer matrix $G_p C$ is

very different from the nominal (GC) and poor performance or even instability is expected when $\Delta_I \neq 0$. We see from (20) that controllers with large RGA-elements should always be avoided, because otherwise the closed-loop system is very sensitive to input uncertainty.

It should be added that it is the behavior of $G_p C$ at frequencies close to the closed-loop bandwidth (where $\sigma_i(G_p C) \approx 1$) which is of primary importance for the stability of the closed-loops system. Therefore, it is particularly bad if the controller has large RGA-elements in this frequency range.

Inverse-Based Controller. To have “tight” control it is desirable to use an inverse-based controller $C(s) = c(s)G^{-1}(s)$ where $c(s)$ is a scalar. In this case $\Lambda(C) = \Lambda(G^{-1}) = \Lambda^T(G)$ and the controller will have large RGA-elements whenever the plant has. Consequently, inverse-based controllers should always be avoided for plants with large RGA-elements. In particular, this applies to LV-control of high-purity distillation columns which always yields large RGA-elements.

Control of Plants with Large RGA-Elements. We clearly should not use an inverse-based controller for a plant with large RGA-elements. On the other hand, a diagonal controller is insensitive to uncertainty ($C^{-1}\Delta_I C = \Delta_I$), but is not able to correct for the strong directionality of the plant, which implies that performance has to be sacrificed. This is confirmed by the results presented below.

4. FORMULATION OF THE CONTROL PROBLEM

4.1 Performance and Uncertainty Specifications

The uncertainty and performance specifications are the same as those used by Skogestad and Morari (1986a).

Uncertainty. The only source of uncertainty considered is uncertainty on the manipulated inputs (L and V) with a magnitude bound

$$w_I(s) = 0.2 \frac{5s + 1}{0.5s + 1} \quad (21)$$

The possible perturbed plants G_p are obtained by allowing any $dL = dL_c(1 \pm |w_I|)$ and $dV = dV_c(1 \pm |w_I|)$. (Actually, the perturbations are allowed to be complex, mainly for mathematical convenience). (21) allows for an input error of up to 20% at low frequency as is used in the simulations (9). The uncertainty in (21) increases with frequency. This allows, for example, for a time delay of about 1 min in the response between the inputs, L and V, and the outputs, y_D and x_B . In practice, such delays may be caused by the flow dynamics. Therefore, although flow dynamics are not included in the models or in the simulations, they are partially accounted for in the μ -analysis and in the controller design.

Performance. Robust performance is satisfied if

$$\bar{\sigma}(S_p) = \bar{\sigma}((I + G_p C)^{-1}) \leq \frac{1}{|w_p|} \quad (15)$$

is satisfied for all possible plants, G_p . We use the performance weight

$$w_p(s) = 0.5 \frac{10s + 1}{10s} \quad (22)$$

A particular S which exactly matches the bound (15) at low frequencies and satisfies it easily at high frequencies is $S = 20s/20s + 1$. This corresponds to a first-order response with closed-loop time constant 20 min.

4.2 Analysis of Controllers

Comparison of controllers is based mainly on computing μ for robust performance (μ_{RP}). Simulations are used only to support conclusions found using the μ -analysis. The main advantage of using the μ -analysis is that it provides a well-defined basis for comparison. On the other hand, simulations are strongly dependent on the choice of setpoints, uncertainty, etc.

The value of μ_{RP} is indicative of the worst-case response. If $\mu_{RP} > 1$ then the “worst case” does not satisfy our performance objective, and if $\mu_{RP} < 1$ then the “worst case” is better than required by our performance objective. Similarly, if $\mu_{NP} < 1$ then the performance objective is satisfied for the nominal case. However,

this may not mean very much if the system is sensitive to uncertainty and μ_{RP} is significantly larger than one. We will show below that this is the case, for example, if an inverse-based controller is used for our distillation column.

4.3 Controllers

We will study the distillation column using the following six controllers:

- 1) Diagonal PI-controller.

$$C_{PI}(s) = \frac{0.01}{s}(1 + 75s) \begin{pmatrix} 2.4 & 0 \\ 0 & -2.4 \end{pmatrix} \quad (23)$$

This controlled was studied in Skogestad and Morari (1986a) and it was tuned in order to achieve as good a performance as possible while maintaining robust stability (also see Fig. 6).

- 2) Steady-state decoupler plus two PI-controllers.

$$C_{oiv}(s) = 0.7 \frac{(1 + 75s)}{s} G^S(0)^{-1} = \frac{0.01(1 + 75s)}{s} \begin{pmatrix} 27.96 & -22.04 \\ 27.60 & -22.40 \end{pmatrix} \quad (24)$$

This controller was tuned to achieve good nominal performance. However, the controller has large RGA-elements ($\lambda_{11}(C) = 35.1$) at all frequencies and we expect the controller to be extremely sensitive to input uncertainty.

- 3) Inverse-based controller based on the linear model $G_1^S(s)$ for Case 1.

$$C_{inv}(s) = \frac{0.7}{s} G_1^S(s)^{-1} \quad (25)$$

At low frequency this controller is equal to $C_{oiv}(s)$. Note that $C_{inv}(s)$ and $G_1^S(s)^T$ have the same RGA-elements. Therefore from Fig. 1 we expect $C_{inv}(s)$ to be sensitive to input uncertainty at low frequency, but not at high frequency.

- 4, 5 and 6). μ -optimal controllers based on the models $G_0(s)$, $G_1(s)$ and $G_2(s)$.

The controllers are denoted $C_{0\mu}(s)$, $C_{1\mu}(s)$ and $C_{2\mu}(s)$, respectively, and their state-space descriptions are given in the Appendix.

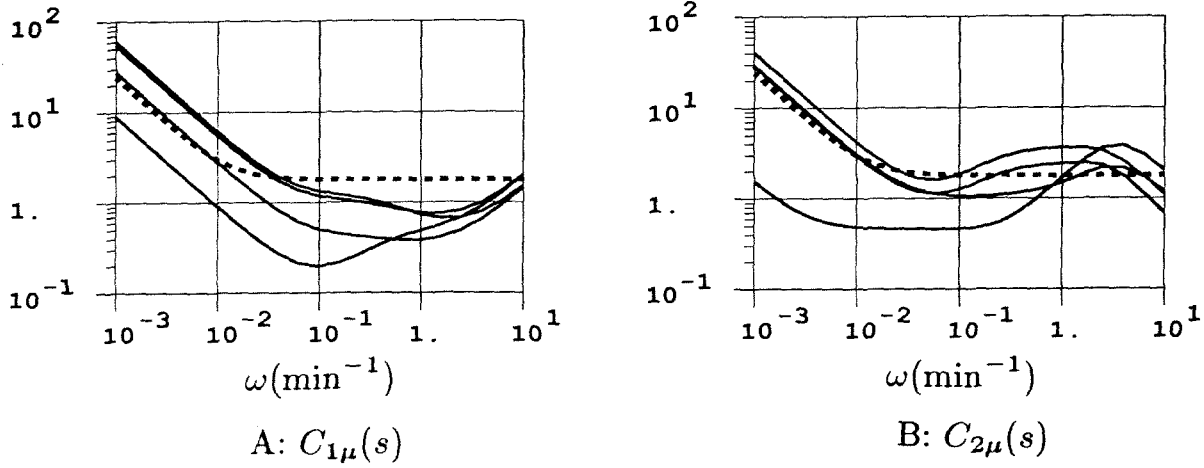


Figure 4. Magnitude plots of elements in μ -optimal controllers $C_{1\mu}(s)$ and $C_{2\mu}(s)$. Dotted line: $C_{PI}(s)$.

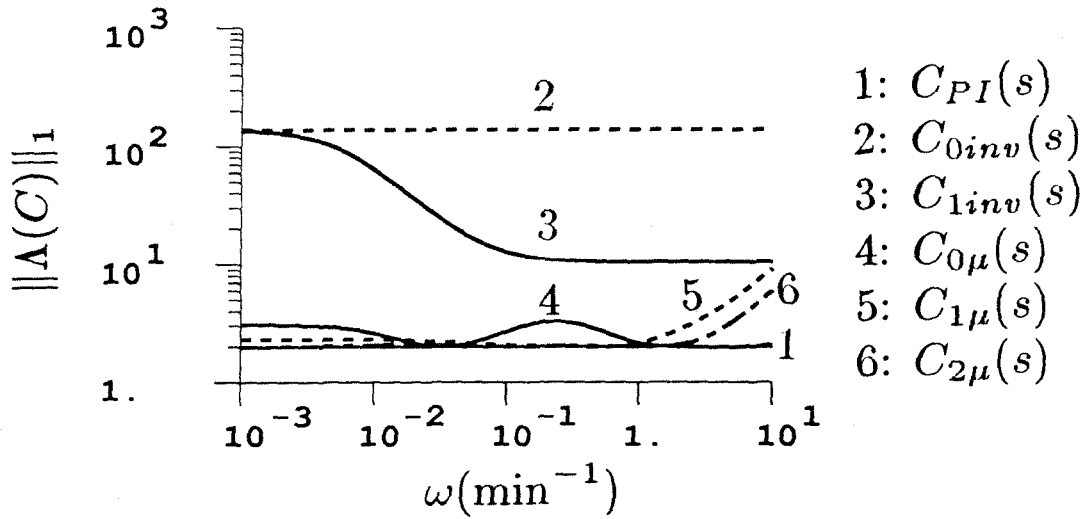


Figure 5. Magnitude of RGA-elements of controllers. $\|\Lambda(C)\|_1 = \sum_{i,j} |\lambda_{ij}(C)|$.

These controllers were obtained by minimizing $\sup_w \mu(N_{RP})$ for each model using the input uncertainty and performance weights given above. The numerical procedure used for the minimization is the same as outlined in Skogestad and Morari (1986a). The μ -plots for RP for the μ -optimal controllers are of particular interest since they indicate the best achievable performance for the plant. Bode-plots of the transfer matrix elements for $C_{1\mu}(s)$ and $C_{2\mu}(s)$ are shown in Fig. 4. Note the similarities between these controllers and the simple diagonal PI-controller (23).

At low frequency ($s \rightarrow 0$) the six controllers are approximately

$$\begin{aligned} C_{PI} &= \frac{0.01}{s} \begin{pmatrix} 2.4 & 0 \\ 0 & 2.4 \end{pmatrix} \\ C_{0inv} = C_{1inv} &= \frac{0.01}{s} \begin{pmatrix} 27.96 & -22.04 \\ 27.80 & -22.40 \end{pmatrix} \\ C_{0\mu} &= \frac{0.01}{s} \begin{pmatrix} 3.82 & -0.92 \\ 1.73 & -3.52 \end{pmatrix} \\ C_{1\mu} &= \frac{0.01}{s} \begin{pmatrix} 6.07 & -0.90 \\ 2.80 & -2.93 \end{pmatrix} \\ C_{2\mu} &= \frac{0.01}{s} \begin{pmatrix} 4.06 & +0.15 \\ 2.85 & -2.93 \end{pmatrix} \end{aligned}$$

The value of $\|\Lambda(C)\|_1$ as a function of frequency is shown for the six controllers in Fig. 5. As expected, the μ -optimal controllers have small RGA-elements, which make them insensitive to the input uncertainty. For example, $C_{2\mu}$ is nearly triangular at low frequency and consequently has $\Lambda \approx I$.

5. RESULTS FOR OPERATING POINT A

In this section we will study how the six controllers perform at the nominal operating point A for the three assumptions regarding condenser and reboiler holdup (corresponding to the models $G_1(s)$, $G_2(s)$ and $G_3(s)$). The μ -plots for the 18 possible combinations are given in Fig. 6. The upper solid line is $\mu(N_{RP})$ computed from (16). The lower solid line is $\mu(N_{NP}) = \bar{\sigma}((I + GC)^{-1})$. The dotted line is $\mu(N_{RS})$ (Eq. (14)). A number of interesting observations can be derived from these

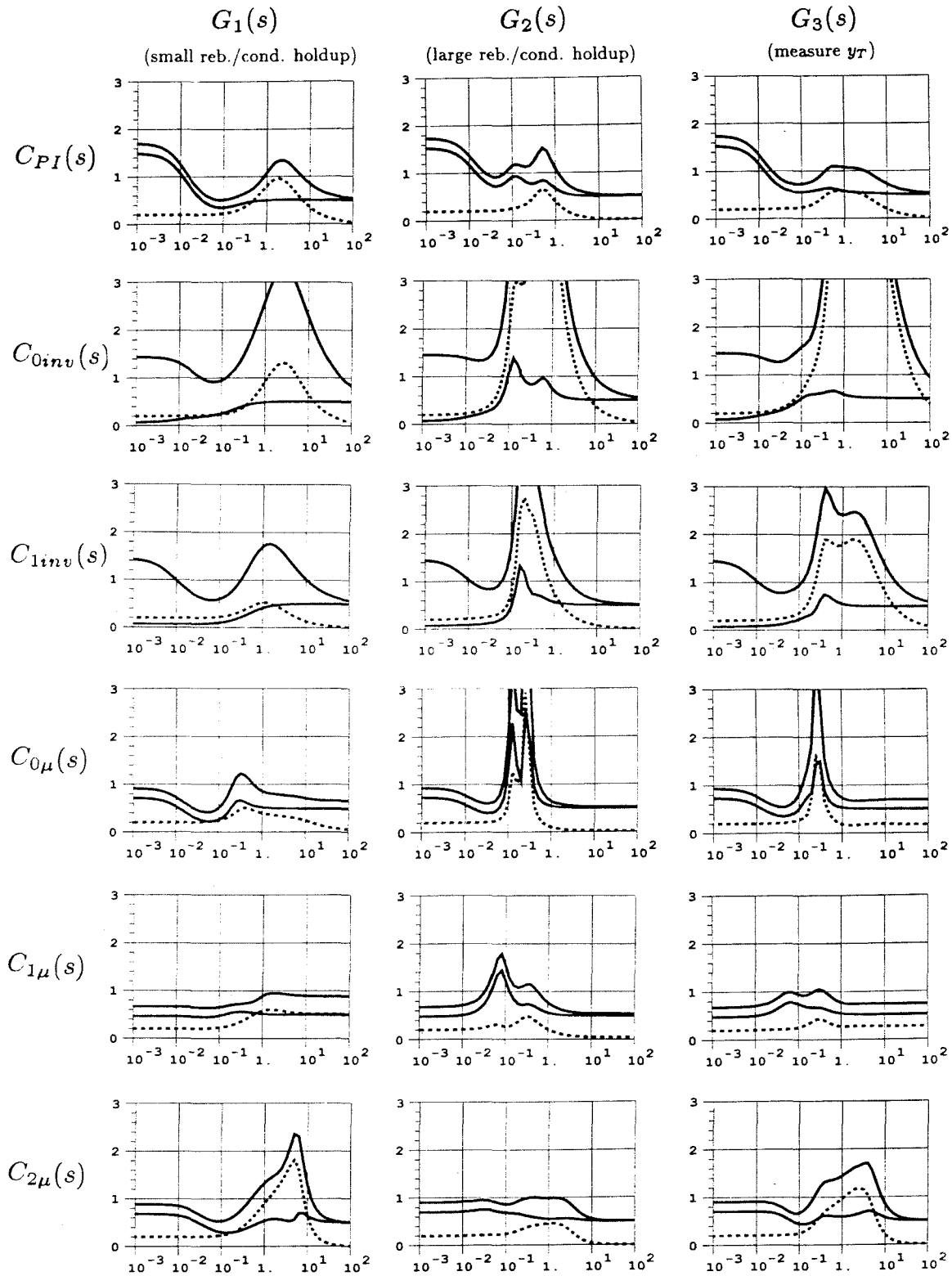


Figure 6. μ -plots for Column A. Upper solid line: $\mu(N_{RP})$ for robust performance; Lower solid line: $\mu(N_{NP})$ for nominal performance; Dotted line: $\mu(N_{RS})$ for robust stability. The RP-, NP- or RS-requirement is satisfied if the corresponding μ -curve is less than one at all frequencies.

plots. These are presented below. In some cases the simulations in Fig. 7-9 are used to support the claims.

Discussion of Controllers

$C_{PI}(s)$. The simple diagonal PI-controller performs reasonably well in all cases. μ_{NP} is higher than one at low frequency, which indicates a slow return to steady-state. This is confirmed by the simulations in Fig. 8 for a feed rate disturbance; after 200 min the column has still not settled. Operators are usually unhappy about this kind of response. The controller is insensitive to input uncertainty and to changes in reboiler and condenser holdup.

$C_{0inv}(s)$. This controller uses a steady-state decoupler. The nominal response is very good for Case 1 (Fig. 7), but the controller is extremely sensitive to input uncertainty. In practice, this controller will yield an unstable system (Skogestad and Morari, 1986a).

$C_{1inv}(s)$. This controller gives an excellent nominal response for Case 1 (Fig. 6). This is also confirmed by the simulations in Fig. 7; the response is almost perfectly decoupled with a time constant of about 1.4 min. Since the simulations are performed with the full-order model, while the controller was designed based on the simple two time-constant model, $G_1(s)$ (8), this confirms that $G_1(s)$ yields a very good approximation of the linearized plant when the reboiler and condenser holdups are small. The controller is sensitive to the input uncertainty as expected from the RGA- analysis. Also note that the controller performs very poorly when the condenser and reboiler holdups are increased. This shows that the controller is very sensitive also to other sources of model-plant mismatch.

$C_{0\mu}(s)$. This is the μ -optimal controller from our previous study (Skogestad and Morari, 1986a) which was designed based on the very simplified model $G_0(s)$. The controller performs surprisingly well on the actual plant ($G_1(s)$) when the holdups are negligible. However, the controller is seen to perform very poorly

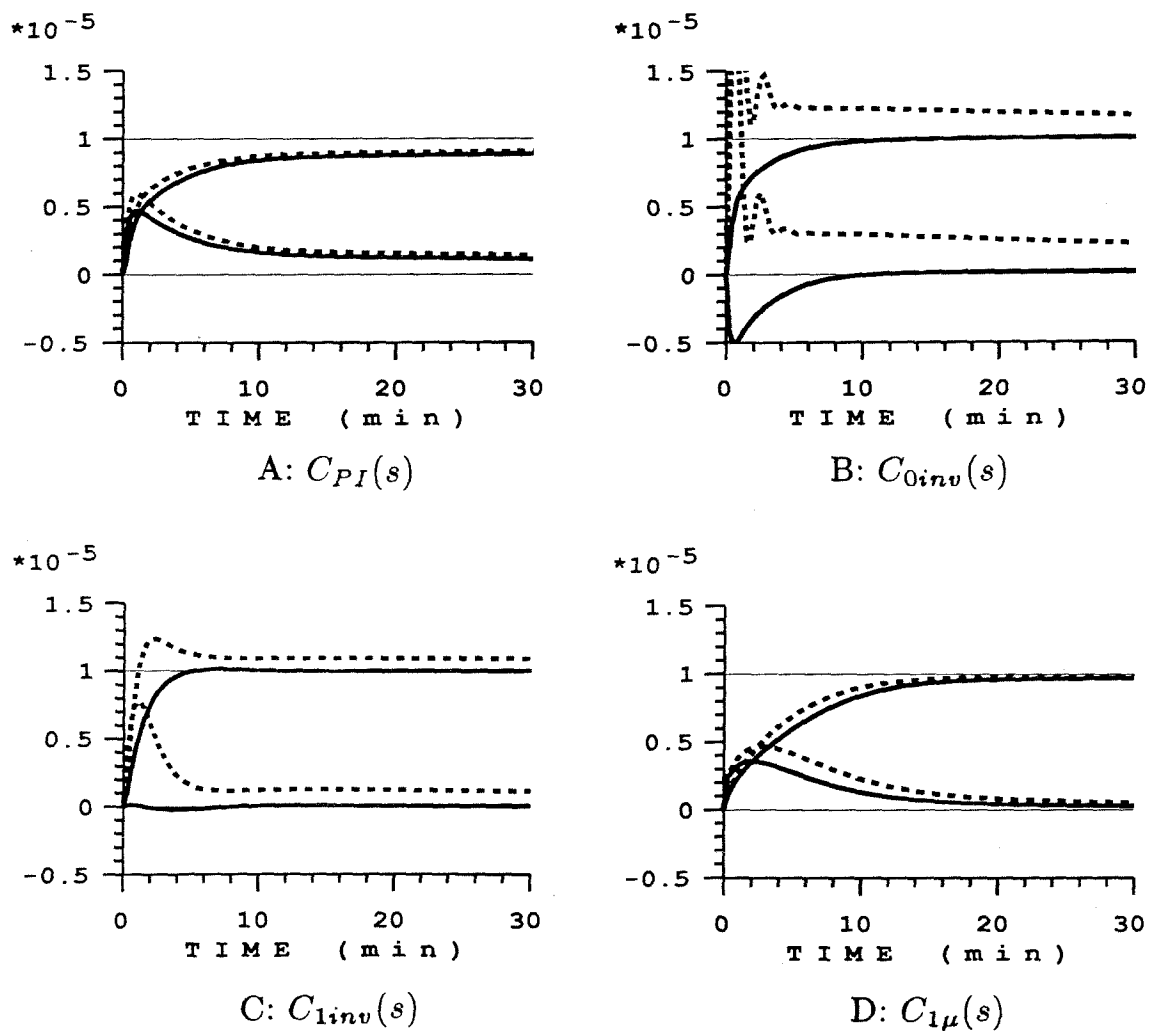


Figure 7. Column A, Case 1. Closed-loop response to small setpoint change in y_D . Solid lines: no uncertainty; Dotted lines: 20% uncertainty on inputs L and V (Eq. 9).

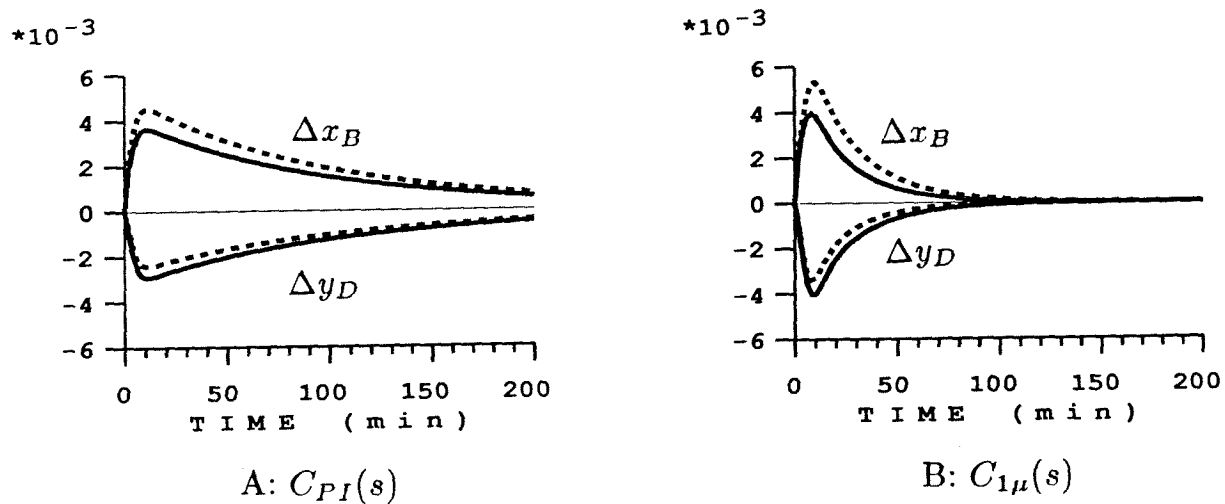


Figure 8. Column A, Case 1. Closed-loop response to a 30% increase in feed rate. Solid lines: no uncertainty; Dotted lines: 20% uncertainty on inputs L and V (Eq. 9).

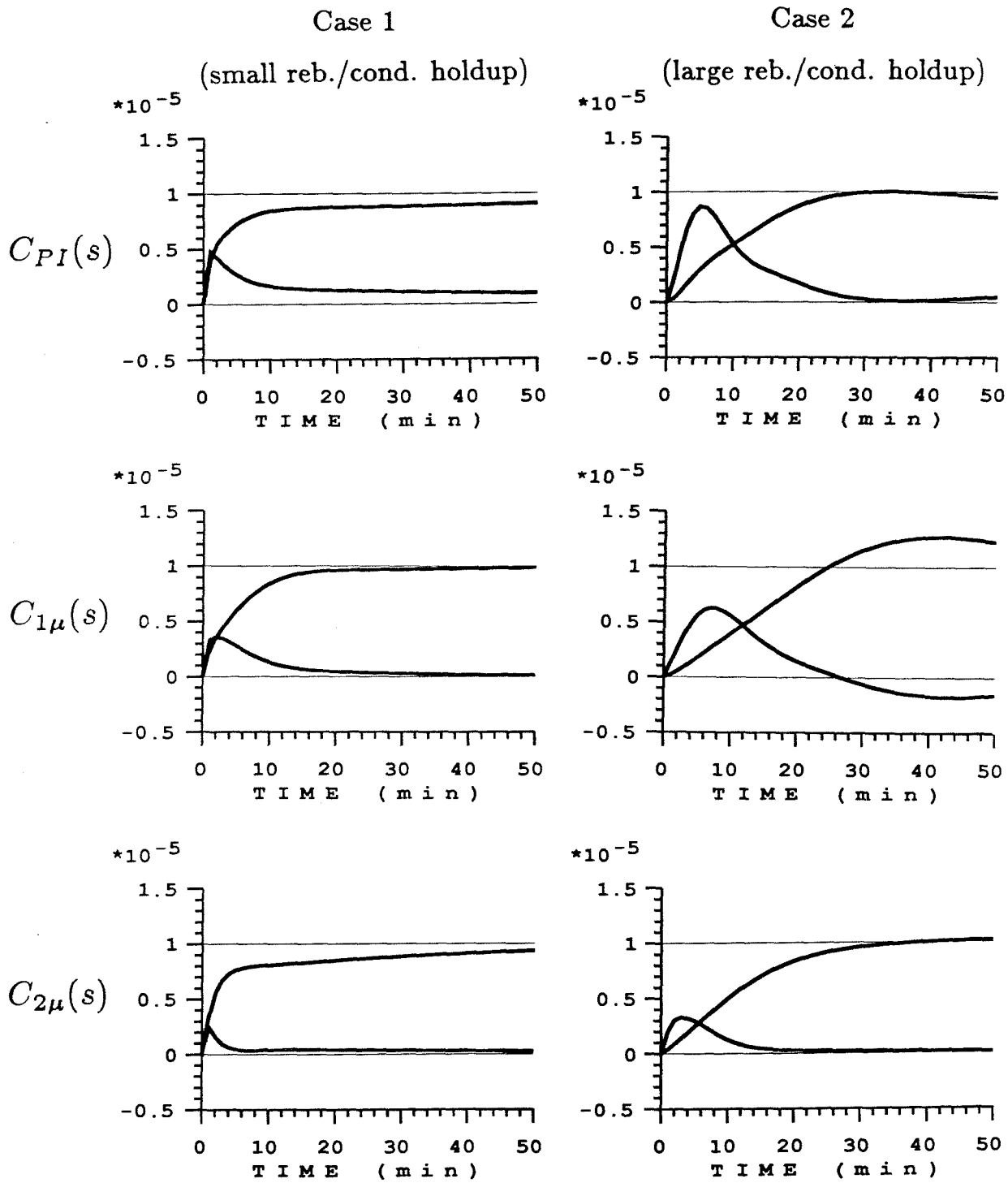


Figure 9. Column A. Effect of reboiler and condenser holdup on closed-loop response.

No uncertainty.

when the holdup in the reboiler and condenser is increased, which shows that the controller is very sensitive to other sources of model inaccuracies (for which it was not designed).

$C_{1\mu}(s)$. This is the μ -optimal controller when there is negligible holdup ($G_1(s)$), and the RP-condition is satisfied for this case since $\mu_{RP} \approx 0.95$. The nominal performance is not as good as for the inverse-based controller $C_{inv}(s)$; we have to sacrifice nominal performance to make the system robust with respect to uncertainty. The controller shows some performance deterioration when the reboiler and condenser holdups are increased (Case 2). This is not surprising since the added holdup makes the response in y_D and x_B more sluggish; the open-loop response for y_D changes from approximately $1/(1 + 194s)$ to $1/(1 + 194s)(1 + 10s)$ (recall discussion following (8)). As expected, the controller is much less sensitive to changes in condenser holdup if overhead composition is measured in the vapor line (Case 3). Overall, this is the best of the six controllers.

$C_{2\mu}(s)$. This is the μ -optimal controller for the case with considerable reboiler and condenser holdup, and with y_D measured in the condenser ($G_2(s)$). $\mu_{RP} \approx 1.00$ for this case. The nominal response is good in all cases (Fig. 6), but the controller is very sensitive to uncertainty when the plant is $G_1(s)$ or $G_3(s)$ rather than $G_2(s)$. This is clearly not desirable since changes in condenser and reboiler holdup are likely to occur during normal operation. The observed behavior is not surprising since the controller includes lead elements at $\omega \approx 0.1$ (Fig.4B) to counteract the lags caused by the reboiler and condenser holdups. If these lags are not present in the plant ($G_1(s)$ or $G_3(s)$), the “derivative” action caused by the lead elements result in a system which is very sensitive to uncertainty.

Conclusions

- The μ -optimal controller $C_{0\mu}(s)$ for the plant $G_0(s)$ has $\mu_{RP} \approx 1.06$ (Skogestad and Morari, 1986a), while the μ -optimal controller $C_{1\mu}(s)$ for the plant $G_1(s)$

has $\mu_{RP} \approx 0.95$. Thus, somewhat surprisingly, the achievable performance is not much better for $G_1(s)$ than for $G_0(s)$, even though $G_0(s)$ is ill-conditioned and has large RGA-elements at all frequencies, while $G_1(s)$ only has large RGA- elements at low frequencies (Fig. 1). This seems to indicate that large RGA- elements at low frequency imply limitations on the achievable control performance and partially justifies the use of steady-state values of the RGA for selecting the best control configuration (Shinskey, 1984).

- However, the use of the more detailed model $G_1(s)$, rather than $G_0(s)$, is still justified since the resulting μ -optimal controller is much less sensitive to changes in reboiler and condenser holdup (which will occur during operation).
- $G_1(s)$ approximates the full-order model very closely as seen from Fig. 7C; the response is almost perfectly decoupled when there is no uncertainty.
- To avoid sensitivity to the amount of condenser and reboiler holdup, the overhead composition should be measured in the overhead vapor, rather than in the condenser. In practice, temperature measurements inside the column are often used to infer compositions, and the dynamic response of these measurements is similar to that when the condenser and reboiler holdup is neglected.
- The simple model $G_2(s)$ is useful for controller design also when the reboiler and condenser holdup is large.
- The main advantage of the μ -optimal controllers over the simple diagonal PI-controller is a faster return to steady-state. This comes out very clearly in Fig. 8 which shows the closed-loop response to a 30% increase in feed rate.

6. EFFECT OF NONLINEARITY (RESULTS FOR OPERATING POINT C)

In this paper we do not treat nonlinearity as uncertainty as was attempted in Skogestad and Morari (1986a). The reason is that this approach is not rigorous and is also easily very conservative because of the strong correlation between all the

parameters in the model which is difficult to account for. Furthermore, we know from the insights presented by Skogestad and Morari (1987a) that the column is actually not as nonlinear as one might expect. Though the steady-state gains may change dramatically, the initial response (the high frequency behavior), which is of principal importance for feedback control, is much less affected. In particular, this is the case if relative (logarithmic) compositions are used (Skogestad and Morari, 1987a). To demonstrate this we compute μ and show simulations for some of the controllers when the "plant" is $G_C^S(s)$ rather than $G^S(s)$.

6.1 Modelling

$G_C(s)$ corresponds to the same column as $G(s)$, but the distillate flow rate ($\frac{D}{F}$) has been increased from 0.5 to 0.555 such that $y_D = 0.9$ and $x_B = 0.002$ (see Table 1). For Case 1 ($M_D/F = M_B/F = 0.5$ min), the following approximate model is derived when scaled compositions ($dy_D/0.1, dx_B/0.002$) are used:

$$G_{C1}^S(s) = \begin{pmatrix} \frac{16.0}{1+\tau_1 s} & \frac{16.0}{1+\tau_1 s} + \frac{0.023}{1+\tau_2 s} \\ \frac{9.3}{1+\tau_1 s} & \frac{-9.3}{1+\tau_1 s} - \frac{1.41}{1+\tau_2 s} \end{pmatrix} \quad \begin{matrix} \tau_1 = 24.5 \text{ min} \\ \tau_2 = 10 \text{ min} \end{matrix} \quad (27)$$

The steady-state gains and time constants are entirely different from those at operating point A (8). Also note that at steady state $\lambda_{11}(G(0)) = 35.1$ for Column A, but only 7.5 for Column C. However, at high-frequency the scaled plants at operating points A and C are very similar. (8) and (27) yield:

$$G_1^S(\infty) = \frac{1}{s} \begin{pmatrix} 0.45 & -0.36 \\ 0.56 & -0.65 \end{pmatrix} \quad \lambda_{11}(\infty) = 3.2 \quad (28a)$$

$$G_{C1}^S(\infty) = \frac{1}{s} \begin{pmatrix} 0.65 & -0.65 \\ 0.38 & -0.52 \end{pmatrix} \quad \lambda_{11}(\infty) = 3.7 \quad (28b)$$

Therefore, as we will show, controllers which were designed based on the model $G^S(s)$ (operating point A) do in fact perform satisfactory also when the plant is $G_C^S(s)$ rather than $G^S(s)$. Recall that the use of a scaled plant is equivalent to using logarithmic compositions (Y_D and X_B). The variation in gains with operating

conditions is much larger if unscaled compositions are used – both at steady-state (Table 1) and at high frequencies:

$$G_1(\infty) = \frac{0.01}{s} \begin{pmatrix} 0.45 & -0.36 \\ 0.56 & -0.65 \end{pmatrix} \quad (29a)$$

$$G_{C1}(\infty) = \frac{0.01}{s} \begin{pmatrix} 6.5 & -6.5 \\ 0.08 & -0.10 \end{pmatrix} \quad (29b)$$

6.2 μ -Analysis

The μ -plots with the model $G_C^S(s)$ and four of the controllers are shown in Fig. 10 (all four controllers yield nominally stable closed-loop systems). At high frequencies the μ -values are almost the same as those found at operating point A. The only exception is the inverse-based controller $C_{inv}(s)$ which was found to be robustly stable at operating point A, but which is not at operating point C. Again, this confirms the sensitivity of this controller to model inaccuracies. Performance is clearly worse at low frequencies at operating point C (Fig. 10) than at operating point A (Fig. 6). This is expected; the controllers were designed based on model A, and the plants are quite different at low frequencies.

The μ -optimal controller $C_{1\mu}(s)$ satisfies the robust performance requirements also at operating point C when the reboiler and condenser holdups are small. Consequently, with the use of scaled (logarithmic) compositions, a single linear controller is able to give acceptable performance at these two operating points which have quite different linear models. The main difference between $C_{1\mu}(s)$ and the diagonal PI-controller is again that the μ -optimal controller gives a much faster return to steady-state. This is clearly seen from Fig. 11A.

6.3 Logarithmic Versus Unscaled Compositions

Fig. 10 shows how controllers designed based on the scaled plant $G^S(s)$ at operating point A, perform for the scaled plant (different scaling factors!) at operating point C; this is equivalent to using logarithmic compositions (Y_D and X_B). However, we know from (29) that the plant model shows much larger changes if

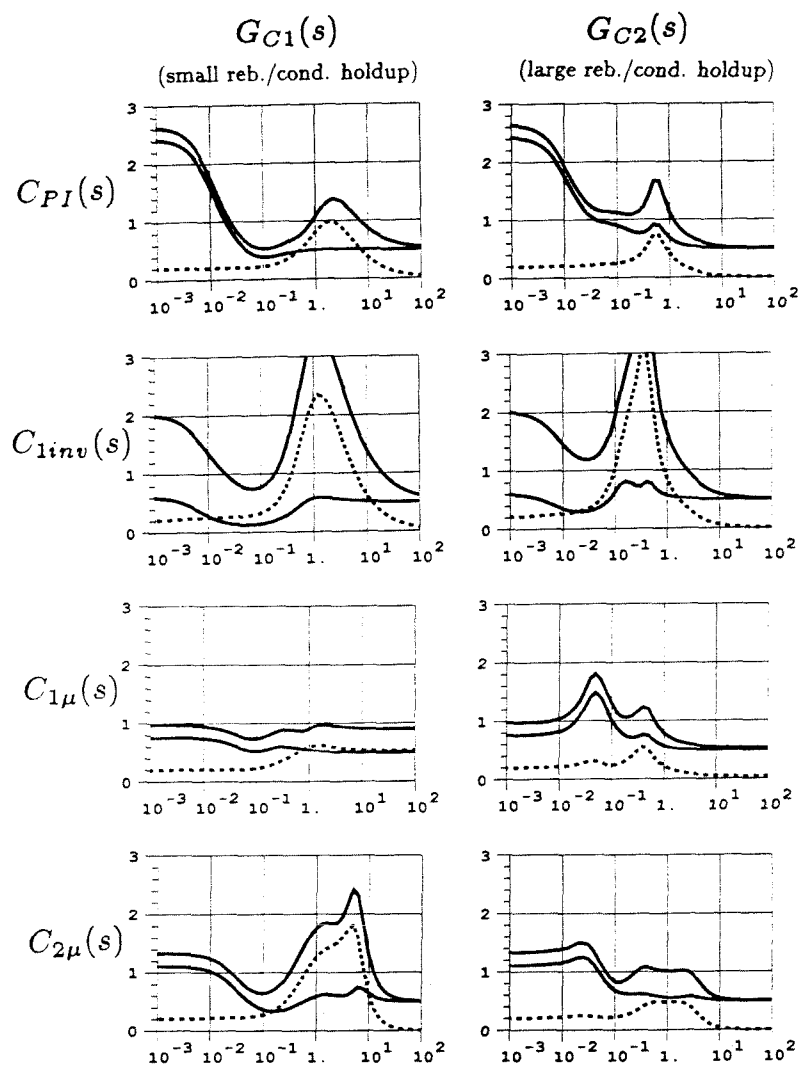


Figure 10. μ -plots for column C. Upper solid line: $\mu(N_{RP})$; Lower solid line: $\mu(N_{NP})$; Dotted line: $\mu(N_{RS})$.

absolute (unscaled) compositions (y_D and x_B) are used. We therefore expect the closed-loop performance to be entirely different at operating points A and C when unscaled (absolute) compositions are used. This is indeed confirmed by Fig. 11B which shows the closed-loop response to a small setpoint change in x_B at operating point C. Fig. 11B should be compared to Fig. 11A which shows the same response, but using logarithmic compositions as controlled outputs. In Fig. 11B (absolute compositions) the response for x_B is significantly more sluggish, and the response for y_D is much faster than in Fig. 11A (logarithmic compositions). This is exactly what we would expect by comparing (29a) and (29b): The high-frequency gain for changes in y_D is increased by an order of magnitude and the gain for changes in x_B is reduced by an order of magnitude. However, recall from (28) that the gain shows very small changes if logarithmic compositions are used.

The simulations in Fig. 12 are with no flow dynamics and in practice we expect the system to be unstable at operating point C if unscaled (absolute) compositions are used; the loop gain for y_D is increased by a factor of about 10 compared to the design conditions at operating point A. Assume we use the diagonal controller $C_{PI}(s)$ and are only controlling top composition (y_D) using reflux (L). Then the analysis reduces to a SISO- problem. At operating point A the loop transfer function for this loop is (unscaled compositions)

$$\text{A : } g_{11}c(s) = \frac{0.878}{1 + 194s} \frac{2.4(1 + 75s)}{s}$$

This corresponds to a closed-loop bandwidth ($|g_{11}c(j\omega_c)| \approx 1$) of about $\omega_c \approx 0.81 \text{ min}^{-1}$. The phase of $g_{11}c$ at this frequency is about -90° . The system will therefore become unstable if $90^\circ = \pi/2$ rad additional phase lag is added at this frequency. Consequently, the maximum allowed deadtime is $\theta_{max} = \frac{1.57}{0.81} = 1.93 \text{ min}$. Next, consider operating point C

$$\text{C : } g_{11}c(s) = \frac{1.6}{1 + 24.5s} \frac{2.4(1 + 75s)}{s}$$

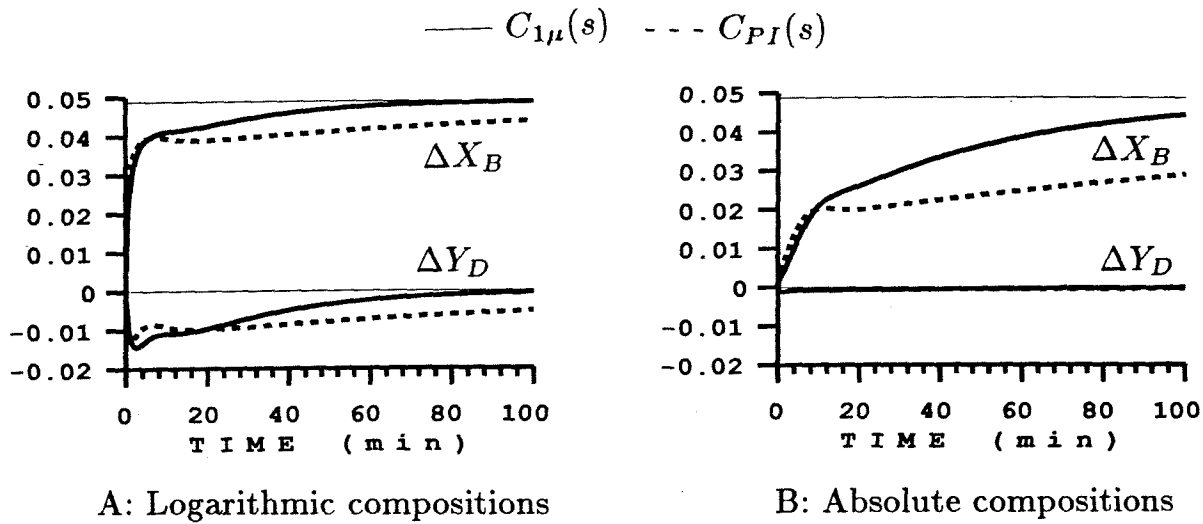


Figure 11. Column C, Case 1. Closed-loop response to small setpoint change in x_B (x_B increases from 0.002 to 0.0021) using diagonal PI-controller (dotted line) and the μ -optimal controller for column A (solid line). Left: logarithmic compositions as controlled outputs (equivalent to using scaled compositions); Right: Absolute (unscaled) compositions as controlled outputs. No uncertainty.

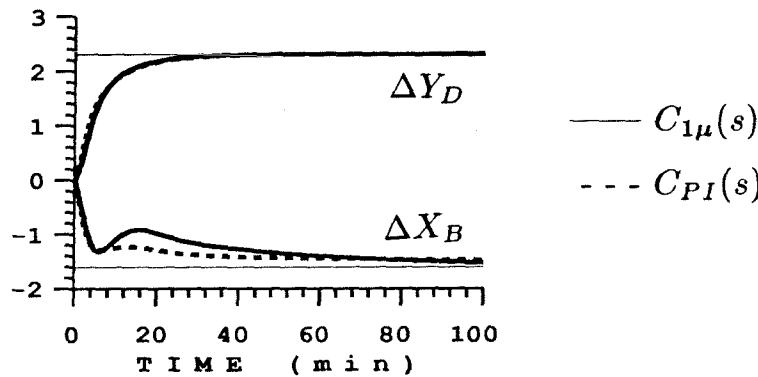


Figure 12. Transition from operating point A to C (Case 1) using controllers $C_{1\mu}$ (solid line) and C_{PI} (dotted line). Logarithmic compositions are used as controlled outputs to reduce the effect of nonlinearity. Desired trajectory is a first-order response with time constant 10 min. No uncertainty.

Here $\omega_c \approx 11.7 \text{ min}^{-1}$ and the phase is again about 90° . This gives a maximum allowed deadtime of only $\theta_{max} = \frac{1.57}{11.7} = 0.13 \text{ min}$. There may not actually be deadtime in the system, but the presence of other sources of phase lag (valve dynamics, measurements dynamics, etc.) will most likely result in an unstable system.

4.3 Transition from Operating Point A to C

Figure 12 shows a transition from operating point A ($Y_D = X_B = 4.605$) to operating point C ($Y_D = 2.303, X_B = 6.215$) using logarithmic compositions as controlled outputs. The desired setpoint change is a first order response with time constant 10 min:

$$\Delta Y_{D,s} = \frac{2.303}{1 + 10s}, \quad \Delta X_{B,s} = \frac{-1.609}{1 + 10s}$$

The closed-loop response is seen to be very good. The diagonal controller $C_{PI}(s)$ and the μ -optimal controller $C_{1\mu}(s)$ give very similar responses in this particular case. (However, the μ -optimal controller generally performs better at operating point C as is evident from Fig. 10 and 11.) This illustrates that a linear controller, based on the nominal operating point A, can perform satisfactory for a large deviation from this operating point when logarithmic compositions are used.

7. CONCLUSIONS

A single linear controller is able to give satisfactory control of this high-purity column at widely different operating conditions. One reason for this is the use of logarithmic compositions which effectively counteract for the nonlinearity in the plant. However, even if absolute compositions are used, a single linear controller performs satisfactory if the deviations from steady-state are reasonably small.

A simple diagonal controller was found to be robust with respect to model-plant mismatch, but gives a sluggish return to steady-state. This particular part of the response is improved using the μ -optimal controller. Inverse-based controllers, and in particular those based on a steady-state decoupler, are very sensitive to model-plant mismatch.

Acknowledgements. Financial support from the National Science Foundation and Norsk Hydro is gratefully acknowledged.

References

- Bristol, E. H., "On a New Measure of Interactions for Multivariable Process Control", IEEE Trans. Automatic Control, **AC-11**, 133-134 (1966)
- Doyle, J. C., "Analysis of Feedback Systems with Structured Uncertainties," IEE Proc., 129, Pt. D, **6**, 242-250 (1982).
- Fuentes, C. and W. L. Luyben, "Control of High-Purity Distillation Columns", Ind. & Eng. Chem. Process Des. & Dev., **22**, 361-366 (1983).
- Grosdidier, P., M. Morari and B. Holt, "Closed-Loop Properties from Steady-State Gain Information," Ind. & Eng. Chem. Fund., **24**, 221-235 (1985).
- Kapoor, N., T. J. McAvoy and T. E. Marlin, "Effect of Recycle Structure on Distillation Tower Time Constant," AIChE Journal, **32**, 3, 411-418 (1986).
- Moczek, J. S., R. E. Otto and T. J. Williams, "Approximation Model for the Dynamic Response of Large Distillation Columns", Proc. 2nd IFAC Congress, Basel (1963). Also published in Chem. Eng. Progress Symp. Series, **61**, 136-146 (1965).
- Moore, B. C., "Principal Component Analysis in Linear System", IEEE Trans. on Automatic Control, **AC-26**, 17-32 (1981).
- Nett, C. N. and V. Manousiouthakis, "Euclidean Condition Number and Block Relative Gain: Connections, Conjectures and Clarifications", submitted to IEEE Trans. of Automatic Control (1986).
- Ryskamp, C. J., "Explicit Versus Implicit Decoupling in Distillation Control", Chemical Process Control 2 Conference, Sea Island, GA, Jan. 18-23 (1981). (Published in T. F. Edgar and D. E. Seborg (eds.), United Engineering Trustees (1982), available from AIChE).
- Shinskey, F. G., Distillation Control, 2nd Ed., McGraw-Hill, New York (1984).
- Skogestad, S. and M. Morari, "Control of Ill-Conditioned Plants: High-Purity Distillation", paper 74a, AIChE Annual Mtg., Miami Beach (1986a).
- Skogestad, S. and M. Morari, "Implications of Large RGA-Elements on Control Performance", paper 6d, AIChE Annual Mtg., Miami Beach (1986b).
- Skogestad, S. and M. Morari, "Understanding the Dynamic Behavior of Distillation Columns", submitted to Ind. & Eng. Chem. Research (1987a).
- Skogestad, S. and M. Morari, "Understanding the Steady-State Behavior of Distillation Columns", in preparation (1987b).
- Skogestad, S. and M. Morari, "Control Configurations Selection for Distillation Columns", submitted to AIChE Journal (1987c).

Appendix. State-space realizations of plants and controllers.

Below are shown state-space realizations of $G(s) = C(sI - A)^{-1}B + D$ using "packed" form

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where D in all cases is a 2×2 matrix. The plant models and controllers are for the scaled plant, that is, correspond to using logarithmic compositions. All controllers were designed based on operating point A and when unscaled (absolute) compositions are used as controlled outputs the controllers should be multiplied by

$$\frac{1}{x_B^0} = \frac{1}{1-y_D^0} = 100.$$

$G_1^S(s)$:

```
-5.155e-03  0.000e+00  7.131e-02 -7.131e-02
0.000e+00 -6.667e-02 -1.556e-17  3.762e-02
6.347e-02  2.481e-02  0.000e+00  0.000e+00
7.821e-02 -2.481e-02  0.000e+00  0.000e+00
```

$G_2^S(s)$:

```
-4.539e-03  0.000e+00  0.000e+00  0.000e+00  0.000e+00  7.110e+00 -7.105e+00
0.000e+00 -3.086e-02  0.000e+00  0.000e+00  0.000e+00  6.297e+00 -7.408e+00
0.000e+00  0.000e+00 -8.580e-02  0.000e+00  0.000e+00  5.845e+00 -2.658e+00
0.000e+00  0.000e+00  0.000e+00 -1.708e-01  0.000e+00  2.852e+00 -5.414e+00
0.000e+00  0.000e+00  0.000e+00  0.000e+00 -8.876e-01 -2.699e+00  2.872e+00
6.361e-02 -5.394e-02 -1.641e-02  2.929e-03  8.189e-03  0.000e+00  0.000e+00
6.981e-02  1.664e-02 -4.683e-02 -5.886e-02  3.864e-02  0.000e+00  0.000e+00
```

$G_3^S(s)$:

```
-4.533e-03  0.000e+00  0.000e+00  0.000e+00  0.000e+00 -6.962e-01  6.958e-01
0.000e+00 -3.184e-02  0.000e+00  0.000e+00  0.000e+00 -5.982e-01  7.055e-01
0.000e+00  0.000e+00 -8.220e-02  0.000e+00  0.000e+00  6.681e-01 -4.407e-01
0.000e+00  0.000e+00  0.000e+00 -2.302e-01  0.000e+00  3.154e-01 -5.217e-01
0.000e+00  0.000e+00  0.000e+00  0.000e+00 -2.402e+00  5.535e-01 -4.607e-01
-6.185e-01  4.041e-01 -1.063e-02  6.331e-02  6.190e-01  0.000e+00  0.000e+00
-7.118e-01 -2.025e-01 -4.653e-01 -5.940e-01 -1.601e-01  0.000e+00  0.000e+00
```

$G_{C1}^S(s)$:

```
-4.082e-02  0.000e+00  7.354e-01 -7.354e-01
0.000e+00 -1.000e-01 -5.659e-17  4.052e-01
8.902e-01  5.604e-03  0.000e+00  0.000e+00
5.175e-01 -3.488e-01  0.000e+00  0.000e+00
```


$G_{C2}^S(s)$:

-1.382e-02	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	5.116e-01	-5.215e-01
0.000e+00	-7.809e-02	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-1.327e+00	1.540e+00
0.000e+00	0.000e+00	-1.328e-01	0.000e+00	0.000e+00	0.000e+00	0.000e+00	1.894e-02	2.508e-01
0.000e+00	0.000e+00	0.000e+00	-1.876e-01	0.000e+00	0.000e+00	0.000e+00	-1.043e+00	1.436e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-4.435e-01	8.790e-02	0.000e+00	3.310e-01	-1.313e-01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-8.790e-02	-4.435e-01	0.000e+00	1.037e+00	-9.500e-01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-3.428e+00	1.957e-01	-2.356e-01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-3.055e-02	0.000e+00	0.000e+00
5.338e-01	2.476e-01	-3.312e-02	-1.014e-01	2.864e-02	-5.006e-02	-2.215e-01	0.000e+00	0.000e+00
3.112e-02	-7.672e-01	-4.200e-01	8.015e-01	-2.502e-01	-1.897e-02	-2.215e-01	0.000e+00	0.000e+00

$G_{C3}^S(s)$:

-1.381e-02	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	1.334e-01	-1.360e-01
0.000e+00	-7.975e-02	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	9.212e-02	-1.076e-01
0.000e+00	0.000e+00	-1.758e-01	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-6.708e-02	7.042e-02
0.000e+00	0.000e+00	0.000e+00	-2.017e-01	0.000e+00	0.000e+00	0.000e+00	1.492e-01	-1.952e-01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-4.795e-01	0.000e+00	0.000e+00	-1.763e-01	1.859e-01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-3.292e+00	0.000e+00	1.229e-01	-9.529e-02
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-1.118e+01	1.212e-01	-1.045e-01
1.771e+00	-9.066e-01	-3.212e-01	-8.381e-01	-1.131e+00	1.510e+00	1.594e+00	0.000e+00	0.000e+00
1.250e-01	1.164e+01	7.279e+00	-3.346e+00	1.007e-01	-2.493e-01	-1.250e-01	0.000e+00	0.000e+00

$C_{PI}(s)$:

-1.000e-06	0.000e+00	1.000e-02	0.000e+00
0.000e+00	-1.000e-06	0.000e+00	1.000e-02
2.400e+00	0.000e+00	1.800e+00	0.000e+00
0.000e+00	-2.400e+00	0.000e+00	-1.800e+00

$C_{0inv}(s)$:

-1.000e-06	0.000e+00	3.994e-01	-3.149e-01
0.000e+00	-1.000e-06	3.943e-01	-3.200e-01
6.999e-01	0.000e+00	2.097e+01	-1.653e+01
0.000e+00	6.999e-01	2.070e+01	-1.680e+01

$C_{1inv}(s)$:

-1.000e-06	0.000e+00	0.000e+00	0.000e+00	-3.647e+00	-4.618e+00
0.000e+00	-1.000e-06	0.000e+00	0.000e+00	4.201e+00	4.175e+00
0.000e+00	0.000e+00	-1.000e+03	0.000e+00	-4.066e+02	-5.200e+02
0.000e+00	0.000e+00	0.000e+00	-1.000e+03	-4.786e+02	-4.705e+02
5.015e-01	5.020e-01	6.167e+01	-6.249e+01	0.000e+00	0.000e+00
5.015e-01	5.011e-01	6.176e+01	-6.112e+01	0.000e+00	0.000e+00

$C_{0\mu}(s)$:

-9.993e-08	0.000e+00	0.000e+00	0.000e+00	0.000e+00	1.530e-02	1.328e-02
0.000e+00	-1.000e-07	0.000e+00	0.000e+00	0.000e+00	2.140e-02	-1.774e-02
0.000e+00	0.000e+00	-4.873e-01	0.000e+00	0.000e+00	5.557e-02	-3.543e-02
0.000e+00	0.000e+00	0.000e+00	-2.267e+03	0.000e+00	-7.372e+01	-9.722e+01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-3.840e+03	4.796e+01	7.539e+01
1.592e+00	1.699e+00	4.751e+00	8.698e+03	-5.536e+03	3.526e+02	4.821e+02
-1.215e+00	2.176e+00	4.566e+00	-8.566e+03	5.777e+03	-3.509e+02	-4.815e+02

$C_{1\mu}(s)$:

-1.002e-07	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-6.513e-01	-9.009e-01
0.000e+00	-3.272e-06	0.000e+00	0.000e+00	0.000e+00	0.000e+00	7.224e-01	9.031e-01
0.000e+00	0.000e+00	-1.510e-01	0.000e+00	0.000e+00	0.000e+00	5.492e-02	-4.394e-02
0.000e+00	0.000e+00	0.000e+00	-9.032e+00	0.000e+00	0.000e+00	-9.086e-01	-1.136e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-5.838e+02	0.000e+00	1.867e+01	-1.494e+01
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-5.868e+02	6.722e+00	8.403e+00
6.564e-01	7.171e-01	4.949e+00	5.033e+00	-1.691e+03	-3.112e+02	5.866e+01	-3.816e+01
6.555e-01	5.425e-01	4.941e+00	-5.040e+00	-1.689e+03	3.116e+02	5.002e+01	-4.878e+01

$C_{2\mu}(s)$:

-9.487e-08	3.248e-09	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00
-3.248e-09	-9.487e-08	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00
0.000e+00	0.000e+00	-1.136e-07	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00
0.000e+00	0.000e+00	0.000e+00	-2.372e-01	0.000e+00	0.000e+00	0.000e+00	0.000e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-2.237e+00	1.977e+00	0.000e+00	0.000e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	-1.977e+00	-2.237e+00	0.000e+00	0.000e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-2.757e+00	2.082e+00
0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	0.000e+00	-2.082e+00	-2.757e+00
-1.091e+00	-3.796e-01	1.952e+00	-6.232e+00	9.384e+00	-1.239e+01	1.924e+01	-2.277e+01
-1.304e+00	-5.861e-01	8.008e-01	-4.432e+00	5.591e+00	-2.126e+01	-1.687e+01	7.900e+00

-1.142e-02	2.193e-02
-4.747e-03	2.594e-02
1.350e-02	1.806e-02
8.986e-02	-4.583e-03
5.918e-01	-3.288e-01
2.734e-01	8.376e-02
4.328e-02	-7.059e-04
-3.761e-01	-1.111e+00
-9.749e-05	-1.833e-04
5.644e-05	1.060e-04

Chapter XIV

CONCLUSIONS

This research project has motivated the need for a systematic approach to distillation column control which yields a control system with robust performance, that is, a system which performs satisfactory also in the presence of model-plant mismatch. Important steps include modelling, selection of the control configuration as well as robust controller design. In particular, a good physical insight about the process is essential to a non-conservative formulation of the control problem. In summary, the following contributions are made in this thesis:

A. Robust Control of linear systems. Simple measures are suggested which prove useful for screening design alternatives with respect to their achievable control performance and it is established that ill-conditioned plants are sensitive to model-plant mismatch. This does not mean that all ill-conditioned plants are difficult to control because the uncertainty is often "structured," and the plant may not be sensitive to the particular uncertainty which is present. However, one source of uncertainty which is *always* present is uncertainty with respect to the true values of the manipulated inputs. It is demonstrated that plants with large elements in the Relative Gain Array (RGA) are always sensitive to this kind of uncertainty, and are therefore generally difficult to control. * This fact is illustrated by the following two extreme choices for the controller: An inverse-based controller yields good nominal performance but is very sensitive to input uncertainty. On the other hand, a diagonal controller is insensitive to uncertainty but does not correct for the strong directionality of the plant (as expressed by the large RGA-elements) and therefore yields poor performance. Hopefully, these results will settle the persistent discussions in the literature concerning whether or not decouplers should be used for distillation columns.

The usefulness of the μ -theory for analysis and controller design is illustrated

* Note that plants with large RGA-elements are always ill-conditioned, but the converse is not necessarily true.

on a simple distillation column example, and it is demonstrated that ill-conditioned multivariable plants display a sensitivity to model uncertainty not observed for SISO plants. Furthermore, this sensitivity only becomes apparent under closed-loop control. Traditional design techniques such as LQG and loopshaping, which do not explicitly take uncertainty into account, will most likely fail for ill-conditioned plants.

Decentralized controllers are very commonly used for distillation columns because they are simpler to tune and to make failure tolerant. A method for designing decentralized controllers is proposed; each loop is designed independently in such a manner that the overall system is guaranteed to satisfy the robust performance specifications. The method provides a generalization of the μ -interaction measure introduced by Grosdidier and Morari.

B. Dynamic and steady-state behavior of distillation columns. Fundamental insight into the dynamic and steady-state behavior of distillation columns is provided. It is shown that high-purity distillation columns are always ill-conditioned because the product compositions are very sensitive to changes in the external material balance but are quite insensitive to changes in the internal flows. This fundamental difference between internal and external flows also manifests itself in the dynamic behavior; the time constant for changes in the external material balance (τ_1) is usually much larger than that for changes in the internal flows (τ_2). A simple, linear two time-constant model which is consistent with this behavior is presented together with simple formulas for estimating the parameters in the model. These formulas are *analytic* and therefore also yield insight into the nonlinear behavior of the column. It is shown that the response to any single disturbance or input is approximately first order with time constant τ_1 . However, to capture the multivariable effects, for example the *simultaneous* increase in reflux and boilup, the two time-constant model is required. Distillation columns are strongly nonlinear and

this may present a problem when employing linear controllers. This work shows that the initial response (which is of primary importance for feedback control) is much less markedly affected by nonlinearity and changes in operating conditions than is the steady-state behavior. This holds both when unscaled compositions are used and even more so when logarithmic compositions are utilized as controlled outputs. Logarithmic compositions should be used whenever large deviations from the nominal operating point are expected because they substantially counteract the effect of nonlinearity caused by changes in operating conditions.

C. Robust control of distillation columns. Most distillation columns can be described as a 5×5 plant but the flow and pressure dynamics are usually so fast that the composition control problem is essentially 2×2 . An important issue in distillation control is the selection of the two independent combinations of the five manipulated inputs to be used for composition control. Each configuration may yield entirely different control performance. For example, the LV-configuration (where L and V are used for composition control) yields large RGA-elements for high-purity columns and is therefore sensitive to input uncertainty, whereas the DV-configuration always has small RGA-elements. However, there are also other, generally conflicting, considerations besides uncertainty which have to be addressed. These include dynamic response, rejection of flow disturbances by the level loops, one- and two-point composition control, changes between manual and automatic control as well as constraints. It is shown that ratio configurations yield complex multivariable controllers which in some cases provide improved flow disturbance rejection. However, their linearizing effect on the plant does not prove significant. Overall, weighing all the above-mentioned considerations, the $\frac{L}{D} \frac{V}{B}$ -configuration seems most versatile.

To verify the general results found on the dynamic behavior of distillation columns, a realistic control study for a high-purity column using the LV-

configuration was performed. The column was successfully controlled with a linear controller designed on the basis of a linear model of the nominal operating point. The effects of model uncertainty, nonlinearity due to changes in operating conditions as well as variations in reboiler and condenser holdup were studied. The primary conclusion is that a linear controller is indeed able to provide satisfactory control for this strongly nonlinear plant despite the above-mentioned sources of model-plant mismatch. For large deviations from steady state logarithmic compositions had to be used in order to reduce the effect of nonlinearity. A simple diagonal controller was found to be robust with respect to model-plant mismatch but produced a very sluggish return to steady state. This particular part of the response (the low-frequency behavior) was improved with the μ -optimal controller. The inverse-based controllers, in particular those based on steady-state decouplers, were found to be very sensitive to model-plant mismatch and therefore did not offer robust performance.

Suggestions for future research.

Robust control. There remains a substantial number of outstanding research issues in the area of linear robust control (obviously, if nonlinear systems are included the list becomes endless). For example, for the application of the μ -theory, the formulation of the control problem in terms of performance and uncertainty weights is essential, and no general guidelines exist for their selection.

One particular research area of interest to process control is the design of decentralized controllers. The μ -theory provides a means for evaluating the achievable performance bound for a given plant using any multivariable controller (by actually determining the μ -optimal controller). A similar target on the achievable performance for cases when the controller is restricted to diagonal would be very helpful; the engineer could then verify immediately whether acceptable performance were indeed possible with a diagonal controller.

Distillation. In the area of distillation control, one of the most important outstanding research areas is the measurement selection and estimation of product compositions: in most cases accurate on-line measurements of the product compositions are not available and one has to rely on measurements of temperatures inside the column. It seems likely that the best results will be obtained by using a nonlinear model which estimates the product compositions based on the entire temperature profile. In some cases only two temperature measurements are used and it is not at all clear how to select the best location of these measurements.

Another important issue which was only briefly discussed in this thesis (Chapter XII) is constraints on the manipulated inputs. It is clearly of utmost importance that the control system deal with constraints in an effective manner.

A number of simplifying assumptions were made when deriving specific results regarding the behavior of distillation columns in this thesis. Though it is believed that the results presented will carry over to more complicated mixtures and column configurations, these issues should obviously be investigated. Furthermore, there is a need for further simulation studies and μ -analysis of other configurations in addition to the LV-configuration. Some topics suggested for future considerations are summarized below:

- Mixtures with non-ideal VLE and with non-constant molar flows.
- Multicomponent mixtures.
- Columns with more than one feed (trivial extension) and more than two products (not as obvious).
- Effect of flow dynamics.
- Columns with high-purity products and large reflux for which the worst control problems are expected when using the LV-configuration.
- Simulation studies and μ -analysis of other control configurations (for example, the DL- or $\frac{L}{D} \frac{V}{B}$ -configurations). The tuning of the level loops is important for

configurations which use D or B for composition control.

Finally, it would be most desirable to perform experiments and implement controllers on a real industrial column. This might reveal new issues of importance which were not covered here.