INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

PRODUCT QUALITY CONTROL

IN

REDUCED DIMENSION SPACES

BY

TRACY LEE CLARKE-PRINGLE, B.Sc., M.Eng.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

McMaster University

....to my parents, Howard and Judy Clarke

PRODUCT QUALITY CONTROL IN REDUCED DIMENSION SPACES

DOCTOR OF PHILOSOPHY (1999)
(Chemical Engineering)

MCMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Product Quality Control in Reduced Dimension Spaces

AUTHOR: Tracy Lee Clarke-Pringle
B. Sc. (Queen's University)
M. Eng. (McMaster University)

SUPERVISOR: Professor John F. MacGregor

NUMBER OF PAGES: xiv, 171

Abstract

In this thesis, the problem of product quality control is addressed. The primary goal of many processes is to produce a consistent product, that is, to minimize the variability in the product quality. This thesis addresses two particular quality control problems, with the focus being on the second issue. First, the problem of incorporating an approximate fundamental model into an online feedback quality control scheme is addressed. The second problem investigated in the thesis relates to the dimensionality issues in product quality control. The issue of controlled and manipulated variable selection while minimizing the overall variability in the product quality is addressed.

In the first part of the thesis, control of the full molecular weight distribution (MWD) in a semi-batch polymerization reactor is considered. A new batch-to-batch optimization methodology for producing a desired MWD using an approximate fundamental model is presented. The optimization approach is also extended for use as an on-line control method by incorporating a multivariable statistical process control (MSPC) monitoring scheme. The combined MSPC/batch-to-batch optimizer is demonstrated on a simulated semi-batch polystyrene reactor and is shown to be very effective in reacting to large process upsets. Despite significant process/model mismatch, the batch-to-batch optimizer is able to adjust the process and produce the desired MWD within several batches following a process upset.

In the remainder of the thesis, dimensionality issues in product quality control are addressed. First, indirect control of the full MWD by directly controlling only the average of the distribution is illustrated. It is shown that the choice of manipulated variables has an important impact on the controller performance when the full MWD is considered, and controlling a single average chain length sometimes causes the MWD to degrade. A simple analysis tool, called the Disturbance Inflation Factor (DIF), is introduced to evaluate which controlled and manipulated variables result in the best overall control of

the full MWD. It is shown that with prudent choices of manipulated and controlled variables, simple single variable control can provide significant improvements in the full MWD. Controlling new linear combinations of the original inputs and outputs is also shown to be a feasible option for minimizing the overall effect of the disturbances.

The dimensionality ideas raised when controlling the MWD are then generalized. A framework for selecting controlled and manipulated variables that minimize the variability in the overall product quality is derived. Given the disturbance directions and process gain matrix, expressions for the optimal directions for control are derived. The role of the number of independent disturbances in determining the number of controlled variables, and the structure of the resulting reduced dimension controller is clearly shown. The framework is then applied to a simulated dynamic Kamyr digester. Two single input, single output Reduced Dimension Controllers (RDCs) are proposed and compared to a Dynamic Matrix Controller (DMC) that controls all outputs and manipulates all inputs. The RDCs performed very well at the conditions for which they were designed and showed only modest degradation when the process operating point was changed. Despite their much simpler structure, their performance is very close to that of the DMC.

Finally, the thesis concludes with a critical review and unification of existing methods for Reduced Dimension Control. It was found that due to the different industries from which the applications originate, there is little or no comparison of similar approaches, thus there is the need for a unification. The existing approaches are characterized as either data based or model based for the discussion. Each of the general methods are discussed, then illustrated with specific published examples. Where appropriate, the methods and examples were placed within the context of the RDC framework derived earlier in the thesis, and the situations in which the methods are expected to perform well or be unreliable are discussed. An overall approach to dealing with RDC problems using a combination of the methods is also discussed.

Acknowledgments

Many people have touched me in my time at McMaster, and it is impossible to thank everyone. However, a few people deserve extra thanks and I mention them here.

To begin, I would like to thank my supervisor, Dr. John MacGregor, from the bottom of my heart. His constant and unwavering guidance and support have made this work possible. Through his positive encouragement and creative ideas, I have learned and my confidence has grown. In the tough times, Dr. MacGregor reminded me of the big picture and continued to encourage me. I will remember these years with great fondness thanks to him. I would also like to thank Dr. Tom Harris from Queen's University, without whom I would have never journeyed to McMaster, and who continues to be a mentor for me. I would like to thank NSERC and the Department of Chemical Engineering for the funding that made it possible for me to be a student at McMaster.

To my closest friends, Liz, Paula, Melissa, Christine, Michelle, Becky, Manish and Carl: 'Many people will walk in and out of your life, but only true friends leave footprints in your heart'...Over the years, you have each touched me in unique ways. I thank you for your true, unconditional friendship and the inspiration you continue to give me.

For my colleagues and friends in the Penthouse: Jose, Ken, Ali, Jason, Seongkyu, Song, Laura, Ruijie, San, Manish and Carl, I am very happy to have gotten to know each of you better during my stay at McMaster and I thank you for welcoming me into your group. The discussions and good times we had will never be forgotten by me. I cherish the memories that I leave with, and will carry them with me as I go.

I want to especially thank my mother and father, for doing everything right (even when I was too young to see it). I credit my parents with where I am and I thank my mom and dad for the sacrifices they have made to help me reach for my full potential.

And finally, I thank my husband Matt with all my heart, for his love and support throughout my thesis, and particularly in my final year. You have always believed in me and stood by me, and to be able to share this success with you is sweet.

One last, parting thought: '*This is not the end, this is not even the beginning of the end, it is the end of the beginning' (Winston Churchill)*...

Table of Contents

List of Figures

List of Tables

1. Introduction

In this thesis, the problem of product quality control is addressed. For many processes, the final product quality is of paramount importance. The primary goal of many such processes is to operate within specifications while minimizing the variability in the product quality. This thesis addresses two particular quality control problems, with the focus being on the second issue. First, the problem of incorporating an approximate fundamental model into an online quality control scheme is addressed. In particular, control of the full molecular weight distribution in a semi-batch polymerization reactor is addressed. The second problem investigated in the thesis relates to the dimensionality issues in product quality control. Specifically, product quality is often characterized by numerous, highly correlated quality variables, and these variables usually outnumber the variables available for manipulation. The issue of controlled and manipulated variable selection while minimizing the overall variability in the product quality is addressed in the thesis.

In Chapter 2, achieving acceptable product quality in the absence of complete fundamental models is investigated. Approximate models are often available, and it could be beneficial to incorporate these into a control scheme. To this end, a batch-to-batch optimization methodology for producing a desired molecular weight distribution (MWD) in a polymerization reactor using an approximate fundamental model is developed. The method uses fundamental polymer knowledge in order to simplify an otherwise complicated optimization problem and provide significant freedom in manipulated variable selection. The optimization approach is also extended for potential use as an on-line control method. A multivariable statistical process control (MSPC) monitoring scheme is implemented in conjunction with the optimizer for deciding when a new batch correction is required. The optimizer remains on, but dormant while the desired MWD is

being produced, and re-optimizes the process if the process changes and poor quality polymer is produced. The combined MSPC/batch-to-batch optimizer is demonstrated on a simulated semi-batch polystyrene reactor.

Directly controlling the full MWD requires substantial process information. Due to the lack of sufficient process knowledge, indirect control of the full MWD, for example by controlling an average of the distribution, is often practiced in industry instead. In Chapter 3, indirect control of the full MWD, by controlling the average of the MWD, is addressed. Unfortunately, good control of the distribution average does not necessarily result in a better distribution (and hence better polymer quality). In Chapter 3, it is shown that as a controller eliminates a disturbance in the controlled variables (for example, the average chain length), it transfers and can possibly inflate the disturbance in the remaining quality variables (the full MWD). Therefore, while it may appear that good control is being achieved (the average is at its target), the polymer quality has in fact degraded. A simple steady state analysis tool, called the Disturbance Inflation Factor (DIF), is introduced to quantify this effect and provide guidance in selecting controlled and manipulated variables. The DIF is used to predict which manipulated variables result in the best control of the full MWD while acting only on a single measured variable such as the weight-average chain length. It is further applied to evaluate if control of the full distribution may be improved by considering other controlled variables, such as the number-average chain length, or other manipulated variables, such as combinations of the existing manipulated variables. The ideas are illustrated on the same simulated polystyrene reactor.

The final two chapters of the thesis deal more generally with the dimensionality issues that are raised in the third chapter. Often, when product quality is characterized by numerous, highly correlated variables, a choice is made to control a subset of the variables, in order to improve the overall product quality. However, the choice of the variables to control and manipulate is critical to improving the overall product quality. In Chapter 4, a framework for the optimal reduction of the dimension of the controlled

system is derived from minimum variance theory. The framework for Reduced Dimension Control (RDC) shows explicitly how process and disturbance directions impact the selection and required number of controlled and manipulated variables. Furthermore, the framework highlights the assumptions that are made about the process when selecting the subsystem for control. The framework is then demonstrated on a simulated, dynamic Kamyr digester for controlled and manipulated variable selection. The digester was selected for the case study since it is an example of a process that violates many of the framework assumptions. Therefore, the impact of applying the RDC framework under these circumstances is evaluated using the digester.

In Chapter 5, a unification of existing applications in the area of Reduced Dimension Control is given. Quite often, RDC approaches are application driven and therefore solutions for controlled and manipulated variable selection have been independently proposed by researchers in many different fields. In order to increase awareness and understanding in this promising field, Chapter 5 provides a unifying view of the existing methodologies and applications of reduced dimension control. Existing methods are characterized as either Data Based or Model Based Approaches. Data based approaches use plant data directly, whereas Model based approaches require the identification of a plant model first (usually in transfer function form). The methods are discussed generally, then illustrated with specific examples from the literature. Where appropriate, each method is placed within the framework of Chapter 4 in order to provide insight into the method and evaluate the effectiveness. Situations in which a method is expected to be unreliable are also discussed. The Chapter concludes by discussing a comprehensive approach to RDC and process dimensionality analysis.

Concluding remarks and contributions of the thesis are outlined in the final chapter.

2. Optimization of Molecular Weight Distribution Using Batch-to-Batch Adjustments

2.1 Introduction

Research can often be broadly classified as either optimization or control. In batch process optimization, one wishes to calculate the optimal set of initial conditions and manipulated variable trajectories for a non-time varying situation, subject to process constraints. In contrast, batch process control implies some form of feedback to adjust for time varying disturbances and process changes. The two approaches are complementary, and the distinction between them is often somewhat artificial.

The two areas can be further subdivided. Most batch reactor control approaches can be classified into two areas: Within-batch control and batch-to-batch control. Within-batch control includes any action taken during the operation of the batch, for the purpose of controlling the quality of the final polymer from that batch. This may involve continuous on-line control (Kozub and MacGregor (1992), Soroush and Kravaris (1992)) or mid-batch corrections based on intermittent quality measures (Yabuki and MacGregor (1996), Tsen et al (1996)). Batch-to-batch control implies any correction made for the next batch, based on the results of previous batches. Two such examples may be found in Box and Jenkins (1976) and Vander Wiel et al (1992).

Batch optimization can also be broken down into two main areas: classic approaches and batch-to-batch optimization. The classic approach to batch reactor optimization, often referred to as optimal control, involves the use of fundamental models and constrained optimization methods in an off-line optimization (e.g. Cawthon and Knaebel (1989), Choi and Butala (1991)). The optimization criterion is typically some function of average molecular weight, polydispersity, batch time and/or final conversion and the result is one or more trajectories that are to be implemented in open-loop fashion

during the batch. The limitation of optimal control methods is that a very good process model is required. The direct implementation of the optimal trajectory calculated off-line will not result in the expected polymer if there is any process/model mismatch or unmeasured disturbances. The emergence of the second area, batch-to-batch optimization, which is the focus of this work, is a direct result of this deficiency. The general idea is to use the best available knowledge (reasonable model and results from previous batches), and converge to a good set of operating conditions using the iterative nature of batches. Several applications have appeared in the literature and these are briefly discussed below.

Filippi-Bossy et al (1989) applied the concept of tendency models (as introduced in previous papers by this research group) to optimize batch reactors. They assumed that the kinetics of the process are not well-known, and proposed a structure that could approximate the kinetics. By alternately solving an optimization algorithm for a manipulated variable trajectory and updating the model parameters based on the actual batch results, they arrive at a good manipulated variable trajectory within several batches. Optimization objectives considered are maximizing a certain product while minimizing batch time. Recent work by this group looks at assigning confidence regions to tendency models (Fotopoulos (1996)).

The optimization approach published by Zafiriou and Zhu (1990) and Dong et al. (1996) uses batch data (trajectories and endpoint) in the gradient calculation of the optimization scheme. The idea is based on an analogy between numerical optimization and batches and was first demonstrated by Zafiriou and Zhu (1990) with a first principles model. In that paper, the authors replaced the model with the actual plant for the forward integration step of the gradient calculation and the model was used only for the solution of the adjoint equations. The method was applied to a bulk polystyrene polymerization process in which the batch temperature profile is adjusted to maximize an objective function of number and weight averages and conversion. In the more recent work, an empirical NNMPLS model was incorporated in order to predict the values of the states (during the batch) required for the gradient calculation. Furthermore, the use of the

empirical model allowed for an analytical gradient calculation. In the latter paper, two case studies are presented. First, the polystyrene example in Zafiriou and Zhu (1990) was repeated. Second, the feedrate of a reagent is optimized to maximize the production of a certain product in a bio-reactor.

Another nice application of batch-to-batch optimization for polymerization processes is that of Gugliotta et al (1995). The objective was to produce a compositionally uniform copolymer in the minimum batch time. The monomer feed was available to manipulate. Using a simple model and heat release data from the last batch, the feedrate profile for the next batch was calculated. In the experimental example shown, an 'optimal' feed profile was obtained after only four batches.

In this chapter, a novel batch-to-batch optimization methodology for producing a desired MWD using an approximate model is presented. The method uses fundamental polymer knowledge in order to simplify an otherwise complicated optimization problem and to provide significant freedom in manipulated variable selection. A measurement of the MWD at the end of the batch is used to update manipulated variable trajectories for the next batch, thus iterating into a good operating policy. The MWD has long been recognized as a desirable variable to control. Recently, some attempts have been made to estimate or control the full MWD in batch reactors (Takamatsu et al. (1988), Chang and Lai (1992), Ellis et al (1994) and Crowley and Choi (1997), Yoon et al (1998), Chang and Liao (1999)). All applications require a detailed polymerization model and consider linear polymers, however none have attempted to incorporate batch-to-batch corrections. Instead, all rely on the soundness of their models to calculate the required manipulated variable trajectories correctly the first time.

The chapter is structured as follows: in the second section, an overview of the strategy is presented. This will include the division of all linear polymers into two general classes. In the third section, the methodology is outlined for one class of linear polymers. A case study, using a polystyrene simulation, is also presented. The fourth section presents the method and a case study for the second class of linear polymers. Initially the

method is demonstrated with noise-free MWD measurements. However, in the fifth section, realistic noise is added to the true MWD in order to simulate an industrial environment and a combined monitoring and batch-to-batch optimization scheme is presented. Finally, a discussion of the methodology and the results may be found in the final section.

2.2 Overview of Strategy

Many polymer reactor control problems are demonstrated using the well known polymethylmethacrylate system. PMMA is just one of many linear polymers, all of which have certain characteristics that allow the MWD problem to be simplified significantly. Specifically, once a linear polymer is formed, it remains 'dead' until the end of the batch. Because of this fact, the MWD at the end of the batch is simply a weighted average of the MWD of the polymer formed at each instant during the batch:

$$W(r)_{cumulative} = \sum_{i} W(r)_{\text{formed in instant } i} * f_i \tag{2.1}$$

where f_i is the fraction of the polymer made at instant 'i'

The instantaneous MWD distribution of linear polymers is characterized by at most two parameters (Hamielec and Tobita (1992)). These parameters, τ and β, are ratios of the various kinetic rates. For the most general linear polymer (notation is provided at the end of the chapter):

$$\tau = \frac{R_{td} + R_{fm} + R_{fT} + R_{fi} + R_{fs}}{R_p}$$

$$\beta = \frac{R_{tc}}{R_p} \tag{2.2}$$

and,

$$W(r)_{instantaneous} = (\tau+\beta)\left(\tau+\frac{\beta}{2}(\tau+\beta)(r-1)\right)r\left(\frac{1}{1+\tau+\beta}\right)^{r+1} \qquad (2.3)$$

Therefore, the distributions made at each instant during the batch can be characterized by the values of only two parameters, τ and β. The trajectories of these two parameters during the batch (along with a knowledge of f_i) are enough to describe the final (cumulative) MWD. If τ and β are held constant for the entire batch, the cumulative MWD (2.1) will be the same as the instantaneous one (2.3). An extension of this idea forms the basis for the proposed methodology: if one could 'control' τ and β at a few desired levels during the batch, one has the ability to control the shape of the MWD that results.

With this in mind, the overall strategy can be viewed as a series of three steps:

1/ Decompose the (cumulative) desired MWD into a series of instantaneous distributions, each given by (2.3) and characterized by a particular choice of τ and β. This will provide initial trajectories of τ and β to be implemented during the first batch.

2/ Implement the specified τ and β trajectories in the batch reactor and measure the cumulative MWD of the final polymer produced.

3/ Compare the measured MWD to the desired distribution and update the τ and β trajectories accordingly. Repeat steps 2/ and 3/ until the distribution being produced is sufficiently close to the desired one.

For many linear polymers, one of the parameters is much larger than the other. The first class of polymers considered in this paper are those in which one parameter dominates the MWD. The second class are those in which both parameters are significant.

2.3 Methodology for Linear Polymers

The three step methodology for linear polymers will first be discussed for a system in which β is much larger than τ. The results are easily extended to systems in which $\tau >> \beta$ but are not shown in order to conserve space.

2.3.1 Decomposition of the Desired MWD (Step 1)

As discussed in the previous section, the distribution at the end of the batch is a weighted sum of all the distributions formed instantaneously:

$$W(r)_{cumulative,desired} = \sum_{j=1}^{\infty} f_j \times W_j(r,\beta_j) \tag{2.4}$$

The LHS of (2.4) is specified: it is the desired distribution. The RHS is unknown; we do not know what weighted sum will equal the desired distribution. However, the desired MWD can usually be approximated by a finite number of instantaneous distributions. Equation (2.4) becomes:

$$W(r)_{cumulative,desired} = f_1 \times W_1(r,\beta_1) + f_2 \times W_2(r,\beta_2) + \ldots + f_n \times W_n(r,\beta_n) \tag{2.5}$$

Rewriting (2.5) in vector form:

$$\mathbf{w_D} = \mathbf{W_I f} \tag{2.6}$$

The vector $\mathbf{w_D}$ contains the specified values of $W(r)_{cumulative,desired}$ at a finite number of chain lengths r. Each column of the $\mathbf{W_I}$ matrix contains an instantaneous MWD (that is characterized by a single value of β). Both $\mathbf{W_I}$ and **f** are unknown. However, since the desired distribution ($\mathbf{w_D}$) is known, it is very easy to generate a group of instantaneous distributions that, summed together, *could* approximate the desired MWD. One could generate as many as 10 or 15 instantaneous distributions using (2.3) with $\tau = 0$ (at equispaced values of β, for example) for the columns of the $\mathbf{W_I}$ matrix. Then, a very simple optimization can be performed to calculate the weight fractions f_i associated with each instantaneous distribution:

$$\min_{f} (w_D - W_I f)^2$$

$$\text{subject to} \sum_i f_i = 1 \text{ and } f_i > 0 \quad \forall i \tag{2.7}$$

The optimization returns a value for the vector **f**. Distributions in W_I with an associated non-zero value of f_i are constituent distributions. Any distribution in W_I that does not contribute will have an associated fraction that is very small or even equal to zero. Distributions with small fractions can be eliminated (delete the associated column in W_I) and the optimization (2.7) repeated with the new W_I matrix. In this way, the number of instantaneous distributions that must be made can be minimized. Furthermore, since each instantaneous distribution is characterized by a single value of β, the decomposition provides the 'n' levels of β at which the process must be sequentially operated in order to produce the desired MWD.

Of course, the solution to the decomposition is not unique and will depend on the set of instantaneous distributions one starts with, and the instantaneous distributions one chooses to eliminate. Different sets of distributions (different W_I matrices) should be attempted in order to find the best approximating set of constituent distributions.

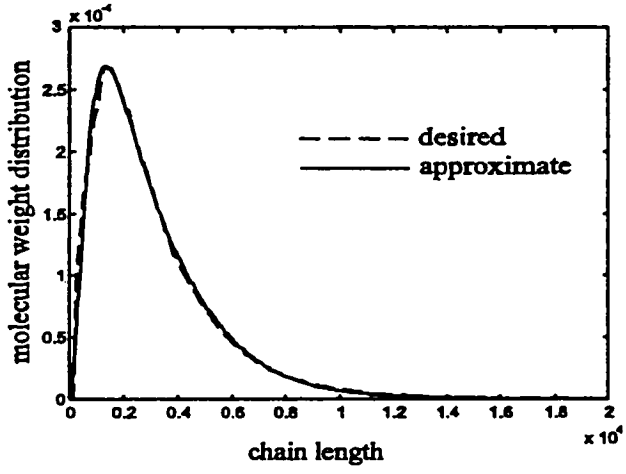

Figure 2.1. Decomposition of the desired MWD

Consider as an example the distribution shown in Figure 2.1 (dashed line). This distribution is part of a published data set (Dube and Penlidis, 1995), and thus represents an arbitrary distribution. Assume this is the desired MWD and the polymer is one for which $\beta >> \tau$. The desired weight average chain length is approximately 3350, and so a good range of instantaneous

distributions are those characterized by β_i values from 0.0004 to 0.0022.

The first $\mathbf{W}_I$ matrix had 10 columns, for the instantaneous distributions characterized by $\beta = 0.0004$, $\beta = 0.0006$, ... $\beta = 0.0022$ (the columns were generated using (2.3) with $\tau = 0$). After the first optimization, the vector **f** had several elements that were very small. The associated columns in $\mathbf{W}_I$ were deleted and the calculation (Equation (2.7)) repeated. After several trial and error attempts, a good fit was found. Table 2.1 contains the results of the final attempt.

Table 2.1. Decomposition of example distribution

Set of instantaneous distributions (specified by β value)	Result of optimization (f_i)
$\beta_1 = 0.0006$	0.1551
$\beta_2 = 0.0008$	0.2421
$\beta_3 = 0.001$	0.3084
$\beta_4 = 0.002$	0.2984

The operating policy calculated by the optimization indicates that 15.5% of the polymer should be made while operating at a constant β value of 0.0006, 24.2 % of the polymer made while operating at a constant β value of 0.0008, and so on. Perfect implementation of this operating policy would give the distribution shown as a solid line in Figure 2.1.

Note that the optimization does not specify the order in which the instantaneous distributions should be produced. With respect to making the desired distribution, the order does not matter. Other factors, such as ease of implementation, ability to execute changeovers, batch time or final conversion will dictate which order is appropriate.

2.3.2 Production of Constituent Distributions (Step 2)

The second step in the methodology involves producing the instantaneous distributions (or equivalently, operating at the associated β values) that were calculated in the first step. The desired β trajectory will consist of 'n' constant sections, such as those shown in Table 2.1 (with a step change between each section). To implement this desired trajectory, one option is to estimate the value of β online (during the batch), and manipulate a variable such as a flowrate or temperature so that the estimated value β follows the desired trajectory. Recall the expression for β:

$$\beta(t) = \frac{R_{tc}}{R_p} = \frac{k_{tc}[R\bullet]^2}{k_p[M][R\bullet]} = \frac{k_{tc}[R\bullet]}{k_p[M]} \tag{2.8}$$

In order to estimate β online, the following information is needed: values of k_p and k_{tc}, an estimate of the monomer concentration ([M]) and estimate of the radical concentration ([R•]). It is expected that approximate (not exact) values of k_p and k_{tc} are known. It is unlikely that the monomer concentration is measured online, however it is reasonable to expect an online measurement of the heat released (Q_r). Therefore, the monomer concentration can be estimated using an online mole balance:

$$\frac{dM_{estimated}}{dt} = F_{monomer} - \frac{Q_{r,measured}}{H_r}$$
$$[M]_{estimated} = M_{estimated} / V \tag{2.9}$$

where H_r is the heat of propagation per mole of monomer. Furthermore, the measured value of the heat released can be used to infer radical concentration:

$$Q_r = R_p VH_r = k_p[M][R\bullet]VH_r$$

$$[R\bullet]_{estimated} = \frac{Q_{r,measured}}{k_p[M]_{estimated} VH_r} \tag{2.10}$$

Substituting these expressions into (2.8):

$$\beta_{estimated}(t) = \frac{k_{tc} Q_{r,measured}}{(k_p[M]_{estimated})^2 VH_r} \tag{2.11}$$

Equation (2.11) provides an online estimate of β during the batch. To control $\beta_{estimated}$, a standard feedback controller can be implemented. The desired β trajectory calculated from the decomposition step (consisting of the 'n' constant values of β) is the setpoint for this controller. Depending on the process, a variable such as batch temperature, monomer flowrate or initiator flowrate can be used as the manipulated variable for the controller. The control algorithm can be as simple as the classic PID, or more complex as the application demands.

If there is no model mismatch and no measurement error in Q_r, then the estimated value of β (calculated from (2.11)) will be exactly equal to the true value of β in the process (2.8). Furthermore, if perfect control is achieved ($\beta_{estimated}$ is maintained at its setpoint without deviation), then the measured molecular weight distribution at the end of the first batch will be the desired distribution. However, this is never the case and therefore the third step in the methodology is a procedure for modifying the operating conditions for the next batch based on the measured MWD at the end of the current batch.

2.3.3 Batch-to-Batch Updating (Step 3)

Of course, the trajectory that the true value of β follows will not be exactly equal to the estimated β trajectory due to process/model mismatch and the effect of measurement errors. The result is that the MWD at the end of the first batch will not be

the desired one. Therefore, the methodology must provide some mechanism for modifying the β_{sp} trajectory for the next batch.

One modification that can be easily made from batch-to-batch is to adjust the individual levels of the 'n' values of β in the desired β trajectory. Consider a Taylor series expansion of the molecular weight distribution about the setpoint trajectory (the 'n' constant segments $[\beta_1\ \beta_2 \ldots \beta_n]$) from the B^{th} batch in a sequence of batches:

$$W(r)_{cumulative,B+1} = W(r)_{cumulative,B} + \left[\frac{\partial W(r)_{cumulative}}{\partial \beta_1}\ \frac{\partial W(r)_{cumulative}}{\partial \beta_2} \ldots\right]\begin{bmatrix}\Delta\beta_1\\ \Delta\beta_2\\ \vdots\end{bmatrix} \tag{2.12}$$

where an expression for the derivative can be derived from (2.3), assuming that only the level of the setpoint (not the mass fraction of polymer made at each condition) is adjusted:

$$\frac{\partial W(r)_{cumulative}}{\partial \beta_i} = f_i \times \frac{\partial W(r)_{instantaneous}}{\partial \beta_i} = f_i \times \left(\frac{3}{\beta_i} - r\right) W(r,\beta_i)_{instantaneous} \tag{2.13}$$

The derivation of (2.13) may be found in Appendix 2.1 at the end of this Chapter. A similar expression can be derived for polymers with $\tau >> \beta$.

Rewriting the expressions in (2.12) and (2.13) as vectors and matrices:

$$\mathbf{w}_{B+1} = \mathbf{w}_B + \mathbf{D}\,\Delta\beta \tag{2.14}$$

Feedback is introduced into (2.14) by using the measured molecular weight distribution (from batch 'B') for $\mathbf{w}_B$. The derivative matrix ($\mathbf{D}$) is evaluated using (2.13) assuming perfect setpoint tracking. The MWD for batch 'B+1' is specified as the desired MWD

($w_{B+1} = w_D$) and (2.14) can be solved for the changes in the setpoint trajectory ($\Delta\beta$) so that the MWD produced in the next batch is the desired distribution.

In (2.14), the MWD for the next batch is specified as the desired MWD. However, this could result in unacceptably large changes for the next batch, especially if the error in the MWD is large. Furthermore, model error may result in overcompensation (and oscillations in $\Delta\beta$ over many batches). For both these situations, a tuning factor, d_1, can be included to moderate the corrections:

$$d_1(w_D - w_B) = D\Delta\beta \qquad d_1 < 1 \tag{2.15}$$

Thus, the third step in the methodology calculates a new setpoint trajectory for the feedback controller. The modified setpoint is implemented in the next batch, using the feedback controller described in step 2. Again, because of process/model mismatch and the simplifying assumptions of (2.14), the calculated change may not be exactly correct. Therefore, the MWD at the end of the next batch may not be the desired one (although one would expect it has improved). Steps 2 and 3 are repeated until a MWD sufficiently closed to the desired one is being produced.

2.3.4 Case Study: Polystyrene

2.3.4.1 Description of the Process Simulation

The full methodology will be demonstrated on a simulated polystyrene semi-batch reactor. The kinetics are as follows: termination occurs by combination only, and no chain transfer to solvent or an external agent occurs. A small amount of transfer to monomer occurs. An unknown but constant disturbance, impurities entering the batch with the monomer, is present. This leads to chain transfer to impurities as a source of chain termination. For this system's kinetics, β dominates the MWD development ($\beta >> \tau$).

The simulation equations are the mole balances for each of the species (monomer, initiator, solvent, and impurities). The quasi-stationary state hypothesis is assumed for the radical population. Kinetic parameters used for the process simulation are summarized in Appendix 2.2. The gel effect is simulated with a gel factor ($k_{tc,gel\ effect} = g_t \times k_{tc}$, with $g_t \leq 1$) based on the free volume theory (Marten and Hameliec, 1992). Reaction volume shrinkage is also included in the simulation. Normally distributed random noise is added to the heat release measurement.

Calculation of the true molecular weight distribution is as follows. At regular small intervals (Δt) in the simulation, the values of the parameters β and τ (2.2) and mass of polymer produced in the interval ($MW_{mon} R_p V \Delta t$) are calculated. The cumulative MWD is evaluated at the end of the batch by calculating the instantaneous distributions from the values of β and τ (2.3) and then taking a weighted sum using the mass of polymer produced in each instant. The MWD is assumed to be measured at 100 equispaced chain length values.

2.3.4.2 Application of the Methodology

Several case studies have been simulated, however only one representative case will be discussed here. The decomposition step will be omitted in this example since the method was previously demonstrated. It is assumed that $\tau = 0$ in the decomposition (this is equivalent to assuming that chain transfer to monomer and impurities are negligible). The steps followed for decomposing the desired distribution are exactly as described in the previous section and will not be repeated here. Table 2.2 summarizes the operating policy for the first batch.

Due to the assumption that $\tau = 0$, the initial decomposition will provide incorrect instantaneous distributions even if $\beta_{estimated} = \beta_{true}$ and perfect setpoint tracking is achieved. A second source of error is introduced with kinetic parameter mismatch. There is a 20% error in the value of the k_p rate constant used for estimating β online (2.11). As well, changes in the radical termination rate due to the gel effect in the process are not exactly

known. A crude approximation to the gel effect is included in the parameter k_{tc} (used to calculate $\beta_{estimated}$). These two forms of parameter mismatch cause $\beta_{estimated}$ to deviate from β_{true}.

Table 2.2. Results of Decomposition for the Polystyrene Example

Set of instantaneous distributions (specified by β value)	Mass fraction of Polymer
$\beta_1 = 0.005$	0.4
$\beta_2 = 0.015$	0.6

The decomposition provides the setpoint trajectory for the first batch (Table 2.2). Before proceeding, several operating decisions must be made. First, the order in which the β's will be implemented must be specified. Since β is inversely proportional to monomer concentration, all batches will be operated from low β_{sp} to high β_{sp} to maximize conversion. Second, the manipulated variable must be selected. The choice of a convenient manipulated variable depends on the specific process. For this system, the batch will be operated isothermally (T = 75 °C) with only the monomer and initiator flowrates available as manipulated variables. Note that the choice of manipulated variable is a process specific decision. If it is preferred to operate in pure batch mode, the temperature could be used as the sole manipulated variable.

The flowrate of monomer was selected as the primary manipulated variable to control the value of β. However, in order to execute the changeover from $\beta_{sp} = 0.005$ to $\beta_{sp} = 0.015$, monomer concentration must decrease. Since monomer flowrate has a physical lower limit of zero, the transition may be too slow. Therefore, the flowrate of initiator will be used to execute the transition between the two β levels. This amounts to adding a slug of initiator once 40% of the polymer has been made.

The third decision that must be made is the mass of polymer to be produced. Fundamentally, the final distribution will depend on the fraction of the polymer made at

each condition. Changeovers (between β levels) are executed based on the mass fraction of polymer made (which is estimated using the online heat released measurement). For this example, a total polymer mass of 400 kg is desired (final conversion of around 85%) therefore the changeover between setpoint levels will be executed after 160 kg are made. The initial monomer and initiator charges are 5.5 and 0.03 mol/L, respectively. The initial reaction volume is 500L.

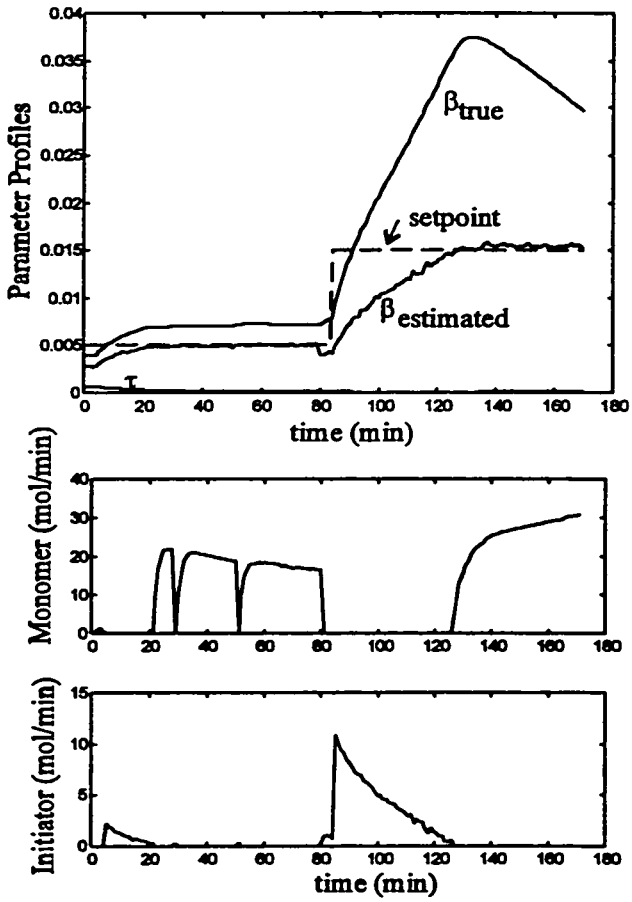

Figure 2.2. Selected profiles from the first batch

A geometric nonlinear controller (Appendix 2.3) is used to control $\beta_{estimated}$ using the flowrate of monomer (Soroush and Kravaris (1992); Clarke-Pringle and MacGregor (1997)). While a PI controller has also been used successfully, the nonlinear controller provides an eloquent solution for this system. Its main advantage is its easy implementation; only one set of tuning parameters are needed, and these are set based on the desired closed loop response.

The initiator flowrate is adjusted using a simple banded Proportional-only controller. If the error (β_{sp}-β) exceeds a certain preset value, the Proportional-only controller gives a shot of initiator. In this manner, large negative deviations are eliminated and the changeovers are executed quickly and efficiently.

Figure 2.2 shows some selected profiles from the first batch. The true value of β is time varying due to the gel effect, and the noise on $Q_{r,meas}$ is visible in the estimated value

of β. Note that the selected reactor charge does not give the desired value of β initially (this is another source of error that must be compensated for). Final conversion (not shown) is around 85%.

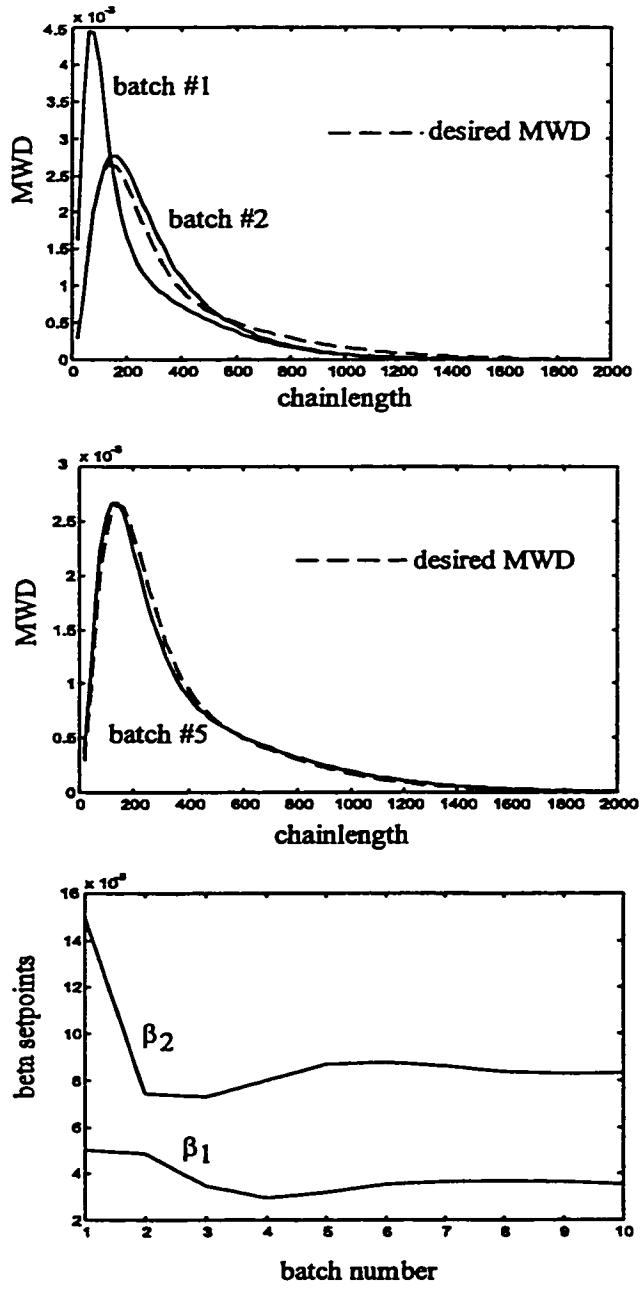

Figure 2.3. Results of batch-to-batch updating

Figure 2.3 shows the MWD from a number of successive batches. The top plot shows the MWD after the first and second batches, as well as the desired MWD. Clearly, due to the various errors, the MWD after the first batch is very different from the desired distribution. Using this MWD, the batch-to-batch optimizer calculates new values for the two β levels ($\Delta\beta$ from (2.14)), thus providing a new setpoint trajectory for the second batch. The MWD from the second batch shows much improvement, however there is still some error and so another batch correction is calculated. After only five batches, the MWD being produced is very close to the desired MWD, and the batch-to-batch corrections are small. Figure 2.3 also shows the batch-to-batch adjustments of the two β_{sp} segments (shown for 10 batches).

A second interesting case that was simulated will be briefly discussed here. Perhaps a conversion higher than 85% is desired, however it is difficult (if not impossible) to achieve while feeding monomer. An alternative policy may be as follows. Operate as described above (manipulate monomer and initiator feeds so that $\beta_{estimated}$ follows the desired trajectory). However, once the required trajectory is finished, instead of ending the batch, turn the monomer and initiator feeds off and run the reactor until the desired conversion is achieved. This action introduces further error in the final MWD. The batch-to-batch optimizer was implemented, successfully, to correct for this type of error. The methodology's success can be attributed to the fact that it is based on a very approximate model, therefore it is able to compensate for a wide range of observed errors in the MWD, regardless of their source.

2.4 Methodology for Linear Polymers with Two Dominant Parameters

2.4.1 Decomposition

The methodology for the second class of linear polymers is an extension of that for the first class. Once again, the desired MWD can be decomposed as a series of instantaneous distributions. The constituent distributions are now functions of both τ and β:

$$W(r)_{cumulative,desired} = f_1 \times W_1(r,\beta_1,\tau_1) + f_2 \times W_2(r,\beta_2,\tau_2) + \ldots + f_n \times W_n(r,\beta_n,\tau_n) \quad (2.16)$$

More degrees of freedom are introduced when τ is significant and the optimization becomes more general:

$$\min_{f,\tau,\beta}\left(W_D - \sum_{i=1}^{n} f_i W_i(\tau_i,\beta_i)\right)$$
$$\text{s.t} \quad \sum f_i = 1, \quad \tau_i,\beta_i > 0 \quad (2.17)$$

The number of constant segments must be pre-specified. The optimization returns the values of f, τ and β that provide the best set of constituent distributions. Different numbers of segments (n = 2, 3, ...) can be attempted to obtain the best decomposition.

Consider again as a brief example the distribution that was decomposed in the third section. Table 2.3 summarizes the results when the optimization finds the best τ *and* β trajectories to achieve the desired distribution:

Table 2.3. Decomposition of MWD for Linear Polymer with Two Parameters

Number of Intervals (preset)	Optimal fractions	Converged value of τ	Converged value of β	Value of Optimization Criteria
2	0.398 0.602	0.0002813 0.0009675	0.0001211 0.0001	1.0E-4

This optimization exhibits many local minima. As a result, starting point and lower and upper bounds will affect the 'optimal' solution. Prior to arriving at the numbers in Table 2.3, many different starting points were attempted. When more than two intervals were preset in the optimization, at least one fraction converged to zero.

Therefore, this distribution is best represented by only two constituent distributions. Figure 2.4 shows the desired distribution (dashed) and that from Table 2.3 (solid):

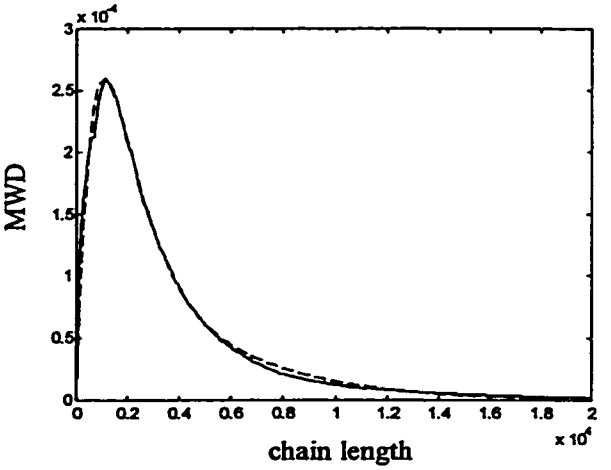

Figure 2.4. Decomposition of desired MWD for two parameter polymer

2.4.2 Production of Constituent Distributions

When both τ and β are significant, there are two control loops to implement. This is a direct extension of the one parameter method. For example, flowrate of monomer can be used to control $\beta_{estimated}$ and the flowrate of chain transfer agent can control $\tau_{estimated}$.

2.4.3 Batch-to-Batch Updating

The third step is batch-to-batch updating. Either the τ_{sp} or β_{sp} trajectories, or both, can be updated. If τ_{sp} is updated:

$$W(r)_{cumulative,B+1} = W(r)_{cumulative,B} + \left[\frac{\partial W(r)_{cumulative}}{\partial \tau_1} \quad \frac{\partial W(r)_{cumulative}}{\partial \tau_2} \quad \cdots\right]_{\beta_{sp},\tau_{sp}} \begin{bmatrix} \Delta\tau_1 \\ \Delta\tau_2 \\ \vdots \end{bmatrix} \tag{2.18}$$

and solve for $\Delta\tau$. If β_{sp} is updated:

$$W(r)_{cumulative,B+1} = W(r)_{cumulative,B} + \left[\frac{\partial W(r)_{cumulative}}{\partial \beta_1} \quad \frac{\partial W(r)_{cumulative}}{\partial \beta_2} \quad \cdots\right]_{\beta_{sp},\tau_{sp}} \begin{bmatrix} \Delta\beta_1 \\ \Delta\beta_2 \\ \vdots \end{bmatrix} \tag{2.19}$$

and solve for $\Delta\beta$. If both are updated:

$$W(r)_{cumulative,B+1} = W(r)_{cumulative,B} +$$

$$\left[\left.\frac{\partial W(r)_{cumulative}}{\partial \beta_1}\right|_{\tau\,constant} \cdots \left.\frac{\partial W(r)_{cumulative}}{\partial \tau_1}\right|_{\beta\,constant} \cdots\right]_{\beta_{sp},\tau_{sp}} \begin{bmatrix} \Delta\beta_1 \\ \vdots \\ \Delta\beta_n \\ \Delta\tau_1 \\ \vdots \\ \Delta\tau_n \end{bmatrix} \tag{2.20}$$

and solve for $[\Delta\beta\ \Delta\tau]^T$. Expressions for the derivatives can be easily derived from (2.3) and are given in Appendix 2.4.

Although any of the three options for updating are possible, simulations have indicated that updating only one of the β_{sp} or τ_{sp} trajectories at a time is preferable. Since the model is only approximate, in most cases it is difficult to extract enough good information to update both setpoint trajectories simultaneously.

The methodology for a polymer with two dominant parameters is illustrated with a case study in the next section.

2.4.4 Case Study: Polystyrene with Significant Chain Transfer to Solvent

The same polystyrene system will be used to demonstrate the two-parameter methodology. The kinetics of the system are those described earlier, but now significant chain transfer to solvent is added to the simulation equations. Flowrate of chain transfer agent is now available as another manipulated variable. The structures of the two true parameters are:

$$\beta = \frac{R_{tc}}{R_p} \qquad \tau = \frac{R_{fm} + R_{fS} + R_{fimp} + R_{fcta}}{R_p} \tag{2.21}$$

Assume for this example that the desired distribution can be decomposed into two constituent distributions:

$$W(r)_{cumulative,desired} = 0.6 * W_1(r, \beta = 0.001, \tau = 0.001) + 0.4 * W_2(r, \beta = 0.004, \tau = 0.001) \tag{2.22}$$

Therefore, the initial operating policy is: make 60% of the polymer while controlling β at 0.001 and τ at 0.001, and make 40% of the polymer with β = 0.004 and τ = 0.001. The total amount of polymer to be made is 500 kg.

The same nonlinear controller is used to control $\beta_{estimated}$ using the flowrate of monomer, and initiator is used for the changeovers. The β_{sp} trajectory is implemented from low to high β (again, to maximize conversion).

The flowrate of chain transfer agent is the logical choice for controlling $\tau_{estimated}$:

$$\tau_{estimated} = \frac{R_{fs} + R_{fcta}}{R_p} = \frac{R_{fs}}{R_p} + \frac{k_{fcta}[CTA]}{k_p[\hat{M}]} \tag{2.23}$$

The online heat released measurement is used to estimate monomer concentration. It is assumed that little will be known about the kinetics of chain transfer to solvent (other than it exists), and so a constant ratio of $R_{fs}/R_p = 0.0005$ is used to calculate $\tau_{estimated}$ online. It is assumed that some information about k_{fcta}/k_p and [CTA] is available.

A PI controller ($K_c = 1000$ mol, $\tau_I = 5$ min^{-1}) is used to control $\tau_{estimated}$. The process/model mismatch considered for the case study is as follows: the impurities are unknown and chain transfer to monomer is assumed negligible; the expression for $\beta_{estimated}$ contains an approximate empirical gel effect: $g_{t\,model}k_{tc}$ where $g_{t\,model} = 1 - 2.5e^{-1/X}$; there is a 30% error in the kinetic parameter k_{tc}; the ratio R_{fs}/R_p is set to a constant value (0.0005) in $\tau_{estimated}$ while its true value varies from 0.0004 to 0.0008 over the course of the batch; a 30% error in the value of $\tau_{estimated}$ is included to simulate error in various kinetic parameters. That is,

$$\tau_{true} = \frac{R_{fimp} + R_{fm} + R_{fs} + R_{fcta}}{R_p} \qquad \tau_{estimated} = 1.3\left(\frac{k_{fcta}[CTA]}{k_p[\hat{M}]} + 0.005\right) \tag{2.24}$$

Other mismatch includes an incorrect initial reactor charge; the selected charge does not give $\beta = 0.001$ or $\tau = 0.001$. Normally distributed random measurement noise is added to the heat released term.

Figure 2.5 shows some selected profiles from the first batch. True and estimated values of τ and β, along with the desired trajectories, are shown. The effect of process/model mismatch is clearly seen in the difference between the true and estimated trajectories of the parameters. Both controllers do a good job of tracking the τ_{sp} and β_{sp} trajectories.

Final conversion is just above 60%. Figure 2.6 shows the results of the batch-to-batch updating if the β_{sp} trajectory is updated. By the third batch, the MWD being produced is close to the desired and by the fifth batch, it is almost perfect. The β_{sp} updates level off nicely. Figure 2.7 shows the results of updating the τ_{sp} trajectory. Again the third

batch MWD is good, with small improvements up to the fifth batch. Note the MWD resulting from updating the τ_{sp} trajectory is slightly inferior to that resulting from updating β_{sp}. This is not surprising given that there is more error in $\beta_{estimated}$ than in $\tau_{estimated}$. Therefore, this method can provide some insight into where the process/model mismatch occurs, and perhaps an improved model can eventually result.

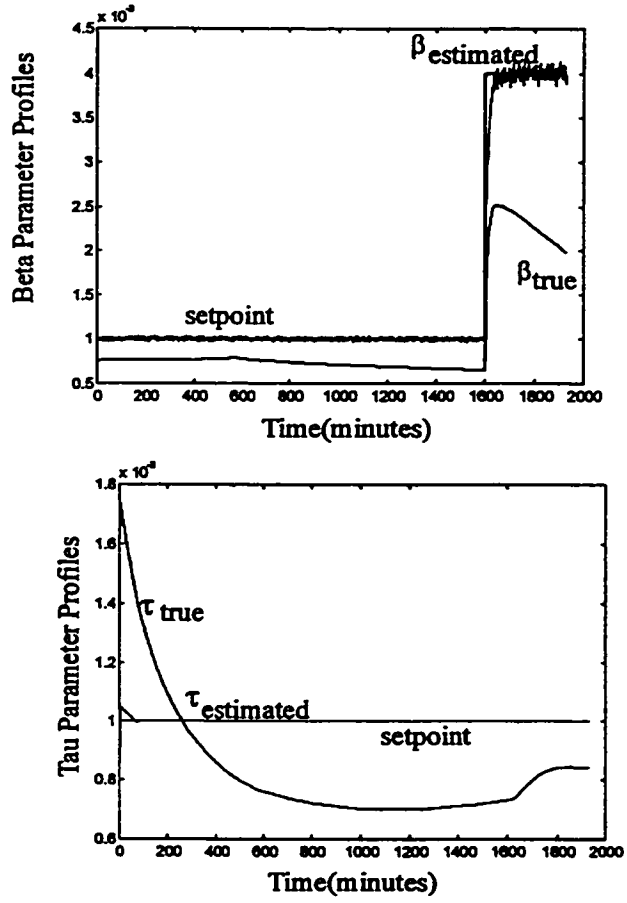

Figure 2.5. Parameter profiles for the two parameter polymer, first batch

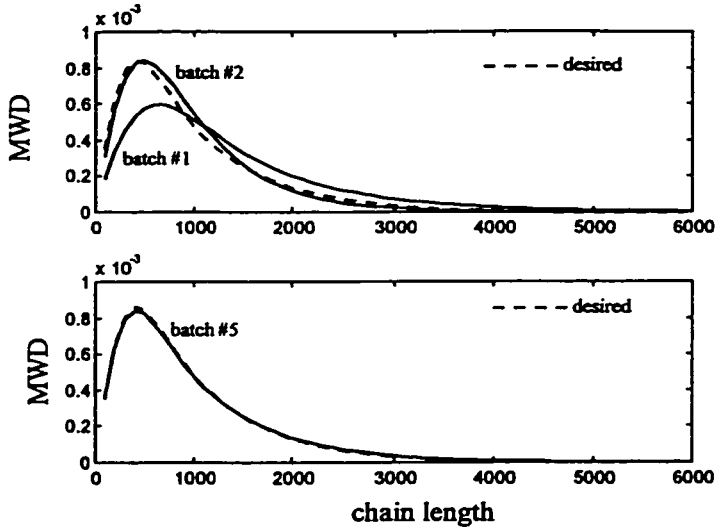

Figure 2.6. Results of batch-to-batch updating for two parameter polymer when beta is updated

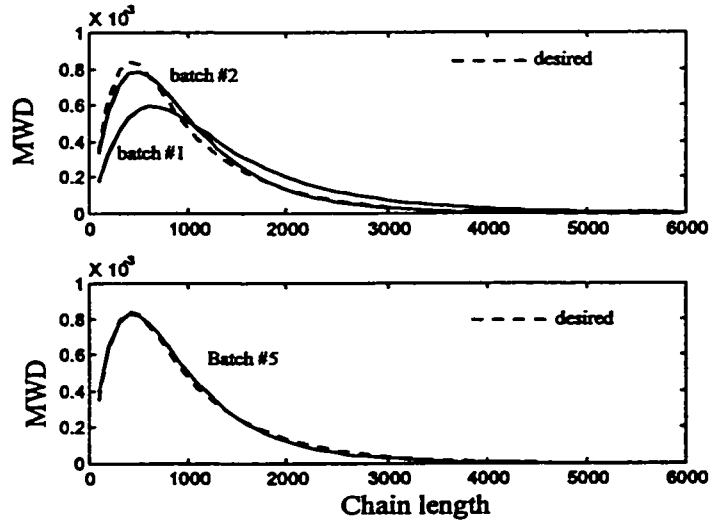

Figure 2.7. Results of batch-to-batch updating when tau is updated

Finally, Figure 2.8 shows an attempt at updating both the β_{sp} and τ_{sp} trajectories simultaneously. After 3 batches, the method has converged to an incorrect distribution.

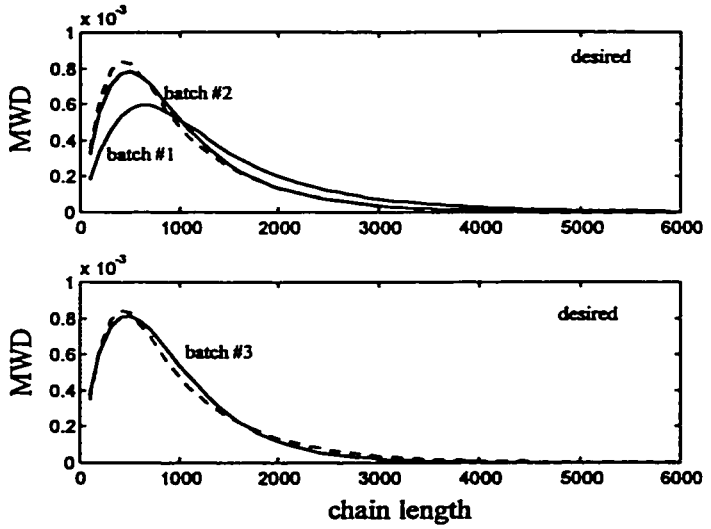

Figure 2.8. Results of batch-to-batch updating, when both tau and beta parameters are updated

This result is typical of the case studies simulated when both the β_{sp} and τ_{sp} trajectories are updated. Given the freedom to update more variables, the optimization in effect is misled by the model. That is, there is not enough information in the approximate model to correctly update both trajectories simultaneously, and as a result the optimization converges to a local minimum.

2.4.5 Simulation of Realistic MWD Measurement Noise

In the case studies presented above, there was no measurement noise on the MWD. This will never be the case in practice, and therefore the realistic generation of noise error on the MWD measurements must be addressed.

Noise on a MWD measurement is correlated noise, that is, it has structure. One cannot simply add independent random error to the molecular weight measurements at each chain length 'r'. The noise has a structure that depends on the instrument providing the measurements.

In order to realistically simulate MWD noise, replicate GPC data were obtained from Dube and Penlidis (1995). Principal Components Analysis (Wold et al (1987)) was

used to analyze the replicate MWD data and uncover the underlying noise structure. One hundred points at equally spaced chain length values were extracted from each of thirty two replicate distributions. Therefore, the matrix to be analyzed with PCA is 32x100. The data was mean centered but not scaled. The first five principal components explained 97% of the variation in the data. Table 2.4 summarizes the analysis.

Table 2.4. Summary of PCA analysis of MWD data

Principal Component	Eigenvalue	% Variation explained (cumulative)
1	8.1662 E-14	51.6
2	4.8736 E-14	82.5
3	1.5783 E-14	92.4
4	5.2331 E-15	95.7
5	2.7489 E-15	97.9

At the end of each simulated batch, noise with the a structure similar to that observed in the above real data must be added to the simulated MWD. New MWD noise is generated as follows. Recall that PCA decomposes a given matrix (here, the noise matrix) into two smaller matrices:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{T} + \mathbf{E} = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_a \end{bmatrix} \begin{bmatrix} \mathbf{p}_1^{T} \\ \mathbf{p}_2^{T} \\ \vdots \\ \mathbf{p}_a^{T} \end{bmatrix} + \mathbf{E} \qquad (2.25)$$

where a = 5 for this data. Each row in **X** will have an associated row in the score matrix (**T**). New rows of **X** (i.e. noise for the MWD of new batches) can be simulated by generating new scores using a random number generator, and using the structure established by the principal components ($\mathbf{p}_i$ vectors). Since the eigenvalues from the PCA

(λ_k) are the variances of the scores, new noise model scores can be easily generated as $t_{k,new} \sim N(0,\lambda_k)$. Then

$$\mathbf{x}_{new} = t_{1,new}\mathbf{p}_1^T + t_{2,new}\mathbf{p}_2^T + \cdots + t_{5,new}\mathbf{p}_5^T \tag{2.26}$$

In this manner, realistic noise can be added to the MWD. If one wishes to increase or decrease the absolute value of noise being added, the RHS of (2.26) can be multiplied by an amplification factor.

Figure 2.9 shows sixteen of the replicate GPC MWD data (Dube and Penlidis (1995)) and sixteen simulated replicate MWDs. The simulated MWDs were obtained by adding sixteen different noise realizations, each calculated with (2.26) as described above, to a constant MWD. The underlying noise structure appears to have been modeled well. All the following case studies in this paper will have noise added to the simulated MWD in this manner.

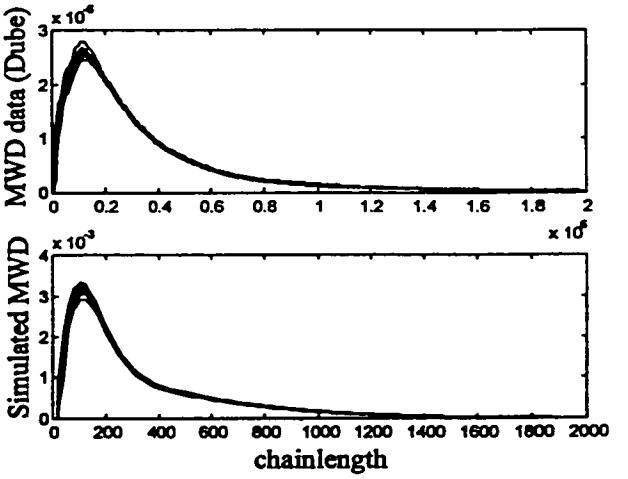

Figure 2.9. Generation of correlated MWD noise

2.5 Online Batch-to-batch Control of the MWD

2.5.1 Supervisory Statistical Process Control

In the above case studies, the process was considered optimized when the measured MWD was arbitrarily close to the desired MWD (error close to zero), and the trajectory updates leveled off. With noise on the MWD measurements, there will always be error in the MWD, and, without modification, the batch-to-batch optimization

methodology will react to the noise. This is an undesirable situation given that the noise is unpredictable from batch to batch. Attempting to correct for it will only introduce more variability into the MWD. Here, a multivariable statistical process control (MSPC) system is implemented to decide whether a given error is significant, and if a batch correction is required.

Kourti and MacGregor (1995) provide an overview of MSPC, with emphasis on problems with correlated multivariable data. A multivariable control chart based on Hotelling's T^2 statistic is implemented here.

Recall that the MWD is measured at several different chainlengths (for example, 100 for the first polystyrene case study). However, clearly these are not 100 independent variables. The true variance-covariance matrix will be of much lower dimension (rank <<100). Because of the correlated data, it is more meaningful to monitor only the significant linear combinations (scores) of the 100 MWD variables. Significant linear combinations are established by doing a Principal Components Analysis on mean-centered 'representative noise data' (the selection of representative data is discussed in more detail in the next sub-section). This analysis provides a reduced dimensional noise model against which future MWD measurements can be tested. Note that scaling of the representative data is not required since all MWD variables are assumed to have the same relative importance.

PCA provides the 'q' largest directions of variance in the data (loadings). At the end of each batch, the new MWD observation is projected onto the plane defined by the loadings to obtain the score values (t_1, t_2 ...t_q) for the batch. The Hotelling T^2 statistic is calculated for each batch from the batch scores:

$$T^2 = \sum_{i=1}^{q} \frac{t^2_{i,\text{batch b}}}{s_i^2} \tag{2.27}$$

where s_i^2 is the estimated variance of t_i. The upper control limit for T^2 is given by:

$$T^2_{UCL} = \frac{(n-1)(n+1)q}{n(n-q)} F_\alpha(q, n-q) \tag{2.28}$$

where n is the number of observations used to calculate the score variances, q is the number of variables being monitored and $F_\alpha(q,n-q)$ is the upper 100α % point of the F-distribution with q and n-q degrees of freedom. Selection of the α value is discussed in the next sub-section.

By monitoring the T^2 statistic, we are checking whether the batch scores lay within an elliptical region on the plane defined by the PCA loadings. However, we must also continue to monitor the other dimensions. This is done by evaluating the squared prediction error (SPE) of the PCA model relative to a pre-set upper bound. This amounts to monitoring the squared perpendicular distance between the MWD observation and the plane defined by the PCA loadings. The SPE is defined as:

$$SPE_{batch\ b} = \sum_{i=1}^{k} (y_{i,batch\ b} - \hat{y}_{i,batch\ b})^2 \tag{2.29}$$

where k is the number of variables (k = 100 for MWD example), y_i is the observed MWD variable and $\hat{y}_i$ is the value predicted by the PCA model, for batch 'b'. An approximate distribution, and corresponding upper control limit, for the SPE was derived by Jackson and Mudholker (1979). The upper control limit is given by (2.30), where z_α is the normal variable with the same sign at h_0 and λ_i are the eigenvalues of the variance covariance matrix of the representative noise data. Jackson and Mudholker (1979) provide an expression for accurately calculating these eigenvalues.

$$Q_\alpha = \theta_1 \left[1 - \frac{\theta_2 h_0 (1-h_0)}{\theta_1^2} + z_\alpha \frac{2(\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0}$$

$$\begin{aligned} \theta_1 &= \sum_{i=q+1}^{n} \lambda_i \\ \theta_2 &= \sum_{i=q+1}^{n} \lambda_i^2 \\ \theta_3 &= \sum_{i=q+1}^{n} \lambda_i^3 \\ h_0 &= 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \end{aligned} \qquad (2.30)$$

The supervisory SPC is implemented over the batch-to-batch optimizer. The following steps are executed at the end of each batch:

1/ Measure the MWD at the end of batch 'b', subtract the mean (calculated from the representative noise data) and project the observation vector onto the PCA loadings:

$$\mathbf{t}_{\text{batch b}} = \mathbf{P}^T \mathbf{y}_{\text{batch b}} \qquad (2.31)$$

2/ Calculate the SPE for batch 'b':

$$\begin{aligned} \text{SPE} &= \sum_{i=1}^{n} \left(\mathbf{y}_{\text{batch b}} - \hat{\mathbf{y}}_{\text{batch b}} \right)^2 \\ &= \sum_{i=1}^{n} \left(\mathbf{y}_{\text{batch b}} - \mathbf{P}\mathbf{t}_{\text{batch b}} \right)^2 \end{aligned} \qquad (2.32)$$

3/ Calculate the Hotelling's T^2 for the batch 'b' scores:

$$T^2 = \sum_{i=1}^{q} \frac{t^2_{i,\text{batch b}}}{s_i^2} \tag{2.33}$$

4/ Compare the SPE and T^2 for batch 'b' to the control limits Q_a and T^2_{UCL}, respectively. If either SPE > Q_a or $T^2 > T^2_{UCL}$, the error in the MWD is more than common cause variation and an adjustment is calculated for batch 'b+1'. If SPE < Q_a and $T^2 < T^2_{UCL}$, no adjustment for the next batch is made.

2.5.2 Implementation of a Combined SPC and Batch-to-batch Optimizer.

In an industrial setting, it is expected that things will change periodically. For example, after several batches, the feedstock may suddenly be switched to a new feed tank. This introduces a new disturbance (e.g. impurity concentration) into the process. The advantage of the combined SPC/optimization is that the optimizer does not need to be turned off. While the desired product is being produced, the observed MWD will not be different from common cause variation. If a new disturbance occurs that causes an error in the MWD, the SPC scheme will indicate that a significant error has occurred and that a batch correction is in order. In this section, the combined SPC/optimizer will be demonstrated on the β-dominant polymer shown earlier.

To implement the combined scheme, the PCA model and the control limits for T^2 and SPE must be obtained from historical data which represent common cause variation. For the example in this section, the representative data is generated by simulating 50 batches in which the β_{sp} trajectory is constant (i.e. not updated between batches). Noise on the Q_r measurements is added during the batch, and correlated measurement noise is added to the MWD after the batch. Therefore, the variation in the data from these 50 batches represents variation due to MWD measurement error and common cause process variation. In practice, the SPC charts would have to be based at first on whatever data were available that were representative of 'good' operation in the past, and then updated once sufficient data were collected under the new control scheme.

A PCA of the representative data set described above (50 simulated batches) show two significant directions. Therefore, only two scores are monitored. A decision must also be made for the significance level (α) used for the control limits on the T^2 and SPE charts. A larger significance level (α = 0.05 vs. 0.01) implies that one is taking a larger chance in deciding to make a correction when none is needed. At the same time, one is reducing the probability of not calling for a correction when one is needed. The choice obviously depends upon the current circumstances. If one is in an optimization phase, one may prefer to accept a few unnecessary corrections in order to ensure that one gets close to the optimum. However, once the process is optimized, and one is in monitoring mode, it is preferable not to make unnecessary corrections. These ideas are summarized below in the steps for implementing the combined SPC/optimization scheme:

1/ Start with α = 0.05 (optimization mode);

2/ If either of the control charts exceed the upper control limit, update the β_{sp} trajectory for the next batch;

3/ If three sequential batches call for no correction, indicating the process is close to its optimum, adjust the α value to 0.01 (monitoring mode);

4/ If a batch with a significant error occurs, recommence the optimization phase and reset the α value to 0.05. Repeat steps 2-4.

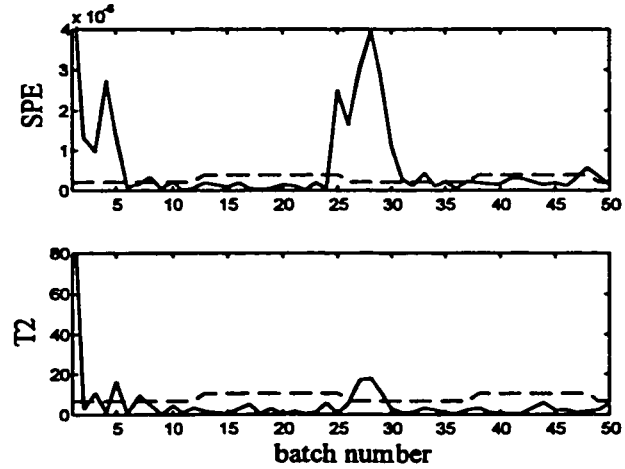

Figure 2.10. Combined SPC/optimization scheme

The previous case study for polystyrene with $\beta >> \tau$ is used here to illustrate the combined SPC/optimization strategy. Measurement noise is now added to the MWD. Fifty successive batches are simulated, and after batch 24, the feedstock is changed to a new monomer feed tank with a different level of impurities, thus introducing a new disturbance into the system. Figure 2.10 shows the SPE, T^2 and their limits for the 50 batches. The first batch shows significant error due to the initial process/model mismatch. Within nine batches, both the SPE and T^2 indicate only common cause variation, and after twelve batches, the control limits increase (α goes from 0.05 to 0.01). No batch-to-batch adjustments are made from batches 12 to 24. At batch 25, the feedstock is changed. This new disturbance is clearly seen in the two plots. By batch 32, the process is once again optimized and the control limits increase after batch 38. Figure 2.11 show the batch-to-batch β_{sp} trajectory updates.

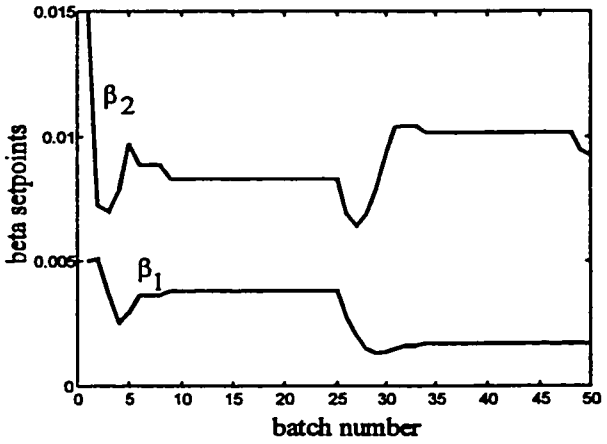

Figure 2.11. Setpoint adjustments using SPC/optimizer

2.6 Discussion of Results

In this chapter, a new batch-to-batch optimization methodology for producing a desired MWD has been presented. The advantages of the proposed batch-to-batch optimization methodology are its simplicity and flexibility. Fundamental polymerization knowledge is used to simplify the optimization problem, thereby avoiding the need for a complex on-line numerical optimization routine and issues relating to its convergence. The approach shows that for linear polymers, any desired MWD can be approximated by a combination of a few constituent distributions. Each of these component distributions can be produced by controlling one (or possibly two) parameters (τ,β) representing the ratios of kinetic rates. The method allows for great flexibility in choosing the variables to

manipulate (batch temperature, monomer or initiator flowrates) since any combination of them can be used to achieve a desired value of τ or β. These decisions can therefore be made with the specific process and its constraints in mind. Furthermore, this method has many desirable features when compared to other batch-to-batch optimization or MWD control applications: only an approximate fundamental model is required; historical plant test data is not needed and on-line (within batch) measurements of the full MWD are not required.

For the case studies presented, only two constant segments are discussed, however up to four has been simulated successfully. Extension to more is not expected to be a problem.

With respect to process/model mismatch, the robustness of the adjustments to model mismatch decreases with increasing level changes in the τ and/or β setpoint trajectory. That is, to make a bimodal distribution (e.g. two very different β levels), a better model is needed to converge to the desired MWD. Transitions must be very sharp and clear (for example, one may want to manipulate two variables, such as temperature and flowrate of initiator, to achieve a sharp transition).

When decomposing the desired MWD, it was assumed that all parts of the distribution are equally important. Perhaps for a given process, it is critical to match the long tail, and less important to match the short chains. To account for this in the methodology, one can weight the various chainlengths accordingly. The optimization would have the general form:

$$\min_{\underline{f}} \left(\mathbf{S}(\mathbf{W}_d - \mathbf{W}_i \mathbf{f}) \right)$$

$$\text{s.t.} \quad \sum_i f_i = 1 \tag{2.34}$$

where S is a diagonal weighting matrix. More important sections of the MWD would have a higher weight. This is extended to the updating by:

$$\mathbf{Se} = \mathbf{SD}\Delta\beta \tag{2.35}$$

Therefore if it was important to have a significant tail and the small chains were less critical, one could weight the distribution accordingly in order to focus the batch-to-batch corrections.

2.7 Concluding Remarks

In this chapter, a new approach to controlling the MWD of linear polymers in batch and semi-batch reactors is presented. The approach combines a new method for batch-to-batch optimization with a multivariable statistical process control (MSPC) scheme. The methodology is shown, through case studies, to efficiently return the process to the desired MWD within a few batches for a wide range of process/model mismatch.

By combining the batch-to-batch optimization with a multivariate SPC monitoring scheme, the optimizer can always be ready, but used only when a significant deviation from the desired MWD is detected. The SPC scheme is developed specifically to handle realistic, highly correlated measurement errors typically observed in MWD measurements, and avoids unnecessary batch-to-batch corrections in response to process noise and measurement errors. The combined SPC/optimization methodology is demonstrated on a system with realistic MWD measurement errors. The SPC scheme is shown to easily detect both when a re-optimization is needed and when the optimization can be terminated.

2.8 Nomenclature

D: derivative matrix

E: PCA modeling error matrix

$F_{monomer}$ = flowrate of monomer, moles/min

f = mass fraction of polymer

g_t = gel effect factor

H_r = heat of propagation (cal/mole)

[I] = initiator concentration (mole/L)

k_p = propagation rate constant (L/min mol)

k_d = initiator decomposition rate constant (min^{-1})

k_{tc} = radical termination by combination rate constant (L/mol min)

k_{td} = radical termination by disproportionation rate constant (L/mol min)

k_{fimp} = chain transfer to impurities rate constant (L/mol min)

[M] = monomer concentration (mole/L)

M = total moles of monomer (moles)

MW_{mon} = molecular weight of the monomer (g/mol)

P = principal component matrix

$\mathbf{p}_i$ = principal component

Q_r = total heat released (cal/min)

r = chain length

[R•] = radical concentration (moles/L)

R = universal gas constant (= 1.987 cal/mol K)

R_p = rate of propagation (mol/L min)

R_{fm} = rate of chain transfer to monomer (mol/L min)

R_{fs} = rate of chain transfer to solvent (mol/L min)

R_{td} = rate of termination by disproportionation (mol/L min)

R_{tc} = rate of termination by combination (mol/L min)

R_{fimp} = rate of chain transfer to impurity (mol/L min)

R_{fT} = rate of chain transfer to a small molecule (mol/L min)

SPE = squared prediction error

s_i^2 = estimated variance of variable 'i'

T = temperature (°C)

T = score matrix

T^2 = Hotelling's T^2 statistic

t = time (min)

$\mathbf{t_I}$ = score vector

t_i = value of score i

V = reaction volume (L)

X = MWD data matrix

W(r) = mass fraction of polymer of length r

$\mathbf{w_B}$ = vector of molecular weight distribution values from batch B

$\mathbf{w_D}$ = vector of desired molecular weight distribution values

$\mathbf{W_I}$ = matrix of instantaneous distributions

β: dimensionless kinetic parameter

τ: dimensionless kinetic parameter

α: significance level

Appendix 2.1. Derivation of Derivative Term

If $\beta >> \tau$, the instantaneous distribution for any given value of β_i is:

$$W_i(r,\beta_i) = \frac{\beta_i{}^3}{2}(r-1)r\left(\frac{1}{1+\beta_i}\right)^{r+1} \tag{A.1}$$

The cumulative distribution is a sum of 'n' instantaneous distributions:

$$W(r)_{cumulative} = f_1 \times W_1(r,\beta_1) + f_2 \times W_2(r,\beta_2) + \ldots + f_n \times W_n(r,\beta_n) \tag{A.2}$$

Therefore,

$$\frac{\partial W(r)_{cumulative}}{\partial \beta_i} = f_i \frac{\partial W_i(r,\beta_i)}{\partial \beta_i}$$

$$= f_i \frac{1}{2}(r-1)r\left(\beta_i{}^3 \frac{\partial\left(\frac{1}{1+\beta_i}\right)^{r+1}}{\partial \beta_i} + \left(\frac{1}{1+\beta_i}\right)^{r+1}(3\beta_i{}^2)\right) \tag{A.3}$$

$$= f_i \frac{1}{2}(r-1)r\left(-\beta_i{}^3(r+1)\left(\frac{1}{1+\beta_i}\right)^{r+2} + \left(\frac{1}{1+\beta_i}\right)^{r+1}(3\beta_i{}^2)\right)$$

Rearranging, with (r+1) ~ (r-1) and (r+2) ~ (r+1) when r is large,

$$= f_i \frac{1}{2}(r-1)r\left(-\beta_i{}^3(r+1)\left(\frac{1}{1+\beta_i}\right)^{r+2} + \left(\frac{1}{1+\beta_i}\right)^{r+1}(3\beta_i{}^2)\right)$$

$$= f_i(r-1)\left(-\frac{\beta_i{}^3}{2}(r+1)r\left(\frac{1}{1+\beta_i}\right)^{r+1}\right) + f_i\left(\frac{3}{2}\beta_i{}^2(r+1)r\left(\frac{1}{1+\beta_i}\right)^{r+1}\left(\frac{\beta_i}{\beta_i}\right)\right) \quad (A.4)$$

$$= f_i(r-1)(-W_i(r,\beta_i)) + f_i\left(\frac{3}{\beta_i}\right)W_i(r,\beta_i)$$

$$= f_i\left(\frac{3}{\beta_i} - (r-1)\right)W_i(r,\beta_i)$$

A similar expression results when $\tau >> \beta$.

Appendix 2.2. Summary of Simulation Parameters, Polystyrene

Parameter	Value
$k_{p\|60°C}$	8700 L/mol min (1)
E_p	6.4 kcal/mol (1)
initiator efficiency	0.5 (3)
$k_{d\|60\ °C}$	8.5×10^{-6} min^{-1} (1)
E_d	29.5 kcal/mol (1)
$k_{tc\|60\ °C}$	6×10^{8} L/mol min (3)
k_{td}	0
E_t	1 kcal/mol (3)
k_{fimp}	$10k_{p\|60\ °C}$
k_{fm}	1 L/mol/min (2)
E_{fm}	6.36 kcal/mol (2)
H_r	17 kcal/mol (1)

1. Hameleic, A.E. and Friis, N., *Introduction to Chain Polymerization Kinetics*, Short Course Notes, Department of Chemical Engineering, McMaster University, 1975.

2. Marten, F.L.; Hameliec, A.E; High-Conversion Diffusion-Controlled Polymerization of Styrene I, *J. of App. Poly. Sci*, **1982**, *27*, 489.

3. Zhu, S., Chemical Engineering Course notes, CHEE 6B3, 1993

Appendix 2.3. Derivation of Nonlinear Controller

This system is relative order one, meaning that the first derivative, $d\beta_{estimated}/dt = f(F_{monomer})$. Therefore, an appropriate expression for the closed loop behavior of the system is given by:

$$\frac{de}{dt} + d_1 e + d_2 \int e\,dt = 0$$

$$\text{where } e = \beta_{sp} - \beta_{estimated} \tag{C.1}$$

To derive the controller equation for the flowrate of monomer, an expression for de/dt must be obtained:

$$\frac{de}{dt} = \frac{d(\beta_{sp} - \beta_{estimated})}{dt} = -\frac{d\beta_{estimated}}{dt} \tag{C.2}$$

Recall the expression for $\beta_{estimated}$:

$$\beta_{estimated} = \frac{g_{t\,model} k_{tc} Q_{r,meas}}{(k_p[\hat{M}])^2 V H_r} \tag{C.3}$$

So, assuming isothermal operation,

$$\frac{d\beta_{estimated}}{dt} = \frac{k_{tc}}{k_p^2 H_r} \frac{d\left(g_{t\,model} Q_{r,meas} / [\hat{M}]^2 V\right)}{dt} \tag{C.4}$$

After a long but straightforward derivation, using the expressions in (C.5), one can write the derivative (C.4) as a function of the flowrate of monomer.

$$[\hat{M}]^2 V = \hat{M}\left(\frac{\hat{M}}{V}\right)$$

$$\frac{d\hat{n}_m}{dt} = F_{monomer} - \frac{Q_{r,meas}}{H_r}$$

$$\frac{dV}{dt} = \frac{F_{monomer} MW_{monomer}}{\rho_{monomer}} + MW_{monomer} \frac{Q_{r,meas}}{H_r}\left(\frac{1}{\rho_{polymer}} - \frac{1}{\rho_{monomer}}\right)$$

$$\frac{dQ_{r,meas}}{dt} = \frac{Q_{r,meas}\big|_{t+1} - Q_{r,meas}\big|_t}{\Delta t}$$

$$\frac{dg_{t\,model}}{dt} = \frac{g_{t\,model}\big|_{t+1} - g_{t\,model}\big|_t}{\Delta t} \tag{C.5}$$

Plugging the full expression for the derivative into (C.1) and rearranging for $F_{monomer}$ gives:

$$F_{monomer} = \frac{1}{\left(2[\hat{M}] - [\hat{M}]^2 \frac{MW_{monomer}}{\rho_{monomer}}\right)}\left[\frac{-(V[\hat{M}]^2)^2}{Q_{r,meas}}\left(\frac{k_p{}^2 H_r}{g_{t\,model} k_{tc}}\left(d_1 e + d_2 \int e\,dt \quad \ldots\right.\right.\right.$$

$$\left. - \left(\frac{k_{tc}}{k_p{}^2 H_r}\right)\left(\frac{Q_{r,meas}}{(V[\hat{M}]^2)}\right)\left(\frac{dg_{t\,model}}{dt}\right)\right) - \frac{1}{V[\hat{M}]^2}\frac{dQ_{r,meas}}{dt}\Bigg) + \frac{2[\hat{M}]Q_{r,meas}}{H_r} \quad \ldots$$

$$\left. + [\hat{M}]^2\left(\frac{Q_{r,meas}}{H_r} MW_{monomer}\left(\frac{1}{\rho_{solvent}} - \frac{1}{\rho_{monomer}}\right)\right)\right] \tag{C.6}$$

Although the expression is cumbersome, no information beyond that required to estimate β online is used. The parameters d_1 and d_2 are the tuning parameters and are set according to the desired closed loop response. They are constant during the batch and so tuning the controller is very simple. An equivalent PI may require different sets of tuning parameters to achieve the same level of control for the entire batch.

It is assumed that the heat released is measured every minute, therefore the controller executes once a minute (Δt = 1min). The values of d_1 and d_2 used in the simulations are 0.02 and 0.0001 respectively.

Appendix 2.4. Expressions for the Derivatives

$$\left.\frac{\partial W_i(r,\beta_i,\tau_i)}{\partial \tau_i}\right|_{\beta_i \text{ constant}} = \frac{W_i(r,\beta_i,\tau_i)}{(\beta_i+\tau_i)} - (r+1)W_i(r,\beta_i,\tau_i) + (\beta_i+\tau_i)\left(1+\frac{\beta_i}{2}(r-1)\right)r\left(\frac{1}{1+\beta_i+\tau_i}\right)^{r+1} \tag{D.1}$$

$$\left.\frac{\partial W_i(r,\beta_i,\tau_i)}{\partial \beta_i}\right|_{\tau_i \text{ constant}} = \frac{W_i(r,\beta_i,\tau_i)}{(\tau_i+\beta_i)} - (\tau_i+1)(r+1)W_i(r,\beta_i,\tau_i) + r(\tau_i+\beta_i)\left(\frac{\tau_i}{2}+\beta_i\right)(r-1)\left(\frac{1}{1+\tau_i+\beta_i}\right)^{r+1} \tag{D.2}$$

3. Polymer Product Quality Control in Reduced Dimensional Spaces

3.1 Introduction

Control of the full MWD was addressed in the previous chapter and has been the subject of several publications. The approaches can be categorized into two groups: explicit and implicit control of the MWD. As is evident by the names, the former classification implies that the full distribution is controlled directly, and typically involves measuring or estimating the full MWD. Implicit control of the MWD involves controlling the full distribution indirectly, by acting on one or more averages and/or the polydispersity.

Control of the full MWD (or equivalently, the entire quality variable space) is not a trivial problem, although some solutions have been proposed, as discussed in the last chapter. Many proposed solutions require a detailed mechanistic model, and treat the simplest problem of linear polymers. Furthermore, while it would be desirable to control the full MWD, it is not desirable to measure it online. Measurements such as the average or breath of the distribution are much easier to obtain. Implicit control of the MWD therefore is the most common approach in industry, and has been the subject of many publications (Hicks et al (1969); Choi and Butala (1991); Ellis et al (1994); Houston and Schork (1987)) . The effect of the many choices that must be made for implicit control of the MWD is the subject of this chapter.

While control of a distribution average is a common approach for controlling the MWD, good control of the average does not necessarily result in good control of the full distribution. In fact, it can sometimes result in a less desirable distribution. It is shown in this chapter that as a controller eliminates a disturbance in the controlled variables (for example, the weight average chain length), it transfers and can possibly inflate the disturbance in the remaining quality variables (e.g. the full MWD). Therefore, while it

may appear that good control is being achieved (the controlled average is maintained close to its target), the polymer quality has in fact degraded as a result of control. A simple analysis tool, called the Disturbance Inflation Factor (DIF), is introduced to quantify this effect. The DIF is used to predict which manipulated variable results in the best control of the full MWD while acting only on a single measured quality variable such as the weight-average chain length. It is further applied to show how control of the full distribution may be improved by considering other controlled variables, such as the number-average chain length, or other manipulated variables, such as a ratio of variables.

The layout of the chapter is as follows. First, batch-to-batch control of the weight-average chain length will be considered as an illustrative example. The Disturbance Inflation Factor is then derived for a single disturbance, and applied to explain the results of the example. Interpretation of the DIF and its application to the selection of effective manipulated and controlled variables is illustrated. The results are fully generalized to multiple disturbances and controlled variables, and the impact of scaling on the analysis is discussed.

3.2 Batch-to-Batch Control of the Weight-Average Chain Length

Consider batch-to-batch control of the molecular weight distribution of a polymer produced in a batch/semi-batch reactor and subject to a batch-to-batch disturbance. Measuring and controlling the full MWD is not a trivial problem and so the weight average chain length (p_w), is used as the controlled variable instead. The weight-average chain length is controlled by adjusting the initial reactor conditions (e.g. monomer or initiator charge) or the isothermal temperature at which the batch operates. This reduces a complicated multivariable problem to a simple single-input single-output (SISO) problem.

The example system is polystyrene produced in a simulated solution semi-batch reactor. The simulation equations are the mole balances for each of the species (monomer, initiator, solvent, and impurities). The quasi-stationary state hypothesis is

assumed for the radical population. The gel effect is simulated with a gel factor based on the free volume theory. Reaction volume shrinkage is also included in the simulation. Calculation of the full MWD is based on the instantaneous method, for linear polymers (Hameliec and Tobita (1992)). Information regarding the simulation equations and parameters was given in the previous chapter.

The reactor is operated as follows: monomer, initiator and solvent are initially charged (nominal values are 5.5 mol/L, 0.01 mol/L and 3.459 mol/L respectively). The batch is operated isothermally (nominal value 60 °C) and monomer is fed to the reactor using a pre-specified flowrate profile. The same monomer flowrate profile is implemented for every batch (note that the flowrate profile was determined via the optimization method given in the last chapter). At the end of the batch, a polymer sample is removed and weight average chain length determined. In the absence of batch-to-batch disturbances (variations), the same polymer would be produced in every batch. However, this is rarely the case. In this example, batch-to-batch variations are observed due to changes in the impurities. Impurities enter the system with the monomer charge and the monomer feed (simulated as: mols impurities = f × mols monomer); the impurities fraction (f) changes from batch-to-batch in a random walk. Figure 3.1 (top plot, dashed line) shows the variation in the weight-average chain length due to impurities for 50 simulated batches when no control action is taken (variation is also observed in the full MWD however only the average is shown).

In order to reduce the variability in the MWD, batch-to-batch control of the weight average is considered. At the end of each batch, the weight average is measured. Adjustments for the next batch are made on the basis of this measurement. There are three potential manipulated variables: the initial monomer charge, the initial initiator charge and the constant temperature at which the batch operates. Alternatively, charge of chain transfer agent could also be considered as another manipulated variable, however we limit the illustration to the former three options.

SISO batch-to-batch control is straightforward. There are no 'dynamics' to

consider (the initial conditions and batch temperature affect the weight average of the current batch only, not of future batches). To calculate the controller parameters, the gain between the weight average and each manipulated variable is required. These can be identified from a model, or from an identification experiment on the plant. For this example, simple identification experiments were simulated (the manipulated variables were perturbed independently) and the process gains estimated. These identification experiments also confirmed that the relationships between the manipulated variables and the weight-average are approximately linear, therefore linear controllers should suffice. Table 3.1 summarizes a series of batches with step changes in the inputs to identify the 3 input-output models:

Table 3.1. Summary of Process Gain Identification Batches

$[M]_0$ (mol/L)	$[I]_0$ (mol/L)	T_{batch} (°C)	weight average chain length	process gain
5.5	0.01	60	1790 (base case)	-
6.5	0.01	60	2094	305
4.5	0.01	60	1540	249
5.5	0.014	60	1610	- 45000
5.5	0.006	60	1992	- 50500
5.5	0.01	65	1461	- 65
5.5	0.01	55	2224	- 87

The resulting batch-to-batch controllers are:

$$[M]_{0\ batch,k+1} = [M]_{0\ batch,k} + \frac{1}{275}(p_{w,desired} - p_{w,k})$$
$$[I]_{0\ batch,k+1} = [I]_{0\ batch,k} - \frac{1}{47000}(p_{w,desired} - p_{w,k}) \qquad (3.1)$$
$$T_{batch,k+1} = T_{batch,k} - \frac{1}{75}(p_{w,desired} - p_{w,k})$$

where the controller parameters (e.g. 1/275) are the inverse of the process gain. Note that the algorithms in (3.1) are pure integral controllers and are optimal (in a minimum variance sense) for a random walk disturbance.

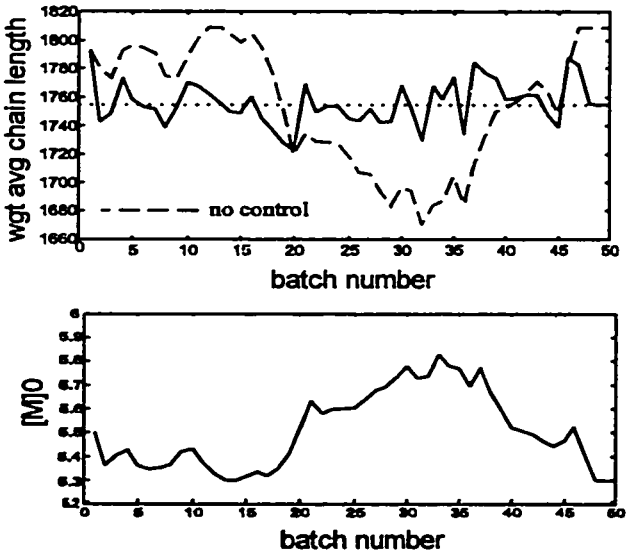

Figure 3.1. Control of the weight-average chain length

Figure 3.1 shows the batch-to-batch control of the weight-average chain length by adjusting initial monomer concentration. The setpoint for the weight-average chain length is 1755 (this is the weight-average for the desired MWD) and is also shown on the plots. The weight-average chain length plots for the other two control schemes (using initiator and temperature) are not shown but give essentially identical results.

Therefore, all three manipulated variables are capable of good control of the weight-average chain length; each achieves the theoretical lower bound of minimum variance. In terms of controlling the weight average chain length, all three are equivalent. However, the three manipulated variables are not equivalent if the full MWD is considered. Assume that the MWD from the simulated batches was also measured, at 100 different chain lengths, and the desired distribution (whose weight-average is 1755) is known.

The root mean squared error in the molecular weight distribution from the batches can then be plotted:

$$\text{MWD error} = \left(\sum_{r} (W(r)_{desired} - W(r)_{measured})^2 \right)^{1/2} \tag{3.2}$$

One may prefer to use a weighted error, or a relative error (Crowley and Choi (1997)), to emphasize critical parts of the MWD, and this can be incorporated using scaling (this will be discussed in a later section). Figure 3.2 shows the error in the MWD plotted against batch number, when the weight-average molecular weight is being controlled, using each of the manipulated variables. While all three manipulated variables control the average well, in this case only the initiator reduces the variability in the full distribution. Manipulating monomer and temperature actually inflates the effect of the disturbance on the MWD and leads to increased variability. Clearly, it would be very desirable to predict this result a priori so that the appropriate manipulated variable is used and poor quality polymer is not produced.

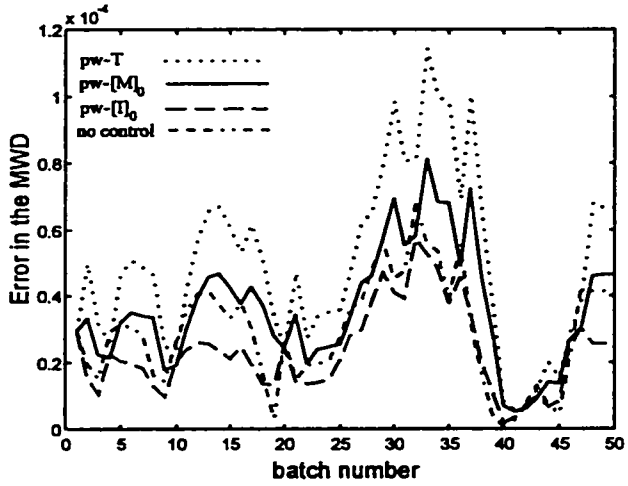

Figure 3.2. Error in the full MWD while controlling the weight-average chain length

In order to understand what is happening, consider first a generic system with two output variables y_1 and y_2. These can be viewed as in Figure 3.3, with the origin as the desired values of these variables. Also shown in Figure 3.3 are the gain vectors associated with two manipulated variables (u_1 and u_2) and the direction of a disturbance (d) (in this example the 'gain vector' of $u_2 = [1\ 2]^T$, implying that a unit change in u_2 causes a change of 1 in y_1 and a change of 2 in y_2. Each manipulated variable will have its own characteristic gain vector).

If a control loop with integral action is implemented between y_1 and u_1, the controller will adjust the process until $y_1 = 0$. The manipulated variable u_1 can move the process only in the direction of its gain vector, assuming a linear system. The dashed line shows the path that this controller will take in controlling y_1 in response to the disturbance. The result of this action is to reduce the error in y_1 to zero but leave an error d_{c1} in y_2.

Since $\|d_{c1}\| \lhd \|d\|$ there is an overall effect of reducing the magnitude of the error in the (y_1,y_2) space. If, however, a control loop is set up between y_1 and u_2, the controller still reduces the error in y_1 to zero but actually inflates the effect of the disturbance (d_{c2}) in the space of y_2. Both appear to be functioning well if only y_1 is considered.

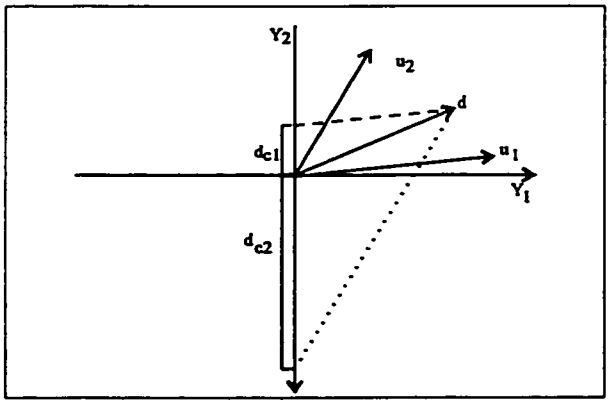

Figure 3.3. Illustrative Example

To illustrate that this is what has happened in the polystyrene system, a Principal Components Analysis on the MWD data can be performed (Wold et al (1987)). Fifty batches, with a random walk disturbance in the impurities and some independent random perturbations on $[M]_0$, $[I]_0$ and T_{batch}, are simulated and the MWD measured at 100 different chain lengths for each batch. A PCA analysis of the MWD data (centered about the desired distribution) indicates that 93% of the variation falls in the first two dimensions. Therefore, one can gain an approximate understanding of the process by looking only at this 2-D plane.

Figure 3.4 shows the two dimensional plane, with the axes being the first two principal components of the MWD data. Since the data was centered by subtracting the desired MWD from each measured MWD, the origin represents the desired location and the distance from the origin is the magnitude of the error. The manipulated variable gain vectors and the disturbance direction are also shown on the figure (these are actually projections of these vectors onto the plane of the process). In this example, the controlled variable is the weight average molecular weight, which must be represented geometrically:

$$p_w = \sum_r rW(r) = [r_1 \quad r_2 \quad \cdots] \begin{bmatrix} W(r_1) \\ W(r_2) \\ \vdots \end{bmatrix} \tag{3.3}$$

$$= \mathbf{r}_w^T \mathbf{y}$$

The weight-average chain length is a linear combination of the MWD variables. Note that since the data is centered about the desired MWD, the dot product is actually equal to the error in the average ($\mathbf{r}_w \cdot \mathbf{y} = p_w - p_{w,setpoint}$). From the geometric interpretation of the dot product, the error in the weight average chain length is the projection of the observed MWD vector (y) onto the defining vector $\mathbf{r}_w$. Since the process varies primarily in the $\{\mathbf{p}_1, \mathbf{p}_2\}$ plane, the projection of $\mathbf{r}_w$ into $\{\mathbf{p}_1, \mathbf{p}_2\}$ is shown in Figure 3.4. This vector defines the axis for the chain weight (just as [1 0] defined the axis for the variable y_1 in the earlier example). An axis perpendicular to $\mathbf{r}_w$ is also included in Figure 3.4. If the process is at any point along this perpendicular axis, p_w is at its setpoint ($\mathbf{r}_w \cdot \mathbf{y} = 0$).

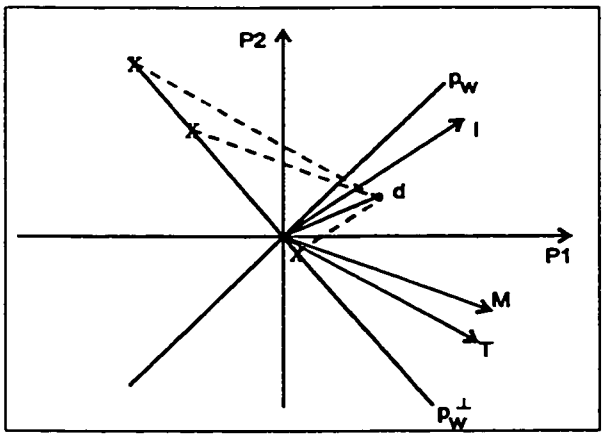

Figure 3.4. The plane of the process

However, the desired MWD is produced only if the process is at the origin.

As in the earlier y_1-y_2 example, each p_w controller moves the process to the p_w setpoint, in a direction parallel to the gain vector of its manipulated variable, with the movements of the controllers shown as dashed lines. The error in the MWD *after* controlling the weight-average chain length is given by the length of the residual disturbance vector along the axis perpendicular to $\mathbf{r}_w$. From the figure, in this study it is clear that using the initiator as the manipulated variable reduces the effect of the disturbance while using the temperature and monomer inflate the effect, with temperature being the worst. This is exactly what was observed in the simulation study, and what would be desirable to predict.

Of course, not all processes can be approximated by a plane that can be analyzed visually. A generalization of the ideas presented above is discussed in the next section.

3.3 The Disturbance Inflation Factor for a Single Disturbance

3.3.1 The Disturbance Inflation Factor

In this section, the Disturbance Inflation Factor (DIF) will be formally introduced for a single disturbance and single controlled variable. Extensions to higher dimensions are considered later. Using disturbance direction, manipulated variable gain vector directions and knowledge of the controlled variable, the DIF can be calculated. The DIF predicts the impact of controlling one variable on the full quality space (e.g. the effect of controlling the weight-average on the full MWD). To calculate the DIF, the effects of the manipulated variables and disturbance on the full MWD (i.e. the entire quality space) must be known. This implies that extra measurements may be temporarily required in order to analyze the process. However, a detailed mechanistic model is not required; all the information can be obtained from plant data.

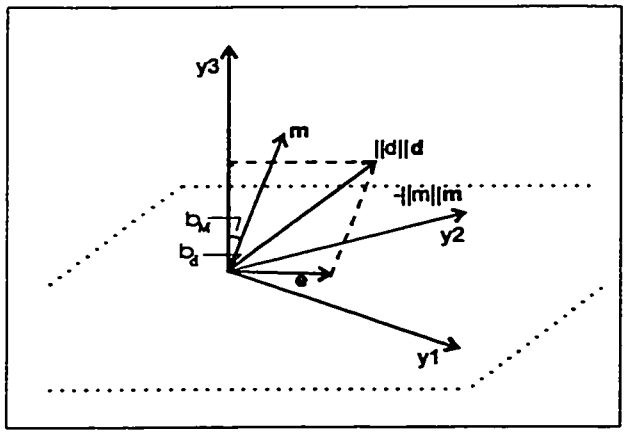

Figure 3.5. Visual interpretation of the DIF

Consider a process with 'n' important quality variables. Figure 3.5 shows visually the case where n = 3. A PCA of disturbance data provides the disturbance direction; in this case, it is a single direction represented by a vector (**d**). The normalized gain vector, **m**, of an arbitrary manipulated variable is also shown. To visually simplify the situation, assume that the variable y_3 is the controlled variable, represented by the axis vector $\mathbf{c} = [0\ 0\ 1]^T$. All variables are assumed to be centered about their desired values.

In the analysis, the orthogonal complement (OC) space (Golub and Van Loan (1989)) is very useful. The OC space of a vector (or set of vectors) is the set of all vectors which are perpendicular to the original set. The orthogonal complement space of **c**, given above, contains all vectors that are in the plane of $\{[1\ 0\ 0]^T, [0\ 1\ 0]^T\}$. The OC space is important because any controller (with integral action) acting on **c** will move the process to the orthogonal complement space of **c**. This is a direct extension of the 2-D situation discussed earlier without introducing vector spaces. The orthogonal complement space and other key algebraic concepts are reviewed in Appendix 3.1 at the end of this Chapter.

When a disturbance **d** perturbs the process, the manipulated variable will move its effect into the orthogonal complement space along the direction of the gain vector (**m**). The control action results in $y_3 = y_{3,setpoint}$ with an error vector **e** in the OC space (Figure 3.5). The error vector will be the vector sum of the disturbance and the manipulated variable move:

$$\mathbf{e} = -\|\mathbf{m}\|\mathbf{m} + \|\mathbf{d}\|\mathbf{d} \tag{3.4}$$

In order to calculate **e**, one needs the magnitude of the change in the manipulated variable, ||m||. The error in the y_3 due to the disturbance is:

$$\text{error in } y_3 = \|d\|\cos(\beta_d) \tag{3.5}$$

where β_d is the angle between the disturbance vector and controlled variable axis. In order to eliminate the error, the manipulated variable move must be:

$$\|m\|\cos(\beta_M) = \|d\|\cos(\beta_d) \tag{3.6}$$

where β_M is the angle between the manipulated variable gain vector and the controlled variable axis. Solving for ||m|| and substituting into (3.4):

$$\begin{aligned} \mathbf{e} &= -\frac{\|d\|\cos(\beta_d)}{\cos(\beta_M)}\mathbf{m} + \|d\|\mathbf{d} \\ &= \left(-\frac{\cos(\beta_d)}{\cos(\beta_M)}\mathbf{m} + \mathbf{d}\right)\|d\| \end{aligned} \tag{3.7}$$

The Disturbance Inflation Factor is defined as:

$$\text{DIF} = \|\mathbf{e}\| / \|\mathbf{d}\| = \left\| -\frac{\cos(\beta_d)}{\cos(\beta_M)}\mathbf{m} + \mathbf{d} \right\| \tag{3.8}$$

The DIF is positive and indicates to what extent a manipulated variable inflates or deflates the effect of the disturbance in the uncontrolled quality space. A value greater than one indicates the effect of the disturbance on the quality space is inflated, a value less than one indicates the effect has been deflated and a value of one means the manipulated variable simply transfers the disturbance from the controlled variable to the

uncontrolled quality space. Each manipulated variable will have a characteristic DIF, for a given controlled variable and disturbance.

Returning to the MWD example, the DIF can be calculated for each manipulated variable, given only their gain vector directions and the disturbance direction. The controlled vector $\mathbf{c}$ (= $\mathbf{r}_w$) (normalized to unit length) is:

$$\mathbf{c} = [100\ 200\ 300 \cdots\ 10000]^T / \|\mathbf{c}\| \tag{3.9}$$

Since there are 100 quality variables, each of the gain vectors and disturbance direction vectors have a dimension of 100 x 1. Table 3.2 summarizes the Disturbance Inflation Factors.

The DIF correctly shows that, all other things being equal, the $[I]_0$ is the best manipulated variable to control the weight average chain length, if the entire MWD was important. Note that this is not a general result for polymerization, but only for the particular polystyrene process and operating policy of this example.

Table 3.2. DIF for MWD example

Manipulated Variable	DIF
Initial Initiator Concentration	0.68
Initial Monomer Concentration	1.10
Batch Temperature	1.50

Note that in order to calculate the DIF, all quality variables (e.g. the full MWD) will need to be measured for a period of time in order to capture the disturbance and manipulated variable effects in the full quality space. However, once the controller (for example p_w-$[I]_0$) has been implemented, the uncontrolled quality variables (e.g. full MWD) do not need to be measured.

There are two situations that merit attention. First, consider the case where $\mathbf{m} = \mathbf{d}$ (**d**

is 'in the span of' **m**). In this very special case, DIF = 0. Therefore, it is possible to completely eliminate the disturbance in all the quality variables (e.g. the full MWD) by only controlling one (e.g. weight-average chain length). Second, consider the case where $\beta_d = 90°$, that is, the disturbance is perpendicular to the controlled variable vector. The error vector will be equal to the original disturbance and DIF = 1. This would be an unfortunate choice of controlled variable to reject this disturbance, since the disturbance does not affect the controlled variable (it is unobservable in the controlled variable). These interpretations lead to the key point that an improvement in the full MWD can be obtained by selecting 'better' manipulated and controlled variables, without necessarily resorting to a more complex controller. This is discussed in the next two sections.

3.3.2 Choice of Manipulated Variable

As indicated above, the DIF provides some insight into selecting a manipulated variable. The perfect manipulated variable is one that is aligned exactly with the disturbance. Therefore, for a given controlled variable, the closer a manipulated variable gain vector is to the disturbance direction, the better it *may* be expected to perform. To this end, ratio control is proposed to create 'new' manipulated variables from combinations of existing ones. That is, two or more manipulated variables are adjusted in a fixed ratio to directly control the weight-average chain length (and indirectly, the full MWD). The ratio is chosen so that the gain vector of the 'new' manipulated variable is closely aligned to the disturbance. The ideas are illustrated for the polystyrene reactor presented earlier.

With three adjustable variables ($[M]_0$, $[I]_0$ and T_{batch}), there exists four possible ratio controllers: $[M]_0$-$[I]_0$, $[M]_0$-T, $[I]_0$-T_{batch} and all three together. The appropriate ratio for each is calculated by regressing the disturbance direction vector ($\mathbf{P_d}$) onto the manipulated variable gain vectors (**M**) to obtain a new manipulated variable:

$$\alpha = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{P_d} \tag{3.10}$$

This regression provides the linear combination of the manipulated variable vectors that are closest, in the least squares sense, to the disturbance vector. The associated manipulated variable direction that is most aligned with the disturbance direction is given by $\mathbf{M}^* = \mathbf{M}\alpha$. Table 3.3 summarizes the predicted DIF for the four possible ratio controllers.

The DIF indicates that if the two manipulated variables, $[M]_0$ and T_{batch}, are used in the appropriate ratio to control the weight average chain length, the full MWD should be significantly improved (using all three variables is possible too, although this results in a more complicated ratio controller, and provides marginal improvement over $[M]_0$-T).

Table 3.3. DIF for possible ratio controllers

Manipulated Variable	DIF
$[M]_0$ and $[I]_0$	0.60
$[M]_0$ and T_{batch}	0.26
$[I]_0$ and T_{batch}	0.63
$[M]_0$, $[I]_0$ and T_{batch}	0.22

For the polystyrene reactor, $\alpha = [1;-0.41]$ when $\mathbf{M} = [\mathbf{M}_{T,batch}\ \mathbf{M}_{M0}]$. Therefore, for each temperature increase of one degree, the initial monomer concentration should be decreased by 0.41 mol/L. Note that to calculate this ratio, the effect of the manipulated variables and disturbance on the full MWD is needed. However, once the ratio is calculated, only the weight-average will be measured and controlled. As before, a deadbeat controller is implemented:

$$T_{batch,k+1} = T_{batch,k} + \frac{1}{32}(p_{w,desired} - p_{w,k})$$
$$[M]_{0\ batch,k+1} = [M]_{0\ batch,k} - 0.41(T_{batch,k+1} - T_{batch,k}) \tag{3.10}$$

In (3.10), the number 32 is the new gain identified between T_{batch} and p_w, when the monomer charge is ratioed to changes in T_{batch}. Figure 3.6 shows the error in the MWD resulting from control of the weight average using the $[M]_0$-T_{batch} ratio controller. Also shown for reference is the no control error, and the error from the $[I]_0$-p_w batches. Clearly, ratio control has provided an effective option for improving control over the full MWD, while retaining the simplicity of a SISO controller.

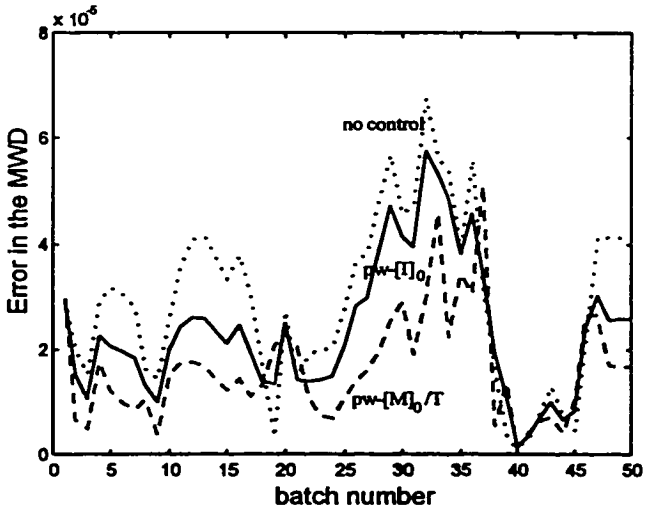

Figure 3.6. Ratio control results

3.3.3 Choice of Controlled Variable

A less obvious choice for improving control over the MWD is to select a new controlled variable. For example, controlling the number-average chain length (instead of the weight-average) may improve control over the full MWD. This would be particularly appealing if the number-average measurement is more readily accessible (e.g. through osmometry). In this section, the DIF is applied first to establish the 'ideal' control

variable, and then to provide a guide for the selection of alternative measurable controlled variables.

Before identifying the ideal control variable, a small note regarding achievable performance is in order. When dealing with systems of high dimension, there is a lower bound of achievable performance. A given manipulated variable can eliminate, at best, the portion of the disturbance that is co-linear with its gain vector. A manipulated variable whose gain vector is perfectly aligned with a disturbance can completely eliminate it; a MV whose vector is perpendicular to a disturbance cannot eliminate any of it. Most manipulated variables fall somewhere in between.

Assume that we are free to select any controlled variable (defined by its associated axis, **c**, as before) and so select $\mathbf{c} = \mathbf{m}$. That is, the controlled variable axis is chosen to be aligned with the manipulated variable gain vector. Therefore, $\beta_m = 0$ and (8) becomes:

$$\mathbf{e} = \left(-\cos(\beta_d)\|\mathbf{d}\|\mathbf{m} + \|\mathbf{d}\|\mathbf{d}\right) \tag{3.11}$$

Since $\mathbf{c} = \mathbf{m}$, β_d is also the angle between the manipulated variable and the disturbance. The coefficient for **m** in (3.11) is the magnitude of the projection of **d** onto **m**. Therefore, if the controlled variable axis is chosen as above, the manipulated variable eliminates exactly the portion of the disturbance that is co-linear with its gain vector. That is, the manipulated variable is achieving its lower bound of performance, for the given disturbance. It can be concluded that for a given manipulated variable, the 'ideal' controlled variable is one whose axis aligns with the manipulated variable gain vector.

In theory, this is very appealing. One can define a new linear combination of the quality variables, and control this new 'latent' variable. The lower bound of achievable performance (for a given manipulated variable) is attained by simple SISO control of the latent variable and no advanced control theory or expertise is needed. In practice, latent variable control may be less appealing because it would require measuring the entire quality space (i.e. the full MWD). From a MWD measurement, the latent variable is

calculated easily:

$$t = \mathbf{m}^T\left(\mathbf{MWD} - \mathbf{MWD}_{desired}\right) \tag{3.12}$$

where **m** is the manipulated variable unit gain vector.

If no other feasible alternatives exist, latent variable control is an easily implemented option (aside from requiring the entire quality space be measured). Table 3.4 summarizes the DIFs for each manipulated variable for the polystyrene process considered earlier, assuming that the ideal latent variable is used in each case (**c** = **m**).

Table 3.4. DIFs for latent variable control

Manipulated Variable	DIF
$[M]_0$	0.72
$[I]_0$	0.67
T_{batch}	0.80

Comparing the DIFs in this table with those in Tables 3.2 and 3.3 (in which the controlled variable was the weight-average chain length) shows that none of the ideal latent variable controllers significantly out-perform the $[I]_0$-p_w or $[M]_0/T$-p_w ratio controllers. This was also confirmed by simulations. For this system, latent variable control provides no improvement, and would not be worth implementing since excellent control possibilities exist by controlling p_w.

Another option for improving control of the full MWD may be to select another measurable control variable. One possibility is the number-average chain length (p_n). As with the weight-average, the number-average can be expressed as a linear combination of the MWD variables:

$$\frac{1}{p_n} = \sum \frac{1}{r} W(r) \tag{3.13}$$

Therefore, the defining axis for number-average is: $\mathbf{c}$ = [0.01; 0.005 ...0.0001]/||c||. The DIF can be calculated for $[M]_0$, $[I]_0$ and T_{batch}, given the impurities disturbance and number-average controlled variable (p_w control is included in Table 3.5 for reference).

For all three manipulated variables, SISO control of the number-average improves the full MWD (DIF < 1). The DIFs for $[M]_0$ and T_{batch} are significantly lower compared to the case in which the weight-average is controlled. Simulations confirm the numbers. Furthermore, $[I]_0$ is the best single manipulated variable, and either the number- or the weight-average can be used as the controlled variable. On the other hand, if $[M]_0$ or T_{batch} was to be used as the manipulated variable, the number-average must be used as the controlled variable, not the weight-average chain length.

Table 3.5. DIF for number-average molecular weight control

Manipulated Variable	DIF (p_n control)	DIF (p_w control)
$[I]_0$	0.73	0.68
$[M]_0$	0.77	1.10
T_{batch}	0.88	1.50

Unfortunately, none of the number-average controllers outperform the $[I]_0$-p_w controller, nor do they approach the performance of the $T_{batch}/[M]_0$-p_w ratio controller. Of course, these results are process specific, and other systems will have different results. Furthermore, many other combinations of controllers have not been examined here (for example, $T_{batch}/[M]_0$-p_n control). However, it is clear from the analysis presented that significant improvements in the full MWD can be achieved with SISO control, and many options exist to do so.

Like well-known tools such as the Relative Gain Array and the Singular Value Decomposition, the DIF can be used to select controlled and manipulated variable pairings. However, unlike these tools, the DIF explicitly considers disturbances. In this sense, the DIF is very similar to another linear analysis tool, the Partial Disturbance Gain (Havre and Skogestad (1996); Zhao and Skogestad (1997)). Further discussion of the relationship between the DIF and other disturbance analysis tools may be found in Appendix 3.2.

3.4 Control of Multiple Disturbances

3.4.1 Characterization of Disturbance Space

In the previous section, the disturbance was characterized by one direction only. However, many processes have two or more disturbances, and if these have different directions, the total effect of the (uncontrolled) disturbances may move the process in two or more dimensions. The first step in characterizing multiple disturbances is to establish the space in which the uncontrolled disturbances move the process. This can be approached three different ways (note that these approaches to characterizing the disturbance space/direction apply for the case in which a single disturbance perturbs the process as well). These are briefly outlined here, with more detail being provided in Appendix 3.3 at the end of the Chapter. First, there may exist a database in which MWD data was collected in the absence of feedback control. In such a case, the data represents disturbance data, and a Principal Components Analysis of this data will provide the defining directions of the disturbance space. Second, MWD data may exist for a time period in which known manipulated variable corrections were made. In this case, an input/output model could be identified, and the disturbance effect backed out as **D** = **Y**-{effect of input moves}. The reconstructed disturbance effect could then be analyzed by PCA as before. Finally, one may have a deterministic model (empirical or fundamental) from which the individual disturbance effects are known. These individual disturbance gain vectors would be the defining vectors of the disturbance space.

In the following section, the DIF concept will be extended to handle multiple disturbance directions. The ideas will first be presented for a single controlled variable. Simulation results from the polystyrene system will be presented for the case in which two independent disturbances perturb the process. Extensions to multiple controlled variables will also be discussed and an example presented.

3.4.2 Single Controlled Variable

When a controller rejects a disturbance, it is reacting to offset in the controlled variable. Whether this offset is created by one or more disturbances is irrelevant. Therefore, while a disturbance may be characterized by many directions, the controller will only react to one realization of the disturbance at a time. The primary difference between single and multiple disturbance directions is that with multiple directions, the observed realization of the disturbance can be any linear combination of the defining disturbance vector directions. With many disturbance directions, the performance of the manipulated variables depends on the exact realization of the disturbance. Therefore, it is only possible to calculate the minimum and maximum possible Disturbance Inflation Factors, and the associated directions of the disturbance.

Consider the general situation in which there are 'n' disturbances (defined by the unit direction vectors $\{\mathbf{d}_1, \mathbf{d}_2 ..\mathbf{d}_n\}$), one controlled variable (defined by the unit vector **c**) and one manipulated variable (unit gain vector **m**). A particular observed realization of the disturbance will be a linear combination of the individual disturbances:

$$\mathbf{d}_{arbitary} = a_1\mathbf{d}_1 + a_2\mathbf{d}_2 + ... + a_n\mathbf{d}_n \tag{3.14}$$

The Disturbance Inflation Factor is (for $\|d_{arb}\|=1$):

$$\mathrm{DIF}_{\text{arbitrary}} = \left\| \mathbf{e}_{\text{arbitrary}} \right\|$$

where (3.15)

$$\mathbf{e}_{\text{arbitrary}} = \left(-\frac{\cos\beta_{\text{d,arb}}}{\cos\beta_{\text{M}}} \mathbf{m} + \mathbf{d}_{\text{arbitrary}} \right)$$

Replacing the numerator term with vectors:

$$\mathbf{e}_{\text{arbitrary}} = \left(-\frac{\mathbf{c}^{\text{T}}\mathbf{d}_{\text{arbitrary}}}{\cos\beta_{\text{M}}} \mathbf{m} + \mathbf{d}_{\text{arbitrary}} \right)$$

$$= \left(-\frac{\mathbf{c}^{\text{T}}(a_1\mathbf{d}_1 + a_2\mathbf{d}_2 + \ldots + a_n\mathbf{d}_n)}{\cos\beta_{\text{M}}} \mathbf{m} + (a_1\mathbf{d}_1 + a_2\mathbf{d}_2 + \ldots + a_n\mathbf{d}_n) \right) \tag{3.16}$$

$$= a_1\mathbf{e}_{\text{d1}} + a_2\mathbf{e}_{\text{d2}} + \ldots + a_n\mathbf{e}_{\text{dn}}$$

$$= \left[\mathbf{e}_{\text{d1}}\ \mathbf{e}_{\text{d2}} \ldots \mathbf{e}_{\text{dn}} \right] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Therefore, the error (after control) for any arbitrary disturbance realization can be expressed as a linear combination of the errors for each of the defining disturbance vectors. Each of the these errors ($\mathbf{e}_{\text{di}}$) can be easily calculated knowing the individual $\mathbf{d}_i$'s, **m** and **c**. The maximum and minimum Disturbance Inflation Factors ($\mathrm{DIF}_{\text{min}}$ and $\mathrm{DIF}_{\text{max}}$) correspond to length of the largest and smallest vectors $\mathbf{e}_{\text{arbitrary}}$ (because of the assumption that $\|\mathbf{d}_{\text{arbitrary}}\| = 1$). The problem of finding $\mathrm{DIF}_{\text{min}}$ and $\mathrm{DIF}_{\text{max}}$ can be stated mathematically as:

$$\max_{\mathbf{a}} \left\| \mathbf{e}_{\text{arbitrary}} \right\| \tag{3.17}$$

$$\text{given: } \mathbf{e}_{\text{arbitrary}} = \mathbf{Ea} \quad \text{and subject to: } \left\| \mathbf{d}_{\text{arbitrary}} \right\| = \|\mathbf{Da}\| = 1$$

where $\mathbf{D} = [\mathbf{d}_1\ \mathbf{d}_2\ \ldots \mathbf{d}_n]$. Conceptually, we wish to search all directions in the disturbance space (all vectors $\mathbf{d}_{\text{arbitrary}}$) in order to find the disturbance realizations which result in the largest and smallest errors (after control).

If the defining disturbance vectors are orthogonal, **D** is an orthonormal matrix, $\|\mathbf{Da}\| = \|\mathbf{a}\|$, and the constraint simplifies to $\|\mathbf{a}\| = 1$. In this case, (3.17) is the mathematical statement of the singular value decomposition of the error matrix (**E**) and therefore an SVD of the matrix $[\mathbf{e}_{d1} \ldots \mathbf{e}_{dn}]$ provides the desired information ($\mathbf{E} = \mathbf{USV}^T$) (Golub and Van Loan (1989)). The largest and smallest singular values (largest and smallest elements of **S**) are equal to the minimum and maximum DIFs. The best and worst disturbance realizations are given by the corresponding right singular vectors (columns of **V**): $\mathbf{d}_{\text{arbitrary,best}} = \mathbf{Dv}_{\min}$ and $\mathbf{d}_{\text{arbitrary,worst}} = \mathbf{Dv}_{\max}$. The values of DIF_{min} and DIF_{max} provide lower and upper bounds for the DIF, given the disturbance space and the controlled and manipulated variable. The realizations of the disturbance associated with DIF_{min} and DIF_{max} are the particular directions of the disturbance that are the best and worst directions, respectively, for the given manipulated variable and controlled variable.

In order to solve (3.17) using the singular value decomposition, the defining disturbance vectors must be orthogonal. There are two ways in which a set of orthogonal disturbance direction vectors (i.e. an orthonormal **D** matrix) can be generated. For example consider a process with 'n' physical disturbances with direction vectors $\{\mathbf{d}_1, \mathbf{d}_2 \ldots \mathbf{d}_n\}$. If the disturbance space is characterized by doing a PCA on disturbance data (as described earlier), the resulting principal component vectors $\{\mathbf{d}_1^*, \mathbf{d}_2^*, \ldots \mathbf{d}_n^*\}$ are an orthogonal basis for the physical disturbance vectors $\{\mathbf{d}_1, \mathbf{d}_2 \ldots \mathbf{d}_n\}$. By default, **D** is orthonormal. Second, if the physical disturbance directions $\{\mathbf{d}_1, \mathbf{d}_2 \ldots \mathbf{d}_n\}$ are known (for example from a model), an orthogonal basis for the space spanned by $\{\mathbf{d}_1, \mathbf{d}_2 \ldots \mathbf{d}_n\}$ can

be calculated directly. The new set of orthogonal disturbance vectors, $\{\mathbf{d}_1^*, \mathbf{d}_2^*, \ldots \mathbf{d}_n^*\}$, are used to generate the individual error vectors for the matrix $[\mathbf{e}_{d1^*} \quad \mathbf{e}_{d2^*} \quad \ldots \quad \mathbf{e}_{dn^*}]$. A singular value decomposition of this error matrix provides the minimum and maximum DIFs and the best and worst disturbance directions.

3.4.3 Process Example

This ideas can be demonstrated on the same batch polystyrene example presented earlier. There are now two disturbances perturbing the process: the impurities and the initiator efficiency vary from batch-to-batch. Assuming the individual direction vectors are unknown, a PCA of MWD data from 50 uncontrolled batches provides two orthogonal defining direction vectors ($\mathbf{d}_1^*$ and $\mathbf{d}_2^*$). Consider again the case in which the three potential manipulated variables are $[I]_0$, $[M]_0$ and T_{batch} and the batch-to-batch controllers (3.1) are used to control the weight average chain length. Figure 3.7 shows the error in the full MWD, with no control and control using each of the manipulated variables. Table 3.6 summarizes the minimum and maximum Disturbance Inflation Factors for each of the manipulated variables, given the weight-average chain length as the controlled variable.

Table 3.6. Best and worst disturbance directions with controlled variable p_w

MV	Best Disturbance Direction	Worst Disturbance Direction	DIF_{min}	DIF_{max}
$[M]_0$	$0.9978\mathbf{d}_1^*-0.0666\mathbf{d}_2^*$	$0.0666\mathbf{d}_1^*+0.9978\mathbf{d}_2^*$	0.03	1.50
$[I]_0$	$0.8478\mathbf{d}_1^*+0.5303\mathbf{d}_2^*$	$0.5303\mathbf{d}_1^*-0.8478\mathbf{d}_2^*$	0.68	1.01
T_{batch}	$0.9774\mathbf{d}_1^*-0.2115\mathbf{d}_2^*$	$0.2115\mathbf{d}_1^*+0.9774\mathbf{d}_2^*$	0.05	1.82

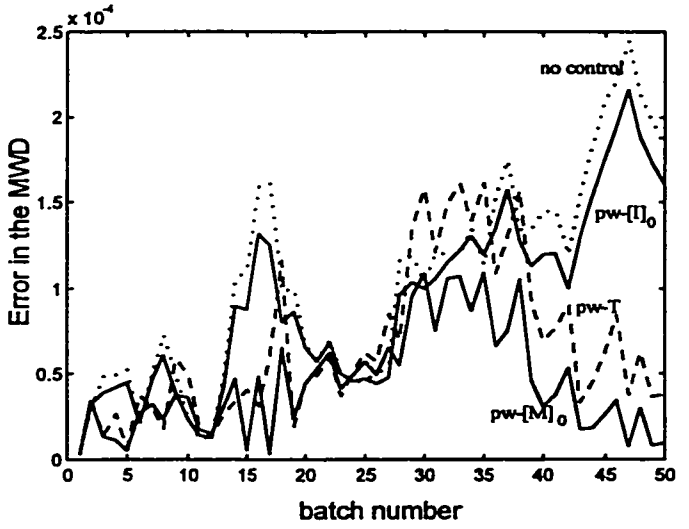

Figure 3.7. Error in the MWD with multiple disturbances

The information in Table 3.6 can also be represented visually. Figure 3.8 is a plot of the two dimensional disturbance plane (the plane defined by vectors $\mathbf{d}_1^*$ and $\mathbf{d}_2^*$). The MWD disturbance data (centered about the desired MWD) from the 50 uncontrolled batches are plotted in the plane $\{\mathbf{d}_1^*, \mathbf{d}_2^*\}$ with batch number indicated beside several points. Also shown on the plot are the best and worst directions for the initial monomer concentration (Table 3.6). It is clear that many of the disturbance realizations that occur in the 50 batches align with the direction for which $[M]_0$ can compensate well. Therefore, the DIF predicts

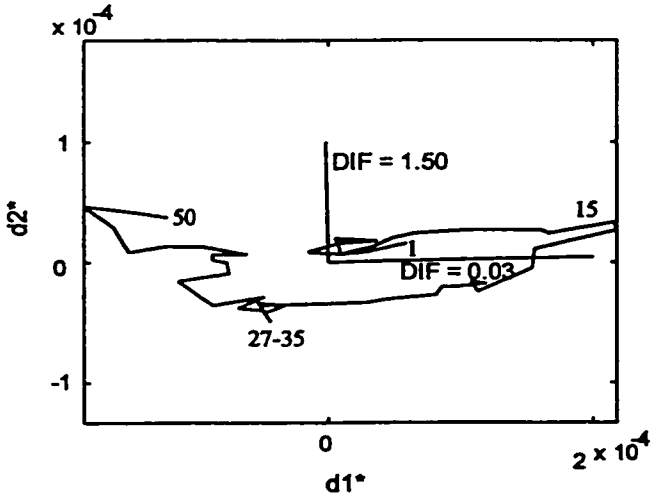

Figure 3.8. Best and worst disturbance directions for initial monomer concentration

that $[M]_0$ should perform well for many of the batches, particularly around batch 15. Weaker performance is expected around batches 27-35. These trends were evident in

Figure 3.7. Similar plots for initiator and temperature can be constructed.

The information Table 3.6, therefore, can be used to determine which manipulated variable is best. From this example, the $[M]_0$ appears most favorable: its best direction is closely aligned with the largest and most common disturbance realizations and its minimum DIF is very small. Of course, if it was important to never inflate the disturbance, $[I]_0$ would be the prudent choice of manipulated variable.

Table 3.7. Performance of latent variable controllers

MV	CV	Best Disturbance Direction	Worst Disturbance Direction	DIF_{min}	DIF_{max}
$[M]_0$	t_m	$0.9976\mathbf{d}_1^*-0.0685\mathbf{d}_2^*$	$0.0685\mathbf{d}_1^*+0.9976\mathbf{d}_2^*$	0.03	1.00
$[I]_0$	t_i	$0.8573\mathbf{d}_1^*+0.5148\mathbf{d}_2^*$	$0.5148\mathbf{d}_1^*-0.8573\mathbf{d}_2^*$	0.67	1.00
T_{batch}	t_t	$0.9827\mathbf{d}_1^*-0.1849\mathbf{d}_2^*$	$0.1849\mathbf{d}_1^*+0.9827\mathbf{d}_2^*$	0.04	1.00

Many other options exist for further improving MWD control. Consider again univariate latent variable control. Table 3.7 summarizes the DIFs of the latent variable controller for each manipulated variable.

Both monomer and temperature show promise for latent variable control. Figure 3.9 shows the error in the MWD when the latent variable t_m ($=\mathbf{m}_m^T(\mathbf{MWD}-\mathbf{MWD}_d)$) is controlled by the initial monomer concentration. Also shown for reference are the no control error and the error from the $[M]_0$ - p_w controller. For this particular system, latent variable control does provide an effective option for improving control over the full MWD while retaining the simplicity of a SISO controller. Of course the disadvantage is that the full MWD must be measured.

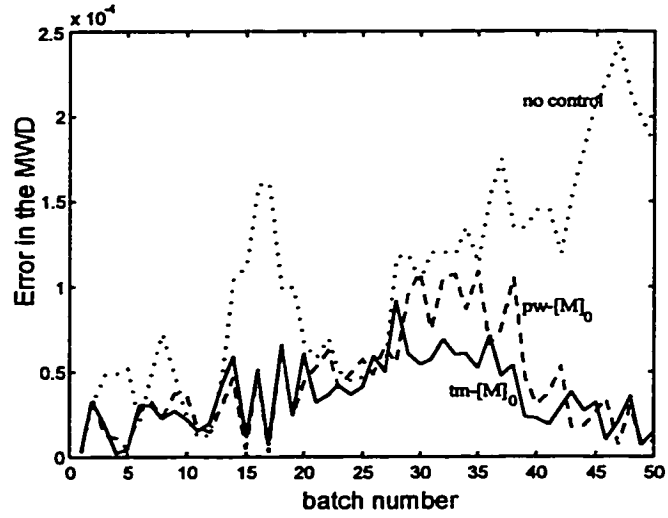

Figure 3.9. Latent variable control results

3.4.4 Many Controlled Variables

The calculation of the DIF can be easily extended to processes in which there are many controlled and manipulated variables (but still less than the total number of quality variables). Consider a system with 'k' controlled and manipulated variables. As with single variables, the controller will drive the controlled variables to their setpoints. That is, the manipulated variables will eliminate the disturbance in the space of the controlled variables. For a given defining disturbance vector, the change in the manipulated variables needed will satisfy (3.18):

$$\mathbf{d}^*_{i,proj} = \Delta m_1 \mathbf{m}_{1,proj} + \cdots + \Delta m_k \mathbf{m}_{k,proj} = \mathbf{M}_{proj} \Delta \mathbf{m} \tag{3.18}$$

where $\mathbf{d}^*_{i,proj}$ is the projection of vector $\mathbf{d}^*_i$ into the controlled variable space, and can be calculated using standard matrix projection algebra. Equation (3.18) can be solved for the multivariable change in the manipulated variables, $\Delta\mathbf{m}$, required to eliminate the disturbance. Then, the error vector equation introduced earlier becomes:

$$\mathbf{e}_{di} = -\mathbf{M}\Delta\mathbf{m} + \|\mathbf{d}^*_i\|\mathbf{d}^*_i \tag{3.19}$$

where $\mathbf{M} = [\mathbf{m}_1 \ \ldots \ \mathbf{m}_k]$. The error after control ($\mathbf{e}_{di}$) can be calculated for each defining disturbance vector $\mathbf{d}_i^*$, and the minimum and maximum DIFs, and the corresponding disturbance directions are then calculated as before.

Consider again the polystyrene process with two disturbances. A second possible control variable for this process is the number-average chain length. The Disturbance Inflation Factor can be calculated when both the number- and weight- average chain lengths are controlled, using two of the three manipulated variables. Table 3.8 summarizes the minimum and maximum DIFs.

Table 3.8. DIF for weight- and number-average control

MVs	DIF_{min}	DIF_{max}	Best Disturbance Direction	Worst Disturbance Direction
$[M]_0$ and $[I]_0$	0.05	0.85	$0.9977\mathbf{d}_1^*-0.0671\mathbf{d}_2^*$	$0.0671\mathbf{d}_1^*+0.9977\mathbf{d}_2^*$
$[M]_0$ and T_{batch}	0.03	0.52	$0.9927\mathbf{d}_1^*-0.1204\mathbf{d}_2^*$	$0.1204\mathbf{d}_1^*+0.9927\mathbf{d}_2^*$
$[I]_0$ and T_{batch}	0.02	0.83	$0.9878\mathbf{d}_1^*-0.1558\mathbf{d}_2^*$	$0.1558\mathbf{d}_1^*+0.9878\mathbf{d}_2^*$

The Disturbance Inflation Factors indicate that $[M]_0$ and T_{batch} are the best pair of variables to control p_n and p_w. However, examination of the three gain vectors show that T_{batch} lies very close to $[M]_0$, and therefore some conditioning problems may be anticipated. Therefore, $[M]_0$ and $[I]_0$ are selected as the two most promising manipulated variables for controlling p_n and p_w in this system. A multivariable deadbeat controller (consistent with earlier SISO controllers) is implemented:

$$\begin{bmatrix} [M]_0 \\ [I]_0 \end{bmatrix}_{batchk+1} = \begin{bmatrix} [M]_0 \\ [I]_0 \end{bmatrix}_{batchk} + \begin{bmatrix} 163 & -12900 \\ 270 & -50000 \end{bmatrix}^{-1} \begin{pmatrix} p_{n,sp} - p_n \\ p_{w,sp} - p_w \end{pmatrix} \tag{3.20}$$

Figure 3.10 shows the results of controlling p_w and p_n with $[M]_0$ and $[I]_0$. Also shown for reference are the no control error and the error from the p_w-$[M]_0$ controller. The numbers in Table 3.8 accurately represent the performance of the multivariable controller; as predicted, its performance is better than that of the SISO p_w-$[M]_0$ controller. It is comparable to that of the SISO latent variable controller (Figure 3.9). Although not shown, control of the individual averages is very good, and comparable to the results in Figure 3.1.

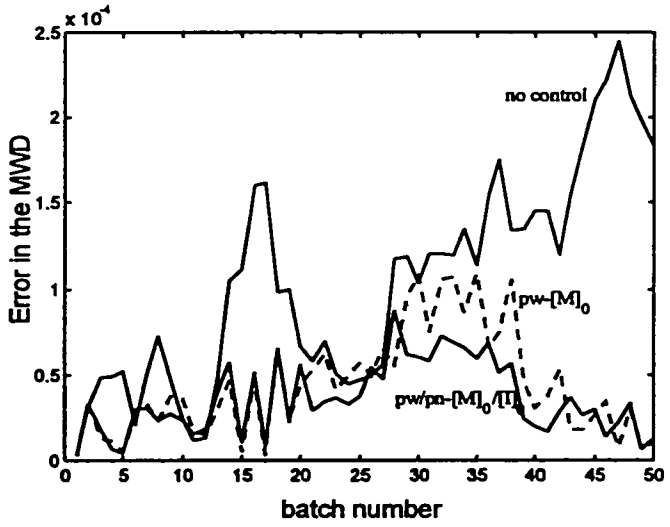

Figure 3.10: Multivariable control results

Therefore there are two feasible options for improving the full MWD using simple control structures. First, one could measure the full MWD and implement a SISO latent variable controller. Second, one could measure only the number and weight average chain lengths, and implement a 2-input, 2-output multivariable controller. If it is easier to measure the averages, the latter may be the preferred choice.

3.5 Impact of Scaling on DIF

One issue that plays a role in most processes is scaling. Scaling of quality variables is done for many reasons, for example if variables have widely different units or levels of importance. Scaling would be particularly useful for MWD control if one has some prior information about which part of the distribution predominantly influences polymer properties. Therefore, the impact of scaling on the calculation of the DIF must be addressed.

The original disturbance, gain vectors or the calculation of the error vector (**e**) do not change when scaling is a factor. Scaling is simply used to incorporate one's process knowledge. Instead of taking the ratio of the error vector norm to the original disturbance norm, it is more meaningful to use:

$$\mathrm{DIF} = \frac{\left\| \underline{\underline{\mathbf{S}}}\mathbf{e} \right\|}{\left\| \left(\underline{\underline{\mathbf{S}}} \|\mathbf{d}\| \mathbf{d} \right) \right\|} \tag{3.21}$$

where $\|\mathbf{d}\|\mathbf{d}$ is the disturbance vector in the unscaled space and **S** is the scaling matrix. The DIF now defines the ratio of the scaled error to the scaled disturbance, thus reflecting the relative importance of each of the quality variables before and after control. Equivalently, one could scale the output variables (plant data) first, obtain the gain and disturbance vectors in the scaled space and use the DIF expression presented previously.

3.6 Concluding Remarks

In this chapter, implicit control of the full MWD has been considered. It was seen that the choice of manipulated and controlled variables have an important impact on the controller performance when the full MWD is considered. Controlling a single average chain length can sometimes cause the MWD to degrade. A simple linear analysis tool, the Disturbance Inflation Factor, was introduced to predict this effect.

The DIF was applied to quickly and effectively predict the impact, on the full

MWD, of a large number of control options. It was seen that with prudent choices of manipulated and controlled variables, simple SISO control can provide significant improvement in the full MWD. Ratio and latent variable control were two options suggested. Ratio control (in which a ratio of two manipulated variables is used to control a molecular weight average) was seen to be a feasible option for the polystyrene system in which one disturbance perturbed the process. Latent variable control was successful for the process with two disturbances. Also shown was a multivariable controller in which both number- and weight- averages are controlled. This provided an effective and easily implemented controller for the implicit control of the full MWD.

It should be noted that while the focus of this chapter has been on the implicit control of the MWD, the ideas are applicable to any system with many quality variables. When a process has a lot of quality variables, typically only a small subset is chosen for control. Others may only be monitored, or not even measured at all. The DIF can be applied to select the best manipulated and/or controlled variables in this situation. It is directly applicable to continuous processes as well, since quite often long time delays in measurements mean that steady state considerations still dominate.

Finally, it was seen that only modest process information (disturbance and manipulated variable directions) is required to calculate the DIF. Once this information is obtained, a large number of control schemes can be evaluated quickly and effectively. Furthermore, excellent insight is gained by looking at process directions using the DIF methodology, and this may ultimately lead to further process improvements.

3.7 Nomenclature

The following nomenclature was used in this Chapter:

$\mathbf{x}$: vector

$\mathbf{X}$: matrix

$\|\mathbf{x}\|$: 2-norm of $\mathbf{x}$ ($\|\mathbf{x}\| = (\sum(x_i)^2)^{1/2}$)

C: controlled variable

d: disturbance

e: error

[I]: concentration of initiator

[M]: concentration of monomer

M: manipulated variable

MWD: molecular weight distribution

p: principal component

p_w: weight average chain length

r: chain length

S: Scaling matrix

T: temperature

u: manipulated variable

W(r): the mass fraction of polymers with chain length r

y: output variable

β_d: angle between controlled variable axis and disturbance vector

β_M: angle between controlled variable axis and manipulated variable vector

λ: element in the RGA matrix

Appendix 3.1. Review of Linear Algebra

Vector Norm. In this paper, the vector norm refers to the 2-norm. The norm of a vector $\mathbf{x}$ is:

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

Physically, this is the length of the vector.

Dot Product. The dot product of two vectors is a scalar:

$$\mathbf{x} \bullet \mathbf{y} = \mathbf{x}^T \mathbf{y} = \sum_i x_i y_i \tag{A.2}$$

The dot product can also be calculated by:

$$\mathbf{x} \bullet \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos(\theta) \tag{A.3}$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Vector Spaces. A vector space is usually expressed in terms of a set of basis vectors. For example, the vector space spanned by $\{\mathbf{x}_1, \mathbf{x}_2\}$ contains all vectors that are linear combinations of $\mathbf{x}_1$ and $\mathbf{x}_2$. Therefore, this vector space is a plane. Note that the set of basis vectors describing a vector space is not unique.

Projection Matrices. For a given space spanned by vectors $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$, the projection matrix is $\mathbf{V}(\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T$, where $\mathbf{V} = [\mathbf{v}_1\ \mathbf{v}_2 \ldots \mathbf{v}_n]$. To obtain the projection of any vector onto the space spanned by $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$, pre-multiply this vector by the projection matrix:

$$\mathbf{x}_{proj} = \mathbf{V}(\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\mathbf{x} \tag{A.4}$$

Similarly, one can also define a projection matrix for the orthogonal complement space (OC) of $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$. By definition, the OC space of $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$ contains all vectors that are perpendicular to $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$. To obtain the projection of any vector onto the orthogonal complement space of $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$, pre-multiply this vector by

$$\mathbf{x}_{proj,OC} = \left(\mathbf{I} - \mathbf{V}(\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\right)\mathbf{x} \tag{A.5}$$

Angle Between two Vectors. The definition of the dot product allows for the calculation of the angle between two given vectors:

$$\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos(\theta)$$
$$\theta = \cos^{-1}\left(\frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}\right) \tag{A.6}$$

The angle will be in radians, where one radian = $360/2\pi$ °.

Angle Between a Vector and a Space. The angle between a vector and a space refers to the angle between a vector and its projection onto the space. Given a vector $\mathbf{x}$ and a space spanned by $\{\mathbf{v}_1, \mathbf{v}_2 \ldots \mathbf{v}_n\}$, one can project $\mathbf{x}$ onto the vector space via the methods explained above. Then, calculate the angle between $\mathbf{x}$ and $\mathbf{x}_{proj}$, also as explained above.

Appendix 3.2. Relationship of DIF to Other Disturbance Analysis Tools

Zhao and Skogestad (1997) consider the problem of partial control of a continuous bioreactor. Partial control is defined as controlling only a subset of the output variables, while allowing the others to float. They use the partial disturbance gain to evaluate the feasibility of the partial control for a 2-input, 2-output system. The partial disturbance gain is described below and compared to the disturbance inflation factor.

Zhao and Skogestad (1997) consider partitioning a system, where $\mathbf{y}_1$ contains the uncontrolled variables, $\mathbf{y}_2$ contains the controlled variables:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}=\begin{bmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix}+\begin{bmatrix}\mathbf{G}_{d1}\\ \mathbf{G}_{d2}\end{bmatrix}\mathbf{d} \tag{B.1}$$

For one disturbance, $\mathbf{G}_{d1}$ will be a column vector. The open loop disturbance effect on $\mathbf{y}_1$ will be $\mathbf{G}_{d1}$ (assuming a value of d=1). The closed loop disturbance effect on $\mathbf{y}_1$ (that is, $\mathbf{y}_2$ controlled perfectly with $\mathbf{u}_2$):

$$\mathbf{P}_{d1}=\mathbf{G}_{d1}-\mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{d2} \tag{B.2}$$

The $\mathbf{P}_{d1}$ (the partial disturbance gain) is equal to the original disturbance vector minus the effect from $\mathbf{u}_2$ controlling $\mathbf{y}_2$. $\mathbf{P}_{d1}$ is a vector (for one disturbance) and is the error in the $\mathbf{y}_1$ space after $\mathbf{y}_2$ has been perfectly controlled. It is relative to the assumed value of 1 for d. Clearly, the concept is similar to the DIF, despite the different form.

For the evaluation of partial control, $\|\mathbf{P}_{d1}\| < \|\mathbf{G}_{d1}\|$ is desired. This differs slightly from how the DIF evaluates the disturbance. Zhao and Skogestad (1997) compare the post-control error to the original error in $\mathbf{y}_1$ due to the disturbance. The DIF compares the

post-control error to the original error in the entire quality space. The reason for the different comparisons is due to the different reasons for using the tool.

The two main (although minor) distinctions between the DIF and the partial disturbance gain are:

1. The DIF directly handles the case in which the controlled variable is a linear combination of the output variable;
2. Because of the assumed value of d=1 for the partial disturbance gain, it implicitly considers disturbance magnitude as well as direction;

Since Zhao and Skogestad (1997) work with the original variables, there is no need to project to the controlled space. This is one difference between the DIF and the partial disturbance gain. The DIF was derived to handle cases in which the controlled variable is a linear combination of the output variables (e.g. an average). This is particularly useful for proposing and analyzing latent variable controllers. The DIF simplifies to the P_d expression if a subset of the output variables is considered.

Second, when calculating the PDG values, the results reflect both disturbance magnitude and direction. The DIF calculation considers only direction; magnitude is considered in the interpretation (see the section on Multiple disturbances). The DIF therefore clearly distinguishes between large disturbances and disturbances that occur in bad directions. This point can be illustrated by comparing the results of the DIF and PDG for the bioreactor process. For the given matrices (section 4.2, pg 701) and the first MV-CV pairing considered, Table B.1 summarizes the DIF and the PDG.

The DIF results are expected due to the disturbance - manipulated variable alignments. The manipulated variable and the three disturbances (denoted with an asterisk) are well aligned. The two other disturbances (denoted with a double asterisk) are poorly aligned with the manipulated variable gain vector. The disturbances Yxs and S_{fd} perturb the process in exactly the same direction, therefore, their DIFs are equal. However, Y_{xs} has an openloop effect that is 3.38 times larger than that of S_{fd}, therefore, the PDG for the Y_{xs} disturbance is 3.38 times larger. The PDG inflates the effect of the

Y_{XS} disturbance because of its original size in the output space. The DIF considers both disturbances equal because it only evaluates direction.

Table B.1. Comparison of the DIF and the PDG

Disturbance	DIF	PDG (table 2, page 702)
μ_m *	0	0
K_s *	0.0311	0
Y_{XS} **	9.375	23.75
D_d *	0.0414	0
S_{fd} **	9.375	7

Another well known disturbance analysis tool is the Relative Disturbance Gain (Stanley et al. 1985). The RDG considers the impact of a disturbance's direction on control loop effectiveness. For a given disturbance, it is defined for controlled variable 'i' paired to manipulated variable 'j' as:

$$RDG_{i,j} = \frac{\Delta M_j \text{ needed to make } y_i = y_{i,sp} \text{ (all other control loops closed)}}{\Delta M_j \text{ needed to make } y_i = y_{i,sp} \text{ (all other control loops open)}} \qquad \text{(B.3)}$$

The relative disturbance gain is typically calculated for a given disturbance and control structure. For example, consider the 2x2 distillation column example in Stanley et al (1985). For the multiloop control scheme x_D-L (loop 1) and x_B-V (loop 2) and a feed composition disturbance, the following two RDG values are calculated:

$$RDG_1 = 0.17 = \frac{\Delta L \text{ needed to make } x_d = x_{d,sp} \text{ (with } x_b \text{ perfectly controlled by V)}}{\Delta L \text{ needed to make } x_d = x_{d,sp} \text{ (no control over } x_b)}$$

$$RDG_2 = 0.82 = \frac{\Delta V \text{ needed to make } x_b = x_{b,sp} \text{ (with } x_d \text{ perfectly controlled by L)}}{\Delta V \text{ needed to make } x_b = x_{b,sp} \text{ (no control over } x_d)}$$

(B.4)

The RDG can be used to assess the impact of adding one or more new control loops. From RDG_1, it is clear that controlling x_b has a favorable impact on the control of x_d, in terms of manipulated variable actions needed. In this way, the RDG could be used to assess how many loops are needed.

In order to directly compare the RDG to the DIF, it is illustrative to rearrange it:

$$RDG_{i,j} = \frac{\Delta M_j \text{ needed to make } y_i = y_{i,sp} \text{ (all other control loops closed)}}{\Delta M_j \text{ needed to make } y_i = y_{i,sp} \text{ (all other control loops open)}}$$

$$= \frac{\left(1/k_{i,j\ \text{closed loop}}\right) e_{y_i} \text{ (all other control loops closed)}}{\left(1/k_{i,j\ \text{open loop}}\right) e_{y_i} \text{ (all other control loops open)}} \qquad \text{(B.5)}$$

$$= \lambda_{i,j} \frac{e_{y_i} \text{ (all other control loops closed)}}{e_{y_i} \text{ (all other control loops open)}}$$

The RDG appears similar in concept to the DIF since both contain ratios of errors. However, the RDG focuses on the error in one output variable, rather than a norm of the error in the full quality space. In its original form, the RDG is most useful for assessing control loop interaction (if one or more extra control loops are advantageous to the existing loop) or for determining if decoupling or full multivariable control is advised. This is not the problem addressed by the DIF.

Appendix 3.3. Discussion of Disturbance Direction

In this appendix, two different approaches for establishing the disturbance direction(s) will be discussed. One approach relies on plant data, the other uses a model to establish the disturbance space. It should be noted that regardless of the approach, in order to calculate the DIF, the effects of the disturbances and manipulated variables on the entire quality space must be known. This implies that all the quality variables must be measured for a period of time (if using data) or predicted from a model (if using a model).

Disturbance Directions from Plant Data

Disturbance direction is a key concept in calculating the DIF. In many processes, the exact disturbance is unknown; it is only observed in the quality variables. There may be one, two or several disturbances perturbing the process at any given time. For these reasons, disturbance 'direction', rather than disturbance 'gain', is a more appropriate term.

In this chapter, processes are classified as having either one or more than one disturbance direction. Principal Components Analysis is used to establish the dominant disturbance directions. In this appendix, a brief overview of the use of PCA for establishing the disturbance directions will be presented. The details of the method can be found elsewhere (Wold et al (1987)).

To begin, it is assumed that in the historical database of the process there exists quality data from times in which no moves in the potential manipulated variables are made (or, data that has had the effect of known manipulated variable moves removed). That is, data contains only information on how the disturbance(s) are perturbing the plant. A PCA is performed on the data centered about the desired values, and a significant number of components are determined using the appropriate cross-validation techniques. For some processes, one principal component will explain most of the variation.

Conceptually, the data in the 'n' dimension quality space falls in an almost perfect line, and the idea of a disturbance 'gain vector' makes sense. If two or more components are significant, the disturbance(s) perturb the process in a plane or higher dimensional space. Then, the principal components form a vector basis describing the disturbance space. A given realization of the disturbance(s) can be any linear combination of the principal components.

Disturbance Directions from a Model

A model can also be used to establish the space in which the disturbance moves the process. If you have some idea of the disturbance(s), these can be simulated and their effect on the entire quality recorded. That is, several different levels for each the disturbance can be systematically simulated, and the disturbance effect observed. There are two options for simulation:

i) simulate the disturbances one at a time

ii) simulate the disturbances simultaneously

Simulating the disturbances one at a time will provide some idea of the individual gain vectors for each. These gain vectors become the defining vectors for the disturbance space. For example, if there are two disturbances, the two gain vectors observed by simulating different disturbances will define a plane in which the process is moved disturbance (if there is only one disturbance, its gain vector is the disturbance space).

If there is some a priori knowledge about the relative magnitudes of the disturbances, simulating them simultaneously is advantageous. By simulating them together, some information about their additive effect is obtained. The data will clearly show what direction is dominant, that is the direction in which the two or more disturbances combine to create a large observed disturbance.

4. Reduced Dimension Control of Dynamic Systems

4.1 Introduction

For many processes, quality of the end product is of paramount importance, therefore the focus of control is product quality. Several common process characteristics often exist in product quality control situations, regardless of the specific industry. First, the quality variables often outnumber the variables available for manipulation. This is particularly true when quality measurements are inexpensive to obtain (such as in paper making processes), or the product quality is characterized by a distribution (such as in crystalization or polymerization processes). Second, the variables that characterize quality are nearly always highly correlated. Similar trends are often observed in many of the measured quality variables. Furthermore, the manipulated variables, regardless of their number, often do not have the ability to move the process in an equal number of independent directions, thus quality control is often characterized by ill-conditioning. And finally, there are usually a limited number of disturbances that perturb the process; even for those processes with many unknown disturbances, it is often observed that they move the process in a much smaller dimension than the number of disturbances.

These characteristics are related to the dimension of the process, and the implications for control are best illustrated through several examples. Consider first the batch-to batch control of the MWD in a semi-batch polymerization reactor, as was discussed in the previous chapter. The polymer quality is characterized by the full MWD, which consists of 100 measured points on the distribution. The MWD is measured at the end of each batch. One can adjust three variables: the initial monomer charge, the initial initiator charge or the constant temperature at which the batch is

operated. For each batch, the monomer is fed in a pre-specified feedrate profile that has been optimized for the given product (Chapter 2).

The main disturbance affecting the MWD is the impurities, which enter into the batch with the monomer feed. The impurities level drifts in a random walk from batch-to-batch, and causes variations in the polymer quality from batch-to-batch. Ideally, one would like to implement batch-to-batch corrections to minimize variability and produce a more consistent product. A typical batch-to-batch control scheme involves measuring the MWD at the end of the batch and based on some measure of MWD error, calculating the change(s) in the adjustable variables for the next batch using a feedback algorithm.

A dimensional analysis of the process highlights the following issues. There are 100 output measurements, therefore the measured output space has a dimension of 100. However, plant tests (consisting of independent random perturbations on the three available inputs) indicate that these variables can effectively move the process in only two dimensions. The disturbance perturbs the process in a single dimension. The combined variation of the disturbances and manipulated variables falls primarily in three dimensions. Therefore, the dimension of the measured outputs is much larger than the dimension of the output space affected by the manipulated variables and the disturbances.

Therefore, the true process dimension is much smaller than 100. Clearly, not all 100 dimensions need to be controlled, since all the observed process variation falls in only three dimensions. Three potential solutions exist:

1. Attempt to directly control the entire output space (the full MWD);
2. Control a small number of more easily measured, representative variables (such as the average molecular weight);
3. Control a subspace of the output variables that is representative of the observed variation in the MWD (e.g. control linear combinations of the MWD variables that are somehow representative of the observed process variation).

Therefore, the choice exists to directly control the full output space (option one) or control a small subspace of the output space (option two or three). There are several motivating reasons for considering the latter two options. Certainly, it would be preferable if only some easily measured averages of the MWD were controlled. In fact, in most cases, good product quality (good control over the full MWD) is achieved by controlling only the average of the distribution. This is a direct result of the fact that disturbances and manipulated variables can only affect the MWD in limited directions.

Alternatively, one could still choose to directly control the full MWD (i.e control 100 output directions) with all three available manipulated variables. However, there could be robustness problems due to the ill-conditioning. This particular problem has been observed in both an ill-conditioned paper making system (Arkun et al. (1998)) and a blown-film process (Featherstone and Braatz (1997)) and will be discussed briefly here. In the case of the blown-film process, film thickness is measured in 45 different locations, and there are 45 actuators for adjusting the local film thickness. However, plant tests indicated that the process variability was observed in only ten dimensions. This is typical of a process in which extra measurements (beyond the true process dimension) are being made, simply because of the availability of sensors and computers.

Featherstone and Braatz (1997) proposed a reduced dimension controller, in which only ten output directions are controlled and manipulations are made in only ten input directions. Thus, ten output directions which somehow summarize the underlying dimension of the process are controlled (this was listed as the third option above). When compared to the case in which all forty five outputs were directly controlled (i.e. a full dimension controller), the ten dimension controller led to a lower overall thickness variability. This was attributed to the lack of controller robustness to process/model mismatch. The full dimension controller attempts to manipulate the input variables in directions that are not well known (due to the ill conditioning) and

the result is an increased output variability. By restricting the controller's dimension, and allowing it to move the process only in strong, well identified directions, the overall output variability is ultimately reduced (Arkun et al (1998), Featherstone and Braatz (1997)).

Situations such as those described above lead to the idea of 'reduced dimension control' (RDC). In the examples discussed, the manipulated variables and the disturbances move the process in a much smaller dimension than the measured output space. In RDC, a distinction is made between the size of the measured output space, and the size of the true underlying process dimension (the subspace within the measured output space in which the process varies). With the increasing computer capability for data acquisition, dimension issues such as these are becoming more frequent. The reduced dimension control approach is first, to evaluate if the process operates in a space that has a smaller dimension than the measured output space; second, to acknowledge that direct control of the full measured output space is not necessary in such cases; and finally to consider control only in a reduced space.

The practical motivation for controlling a reduced dimension includes ease of measurement (i.e not all outputs need to be measured online, if certain quality measurements are more difficult to obtain), controller simplicity, and controller robustness for ill-conditioned processes.

When considering control in a reduced space, the RDC approach leads to the following questions:

1. What subset of the original variables should one manipulate and control?
2. Can one select new, representative linear combinations of the original manipulated and output variables (called *latent variables*) that one can manipulate and control to provide better control in the entire quality space?

The choice of the controlled subsystem has a significant impact on the quality (as characterized by the full quality space). It was observed in the last chapter that if a poor subsystem for control is chosen, the overall effect can be to degrade product

quality, rather than improve it. Several solutions have been proposed in the literature for choosing subsets of the original variables or for choosing linear combinations of the variables for control. A linear analysis tool called the Disturbance Inflation Factor was introduced in the last chapter for evaluating the effect of a selected controlled subsystem (controlled and manipulated variables) on the full quality space. A similar tool, the Partial Disturbance Gain, was applied to a bioreactor for selecting appropriate control loops (Zhao and Skogestad (1997)). Other tools for selecting a subset of the original output variables include the NonSquare RGA (Chang and Yu (1990)). Reduced dimension control of paper making machines has been addressed by several research groups (Arkun et al (1998), Featherstone and Braatz (1997)). Singular Value Decomposition has also been applied for reducing the dimension of the controlled space (Featherstone and Braatz (1997), Lau et al. (1986)). Finally, data-based methods of Principal Components Analysis (Roffel et al. (1989), Eek and Bosgra (1995)) and Canonical Correlations Analysis (MacGregor and Wong (1980)) for selecting controlled and manipulated variables have also been demonstrated and applied in industry. A more thorough discussion of these methods (and the relationships among them) is found in Chapter 5.

In this chapter, a mathematical framework for reducing the dimension of the controlled subsystem is presented, based on minimum variance control theory. Two expressions are derived:

1. the linear combinations of the inputs and outputs that should be controlled to minimize the error in the entire output space (that is, an expression for the optimal controlled subsystem will be derived);
2. the linear combinations of the inputs that should be manipulated to minimize the error in the overall output space if the controlled variables are set (for example, if some outputs are more easily measured);

The framework shows explicitly the role of the disturbances in determining the optimal reduced dimension control system. This has not been explicitly discussed before in the literature.

Based on this framework, a process dimensionality analysis for a simulated dynamic Kamyr digester is demonstrated and two reduced dimension controllers are proposed. The performance of the reduced dimension controllers is compared to that of a full dimension DMC. In comparing the controllers, the emphasis is on the performance of the reduced dimension controllers when the assumptions on which they are based are violated.

4.2 A Framework for Dimension Reduction

In this section, a framework for the optimal reduction of the dimension of a control system is presented. Consider the process represented by:

$$\mathbf{y}_t = \mathbf{G}(B)\mathbf{u}_{t-f-1} + \mathbf{y}_{d_t} \tag{4.1}$$

where $\mathbf{y}_t$ and $\mathbf{u}_t$ are (nx1) and (rx1) vectors of outputs and inputs respectively.

Prior to developing the framework for dimension reduction, a disturbance model (an expression for $\mathbf{y}_{dt}$ in (4.1)) is needed. For the case of k disturbance sources, the overall disturbance can be represented by:

$$\mathbf{y}_{d_t} = \mathbf{p}_{d1}d_{1,t} + \mathbf{p}_{d2}d_{2,t} + \ldots + \mathbf{p}_{dk}d_{k,t} \tag{4.2}$$

Equation (4.2) indicates that the overall observed disturbance at time 't' is a sum of all the individual disturbances, each characterized by their own direction vector, $\mathbf{p}_{di}$, in the output space. Disturbance dynamics can be represented by a time series model for each of the individual disturbances, $d_{i,t}$:

$$\mathbf{y}_{d_t} = \mathbf{p}_{d1}\left(\frac{\theta_1(B)}{\phi_1(B)}\right)a_{1,t} + \mathbf{p}_{d2}\left(\frac{\theta_2(B)}{\phi_2(B)}\right)a_{2,t} + \ldots + \mathbf{p}_{dk}\left(\frac{\theta_n(B)}{\phi_n(B)}\right)a_{k,t} \qquad (4.3)$$

where B is the backwards shift operator (Box and Jenkins (1976)). Substituting (4.3) into (4.1) gives:

$$\mathbf{y}_t = \mathbf{G}(B)\mathbf{u}_{t-f-1} + \mathbf{p}_{d1}\left(\frac{\theta_1(B)}{\phi_1(B)}\right)a_{1,t} + \mathbf{p}_{d2}\left(\frac{\theta_2(B)}{\phi_2(B)}\right)a_{2,t} + \ldots + \mathbf{p}_{dk}\left(\frac{\theta_n(B)}{\phi_n(B)}\right)a_{k,t} \qquad (4.4)$$

In practice, it is unlikely that all disturbances are known; only the total effect of the disturbances is observed. In this case, one can formulate the disturbance model (equation (4.3)) using principal components analysis of data that contains only the effects of the disturbances. Such data, referred to as disturbance data, is either open loop data or data with the effect of known manipulated variable moves removed. In order to formulate the disturbance model in the form of (4.3), the disturbance data is first decomposed using Principal Components Analysis (Wold et al (1987)):

$$\mathbf{Y}_d = \mathbf{T}\mathbf{P}_d^T \qquad (4.5)$$

where the disturbance data is contained in the matrix $\mathbf{Y}_d$ (each column of $\mathbf{Y}_d$ contains sequential observations of a single output, y_i). The k columns of $\mathbf{P}_d$ are orthogonal vectors that form a basis for the k dimensional disturbance space. Each column of the matrix T ($\mathbf{t}_i$) contains sequential observations of new latent variables, where $\mathbf{t}_i = \mathbf{Y}_d\mathbf{p}_{di}$. Disturbance dynamics are modeled by fitting a time series model to each latent variable ($\mathbf{t}_i$) individually, resulting in an expression for the disturbance in the form of (4.3). More details on using PCA to model disturbances may be found in Rigopoulos et al (1997).

For the derivation of the framework, the following simplifications are made:

1. the disturbance is assumed to be represented by a single direction, however the results are later extended to multiple disturbance directions.
2. the deadtime between all inputs and outputs is the same and equal to f periods of delay.

4.2.1 Control using an optimal linear combination of inputs and outputs

In this case we ask what linear combination of inputs and outputs should be manipulated and controlled in order to minimize the overall error in the full output space. Consider the following situation: there are many outputs (n), several inputs (r, r<n) and one disturbance (k=1). The objective function is to minimize the error in the full output space (n dimensional). Mathematically, this corresponds to optimizing the following objective function:

$$\begin{aligned} &\min \; E(\mathbf{e}_{t+f+1}^{T}\mathbf{e}_{t+f+1}) \\ &\text{where } \mathbf{e}_t = \mathbf{y}_t - \mathbf{y}_{sp,t} \end{aligned} \tag{4.6}$$

given the model for the system:

$$\mathbf{y}_{t+f+1} = \mathbf{p}_d\left(\frac{\theta(B)}{\phi(B)}\right)\mathbf{a}_{t+f+1} + \mathbf{G}(B)\mathbf{u}_t \tag{4.7}$$

Assuming $\mathbf{y}_{sp} = \mathbf{0}$, an expression for $\mathbf{e}_{t+f+1}^{T}\mathbf{e}_{t+f+1}$ can be obtained from (4.7) as:

$$\mathbf{y}_{t+f+1}^{T}\mathbf{y}_{t+f+1} = \mathbf{p}_d{}^{T}\mathbf{p}_d\left(\frac{\theta(B)}{\phi(B)}\mathbf{a}_{t+f+1}\right)^2 + 2\mathbf{p}_d{}^{T}\mathbf{G}(B)\left(\frac{\theta(B)}{\phi(B)}\mathbf{a}_{t+f+1}\right)\mathbf{u}_t + (\mathbf{G}(B)\mathbf{u}_t)^{T}(\mathbf{G}(B)\mathbf{u}_t) \tag{4.8}$$

The disturbance model can be expanded into a forecast (terms involving current and past innovations) and a forecast error (terms involving future innovations):

$$\frac{\theta(B)}{\phi(B)}a_{t+f+1} = \Psi(B)a_{t+f+1} + \frac{T(B)}{\phi(B)}a_t \tag{4.9}$$

Substituting (4.9) into (4.8), taking expectations on both sides of equation (4.8) and noting that the autocorrelations ($E(a_{t+j}a_t) = 0$ and $E(a_{t+j}u_t) = 0, j \geq 1$) can be eliminated:

$$E(\mathbf{y}_{t+f+1}^T\mathbf{y}_{t+f+1}) = E\left(\begin{array}{c}\mathbf{p}_d^T\mathbf{p}_d\left(\frac{T(B)}{\phi(B)}a_t\right)^2 + \mathbf{p}_d^T\mathbf{p}_d\left(\Psi(B)a_{t+f+1}\right)^2 + \\ 2\mathbf{p}_d^T\mathbf{G}(B)\left(\frac{T(B)}{\phi(B)}a_t\right)\mathbf{u}_t + (\mathbf{G}(B)\mathbf{u}_t)^T(\mathbf{G}(B)\mathbf{u}_t)\end{array}\right) \tag{4.10}$$

Taking the derivative of the expectation and setting to zero:

$$\frac{\partial E(\mathbf{y}_{t+f+1}^T\mathbf{y}_{t+f+1})}{\partial \mathbf{u}_t} = 0 = 2\,\mathbf{G}(B)^T\mathbf{p}_d\left(\frac{T(B)}{\phi(B)}a_t\right) + 2\,\mathbf{G}(B)^T\mathbf{G}(B)\mathbf{u}_t \tag{4.11}$$

Solving for the input vector:

$$\mathbf{u}_t = -(\mathbf{G}(B)^T\mathbf{G}(B))^{-1}\mathbf{G}(B)^T\mathbf{p}_d\left(\frac{T(B)}{\phi(B)}a_t\right) \tag{4.12}$$

Substituting (4.12) into the expression for $\mathbf{y}_{t+f+1}$, equation (4.7), expanding the disturbance term and rearranging:

$$\mathbf{y}_{t+f+1} = \mathbf{p}_d\Psi(B)a_{t+f+1} + \left[\mathbf{I}_{mxn} - \mathbf{G}(B)\left(\mathbf{G}(B)^T\mathbf{G}(B)\right)^{-1}\mathbf{G}(B)^T\right]\mathbf{p}_d\left(\frac{T(B)}{\phi(B)}a_t\right) \tag{4.13}$$

The second term is a projection matrix; it projects the disturbance direction vector onto the space that is perpendicular to the space spanned by the process gain matrix (please refer to Appendix 3.1 for a review of projection matrices).

Therefore, variability in the outputs comes from two sources: the portion of the disturbance that cannot be corrected for due to process deadtime (first term) and the portion of the disturbance that is perpendicular to the effect of the process inputs (the second term). This analysis implies that the output error can be minimized if control is concentrated only on the space:

$$\mathbf{G(B)}\left(\mathbf{G(B)}^{\mathrm{T}}\mathbf{G(B)}\right)^{-1}\mathbf{G(B)}^{\mathrm{T}}\mathbf{p}_d \tag{4.14}$$

For the general case, this is a polynomial vector. Further insight is gained by considering the following two cases:

1. Steady State Processes
2. Processes in which all dynamics and deadtimes are the same.

Considering the latter, the gain matrix can be expressed as:

$$\mathbf{G(B)} = \frac{w(\mathrm{B})}{\delta(\mathrm{B})}\mathrm{B}^{\mathrm{f}}\mathbf{K} \tag{4.15}$$

Substituting (4.15) into (4.14) and canceling the dynamic terms gives:

$$\mathbf{C}_{\text{optimal}} = \mathbf{K}\left(\mathbf{K}^{\mathrm{T}}\mathbf{K}\right)^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{p}_d \tag{4.16}$$

Equation (4.16) is the projection of the disturbance vector onto the space spanned by the columns of **K** and this is the optimal direction to control. That is, the maximum benefits of control (minimizing the variability in the full output space) are

theoretically possible by controlling the linear combination of the outputs given by $y_{cv} = C_{optimal}^T y$.

From (4.12) and assuming equal deadtimes and dynamics, it can be easily seen that the inputs should be adjusted in the direction given by:

$$M_{optimal} = (K^T K)^{-1} K^T p_d \tag{4.17}$$

Geometrically, $M_{optimal}$ (4.17) defines a direction, in the input coordinates, in which the inputs should always be manipulated:

$$u = M_{optimal} \Delta \tag{4.18}$$

where Δ is a scalar for one disturbance. For a single disturbance, $M_{optimal}$ is a line in the input space which defines the ratio in which the manipulated variables should always be adjusted. A new manipulated variable for reduced dimension control can be defined by the following linear combination of the original inputs:

$$u_{mv} = (M_{optimal}^T M_{optimal})^{-1} M_{optimal}^T u \tag{4.19}$$

The number of independent controlled and manipulated variables required (i.e. the dimensions of y_{cv} and u_{mv} respectively) will be given by min(rank(K),rank(P_d)). In many cases, particularly in the case of regulation only, the number of potential manipulated variables often outnumber the independent disturbances, therefore one should choose to control as many directions as there are independent disturbance directions. If there are k independent disturbance directions defined by the columns of P_d, then the k-dimensional controlled variable space is defined by $C_{optimal} = K(K^T K)^{-1} K^T P_d$. The inputs should be adjusted in the k-dimensional space defined by $M_{optimal}$ =

$(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{P}_d$. A more thorough discussion may be found in Appendix 4.1 at the back of this Chapter.

In order to illustrate the concepts, consider a simple 3-output, 2-input process. Assume that the outputs are measured infrequently (relative to the process dynamics) and the process can be modeled by:

$$\mathbf{y}_t = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \mathbf{u}_{t-1} + \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \frac{1}{1-0.9B} a_t \tag{4.20}$$

The single disturbance direction is given by $\mathbf{p}_d = [-1;1;1]$. This disturbance direction is 31° from the two dimensional manipulated variable space (characterized by the columns of the gain matrix). According to the optimal framework, the direction that should be controlled is:

$$\mathbf{C}_{optimal} = \mathbf{K}(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{p}_d = \begin{bmatrix} -0.18 \\ 0.72 \\ 1.27 \end{bmatrix} \tag{4.21}$$

Practically, this means all three outputs should be measured at every control interval, and the value of the controlled variable, $y_{cv,t}$, calculated:

$$y_{cv,t} = -0.18y_{1,t} + 0.72y_{2,t} + 1.27y_{3,t} \tag{4.22}$$

Equation (4.22) is the expression for the controlled variable of the reduced dimension controller (which will be SISO for this example since there is only one disturbance). The manipulated variable, u_{mv}, for the reduced dimension controller is defined by the following linear combination of the original inputs:

$$\begin{aligned} u_{mv} &= (\mathbf{M}_{optimal}{}^{T}\mathbf{M}_{optimal})^{-1}\mathbf{M}_{optimal}{}^{T}\mathbf{u} \\ &= -0.1475u_1 + 0.8931u_2 \end{aligned} \tag{4.23}$$

To maintain y_{cv} at its setpoint, the process inputs should be manipulated in the direction:

$$\mathbf{M}_{optimal} = (\mathbf{K}^{T}\mathbf{K})^{-1}\mathbf{K}^{T}\mathbf{p}_d = \begin{bmatrix} -0.18 \\ 1.09 \end{bmatrix} \tag{4.24}$$

The relationship between the RDC manipulated variable ($u_{mv,t}$) and the physical inputs to the process ($\mathbf{u}_t$) is given by:

$$\mathbf{u}_t = \mathbf{M}_{optimal}u_{mv,t} \tag{4.25}$$

A 1x1 reduced dimension controller would be implemented as follows. A transfer function model can be identified between u_{mv} and y_{cv} using standard plant tests, for example a step test or PRBS on the input u_{mv}. Assume, for illustrative purposes, a step change of magnitude Δ=2 is made in the input u_{mv}. This corresponds to a change in the physical inputs to the plant $u_{1,t}$ = -0.36 and $u_{2,t}$ = 2.18. The step change response of y_{cv} is modeled by a transfer function (providing the empirical relationship between u_{mv} and y_{cv}). A standard SISO feedback controller can then be implemented between y_{cv} and u_{mv}. The input to the controller is given by (4.22); the output of the controller is $u_{mv,t}$ (the required value of the new manipulated variable). To implement the manipulated variable $u_{mv,t}$ in the process, the corresponding values of the real process inputs are obtained using the relationship $\mathbf{u}_t = \mathbf{M}_{optimal}u_{mv,t}$. Figure 4.1 shows the block diagram of the reduced dimension controller for this process.

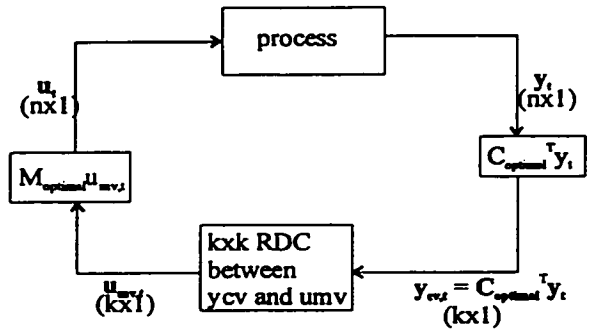

Figure 4.1. Reduced Dimension Controller based on optimal framework

4.2.2 Control of a fixed output by a linear combination of the inputs

In some cases the choice of quality variables to control is not an option. For example, perhaps the full MWD cannot be measured on a routine basis but the weight average molecular weight might be readily available from viscosity measurements. Here, we ask the question if the controlled variable is fixed, what linear combination of the inputs should be manipulated in order to minimize the overall error in the full quality space. Therefore, we have the following situation. There are many outputs (n), several inputs (r, r<n), one disturbance (k=1) and one controlled variable (defined by direction **c**, the controlled variable is $y_{cv} = \mathbf{c}^T\mathbf{y}$). The objective function is defined by:

$$\min_{\mathbf{m}}(\mathbf{e}_{t+1}^T\mathbf{e}_{t+1})$$
$$\text{s.t} \quad \mathbf{u}_t = \mathbf{m}u_{mv,t} \tag{4.26}$$

The single input direction is defined by the vector **m** and is unknown. The process model is:

$$\begin{aligned} \mathbf{y}_{t+1} &= \mathbf{p}_d\left(\frac{\theta(B)}{\phi(B)}\right)a_{t+1} + \mathbf{G}(B)\mathbf{u}_t \\ &= \mathbf{p}_d\left(\frac{\theta(B)}{\phi(B)}\right)a_{t+1} + \mathbf{G}(B)\ \mathbf{m}u_{mv,t} \end{aligned} \tag{4.27}$$

An expression for $u_{mv,t}$ can be derived from (4.27) as:

$$u_{mv,t} = -\frac{\mathbf{c}^T\mathbf{p}_d}{\mathbf{c}^T(\mathbf{Gm})}\left(\frac{T(B)}{\phi(B)}\right)a_t \tag{4.28}$$

Here, $u_{mv,t}$ is the output of the minimum variance controller for y_{cv} in the presence of the disturbances given in (4.27). Substituting (4.28) into (4.27):

$$\mathbf{y}_{t+1} = \mathbf{p}_d\left(\frac{\theta(B)}{\phi(B)}\right) - \mathbf{G}(B)\mathbf{m}\frac{\mathbf{c}^T\mathbf{p}_d}{\mathbf{c}^T(\mathbf{G}(B)\mathbf{m})}\left(\frac{T(B)}{\phi(B)}\right)a_t \tag{4.29}$$

From the above expression, one can derive an expression for $E(\mathbf{y}_{t+1}{}^T\mathbf{y}_{t+1})$, with $\mathbf{y}_{sp} = \mathbf{0}$ assumed. Assuming that the dynamic model is as given in (4.15), taking the derivative of $E(\mathbf{y}_{t+1}{}^T\mathbf{y}_{t+1})$ with respect to $\mathbf{m}$ and setting equal to 0, one gets:

$$0 = \mathbf{K}^T\mathbf{p}_d + \mathbf{K}^T\mathbf{K}\mathbf{m}\frac{\mathbf{c}^T\mathbf{p}_d}{\mathbf{c}^T(\mathbf{Km})} + \mathbf{p}_d{}^T(\mathbf{Km})\frac{\mathbf{K}^T\mathbf{c}}{\mathbf{c}^T(\mathbf{Km})} - (\mathbf{Km})^T(\mathbf{Km})\frac{\mathbf{c}^T\mathbf{p}_d}{(\mathbf{c}^T(\mathbf{Km}))^2}\mathbf{K}^T\mathbf{c} \tag{4.30}$$

Clearly this is not a simple expression for the direction $\mathbf{m}$. However, even when the controlled variable is fixed as $\mathbf{c}^T\mathbf{y}$, it was found, through simulated examples, that the true optimal direction ($\mathbf{m}$, from solving the above equation) is often very close to the direction given by (4.17), in which the controlled variable was free to be specified.

4.3 Process Example

In the previous section, several assumptions were made in order to derive the framework for reducing the dimension of the controller:

1. The process has no dynamics due to the infrequent sampling interval of the product quality, or, if dynamics are present, all input-output relationships follow the same dynamics;
2. The deadtimes between all inputs and outputs are the same;
3. The process is linear and non-time varying.

If the assumptions are not valid, equation (4.14) does not simplify to a constant direction. The implication is that a single linear combination (for one disturbance) is not optimal if the assumptions are not met. While the assumptions are approximately true for many product quality control situations with infrequent sampling, they are usually violated for other process control problems. However, the reduced dimension control vectors will still be valid at steady state, and the non-optimality will occur primarily during the transient periods. This suggests that a reduced dimension control (RDC) approach may still be reasonable even when these assumptions are violated.

To assess this, an RDC approach based on the framework derived above is applied to a simulated Kamyr digester whose dynamic and delay structure deviate greatly from the assumptions listed.

4.3.1 Process Description

A dynamic Kamyr digester simulation will be used as an example of a 'non-ideal' process, in order to illustrate the RDC approach and to assess the effect of the violation of the RDC assumptions. The dynamic simulation of a Kamyr digester has been put forth as a test problem for control studies (Kayihan (1998)). Details of the simulation have been published elsewhere (Kayihan (1997)) and so only the details

necessary for the current problem are described here. Figure 4.2 shows a schematic of the digester and the location of the inputs and outputs. The digester has five outputs that characterize the final pulp quality: three Kappa numbers (κ, effectively a measure of extent of reaction) and two densities (ρ):

$$y = \begin{bmatrix} \kappa_{mcc} \\ \kappa_{emcc} \\ \kappa_{cook} \\ \rho_l \\ \rho_u \end{bmatrix} \tag{4.31}$$

The Kappa number is measured in three different locations along the digester; the subscripts indicate the zone of measurement (cook zone (cook), modified continuous cooking zone (mcc) and extended modified continuous cooking zone (emcc)). The lower extract density (ρ_l) is measured from a liquor stream leaving the mcc zone while the upper extract density (ρ_u) is measured from a liquor stream leaving the cook zone. In practice, only the final Kappa number from the emcc zone (κ_{emcc}) and the two densities are easily measured; the two other Kappa numbers must be inferred (in the simulated process, however, it is assumed that all five outputs are measured online).

There are also five manipulated variables (three temperatures and two flowrates):

$$u = \begin{bmatrix} T_{cook} \\ T_{mcc} \\ T_{emcc} \\ F_{mcc} \\ F_{upextract} \end{bmatrix} \tag{4.32}$$

The temperatures (T_{cook}, T_{mcc} and T_{emcc}) refer to the adjustable temperatures of three different streams entering the digester. The streams enter the digester in different zones, indicated by the subscripts (T_{cook} is the temperature of the stream entering the digester cook zone). The flowrates of two streams can also be adjusted: the upper extract flow ($F_{upextract}$) and the fresh liquor into the mcc zone (F_{mcc}). Five independent disturbances (chip flowrate and moisture, white liquor density, chip lignin density and dilution flowrate) cause the five outputs to drift from their desired values.

The twenty-five process deadtimes are long and very different. They range anywhere from no deadtime to 300 minutes. The dynamics are also widely varying and long, with times to steady state ranging from 3 hours to 18 hours. The outputs are measured every 10 minutes. The digester is also very nonlinear: as the process moves from the desired output values, the gains and dynamics change.

Before doing a dimensionality analysis of the process, some discussion on scaling is needed since any numerical analysis is scaling dependent. In this case, the output variables have different engineering units. We chose to scale using the inverse of the open-loop standard deviations, so that the disturbance is seen as impacting all variables equally. The resulting scaling matrix is:

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1.4 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 4.9 & 0 \\ 0 & 0 & 0 & 0 & 8.7 \end{bmatrix} \tag{4.33}$$

where $y_s = Sy$. Another equally valid choice for scaling is to scale by the inverse of the steady state (or mean) values, which results in a matrix that is very close to the above scaling matrix. For the rest of the chapter and in the figures, it is assumed that all the outputs are scaled in this manner.

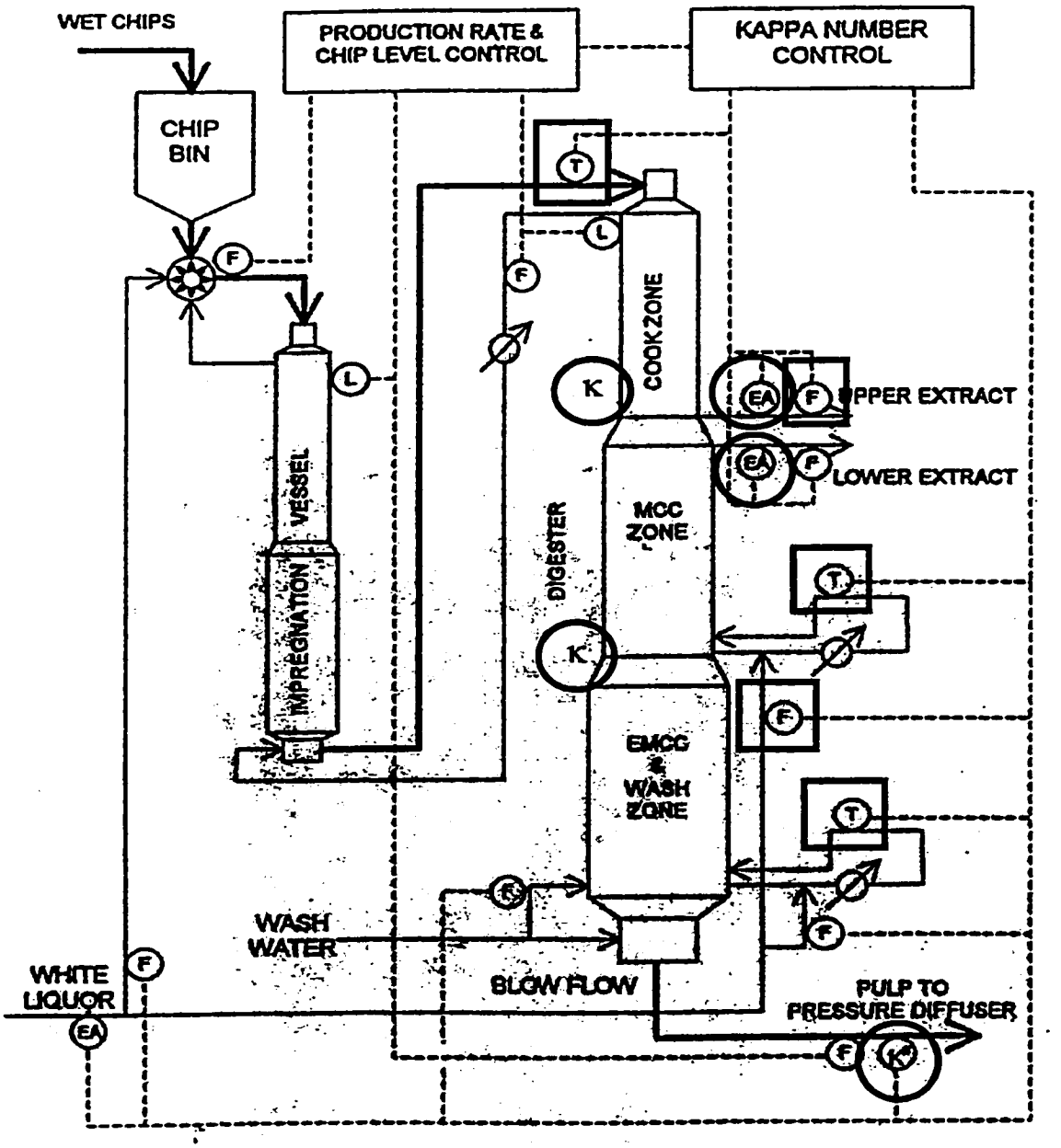

Figure 4.2. Schematic of the Pulp Digester

○ outputs □ inputs

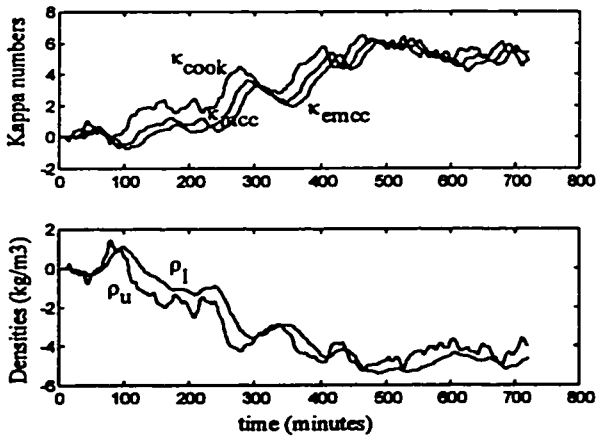

Figure 4.3 Open loop effect of the disturbances

A dimensionality analysis of the digester provides some interesting insight. Figure 4.3 shows the effect of the five disturbances on the five uncontrolled outputs. A PCA of this data indicates that the first principle component explains 98.8% of the variation in the data. The first latent variable is well modeled by an AR(1) time series model. Therefore the disturbance is essentially one dimensional and can be represented as:

$$\mathbf{y}_{d,t} = \begin{bmatrix} -0.42 \\ -0.38 \\ -0.51 \\ 0.42 \\ 0.48 \end{bmatrix} \left(\frac{1}{1-0.9985B} \right) a_t \tag{4.34}$$

Therefore, in theory, to minimize the variability in all five outputs, only one direction (not all five) needs to be controlled.

A dimensionality analysis can also be done on the manipulated variables. The gain matrix identified from step tests in each of the inputs is:

$$\mathbf{K} = \begin{bmatrix} -0.26 & -0.45 & 0.08 & -523 & -442 \\ -0.26 & -0.44 & -0.19 & -506 & -427 \\ -0.37 & 0 & 0 & 0 & -497 \\ 0.05 & -0.38 & -0.15 & 955 & 510 \\ -0.90 & 0 & 0 & 0 & 950 \end{bmatrix} \tag{4.35}$$

Clearly, the effect of the flowrates dominates the matrix because of the different units of the inputs. Therefore, it is more appropriate to scale the gain matrix, **K**, to also reflect the different engineering units of the inputs, $\mathbf{K}_s = \mathbf{K}\mathbf{S}_u$, where $\mathbf{S}_u$ is a diagonal matrix whose elements are the expected relative changes in the inputs:

$$\mathbf{S}_u = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{bmatrix} \tag{4.36}$$

$$\mathbf{K}_s = \begin{bmatrix} -0.78 & -1.34 & 0.40 & -0.52 & -0.44 \\ -0.79 & -1.33 & -0.93 & -0.50 & -0.43 \\ -1.1 & 0 & 0 & 0 & -0.50 \\ -0.15 & -1.16 & -0.73 & 0.96 & 0.51 \\ -2.7 & 0 & 0 & 0 & 0.95 \end{bmatrix} \tag{4.37}$$

A singular value decomposition of the matrix $\mathbf{K}_s$ gives the following singular values: 3.2, 2.2,1.6, 0.85, 0.58. Therefore, while the manipulated variables can move the process more easily in some directions, there are no directions in which they cannot move. That is, the process is reasonably well conditioned.

Finally, the alignment of the manipulated variable steady state gain vectors and the disturbance direction is of interest (Table 4.1). The alignment of the manipulated variable and disturbance directions give an indication of how much of the disturbance can be eliminated by that particular manipulated variable. If a manipulated variable is perfectly aligned with the disturbance, then it can (in theory) eliminate all the disturbance (beyond the process deadtime). Clearly the temperatures

are NOT well aligned with the disturbance, while the upper extract flowrate shows excellent alignment.

Table 4.1. Alignment of Disturbance Direction and MV Gain Vectors

Manipulated Variable	Angle to Disturbance Direction
T_{mcc}	75
T_{emcc}	96
T_{cook}	91
F_{mcc}	47
$F_{upextract}$	16

4.3.2 Latent Variable Control

Using the framework from the previous section, the output direction that should be controlled (based on steady state considerations) is:

$$\mathbf{C}_{optimal} = \mathbf{K}\left(\mathbf{K}^T\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{p}_d = \begin{bmatrix} -0.42 \\ -0.38 \\ -0.51 \\ 0.42 \\ 0.48 \end{bmatrix} \tag{4.38}$$

This is actually the disturbance vector since the five manipulated variables can move the process in all five output dimensions (thus the projection of $\mathbf{p}_d$ onto $\mathbf{K}$ is simply $\mathbf{p}_d$). The optimal input direction (i.e. the ratio in which the manipulated variables should be adjusted) is:

$$\mathbf{M}_{optimal} = \left(\mathbf{K}^T\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{p}_d = \begin{bmatrix} 0.3049 \\ -0.016 \\ -0.1116 \\ -0.00003 \\ 0.0008 \end{bmatrix} \tag{4.39}$$

Therefore, based on the framework from the previous section, when T_{cook} increases by 0.3049°, T_{mcc} should decrease by -0.016°, T_{emcc} should decrease by 0.1116°, and so forth.

To minimize the overall closed loop error, one could implement a SISO controller to control the direction $\mathbf{C}_{optimal}$ (i.e. control the linear combination $y_{cv} = \mathbf{C}_{optimal}{}^T\mathbf{y}$) by adjusting the new latent input, u_{mv} ($u_{mv} = (\mathbf{M}_{optimal}{}^T\mathbf{M}_{optimal})^{-1}\mathbf{M}_{optimal}{}^T\mathbf{u}$). However, the digester is a full rank process (its inputs can move the process in all directions in the output space). Therefore, the effect of adjusting the inputs in the ratio defined by $\mathbf{M}_{optimal}$ is exactly aligned with the effect of the disturbances. That is, the gain vector of the new latent input, u_{mv}, is perfectly aligned with the disturbance direction. It was discussed in the last chapter that for this special case, when the manipulated variable effect is perfectly aligned with the disturbance effect, one can control, in theory, ANY output (not just the ideal linear combination) and achieve the minimum error. This special case arises because the manipulated variables do not introduce any new directions of variation beyond that of the disturbances.

Therefore, instead of measuring all outputs and controlling a linear combination of them, one can, in theory, select the most convenient output and control it. Figure 4.4 shows the dynamic response of the outputs to a step change in the latent input u_{mv} (physically this corresponds to making adjustments to the real process inputs in the ratio given above). Two of the Kappa numbers (κ_{cook} and κ_{mcc})

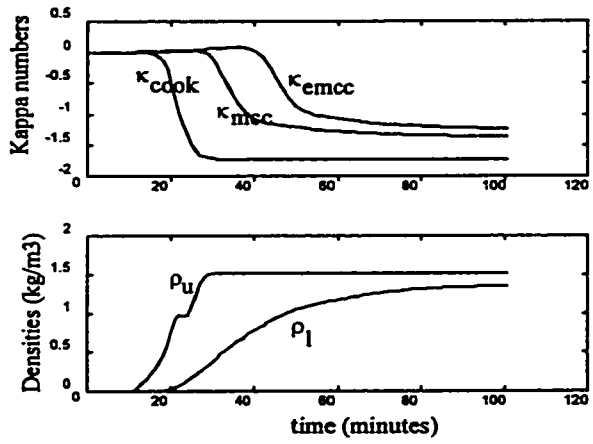

Figure 4.4. Effect of a step change in the manipulated variables in the calculated ratio

are not easily measured in practice, therefore these would not be good controlled variables. Ideally one would like to choose an output with little deadtime between it and the manipulated variable, therefore κ_{emcc} is not a good choice either. Of the remaining two outputs, the upper and lower extract densities, the upper extract density has faster dynamics and less deadtime, therefore from a dynamic control point of view this variable would be preferred.

However, the upper extract density has additional dimensionality issues due to the process nonlinearities. The ratio $\mathbf{M}_{optimal}$ is calculated based on process step changes about the nominal conditions. At these conditions, the effect of the latent input u_{mv} is perfectly aligned with the effect of the disturbances. However, as disturbances perturb the process, the nonlinearities result in changes to the input-output relationships. After 700 minutes, the effect of u_{mv} is no longer perfectly aligned with the disturbance effect, and in fact the gain vector of u_{mv} is 17° from the disturbance direction.

In order to assess the impact of controlling either the upper or lower extract density and manipulating the latent input u_{mv}, a Disturbance Inflation Factor (DIF) analysis can be done. Recall that the DIF is a steady state analysis tool, and requires only the input direction (gain vector of u_{mv}) and the disturbance direction. It is a ratio of two errors: the magnitude of the error after control divided by the magnitude of the error before control. For DIF = 0, the disturbance is completely eliminated in all outputs; for DIF < 1, the control actions attenuate the disturbance; for DIF > 1, the control actions inflate the disturbance; and for DIF = 1, the control actions have only

transferred the disturbance from the controlled output to the uncontrolled outputs with no overall reduction in variability. A low DIF is preferred. Table 4.2 summarizes the Disturbance Inflation Factors for both outputs, assuming u_{mv} is the manipulated variable. The DIF is calculated at time 0 and at time 700 minutes to evaluate the effect of process nonlinearities:

Table 4.2. Effect of process nonlinearities on overall output variability

Controlled Variable	DIF at t = 0 minutes	DIF at t = 700 minutes
ρ_l	0	0.29
ρ_u	0	0.98

Initially, since the gain vector of u_{mv} is perfectly aligned with the disturbance effects, both outputs have a DIF of zero. However, as the gain vector changes directions due to the process nonlinearities, the DIF of both controlled variables increases. From this analysis, it is clear that as the process changes with time, controlling the upper extract density no longer improves the overall variability. Simulations confirm that selecting ρ_u as the controlled variable results in offset in the outputs towards the end of the simulation. Therefore, despite the favorable dynamic response of ρ_u, the lower extract density is chosen as the controlled variable based on the above arguments.

4.3.3 Whole Variable Control

Although the ideal manipulated variable is a ratio of all five original variables, it may be preferable to go with a simpler control structure, such as controlling a single output with a single manipulated variable. A Disturbance Inflation Factor analysis is used screen for the best pair. Table 4.3 summarizes the DIF results (a dash indicates a zero gain between the given input/output pair):

Table 4.3. Disturbance Inflation Factor Analysis of SISO controllers

MV/CV	κ_{mcc}	κ_{emcc}	κ_{cook}	ρ_l	ρ_u
T_{cook}	1.97	1.82	1.78	8.73	1.14
T_{mcc}	1.05	1.03	-	1.44	-
T_{emcc}	1.56	1.17	-	1.18	-
F_{mcc}	0.79	0.77	-	0.75	-
$F_{upextract}$	0.42	0.37	0.5	0.32	0.39

Due to the excellent alignment of the upper extract flowrate ($F_{upextract}$) and the disturbance direction, the DIF for this manipulated variable is low for all controlled variables. Based on the earlier discussion, the lower extract density (ρ_l) was again chosen as the controlled variable, despite the more favorable dynamic response of ρ_u. Note that the lower extract density also has the smallest DIF at the nominal conditions. Therefore, both the latent variable SISO controller and the SISO with the single manipulated variable have the same controlled variable. The latent variable SISO should out-perform the latter since its manipulated variable is better aligned with the disturbance direction.

4.4 Results

In this section, the two SISO controllers will be implemented on the digester and compared to a DMC that controls all outputs by manipulating all inputs. Table 4.4 summarizes the controllers:

The two SISO control algorithms are implemented in IMC form. First order plus deadtime transfer function models were fit between the respective controlled and manipulated variables, and an IMC with a first order filter was implemented (Marlin (1995)). Both needed to be detuned; the filter parameters for the latent SISO and the whole variable SISO are 0.6 and 0.5, respectively (the tuning parameter was established by starting at 1 and backing off until a stable response was achieved).

Table 4.4. Summary of controllers

Controller	Controlled Variable(s)	Manipulated Variable(s)
Latent variable SISO	ρ_1	ratio of manipulated variables
Whole variable SISO	ρ_1	$F_{upextract}$
DMC	all five	all five

A Dynamic Matrix Controller was also implemented on the process. The scaling matrices used were:

$$Q = I_{5x5} \qquad W = 10000 * \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{bmatrix} \tag{4.40}$$

where Q is the output weighting matrix and W is the input weighting matrix. The matrix Q could have also been chosen to reflect the scaling, however it was found that the performance of the DMC was fairly insensitive to the output scaling. The relative input weights were chosen to reflect the different units of the flows and the temperatures. The multiplication factor (10000) was chosen by starting at one and increasing until a stable response was observed. All three controllers are executed every 10 minutes. Figures 4.5 and 4.6 show the results, while Table 4.5 summarizes the mean squared deviations from target.

The output behavior with all three controllers is much better than without control. As expected, the performance of the DMC is superior to the reduced dimension controllers, since it is controlling all the outputs and manipulating all the inputs. However, both reduced dimension SISO controllers perform very well given their much simpler structures. The latent variable controller in particular is very close

in performance to the DMC, despite the fact that it was formulated from steady state principles. Only the upper extract density shows some offset; the remaining four outputs are maintained at their setpoints despite the fact that only the lower density is explicitly controlled. The other SISO controller also does a very good job with only κ_{cook} and the upper extract density showing some offset.

Table 4.5. Controller Performance: mean squared deviation (e_i) from target

Controller	κ_{mcc}	κ_{emcc}	κ_{cook}	ρ_l	ρ_u	overall=$(\Sigma e_i)^{1/2}$
DMC	0.18	0.17	0.22	0.14	0.37	1.04
SISO latent	0.23	0.21	0.69	0.09	1.13	1.54
SISO whole	0.30	0.18	1.41	0.10	3.55	2.35
no control	14.84	13.18	17.52	12.94	12.43	8.42

It is interesting to point out that since the latent and the whole variable SISO have the same controlled variables, the difference in the performance is due entirely to the alignment of the manipulated variable and the disturbance. Since the latent controller has superior alignment, it does a better job of keeping all variables close to their setpoints.

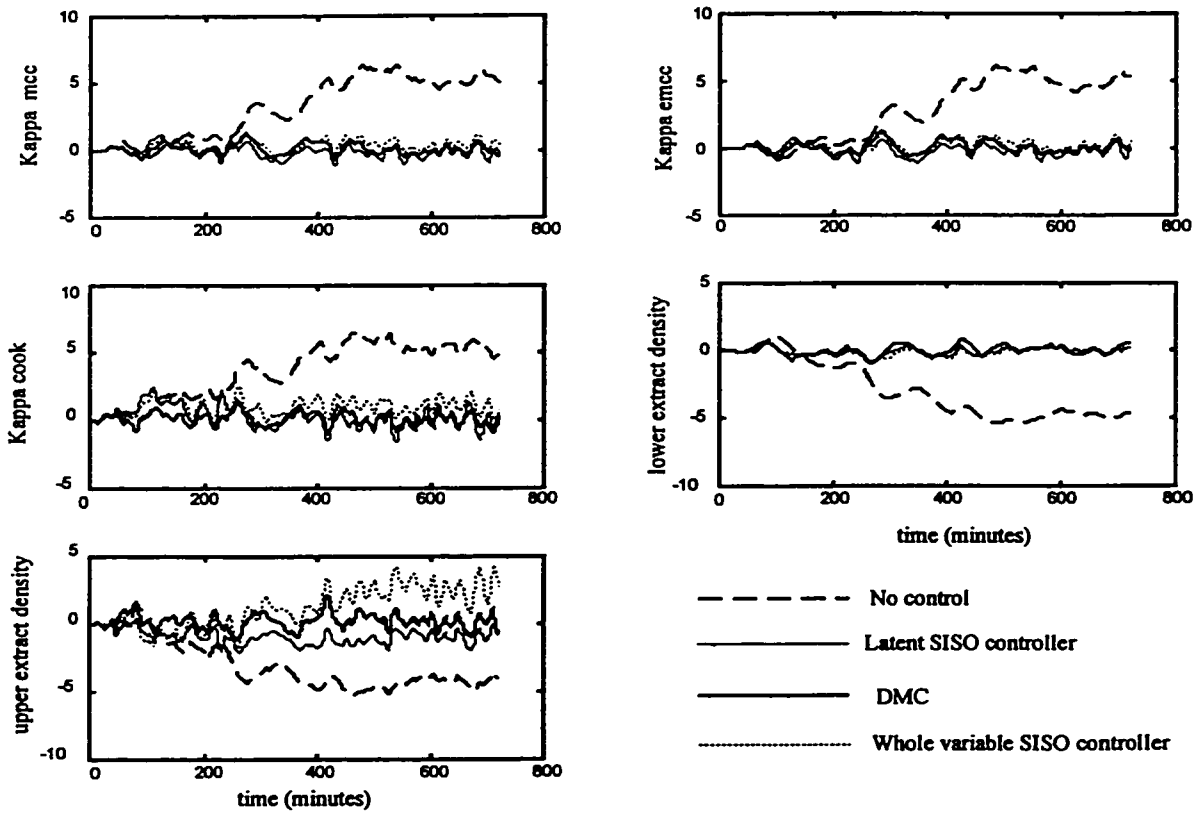

Figure 4.5 Comparison of controller performance

The behavior of the manipulated variables is shown in Figure 4.6. The move suppression is clearly seen in the temperatures, which was required for DMC controller stability. The DMC tends to move the flowrates more aggressively than the latent variable controller.

Despite the better performance of the DMC, it is a much more challenging controller to implement. It requires 25 dynamic process models and these models must be reasonably accurate or it was found that the controller could not easily be stabilized for this process. It is also more difficult to tune since the two weighting matrices must be specified. Furthermore, all output variables must be measured (two will be inferred). For the simulations above, all variables were assumed to be

measured online, therefore the DMC will show some performance degradation when measurement issues are considered.

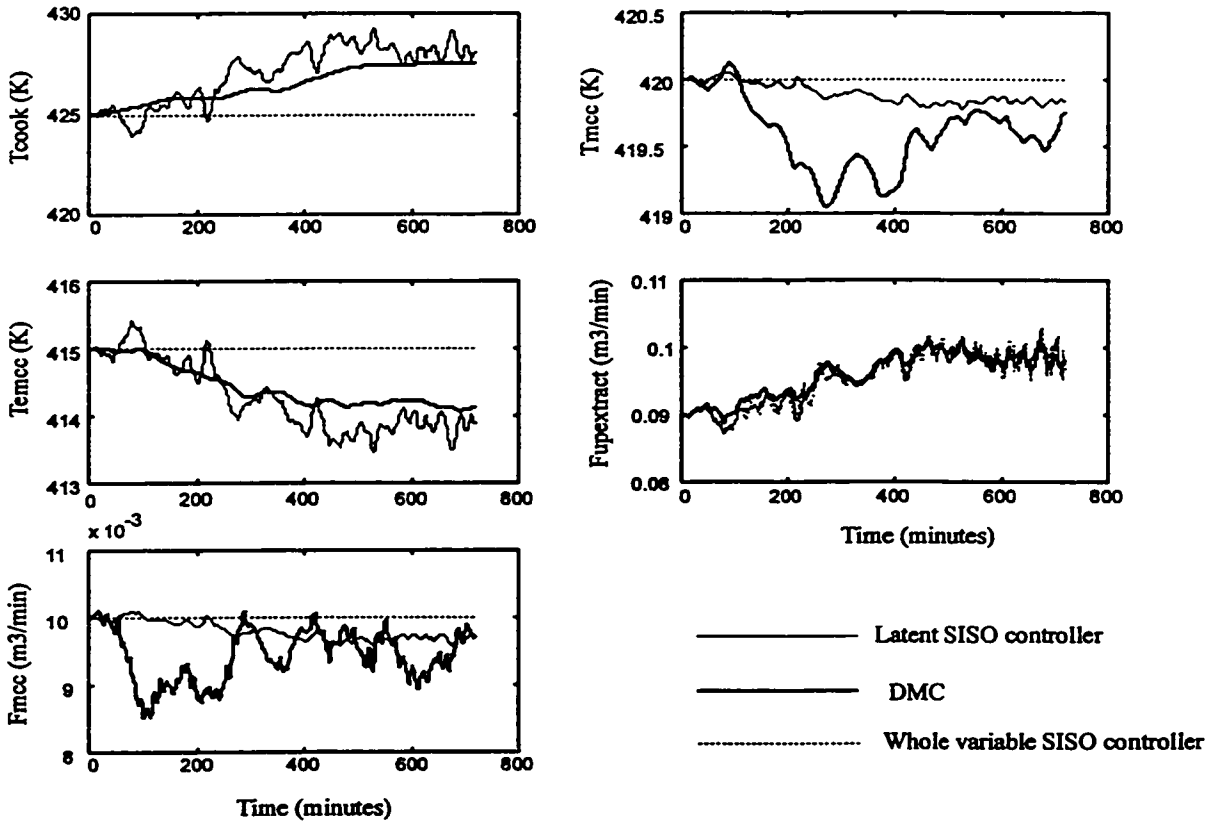

Figure 4.6. Manipulated variable behavior for the three controllers

On the other hand, the SISO controllers required much less process information. Steady state information only was used to calculate the appropriate ratio and only one dynamic transfer function was required. One tuning parameter (the filter time constant) needs to be specified. Due to their simple structure, the SISO controllers were much easier to implement and tune.

There is also the issue of controller robustness. For ill-conditioned multivariable systems, the performance of multivariable controllers such as the DMC is very sensitive to process-model mismatch. This issue was not relevant for the

digester, however for other poorly conditioned processes one might expect the reduced dimension controllers to be less sensitive to process-model mismatch due to their simpler structure. This was indeed shown to be the case in the blown film process of Featherstone and Braatz (1997) and the paper making process of Arkun and Kayihan (1998).

To evaluate the robustness of the two reduced dimension control systems, simulations were also run for different operating conditions. The digester has three different operating conditions. Of particular interest is the upper operating condition, since the process directions change significantly (the temperature effects in particular). The manipulated variable gain vectors do not change much in size, they mainly change direction. Also of note, is that the effect of the five disturbances does not change direction (it only changes by about 4 degrees compared to the original operating point).

The three controllers were implemented, unchanged, at the new operating point. Figures 4.7 and 4.8 and Table 4.6 show the results. All three controllers continue to perform satisfactorily. However, some performance degradation for the reduced dimension controllers is observed (primarily in one output, the upper extract density). This is not surprising since the structure of the RDCs were chosen based on the process directions, and when these change, the performance of the RDCs can be expected to degrade somewhat (although the RDCs do remain surprisingly good). A similar situation would arise if the disturbance direction changes; therefore if the disturbance is to be used in the formulation of the controllers, one must be sure that all representative disturbances are present in the disturbance data. For the digester, the disturbance effect appears to be fairly constant in direction regardless of the operating condition.

Table 4.6. Controller Performance: mean squared deviation from target

Controller	κ_{mcc}	κ_{emcc}	κ_{cook}	ρ_l	ρ_u	overall=$(\Sigma e_i)^{1/2}$
DMC	0.18	0.19	0.22	0.08	0.42	1.04
SISO latent	0.15	0.18	0.54	0.16	17.1	4.27
SISO whole	0.48	0.36	1.31	0.28	5.0	2.72
no control	11.48	11.29	13.54	9.69	13.07	7.69

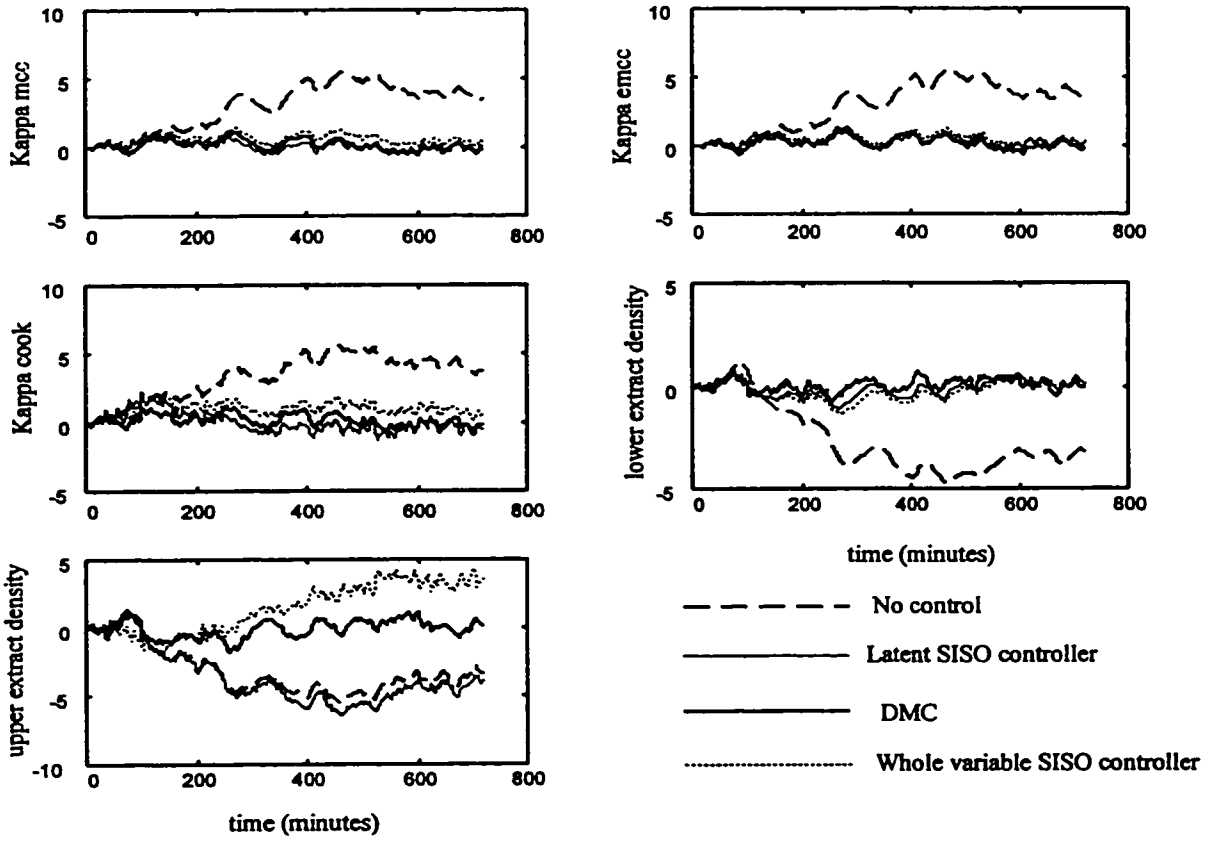

Figure 4.7. Comparison of controller performance at the second operating point

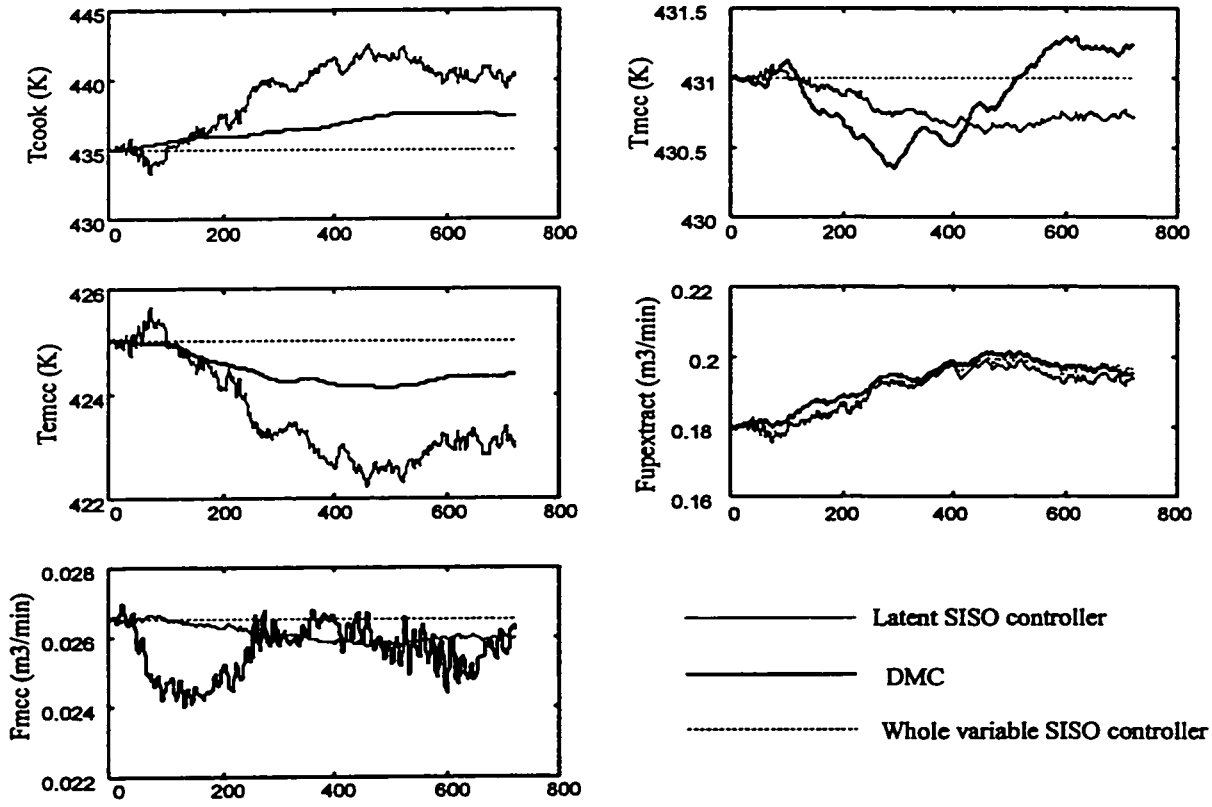

Figure 4.8. Manipulated variable behavior at the second operating point

In summary, the performance of the reduced dimension controllers have been shown to be very good over a wide range of process operations, despite the widely ranging deadtimes and dynamics of the Kaymr digester process. This illustrates that these RDC approaches can be useful even in very non-ideal situations.

4.5 Concluding Remarks

In this chapter, a general framework for reducing the dimension of the control system was presented. Given the disturbance directions and process gain matrix, expressions for the optimal directions for control were derived. The role of the number of independent disturbances in determining the number of controlled

variables, and the structure of the resulting reduced dimension controllers has been clearly shown.

In deriving expressions for the optimal controlled and manipulated variables, several assumptions were made. These included the assumption that the process is sampled infrequently relative to its dynamics or, equivalently, the process dynamics from all inputs to all outputs are the same. In order to illustrate the RDC approach and to evaluate the impact of violating the RDC assumptions, the framework for reducing the dimension of the control system was applied to a nonlinear, highly non-ideal dynamic simulation of a Kamyr digester. Two SISO Reduced Dimension Controllers were compared to a full dimension DMC. The RDCs performed very well at the conditions for which they were designed and showed only modest degradation when the process operating point was changed. It is expected, however, that the major impact on the performance of these RDCs would occur if the disturbance structure changed significantly. Since the structure of the RDCs is chosen in part based on the disturbance directions, it is important that all expected disturbances are characterized in the analysis.

4.6 Nomenclature

The following nomenclature was used in this Chapter:

y: nx1 vector of outputs

u: rx1 vector of inputs

t: time

G(B): process transfer function matrix

$\mathbf{y}_d$: nx1 vector of the overall disturbance

$\mathbf{p}_{di}$: disturbance direction vector

$\mathbf{Y}_d$: matrix of disturbance data

$\mathbf{P}_d$: matrix of disturbance directions

T: score matrix from PCA decomposition

$\mathbf{C}_{optimal}$: optimal directions for control

$\mathbf{M}_{optimal}$: optimal directions for manipulation

u_{mv}: new latent input

y_{cv}: new latent output

K: process gain matrix

$\mathbf{K}_s$: scaled process gain matrix

a_t: white noise

m: optimal input direction for a fixed controlled variable

k: number of independent disturbances

κ: Kappa number

ρ: density

T: temperature (K)

F: flowrate (m^3/min)

S: output scaling matrix

$\mathbf{S}_u$: input scaling matrix

DIF: disturbance inflation factor

Q: output weighting matrix in DMC

W: input weighting matrix in DMC

Appendix 4.1. Physical representation of the optimal controlled variable space and corresponding input directions

The matrix $\mathbf{C}_{optimal}$ (= $[\mathbf{c}_{1optimal}\ \mathbf{c}_{2optimal}\ \cdots\ \mathbf{c}_{k,optimal}]$) defines the *optimal controlled variable space*. This is the space that should be controlled in order to minimize the error in all outputs (not just those being controlled). It is assumed for convenience that each of the columns of $\mathbf{C}_{optimal}$ (i.e. the vectors $\mathbf{c}_{i,optimal}$) are scaled to unit length. Individual controlled variables, $y_{cv,i}$, are defined by the following linear combinations:

$$\begin{aligned} y_{cv,1} &= \mathbf{c}_{1,optimal}{}^{T}\mathbf{y} \\ y_{cv,2} &= \mathbf{c}_{2,optimal}{}^{T}\mathbf{y} \end{aligned} \qquad \text{(A.1)}$$

Rewritten in matrix form:

$$\mathbf{y}_{cv} = \mathbf{C}_{optimal}{}^{T}\mathbf{y} \qquad \text{(A.2)}$$

where the $\mathbf{y}_{cv}$ is a kx1 vector of new latent controlled variables and $\mathbf{y}$ is the nx1 vector of original outputs.

To drive $\mathbf{y}_{cv}$ to its setpoint *and* achieve the minimum overall error, the inputs must also be adjusted in a reduced dimension space. This space is spanned by the columns of the matrix $\mathbf{M}_{optimal}$ (= $[\mathbf{m}_{1,optimal}\ \mathbf{m}_{2,optimal}\ \ldots\ \mathbf{m}_{k,optimal}]$) where

$$\mathbf{M}_{optimal} = (\mathbf{K}^{T}\mathbf{K})^{-1}\mathbf{K}^{T}\mathbf{p}_{d} \qquad \text{(A.3)}$$

The associated input linear combinations (the new latent inputs) are given by:

$$\mathbf{u}_{mv} = (\mathbf{M}_{optimal}{}^{T}\mathbf{M}_{optimal})^{-1}\mathbf{M}_{optimal}{}^{T}\mathbf{u} \qquad \text{(A.4)}$$

Figure A.1 gives a schematic of the inputs for the case in which there are 3 inputs and two latent directions. The rectangular plane is the space spanned by the columns of $\mathbf{M}_{optimal}$. Any adjustments to the manipulated variables must remain within this space, and this will cause a corresponding variation in the outputs that falls in the

space $\mathbf{C}_{optimal}$. Furthermore, manipulating the inputs in the direction $\mathbf{m}_{1optimal}$ results in a change in the direction $\mathbf{c}_{optimal,1}$ in the output space, manipulating the inputs in the direction $\mathbf{m}_{2,optimal}$ results in a change in the direction $\mathbf{c}_{2,optimal}$ in the output space, and so on. The vector $\mathbf{u}_t$ is one particular example of input adjustments that satisfy the input structure requirement (i.e. $\mathbf{u}_t$ is in the space of $\mathbf{M}_{optimal}$).

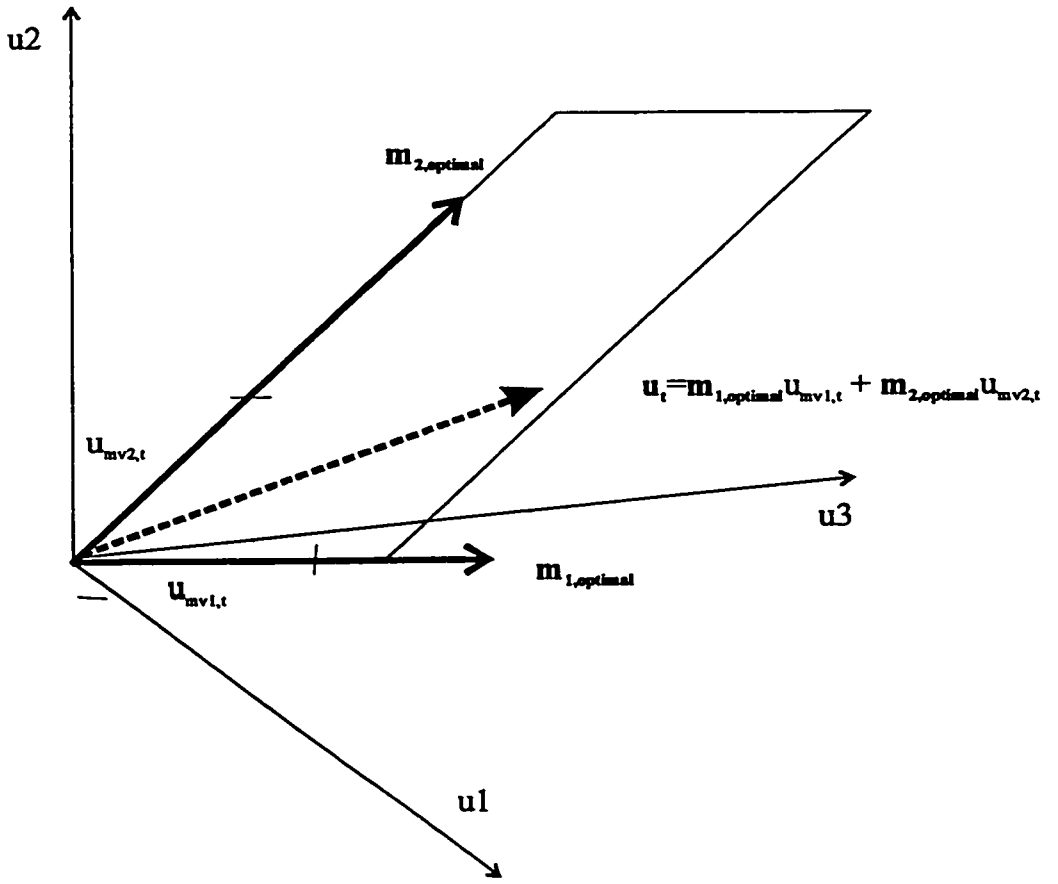

Figure A.1 Representation of the optimal input directions

5. Reduced Dimension Control: a Unification of Methods for Product Quality Control

5.1 Introduction

Reduced Dimension Control (RDC) is a growing field and there has been a number of published applications that can be classified as RDC, as defined in the previous chapter. Quite often, a reduced dimension control solution is problem-driven. Similar dimensionality problems surface in many different industries (for example in paper making processes and polymerization processes), and several independent solutions to the specific problems have been proposed. However, the common issues are often overlooked and therefore new applications are rarely placed within the context of other methods because of the different areas of application. The result is a number of apparently different methodologies for selecting subsystems of controlled and manipulated variables. In order to increase awareness and understanding in this promising field, this chapter will provide a unifying view of the existing methodologies and applications of reduced dimension control.

Processes that are candidates for RDC have one common, defining characteristic: the disturbances and manipulated variables move the process in a much smaller dimension than the measured output space. Certainly, there is no need to control the full output space if all the observed process variation falls into a smaller subspace. This was explicitly shown by the framework derived in the last chapter. However, there are also practical reasons for choosing to control a small dimensional subspace. For example, one can chose to control only the easily measured variables, since only some of the outputs need to be measured for RDC. An example is controlling an average of the MWD distribution, rather than the full MWD. Good control of the full MWD is a consequence of controlling a suitable average because of the low dimension of the process. Another

reason for considering reduced dimension control is improved model-based controller robustness to process/model mismatch. Full dimension controllers tend to have robustness problems for ill-conditioned (non-full rank) processes. This was discussed in the introduction of the last chapter, and will be illustrated through several examples in this Chapter. Finally, reduced dimension controllers have less complicated structures than their full dimensional counterparts. This may be more appealing to plant operators and engineers since they may be more easily maintained and they may prove more reliable in terms of measurement or actuator failure.

The outline of the Chapter is as follows. First, several RDC concepts will be discussed. Many of these concepts were introduced informally in earlier Chapters and will be formalized here. In the body of the Chapter, existing Reduced Dimension Control methods are divided into two main classes, Data Based Methods and Model Based Methods. Data Based Methods can be applied directly to plant data for dimension reduction, whereas Model Based Methods require a plant model (usually in transfer function form) from which the reduction is calculated. The methods are first discussed generally, then illustrated with specific process examples. When appropriate, the methods are related to the RDC framework from the previous Chapter, and situations in which the methods are expected to be unreliable are discussed. The Chapter concludes with a discussion on the merits and pitfalls of each of the methods, and an overall approach to RDC problems is proposed.

5.2 Reduced Dimension Control Concepts

In this section, several important process spaces are formally defined. The framework from the last chapter is also reviewed with respect to the process spaces.

5.2.1 Manipulated Variable Space

Random perturbations of the manipulated variables move the process in a subspace of the output space, which will be referred to as the manipulated variable (MV)

space. Considering only steady state, the manipulated variable space is the space spanned by the gain vectors of the manipulated variables (the gain matrix K). The manipulated variable space therefore defines where one can move the process; the set of achievable setpoints or regulatable disturbances is restricted by the directions and size of the manipulated variable space: only those setpoints that are in the space are actually achievable (feasible).

The effect of dynamics on the manipulated variable space is important and should be pointed out. In the steady state (non-dynamic) case, the dimension of the manipulated variable space will be less than or equal to the number of manipulated variables. However, dynamics can inflate the size of the space. For example, consider three output variables and one manipulated variable. If dynamics from this input to all output variables are the same, dynamic data (e.g. from a step change in the input) falls on the steady state gain vector (one dimension). If dynamics are different, two or even three dimensions may be observed since the dynamics cause the path of data to deviate from the steady state gain vector.

5.2.2 Disturbance Space

Perturbations due to the disturbances also move the process in a subspace of the output space. In general, this will differ from the MV space, and will be referred to as the disturbance space. The disturbance space can be characterized as in the last chapter (equation 4.2). The disturbance space is the subspace spanned by individual disturbance direction vectors ($\mathbf{P_d} = [\mathbf{p_{d1}}\ \mathbf{p_{d2}} \cdots \mathbf{p_{dn}}]$). Setpoint changes can also be considered disturbances; if setpoint changes are to be called for in all the outputs, the corresponding disturbance space is the full output space ($\mathbf{P_d} = \mathbf{I}$).

5.2.3 Controlled Variable Space

Each controlled variable is defined by its controlled variable axis. For example, take (y_1, y_2 ,y_3) as the output variables. If y_2 is chosen as the controlled variable, the

controlled variable axis (vector) is $[0\ 1\ 0]^T$. The span of the controlled variable axes (vectors) is referred to as the controlled variable space (here, $[0\ 1\ 0]^T$ is the controlled variable space). If there are two controlled variables, then the controlled variable space is a plane, if there are more than two controlled variables, the controlled variable space is a hyperplane.

The following two terms will be used frequently in the paper as well, and must be defined and explained.

Latent variable: a latent variable is a new variable defined as a combination of the existing whole variables, which summarizes in some sense the dominant structure or variation in the output space.

Latent direction: each latent variable has a latent direction associated with it. For example, say a latent direction **P** defines the direction of greatest variability in a process, where $\mathbf{P} = [0.7\ 0.2]^T$. The latent variable associated with this latent direction is $z = \mathbf{P}^t\mathbf{y} = 0.7y_1+0.2y_2$.

5.2.4 Framework for Unification

The framework for discussing the methods was derived in the last chapter and is summarized briefly here, using the process space terminology from the previous section. Many of the reduced dimension methods will be related back to this framework.

First, the optimal linear combinations of the inputs and outputs that should be controlled was derived. That is, an expression for the optimal *controlled variable space* was derived as:

$$\mathbf{C}_{optimal} = \mathbf{K}\left(\mathbf{K}^T\mathbf{K}\right)^{-1}\mathbf{K}^T\mathbf{P}_d \tag{5.1}$$

The optimal controlled variable space is that space spanned by the columns of $\mathbf{C}_{optimal}$. For convenience it is assumed that the columns of $\mathbf{C}_{optimal}$ are scaled to unit length. As

discussed in Chapter 4, the corresponding optimal set of controlled variables is given by the linear combinations $\mathbf{y_{cv}} = \mathbf{C_{optimal}}^{T}\mathbf{y}$.

The corresponding space in which the inputs should be manipulated is given by:

$$\mathbf{M_{optimal}} = (\mathbf{K}^{T}\mathbf{K})^{-1}\mathbf{K}^{T}\mathbf{P_d} \tag{5.2}$$

In the remainder of the Chapter, the term 'input directions' will be used to describe the space spanned by the columns of the matrix $\mathbf{M_{optimal}}$. The associated optimal set of manipulated variables is defined by the linear combinations $\mathbf{u_{mv}} = (\mathbf{M_{optimal}}^{T}\mathbf{M_{optimal}})^{-1}\mathbf{M_{optimal}}^{T}\mathbf{u}$.

Therefore, the optimal controlled variable space is a result of projecting the disturbance directions onto the manipulated variable space. The dimension of the controlled variable space (the number of independent directions that must be controlled to minimize the overall error) will be given by either the dimension of the disturbance space or the dimension of the manipulated variable space, whichever is smaller.

Two special cases arise frequently, and merit some further discussion. First, in some processes, one may be interested in setpoint changes, rather than disturbance regulation. Within the framework discussed above, the 'disturbance' is the entire output space, therefore $\mathbf{P_d} = \mathbf{I}$. The optimal controlled variable space becomes:

$$\mathbf{C_{optimal}} = \mathbf{K}\left(\mathbf{K}^{T}\mathbf{K}\right)^{-1}\mathbf{K}^{T}\mathbf{I} \tag{5.3}$$

The projection of the identity matrix onto the manipulated variable space is simply the manipulated variable space itself. Therefore, for the case which setpoint changes are of interest, the optimal controlled variable space is the manipulated variable space (characterized by the span of the gain matrix vectors, $\mathbf{C_{optimal}} = \mathbf{K}$).

A special note should also be made concerning the calculation of $\mathbf{C_{optimal}}$ for ill-conditioned processes. When the process is ill-conditioned, K is not full rank and $\mathbf{K}^{T}\mathbf{K}$

cannot be inverted for the calculation. However, the projection operator $\mathbf{K}(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T$ may still be defined by replacing **K** with a full rank matrix whose columns span the same space as the columns of **K**. A convenient choice for such a matrix is $\mathbf{W}_{nxa}$. The matrix $\mathbf{W}_{nxa}$ is obtained from a singular value decomposition of **K** ($\mathbf{K} = \mathbf{WSV}^T$), where $\mathbf{W}_{nxa}$ are the first 'a' columns of **U** associated with non zero singular values. The optimal controlled variable space for ill-conditioned processes can then be calculated:

$$\begin{aligned} \mathbf{C}_{optimal} &= \mathbf{W}_{nxa}\left(\mathbf{W}_{nxa}{}^T\mathbf{W}_{nxa}\right)^{-1}\mathbf{W}_{nxa}{}^T\mathbf{P}_d \\ &= \mathbf{W}_{nxa}\mathbf{W}_{nxa}{}^T\mathbf{P}_d \end{aligned} \tag{5.4}$$

since $\mathbf{W}_{nxa}{}^T\mathbf{W}_{nxa} = \mathbf{I}$. Note that replacing **K** with the matrix $\mathbf{W}_{nxa}$ does not change the interpretation of $\mathbf{C}_{optimal}$.

5.3 Data Based Methods

In this section, data-based methods for reducing the dimension of the control structure will be discussed. The methods are grouped into two main classes, or approaches: Principal Components (PCA) Based Methods and Canonical Variates Based Methods. The layout of each section is as follows. First, a general description of each approach will be given first, and then the approach will be discussed and illustrated through specific applications found in the literature. The approaches are related back to the optimal framework, and any situations in which these approaches may provide unreliable results are discussed.

5.3.1 PCA-based Methods

Principal Components Analysis (PCA) is a multivariable statistical method that finds the directions of greatest variance within the data, and characterizes them with a set of 'a' orthogonal basis vectors (columns of the PCA matrix **P**, where **P** is an nxa matrix). The latent variables associated with these directions are given by $\mathbf{t} = \mathbf{P}^T\mathbf{y}$. PCA is applied

directly to the plant data with no need to identify a process model first (Wold et al 1987). The application of PCA for model reduction and possibly control space characterization was recognized as a potential approach as early as Moore (1981).

Because it focuses on variability within the data, PCA is very effective for determining the dimension as well as the directions of the process spaces. The 'a' vectors that form the columns of **P** are a basis for the variation observed in the data. The dimension of the space that contains significant variation ('a', where 'a' is less than n, the number of measured outputs) can be determined using various criteria such as cross validation (Wold et al (1987)).

There are several types of output data that one can obtain from a process, and depending on the data, PCA can be used to characterize the different spaces. For example, if one has pure open loop data (data that contains disturbance information only) then an PCA of this data will give the reduced dimension disturbance space ($\mathbf{P} = \mathbf{P}_d$). If data is available that contains only manipulated variable perturbations, the PCA vectors will characterize the manipulated variable space ($\mathbf{P} = \mathbf{K}(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{I}$). If the data are available from open or closed-loop plant tests, in which test signals are introduced in all the inputs, and in which representative disturbances are present, a PCA will uncover the union of disturbance and manipulated variable spaces ($\mathbf{P} = \text{union}(\mathbf{P}_d, \mathbf{K})$). In all three situations, some very useful information is obtained from a PCA analysis of the data. A final type of data that is readily available is regular operating data, or pure feedback data in which no dither signals have been added to the inputs. In this case, there is limited information that can be extracted from the data. The main variation will appear in the variables that are not controlled (depending of course on the effectiveness of the feedback action) and will generally not be indicative of the manipulated variable or disturbance spaces.

As well as providing a method for determining the effective dimension of the process spaces, Principal Components Analysis can also be applied as a technique for controlled variable selection. For example, consider a process with ten quality variables. A PCA of plant test data may indicate only four dimensions of variability. First, this

implies at most four directions need to be controlled, since all the observed variability falls in only four dimensions. Two options exist. First, one may choose to control the four directions that are given by a PCA analysis. This corresponds to controlling the four latent variables associated with the each of the PCA directions ($y_{cv} = P^T y$). Second, one may choose instead to control four physical variables. Both approaches have been suggested and will be discussed using examples below.

One comment is in order prior to discussing the examples. In applying PCA for controlled variable selection, some care must be taken with respect to the data that is analyzed. For example, if one uses PCA to analyze disturbance data, the space characterized by PCA is the disturbance space. If one then proceeds to use the analysis for controlled variable selection, the resulting controlled variable space could be completely different from the controlled variable space derived within the optimal framework since no manipulated variable information is contained within the disturbance data. Therefore, a poor set of controlled variables may result, particularly if a large portion of the disturbance lies outside the manipulated variable space.

The most common type of data that PCA should be applied to is plant test data or open loop data with dithers on the inputs. In this case, a PCA analysis gives the union of disturbance and manipulated variable spaces. While this will not, in most cases, result in the minimum number of controlled variables expected from the optimal framework, the controlled space indicated by PCA (C_{pca}= **union(K,Pd)**) will contain the optimal controlled space from the framework ($C_{optimal} = K(K^TK)^{-1}K^TP_d$).

Two applications that apply PCA for controlled variable selection will be discussed below. One application uses PCA directly, and proposes to control linear combinations of the original variables, and one application proposes a modification to PCA, for selecting a subset of the original variables for control.

Control of Crystal Size Distribution

Principal Components Analysis was applied for reduced dimension control of the crystal size distribution (CSD) in a continuous crystallizer (Eek and Bosgra (1995a and b), Eek et al. (1995, 1996)). The crystal size distribution was characterized by recording 31 locations along the distribution. Three inputs could be varied: fines removal, total heat input and product removal rate. The motivation for pursuing dimensionality arose primarily from the desire to avoid using a complicated model based scheme (Eek et al (1995)), and an awareness that there were clearly limited degrees of freedom in the plant.

To assess the process dimensionality and the potential for dimension reduction, a PCA of plant data was performed. There were 10 unique data sets, containing plant tests, startups, shutdowns and regular operations over 4 years. A PCA analysis was performed on each of the ten data sets. It was found that six dimensions explained between 95 and 99 % of the variation in each of the data sets. These six directions also tended to be fairly consistent from data set to data set. Therefore, there is definite potential for reduction.

Eek et Bosgra (1995a) then applied PCA for selecting controlled variables. They proposed that the first principal component be controlled, that is, control the latent variable $z_1 = \mathbf{p}_1^T\mathbf{y}$, where $\mathbf{p}_1$ is the direction of largest variation and y contains the 31 measured values along the CSD. Physically, z_1 is one particular weighted sum of the CSD variables, and it was chosen because of its high variability (note that another particular weighted sum of the CSD is the average crystal size, which is often chosen due to its physical significance). Following the controlled variable selection, several reduced dimension control structures were tested: a SISO controller, with the controlled variable taken as z_1, and the manipulated variable taken as the fines removal rate, and a 2x2 MIMO controller (the controlled variables were z_1 and the slurry density while the manipulated variables were the fines removal rate and the total heat input). Results were reported for the controlled variables only and in both cases, good control of the selected outputs was reported. However, the impact of the control actions on the full CSD wasn't explicitly given.

In this application, the data analyzed contained both disturbance and manipulated variable information. However, it covered a very wide range of steady state operations and some major dynamics due to start up and shutdown. This might very well have inflated the apparent dimension of the combined disturbance and manipulated variable spaces. All that is known is that the expected process variation appears in six dimensions. It is therefore uncertain what variation is captured in the first principal component. It may be primarily disturbance information or manipulated variable information, or a combination of both. It is unknown whether the direction $\mathbf{p}_1$ is going to be close to optimal controlled space, $\mathbf{C}_{optimal}$ for this process.

Control of Industrial Polybutadiene Reactor

Roffel et al (1989) used Principal Components Analysis for reducing the dimension of a multivariable controller for an industrial polybutadiene reactor. In the butadiene process, three outputs that characterize quality were measured (mooney viscosity, fluids viscosity, and percent solids). Three inputs could be adjusted (flowrate of catalyst, flowrate of modifier and flowrate of promoter).

A PCA of plant test data indicated that 96% of the variation was explained by only two dimensions. Therefore, at most, only two of the three dimensions need to be controlled. The option certainly exists to control the two linear combinations given by the PCA analysis, as was illustrated with CSD control. In the CSD control problem, a linear combination of distribution variables has some physical meaning, being a weighted average of the CSD. However, a linear combination of different variables, such as mooney viscosity, fluid viscosity and percent solids is less meaningful. The operators and engineers opposed controlling linear combinations of the outputs for this reason and preferred that whole outputs be chosen for control instead.

This is typical of many industrial processes in which it may be preferable to control whole outputs rather than linear combinations. Roffel et al. (1989) therefore propose a systematic procedure called Selective Principal Components Analysis (SPCA)

for choosing subsets of the output variables that are well aligned with the variation observed in the data. First, a regular PCA of plant data is performed, in order to decide the number of dimensions that need to be controlled. Then, the output associated with the largest element in the first principal component (or one of the largest elements) is chosen as the first controlled variable. This is the output that is best aligned with the direction of largest variation. After a particular output is chosen as a controlled variable, the data is deflated prior to selecting the next controlled variable. This is accomplished by regressing all other output variables on the chosen controlled variable and then taking residuals. All variation in the chosen variable is removed (corresponding to perfect control) AND all the variation in all other outputs that is correlated to the chosen variable is therefore also removed. A PCA is then performed on the residuals and the procedure is repeated until the desired number of controlled variables have been chosen.

Roffel et al (1989) applied SPCA to plant test data from the polybutadiene reactor, and the analysis indicated that mooney viscosity and percent solids should be controlled. The residuals of SPCA show that less than 7% of the variation remains after controlling these two outputs. Two manipulated variables, the modifier and catalyst flowrates, were chosen based on the process engineer's experience. A 2x2 multivariable controller was implemented on the plant and good overall control (in all three outputs) was reported, even though only two were directly controlled.

A key advantage of the SPCA method is that the resulting controlled variables are complementary due to the nature of the deflation. Deflation removes all variation in the chosen output, as well as variation in the other outputs that is correlated with the chosen output. By perfectly controlling the chosen output, that variation in the remaining outputs that is correlated with the chosen output will also be removed. This is accounted for in SPCA, prior to selecting the next output for control. Therefore, subsequent controlled variables will be complementary to the ones chosen already.

For illustration, in the polybutadiene example, given that only two output dimensions needed to be controlled, the plant operators and engineers wanted to control

the two viscosities, since these were the variables that had the highest variability. The first principal component of the plant data also indicated that these two variables had the largest elements in the first principal component. However, SPCA revealed that once the mooney viscosity was selected as the first control variable, the residual variation in the fluid viscosity was small. Therefore, it was only necessary to directly control one of the viscosities, and the percent solids should be chosen as the second controlled variable. The results of the SPCA were confirmed in practice when the 2x2 multivariable controller was implemented on the system. It was seen that the variability of the fluid viscosity was reduced substantially even though it was not directly controlled. Therefore, the SPCA selected the two most complementary variables for control.

Summarizing, PCA-based methods are effective for evaluating dimensionality and characterizing process spaces from plant data. Both PCA and SPCA are effective for choosing complementary sets of controlled variables, and although the methods will not necessarily result in the theoretical minimum number of controlled variables, some dimension reduction will result. Furthermore, some insight into the process may be obtained by interpreting the linear combinations given by the principal components. The main drawback to PCA based methods is that the data used in the analysis needs to be chosen with caution and must contain manipulated variable information. No guidance into manipulated variable selection is provided since causal relationships between the input and output spaces are not available.

5.3.2 Canonical Variates Based Methods

Canonical Correlation Analysis (CCA) is another multivariate statistical method closely related to PCA (Box and Tiao (1977)). PCA finds the directions of highest variability. CCA, on the other hand, finds directions in dynamic data that are most predictable, and ranks them according to predictability. Box and Tiao (1977) developed this approach in order to reduce the dimension of multivariate time series data. It was

then applied by MacGregor and Wong (1980) for obtaining reduced dimension controllers.

An ideal control application of CCA is disturbance space reduction because CCA uncovers the most predictable linear combinations. Only the predictable portion of the disturbances can be eliminated by feedback control. Therefore, CCA can give insight into the *controllable* disturbance space if applied to data collected under pure open loop conditions. Note that this subspace may be different from that obtained by PCA which obtains the subspace based only on variance regardless of whether it is predictable. Jutan et al. (1984) applied CCA to reduce the dimension of the disturbance space in a packed bed reactor system from seven dimensions to three dimensions.

In some cases, particularly in an industrial environment, obtaining disturbance data is not feasible. However, CCA can still be applied to process output data with manipulated variable moves present. With respect to characterizing the combined disturbance and manipulated variable spaces, CCA is perhaps less useful than PCA under these circumstances. The predictable components will be some combination of the disturbances and the effects of the manipulated variables and the results of the CCA analysis will depend on the input signal as well as process characteristics. However, even in this case CCA can provide useful insight into whether a dimension reduction is possible.

It should be emphasized that CCA is a 'dynamic' tool. In the method, it is assumed that the original time series can be represented by a multivariable AR(k) model, that is, the data must have trends or dynamics. If, for example, the process is sampled infrequently relative to the process dynamic and the inputs are forced by random PRBS test signals, a CCA analysis provides no insight since all components will be unpredictable.

MacGregor and Wong (1981) applied CCA to obtain a reduced dimension control system for a packed bed hydrolysis reactor. The process had three outputs (production rates of butane, propane and hydrogen) and two inputs (flowrates of butane and

hydrogen). Since the plant was open loop unstable, quasi-closed loop plant test data was collected, in which one of the flows was adjusted to maintain the bed hot spot temperature at a setpoint and test signals were added to both inputs. A CCA analysis indicated two highly predictable components and one component that was almost entirely white noise and had a very small variance. The authors chose to control two of the three production rates, rather than the two linear combinations indicated by CCA, since the former was more meaningful. Propane and butane production rates were chosen since they had the highest signal to noise ratio.

5.4 Model Based Methods

5.4.1 Singular Value Decomposition Methods

As a control and analysis tool, Singular Value Decomposition (SVD) was originally developed to design controllers for interacting square multivariable processes. However, it is also useful for the analysis of non-square processes and non-full rank processes due to the insight it provides into the underlying process dimension. In this section, Singular Value Decomposition will be discussed in the context of reduced dimension control. First, the general method will be outlined and briefly related to the RDC framework. Following this, two recent applications will be discussed in order to illustrate the SVD approach and highlight key results.

The SVD method has been applied by many research groups (for example, Hovd et al (1996), Zhu et Jutan (1998), Keller and Bonvin (1987), Lau et al (1985)). The general methodology is based on the singular value decomposition of the input-output gain matrix (assuming n outputs and r inputs):

$$\begin{aligned} \mathbf{y} &= \mathbf{G}(s)\mathbf{u} \\ &= \mathbf{W}(s)\mathbf{S}(s)\mathbf{V}(s)^{T}\mathbf{u} \end{aligned} \tag{5.5}$$

where **W** is an nxn orthogonal matrix, **V** is an rxr orthogonal matrix and **S** is a diagonal nxr matrix. In the general case, **V** and **W** are frequency dependent (but constant at a given frequency). However, one of the following simplifications can often be made (Lau et al (1985)):

1. A particular frequency is established as a 'critical frequency', or most important for the process for a physically significant reason. Therefore, V(s) and W(s) are replaced by constant matrices **V** and **W** at that frequency.
2. The process is a steady state process; therefore, **V** and **W** are constant
3. The process has the same dynamics from all inputs to all outputs; therefore V(s) and W(s) are the same constant values at all frequencies.

In quality control, one of the last two conditions are often true. For illustrative purposes, assume that the third condition is met. The singular value decomposition of the gain matrix becomes:

$$\begin{aligned} \mathbf{y} &= \frac{\omega(s)}{\delta(s)}\mathbf{K}\mathbf{u} \\ &= \frac{\omega(s)}{\delta(s)}\mathbf{W}\mathbf{S}\mathbf{V}^{T}\mathbf{u} \end{aligned} \tag{5.6}$$

Rearranging,

$$\begin{aligned} \mathbf{W}^{T}\mathbf{y} &= \frac{\omega(s)}{\delta(s)}\mathbf{S}\mathbf{V}^{T}\mathbf{u} \\ \mathbf{z} &= \frac{\omega(s)}{\delta(s)}\mathbf{S}\mathbf{m} \end{aligned} \tag{5.7}$$

An inner system is defined between the variables **z** (= $\mathbf{W}^{T}\mathbf{y}$) and **m** (= $\mathbf{V}^{T}\mathbf{u}$). Only the first r variables in **z** (at most) are affected by the 'r' inputs ($\mathbf{z}_r$, where $\mathbf{z} = [\mathbf{z}_r\ \mathbf{z}_{n-r}]$). Furthermore, since the matrix **S** is diagonal, the square, inner system between $\mathbf{z}_r$ and **m** is non-interacting (m_i only affects $z_{r,i}$).

In the context of reduced dimension control, the singular value decomposition is useful for two reasons. First, it is helpful in characterizing the dimension and directions of the manipulated variable space due to the physical interpretation of the vectors of SVD. When the inputs are perturbed along the direction $\mathbf{v}_1$ in the input space (where $\mathbf{v}_1$ is the first column of V), variation is observed in the output space in the direction $\mathbf{w}_1$ (where $\mathbf{w}_1$ is the first column of W). The magnitude of the variation is given by the size of the associated element in S (the elements in S are traditionally ordered from largest to smallest). Therefore, the vectors $\mathbf{u}_1$, $\mathbf{u}_2$... (the columns of the matrix W) are a basis for the directions in which the inputs can move the process. That is, the columns of the matrix W span the manipulated variable space (as defined in the section 5.2). The shape and dimension of manipulated variable space is indicated by diagonal elements of S, s_i. If the s_i element is very close to zero, the manipulated variables have limited ability to move in the associated output direction, $\mathbf{w}_i$ (indicative of ill-conditioning in the process). The number of non-zero diagonal elements gives the dimension of the manipulated variable space. Therefore, SVD is very effective in assessing the manipulated variable dimensions and directions. Note here that a singular value decomposition of the gain matrix, and a PCA of data that contains only manipulated variable moves are in fact equivalent ways of defining the manipulated variable space.

The SVD can also be used for controlled variable selection. As with PCA, one can use SVD to select linear combinations (new latent variables) to control, or for insight into which whole variables should be controlled. Both will be discussed, with emphasis on the former.

Selection of latent variables is very straightforward from the singular value analysis. New latent controlled variables are defined by $\mathbf{z}_r = \mathbf{W}_{nxr}^T\mathbf{y}$ (where the subscript nxr indicates matrix dimension) and the new latent manipulated variables by $\mathbf{m} = \mathbf{V}^T\mathbf{u}$. The controlled variable space is given by the first r columns of W ($\mathbf{C}_{svd} = \mathbf{W}_{nxr}$) and the input directions (as discussed in section 5.2.4 and Appendix 4.1 from Chapter 4) are given by the columns of V. Therefore, the controlled variable space resulting from the

SVD approach is the manipulated variable space. This is exactly equivalent to the optimal controlled variable space derived within the framework ($\mathbf{C}_{optimal}$, equation (5.3)) for the disturbance assumption $\mathbf{P}_d = \mathbf{I}$ (setpoint changes).

The selection of whole variables for control using the SVD is equally straightforward (Luyben (1989), Smith et al (1981). Once an SVD of the gain matrix is performed, select as the first controlled variable the output with the largest element in the vector $\mathbf{w}_1$. This corresponds to selecting the whole output that is best aligned with the largest direction of variation in the manipulated variable space. This output should then be paired with the input with the largest element in $\mathbf{v}_1$. This procedure is repeated for the other SVD components. Since $\mathbf{W}$ is an orthonormal matrix, the columns are perpendicular to each other, the resulting set of controlled variables should be somewhat complementary. However, unlike SPCA, the procedure does not explicitly account for the correlation or interaction effects that controlling a selected output will have on the other outputs. Lau et al (1985) extend the SVD-based selection approach to include a measure of interaction when one chooses to pair individual inputs and outputs rather than controlling and manipulating linear combinations.

Two recent applications will now be discussed to illustrate the SVD method and highlight important results.

Control of a Blown Film process

An application of SVD to reduced dimension control can be found in Featherstone and Braatz (1997). The simulated blown film example addressed in this publication is very typical of observations on large ill-conditioned processes and therefore merits discussion.

In the blown film process, film thickness is measured at forty-five unique locations. There are also forty-five actuators available. The dynamics from all inputs to all outputs is the same. A standard multivariable controller (Quadratic Penality Function method) is widely used in industrial sheet and film processes, however, poor performance

of the QPF controllers are often reported. The authors suspected that dimensionality issues were the root cause and proposed an SVD analysis.

An SVD of 45x45 gain matrix indicated that only ten significant directions (only ten of the diagonal elements of S were statistically significant). This implies a process limitation in that the inputs cannot independently control all the outputs since the true underlying dimension of the manipulated variable space is only ten. Therefore, the authors proposed that only the 'significant' directions be controlled, and a reduced dimension controller be implemented to control the 10 dominant singular directions of the output space ($\mathbf{w}_1$, $\mathbf{w}_2$ ... $\mathbf{w}_{10}$) by adjusting the inputs in the corresponding 10 singular directions of the input space. That is, the following linear combinations are chosen for the reduced dimension controller:

$$\begin{aligned} \mathbf{y}_{cv} &= \mathbf{z}_{10} = \mathbf{W}_{45x10}{}^{T}\mathbf{y} \\ \mathbf{u}_{mv} &= \mathbf{m}_{10} = \mathbf{V}_{10x10}{}^{T}\mathbf{u} \end{aligned} \tag{5.8}$$

This example is particularly insightful as it compares the RDC (which acts on only 10 output directions) to the standard QPF controller (which attempts to directly control all 45 outputs). It is seen that the performance of the SVD based reduced dimension controller is significantly better than the full dimensional QPF, when variability in all the outputs is quantified. This is attributed to that fact that the QPF attempts to adjust the manipulated variables in all directions in the output space (including weak ones), and due to the process model mismatch (particularly in weak directions), deterioration in the overall output performance is seen.

Control of Sheet Forming Processes

Several research groups have also addressed control of paper making machines using a SVD based approach (Featherstone and Braatz (1997,1998), Arkun and Kayihan

(1998), Rigopoulos et al (1997)). In this section, the application of Arkun and Kayihan (1998) to a simulated paper making machine will be discussed.

This paper making process has 125 measurements relating to paper quality and 125 actuators (slice lips) across the sheet of paper, with similar dynamics and deadtimes from all the actuators to all the measured outputs. Control of such paper making processes has been a challenging control problem for industry due to the large scale and ill-conditioned nature of the process. As with the blown film process, the paper making process is an example of a process in which full multivariable model based controllers often exhibit poor performance due to a lack of robustness to process/model mismatch.

A dimensionality analysis of the paper making process gave the following insight. An SVD of the 125x125 gain matrix indicated that the manipulated variable space is only 75 dimensions. Furthermore, a PCA of open loop disturbance data indicated that there were only 2 disturbance directions (details of the PCA analysis can be found in Rigopoulous et al (1997)).

Hence, the authors suggested that one need only control two dimensions. Rather than use the SVD methodology directly (which would result in 75 output directions being controlled), the authors proposed to reduce the controller dimension further, from 75 to 2, by incorporating the disturbance directions into the controller. The two disturbance directions were projected onto the manipulated variable space (defined by the columns of W_{125x75} from the SVD of the gain matrix), and the authors proposed to control only these two directions, with the corresponding input directions. The controller had an additional online disturbance estimation step (details may be found in the publications) in order to handle changing disturbance structures. It can also be shown that the controlled variable space (and corresponding input directions) in this method is equivalent to the controlled variable space derived with the framework earlier. A more detailed explanation of the mathematics of the approach and its relationship to the optimal framework may be found in Appendix 5.1 at the end of this Chapter.

The two dimension RDC was applied to a simulated paper making process and compared to a full dimension IMC (that acts explicitly on all 125 outputs). The RDC, which acted explicitly on only 2 output directions, was shown to outperform a full dimensional IMC controller, when variability in all the outputs were considered. Again, this was attributed to the lack of robustness of model based multivariable controllers when applied to ill-conditioned processes.

5.4.2 Gain Matrix Indices

As was seen in the section on singular value analysis, a lot of information can be obtained from the input-output matrix. In this section, two indices that are calculated directly from the gain matrix will be discussed. One index, the Disturbance Inflation Factor, explicitly considers disturbances, while another, the Non-Square Relative Gain Array does not.

5.4.2.1 Non-Square Relative Gain Array

In chemical engineering, processes with more outputs than inputs are very common. Two approaches to controlling non-square processes are possible. One could implement a non square controller, or one could choose to eliminate some of the outputs and act on the resulting square process. Process control theory (both the analysis and the algorithms) historically caters to square systems, therefore it would be desirable to have a method for paring down the outputs and squaring a non-square system. Chang and Yu (1990) address this problem by proposing a tool based on the Non-Square Relative Gain Array (NSRGA).

In Chang and Yu (1990), an index called the Row Sum, based on the NSRGA, is derived. For details, the reader is refered to Chang and Yu (1990) as only a brief overview is given below. The row sum (the sum of each row of the NSRGA) is calculated from the steady state gain matrix and is evaluated for each output. Those outputs with the

largest row sums are chosen for control. Geometrically, the row sum is the norm of the projection of each output axis ($\mathbf{y}_i$) onto the manipulated variable space:

$$rs_i = \left\| \mathbf{K}(\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{y}_i \right\|^2 \tag{5.9}$$

where K is the steady state gain matrix. The proof of equation (5.9) can be found in Appendix 5.2.

The larger the row sum, the closer the associated output is to the manipulated variable space. Therefore, the row sum is a method for selecting those outputs that are best aligned with the manipulated variable space. Selecting outputs based on the row sum should provide a subsystem that results in a low overall error.

The implicit assumption in calculating the NSRGA and the row sum is that the manipulated variables have already been selected, and a square system is desired. Therefore, the object is to chose as many controlled variables as there are manipulated variables. No insight into the 'true' process dimension is given. Disturbances are not considered (it is assumed that the primary interest is setpoint changes).

Clearly, the concept of finding outputs based on their alignment with the manipulated variable space is very similar to that of Selective Principal Components Analysis. However, the row sum is very much a 'univariate' tool rather than a multivariate tool like SPCA. To illustrate this point, consider the MWD example from Chapter 3. Recall there are 100 MWD (output) variables and 3 input variables (temperature, monomer charge and initiator charge). Using the gain matrix identified between the full MWD (100 outputs) and the three inputs, the row sums can be calculated. Figure 5.1 shows the row sum for each of the outputs.

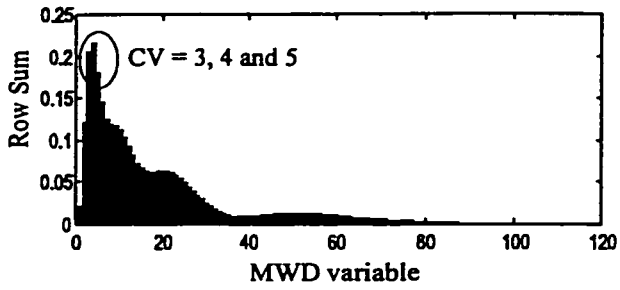

Figure 5.1. Selection of controlled variables using the Row Sum

For this set of manipulated variables, the row sum indicates that variables 3, 4 and 5 should be controlled (these are the outputs with the largest associated row sums). Since variables 3, 4 and 5 are three adjunct points on the MWD, the row sum approach has essentially selected the same variable three times. In reality, perfectly controlling variable 3 would likely result in near perfect control of variables 4 and 5 as well. From a practical standpoint, the result is that the subsystem with variables 3, 4 and 5 as the outputs would be very ill-conditioned. This type of result was also pointed out in Chang and Yu (1990), although no explanation was given.

Compare this result to that from SPCA, which also selects a set of the outputs, but is designed to provide complementary outputs. MWD data is generated in which there are only manipulated variable perturbations (this data also allows a more appropriate comparison to the NSRGA, which uses only manipulated variable information in the selection of controlled variables).

An SPCA analysis indicates that output 16 should be chosen as the first controlled variable, followed by output 41. Note that by perfectly controlling 16, many of the surrounding locations on the MWD will also be partially controlled, therefore it would make little sense for variables 17 or 15 to also be directly controlled. In fact, an analysis of the SPCA residuals (after variable 16 is removed, or, perfectly controlled) shows that over 99% of the variation is also removed from variables 17 and 15. Variable 41 is far from location 16 and is the next best variable to control given that variable 16 has already been chosen.

While the NSRGA is a fast method for picking outputs based on their alignment with the manipulated variable space, it is clearly a univariate tool, in that it selects

controlled variables without considering the controlled variables already selected. This can lead to ill-conditioned systems.

The NSRGA could be modified so that the choice of controlled variables is reflected in future controlled variable decisions. Assume y_i is chosen as the first controlled variable. The gain matrix can be deflated in a manner similar to SPCA. The notation is:

$$\mathbf{K} = \begin{bmatrix} \mathbf{k}_1^T \\ \mathbf{k}_2^T \\ \vdots \\ \mathbf{k}_n^T \end{bmatrix} \tag{5.10}$$

Each *row* in the gain matrix ($\mathbf{k}_j^T$) is regressed on the i^{th} *row* of the gain matrix ($\mathbf{k}_i^T$). Then, the residuals are taken, leaving only that portion of $\mathbf{k}_j^T$ that is perpendicular (uncorrelated) to $\mathbf{k}_i^T$:

$$\mathbf{K}_{new} = \mathbf{K} - \begin{bmatrix} \frac{\mathbf{k}_1^T\mathbf{k}_1}{\mathbf{k}_i^T\mathbf{k}_i} & \frac{\mathbf{k}_i^T\mathbf{k}_2}{\mathbf{k}_i^T\mathbf{k}_i} & \cdots & \frac{\mathbf{k}_i^T\mathbf{k}_n}{\mathbf{k}_i^T\mathbf{k}_i} \end{bmatrix}^T \mathbf{k}_i^T \tag{5.11}$$

Each element in the i^{th} row of $\mathbf{K}_{new}$ is zero, and the portion of the other rows co-linear with the first row of $\mathbf{K}_{new}$ is subtracted to reflect the effect of perfectly controlling y_1. The row sum can then be applied to $\mathbf{K}_{new}$ to select the next controlled variable. The deflation procedure is repeated after each selected variable. Using this modified procedure, which now explicitly considers correlation between output variables, the row sum indicates that output 4, followed by 9 and 20 should be chosen as the controlled variables.

5.4.2.2 The Disturbance Inflation Factor

The Disturbance Inflation Factor (seen earlier in Chapter 3) was derived to address the problem of MWD control, although it is generally applicable to all reduced

dimension control problems. The DIF evaluates whether controlling a subspace of the output space will have a positive effect on the overall output space (disturbance is attenuated in the uncontrolled space) or negative effect on the overall output space (disturbance is inflated in the uncontrolled space). It is calculated for a given set of controlled and manipulated variables, for a particular disturbance structure. The numbers are straightforward to interpret. If the DIF is less than one, the control actions attenuate the disturbance in the overall output space; if the DIF is greater than one, the control actions inflate the disturbance; if the DIF=1, the control actions simply transfer the disturbance from the controlled variables to the uncontrolled output variables. A DIF of 0 means the disturbance is completely eliminated in all outputs while explicitly controlling the given subsystem. A low DIF is preferable.

For the MWD example in the Chapter 3, the DIF was used to evaluate which of the three manipulated variables were best for controlling the average of the distribution, in terms of the overall effect on the MWD. Specific linear combinations of the MWD variables were also tested as potential control variables, to evaluate the potential for improvement. Furthermore, ratios of the manipulated variables were also proposed and tested using the DIF.

Unlike the NSRGA, the DIF considers disturbances. However, since the DIF analysis considers specific disturbances, all representative disturbances should be included in the analysis. Furthermore, because the DIF must be recalculated for each subsystem, a DIF analysis is ideal only for those problems in which there is a limited number of potential control structures (less than 10 were evaluated for the MWD problem). The DIF analysis would also be particularly useful as a post-screening tool. Promising subsets of variables could be chosen based on any of the methods outlined earlier, and then the performance of the various subsystems for expected disturbance structure evaluated using the DIF.

5.4.3 Partial Control Methodology

The situation addressed in partial control is consistent with the reduced dimension control problem outlined earlier: there are many outputs and few inputs, and the control structure must be determined (Arbel et al (1996)). Partial control has been discussed in many different papers such as Arbel et al (1996), Havre and Skogestad (1996), Zhao and Skogestad (1997), Kothare et al (1998), Tyreus (1999). The concept of partial control will be explained in the context of the Fluidized Catalytic Cracking Unit (FCCU) example of Arbel et al (1996). Then, two other examples of partial control will be briefly discussed. Emphasis will be on the dimensionality aspects of partial control.

Partial control is an approach rather than a tool or algorithm. In partial control, the control structure consists of primary and secondary controllers. The primary loop acts on process variables such as temperatures or pressures, and operates quickly. The secondary loop acts on quality variables and adjusts more slowly. The setpoints for the primary loop are set by the secondary loop. Ideally, one wishes to choose 'dominant' outputs for the primary loop. Controlling a 'dominant' output in the primary loop results in partial control of many of the other (uncontrolled) outputs, such as those in the secondary loop. In this manner, many of the product specifications or constraints are met without directly controlling those output variables.

In the FCCU process addressed by Arbel et al (1996), well over fifteen different output variables were available, including quality variables such as conversion and yield and operating variables such as process temperatures. Operating variables are typically considered for the primary (fast) loop, while quality variables are relegated to the secondary (slow) loop. At least seven inputs could be adjusted in the FCCU; four could be adjusted very quickly (for example catalyst flowrate) and three could be adjusted infrequently (for example catalyst type). Arbel et al (1996) address the problem of selecting appropriate operating variables (inputs and outputs) for the primary loop.

In order to maintain many of the outputs within process specifications, a few key process variables will be controlled. Based on current practice in industry and literature, a

limited number of potential control structures (sets of controlled and manipulated variables) was proposed. Through an exhaustive analysis (e.g. time scale of responses, controller stability, input multiplicities, effect on uncontrolled variables when a disturbance occurs, non-linearities), a single control structure (in which two process temperatures were controlled, and two flowrates adjusted) was ultimately chosen for the FCCU.

The key to partial control is the identification and control of dominant outputs. If one controls a dominant output, many of the remaining outputs will be partially controlled. The dominance of an output is clearly related to process dimensions and directions and is determined by the alignment of the output, the disturbance and the manipulated variable spaces. Many outputs are strongly correlated to a dominant process variable, thus directly controlling the dominant output partially controls the others.

In Arbel et al (1996), no tools were provided to help identify potential dominant variables, instead they relied on process knowledge and experience. Recent publications by Tyreus (1999a, 1999b) provide insight into dominance by applying energy balances to uncover sets of potential dominant outputs, particularly when process knowledge is lacking. The approach is illustrated on the Tennessee Eastman plant. In Havre and Skogestad (1996), a tool directly related to dimensionality and output dominance is derived. An ill-conditioned distillation column is considered for control, and the question posed is whether partial control (i.e. direct control of only one of the two compositions) will maintain both compositions within specifications in the face of certain disturbances. A tool called the Partial Disturbance Gain (PDG) is introduced to aid in the selection; as discussed in an Appendix to Chapter 3, the Partial Disturbance Gain and the Disturbance Inflation Factor are closely related. The PDG was used to evaluate which composition was most 'dominant' and therefore should be controlled.

To summarize, the ideas of partial control are closely related to those of reduced dimension control. In fact, output dominance, as defined in partial control, is a direct

result of high correlation among the process output variables, which arises when the true process dimension is less than the size of the measured space.

5.4.4 Modal Methods

Modal analysis was originally proposed as a multivariable controller design method for processes in which there are many more state variables than inputs (Rosenbrock, 1962). It is based on state space representation of the process:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \tag{5.12}$$

A general overview of modal methods and their relationship to reduced dimension control will be given here. For more details the reader is referred to a modal review paper by Bonvin and Mellichamp (1982).

Modal analysis involves the eigen-decomposition of the dynamic matrix (A) in the state space model to give new linear combinations of the states (called modes). The modes are ordered according to their open loop speed, which is given by the associated eigenvalue. Control is applied only to the slowest modes in order to speed up the closed loop process response.

The main difference between modal methods and many of the other methods discussed above is that modal analysis is primarily concerned with the dynamics of the closed loop response. It requires a full dynamic state space model (which would not typically be available in quality control situations) and does not specifically address the reduced rank nature of the input/output space when selecting the modes for control. Selection of modes (linear combinations) for control is based on the dynamic response of the system.

The reduced rank nature of the process could be incorporated into the modal method approach by using subspace identification techniques in the identification of the

state space model to be analyzed. Viberg (1995) and Ljung and McKelvey (1996) provide good overviews of the subspace based identification methods. By applying such methods in the identification of the state space representation of the process, the reduced rank nature of the problem is captured prior to the modal analysis. The standard modal analysis could then be applied to uncover the slowest modes and to further reduce the dimension of the system using dynamic considerations.

5.5 Discussion

In this Chapter, many different methods and approaches for reducing the dimension of the output space have been discussed. Clearly, many of them have the same objectives, and one may be tempted to ask 'which method is best'? The best method depends on the specific problem, and certainly no method comprehensively addresses all the problems. Table 5.1 summarizes the information from this chapter, outlining the methods, the ideal problem for each method and the associated warnings.

Despite the individual methodologies available, a combined approach is probably the most thorough and effective manner of approaching an RDC problem. First, the dimensionality of the process should be evaluated to determine if an RDC situation exists. Does the observed plant variability fall in a space that is much smaller than the measured output/quality space? Has correlation been observed among many of the quality variables? The data-based methods are very effective, and efficient, for evaluating dimensionality on various types of plant data and identifying immediately whether a reduction is possible.

It is also important to establish the type of plant. Most RDC tools either explicitly or implicitly make the assumption of equal deadtimes and dynamics, or infrequent measurements. Are the dynamics and deadtimes different? While a Reduced Dimension Controller may still perform well when the plant has non ideal dynamics (as was observed with the digester in Chapter 4) this type of situation should still be noted. Is the process highly non-linear? The structure of the RDC may change at different operating

regions or the plant may be very nonlinear, and this must be addressed in the reduction. Is the process ill-conditioned? Ill-conditioned plants exhibit robustness problems with full model-based multivariable controllers, therefore RDC has shown a lot of potential for these types of plants. Does the disturbance structure change frequently? If yes, one may wish to consider methods that do not explicitly incorporate disturbances. Are certain measurements less reliable or difficult to obtain regularly? One may wish to avoid controlling linear combinations of the outputs if this is true.

Table 5.1. Summary of RDC Approaches

Method	**Ideal Problem**	**Warning**
Non Square RGA	Selection of a group of outputs for control assuming the manipulated variables are pre-specified	Outputs are not complementary; could end up with ill conditioned subsystem
PCA/SPCA	Characterization of process spaces Selection of complementary set of controlled variables (linear combinations or whole variables)	Data must contain input information
CCA	Selecting the controllable disturbance subspace	Must have open loop or disturbance data
DIF/PDG	Evaluation of a small number of potential controlled subsystems	Must include all representative disturbances in analysis
SVD and SVD-PCA	Input/output dimension reduction for ill-conditioned processes and if the disturbance dimension is low	Must measure all outputs if linear combinations are controlled
Partial Control	Selection of a small number of inputs and outputs when strong correlation exists between process variables and process/product specifications	Limited tools available
Modal Analysis	Dimension reduction using dynamic considerations	Does not explicitly consider reduced dimension spaces

Depending on the data available, it may be possible to characterize the disturbance and manipulated variable spaces independently using PCA directly from the data.

Otherwise, an SVD of the process gain matrix (if available) provides a basis for the manipulated variable space. If one can separate the spaces by removing known manipulated variable adjustments from the data, useful process insight is obtained by determining if the disturbance and manipulated variable spaces overlap. If a portion of the disturbance space lays outside the manipulated variable space, this immediately signals a process limitation with respect to the ability to eliminate all the disturbances. It is also of interest to note whether the disturbance space is much smaller than the manipulated variable space; this indicates increased potential for reduction. If disturbance data is available, PCA and CCA can be used to evaluate the variability and predictability and how much of the disturbance is predictable and therefore controllable.

Once a preliminary evaluation of the process has been completed, a quantitative dimension reduction can begin. Plant test data (data with manipulated variable information) is required regardless of which method is used. It may exist from past plant tests, or new plant tests will need to be run. Once the data is available, there are many options. One can apply PCA or SPCA directly to the data to select linear combinations or whole variables for control. Or, one can use the data to identify a process model, and then apply one of the other methods, such as SVD or the Non Square Relative Gain Array. A prudent approach would be to generate several alternative control structures using the different methods and compare the results to see if they are consistent. Prior to settling on one structure, it would be useful to analyze each potential candidate using the Disturbance Inflation Factor or Partial Disturbance Gain for specific disturbances, so that the performance in the overall quality space is confirmed.

5.6 Concluding Remarks

In this Chapter, a unification of many of the Reduced Dimension Control methods was presented. The approaches were characterized as either data based (for example, PCA, SPCA and CCA) or model based (for example, SVD, DIF, NSRGA and Modal Analysis) for the discussion. Each of the general methods were discussed and then

illustrated with specific published examples. Where appropriate, the methods and examples were placed within the context of the RDC framework from Chapter 4. The examples, and their relationship to the framework, were chosen in part to highlight the issues or shortcomings of each of the approaches. For example, it was seen that the data used in the data-based methods must contain manipulated variable information. It was also illustrated that the model based method of the NSRGA did not chose complementary output variables for control, and as a result an ill-conditioned subsystem could result.

In general, all methods for reducing the dimension of the controlled variable space make the same dynamic assumptions as the framework, but this was rarely stated explicitly in any of the publications. Many of the methods implicitly assumed setpoint changes for the disturbance space; only a minority of methods explicitly consider the disturbances.

The Chapter concluded with a discussion on an overall RDC approach. It was discussed that a combined approach is the most effective and comprehensive way of addressing dimension problems within a process. It was suggested that the data based methods be used first to assess the process dimensionality, after which the model based methods could be used to propose potential control structures. Specific analysis tools such as the Disturbance Inflation Factor or Partial Disturbance Gain could then be used to evaluate performance for expected disturbances.

5.7 Nomenclature

The following nomenclature was used in this Chapter:

$\mathbf{C}_{optimal}$: optimal controlled variable space

$\mathbf{p}_d$: disturbance direction

K: gain matrix

$\mathbf{M}_{optimal}$: optimal input directions

W: left singular matrix from SVD

S: diagonal matrix from SVD

V: right singular matrix from SVD

P: PCA loading matrix

$\mathbf{y}_{cv}$: latent controlled variable

$\mathbf{u}_{mv}$: latent manipulated variable

y: vector of outputs

u: vectors of inputs

G(s): process transfer function gain matrix

rs: row sum

Appendix 5.1. Relationship between SVD-PCA approach and the Optimal Framework

The notation of Arkun and Kayihan (1998) is used in this Appendix. The first step in the method of Arkun and Kayihan (1998) is an SVD of the process steady state gain matrix ($\mathbf{K}=\mathbf{U}\Sigma\mathbf{V}^T$). The columns of **U** are independent orthonormal vectors that define the manipulated variable space, in the full output space. The size of **U** for the paper making process is 125 x 75.

Second, a dimensionality analysis of the disturbance data is performed. The disturbance varies in two dimensions, represented here by the matrix $\mathbf{P}_d$ (= $[\mathbf{p}_{d1}\ \mathbf{p}_{d2}]$), where the size of $\mathbf{P}_d$ is 125x2. Open loop disturbance data transformed:

$$\mathbf{P}_{d,trans} = \mathbf{U}^T\mathbf{P}_d \qquad \text{(B.1)}$$

The size of $\mathbf{P}_{d,trans}$ is 75 x 2. Next, the authors define a matrix Φ, which is an orthonormal basis for $\mathbf{U}^T\mathbf{P}_d$ (size of Φ is 75 x 2). In practice, it is obtained by doing a principal components analysis of the transformed disturbance data. In terms of mathematics, Φ is obtained from an SVD on $\mathbf{U}^T\mathbf{P}_d$:

$$\begin{aligned}\mathbf{U}^T\mathbf{P}_d &= \Phi\Sigma\Psi^T \\ \Phi &= \mathbf{U}^T\mathbf{P}_d(\Sigma\Psi^T)^{-1}\end{aligned} \qquad \text{(B.2)}$$

From the controller description, the outputs that are controlled are given by the following expression:

$$\begin{aligned}\mathbf{y}_{cv} &= (\Phi^T\Phi)^{-1}\Phi^T\mathbf{U}^T\mathbf{y} \\ &= (\mathbf{U}\Phi)^T\mathbf{y}\end{aligned} \qquad \text{(B.3)}$$

where **y** is the 125x1 full measurement vector

Therefore the controlled variable space is defined by:

$$\begin{aligned} \mathbf{C}_{pca} &= \mathbf{U}\Phi \\ &= \mathbf{U}\mathbf{U}^T\mathbf{P}_d(\Sigma\Psi^T)^{-1} \\ &= \mathbf{C}_{optimal}(\Sigma\Psi^T)^{-1} \end{aligned} \qquad \text{(B.4)}$$

where the size of the matrix $\Sigma\Psi^T$ is 2x2

Therefore, the controlled variable space (spanned by the columns of $\mathbf{C}_{pca}$) in this method is simply the optimal controlled variable space ($\mathbf{C}_{optimal}$), rotated. That is, the controlled variable spaces defined by $\mathbf{C}_{pca}$ and $\mathbf{C}_{optimal}$ are equivalent. The corresponding input directions are calculated by inverting the process. A similar analysis to that above can be used to show that the method's input directions are equivalent to the optimal input directions given by (5.2).

Appendix 5.2. Relationship of the row sum to the manipulated variable space

It will be shown in this Appendix that the row sum is the magnitude of the projection of each output vector onto the manipulated variable gain space. The expression for the projection of an arbitrary output, defined by its vector axis $\mathbf{y}_i$, is:

$$\mathbf{y}_{proj} = \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{y}_i \qquad \text{(C.1)}$$

The norm of the projection is:

$$\begin{aligned} \left\|\mathbf{y}_{proj}\right\|^2 &= (\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{y}_i)^T(\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{y}_i) \\ &= \mathbf{y}_i^T\mathbf{G}\left((\mathbf{G}^T\mathbf{G})^{-1}\right)^T\mathbf{G}^T\mathbf{y}_i \qquad \text{(C.2)} \\ &= \mathbf{y}_i^T\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{y}_i \end{aligned}$$

The pseudo-inverse of the gain matrix, $\mathbf{G}^+$, is defined as:

$$\mathbf{G}^+ = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{y}_i \qquad \text{(C.3)}$$

Therefore,

$$\begin{aligned} \left\|\mathbf{y}_{i,proj}\right\|^2 &= \mathbf{y}_i^T\mathbf{G}\mathbf{G}^+\mathbf{y}_i \\ &\mathbf{y}_i^T(\text{i}^{\text{th}}\text{ column of }\mathbf{G}\mathbf{G}^+) \qquad \text{(C.4)} \\ &= \text{ii}^{\text{th}}\text{ element of }\mathbf{G}\mathbf{G}^+ \end{aligned}$$

In the Appendix of Chang and Yu (1990), it is shown that $rs_i = \text{ii}^{\text{th}}$ element of $\mathbf{G}\mathbf{G}^+$. Therefore, the row sum for a given output is equal to the norm of the projection of that output's vector onto the manipulated variable gain space. This result was also confirmed

by calculating the norm for the example processes and comparing it to the reported row sums in Chang and Yu (1990).

6. Summary and Conclusions

In this thesis, several issues associated with the control of product quality have been addressed. First, the problem of incorporating an approximate fundamental model into an online feedback quality control scheme was addressed. The second problem investigated in the thesis related to the dimensionality issues in product quality control. The issue of controlled and manipulated variable selection while minimizing the overall variability in the product quality was addressed. The major conclusions and the main contributions of this thesis are outlined in this chapter.

In the second chapter, a new batch-to-batch optimization methodology for producing a desired MWD has been presented. The advantages of the proposed batch-to-batch optimization methodology are its simplicity and flexibility. Fundamental polymerization knowledge was used to simplify the optimization problem, thereby avoiding the need for a complex on-line numerical optimization routine and issues relating to its convergence. The approach showed that for linear polymers, any desired MWD can be approximated by a combination of a few constituent distributions. Each of these component distributions can be produced by controlling one (or possibly two) parameters (τ,β) representing the ratios of kinetic rates. The method allowed for great flexibility in choosing the variables to manipulate (batch temperature, monomer or initiator flowrates) since any combination of them can be used to achieve a desired value of τ or β. These decisions can therefore be made with the specific process and its constraints in mind. Furthermore, this method has many desirable features when compared to other batch-to-batch optimization or MWD control applications: only an approximate fundamental model is required; historical plant test data is not needed and on-line measurements of the full MWD are not required. The methodology is shown,

through simulated case studies, to efficiently return the process to the desired MWD within a few batches for a wide range of process/model mismatch.

The optimization approach was then extended for use as an on-line control method. To achieve this, it was necessary to address the modeling of highly correlated measurement error typically observed in the MWD measurements. Principal Components Analysis was used to analyze replicate GPC data and formulate a MWD noise model consistent with the structure found in the real MWD data. A multivariable statistical process control (MSPC) monitoring scheme was then applied for deciding when a new batch optimization was required. The optimizer remained on, but dormant while the desired MWD was being produced, and re-optimized the process quickly if the process changed and poor quality polymer was produced. The combined MSPC/batch-to-batch optimizer was demonstrated on a simulated semi-batch polystyrene reactor. It was seen that the optimizer was able to distinguish between measurement error and process upsets, and corrections were made only in response to process upsets. The optimizer responded quickly to process upsets and was able to re-optimize the process within several batches, despite process/model mismatch and measurement error.

The contributions from Chapter 2 are as follows. First, a new batch-to-batch optimization methodology for obtaining manipulated variable trajectories for producing a desired MWD in batch/semi-batch reactors was developed. Second, a method for modeling the correlated multivariable noise typically found on real MWD measurements was presented. Finally, the batch-to-batch optimization methodology was combined with a multivariable statistical method in order to handle noisy MWD measurements.

In the third chapter, indirect control of the full MWD by directly controlling the weight-average chain length of the distribution was investigated. When the full MWD was considered, it was shown that the choice of manipulated variable was critical when controlling the average. For the polystyrene semi-batch reactor, only one of the three potential manipulated variables actually improved the full MWD while directly controlling the average. A new steady state analysis tool, the Disturbance Inflation Factor

was derived to quantify the observed results. To calculate the DIF, the dominant disturbance directions and process gain matrix in the full output space was needed.

The DIF was then applied to quickly and effectively predict the impact, on the full MWD, of a large number of potential control options. It was seen that with prudent choices of manipulated and controlled variables, simple single variable controllers can provide significant improvement in the full MWD. Ratio and latent variable control were two options suggested. Ratio control (in which a ratio of two manipulated variables was used to control the molecular weight average) was seen to be a feasible option for the polystyrene system in which one disturbance perturbed the process. The ratio was chosen so that the gain vector of the new manipulated variable was aligned with the disturbance direction. Excellent disturbance attenuation in the full MWD was achieved. When the disturbance was characterized by two or more directions, the concept of best and worst disturbance directions was introduced. Single variable latent variable control (in which a linear combination of the MWD variables is controlled) was successful for the polystyrene process with two disturbances. Also investigated was a multivariable controller in which both number- and weight- averages are controlled. The multivariable controller provided an effective and easily implemented control strategy for implicit control of the full MWD, when more than one disturbance direction was observed.

The contributions from Chapter 3 are as follows. First, a linear analysis tool for controlled and manipulated variable selection, the Disturbance Inflation Factor, was derived. Second, a methodology, whereby the disturbance and process directions are used for predicting the impact on the full MWD when controlling an average or polydispersity was presented. The method can be applied to any process with many quality variables (not just the MWD) for selecting effective controlled and manipulated variables. Finally, the application of Principal Components Analysis (PCA) for the characterization of dominant disturbance directions was presented.

In Chapter 4, the issues related to dimension and control were generalized. A framework for reducing the dimension of the controlled subsystem was derived based on

minimum variance control theory. Two cases were considered. First, expressions for the linear combinations of the inputs and outputs that should be manipulated and controlled in order to minimize the overall error in the output space was derived. Second, an expression for the linear combination of the inputs that should be manipulated to minimize the overall error in the output space when the controlled variable is fixed was derived. The role of the number of independent disturbances in determining the number of controlled and manipulated variables, and the structure of the resulting reduced dimension controllers was clearly illustrated in the framework.

It was shown that several assumptions were made in the derivation of the optimal directions for control. These were: the process is sampled infrequently relative to the process dynamics, or equivalently, the process dynamics from all inputs to all outputs are the same; there are equal deadtimes from all inputs to all outputs; and the process is linear and non-time varying. For many product quality control situations in which disturbance regulation is the primary goal, these assumptions are approximately true. However, there are processes for which these assumptions are violated. In Chapter 4, the RDC approach (based on the framework) was applied to a non-ideal dynamic digester to evaluate the effect of violating the RDC assumptions. Two single variable Reduced Dimension Controllers, a latent variable controller (that adjusted a ratio of the manipulated variables) and a whole variable controller (that adjusted a single manipulated variable), were proposed. The RDCs were compared to a DMC that adjusted all inputs to control all outputs. At the operating point for which the RDCs were designed, control performance was very good despite the dynamic non-idealities of the process. Even though its structure was much simpler, the latent variable controller performance was comparable to that of the DMC in all outputs but one. When the operating point and process directions changed, the RDCs showed only modest degradation in controller performance (primarily in one output). In general, one would expect the performance of the RDCs to be sensitive to changes in disturbance direction (although this was not an issue for the digester).

The contributions from Chapter 4 are as follows. First, a new framework for selecting the optimal directions to control when the overall quality is important was derived. The framework provided important insight into how the process and disturbance directions impact the controlled subsystem structure. Second, the framework was applied to a dynamic process in order to evaluate the feasibility of RDC under non-ideal circumstances.

In Chapter 5, a unification of existing RDC methods was addressed. The approaches were characterized as either data based (for example, PCA, SPCA and CCA) or model based (for example, SVD, DIF, NSRGA and Modal Analysis). Each of the general methods were discussed then illustrated with specific published examples. Where appropriate, the methods and examples were placed within the context of the RDC framework from Chapter 4. The examples, and their relationship to the framework, were chosen in part to highlight the issues or shortcomings of each of the approaches. For example, it was seen that the data used in the data-based methods must contain manipulated variable information. It was also illustrated that the model based method of the NSRGA did not chose complementary output variables for control, and as a result an ill-conditioned subsystem could result.

In general, all methods for reducing the dimension of the controlled variable space made the same dynamic assumptions as the framework, but this was rarely stated explicitly in any of the publications. It was also seen that many of the methods implicitly assumed setpoint changes for the disturbance space; only a minority of methods explicitly consider the disturbances. The Chapter concluded with a discussion on an overall RDC approach. It was discussed that a combined approach, using many of the methods in sequence or simultaneously, was the most effective and comprehensive way of addressing dimension problems within a process.

The primary contribution of Chapter 5 is insight into the area of reduced dimension control. Because of the different backgrounds of the existing applications, the common issues are often overlooked and therefore new applications are rarely placed

within the context of other methods. The result is a number of apparently different methodologies for selecting subsystems of controlled and manipulated variables. In order to increase awareness and understanding in this promising field, the contribution of this chapter was to provide a unifying view of the existing methodologies and applications.

7. Literature Cited

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

Arkun, Y; Kayihan, F. A Novel Approach to Full CD Profile Control of Sheet-Forming Processes using Adaptive PCA and Reduced-Order IMC Design, *Comp. Chem. Eng.*, **1998**, *7-8*, 945.

Box, G.E.P.; G.C. Tiao, G.C.; A Canonical Analysis of Multiple Time Series, *Biometrika*, **1977**, *64*, 355.

Box, G.E.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control;* Robinson, E., Ed.; Holden-Day: Oakland, 1976.

Bonvin, D.; Mellichamp, D.A.; A Unified Derivation and Critical Review of Modal Approaches to Model Reduction, *Int. J. Contr.*, **1982**, *35*, 829.

Cawthon, G.D.; Knaebel, K.S.; Optimization of Semi-batch Polymerization Reactions, *Comp. Chem. Eng.*, **1989**, *13*, 63.

Chang, J.; Liao, P.; Molecular Weight Control of a Batch Polymerization Reactor: Experimental Study, *Ind. Eng. Chem. Res.*, **1999**, 38, 144.

Chang, J.; Lai, J.; Computation of Optimal Temperature Policy for Molecular Weight Control in a Batch Polymerization Reactor, *Ind. Eng. Chem. Res.*, **1992**, *31*, 861.

Chang, J.; Yu, C., The Relative Gain for Non-Square Multivariable Systems, *Chem. Eng. Sci.*, **1990**, *45*, 1309.

Choi, K.Y.; Butala, D.N.; An Experimental Study of Multiobjective Dynamic Optimization of a Semibatch Copolymerization Process, *Poly. Eng. And Sci.*, **1991**, *31*, 353.

Clarke-Pringle, T.; MacGregor, J.F.; Nonlinear Adaptive Temperature Control of

Multiproduct, Semi-batch Polymerization Reactors, *Comp. Chem. Eng.*, **1997**, *21*, 1395.

Clarke-Pringle, T.L., MacGregor, J.F.; Product Quality Control in Reduced Dimensional Spaces, *Ind. Eng. Chem. Res.*, **1998**, *37*, 3992.

Clarke-Pringle, T.L, MacGregor, J.F.; Optimization of Molecular-Weight Distribution Using Batch-to-Batch Adjustments, *Ind. Eng. Chem. Res.*, **1998**, *37*, 3660.

Christie, O.H.J, Data Laundering by Target Rotation in Chemistry Based Oil Exploration, *J. of Chemo.*, **1996**, *10*, 453.

Crowley, T.J; Choi, K.Y.; Discrete Optimal Control of Molecular Weight Distribution in a Batch Free Radical Polymerization Process, *Ind. Eng. Chem. Res.*,**1997**, *36*, 3676.

Dube, M.A.; Penlidis, A.; Hierarchical Data Analysis of a Replicate Experiment in Emulsion Terpolymerization, *AIChE J*, **1996**, *42*, 1985.

Dong, D.; McAvoy, T.; Zafiriou, E.; Batch to Batch Optimization using Neural Network Models, *Ind. Eng. Chem. Res.*, **1996**, *35*, 2269.

Eek, R.A.; Bosgra, O.H., Controllability of Particulate Processes in Relation to the Sensor Characteristics, *Proc. Control of Particulate Processes IV*, Kananaskis, Alberta, Canada, May **1995a**, Prof. H. Henein (Ed).

Eek, R.A.; Bosgra, O.H., Multivariable Control of a Continuous Crystallization Process Using Experimental Input-Output Models, *Dycord +95*, **1995b**, June 7-9, Copenhagen, p 305.

Eek, R.A., Hoogenboezem, A.J.; Bosgra, O.H.; Design Issues Related to the Control of Continuous Crystallizers, *Comp. Chem. Eng.*, **1996**, *20*, 427.

Eek,R.A.; Pouw, H.A.A.; Bosgra, O.H.; Design and Experimental Evaluation of Stabilizing Feedback Controllers for Continuous Crystallizers, **Powder Tech.**, **1995**, *82*, 21.

Ellis, M.F.; Taylor, T.W; Jensen, K.F; On-line Molecular Weight Distribution Estimation and Control in Batch Polymerization, *AIChE J.*, **1994**, *40*, 445.

Featherstone, A.P.; Braatz, R.D.; Integrated Robust Identification and Control of Large-Scale Processes, *Ind. Eng. Chem. Res.*, **1998**, *37*, 97.

Featherstone, A.P; Braatz, R.D. Control-Oriented Modeling of Sheet and Film Processes, *AIChE J.*, **1997**, *43*, 1989.

Filippi-Bossy, C.; Bordet, J.; Villermaux, J.; Marchal-Brassely, S.; Georgakis, C.; Batch Reactor Optimization by Use of Tendency Models, *Comp. Chem. Eng.*, **1989**, *13*, 35.

Fotopoulos, J.; Georgakis, C.; Stenger, H.G; Effect of Process-Model Mismatch on the Optimization of the Catalytic Epoxidation of Oleic Acid Using Tendency Models, *Chem. Eng. Sci.*, **1996**, *51*, 1899.

Golub, G.H.; Van Loan, C.F. *Matrix Computations*, 2nd Ed, John Hopkins U. Press (1989).

Gugliotta, L.; Leiza J.; Arotcarena, M.; Armitage, P.; Asua, J.; Copolymer Composition Control in Unseeded Emulsion Polymerization Using Calorimetric Data, *Ind. and Eng. Chem. Res.*, **1995**, *34*, 3899.

Hamielec, A.E.; Tobita, H; *Ullmann's Encyclopedia of Industrial Chemistry;* VCH Publishers Inc.: New York, *A21*, 324, 1992.

Havre, K.; Skogestad, S. Input/Output Selection and Partial Control, *Preprints IFAC '96, 13th World Congress of IFAC*, **1996**, Permagon Press: San Francisco, CA, *Vol M*, 181.

Hicks, J.; Mohan, A.; Ray, W.H. The Optimal Control of Polymerization Reactors, *CJChE*, **1969**, *47*, 590.

Hoffman, R.F; Schreiber, S.; Rosen G. Batch Polymerization - Narrowing the Molecular Weight Distribution, *Ind. Eng. Chem. Res.*, **1964**, *56*, 51.

Houston, W.E.; Schork, F.J. Adaptive Predictive Control of a Semibatch Polymerization Reactor, *Poly. Proc. Eng.*, **1987**, *5*, 119.

Hovd, M.; Braatz, R.D.; Skogestad, S.; SVD Controllers for H2-, H∞, and μ optimal Control, *Automatica*, **1997**, *33*, 433.

Jackson, J.E.; Mudholkar. G.S; Control Procedures for Residuals Associated with Principal Component Analysis, *Technometrics*, **1979**, *21*, 341.

Jutan, A., MacGregor, J.F., Wright, J.D., 'Multivariable Computer Control of a Butane Hydrogenolysis, Reactor: Part II. Data Collection, Parameter Estimation and Stochastic Disturbance Identification, *AiChE J*, **1984**, *30*, 220

Kayihan, F. Kappa Number Profile Control for Continuous Digesters, In Nonlinear Model Based Process Control, Berber, R. and Kravaris, C. eds,. NATO ASI Series 1997, E Vol 353, 805

Kayihan, F.; Continuous Digester Benchmark Model Toolbox v1.0, Integrated Engineering Technologies, 1998.

Keller, J.P.; Bonvin, D.; Selection of Inputs for the Purpose of Model Reduction and Controller Design, *Proc IFAC 10th Triennial World Congress*, **1987**, Munich, 209.

Kothare, M.V., Shinnar, R., Rinard, I., Morari, M., On defining the Partial Control Problem: Concepts and Examples, *Proc. ACC*, **1998**, Philadelphia, PA, 2103.

Kozub, D.; MacGregor, J.F. Feedback Control of Polymer Quality in Semi-batch Copolymerization Reactors, *Chem. Eng. Sci.*, **1992**, *47*, 929.

Lau, H.; Alvarez, J.; Jensen, K.F.; Synthesis of Control Structures by Singular Value Analysis: Dynamic Measures of Sensitivity and Interaction, *AIChE J*, **1985**, *31*, 427.

Ljung, L.; McKelvey, T.; Subspace Identification from Closed Loop Data, *Signal Processing*, **1996**, *52*, 209.

Luyben, W.L., Process Modeling, Simulation and Control for Chemical Engineers, McGraw-Hill, Inc., New York, **1989**.

MacGregor, J.F.; Kourti, T.; Statistical Process Control of Multivariate Processes, *Control Eng. Prac.*, **1995**, *3*, 403.

Marlin, T; Process Control: Designing Processes and Control Systems for Dynamic Performance, McGraw-Hill, Inc., New York, NY, **1995**.

MacGregor, J.F., Wong, A.K.L., Multivariate Model Identification and Stochastic Control of a Chemical Reactor, *Technometrics*, **1980**, *22*, 453.

Marten, F.L.; Hameliec, A.E; High-Conversion Diffusion-Controlled Polymerization of Styrene I, *J. of App. Poly. Sci*, **1982**, *27*, 489.

Moore, B.C.; Principal Components Analysis in Linear Systems: Controllability, Observability and Model Reduction, *IEEE Trans. Auto. Control.*, **1981**, *26*, 17.

Nomikos, P.; MacGregor, J.F.; Multivariate SPC charts for monitoring batch processes, *Technometics*, **1995**, *37*, 41.

Rigopoulos, A; Arkun, Y; Kayihan, F., Identification of Full Profile Disturbance Models for Sheet Forming Processes, *AIChE J,* **1997**, *43*, 727

Roffel, J.J.; MacGregor, J.F.; Hoffman, T.W., Design and Implementation of a Multivariable Internal Model Controller for a Continuous Polybutadiene Polymerization Train, *Dycord '89*, Maastricht, The Netherlands, **1989**, Control of Continuous Reactors I, 9.

Rosenbrock, H.H.; Distinctive Problems of Process Control, *Chem. Eng. Prog.*, **1962**, *58*, 43.

Smith, C.R., Moore, C.F., Bruns, D.D.; A Structural Framework for Multivariable Control Applications, JACC (1981)

Soroush, M.; Kravaris, C. Nonlinear Control of a Batch Polymerization Reactor: an Experimental Study, *AIChE J*, **1992**, *38*, 1429.

Takamatsu, T.; Shioya, S.; Okada, Y.; Molecular Weight Distribution Control in Batch Polymerization Reactor, *Ind. Eng. Chem. Res.*, **1988**, *27*, 93.

Tsen, A.Y.; Jang S.; Wong, D.; Joseph B. Predictive Control of Quality in Batch Polymerization Using Hybrid ANN Models, *AIChE J*, **1996**, *42*, 455.

Tyreus, B.; Dominant Variables for Partial Control. 1. A Thermodynamic Method for their Indentification; *Ind. Eng. Chem. Res.*, **1999**, *38*, 1432.

Tyreus, B.; Dominant Variables for Partial Control. 2. Application to the Tennesse Eastman Challenge Process; *Ind. Eng. Chem. Res.*, **1999**, *38*, 1444.

Vander Wiel, S.A.; Tucker, W.T.; Faltin, F.W.; Doganaksoy, N.; Algorithmic Statistical Process Control: Concepts and an Application, *Technometrics*, **1992**, *34*, 286.

Viberg, M.; Subspace-based Methods for the Identification of Linear Time-invariant Systems, *Automatica*, **1995**, *31*,1835.

Wold, S.; Esbensen, K; Geladi, P., Principal Component Analysis, *Chem. and Int. Lab. Sys.*, **1987**, *2*, 37.

Yabuki, Y.; MacGregor, J.F. Product Quality Control in Semi-batch Reactors using Mid-course Correction Policies, *Ind. Eng. Chem. Res*, **1997**, *36*, 1268.

Yoon, J.W.; Ryu, J.H; Cheong, C.; Choi, K.Y.; Calculation of Molecular Weight Distribution in a Batch Thermal Polymerization of Styrene, *Macro. Th. Sim.*, **1998**, 7, 327.

Zafiriou, E.; Zhu, J.; Optimal Control of Semi-batch Processes in the Presence of Modeling Error, Proceedings of 1990 American Control Conference, San Diego, CA, **1990**, 1644.

Zhao, Y.; Skogestad, S. Comparison of Various Control Configurations for Continuous Bioreactors, *Ind. Eng. Chem. Res.*, **1997**, *36*, 697.

Zhu, Z; Jutan, A.; Robust Multivariable Control Using an SVD-based controller, *Chem. Eng. Sci.*, **1998**, 53, 1145.