Modelling, Estimation, and Control of Sheet and Film Forming Processes

by

John Carl Campbell

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Chemical Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

1997

To Victoria, for her unconditional love.

Acknowledgments

Each Ph.D. student follows a unique path to achieve his or her academic goals. There are both challenges and setbacks, fortunately there are also loved ones, friends, and colleagues. This is a written thank you to the people who have walked, crossed, and, yes, even occasionally blocked my Ph.D. path.

As a senior at Illinois, I was struggling with some aspect of my senior research project and I needed the help of one of the graduate students. I approached Dan Klingenberg and was quite impressed with the way he attacked my question. I decided right then that I needed to learn more about learning and shortly afterwards I applied to graduate school. It is only fitting that Dan was able to serve on my final committee. Thank you Dan, I guess you never know whose life you will influence. In addition to Dr. Klingenberg, I would like to thank Professors Harmon Ray, Sang Kim, and Bob Barmish for their time and effort in evaluating my work.

I would like to thank my advisor, James B. Rawlings, for, among other things, being my advisor. I felt like I was constantly leaning from you Jim, whether it was during a class, a group meeting, or a personal research discussion. In fact, I recall several occasions when I was teaching you something and I still managed to learn something myself. In addition, I appreciated and enjoyed the opportunity to get to know you better during our many group activities.

The bulk of my research time was spent with, what is called in some circles, the Rawlings' group. The group changes in character over time, but it seems to always build strong friendships. First I'll recall John Eaton, Steve Miller, Tony Mullins, Scott Meadows, and Ken Muske. This group really defined the Texas Rawlings' group for me. I remember Steve as the master storyteller. I also recall that he was almost always too busy to go have a beer during his last few months of graduate school. Well Steve, I understand now (see Daniel Patience below). I have always considered Tony the wise sage of the group, but that doesn't mean I always believed him. Special thanks to John for not only managing a fine computing environment but also for the many educational discussions about computers and computing. Ken had the most profound effect on my research. A fact that I may not have fully realized when he graduated. At the time I just missed having a running buddy across the street who would join me for a few Keystone tall–boys.

The Texas/Wisconsin Rawlings' group has a distinctly younger and more mischievous reputation. I have always joked about writing a "disacknowledgement" section for those who slowed me down. It would certainly contain this group! For the many pleasant diversions through sport, food, discussion and general shenanigans I offer my heartfelt thanks to Rock Matthews, Pierre Scokaert, Chris Rao, Scott Middlebrooks, and Rahul Bindlish. Rock, you have been the one constant throughout my graduate tenure, and I'd like to return your comment: “Nice sharing the ride with you” because I most certainly did. Chris, I'm sure I'll never meet anyone else like you, but now that I think about it, that may not be such a good thing. I've enjoyed getting to know you and wish you the best. Rahul, somehow I think I'll see you working in Houston someday. Scott, a rare day went by when you couldn't make me laugh. I hope I occasionally returned the favor.

Next, of course, I'd like to mention the purely Wisconsin Rawlings' group. Rolf and Peter Findeisen, Jianfeng Zhang, and Daniel Patience. It is at this point I must apologize to Daniel for turning down what seemed to be hundreds of well meaning invitations (see Steve Miller above). I guess I owe you some beers! Thanks also to Shrikar Chakravarti. Due to an abrupt arrival in Madison it was helpful to have someone to teach our group the ropes at UW.

Special thanks to anyone and everyone who took me in during the rediscovery of my bachelorhood. No offense guys, but I want to live with Vicki again. Pierre, Chris, Rock, Scott, Olav Slupphaug, and especially Peter Findeisen. I regret that Peter and I were both “finishing up” when we were living together. Peter, I promise I'm not always like this, thanks for being so generous during a busy, busy time.

A big industrial thanks to Tunde Ogunnaike from Dupont. It is hard to express the impact that senior industrial researchers have on young graduate students. I am thankful to have benefited from interactions with both Tunde and Dupont.

Special thanks to Tom Badgwell who started out as my office mate at Texas, and has turned out to be a mentor of sorts. Tom, I have enjoyed working on your short course and I have always appreciated your willingness to provide valuable advice about our profession.

I appreciate all of the support I have received from afar, a good deal of it coming from my parents. Mom and Dad, I found your heartfelt words of encouragement and unfailing confidence in me to be inspirational during the tough times. Thanks for all the support.

Lastly, I would like to thank my lovely wife Vicki. I recognize that the last year and a half apart has been difficult. Your strength, independence, and patience never ceased to amaze me. Nothing makes me happier then to think about spending the rest of my life with you.

John Carl Campbell

University of Wisconsin–Madison

May 1997

Modelling, Estimation, and Control of Sheet and Film Forming Processes

John Carl Campbell

Under the supervision of Professor James B. Rawlings

At the University of Wisconsin–Madison

The purpose of this study is to use new and existing control technology to improve product properties influenced by film forming processes. In any film forming process, there is an area at which the sheet of film is extruded through some flow controlling actuators and then pinned to roller conveyers. At some point a scanning sensor measures some properties of the film before it is rolled up on a drum as the final product. The sensor is mounted on a frame and scans back and forth over the film while the film is moving underneath it, and provides measurements in a zig–zag pattern along the sheet.

The work presented in this thesis advocates a constrained infinite horizon controller. The infinite horizon controller uses a process model and an initial state estimate, thus the modelling and estimation issues are of critical importance. A model

structure is proposed that captures the main characteristics of the process including the scanning sensor and large time–delay. Model parameters are identified from process data, and symmetry is enforced to improve the parameter estimates.

The thickness of the film is estimated from sparse measurements using a periodic Kalman filter. The covariances of the estimates are used to evaluate scanning patterns. Since constant disturbance models can be used by the controller to achieve offset free control, a general disturbance model and a measured input disturbance model are presented. Lifted systems are constructed for both disturbance models. The lifted systems are useful for discussing the issues of observability and detectability. Conditions of detectability are established for both of the disturbance models.

Once the state of the film can be adequately reconstructed from the sensor data, a control strategy can be implemented to improve film properties. One important control contribution can be made by having the controller handle hard constraints on the actuators without employing clipping of the controller output signal. Determining targets and dealing with with regulator infeasibility are discussed as well. Finally, simulations bring modelling, estimation, and regulation together in an effort to better understand the periodic filter and disturbance models that are key to improved gage control.

Contents

List of Tables

List of Figures

Notation

Some of the following symbols are used for a number of different purposes in the text.

The local meaning should be clear from the context.

Lower Case Letters

d — time delay or input constraint vector

e — reconstruction error vector

h — state constraint vector

k_i — ith row of the gain matrix K

k_2 — final binding constraint index

$k_{i,j}$ — element i, j of the steady–state actuator gain matrix

l — number of measured lanes

m — number of inputs

m_x — mean of the random variable x

n_e — number of identification experiments

p — number of outputs

p_ξ — probability density of random variable ξ

$p_{\xi,\eta}$ — joint probability density of random variables ξ and η

$p_{\xi|\eta}$ — conditional probability density of random variables ξ given η

q — period of scanning sensor

s_ω — number of disturbance states, ω

s_z — number of disturbance states, z

v — open–loop input vector

x — state vector

$\hat{x}$ — estimated state vector

$\hat{x}_{k+1|k}$ — estimated state vector at time $k+1$ given data at time k

$\tilde{x}$ — augmented state vector

x_{LS} — least squares solution

u — input vector

u_s — steady–state input target

y — output vector

y_t — output setpoint

y_s — steady–state output target

y^* — vector of outputs from a single lane (multiple experiments)

y^m — output vector of measured inputs

z — disturbance state vector or open–loop state vector

Upper Case Letters

A^+ — pseudo–inverse of A

A, B, C, D — state–space model matrices

$\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$ — augmented state–space model matrices

B_r — region of guaranteed infinite horizon feasibility

B_R — subregion of feasible states

D — input constraint matrix

E, F, G, H, T — observability decompostion matrices

G_m — umeasured state disturbance matrix

G_p — umeasured output state disturbance matrix

G_ω — state disturbance matrix

$\tilde{G}_\omega$ — augmented state disturbance matrix

H — state constraint matrix

K — steady–state actuator gain matrix or LQR gain

K_{centro} — steady–state gain matrix with centrosymmetric structure

$K_{Toeplitz}$ — steady–state gain matrix with Toeplitz structure

L — Kalman filter gain matrix

M — prediction horizon

N — control horizon

$N(m, P_\xi)$ — normally distributed variable with mean m and covariance P_ξ

P — discrete Riccati matrix

Q — open–loop state penalty matrix

$\tilde{Q}$ — terminal state penalty matrix

Q_s — state target tracking penalty matrix

Q_0 — covariance matrix of x_0

Q_ω — covariance matrix of ω

Q_ξ — covariance matrix of ξ

R_ν — covariance matrix of ν

R — open–loop input penatly

R_s — input target tracking penalty matrix

S — open–loop input rate of change penalty matrix

S_s — input rate of change target tracking penalty matrix

S_1 — soft constraint quadratic penalty

S_2 — soft constraint linear penalty

T — Schur matrix

T_{11}, T_{12}, T_{22} — partitioned Schur matrices

U — steady–state input data matrix, Schur matrix

U_1, U_2 — partioned Schur matrices

V — SVD matrix

W — SVD matrix

Y — steady–state output data matrix

Greek Letters and Mathmetical Symbols

$\mathbb{R}^n$ — Euclidean space of dimension n

$\mathbb{R}^{n\times m}$— Matrix of size $n \times m$

$\mathbb{C}$ — the complex numbers

Δv — change in the open–loop input vector

Φ — objective function

χ^2 — Chi–squared distribution

α — confidence level of χ^2

ϵ — soft constraint

λ — eigenvalue

σ_i — singular value

σ — vector of soft constraints

ρ_{LS} — least squares residuals

π — future input vector

ξ — zero–mean, normal disturbance state noise vector

ν — zero–mean, normally distributed output noise vector

ω — zero–mean, normally distributed state noise vector

ξ — zero–mean, normally distributed disturbance state noise vector

ν — zero–mean, normal output noise vector

Caligraphic Letters

$\mathcal{A}$ — lifted state matrix

$\mathcal{B}$ — lifted input matrix

$\mathcal{C}$ — lifted output matrix

$\mathcal{D}$ — lifted matrix

$\mathcal{E}$ — lifted vector or expectation operator

$\mathcal{G}_\omega$ — lifted noise vector

$\mathcal{G}_p$ — lifted output disturbance matrix

$\mathcal{G}_s$ — lifted disturbance matrix

$\mathcal{G}_c$ — lifted disturbance matrix

$\mathcal{G}_m$ — lifted state disturbance matrix

$\mathcal{H}$ — lifted matrix

$\mathcal{M}$ — lifted disturbance matrix

$\mathcal{N}$ — lifted noise vector

$\mathcal{O}$ — Observability matrix

$\mathcal{P}$ — lifted vector

$\mathcal{Q}$ — lifted covariance matrix

$\mathcal{R}$ — lifted covariance matrix

$\mathcal{T}$ — lifted covariance matrix

$\mathcal{U}$ — lifted input

$\mathcal{V}$ — lifted output noise vector

$\mathcal{W}$ — lifted state noise vector

$\mathcal{X}$ — lifted state

$\mathcal{Y}$ — lifted output

$\mathcal{Y}^m$ — lifted measured input

$\mathcal{Z}$ — lifted disturbance state

Chapter 1

Introduction

If enough stapled handouts are stacked in a pile, the small imperfections in sheet thickness cause the pile to fall over. Yet without the staples, the sheets stack with ease. Millions of dollars are lost every year due to ineffective control of paper and polymer film thickness. Homogeneity of sheet thickness is critical in the paper and film industries because imperfections cause difficulties during final product rolling. Moreover, for thin specialty films used in the microelectronics industries even small imperfections may keep the product from meeting acceptance criteria. Consequently, the affected product must be discarded or recycled. With appropriate process control during production, quality films can be mass produced without these flaws.

The purpose of this study is to use new and existing control technology to improve certain product properties in the paper and film industries. While there are many stages in the paper forming process, there is one stage that determines the paper's final properties: thickness, basis weight and moisture content. The sheet forming part of the paper process involves depositing a pulp slurry onto a fourdrinier wire. The extrusion process used in the production of polymer films also determines final product properties. The paper making and film forming processes come from different industries, yet are remarkably similar and can benefit greatly from advanced control.

1.1 The Process

Polymer Film In one stage of the film forming process, the sheet of film is extruded through a die lip and pinned to roller conveyers. Then the film is exposed to a variety of external forces that can stretch, heat, dry, shrink, and coat the film. Further down the line, a scanning sensor measures the thickness of the film before it is rolled up on a drum as the final product. Figure 1.1 shows the main features of a generic film forming process.

Paper As with polymer films, Figure 1.1 is an appropriate schematic of the generic paper forming process. In paper production thick stock, consisting of treated wood fibers and water, is mixed with white liquor and pumped into a headbox. The headbox pressure is controlled to deposit a constant amount of material through a slice lip onto a moving fourdrinier wire. Some material passes down underneath the wire and is recycled as white liquor while the rest of the material is dried and pressed to form paper. A significant distance from the headbox, a scanning sensor measures some of

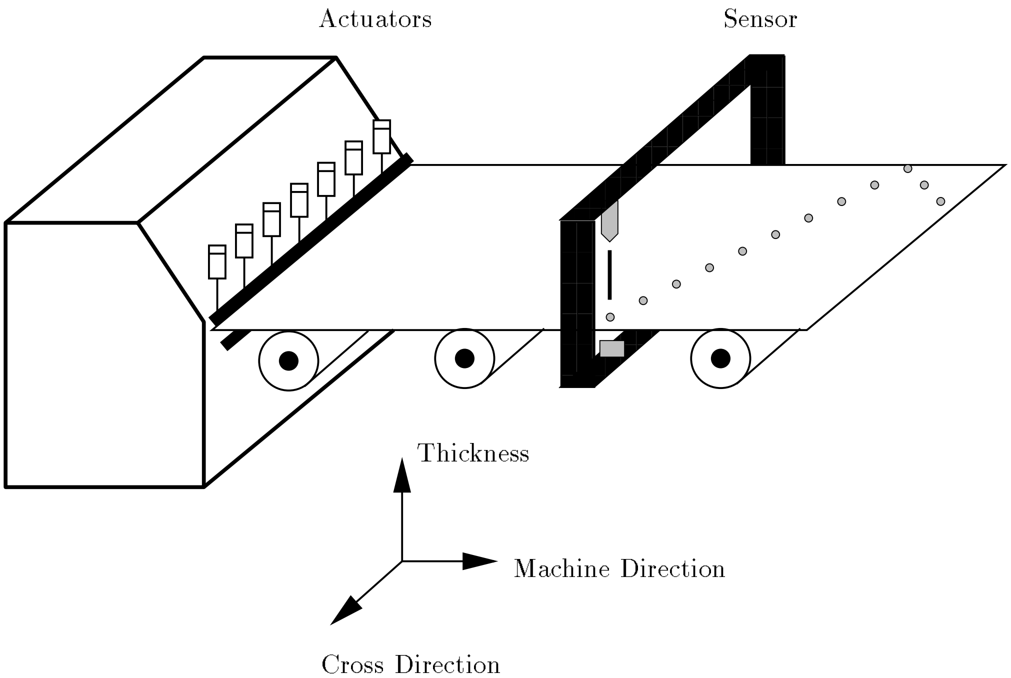

Figure 1.1: The film and sheet forming process.

the paper's properties such as thickness , basis weight, and moisture content. Manipulated variables such as actuators on the slice lip and stock flow into the headbox can be used to control the paper properties.

Both of the preceding processes, generically referred to as the *film process* in this document, have actuators that act on the die (or slice) lip. The die is a large piece of metal that can be slightly deformed by the actuators. Limited movement of the die lip allows fine tuning of the film thickness. The settings that control the base thickness of the film include die lip position, extruder screw speed and headbox pressure. Control of thickness variation, rather than control of thickness itself, is the main objective of this study.

The downstream sensor consists of a radiation emitting source and a detector. The sensor is mounted on a frame and scans back and forth over the film while the film is moving underneath it, and provides measurements in a zig–zag pattern along the sheet. The sensor measurement divides the film into lanes, and only one of the lanes is measured at each time step. The sparse measurement pattern leads to one of the most important unresolved issues: how does one determine the gage of the entire film using potentially noisy measurements from a smaller portion of the film.

1.2 Motivation

The polymer film and paper sheet industries are extensive. Capitalization of industries relying on coating technology was estimated at $500 billion worldwide in 1990 [33]. The statistic on the coating industry includes the paper and film industries as well as more classic coating applications – all of which share similar generic processes. Table 1.1 provides a brief list of some common products that are manufactured using film processes. One reason the current control methods for sheet and film forming processes do not take full advantage of modern control theory is the unusual pattern of the measurements available during run time. By applying current technologies in system identification and feedback control, along with developing new algorithms to handle time–varying estimation problems, significant contributions can be made to quality improvement in the polymer film and paper industries. Clearly, there is a compelling financial motivation for the industrial community.

Polymer Films		Paper Products
motor insulation	Tedlar®	food packaging
wire insulation	Mylar®	printer paper
cash cards	Melinex®	newspaper
garbage bags	Kapton®	posters
safety glass coatings	Teflon®	liner board
metalized film for the space shuttle	x-ray imaging film	typing paper
telephone cards	capacitors	toilet paper
print heads for Inkjet printers	Ryder truck coatings	label printer ribbons
video tape, 35 mm film	circuit boards	general packaging
aircraft interiors	green houses	signs

Table 1.1: Products manufactured with the film process

The academic community also has an interest in the film process. There is an interest in improving the time–varying estimator and the parameter estimation used to carry out detailed system identification. Continued development of model predictive control requires new methods of constraint handling, infinite horizon formulation, and optimal target calculation. Certainly, the film problem provides a rich environment for application of current theoretical issues.

The work presented in this thesis is intended to provide a theoretical foundation for methods to attack the current inadequacies in the film and paper industries. Nominal stability of the controller and globally optimal control laws provide an important theoretical base for this problem, but constraint handling, on–line solution time, and accurate model parameters are other issues that are critical to practitioners. All of these issues are addressed in this thesis.

It is hoped that the industrial practitioner will apply the ideas presented here, and that the academic will find a better starting place for further, perhaps related, estimation and control problems.

1.3 Dissertation Overview

This thesis focuses on the modelling, estimation, and control of film processes. The estimation topic is divided into chapters on Kalman filtering and disturbance modelling. A separate chapter is dedicated to simulations that illustrate all of the main topics. The chapters are summarized as follows.

Chapter 2 provides a thorough literature review for work related to film processes. The sections review each of the three major aspects of this problem: modelling, estimation, and control. A table summarizes the earlier work accomplished in control of film processes.

Chapter 3 describes the modelling issues and results. A model structure that includes the important characteristics of the process is proposed. The characteristics of the model include a significant time–delay, zero dynamics ($A = 0$), and a sparse measurement pattern. Model parameters are identified using data obtained from a 3M pilot plant. An incomplete data set raises the issue of model symmetry and leads to further model development and model evaluation. The result is a periodic, symmetric model that is used in the Chapter 7 simulations.

Chapter 4 begins with an introduction to estimation history and the theory behind the well–known Kalman filter. The periodic Kalman filter is presented and the stability of the estimate is established. Lifted systems are introduced as a method of analyzing periodic systems and calculating filter gains. One of the key steps in the formulating of the Kalman filter gain involves determining the covariance matrix of the state. The covariance matrix provides a level of confidence in the state estimate and can be used to evaluate sensor patterns and disturbance models.

Chapter 5 introduces a general disturbance model that can be used with a feedback controller to provide offset free control. The well–known output disturbance model is contained within this general disturbance model. Of special interest is the so–called input disturbance model, which is also contained in the general disturbance model. An additional model is formulated that has measured inputs. There is compelling evidence that measuring the actuator positions provides a significant control advantage in a system with a large time–delay. Lifted systems are constructed for the general disturbance model and the measured input disturbance model. The lifted systems are useful when discussing the issues of observability and detectability. Definitions and methods for checking observability and detectability are discussed in detail. The Hautus observability (detectability) lemma is presented for time–invariant systems and is applied to the periodic systems. Lastly, detectability conditions for both the general and measured disturbance models are presented.

Chapter 6 is a broad chapter on model predictive control. Several problem statements involving finite and infinite horizon control strategies are reviewed. The stability of the combined linear estimator/non–linear regulator is established. Several methods of determining targets for the regulator are discussed. This discussion focuses on the determination of targets given setpoints rather than on the non–trivial task of choosing appropriate setpoints. Next, conditions of existence and uniqueness for the target tracking solution are presented. Constraint feasibility issues such as shedding, softening, and violating constraints are discussed in the context of both the target tracking problem and the regulator problem. The specific optimization problem formulated in the model predictive control regulator lends itself well to significant speed up via careful structuring of the variables. A short discussion of structured interior point methods is presented along with some results of calculations using a film forming model. Lastly, the software used to generate the simulations in Chapter 7 is briefly discussed.

Chapter 7 consists of four simulation studies. Simulation Study I illustrates the potential gain from using the scanning data as it is acquired. Simulation Study II discusses the use of several different disturbance models. Simulation Study III addresses the problems associated with malfunctioning actuators. The measured disturbance model is proposed as a potential, albeit expensive, solution to this problem. Simulation Study IV provides some assurance that the combined periodic linear estimator and non–linear constrained regulator are robust to typical plant–model mismatch.

Chapter 8 summarizes the thesis and comments on the future direction of this work. Appendix A supplements Chapter 4 by providing a probabilistic derivation of the Kalman filter. Appendix B provides some of the longer proofs from Chapters

3 and 5. Appendix C is a manual/overview of the `lmpc` code which is used for all calculations in this study.

Chapter 2

Published Work: Modelling, Estimation, and Control of Sheet and Film Processes

Much of the literature on sheet and film forming processes is related to the paper and the sensor and detection industries. A general overview of the entire paper making process is provided by Brewster and Bjerring [21]. Dumont [39], in a more recent review paper, covers control schemes used throughout the paper making process. There are also a number of articles showing the financial gain obtained by improving the various stages of paper making production [4, 33, 74, 100, 114]. Several authors recommend process changes *before* an advanced control scheme is attempted [10, 56, 58, 71, 99, 115]. The paper industry is a mature industry, but it has only been in the last 15 to 20 years that engineers have proposed process changes that utilize advanced control. The recent enthusiasm for improved control coupled with the availability of inexpensive, powerful computers presents a unique opportunity to solve some of the previously unapproachable paper control problems.

A common theme throughout the literature is the idea that two distinct control problems exist. One control problem involves removing film variations in the direction the film is moving. This direction is referred to as the machine direction (MD). A variety of control strategies are promoted for removing machine direction variations as evidenced by the participation of seven groups in a recent paper–making benchmark problem [64]. The other control problem removes film variations in the direction in which the scanner moves. This direction is referred to as the cross direction (CD). There are statistical methods to distinguish cross direction variations from machine direction variations [23, 35, 89, 100]. Often, several control systems are employed independently to remove the various directional variations [10, 80]. Independent controllers can lead to confusion on the part of the control engineers and cannot handle the correlation between the cross and machine directional disturbances. Some investigations separate the two disturbance types and emphasize the effect process characteristics have on the cross directional variation in particular [17, 107]. Although the process characteristics and disturbances that effect machine direction and cross direction variation may be uncorrelated, these problems can and should be considered simultaneously.

Brewster states [20].

> "MD control should be considered as an extension of CD control, involving the profile average and an additional actuator."

Although much of the research presented in this document refers to cross directional control, the general philosophy that machine direction can be controlled simultaneously with additional actuators is supported. Zhang [124] discusses including CD and MD actuators simultaneously.

The remainder of the sheet and film forming literature of interest here can be divided into the following three categories: modelling, state estimation and regulation. Although many of the papers address more than one of these issues, it is convenient to summarize them separately. Table 2.1 categorizes the literature directly addressing work in the area of gage control.

2.1 Modelling

The complexity of any modelling effort often depends on the final use of the model, and modelling sheet and film forming processes is no different. Some models are used to better understand the characteristics and limits of the system. First principles models may be used to help design die lips or to decide optimal operating regions for the process. These fundamental models can also determine how to produce thinner film best. Modelling for real time control, however, requires that models be accurate within the operating region of choice while being solved quickly. Consequently, simple linear models are frequently adequate for implementation of feedback control. Efforts

Topic	Subtopic	#	Authors
Review	Paper making process	2	Brewster and Bjerring [21], Dumont [39]
	Suggested process improvements *before* control implementation	6	Bialkowski [10], Hadra [56], Haverinen [58], Kan [71], Roth [99], Wallace [115]
	Potential financial gain obtained by using advanced control	5	Atkins *et al.* [4], Cohen [33], King [74], Saucier and Steinkirchner [100], Wallace [114]
Modelling	First principle polymer models	5	Chambon [29], Kajiwara *et al.* [68], Liu, Wen, and Liu [82], Smith *et al.* [108], Soucy and Holt [111]
	Linear models	7	Åström [2], Bialkowski [9], Boyle [15], Heaven *et al.* [61], Rawlings and Chien [96], Wilhelm Jr. and Fjeld [118], Wilkinson and Hering [119], Xia, Rao, and Qian [123]
	System symmetry	2	Braatz *et al.* [18], Featherstone and Braatz [44]
	Identifying CD and MD variations	4	Burns [23], Dahlin [35], Pfeifer and Wilhelm [89], Saucier and Steinkirchner [100]
	Effects of CD variations	2	Boyle and Hamby [17], Siler [107]
Estimation	Filtering to separate CD and MD variations	1	Chen [30]
	Kalman Filtering	7	Bergh and MacGregor [8], Campbell and Rawlings [24] Dumont *et al.* [40], Duncan [41], Fjeld [47], Rawlings and Chien [96], Tyler and Morari [113]
	Principal Component Analysis	1	Rigopoulos *et al.* [98]
Regulation	Controlling CD and MD variations independently	3	Bialkowski [10], Lebrasseur [80], Zhang [124]
	MD control	2	Isaksson *et al.* [64], Wang, Dumont and Davies [116]
	Adaptive schemes	9	Beecher and Bareiss [7], Cegrell and Hedqvist [28], D'Hulster [38], Guo, Ni, and Zheng [55], Halousková, Kárný, and Nagy [57], He [59], Liang *et al.* [81], Nagy and Dušek [87] Rastogi [95]
	Measured inputs	3	Karlsson [72], Nuyan [88], Campbell and Rawlings [27]
	Constrained QP	3	Boyle [15, 16], Heath [60], Rawlings and Chien [96],
	Various constrained schemes	4	Braatz *et al.* [19], Chen and Wilhelm [31], Kristinsson and Dumont [76], Keyes and Kaya [73]
	Robust Control	2	Duncan [42], Laughlin, Morari, and Braatz [79]

Table 2.1: Summary of Literature

have been made to obtain models for both understanding and control.

In an effort to better design die lips, Liu, Wen, and Liu [82] use finite element methods to compare two different die lips. Their results identify stagnant flow regions in one design. Further modelling provides alternative designs that provide more uniform flow characteristics. Kajiwara *et al.* [68] model stress distribution through a thin slit die lip. Experimental results support the validity of the model. Although the full three–dimensional problem is expensive to solve, the result is a better understanding of the stresses on the die lip. Smith *et al.* [108] use sensitivity analysis to design die lips with constant exit velocity – another example of a detailed model that helps determine a better process design. Soucy and Holt [111] use extended Kalman filtering in conjunction with a heat transfer and kinetic behavior model to determine extent of reaction. This work can be used to better understand the curing of thermosetting polymer composites with the ultimate goal of achieving product consistency. Visualization techniques advocated by Scriven and Suszynski [105] provide insight into the complex flow present in coating processes. Chambon [29] uses first principle modelling to determine speeds and widths at which certain thin films break. The fundamental modelling effort throughout the scientific community provides valuable information for both design and operation of film processes. However, even with the currently available computing power and sensor technology, fundamental models are difficult to use for real time control. Consequently, the literature is rich with alternative models for use in control algorithms.

A number of linear and non–linear models relating actuator movement to gage properties are proposed in the literature. Among these models are steady–state gain matrices which assume the dynamics of the process due to actuator movement can be neglected. Boyle [15] discusses and Wilhelm Jr. and Fjeld [118] elaborate on the diagonally dominant steady–state gain matrix model. Wilkinson and Hering [119] suggest a self tuning or updating gain matrix for their steady–state model. This model attempts to keep up with changes in the process due to changing product specifications or changing operating conditions. Åström [2] uses a linear first order ARMAX model identified from process data. Heaven *et al.* [61] propose a similar linear first order model with a time–delay. Wang, Dumont, and Davies [117] propose a more complex non–linear statistical model that couples the machine direction and cross direction variations in the sheet properties in order to estimate the variances in either direction. The articles cited above ignore the time–delay even though it is a dominant process characteristic. An impulse response model which handles large time–delays is used in the work of Xia, Rao, and Qian [123]. Bialkowski [9] stores past states in order to model the time–delay and, in addition, uses a time–delay compensator to update the estimate of the time–delay. There is strong support in the literature for a model containing both a steady–state gain matrix and a time–delay, although the two ideas have not been used together.

A discussion of the types of symmetry common to paper and film forming processes is found in Braatz *et al.* [18]. In addition, Featherstone and Braatz [44] discuss gain directionality for design experiments to identify circulant symmetric processes. Lastly, there is modeling work concerning identification from sparse data. Jones [66] uses a Markovian representation to fit ARMA models in the case of missing observations. The inherent symmetry in these processes can also be used to help overcome incomplete data, and is discussed in Section 3.4.

2.2 Estimation

The next major section of literature pertains to state estimation. The scanning sensor creates a challenging state estimation problem and leads to interesting filtering methods. Filtering techniques that decouple the machine direction and cross direction variations are presented by Chen [30]. Dumont *et al.* [40] obtain on–line estimates for the cross direction moisture content profile via Kalman filtering in conjunction with a recursive least–squares algorithm. This algorithm updates the current lane and adjacent lane values of moisture content with each measurement. A slightly different approach is taken by Rigopoulos *et al.* [98] to estimate disturbance modes using principal component analysis. Fjeld [47] computes time–invariant Kalman filter gains for a model with a time–delay. Bergh and MacGregor [8] are the first to present the idea of using a time–varying measurement matrix in conjunction with a Kalman filter for this process. Rawlings and Chien [96] elaborate on this idea and demonstrate how the covariance of the estimates can be used to evaluate different scanning patterns. Tyler and Morari [113] also discuss evaluating the differences between scanning and fixed sensors. Duncan [41] designs a time–varying periodic Kalman filter with a period of two (counting one for each scan direction). Duncan incorporates a second stationary scanner in order to help differentiate between cross and machine directional variations.

It is useful to have a general disturbance model because disturbances can originate from many parts of the process. One appeal of the Kalman filter is the wide range of disturbance models that can be used to obtain accurate estimates of the disturbances that affect the system. The scanning sensor forces engineers to implement complex estimation techniques. Some of these methods are standard optimal estimation techniques while others have open theoretical questions. Very little is published about the properties of the estimates and the estimators used in the sheet and film industry.

2.3 Regulation

The last major section of literature covers regulation. Paper processes are amenable to adaptive control because of frequent changes in product grade and raw materials. As a consequence, several control schemes using adaptive techniques are used [28, 38, 55, 95]. Nagy and Dušek [87] use adaptive techniques to re–identify the model for control, but a robot that manipulates one actuator at a time increases the complexity of their unconstrained control algorithm. Halousková, Kárný, and Nagy [57] continue this work and outline a detailed identification and control algorithm. Xiaohua George

He [59] tests his adaptive controller design through simulations with lane drift or actuator/measurement misalignment. None of these adaptive techniques constrain the inputs or the outputs.

The frequency of the control action is currently limited by the sensing systems which only provide data after one or more complete scans. Beecher and Bareiss [7] use full sensor scans to obtain estimates of the thickness of the film before taking control action. Keyes and Kaya [73] use statistical quality control to determine optimal set–points, which are then sent to a controller. Liang *et al.* [81] directly address the problem of control *during* grade changes using feedforward control and gain scheduling. By gain scheduling the authors are successfully able to minimize the time between grade changes. Nuyan [88] and Karlsson [72] both work with systems in which the inputs are measured. Potential problems are avoided by monitoring the difference between the desired input and the actual input. Again, these efforts do not fully address the need for input constraints and several of the algorithms advocate clipping. However, clipping can have undesirable effects as shown by Rawlings and Chien [96].

Boyle [15, 16] first proposed solving a quadratic program to handle hard actuator constraints on a steady–state process model, but this was nearly impossible to implement at the time due to the computing limitations. Later, Chen and Wilhelm [31] used a constrained algorithm on the steady–state problem. Bergh and MacGregor [8] use an LQG controller and avoid constraints by tuning the controller accordingly. Wang, Dumont and Davies [116] use Generalized Prediction Control

(GPC) to control the machine direction variations of paper basis weight. They are able to consider the time delay, but do not address constraints. Braatz *et al.* [19] assume a stationary bank of sensors and scale the input in order to handle input constraints. Kristinsson and Dumont [76] use Gram polynomials to optimize the shape of the slice lip to best control the cross directional variations. This method does handle constraints, but is not a predictive controller. Heath [60] provides controllability conditions for any orthogonal basis functions including Gram polynomials. A key feature of the basis functions approach is that high order systems with large numbers of inputs and outputs can be represented by a much smaller number of coefficients than a full order model. Heath also solves a constrained QP to determine the optimal control profile although it is assumed that the entire film profile is available at each time step (*e.g.* no scanning sensor).

There has been relatively little work in the area of robust control for sheet forming processes. The inherent symmetries and special properties of the systems are used to simplify the design of a robust controller. Laughlin, Morari and Braatz [79] address robust control issues by designing robust SISO controllers that can be scaled to robust MIMO controllers. Constraints are handled via scaling methods. Duncan [42] expands the robustness work in [79] by handling the detuning in the pre–compensator rather than in the controller, thus providing fewer adjustments to the actuators away from the vicinity of the error.

2.4 Summary

It is clear from this review of the literature that there is still much room for much improvement in the estimation and control of paper and film processes. The search for a better predictive controller begins with the modelling effort. The fundamental models find application in the design of die lips and in the search for optimal operating regions. The empirical models, however, are developed for advanced process control. Some of the simplifying assumptions accepted by the research community include zero dynamics and a diagonally dominant steady–state gain matrix. These assumptions are used to create a complete linear model. The model can then be used with a constrained infinite horizon model predictive controller.

The infinite horizon predictive controller needs not only a model, but also an estimate of the state. Therefore, there is need for a sound theoretical foundation on which to base the design of an estimator. Choosing the appropriate disturbance model for the sheet and film forming process is a difficult task. The calculation of the filter gains for time–varying systems is complicated. Determining the observability and detectability of the estimates is complicated by the sparsity of the measurements. Stability of the combined estimator and regulator is important as well. There are still many unresolved issues in the control of sheet and film processes to date. Several of these modelling, estimation and control issues are addressed in the following chapters.

Chapter 3

Modelling

There are many questions which must be addressed when evaluating the accuracy of a model: in which operating regions is the model accurate, how long does it take to get an accurate prediction, are the parameters that are predicted accurately the parameters of interest? Since the goal of this project is to remove variations in gage thickness, the model should be accurate within a specific operating region. This model need only approximate the real system because it is used in conjunction with a feedback controller. The model is not required to predict the thickness from startup to grade change to shutdown. The model is required to handle the system's key characteristics effecting thickness of the operating region in question. This chapter discusses the key characteristics of the gage model, how these particular characteristics shape the model, and the numerical methods used to determine the model parameters.

3.1 Modelling Paper Versus Modelling Film

The film and paper processes are similar enough to share the common model structure described in the following sections. However, there are some differences worth mentioning. The paper processes tend to be larger in dimension. There may be over 100 actuators rather than 10-30 as in the polymer processes [118]. The paper sheets may be wider, and consequently, the number of sensed lanes can increase dramatically. A film process may have up to 150 measurements while a paper process can have more than 400 [76]. It is important to note that these differences increase the size of the model significantly. Although a large model may pose difficulties when computing optimal control strategies, it does not affect the identification of the model parameters. The rest of this chapter focuses on the film forming process. However, parallel conclusions can be drawn for paper forming processes by remembering that paper production terms, slice lip, sheet, and basis weight, can be interchanged with the film production terms, die lip, film, and gage, in the following discussion.

3.2 Model Structure

This section discusses the assumptions used to determine the structure of the film process model. Linear dynamics, sparse measurements, and a large time–delay are the key process characteristics incorporated into the model.

3.2.1 Linear Dynamics

There is strong evidence that the dynamics caused by actuator movement are negligible compared to the process time–delay. This claim supports the use of a steady–state gain model that relates actuator position to film thickness. When actuator dynamics are important, for example for slower moving films, a low order linear model has been successful in approximating the dynamics [2, 61]. The simplest model to consider for this process is a linear, finite–dimensional, time–invariant dynamic model. In discrete time, this model has the form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned} \tag{3.1}$$

In this model k is the time index, x is a vector of film thicknesses in the cross direction, u is a vector of actuator positions, and y is the scanner measurement(s). The state in a state–space model is not required to have physical meaning, but choosing the thickness in each lane as the state is a logical choice for this model. In general, $x \in \mathbb{R}^n$ where n is equal to the number of lanes, $u \in \mathbb{R}^m$ where m is equal to the number of actuators, and $y \in \mathbb{R}^p$ where p is equal to the number of measurements. Since most processes have a mounted scanning sensor, y is a scalar. In addition, the model matrices $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$ are referred to as the state matrix, input matrix and measurement matrix, respectively.

3.2.2 Measurement Matrix

Since the scanner reports only one measurement at every time step (see Figure 3.1) a single measurement matrix is not appropriate. C is replaced by a set of measurement matrices. These measurement matrices are denoted C_j with j ranging from zero to $q-1$ where q is the period of the scanning sensor. For the scanning pattern shown in Figure 3.1 the period is $2l-1$ where l is the number of lanes. Since the state is a vector of film thicknesses a typical C_j matrix contains zero elements except in the lane corresponding to the sensor.

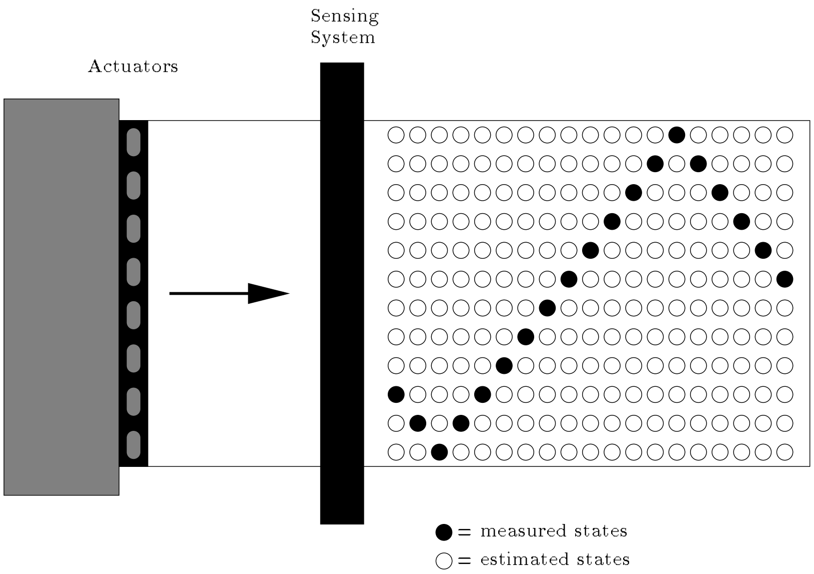

Figure 3.1: Schematic of scanned measurements.

Eq. 3.2 illustrates the full period of the measurement matrix for a simple four lane example. It is important to note that the period is six, not four or eight. The period is six because a measurement obtained from the second lane when the scanner is moving to the third lane is considered different from a measurement acquired in the second lane when the scanner is moving to the first lane. The edge measurements are included only once during the period.

$$
\begin{aligned}
y_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} x_1 \\
y_2 &= \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} x_2 \\
y_3 &= \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} x_3 \\
y_4 &= \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} x_4 \\
y_5 &= \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} x_5 \\
y_6 &= \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} x_6
\end{aligned}
\tag{3.2}
$$

This pattern then repeats and

$$
y_7 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} x_7
$$

Certainly this model can be modified to include other measurement arrangements. For each additional sensor measurement there is an additional row in the measurement matrix. For a stationary sensor this additional row has a one in the column corresponding to the sensor's particular lane. For an additional scanning sensor with the same period as the first scanning sensor, the indexing follows in a straightforward manner. For scanning sensors with different periods, the overall period of the measurement matrix is equal to the least common multiple of the individual periods.

3.2.3 Time–Delay

If actuator dynamics are fast, then $A = 0$ and $B = K$. The assumption of fast dynamics arises because of the significant time–delay between actuator–induced change and sensor detection. Certainly the film has a dynamic response to the actuator movement, but it is futile for the control algorithm to attempt to control a transient response on the order of seconds when the time–delay is on the order of minutes. Furthermore, the scanner may be on the opposite side of the sheet when the transient passes under the sensor. Thus, it is a necessary assumption to ignore the brief transient that would be both difficult to control and to detect.

The time–delay, along with fast dynamics and a scanning sensor, provides a more specific linear model

$$\begin{aligned} x_{k+1} &= K u_{k-d} \\ y_k &= C_j x_k \qquad j = (k+q) mod(q) \end{aligned} \tag{3.3}$$

where the notation $(k+q)mod(q)$ refers to the remainder of $(k+q)/q$. The time–delay, d, can be handled by enlarging the state. Either past states or past inputs can be stored. In this case, there are fewer actuators than lanes so past inputs are stored to keep the state vector a manageable size. The state vector becomes

$$\tilde{x}_k = \begin{bmatrix} x_k \\ u_{k-d} \\ \vdots \\ u_{k-2} \\ u_{k-1} \end{bmatrix} \tag{3.4}$$

The model given by Eq. 3.3 can then be written in the following state–space form:

$$\begin{aligned} \tilde{x}_{k+1} &= \tilde{A}\tilde{x}_k + \tilde{B}u_k \\ y_k &= \tilde{C}_j\tilde{x}_k \qquad j = (k+q)mod(q) \end{aligned} \tag{3.5}$$

where $\tilde{A}$, $\tilde{B}$, and $\tilde{C}_k$ are defined as follows (A is included for completeness):

$$\underbrace{\begin{bmatrix} x_{k+1} \\ u_{k-d+1} \\ \vdots \\ u_{k-1} \\ u_k \end{bmatrix}}_{\tilde{x}_{k+1}} = \underbrace{\begin{bmatrix} A & B & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \\ & & & & \end{bmatrix} \begin{bmatrix} x_k \\ u_{k-d} \\ \vdots \\ u_{k-2} \\ u_{k-1} \end{bmatrix}}_{\tilde{A}\tilde{x}_k} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix} u_k}_{\tilde{B}u_k} \tag{3.6}$$

$$y_k = \underbrace{\begin{bmatrix} C_j & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_k \\ u_{k-d} \\ \vdots \\ u_{k-1} \end{bmatrix}}_{\tilde{C}_j\tilde{x}_k} \qquad j = (k+q)mod(q) \tag{3.7}$$

3.2.4 Variable Time–Delay

If the film speed is changing, the time–delay is not constant. Many processes measure either film speed or roller speed which can be related to film speed. The optimal way to handle variable time–delay depends on the frequency of the variation.

- **High frequency variations in film speed:** The large time–delay in the film process makes it impossible to handle high frequency disturbances. A frequently changing film speed is a high frequency disturbance, thus it cannot be handled by this model. If the process cannot be altered to run at a fixed line speed then an average film speed must be chosen. Feedback control can be used to handle the mismatch introduced by using the average film speed rather than the exact film speed. If the frequency of control action is reduced, the effect of film speed variation is lessened and the control approaches steady–state control.

- **Low frequency variations in film speed:** The solution to this problem is to update the model online. The controller predicts that the current film speed is constant in the future. When the film speed measurement exceeds some predetermined bounds the input matrix, $(\tilde{B})$, is altered. Changing the model online causes intermittent inaccuracies due to the previously stored inputs, but after a short period of time the stored inputs work their way out of the state vector and the new model correctly models the time–delay. A good strategy for this solution is to assume a constant sample time and start with a state vector, $\tilde{x}$, large enough to handle the longest possible time–delay. Figure 3.2

illustrates the potential changes in the model as the film speed changes. With this model structure the input matrix essentially stores the current input in the state variable at a specific location. As the time–delay changes the input matrix alters the storage location of the current input.

$$\tilde{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \\ 0 \end{bmatrix} \xrightarrow{\text{Increase in film speed}} \tilde{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ I \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\text{Decrease in film speed}} \tilde{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ I \end{bmatrix}$$

Figure 3.2: Adaptation of model to changes in film speed

3.3 Model Identification

There are two parts of the model that need to be determined from data. The first is the time–delay. The time–delay is assumed to be constant and can be measured during routine operation. The second part of the model determined from process data is the gain matrix. If the steady–state assumption is valid (*e.g.* the film reaction to actuator movement is virtually instantaneous) then the scanning sensor does not make the identification more difficult. Data can be obtained by injecting a unit step into an actuator and waiting for the sensor to measure all lanes. The problem of determining a gain matrix from the input/output data is a least squares problem. This section presents a conventional statement of the least squares problem before proposing the gain matrix identification problem. Since the solution technique uses singular value decomposition, it is discussed first.[1]

3.3.1 Singular Value Decomposition

Theorem 3.1 (Singular value decomposition (SVD)) *If A is a real $m \times n$ matrix with rank$(A) = r$ then there exist orthogonal matrices*

$$W = \begin{bmatrix} w_1, & \dots, & w_m \end{bmatrix} \in \mathbb{R}^{m \times m} \qquad \textit{and} \qquad V = \begin{bmatrix} v_1, & \dots, & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$$

such that

$$A = W^T \Sigma V$$

where

$$\Sigma = diag\{\sigma_1, \dots, \sigma_p\} \in \mathbb{R}^{m \times n} \qquad p = \min\{m, n\}$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0$$

Proof: See Appendix B.1. □

Comments on the SVD

1. $\sigma_1, \dots, \sigma_r$ are the singular values of A.

2. W and V are unitary matrices

[1] For this section the previous definitions of the variables A and x as linear model parameters are discarded. Traditional linear algebra notation, in which capital letters denote matrices and lower case letters denote vectors, applies.

3. w_i and v_i are known as the ith left singular vector and the ith right singular vector respectively.

4. One interpretation of the singular vectors is that A transforms v_i into a vector with gain σ_i in the direction w_i. ($Av_i = \sigma_i w_i$ and $A^T w_i = \sigma_i v_i$ for $i = 1 : \min\{m, n\}$).

5. $\|A\|_2 = \sigma_1$

6. $\|A\|_F = \sigma_1^2 + \ldots + \sigma_p^2$ (for more on the Frobenius norm see remark 3.4).

It is next necessary to define the pseudo–inverse, which conveniently follows from the singular value decomposition.

Definition 3.1 (Pseudo–Inverse) *Define the matrix $A^+ \in \mathbb{R}^{n\times m}$ as $A^+ = V\Sigma^+W^T$ where*

$$\Sigma^+ = diag\left(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0\right) \in \mathbb{R}^{n\times m} \tag{3.8}$$

If rank$(A) = n$ then $A^+ = (A^TA)^{-1}A^T$, while if $m = n =$ rank(A), then $A^+ = A^{-1}$. Frequently, for numerical reasons, a tolerance ϵ is chosen whereby for every $\sigma_i < \epsilon$, $\sigma_i \equiv 0$. A^+ is the pseudo–inverse of A.

3.3.2 Least Squares

With the SVD and the pseudo–inverse defined, the least squares problem and solution follows.

Problem 3.1 (Least Squares) *Consider the problem of finding a vector $x \in \mathbb{R}^n$ that minimizes the 2–norm ($\|Ax - b\|_2$) where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.*

Remark 3.1 *For the underdetermined case ($m < n$) in which there are more unknowns than equations the system either has no solution or an infinite number of solutions. This case is not relevant to the gain matrix identification problem.*

Remark 3.2 *For the square case ($m = n$) in which rank(A) $= r$ there is an exact solution for $r = m = n$. The solution can be found numerically provided that the problem is well conditioned. The solution is shown in Eq. 3.9. If A is singular or the problem is ill–conditioned because A^{-1} is difficult to calculate, then the pseudo–inverse may be used.*

$$x_{LS} = (A^T A)^{-1} A^T b = A^+ b \tag{3.9}$$

For $r < m$ an infinite number of solutions exist, however, any of several orthogonal transformations of A can provide a reliable solution to the minimization of the 2–norm. The solution shown here utilizes the SVD.

$$x_{LS} = \sum_{i=1}^{r} \frac{w_i^T b}{\sigma_i} v_i \tag{3.10}$$

where the square of the residual ρ_{LS} is defined as

$$\rho_{LS}^2 = \|Ax_{LS} - b\|_2^2 = \sum_{i=r+1}^{m} w_i^T b \tag{3.11}$$

Remark 3.3 *For the overdetermined case ($m > n$) in which there are more equations than unknowns there is usually no exact solution (only if b is an element of the range* *of A). If $r = n$ and the problem is well conditioned then the solution is again given by Eq. 3.9, but for the rank deficient problem an orthogonal decomposition is needed to provide a reliable solution to the minimization of the 2–norm. Again the SVD may be used and the solution is given by Eq. 3.10.*

The next step is to tailor the least squares problem to fit the data from the film forming process. Figure 3.3 illustrates the relationship between the actuator positions and the thickness in a given lane. Since the goal of this identification is to determine the gain matrix from a series of experiments consider Figure 3.4 which relates a series of actuator movements to the thickness measurements in a single lane. Now Problem 3.1 can be rewritten to reflect the actuator position–lane thickness relationship.

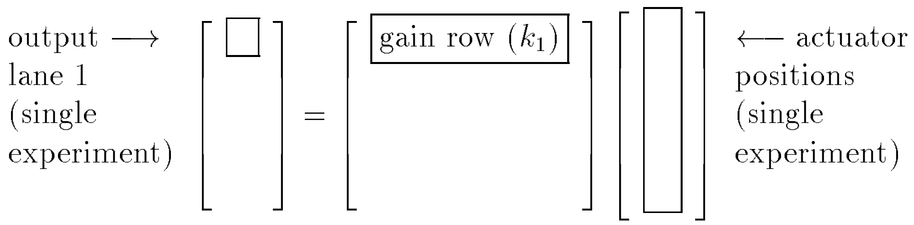

Figure 3.3: The steady–state gain equation ($y = Ku$). The boxes illustrate the relationship between the thickness of the film in lane 1 and the actuator positions.

Problem 3.2 (Gain Vector) *The problem is to determine the effect that each of the m actuators has upon the thickness of an individual lane given the data from n*

$$\begin{array}{l}\text{output} \longrightarrow\\ \text{lane 1}\\ \text{(multiple}\\ \text{experiments)}\end{array}\begin{bmatrix}\square\\ \square\\ \vdots\end{bmatrix} = \underbrace{\begin{bmatrix}\boxed{\text{Experiment 1}}\\ \boxed{\text{Experiment 2}}\\ \vdots\end{bmatrix}}_{\substack{\text{actuator positions}\\ \text{(multiple experiments)}}}\begin{bmatrix}\boxed{}\end{bmatrix}\begin{array}{l}\longleftarrow \text{gain}\\ \text{vector}\\ (k_1^T)\end{array}$$

Figure 3.4: The gain vector identification equation ($y^* = Uk_1^T$). y^* is a vector of lane 1 measurements from multiple experiments.

individual steady–state experiments. The actuator data matrix is defined as follows

$$U = \begin{bmatrix} \tilde{u}_1 \\ \vdots \\ \tilde{u}_{n_e} \end{bmatrix} \in \mathbb{R}^{n \times m} \tag{3.12}$$

where $\tilde{u}_i \in \mathbb{R}^{1\times m}$ *represents the positions of the* m *actuators during the ith experiment (*n_e *total experiments). The sensor data vector for lane 1 is given by* $y^* \in \mathbb{R}^{n\times 1}$ *and the gain vector relating all inputs to lane 1 is given by* $k_1 \in \mathbb{R}^{m\times 1}$*, thus the following least squares problem is proposed*

$$\min \|Uk_1 - y^*\|_2 \tag{3.13}$$

A series of least squares problems arise from each set of n experiments. Each individual problem calculates the effect of all m actuators on a single lane. Multiple dependent variables can be calculated using a single least squares problem. Figure 3.5 illustrates the relationship of the actuator positions from several experiments to the corresponding output in a given lane. The matrix least squares problem is now defined.

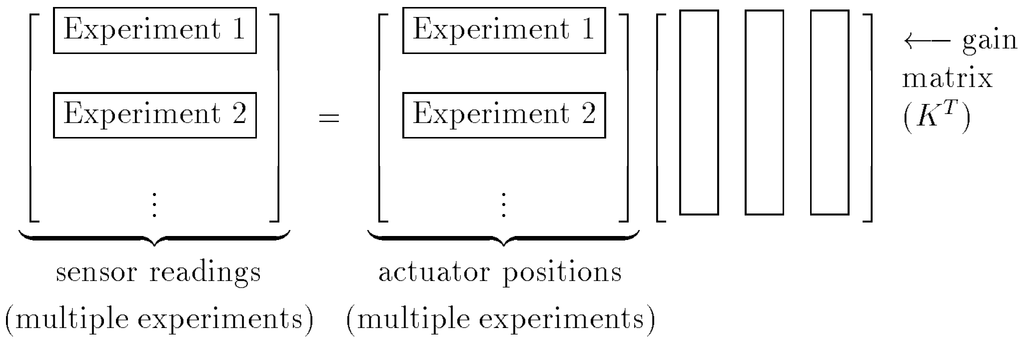

Figure 3.5: The gain matrix identification equation ($Y = UK^T$).

Problem 3.3 (Gain Matrix) *Given an actuator data matrix and a sensor data matrix as follows*

$$U = \begin{bmatrix} \tilde{u}_1 \\ \vdots \\ \tilde{u}_{n_e} \end{bmatrix} \in \mathbb{R}^{n \times m} \qquad \textit{and} \qquad Y = \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_{n_e} \end{bmatrix} \in \mathbb{R}^{n \times p} \tag{3.14}$$

where $\tilde{u}_i \in \mathbb{R}^{1 \times m}$ *represents the positions of the* m *actuators during the ith experiment and* $\tilde{y}_i \in \mathbb{R}^{1 \times p}$ *represents the thickness of the* p *lanes during the ith experiment, solve the following problem*

$$\min \|UK^T - Y\|_F \tag{3.15}$$

for the gain matrix $K \in \mathbb{R}^{p \times m}$.

Remark 3.4 *The matrix least squares problem shown in Problem 3.3 minimizes the Frobenius norm of* $(UK^T - Y)$. *The Frobenius norm for* $A \in \mathbb{R}^{m \times n}$ *is given as follows,*

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_r^2} \tag{3.16}$$

The minimization of the Frobenius norm rather than the 2–norm is a consequence of combining the individual least squares problems, and the properties of the individual solutions are preserved (e.g.the matrix problem is equivalent to solving Problem 3.2 repeatedly).

3.4 Gain Matrix Symmetry

Identification Example A properly designed identification experiment provides a well conditioned input matrix. The input matrix and its corresponding output matrix are used to determine the gain matrix via Eq. 3.9. In the case of missing input data, as in the following example, the singular value decomposition can be used to determine a solution. The following calculation uses data from a 12 input, 12 output polymer film line at 3M in which the third actuator is disabled. Four of the 36 steady–state experiments are shown in Figure 3.6. Absence of data relating the third actuator to the outputs generates a singular data matrix U. The gain matrix is obtained by taking the singular value decomposition of U and inverting only the non–zero singular values. This gain matrix is shown as a shaded picture in Figure 3.7. Larger values are shaded darker. The figure displays the diagonal dominance of the gain matrix. The third column is identically zero due to lack of information about the third actuator.

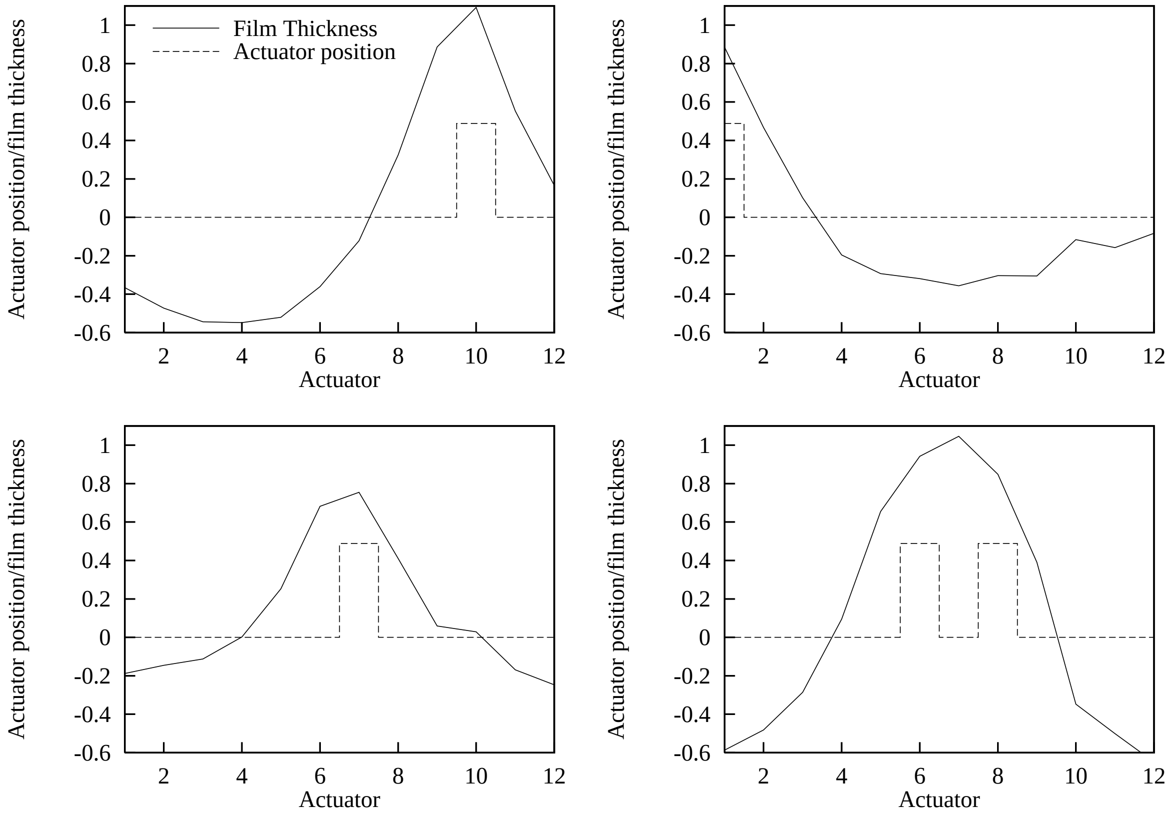

Figure 3.6: Normalized step test data for the 3M polymer film forming process.

Despite the third column, it is clear that the gain matrix in Figure 3.7 has a symmetrical structure. Braatz *et al.* [18] summarize several symmetries that are relevant to film forming processes. Based on symmetry they define several structured gain matrices. Of note are centrosymmetric and Toeplitz model structures.

A model that is centrosymmetric has symmetry about the middle of the sheet. The film's response to the first actuator is symmetric to the film's response to the last actuator. Similar relationships exist between the second actuator and the second to last actuator, *etc.* Centrosymmetry is shown in Figure 3.8 through a series of unlikely

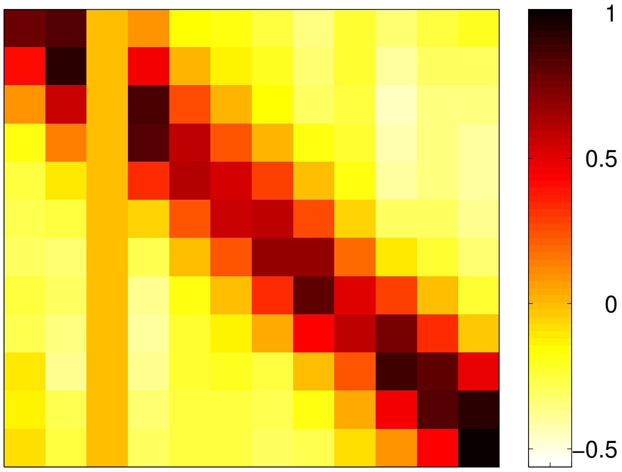

Figure 3.7: Gain matrix, no symmetry enforced; notice the elements corresponding to the third actuator are identically zero.

but illustrative step changes in the input. The corresponding gain matrix for a system with centrosymmetry is shown in Eq. 3.17 for a system with m inputs and p outputs. Since the matrix has $mp/2$ unique entries the variable h is used to simplify notation ($h = p/2$).

$$K_{\text{centro}} = \left[\begin{array}{cccc} k_{1,1} & k_{1,2} & \cdots & k_{1,m} \\ k_{2,1} & k_{2,2} & \cdots & k_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ k_{h,1} & k_{h,2} & \cdots & k_{h,m} \\ \hline k_{h,m} & k_{h,m-1} & \cdots & k_{h,1} \\ \vdots & \vdots & \ddots & \vdots \\ k_{2,m} & k_{2,m-1} & \cdots & k_{2,1} \\ k_{1,m} & k_{1,m-1} & \cdots & k_{1,1} \end{array}\right] \tag{3.17}$$

The parameters for the centrosymmetric model are calculated by reorganizing the data and then solving Problem 3.3. The reduced problem has only $m/2$ unique

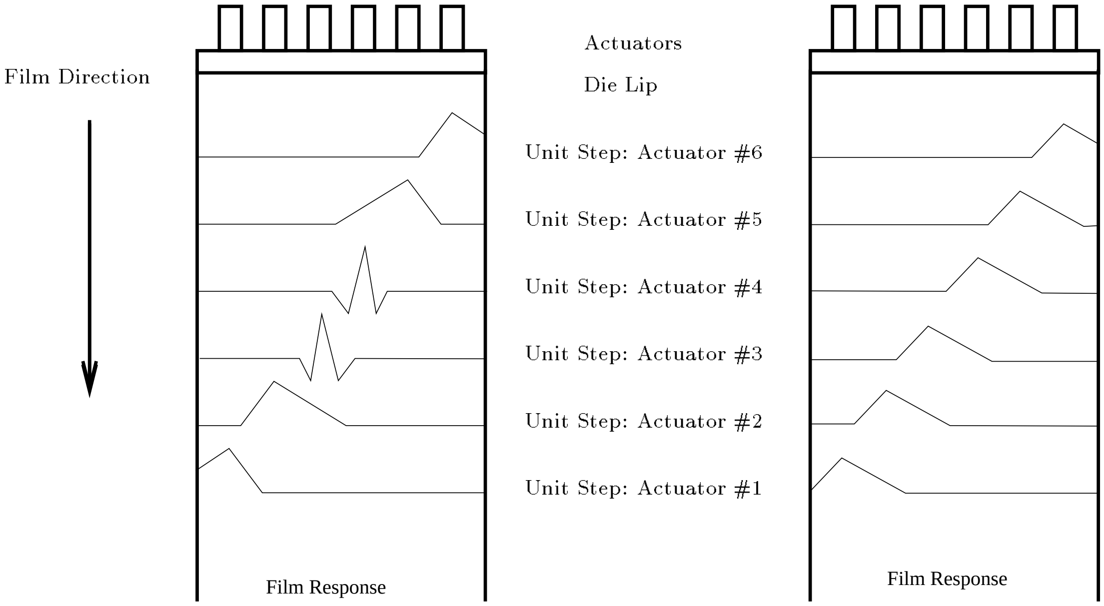

Figure 3.8: Overhead view of film sheets with profiles corresponding to systems that exhibit *centrosymmetric* structure (left) and *Toeplitz* structure (right).

actuators, thus the data from actuators $m/2+1$ through m are considered equivalent to data from actuators $m/2$ through 1 respectively. In the spirit of Figure 3.5, Figure 3.9 shows how the data is organized. Figure 3.10 shows the result of the least squares calculation of a centrosymmetric gain matrix using the 3M data set.

Another type of symmetry arises when all actuators produce shifted versions of the same response. The film's response to the first actuator is the same as its response to the second actuator although the response is shifted by one lane. A consequence of this shifted film response is that the elements of the gain matrix are constant along the diagonals. A square matrix with constant elements along the diagonals is called a Toeplitz matrix. Figure 3.8 shows Toeplitz structure through a set of identical but shifted responses to step changes (notice that the individual responses are not

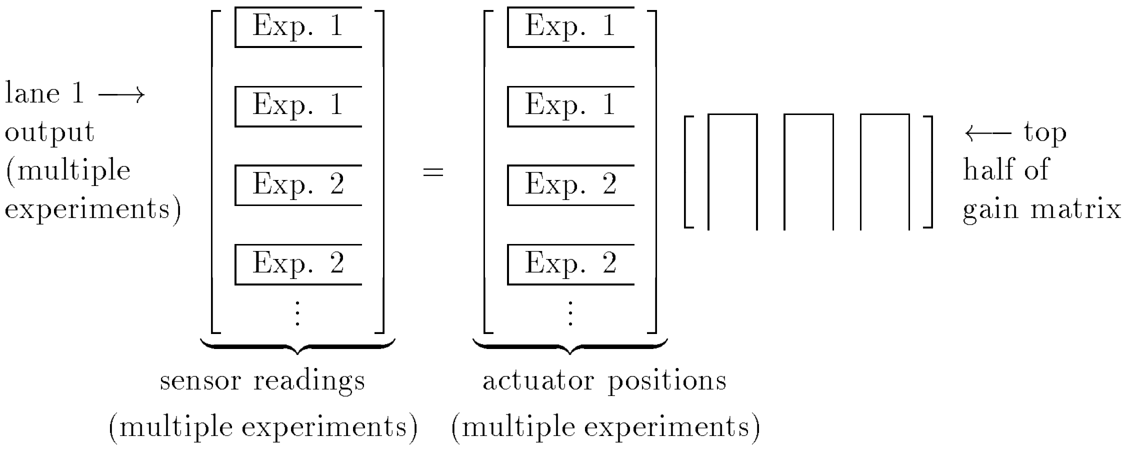

Figure 3.9: Calculating a centrosymmetric matrix requires restructuring the data.

symmetrical). Eq. 3.18 illustrates the structure of a Toeplitz matrix for a system with m inputs, p outputs, and $2m-1$ unique parameters. Enforcing symmetry about the major diagonal leads to a Toeplitz symmetric gain matrix. A Toeplitz symmetric matrix corresponds to $k_i = k_{-i}$ for all i in Eq. 3.18 for a total of m unique parameters. The data can again be arranged such that the parameters are correctly calculated. In essence each actuator is assumed to be at the center of a sufficiently large sheet. Neither Toeplitz model favorably captures edge effects as well as the centrosymmetric model. Figure 3.11 shows the least squares calculation of all four gain matrices discussed in this section.

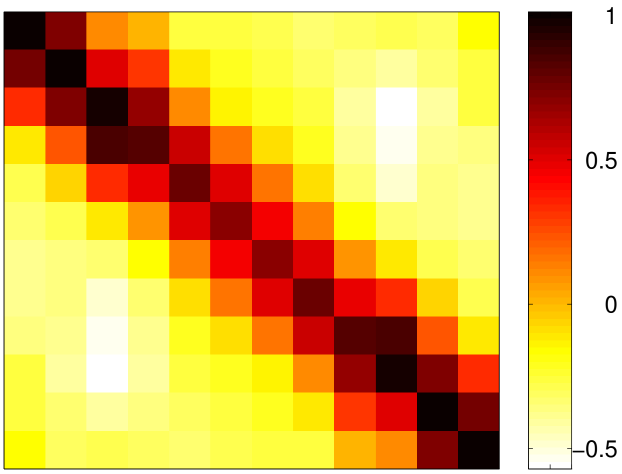

Figure 3.10: Gain matrix, centrosymmetry enforced.

$$K_{\text{Toeplitz}} = \begin{bmatrix} k_0 & k_1 & \cdots & k_j & \cdots & k_m \\ k_{-1} & k_0 & k_1 & \cdots & k_j & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & k_j \\ k_{-j} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & k_{-j} & \cdots & k_{-1} & k_0 & k_1 \\ k_{-m} & \cdots & k_{-j} & \cdots & k_{-1} & k_0 \end{bmatrix} \tag{3.18}$$

Certainly there is some symmetry associated with this process, although care must be taken not to enforce symmetry that is not present. Figure 3.12 shows the cumulative square error of the residuals from the 36 experiments used to determine the various gain matrices. In the absence of noise and process symmetry one would expect the residuals to be larger as more and more symmetry is enforced. Consequently, if the process exhibits strong symmetrical behavior then each model containing the symmetry as a subset of its parameter space would have equivalent residuals. For example, a free parameter model with no symmetry contains both centrosymmetric

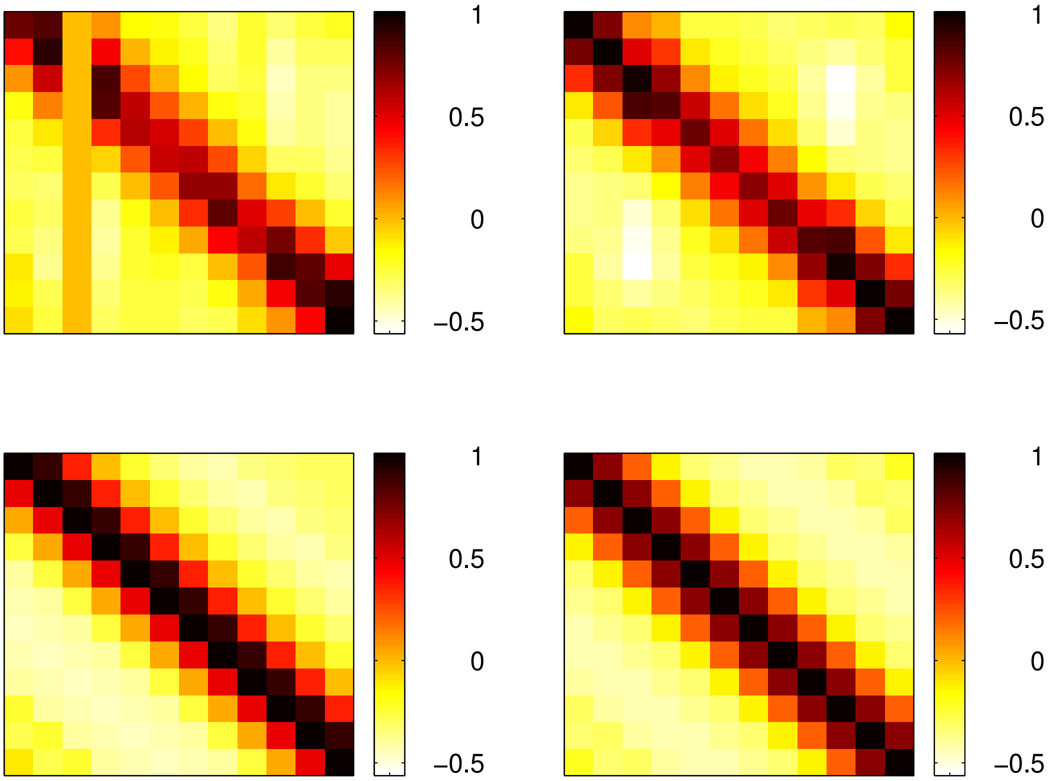

Figure 3.11: Gain matrices calculated from the 3M data set. Shown here are the gains calculated with no symmetry enforced (upper left), centrosymmetry enforced (upper right), Toeplitz structure enforced (lower left), and Toeplitz symmetric structure enforced (lower right).

and Toeplitz symmetric alternatives as part of its solution set. In the presence of noise, however, the true symmetry of the process may be difficult to determine.

In the case of the 3M data, the residuals follow an interesting pattern. The centrosymmetric model with 72 free parameters provides higher residuals than the Toeplitz structure with 23 free parameters. There are several interpretations of this result. First in calculating the centrosymmetric model the general strategy involves estimating the response of the first 6 actuators and mirroring the data from actuators 7 through 12 onto 1 through 6. In determining the Toeplitz structure the general strategy is to shift each of the actuator responses on top of each other. Perhaps the mirroring enforces symmetry that is not present, thus the residuals are higher. A system that does not contain centrosymmetry may have a die lip problem or a flow problem. The magnitudes of the residuals might also indicate that the data is noisy and that 36 samples are not enough to obtain quality parameter estimates.

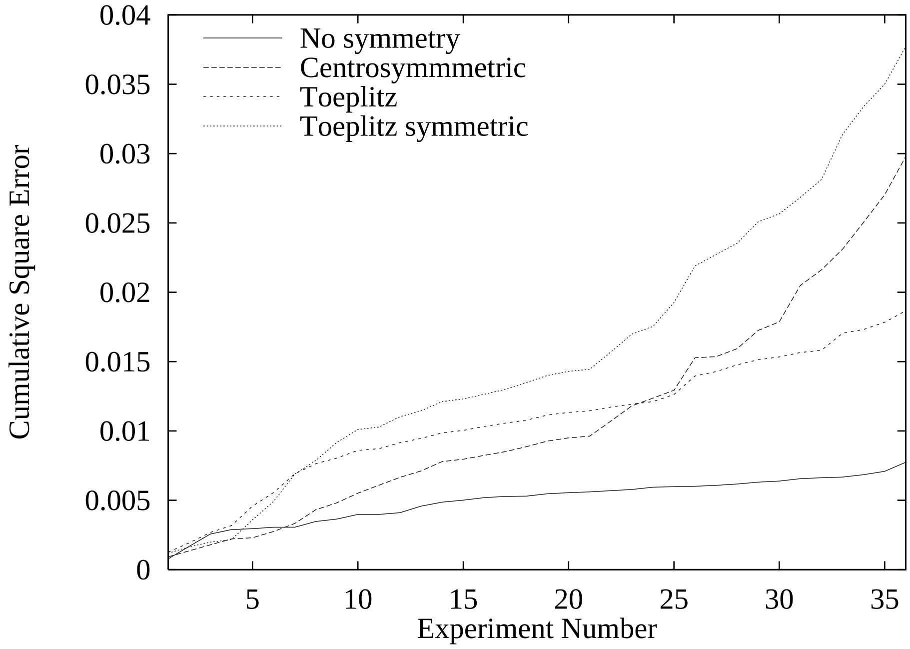

Figure 3.12: Residuals for the various model types.

3.5 Summary

This chapter proposed a model structure that handles large time–delay, zero dynamics and a scanning sensor. The model parameters were determined using data from a 3M pilot plant. Data from all the actuators provides a well conditioned problem that can be treated as a least squares parameter estimation problem. In this study, symmetry was enforced to overcome the problem of missing data.

Chapter 4

State Estimation

In model predictive control calculations, one of the key initial conditions is the initial value of the state. The state may be a mathematical construction as in the case of a realized transfer function, the state may be unmeasured but have physical meaning, or the state may be available directly as a measured quantity. In any of these cases, determining the true value of the state is impossible and some estimation technique is required.

Solving the estimation problem is a key step in the control of gage processes. Filtering in not merely used to counteract the effects of noise and provide a smooth output estimate, rather the entire state of the film must be reconstructed from the sparse measurements. Often this reconstruction entails estimating a cross directional profile from a single measurement and a past value of the state. It is important that the chosen filtering strategy take into account the periodic nature of the sensor. The stability properties of the estimate are also important. This chapter presents the

Kalman filter as a potential solution to the gage estimation problem. Lifted systems are presented as a potential method to calculate periodic Kalman filter gains. A lifted system is a time–invariant equivalent to a periodic or multirate system. The stability properties of the estimate error are also discussed. Lastly, an analysis of the covariance of the estimates is presented providing some interesting results. The calculation of state estimates from data is a classic one and the evolution of estimation science is quite interesting. Therefore, this chapter begins with a brief discussion of the history of estimation theory.

4.1 Estimation History

In 1795 Karl Friedrich Gauss used the least–squares method as part of his astronomical studies. In studying the motion of planets and comets, Gauss, then 18 years old, used least–squares estimation to deduce the six parameters that completely characterized the motion of the planets. Gauss later published a detailed description of the least–squares estimation method in his book *Theoria Motus Corporum Coelestium.* The following quote is from Gauss' *Theoria Motus* which was published in 1809 [110].

> "If the astronomical observations and other quantities on which the computation of orbits is based were absolutely correct, the elements also, whether deduced from three or four observations, would be strictly accurate (so far indeed as the motion is supposed to take place exactly according to the laws of Kepler) and, therefore, if other observations were

> used, they might be confirmed but not corrected. But since all our measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all computations made concerning concrete phenomena must be to approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the unknown quantities. This problem can only be properly undertaken when an approximate knowledge of the orbit has been already attained, which is afterwards to be corrected so as to satisfy all the observations in the most accurate manner possible."

In this quote Gauss mentions several key aspects of estimation theory that are still studied today. Redundant data, dynamic modelling, and minimizing residuals are all issues touched by Gauss. In addition some of the more elegant derivations in estimation theory arise from the probabilistic considerations alluded to in this quote. So Gauss provides a succinct statement of what has become an entire branch of theory.

R.A. Fischer introduced the maximum likelihood method in 1912. In this theory the problem becomes one of determining the most probable estimate based on a known distribution of values. Gauss originally rejected the maximum likelihood method in favor of minimizing the residuals. Gauss derived least squares independent of probability theory although it falls neatly out of the probabilistic derivation of maximum likelihood. In 1941 Kolmogorov, and in 1942 Wiener, independently developed a linear minimum mean–square estimation technique. The advances made by Wiener and Kolmogorov included the ability to handle both continuous and discrete time data in addition to non constant signals (something that did not concern Gauss). The solution method becomes unwieldy for large numbers of measurements, and it assumes an infinite data set. In 1955 J.W. Follin suggested a recursive approach based on the idea that the previous estimate contains the information regarding the previous measurements. This work immediately preceded what we now call the Kalman filter [110].

4.2 The Kalman Filter

The Kalman filter can be derived in several different ways. Originally Kalman derived it geometrically using an orthogonal projection argument [70]. Probability densities, which lead to optimal estimates, can be calculated using a maximum likelihood or Bayesian point of view [65]. Probabilistic notions can be ignored and the problem can be attacked statistically as either a least squares problem or a recursive least squares problem [65]. See Appendix A for more discussion regarding the derivation of the Kalman filter.

This problem is traditionally cast as follows: given a linear disturbance model, information describing the noise distributions, and data up to time k, determine the optimal estimate of the state at either time k or time $k+1$. For the Kalman filter, the calculations and many properties of the solution are still valid even when the system is time–varying. The details and proofs are available in numerous books on stochastic control and optimal filtering (e.g. Bryson and Ho; Åström [22, 3]).

Consider the following time–invariant state–space model

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + G_\omega \omega_k \\ y_k &= Cx_k + \nu_k \end{aligned} \tag{4.1}$$

which consists of two noise terms added to a linear deterministic model. The state noise, ω_k, is a vector in $\mathbb{R}^n$. The measurement or sensor noise, ν_k, is a vector in $\mathbb{R}^p$. Typically ω_k and ν_k are assumed to be mutually–independent, zero–mean white–noise random variables with covariance,

$$\mathcal{E}\left\{ \begin{bmatrix} \omega_k \\ \nu_k \end{bmatrix} \begin{bmatrix} {\omega_k}^T & {\nu_k}^T \end{bmatrix} \right\} = \begin{bmatrix} Q_\omega & 0 \\ 0 & R_\nu \end{bmatrix} \tag{4.2}$$

In addition, the initial state, x_0, is also an independent normally–distributed variable with covariance Q_0. It can be shown that with these assumptions the discrete Kalman filter provides the optimal or minimum variance estimate of the state [65].

Eq. 4.3 provides the expression for the updated state. This equation introduces new notation in which $\hat{x}$ is the estimate of x and the indexing $\hat{x}_{k|k}$ is read as follows, "the estimate of x at time k given the information up to time k."

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k \left(y_k - C\hat{x}_{k|k-1} \right) \tag{4.3}$$

To provide the optimal estimate, the filter gain, L_k, scales a term that represents model error. This error is the difference between the measurement and the model's prediction of the measurement. Eq. 4.3 can be combined with the state evolution equation (Eq. 4.1) to obtain the following:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + AL_k\left(y_k - C\hat{x}_{k|k-1}\right) \tag{4.4}$$

The filter gain matrix, L_k, can be computed efficiently in terms of the estimate error covariance matrix, P_k,

$$L_k = P_k C^T\left[CP_kC^T + R_\nu\right]^{-1} \tag{4.5}$$

in which P_k can be computed efficiently from the following Riccati iteration with initial condition, $P_0 = Q_0$,

$$P_{k+1} = AP_kA^T + G_\omega Q_\omega {G_\omega}^T - AP_kC^T\left[CP_kC^T + R_\nu\right]^{-1} CP_kA^T \tag{4.6}$$

P_k is the covariance matrix of the estimate and provides a quantitative means of evaluating the quality of the estimate.

The Riccati equation converges for large k to provide a steady–state covariance matrix which can, in turn, be used to calculate a time–invariant steady–state discrete Kalman filter gain, L. The Kalman filter is stable provided (A, C) is detectable, $(A, {Q_\omega}^{1/2})$ is stabilizable, R_ν is positive definite, and Q_0 is positive definite [11]. For the nominal case, in the absence of state or measurement noise, a stable estimator drives the reconstruction error to zero. The reconstruction error is the difference between the actual state and the estimated state.

$$e_{k|k-1} = x_k - \hat{x}_{k|k-1} \tag{4.7}$$

The estimate error evolves as follows.

$$e_{k+1|k} = (A - ALC)e_{k|k-1} \tag{4.8}$$

Consequently, a stable filter gain provides a time–invariant error evolution matrix $(A - ALC)$ that is Hurwitz, and the estimate error is exponentially stable.

4.2.1 Kalman Predictor

It is worth noting the semantic differences between the Kalman filter and Kalman predictor. According to Bitmead, Givers, and Wertz [11] the Kalman *predictor* provides the optimal estimate given the inputs $\{u_{k-1}, u_{k-2}, \ldots, u_0\}$ and the outputs $\{y_{k-1}, y_{k-2}, \ldots, y_0\}$. Thus the recursive form of the Kalman predictor is as follows,

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + AL_k\left(y_k - C\hat{x}_{k|k-1}\right) \tag{4.9}$$

The Kalman *filter* provides the optimal estimate given the inputs $\{u_{k-1}, u_{k-2}, \ldots, u_0\}$ and the outputs $\{y_k, y_{k-1}, \ldots, y_0\}$. Therefore the Kalman filter yields

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k\left(y_k - C\hat{x}_{k|k-1}\right) \tag{4.10}$$

which if advanced one step becomes

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + L_{k+1}\left(y_{k+1} - C\hat{x}_{k+1|k}\right) \tag{4.11}$$

Now, using the following state evolution equation

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k \tag{4.12}$$

and substituting into Eq. 4.11 provides the final form of the recursive Kalman filter formula

$$\hat{x}_{k+1|k+1} = A\hat{x}_{k|k} + Bu_k + L_{k+1}\left(y_{k+1} - CA\hat{x}_{k|k} - CBu_k\right) \tag{4.13}$$

Noting that the chosen way to present this problem in Section 4.2 used the Kalman filter equation (Eq. 4.3) and the Kalman predictor equation (Eq. 4.4) one can see where some confusion may arise.

4.3 Time–Varying Kalman Predictor

The model given by Eq. 4.1 can be rewritten for a time–varying C_k matrix. The analogs to Eqs. 4.1 through 4.6 are outlined here starting with the model.

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + G\omega_k \qquad (4.14) \\ y_k &= C_k x_k + \nu_k \end{aligned}$$

This provides a time–varying filter gain L_k and covariance matrix P_k, and is completely analogous to the time–invariant case.

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + L_k\left(y_k - C_k\hat{x}_{k|k-1}\right) \tag{4.15}$$

The filter gain matrix, L_k, can be computed efficiently in terms of the estimate error covariance matrix, P_k,

$$L_k = AP_kC_k^T \left[C_kP_kC_k^T + R_v\right]^{-1} \tag{4.16}$$

in which P_k can be computed efficiently from the following Riccati iteration,

$$P_{k+1} = AP_kA^T + GQ_\omega G^T - AP_kC_k^T \left[C_kP_kC_k^T + R_\nu\right]^{-1} C_kP_kA^T \tag{4.17}$$

Due to the periodic nature of the film and sheet problem induced by the measurement device, a finite number of Kalman filter gains and covariance matrices are computed. In addition, depending on the scanning pattern, there may also be some symmetry associated with these gains.

4.3.1 Stability

The filter error for the time varying case can be analyzed as follows. Let $P(k; P_0, k_0)$ be the solution to Eq. 4.17 with initial data $P_{k_0} = P_0$. As stated in Kwakernaak and Sivan [77] it follows that if the system is reconstructible, then $P(k; P_0, k_0)$ converges to a unique asymptotic solution, $P^*(k)$, as $k_0 \to -\infty$ for every $P_0 \geq 0$. This solution also satisfies Eq. 4.17. The periodic C_k matrix makes P_k^* periodic with period q also, which is easily established. Since in Eq. 4.17 the C_k matrix is periodic with period q, it follows that

$$P(k; P_0, k_0) = P(k + q; P_0, k_0 + q), \qquad k \geq k_0$$

Taking limits as $k_0 \to -\infty$ and the uniqueness of P_k^* yields $P_k^* = P_{k+q}^*$ for every k.

The periodic P_k^* then produces periodic asymptotic filter gains L_k^* via Eq. 4.16. The asymptotic filter is then shown to be exponentially stable. From Eqs. 4.16–4.17 (dropping the asterisk denoting optimality),

$$(A - L_0C_0)P_0(A - L_0C_0)^T = P_1 - H_0 \tag{4.18}$$

in which $H_0 = Q + L_0RL_0^T$ is positive definite. Multiplying from the right and left by $(A - L_1C_1)$ gives

$$(A - L_1C_1)(A - L_0C_0)P_0(A - L_0C_0)^T(A - L_1C_1)^T = P_2 - H_1 \tag{4.19}$$

in which $H_k = (A - L_kC_k)H_{k-1}(A - L_kC_k)^T + Q + L_kRL_k^T$ remains positive definite. If we continue in this way for one period, the periodicity of P_k produces finally

$$\prod_{i=0}^{i=q-1}(A - L_iC_i)\, P_0 \prod_{i=0}^{i=q-1}(A - L_iC_i)^T = P_0 - H_{q-1} \tag{4.20}$$

We conclude the product, $\prod_{i=0}^{i=q-1}(A - L_iC_i)$, is exponentially stable since it satisfies an appropriate Lyapunov equation. This result establishes that the norm of the reconstruction error with the asymptotic filter can be bounded above by an exponentially decreasing sequence. This result can be relaxed for *detectable* periodic systems and is discussed in Section 6.3 after lifted systems are presented.

4.4 Calculation of Periodic Filter Gains

The set of filter gains for a periodic system can be calculated in two ways. First, iteration of the Riccati equation given in Eq. 4.17 along with dutiful accounting provides the set of q covariance matrices. A second method, which is also a useful theoretical tool, involves reformulating the problem as a time–invariant problem. The new system is called a *raised* of *lifted* system. The general idea is that the output and input vectors are enlarged and the state is only considered at the end of each period. Many properties of time–invariant systems apply to the lifted system and thus to the periodic system. These results are presented as needed. Lifted systems are frequently used for multirate systems as well as periodic systems [1]. The gage problem is essentially a multirate problem since the state matrix, A, is time–invariant. Presented here is the lifted system for the gage model as given by Tyler and Morari [113].

Theorem 4.1 (Lifted Systems) *Given the following periodic model*

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + G_\omega \omega_k \\ y_k &= C_k x_k + \nu_k \end{aligned} \tag{4.21}$$

in which $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y, \nu \in \mathbb{R}^p, \omega \in \mathbb{R}^{s_\omega}, p$ *is the number of sensors and* C_k *has period* q *then there is a corresponding lifted system*

$$\begin{aligned} \mathcal{X}_{j+1} &= \mathcal{A}\mathcal{X}_j + \mathcal{B}\mathcal{U}_j + \mathcal{G}_\omega \mathcal{W}_j \\ \mathcal{Y}_j &= \mathcal{C}\mathcal{X}_j + \mathcal{D}\mathcal{U}_j + \mathcal{M}_j \end{aligned} \tag{4.22}$$

Proof: In order to remove the time dependence in the model the lifted state is defined to be the qth occurrence of the state given in Eq. 4.21.

$$\mathcal{X}_j = x_{jq} \tag{4.23}$$

thus the corresponding state matrix is given by

$$\mathcal{A} = A^q \tag{4.24}$$

Next the following vectors are needed to construct the system.

$$\mathcal{U}_j = \begin{bmatrix} u_{jq} \\ u_{jq+1} \\ \vdots \\ u_{jq+q-1} \end{bmatrix} \qquad \mathcal{Y}_j = \begin{bmatrix} y_{jq} \\ y_{jq+1} \\ \vdots \\ y_{jq+q-1} \end{bmatrix} \tag{4.25}$$

$$\mathcal{V}_j = \begin{bmatrix} \nu_{jq} \\ \nu_{jq+1} \\ \vdots \\ \nu_{jq+q-1} \end{bmatrix} \qquad \mathcal{W}_j = \begin{bmatrix} \omega_{jq} \\ \omega_{jq+1} \\ \vdots \\ \omega_{jq+q-1} \end{bmatrix} \tag{4.26}$$

Advancement of Eq. 4.21 and collection of terms leads to the following matrices

$$\mathcal{B} = \begin{bmatrix} A^{q-1}B & A^{q-2}B & \dots & B \end{bmatrix} \tag{4.27}$$

$$\mathcal{G}_\omega = \begin{bmatrix} A^{q-1}G_\omega & A^{q-2}G_\omega & \dots & G_\omega \end{bmatrix} \tag{4.28}$$

$$\mathcal{C} = \begin{bmatrix} C_0 \\ C_1 A \\ \vdots \\ C_{q-1}A^{q-1} \end{bmatrix} \tag{4.29}$$

$$\mathcal{D} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ C_1 B & 0 & 0 & \cdots & 0 \\ C_2 AB & C_2 B & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ C_{q-1}A^{q-2}B & C_{q-1}A^{q-3}B & \cdots & C_{q-1}B & 0 \end{bmatrix} \tag{4.30}$$

$$\mathcal{H} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ C_1 G_\omega & 0 & 0 & \cdots & 0 \\ C_2 AG_\omega & C_2 G_\omega & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ C_{q-1}A^{q-2}G_\omega & C_{q-1}A^{q-3}G_\omega & \cdots & C^{q-1}G_\omega & 0 \end{bmatrix} \tag{4.31}$$

It is next convenient to define $\mathcal{M} = \mathcal{H}\mathcal{W} + \mathcal{V}$. Notice that this is a linear combination of stochastic variables. The covariance of the lifted stochastic variables is given by

$$\mathcal{E}\left\{ \begin{bmatrix} \mathcal{G}_\omega \mathcal{W} \\ \mathcal{M} \end{bmatrix} \begin{bmatrix} (\mathcal{G}_\omega \mathcal{W})^T & \mathcal{M}^T \end{bmatrix} \right\} = \begin{bmatrix} \mathcal{Q} & \mathcal{T} \\ \mathcal{T}^T & \mathcal{R} \end{bmatrix} \tag{4.32}$$

in which

$$\mathcal{Q} = \mathcal{G}_\omega \begin{bmatrix} Q_\omega & 0 & \cdots & 0 \\ 0 & Q_\omega & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_\omega \end{bmatrix} \mathcal{G}_\omega^T \tag{4.33}$$

$$\mathcal{R} = \mathcal{H} \begin{bmatrix} Q_\omega & 0 & \cdots & 0 \\ 0 & Q_\omega & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_\omega \end{bmatrix} \mathcal{H}^T + \begin{bmatrix} R_\nu & 0 & \cdots & 0 \\ 0 & R_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & R_\nu \end{bmatrix} \tag{4.34}$$

$$\mathcal{T} = \mathcal{G}_\omega \begin{bmatrix} Q_\omega & 0 & \cdots & 0 \\ 0 & Q_\omega & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_\omega \end{bmatrix} \mathcal{H}^T \tag{4.35}$$

The lifted system is now completely specified. □

Using the lifted system in Eq. 4.22 and the Riccati iteration given by Eq. 4.6 the first of the steady–state covariance matrices, P_k can be calculated. The remainder of the covariance matrices are calculated by iterating Eq. 4.17.

4.5 Covariance Matrices

It is illustrative to analyze the covariance matrices in the context of the gage estimation problem in order to gain some insight about the quality and characteristics of the estimates. Tyler and Morari [113] define an effectiveness factor for the process in an effort to determine the potential gain of adding sensors. While no effectiveness factors are used here, the general idea of examining covariance matrices in detail is illustrated.

The Kalman filter is based on the assumption that the estimate is normally distributed. The following expression of the error is distributed χ^2 with n degrees of freedom [6].

$$(x - \hat{x})^T P^{-1} (x - \hat{x}) \tag{4.36}$$

In terms of the parameter α the probability of an event can be specified. Eq. 4.37 specifies the calculation of the probability of an event.

$$\Pr\{(x - \hat{x})^T P^{-1} (x - \hat{x}) \leq \chi_n^2(\alpha)\} = 1 - \alpha \tag{4.37}$$

Regions of constant probability over the parameter space are defined by hyperellipsoids with dimension equal to the number of parameters. Consider the situation in which two states are estimated. The corresponding covariance matrix, P, defines an elliptical region of constant probability.

$$(x - \hat{x})^T P^{-1} (x - \hat{x}) = \chi_n^2(\alpha) \tag{4.38}$$

The directions of the elliptical axes are defined by the eigenvectors of P. The lengths of the elliptical axes are defined by the eigenvalues of P. A convenient, yet somewhat conservative, method of presenting confidence intervals on parameters is based on calculating the hyperbox that circumscribes the hyperellipsoid [45]. Thus it is useful to know that the half width of the side of the box that circumscribes the elliptical region is proportional to the square root of the diagonal elements of the covariance matrix. For this study, it is sufficient to realize that the $P_{i,i}$ element of the covariance matrix provides a measure of the certainty of the ith parameter.

Figure 4.1 illustrates a region of constant probability for a two parameter system [83]. Analysis of the diagonal elements of the covariance matrix provides

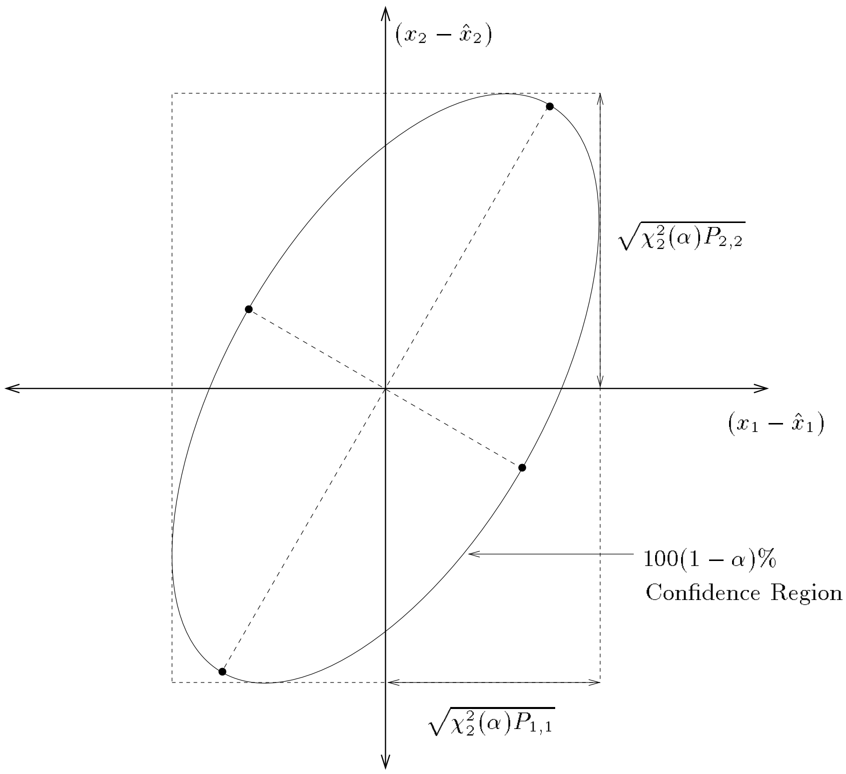

Figure 4.1: Ellipse of constant probability density

some useful insight into the effect of the scanning sensor. Consider the following examples.

Example I: Consider a 12 input 12 output model. The time–delay is not considered here as it does not effect the results of the optimal gain calculation unless an input disturbance model is used[1]. The scanner follows a traditional path starting in lane 1, traveling to lane 12, and then returning to lane 1. Thus there are 22 filter gains to calculate, 12 in the forward direction and 10 in the return direction.

The model in Eq. 4.39 with $A = .9I$ is used. The normal assumption in gage problems is zero dynamics, but analyzing non–zero dynamics in this example is insightful and leads to a better understanding of the next examples.

$$\begin{aligned} x_{k+1} &= Ax_k + G_\omega \omega_k \\ y_k &= C_k x_k + \nu_k \end{aligned} \tag{4.39}$$

Covariances for the noise are set to $Q_\omega = I$ and $R_\nu = 1$. The structure chosen for the state disturbance matrix, G_ω, is diagonally dominant and Toeplitz symmetric. This follows the assumption that disturbances that affect a given state, where the state is the thickness in a given lane, most likely effect neighboring states. The following

[1]A peculiar consequence of storing past inputs in the state vector when using the input disturbance model is that the filter updates the stored inputs. See Chapter 5 for more details on disturbance models.

disturbance matrix is used

$$G_\omega = \begin{bmatrix} 1.0 & .8 & .6 & .2 & & & & & & & & \\ .8 & 1.0 & .8 & .6 & .2 & & & & & & & \\ .6 & .8 & 1.0 & .8 & .6 & .2 & & & & & & \\ .2 & .6 & .8 & 1.0 & .8 & .6 & .2 & & & & & \\ & .2 & .6 & .8 & 1.0 & .8 & .6 & .2 & & & & \\ & & .2 & .6 & .8 & 1.0 & .8 & .6 & .2 & & & \\ & & & .2 & .6 & .8 & 1.0 & .8 & .6 & .2 & & \\ & & & & .2 & .6 & .8 & 1.0 & .8 & .6 & .2 & \\ & & & & & .2 & .6 & .8 & 1.0 & .8 & .6 & .2 \\ & & & & & & .2 & .6 & .8 & 1.0 & .8 & .6 \\ & & & & & & & .2 & .6 & .8 & 1.0 & .8 \\ & & & & & & & & .2 & .6 & .8 & 1.0 \end{bmatrix} \qquad (4.40)$$

The square root of the diagonal elements of the covariance matrix, (P_0), are plotted in Figure 4.2. This Figure shows the relative confidence regions on the 12 state estimates when the scanner is in the first lane. As expected the estimate with the tightest confidence region is in the lane where the sensor is located. It would be logical to assume that while the sensor is in the first lane the uncertainty of the state estimates would increase as the distance from the sensor increases. This is not the case as the confidence regions of the 11th and 12th lanes are noticeably lower than the confidence region of the 10th lane. Figure 4.3 shows the confidence regions associated with P_0, P_1, and P_2. Again, the lane that is home to the sensor provides the tightest confidence region. One additional way to view this data is to watch the confidence regions for a given lane change over time. This is shown in Figure 4.4 for the first three estimates. Notice that the lowest values of each curve correspond to the presence of the sensor. For example, at times two, 21, 23, and 43 the sensor is in lane two. Here it is clear that there is almost always more confidence in the value of the first state.

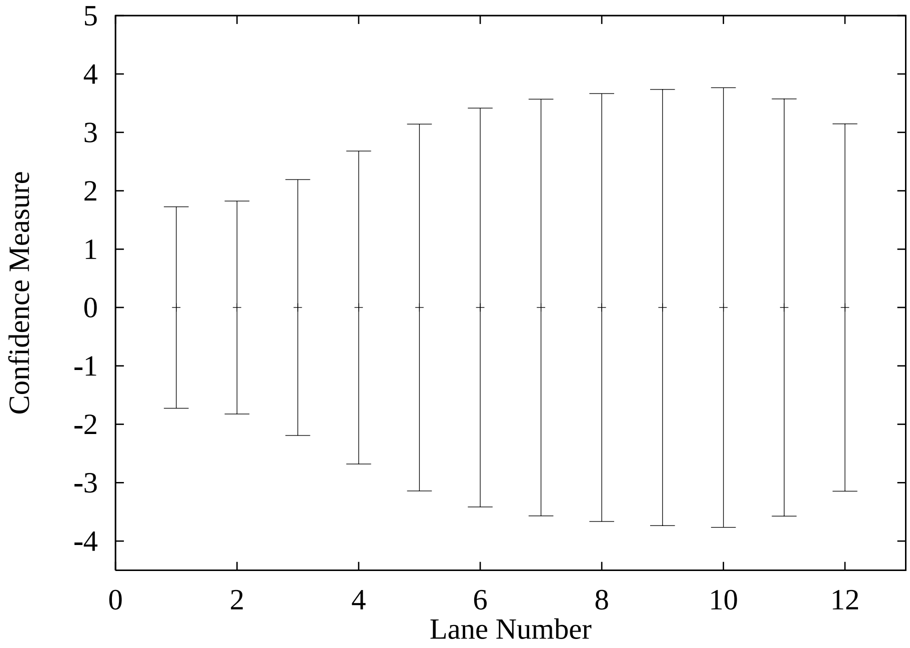

Figure 4.2: Confidence Regions ($\sqrt{P_{ii}}$) for the 12 state estimates when the the sensor is in lane one.

The counterintuitive values of the confidence region relating to the 12th lane in Figures 4.2 and 4.3 and the consistently lower confidence regions of state one (relative to state two and three) in Figure 4.4 can be explained by the model. When there is an error between the measurement and the model prediction the optimal estimation scheme basically finds the most probable origin of the error (*e.g.*the source

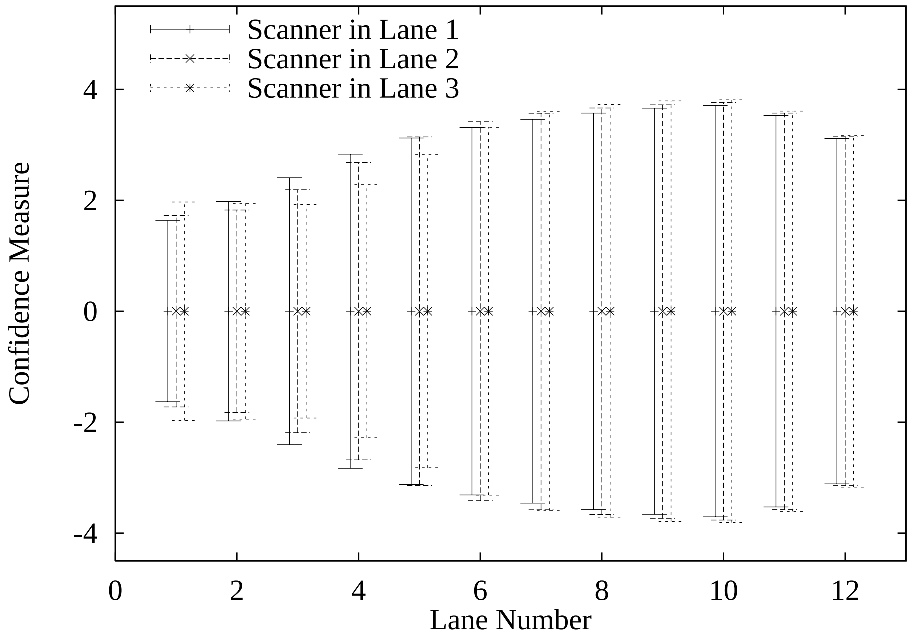

Figure 4.3: Confidence Regions ($\sqrt{P_{ii}}$) for the 12 state estimates when the the sensor is in lanes one, two, three.

that minimizes the error in some prescribed manner). The model in Eq. 4.39 shows that noise can add to the state through a dynamic matrix, G_ω or it can add directly to the output. Based on the structure of G_ω it should be clear that only disturbances from $\omega_9 - \omega_{12}$ can effect the state x_{12}, however, the disturbance states $\omega_7 - \omega_{12}$ can effect state x_{10}. The states in the middle of the sheet are effected by many more of the unknown stochastic variables and, therefore, have more uncertainty associated with them. This is a realistic model if disturbances truly come through the system in such a correlated manner.

This example raises another question which can be answered by simulation: if the edges have adequate confidence regions associated with their state estimates, can the edge measurements be skipped altogether.

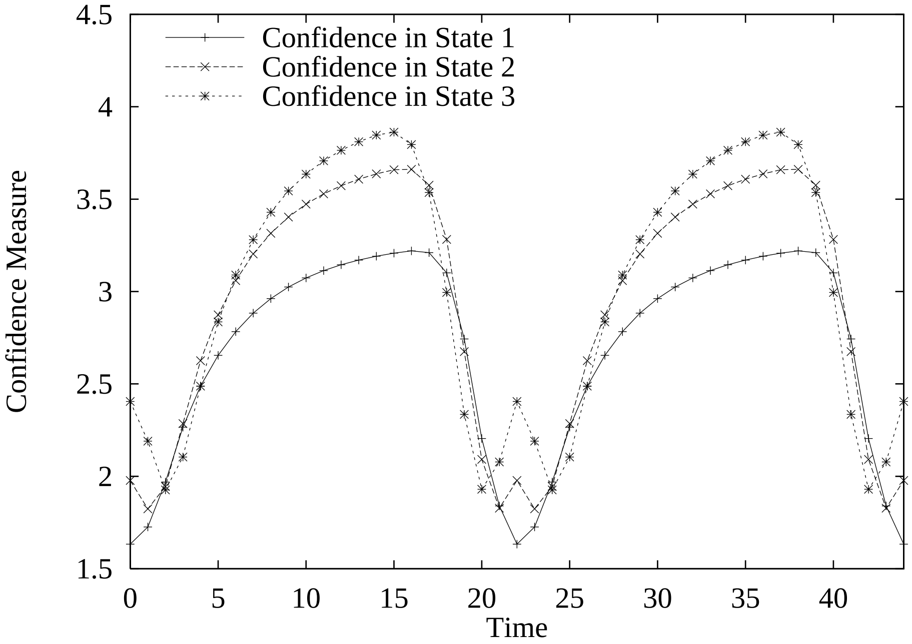

Figure 4.4: Confidence Regions ($\sqrt{P_{ii}}$) for the first three states versus time. Note that the period is 22.

Example II: Motivated by the previous example, the scanning pattern is now changed to exclude lanes one and 12. Thus there are now 18 Kalman gains and covariance matrices. Plots of the type shown in Example I show that this scanning pattern exhibits similar behavior to the previous scanning pattern with slightly lower confidence regions. This is best shown by calculating the volume of the hyperbox surrounding the hyperellipsoid. Figure 4.5 shows the normalized confidence regions over several periods. By not scanning all the way to the edge, the sensor travels to each lane more frequently and the confidence regions are smaller. Thus it may be of interest to determine the proper balance between full scans and parking the sensor, but this does not take into account integrating modes. Integrating modes are the topic of the next example.

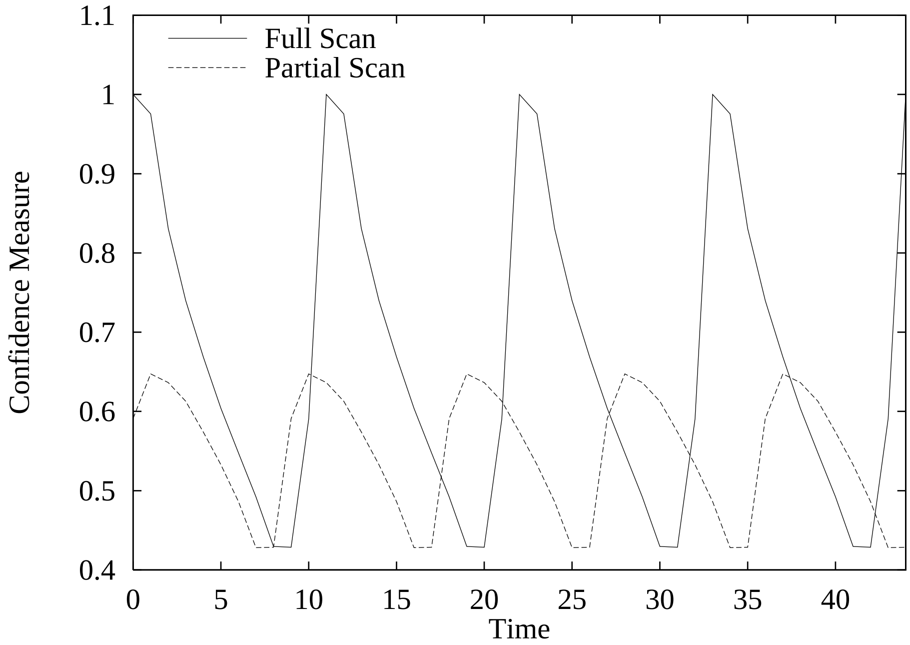

Figure 4.5: Normalized volume of the 12–dimensional hyperbox for a full 22 sample period scan and a partial 18 sample period scan that ignores lanes one and 12.

Example III: The use of integrating modes is covered in detail in Chapter 5. For now it is sufficient to say that integrating modes are desired to provide offset free control and that the reader is encouraged to revisit this example after the next chapter. Given the following model with an output disturbance state,

$$\begin{aligned} x_{k+1} &= Ax_k + G_\omega \omega_k \\ z_{k+1} &= z_k + \xi_k \\ y_k &= Cx_k + G_p z_k + \nu_k \end{aligned} \tag{4.41}$$

and let $A = 0$, $G_p = I$, $G_\omega = I$, $Q_\omega = I$, $Q_\xi = 0$, and $R_\nu = 1$. An example with similar tuning is provided by Rawlings and Chien [96]. The confidence regions on the states (x_k) are constant because there are no dynamics in the process. The confidence regions on the integrating disturbance states, however, are integrating. Figure 4.6 shows the evolution over time of the confidence regions in each of the first three lanes. The uncertainty grows unchecked until the sensor passes through each lane. It is clear that if the disturbance model given here is used with the modified scanning pattern in Example II the estimates would be unstable. This raises the issue of detectability which is discussed in the next chapter. Example III provides the more intuitive trends associated with confidence regions: the longer the time between measurements the larger the uncertainty in the estimate.

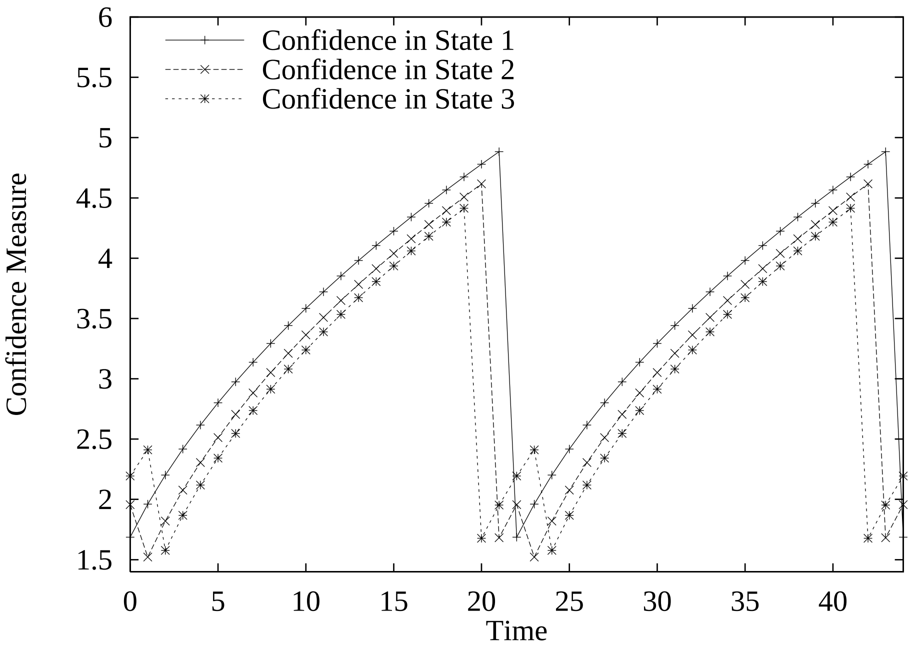

Figure 4.6: Confidence Regions ($\sqrt{P_{ii}}$) for the first three disturbance states versus time. Note that the period is 22.

4.6 Summary

The Kalman filter is a convenient optimal estimation tool to use for gage control. The periodic nature of the scanning sensor provides an estimation problem that is easily solved via standard Kalman filtering techniques. Exponential stability of the estimate error is guaranteed under conditions of detectability.

The actual filter gains can be calculated by iterating the steady–state Riccati equation directly or by lifting the system and then solving the steady–state Riccati equation. A lifted system is an equivalent time–invariant representation of the same system.

The covariance of the state estimates can be analyzed by looking at the diagonal elements. These elements are related to the hyperbox that contains the hyperellipsoid of constant probability. Analyzing the covariance matrices of gage control problems is one way of evaluating potential sensor patterns. Example III in this chapter raises the issue of detectability. The next chapter addresses many of the issues regarding the detectability of periodic systems.

Chapter 5

Disturbance Models

Integrating disturbance models are used with model predictive controllers to obtain offset free control. This chapter introduces a general disturbance model. Contained within this disturbance model is the output disturbance model used by QDMC controllers. Also contained in the general model is the input disturbance model which can be used to model upstream disturbances. A measured input disturbance model is presented as well. An augmented model is constructed to calculate the Kalman filter gains. The combined system should be detectable to ensure that the unstable integral states in the disturbance model are observable. For periodic systems, lifting the system is useful for analyzing system observability and detectability.

5.1 Disturbance models

In order for the controller to remove steady–state offset in the face of modelling errors or nonzero disturbances, a number of integrating disturbance models may be employed. A generic disturbance model is given in Eq. 5.1. This model has an integrating disturbance z that can effect either the state or the output.

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + G_m z_k + G_\omega \omega_k \\
z_{k+1} &= z_k + \xi_k \\
y_k &= C_k x_k + G_p z_k + \nu_k
\end{aligned}
\tag{5.1}
$$

By considering z to be a continuation of the state, an augmented state vector can be used to put the system into the general form of Eq. 4.14.

$$
\underbrace{\begin{bmatrix} x_{k+1} \\ z_{k+1} \end{bmatrix}}_{\tilde{x}_{k+1}} = \underbrace{\begin{bmatrix} A & G_m \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ z_k \end{bmatrix}}_{\tilde{A}\tilde{x}_k} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix} u_k}_{\tilde{B}u_k} + \underbrace{\begin{bmatrix} G_\omega & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \omega_k \\ \xi_k \end{bmatrix}}_{\tilde{G}_\omega \tilde{\omega}_k}
\tag{5.2}
$$

$$
\underbrace{y_k}_{y_k} = \underbrace{\begin{bmatrix} C_k & G_p \end{bmatrix}}_{\tilde{C}_k} \underbrace{\begin{bmatrix} x_k \\ z_k \end{bmatrix}}_{\tilde{x}_k} + \underbrace{\nu_k}_{\nu_k}
\tag{5.3}
$$

The covariance is now given by

$$
\mathcal{E}\left\{ \begin{bmatrix} \omega_k \\ \xi_k \\ \nu_k \end{bmatrix} \begin{bmatrix} {\omega_k}^T & {\xi_k}^T & {\nu_k}^T \end{bmatrix} \right\} = \begin{bmatrix} Q_\omega & 0 & 0 \\ 0 & Q_\xi & 0 \\ 0 & 0 & R_\nu \end{bmatrix}
\tag{5.4}
$$

The filter gains for the augmented system can be calculated using the equations in Section 4.2.

Output disturbance model The output disturbance model is the most frequently used model in industrial applications [91]. An output disturbance model corresponds to $G_m = 0$ in Eq. 5.1. In addition G_p is often chosen such that there is an integrating state in each of the desired outputs. One popular choice of noise characteristics corresponds to a "reliable" sensor and an "unreliable" state (*i.e.* $Q_\omega/R_\nu >> 1$, $Q_\omega/Q_\xi >> 1$). Calculation of the Kalman filter gains provides the following state update equations for the augmented system,

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + AL(y_k - C\hat{x}_{k|k-1}) \tag{5.5}$$

partitioning the augmented system provides the following equations.

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + Bu_k + AL_x(y_k - C\hat{x}_{k|k-1} - \hat{z}_{k|k-1}) \\ \hat{z}_{k+1|k} &= \hat{z}_{k|k-1} + IL_z(y_k - C\hat{x}_{k|k-1} - \hat{z}_{k|k-1}) \end{aligned} \tag{5.6}$$

For $L = [0,\ I]^T$ (a consequence of $Q_\omega/R_\nu >> 1$, $Q_\omega/Q_\xi >> 1$) this reduces to

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + Bu_k \\ \hat{z}_{k+1|k} &= y_k - C\hat{x}_{k|k-1} \end{aligned} \tag{5.7}$$

Eq. 5.7 is the QDMC [49] estimator where at each sample time z is the difference between the measurement and the model estimate. This value is then predicted to be constant in the future for any control calculations.

Input disturbance model The output disturbance model is a popular and easily understand solution to the steady–state offset problem. It is a design choice, however, and, according to Shinskey [106], not a very good one ...

> "The DMC is capable of outperforming the PID controller on setpoint changes but not on load changes introduced upstream of a dominant lag.
>
> In fact, the load always enters upstream of the dominant process time constant and, in many cases, at the same point as the manipulated" variable.

–F.G.Shinskey

An input disturbance model assumes that the error between the measurement and the model's prediction of the output is due to a constant disturbance in the input. An input disturbance model can account for disturbances due to perturbations in feed stock quality, valve biases, and upstream process inconsistencies. In terms of Eq. 5.1, an input disturbance model has $G_p = 0$ and $G_m = B$. The Kalman filter update equations are slightly more complicated (and less intuitive?) than the output disturbance update equations given in 5.5. Perhaps this is one of the reasons it is seldom used in practice.

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + Bu_k + AL(y_k - C\hat{x}_{k|k-1}) \tag{5.8}$$

partitioning the augmented system provides the following equations

$$\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + B(\hat{z}_{k|k-1} + u_k) + L_x(y_k - C\hat{x}_{k|k-1}) && (5.9) \\ \hat{z}_{k+1|k} &= \hat{z}_{k|k-1} + L_z(y_k - C\hat{x}_{k|k-1}) && (5.10) \end{aligned}$$

The output observer, tuned in a specific way, provides the observer used in QDMC. The input observer is not as renown, but can be used to improve performance for certain types of disturbances.

Certainly other choices of G_p and G_m provide a wide range of disturbance models. One other disturbance model of use in gage control is the measured input disturbance model. However it does not fit into the structure presented in Eq. 5.1 and is therefore presented as a separate model.

5.1.1 Measured Input Disturbance Model

The measured input disturbance model is designed for input biases. The disturbance states, in the case of a disabled actuator, grow monotonically. In order to contain the disturbance states so that the disabled actuator can be identified, the estimator must be constrained or used in conjunction with a constrained model predictive controller. The generic disturbance model defined above requires a number of design decisions in order to determine the best place to put the integrating states. The measured input disturbance model, however, requires additional measurements that may not be available to the controller. The actual measurement required is a measurement of the *input*. These measurements can prove to be extremely useful especially in the case of large time–delay systems.

The measured input disturbance model is shown below with z as the estimate of the input bias and $y^m \in \mathbb{R}^m$.

$$\begin{aligned} x_{k+1} &= Ax_k + B(u_k + z_k) + G_\omega \omega_k \\ z_{k+1} &= z_k + \xi_k \\ y_k &= C_k x_k + \nu_k \\ y_k^m &= (u_k + z_k) + \eta_k \end{aligned} \tag{5.11}$$

By considering z to be a state, an augmented state vector can be used to put the system into the general form of Eq. 4.14.

$$\underbrace{\begin{bmatrix} x_{k+1} \\ z_{k+1} \end{bmatrix}}_{\tilde{x}_{k+1}} = \underbrace{\begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ z_k \end{bmatrix}}_{\tilde{A}\tilde{x}_k} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix} u_k}_{\tilde{B}u_k} + \underbrace{\begin{bmatrix} G_\omega & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \omega_k \\ \xi_k \end{bmatrix}}_{\tilde{G}_\omega \tilde{\omega}_k} \tag{5.12}$$

$$\underbrace{\begin{bmatrix} y_k \\ {y^m}_k \end{bmatrix}}_{\tilde{y}_k} = \underbrace{\begin{bmatrix} C_k & 0 \\ 0 & I \end{bmatrix}}_{\tilde{C}_k} \underbrace{\begin{bmatrix} x_k \\ z_k \end{bmatrix}}_{\tilde{x}_k} + \underbrace{\begin{bmatrix} 0 \\ I \end{bmatrix} u_k}_{\tilde{D}\tilde{u}_k} + \underbrace{\begin{bmatrix} \nu_k \\ \eta_k \end{bmatrix} u_k}_{\tilde{\nu}_k} \tag{5.13}$$

Notice that the augmentation produces a D matrix which, similar to the B matrix, can be ignored during the calculation of the Kalman filter gains.

5.2 Lifted Disturbance Models

It is useful to have the corresponding lifted system for both of the aforementioned disturbance models. The equivalent lifted models are of use in both calculating the filter gains and in analyzing the properties of the models.

Theorem 5.1 (Lifted Systems with Disturbances) *Given the following periodic model with a disturbance state*

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + G_m z_k + G_\omega \omega_k \\
z_{k+1} &= z_k + \xi_k \\
y_k &= C_k x_k + G_p z_k + \nu_k
\end{aligned}
\tag{5.14}
$$

in which $x \in \mathbb{R}^n, u \in \mathbb{R}^m, z, \xi \in \mathbb{R}^{s_z}, \nu, y \in \mathbb{R}^p, \omega \in \mathbb{R}^{s_\omega}, p$ is the number of sensors and C_k has period q, then there is a corresponding lifted system

$$
\begin{aligned}
\mathcal{X}_{j+1} &= \mathcal{A}\mathcal{X}_j + \mathcal{B}\mathcal{U}_j + \mathcal{G}_m \mathcal{Z}_j + \mathcal{G}_\omega \mathcal{W}_j \\
\mathcal{Z}_{j+1} &= \mathcal{Z}_j + \mathcal{E}_j \\
\mathcal{Y}_j &= \mathcal{C}\mathcal{X}_j + \mathcal{D}\mathcal{U}_j + \mathcal{G}_c \mathcal{Z}_j + \mathcal{M}_j
\end{aligned}
\tag{5.15}
$$

Proof: This follows the proof in Theorem 4.1 with the following additions.

$$
\mathcal{Z}_j = \begin{bmatrix} z_{jq} \\ z_{jq+1} \\ \vdots \\ z_{jq+q-1} \end{bmatrix} \qquad \mathcal{E}_j = \begin{bmatrix} \xi_{jq} \\ \xi_{jq+1} \\ \vdots \\ \xi_{jq+q-1} \end{bmatrix}
\tag{5.16}
$$

$$
\mathcal{G}_m = \begin{bmatrix} A^{q-1}G_m & A^{q-2}G_m & \dots & G_m \end{bmatrix}
\tag{5.17}
$$

It is convenient to separate the the cross term matrix, $\mathcal{G}_c$, into output disturbance contributions ($\mathcal{G}_p$), and state disturbance contributions ($\mathcal{G}_s$) such that $\mathcal{G}_c = \mathcal{G}_p + \mathcal{G}_s$

$$\mathcal{G}_p = \begin{bmatrix} G_p & 0 & \cdots & 0 \\ 0 & G_p & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & G_p \end{bmatrix} \tag{5.18}$$

$$\mathcal{G}_s = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ C_1 G_m & 0 & 0 & \cdots & 0 \\ C_2 A G_m & C_2 G_m & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ C_{q-1} A^{q-2} G_m & C_{q-1} A^{q-3} G_m & \cdots & C_{q-1} G_m & 0 \end{bmatrix} \tag{5.19}$$

The lifted system is now completely specified. □

Theorem 5.2 (Lifted Systems with Measured Input Disturbances) *Given the following periodic model with an input disturbance state*

$$\begin{aligned} x_{k+1} &= A x_k + B(u_k + z_k) + G_\omega \omega_k \\ z_{k+1} &= z_k + \xi_k \\ y_k &= C_k x_k + \nu_k \\ y_k^m &= (u_k + z_k) + \eta_k \end{aligned} \tag{5.20}$$

in which $x \in \mathbb{R}^n, u, z, \xi \in \mathbb{R}^m, y, v \in \mathbb{R}^p, \omega \in \mathbb{R}^{s_\omega}, p$ *is the number of sensors and* C_k *has period* q*, then there is a corresponding lifted system*

$$\begin{aligned} \mathcal{X}_{j+1} &= \mathcal{A}\mathcal{X}_j + \mathcal{B}(\mathcal{U}_j + \mathcal{Z}_j) + \mathcal{G}_\omega \mathcal{W}_j \\ \mathcal{Z}_{j+1} &= \mathcal{Z}_j + \mathcal{E}_j \\ \mathcal{Y}_j &= \mathcal{C}\mathcal{X}_j + \mathcal{D}(\mathcal{U}_j + \mathcal{Z}_j) + \mathcal{M}_j \\ \mathcal{Y}_j^m &= (\mathcal{U}_j + \mathcal{Z}_j) + \mathcal{N}_j \end{aligned} \tag{5.21}$$

Proof: This follows the proofs in Theorems 4.1 and 5.1 with the following additions.

$$\mathcal{Y}_j^m = \begin{bmatrix} y_{jq}^m \\ y_{jq+1}^m \\ \vdots \\ y_{jq+q-1}^m \end{bmatrix} \qquad \mathcal{N}_j = \begin{bmatrix} \eta_{jq} \\ \eta_{jq+1} \\ \vdots \\ \eta_{jq+q-1} \end{bmatrix} \tag{5.22}$$

The lifted system is now completely specified. □

5.3 Observability and Detectability

Whenever the output measurements are used to reconstruct the state the issues of reconstructibility, observability and detectability arise. Observability allows the determination of the state x from the measurement y. There is a distinction between reconstructibility, which is the ability to determine the state from past measurements, and observability, which is the assurance that the present state can be determined from future measurements and future inputs. It is worth mentioning that this distinction does not hold for continuous–time constant realizations since the continuous state matrix e^{At} is nonsingular [67]. For discrete–time systems reconstructibility need not imply observability, however for non–singular A and time–invariant systems, one implies the other[1]. In the context of this work, the more strict definition of observability is used. A precise definition of observability is given in Kwakernaak and Sivan [77](Theorem 1.31 applied to discrete–time systems) and is shown here in Definition 5.1.

Definition 5.1 (Observability) *Let $y(k; k_0, x_0, u)$ (y at time k given k_0, x_0, u) denote the response of the output variable y_k of the linear difference equation*

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k \\ y_k &= C_k x_k \end{aligned} \tag{5.23}$$

to the initial state $x_{k_0} = x_0$. Then the system is called completely observable if for all k_0 there exists a k_1 with $k_0 < k_1 < \infty$ such that

$$y_{k;k_0,x_0,0} = y_{k;k_0,x_0',0}, \qquad k_0 \leq k \leq k_1, \tag{5.24}$$

implies that $x_0 = x_0'$.

In order to determine if a linear time–invariant system is observable several methods or tests are used. One of which is a rank condition on the observability

[1]There is, of course, the dual counterpart pertaining to controllability and reachability where controllability is the ability to zero the state in finite time from any initial condition, and reachability is the ability to achieve any state in finite time from the origin. Again controllability need not imply reachability.[67]

matrix. This is shown in Lemma 5.1 for time–invariant systems and in Lemma 5.2 for the particular periodic systems of interest.

Lemma 5.1 (Observability Matrix) *The following matrix is the observability matrix of the n–dimensional time–invariant linear system (A, C).*

$$\mathcal{O}(A,C) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{5.25}$$

If $\mathcal{O}(A,C)$ is full rank then the linear system is considered to be observable.

This follows from the Cayley–Hamilton theorem. See Theorem 2 in Sontag [109] which proves the dual problem.

Lemma 5.2 (Periodic Observability Matrix) *The following matrix is the observ-*

ability matrix of the n–dimensional periodic linear system (A, C_k).

$$\mathcal{O}(A, C_k) = \begin{bmatrix} C_0 \\ C_1 A \\ \vdots \\ C_{q-1} A^{q-1} \\ ------- \\ \vdots \\ ------- \\ C_0 A^{q(n-1)} \\ C_1 A^{q(n-1)+1} \\ \vdots \\ C_{q-1} A^{qn-1} \end{bmatrix} \tag{5.26}$$

If $\mathcal{O}(A, C_k)$ *is full rank then the linear system is considered to be observable.*

This follows from the Cayley–Hamilton theorem and Theorem 2 in Sontag. Although this observability matrix is more formidable than the one in Lemma 5.1 the rank condition still ensures that there is no state which cannot be uniquely determined from nq outputs. See also Grasselli and Lampariello [54] for discrete–time periodic conditions of reconstructibility involving the null space of the periodic observability matrix given in Eq. 5.26.

In cases where a system is not observable, the system may still be detectable. Detectability ensures that the unobservable modes are stable. A more precise definition is given in Definitions 5.2 and 5.3.

Definition 5.2 (Detectability) *The linear time–invariant system*

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned} \tag{5.27}$$

is detectable if its unobservable subspace is contained in its stable subspace.

Definition 5.3 (Periodic Detectability) *A similar statement holds true for periodic systems although it may be more precise to say that the system is detectable if its unobservable part is asymptotically stable[12]. While this may seem ambiguous the canonical form introduced next clearly defines the unobservable part of a periodic system.*

One convenient way to calculate detectability is to use a similarity transform to construct the observability canonical form of the system. This is shown in Lemmas 5.3 and 5.4.

Lemma 5.3 (Observability Decomposition) *Assume that the n–dimensional discrete–time linear system (A, C) is not observable. Let dim $\mathcal{O}(A, C) = r < n$. Then there exists an invertible matrix T such that the matrices $\tilde{A} := T^{-1}AT$ and $\tilde{C} := CT$ have the block structure*

$$\tilde{A} = \begin{bmatrix} E & 0 \\ F & G \end{bmatrix}, \qquad \tilde{C} = \begin{bmatrix} H & 0 \end{bmatrix} \tag{5.28}$$

where $E \in \mathbb{R}^{r\times r}$ and $H \in \mathbb{R}^{p\times r}$ and (E, H) is an observable pair.

A proof is given in Sontag [109], Lemmas 3.3.3 and 3.3.4 for the controllability decomposition.

Lemma 5.4 (Periodic Observability Decomposition) *Assume that the n–dimensional periodic discrete–time linear system (A, C_k) is not observable. Let dim $\mathcal{O}(A, C_k) = r < n$. Then there exists an invertible matrix T_k (also with period q) such that the matrices $\tilde{A}_k := T_k^{-1} A T_k$ and $\tilde{C}_k := C_k T_k$ have the block structure*

$$\tilde{A}_k = \begin{bmatrix} E_k & 0 \\ F_k & G_k \end{bmatrix}, \qquad \tilde{C}_k = \begin{bmatrix} H_k & 0 \end{bmatrix} \tag{5.29}$$

where all $E_k \in \mathbb{R}^{r \times r}$ and all $H_k \in \mathbb{R}^{p \times r}$ and (E_k, H_k) is an observable pair.

Lemma 5.4 is proved for the controllable and observable parts of the full periodic Kalman canonical decomposition by both Bittanti and Bolzern, and Grasselli [13, 53]. An interesting discussion by Bittanti and Bolzern [12] points out that the transformation T_k need not be periodic. However, Doležal's theorem ensures that a periodic transformation can be found.

Two other methods of determining observability and detectability can be attributed to Hautus and are shown in Lemmas 5.5 and 5.7. The corresponding periodic tests are shown in Lemmas 5.6 and 5.8. All four proofs are provided in Appendix B and make use of the above definitions and lemmas.

Lemma 5.5 (Hautus Observability Lemma) *The n–dimensional discrete–time linear system (A, C) is observable if and only if*

$$rank \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C} \tag{5.30}$$

Proof: See Appendix B.2. □

Lemma 5.6 (Periodic Observability Lemma) *The n–dimensional periodic discrete–time linear system (A, C_k) with period q is observable if and only if*

$$rank \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C} \tag{5.31}$$

Proof: See Appendix B.3. □

Lemma 5.7 (Hautus Detectability Lemma) *The n–dimensional discrete–time linear system (A, C) is asymptotically observable (detectable) if and only if*

$$rank \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C}, \quad |\lambda| \geq 1 \tag{5.32}$$

Proof: See Appendix B.4. □

Lemma 5.8 (Periodic Detectability Lemma) *The n–dimensional periodic discrete–time linear system (A, C_k) with period q is asymptotically observable (detectable) if and only if*

$$rank \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C}, \quad |\lambda| \geq 1 \tag{5.33}$$

Proof: See Appendix B.5. □

The periodic version of the Hautus Lemmas for observability and detectability provide structural insight into the meaning of periodic observability and detectability. Consider a film or sheet forming process with a scanning sensor. Lemmas 5.6 and 5.8 illustrate that at times when A^q loses rank only *one* of the corresponding sensor positions need ensure that the system is observable. While this insight is useful, a better test involves the lifted system and the time–invariant Hautus test. The next two theorems, the first of which can be found in [1], prove that the lifted system is indeed equivalent to the periodic systems in terms of observability and detectability.

Theorem 5.3 (Congruency of the Lifted and Periodic Systems (Observability)) *Given the periodic model from Eq. 4.21 (ignoring the stochastic terms).*

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= C_k x_k \end{aligned} \tag{5.34}$$

in which $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p, p$ *is the number of sensors and* C_k *has period* q*, and given the corresponding lifted system (as defined in Theorem 4.1)*

$$\begin{aligned} \mathcal{X}_{j+1} &= \mathcal{A}\mathcal{X}_j + \mathcal{B}\mathcal{U}_j \\ \mathcal{Y}_j &= \mathcal{C}\mathcal{X}_j + \mathcal{D}\mathcal{U}_j \end{aligned} \tag{5.35}$$

then (A, C_k) *is observable if and only if the corresponding lifted system* $(\mathcal{A}, \mathcal{C})$ *is observable.*

Proof: Calculation of the periodic analog of the observability matrix shown in Lemma 5.2 yields the following matrix,

$$\mathcal{O}(A, C_k) = \begin{bmatrix} C_0 \\ C_1 A \\ \vdots \\ C_{q-1}A^{q-1} \\ ------- \\ \vdots \\ ------- \\ C_0 A^{q(n-1)} \\ C_1 A^{q(n-1)+1} \\ \vdots \\ C_{q-1}A^{qn-1} \end{bmatrix} \tag{5.36}$$

Calculation of the observability matrix of the lifted system from Eqs. 4.24 and 4.29 provides the same matrix. Since $\mathcal{O}(A, C_k) = \mathcal{O}(\mathcal{A}, \mathcal{C})$ the proof is complete. □

Theorem 5.4 (Congruency of the Lifted and Periodic Systems (Detectability)) *Given the periodic model from Eq. 4.21 (ignoring the stochastic terms).*

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= C_k x_k \end{aligned} \tag{5.37}$$

in which $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p, p$ *is the number of sensors and* C_k *has period* q, *and given the corresponding lifted system (as defined in Theorem 4.1)*

$$\begin{aligned} \mathcal{X}_{j+1} &= \mathcal{A}\mathcal{X}_j + \mathcal{B}\mathcal{U}_j \\ \mathcal{Y}_j &= \mathcal{C}\mathcal{X}_j + \mathcal{D}\mathcal{U}_j \end{aligned} \tag{5.38}$$

then (A, C_k) *is detectable if and only if the corresponding lifted system* $(\mathcal{A}, \mathcal{C})$ *is detectable.*

Proof: The periodic model in Eq. 5.37 is detectable when the condition from Lemma 5.8 is satisfied. This condition is shown in Eq. 5.39.

$$\operatorname{rank} \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C}, \quad |\lambda| \geq 1 \tag{5.39}$$

Eq. 5.39 is also the time–invariant Hautus condition for the lifted system (Lemma 5.7). Therefore, the detectability condition is identical for the two systems and the proof is complete. $\square$

5.4 Detectability of Disturbance Models

With the many definitions, tests, and theorems already presented in this chapter, there is now sufficient background to approach some questions of specific interest to the gage estimation problem. Both a general periodic disturbance model and a specific periodic measured input disturbance model have been presented. Definitions of observability and detectability for both time–invariant and periodic systems have been presented, and in the previous section, the periodic system and its corresponding lifted counterpart were shown to be equivalent in terms of observability and detectability. To reiterate, the original goal was to define conditions under which the disturbance models would be detectable. Detectability of the various disturbance models ensures that the integrating augmented states are observable.

Theorem 5.5 (Detectability of the Disturbance Augmented System) *The periodic disturbance model presented in Eq. 5.1 is detectable if and only if (A, C_k) is detectable and*

$$\begin{bmatrix} I_n - \mathcal{A} & \mathcal{G}_m \\ \mathcal{C} & \mathcal{G}_c \end{bmatrix} \tag{5.40}$$

is full column rank.

Proof: First, the lifted system is constructed without any loss of generality since observability (detectability) of the periodic system is equivalent to observability (detectability) of the lifted system (Lemmas 5.6 and 5.8). Next, an augmented system using Eqs. 5.2 and 5.3 is constructed and the Hautus Lemma of time–invariant systems is applied to the new system.

$$\begin{bmatrix} \lambda I_n - \mathcal{A} & \mathcal{G}_m \\ 0 & (\lambda - 1)I_p \\ \mathcal{C} & \mathcal{G}_c \end{bmatrix} \tag{5.41}$$

Proving the sufficient condition first, assume that the disturbance model presented in Eq. 5.1 is detectable. Then Eq. 5.41 must be rank $(n + s_z)$ for all $|\lambda| \geq 1$. It is clear that for all $\lambda > 1$ this condition is met as the first n columns are rank n due to the Hautus condition on the lifted system and the next s_z columns are rank s_z from the identity matrix that does not coincide with the first n columns. For $\lambda = 1$ full rank of Eq. 5.41 reduces to full rank of Eq. 5.40. The necessary condition follows as the arguments above can be reversed. □

Theorem 5.5 provides a starting point for the analysis of several disturbance models. One immediate consequence of Eq. 5.40 is that no disturbance model can have more disturbance states than the number of outputs and retain detectability. Some of the more popular disturbance models are based on a subset of the model described in Eq. 5.1. Before continuing on to the measured input disturbance model, it is worth looking at some of these models in the context of Theorem 5.5. Consider first the output disturbance model, which is the industrial standard for integral control, and then the input disturbance model.

Output Disturbance Model For the output disturbance model (*i.e.* $G_m = 0$), consider the rank condition in Eq. 5.40.

$$\begin{bmatrix} I_n - \mathcal{A} & 0 \\ \mathcal{C} & \mathcal{G}_p \end{bmatrix} \tag{5.42}$$

This condition limits the number and orientation of the integrating disturbance modes. In other words, there cannot be integrating states that are hidden by the integrating modes of the system. In addition to the requirement that there are less disturbance states than outputs, care must be taken to ensure that the remaining disturbance states can be differentiable from the system's integrators. See section B.6 for further comments regarding this condition.

Input Disturbance Model For the input disturbance model (*i.e.* $G_p = 0, G_m = B$), consider the condition in Eq. 5.40 for $\lambda(A) = 1$.

$$\begin{bmatrix} I_n - \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \tag{5.43}$$

The comments about the output step disturbance model apply to the input step disturbance model as well. The integrating modes must be distinguishable from the disturbance states. Since the integrating modes enter the system at a different point in the input disturbance model, detectability of the output disturbance model does not imply detectability of the input disturbance model. See section B.7 for further comments regarding this condition.

Theorem 5.6 (Detectability of the Measured Disturbance Augmented System) *The periodic disturbance model presented in Eq. 5.11 is detectable if and only if (A, C_k) is detectable.*

Proof: First, the lifted system is constructed without any loss of generality since observability (detectability) of the periodic system implies and is implied by observability (detectability) of the lifted system (Lemmas 5.6 and 5.8). Next, an augmented system using Eqs. 5.12 and 5.13 is constructed and the Hautus Lemma of time–invariant systems is applied to the new system.

$$\begin{bmatrix} \lambda I_n - \mathcal{A} & \mathcal{B} \\ 0 & (\lambda - 1) I_m \\ \mathcal{C} & \mathcal{D} \\ 0 & I_m \end{bmatrix} \tag{5.44}$$

Proving the sufficient condition first, assume that the disturbance model presented in Eq. 5.11 is detectable. Then Eq. 5.44 must be rank $(n + m)$ for all $|\lambda| \geq 1$. It is clear that for all $\lambda \geq 1$ this condition is met as the first n columns are rank n due to the Hautus condition on the lifted system and the next m columns are rank s_z from the identity matrix that coincides with the zero matrix in the first n columns. The necessary condition follows as the arguments above can be reversed. $\square$

5.5 Summary

A general disturbance model is presented in this chapter. Integrating disturbance models are used in conjunction with model predictive controllers to obtain offset free control. Several models of interest are contained in the general model. One of these is the well–known output disturbance model. The output disturbance model is used in industrial MPC applications. A second model of interest is the input disturbance model, which promises improved response to the disturbances that come through the dynamics of the system. A measured input disturbance model is presented as well. Unlike the output and input disturbance models a measured input disturbance model requires additional measurements.

Next, in continuing with the periodic gage control problem, lifted disturbance models are constructed. This greatly aids the analysis of the gage system. Many of linear time–invariant results apply to periodic systems. This fact is evident when the lifted system is constructed. Definitions of observability and detectability of both time–invariant systems and periodic systems are provided. Tests to determine observability and detectability are also provided for both time–invariant and periodic systems. Of particular use is the Hautus Lemma, which provides a simple test for observability and detectability. A periodic Hautus Lemma is presented as a useful analog for the gage control system.

These definitions and tests facilitate analysis of disturbance models for periodic systems. Conditions of detectability, implying observability of the disturbance states, for the various disturbance models are presented. The number of disturbance states are limited to the number of outputs; furthermore, detectability is lost if any of the disturbance states mask any of the integrating modes of the system.

Chapter 6

Model Predictive Control

This chapter provides a discussion of the regulator used in this work. Industrial interest and ever increasing computing power have helped shape the theoretical efforts of MPC over the last 20 years. The first part of this chapter discusses the evolution of the current state of the art model predictive control theory. In the next major section, stability in the context of the combined estimator and regulator is discussed. The non–trivial task of determining targets is discussed third. Determination of feasible targets based on actual set points is the focus (as opposed to the problem of actually determining the setpoints). Lastly several of today's remaining computation issues are discussed. Consider first the definition of the problem of interest: the model predictive control problem.

6.1 The Model Predictive Control Problem

The model of choice in this work is the state space model; however, a variety of model types have been used in model predictive control schemes. Common to all predictive control schemes is the optimization of future inputs. The first move of the optimal input profile is then used for control.

Constraints, prediction horizons, control horizons, objective functions, and tuning parameters are all parameters that have to be specified in order to determine the MPC control law. The following is a very basic mathematical statement of the model predictive control problem.

$$\min_{\pi} \Phi(y_k, u_k, k) \tag{6.1}$$

subject to:

$$y_k = f(u_k, u_{k-1}, \ldots, y_{k-1}, y_{y-2}, \ldots) \tag{6.2}$$

where π is a set of decision variables, u represents the inputs, y represents the outputs, and k is a time index. The first move in the optimal input profile is implemented and, with new data to update the state, the problem is resolved at the next time step. In this study linear state space models are of interest, thus the large class of problems represented by Eq. 6.1 is narrowed. The next section presents a description of the most basic MPC controller and then discusses several advances that have led to the current technology.

6.2 From Infinity and Back

6.2.1 Infinite Horizon Control: Unconstrained

Kalman solved the unconstrained linear quadratic regulator (LQR) problem for continuous time systems in 1960 [69]. The LQR problem is stated for discrete time systems in Eq. 6.3.

Infinite Horizon Control: Unconstrained

$$\min_{\pi} \Phi_\infty(x_0, \pi)$$

subject to:

$$\begin{aligned} \Phi_\infty(x_0,\pi) &= \sum_{j=0}^{\infty} z_j^T Q z_j + v_j^T R v_j \qquad Q \geq 0, R > 0 \\ z_{j+1} &= A z_j + B v_j \\ z_0 &= x_0 \end{aligned} \tag{6.3}$$

where $\pi = \{v_0, v_1, \ldots\}$ and Q and R are the state and input penalty matrices, respectively. For simplicity, perfect state measurement is assumed. The solution to this problem is the well–known linear feedback law,

$$u_k = -K x_k \tag{6.4}$$

where K is calculated using the solution of the discrete algebraic Riccati equation (P).

$$K = -(B^T P B + R)^{-1} B^T P A \tag{6.5}$$

$$P_{k+1} = A^T (P_k - P_k B (B^T P_k B + R)^{-1} B^T P_k) A + Q \tag{6.6}$$

The solution as $k \to \infty$ converges for (A, B) stabilizable, $(A, Q^{1/2})$ detectable, R positive definite, and P_0 positive–semi definite. The solution provides a nominally stabilizing feedback gain. The undisturbed state evolves according to Eq. 6.7 which combines Eq. 6.4 with the state evolution equation.

$$x_{k+1} = (A - BK)x_k \tag{6.7}$$

The LQR was not used in the chemical process industry because of its non–intuitive tuning parameters and lack of constraints capabilities. The aerospace industry, however, embraced this idea because it required little computer storage and required a single multiplication to calculate a new control move.

6.2.2 Finite Horizon Control: Constrained

As far as the chemical and petroleum industries were concerned, MPC became much more useful when constraints were included in commercial algorithms. The first MPC formulations in industry used the finite impulse response model rather than the state space model. IDCOM [97], DMC [34], and QDMC [49] all brought some level of constraint handling to the predictive control problem. Qin and Badgwell provide a thorough review of the industrial evolution of model predictive control [91]. Their review provides MPC usage statistics as well as information on the latest commercial products and their features.

The finite horizon control problem in shown in Eq. 6.8. Notice that there is a distinction between the control horizon, N, and the prediction horizon, M. The control horizon specifies the number of future inputs used in the optimization. The prediction horizon specifies the duration of the state penalty. N is usually smaller than M since it is computationally expensive to add decision variables to the optimization. The decision variable are denoted using $\pi = \{v_0, v_1, \ldots, v_{N-1}\}$. Constraints on both the state and the output are enforced throughout the prediction horizon. State space models allow the use of unstable and integrating models (tanks for instance). The impulse response models used for control do not have this capability.

Finite Horizon Control: Constrained

$$\min_{\pi} \Phi_{M,N}(x_0, \pi)$$

subject to:

$$\begin{aligned} \Phi_{M,N}(x_0,\pi) &= \sum_{j=0}^{M-1} z_j^T Q z_j + \sum_{j=0}^{N-1} v_j^T R v_j \qquad Q \geq 0, R > 0 \\ z_{j+1} &= Az_j + Bv_j \\ z_0 &= x_0 \\ Hz_j &\leq h \qquad j = 0, 1, \ldots, M-1 \\ Dv_j &\leq d \qquad j = 0, 1, \ldots, N-1 \\ v_j &= 0 \qquad j \geq N \end{aligned} \tag{6.8}$$

The quadratic program in Eq. 6.8 is not difficult to solve, although an optimal solution does not guarantee a stable closed loop controller for all tuning parameters. This theoretical weakness did not slow the use of finite horizon controllers in industry. The success stories are well documented [90, 97], and the improvements in control far outweigh the extra care needed to ensure stability.

Endpoint Constraints A logical way to handle the stability problem associated with Eq. 6.8 is to require that the final state equal zero. Stability results for finite horizon control that require endpoint constraints are given by Kleinman, and Kwon and Pearson [75, 78]. The final state constraint is a restrictive condition (in fact, controllability is required as well) and can lead to aggressive control for short horizons. An endpoint stability proof for the general nonlinear infinite horizon problem, which includes the finite horizon problem, is provided by Findeisen [46] and Meadows [84].

6.2.3 Infinite Horizon Control: Constrained

The next improvement to the MPC algorithm involve a limited version of the infinite horizon problem that incorporates the idea of endpoint constraints [86]. First, consider the following expression of the infinite horizon objective function with $v_j = 0$, $j \geq N$. Figure 6.1 illustrates the contributions to the objective function. There are a finite number of inputs yet an infinite number of states that contribute to the objective function.

$$\Phi_N(x_0, \pi) = \sum_{j=0}^{N-1} (z_j^T Q z_j + v_j^T R v_j) + \sum_{j=N}^{\infty} z_j^T Q z_j \tag{6.9}$$

In order to compute the infinite sum consider that after the control horizon, N, the states evolve according to $z_{N+j+1} = A z_{N+j}$, thus

$$\sum_{j=N}^{\infty} z_j^T Q z_j = z_N^T (Q + A^T Q A + A^{2T} Q A^2 + \ldots) z_N \tag{6.10}$$

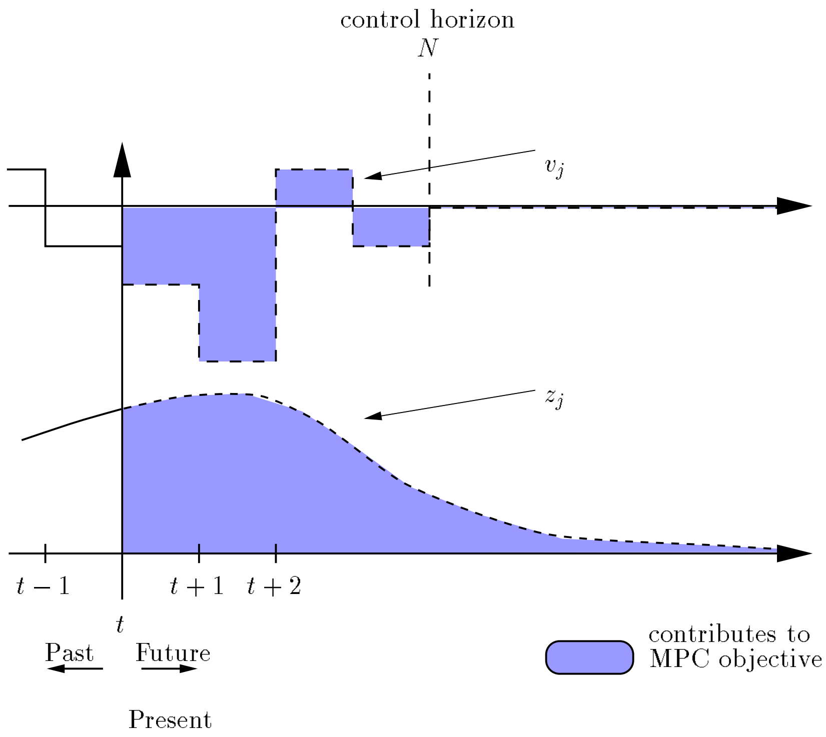

Figure 6.1: Infinite Horizon Control with $v_k = 0$ for $k \geq N$.

Defining $\tilde{Q}$ as follows

$$\tilde{Q} = Q + A^T Q A + A^{2T} Q A^2 \ldots \tag{6.11}$$

multiplying both sides by $A^T(\cdot)A$

$$A^T \tilde{Q} A = A^T Q A + (A^T)^2 Q (A)^2 + \ldots \tag{6.12}$$

and, finally, subtracting Eq. 6.11 from Eq. 6.12 provides a Lyapunov equation $\tilde{Q} = Q + A^T \tilde{Q} A$ for which $\tilde{Q}$ is the positive definite solution. The Lyapunov equation is only valid for stable A, thus for unstable plants the Schur decomposition is used. The Schur decomposition, shown in Eq. 6.13, partitions the eigenvalues such that the stable eigenvalue are contained in T_{11} and the unstable eigenvalue are in T_{22}.

$$U^T A U = T \tag{6.13}$$

$$A = \begin{bmatrix} U_s & U_u \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} U_s^T \\ U_u^T \end{bmatrix} \tag{6.14}$$

The idea of zeroing the state at the end of the control horizon is relaxed somewhat by requiring that only the *unstable* modes equal zero at the end of the control horizon. The stable part ($A_s = U_s T_{11} U_s^T$) is used in the determination of the infinite penalty, $\tilde{Q}$, and an equality constraint is constructed from the unstable part ($U_U^T z_N = 0$). Since the constraints on z_j may be violated for $j > N$, some check is required.

For bounded sets, Gilbert and Tan calculate a time index after which constraints are not violated [50]. The constraint can be found *a priori* and is calculated through a series of linear programs. If the state constraints are satisfied before this time index they are satisfied for all time; k_2 denotes the final constraint index. The complete problem is presented in Eq. 6.15.

Infinite (in the State) Horizon Control: Constrained

$$\min_{\pi} \Phi_N(x_0, \pi)$$

subject to:

$$\begin{aligned}
\Phi_N(x_0, \pi) &= \sum_{j=0}^{N-1} (z_j^T Q z_j + v_j^T R v_j) + z_N^T \tilde{Q} z_N \qquad Q \geq 0, R > 0 \\
z_{j+1} &= A z_j + B v_j \\
z_0 &= x_0 \\
U_u^T z_N &= 0 \qquad j = 0, 1, \ldots, N-1 \\
H z_j &\leq h \qquad j = 0, 1, \ldots, N-1 \\
D v_j &\leq d \qquad j = 0, 1, \ldots, N-1 \\
v_j &= 0 \qquad j \geq N \\
H A^j z_N &\leq h \qquad j = 0, 1, \ldots, k_2
\end{aligned} \tag{6.15}$$

The big advantage of the constrained MPC formulation given by Eq. 6.15 (also known as the Rawlings–Muske controller) over the constrained finite horizon formulation in Eq. 6.8 is that the regulator is nominally stable for all $Q \geq 0$, $R > 0$. A drawback of both constrained controllers is that the closed loop and open loop trajectories are different. The next step involves regaining Kalman's full infinite horizon formulation, including stability, in the presence of constraints.

Parameterized Inputs The previous problem is infinite in the stable modes, but not the inputs. Several researchers have presented the idea of infinite input parameterization [112, 32, 104]. Figure 6.2 illustrates the contributions to the objective function. While an infinite number of inputs are considered, only a finite number of them are decision variables in the quadratic program. A finite QP is formulated by setting $v_k = Kz_k$ for $k \geq N$ where K is the solution to the unconstrained LQR (Eq. 6.3). The infinite penalty must be modified to penalized the inputs.

$$\begin{aligned} \sum_{j=N}^{\infty} v_j^T R v_j + z_j^T Q z_j &= \sum_{j=N}^{\infty} z_j^T (K^T R K + Q) z_j \\ &= \sum_{j=0}^{\infty} z_N^T (A - BK)^{jT} (K^T R K + Q)(A - BK)^j z_N \end{aligned} \tag{6.16}$$

The infinite sum in Eq. 6.16 satisfies the following Lyapunov equation.

$$\tilde{Q} = (K^T R K + Q) + (A - BK)^T \tilde{Q} (A - BK) \tag{6.17}$$

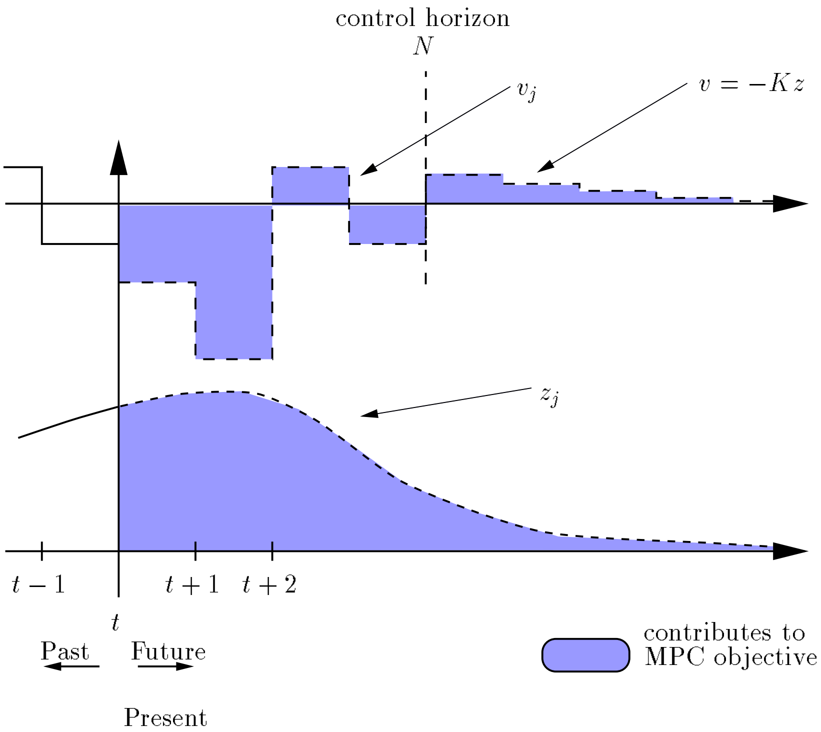

Figure 6.2: Infinite Horizon Control with $v_k = -Kx_k$ for $k \geq N$.

Special consideration is given to the constraints on the infinite horizon since the optimization only ensures that the constraints are satisfied over the first N time steps. Figure 6.3 helps illustrate the method used to check the constraints on the infinite horizon. Several regions are defined in Figure 6.3, the large circular region with radius B_R is the largest circular region containing exclusively allowable states. It can be shown that there is an inner circle (denoted B_r) such that once the state enters the smaller circle it does not leave the outer circle where it could potentially violate constraints. In other words, $x_k \in B_r \Longrightarrow A^j x_k \in B_R$ for all $j \geq 0$ [102]. The quadratic program must be solved before all of the constraints can be checked. If there are constraint violations, then N is increased and the problem is resolved. Scokaert and Rawlings show that there exists a finite N which satisfies the constraints over the infinite horizon [103]. The full infinite horizon problem is presented in Eq. 6.18.

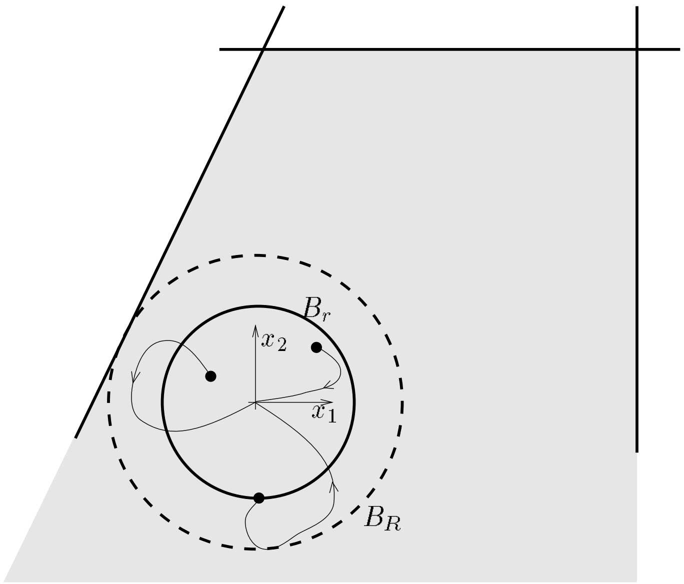

Figure 6.3: Illustration of the two regions involved in checking constraints on the infinite horizon. The lines bounding the shaded region represent constraints.

Infinite Horizon Control: Constrained

$$\min_{\pi} \Phi_N(x_0, \pi)$$

subject to:

$$\begin{aligned}
\Phi_N(x_0, \pi) &= \sum_{j=0}^{N-1} z_j^T Q z_j + v_j^T R v_j + z_N^T \tilde{Q} z_N \qquad Q \geq 0, R > 0 \\
z_{j+1} &= A z_j + B v_j \\
z_0 &= x_0 \\
H z_j &\leq h \qquad j = 0, 1, \ldots, N-1 \\
D v_j &\leq d \qquad j = 0, 1, \ldots, N-1 \\
v_j &= -K z_j \qquad j = N, \ldots, \infty
\end{aligned} \tag{6.18}$$

Additional Comments The four control formulations presented in this section show the evolution of the constrained infinite horizon control problem. Some common variations on the previous control problems include penalties on the rate of change of the inputs and constraints on the rate of change on the inputs. To penalize the input rate of change the term $\sum_{j=0}^{N-1} \Delta v_j^T S \Delta v_j$ (where $\Delta v_j = v_j - v_{j-1}$) is added to the objective function. Solution to the full infinite horizon problem is handled by augmenting the original state with v_{j-1} [94]. Constraints on the rate of change of input are handled when the constraint matrix and vector is constructed for π. In the context of the film problem, spatial rate of change constraints are specified as well. There are physical limits to the bending of the die lip. Excessive bending is handled by placing constraints on the travel between adjacent inputs. Spatial constraints are specified using the existing structure of D and d.

$$d_{\min} \leq D u_j \leq d_{\max} \tag{6.19}$$

in which $d_{\min} = -d_{\max}$ are vectors of dimension $m-1$ (where m is the number of inputs) and

$$D = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{6.20}$$

This discussion concludes the section on the regulator formulation. The stability results for the regulator are now applied to the specific structure of the controller used in the film process.

6.3 Controller/Observer Stability

This section addresses the issue of nominal stability of the combined linear periodic estimator and the non–linear regulator. The regulator is non–linear because of the constraints. To show stability some previous results are reviewed.

Meadows [84] shows that for feasible constraints the stabilizing state feedback control law that solves Eq. 6.15 is a Lipschitz continuous function, $u(x)$. Scokaert and Rawlings [101] use the Lipschitz condition to establish exponential stability of the control law.

In Section 4.3.1 the periodic Kalman filter for a reconstructible system was shown to be exponentially stable. Exponential stability can also be shown for a periodic detectable system. Kwakernaak and Sivan [77] state that for detectable time–invariant systems the Riccati iteration in 4.17 converges to a unique asymptotic solution. The lifted system in Theorem 4.1 can be used to calculate a set of unique covariance matrices. The argument showing exponential stability from Section 4.3.1 applies and the detectable system is shown to be exponential stable.

Scokaert and Rawlings [101] establish the asymptotic stability of a nominally exponentially stable system perturbed by an asymptotically decaying perturbation. Thus the interconnection between the periodic linear estimator and the non–linear regulator satisfies the criteria set forth by Scokaert and Rawlings.

6.4 Target Tracking

In order to simplify the discussion of the controllers in section 6.2 zero targets were assumed. It was assumed that a steady–state target could be calculated and then deviation variables could be used to recast the problem in the forms presented in Eqs. 6.3, 6.8, 6.15, and 6.18. Presented here are two methods of determining the targets from setpoints u_{sp} and y_{sp} and input and output constraints. First a multistep method used in the `lmpc` code (see Section 6.5.3) is outlined. Next, a more streamlined solution using exact soft constraints is presented.

Since targets are not always needed for all outputs an alternative measurement matrix ($\acute{C}$) can be specified. Full row rank of $\acute{C}$ is required because redundant set points or unachievable set points (zeros rows) are unimportant cases. Additional system requirements include $(A, \acute{C})$ detectable and (A, B) stabilizable. Calculating targets for undetectable or unstabilizable systems is possible in certain circumstances, but makes little sense from a practical point of view. The steady–state targets are denoted x_t and u_t. The steady-state linear equations are given by

$$\begin{aligned} x_t &= Ax_t + Bu_t \\ y_{sp} &= \acute{C}x_t \end{aligned} \tag{6.21}$$

Before presenting the multistep target tracking problem the following result is needed.

Theorem 6.1 (Existence of $\mathbf{x_t}, \mathbf{u_t}$) *Consider an eigenvalue decomposition of the state matrix from Eq. 6.21 that partitions the eigenvalues into integrators (i) and non–integrators (n).*

$$A = U\Lambda U^{-1} = \begin{bmatrix} U_i & U_n \end{bmatrix} \begin{bmatrix} I_i & 0 \\ 0 & \Lambda_n \end{bmatrix} \begin{bmatrix} \tilde{U}_i \\ \tilde{U}_n \end{bmatrix} \tag{6.22}$$

letting $z = U^{-1}x_t$ in Eq. 6.21 the system can be rewritten as follows

$$\begin{aligned} z &= \Lambda z + \tilde{B}u_t \\ y_{sp} &= \tilde{C}z \end{aligned} \tag{6.23}$$

where

$$\tilde{B} = U^{-1}B = \begin{bmatrix} \tilde{B}_i \\ \tilde{B}_n \end{bmatrix} \tag{6.24}$$

$$\tilde{C} = \acute{C}U = \begin{bmatrix} \tilde{C}_i & \tilde{C}_n \end{bmatrix} \tag{6.25}$$

and let $\mathcal{N}$ be the null space of $\tilde{B}$ then a solution to the system in Eq. 6.21 exists if Eq. 6.26 is full row rank.

$$\begin{bmatrix} \tilde{C}_i & \tilde{C}_n(I_n - \Lambda_n)^{-1}\tilde{B}_n\mathcal{N}_i \end{bmatrix} \tag{6.26}$$

Proof: This follows directly from partitioning Eq. 6.23

$$\begin{bmatrix} 0 & 0 & -\tilde{B}_i \\ 0 & I - \Lambda_n & \tilde{B}_n \\ \tilde{C}_i & \tilde{C}_n & 0 \end{bmatrix} \begin{bmatrix} z_i \\ z_n \\ u_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ y_{sp} \end{bmatrix} \tag{6.27}$$

and then eliminating z_n leaving the following linear system.

$$\begin{bmatrix} 0 & -\tilde{B}_i \\ \tilde{C}_i & \tilde{C}_n(I_n - \Lambda_n)^{-1}\tilde{B}_n \end{bmatrix} \begin{bmatrix} z_i \\ u_t \end{bmatrix} = \begin{bmatrix} 0 \\ y_{sp} \end{bmatrix} \tag{6.28}$$

For a solution to exist u_t must be in the null space of $\tilde{B}_i, (\tilde{B}_i\mathcal{N}_i u_t = 0)$. It follows that Eq. 6.29 must be full row rank for a solution to exist for every y_{sp}.

$$\begin{bmatrix} \tilde{C}_i & \tilde{C}_n(I_n - \Lambda_n)^{-1}\tilde{B}_n\mathcal{N}_i \end{bmatrix} \tag{6.29}$$

□

6.4.1 Multistage Target Determination

Perfect Target Tracking The first pass at calculating targets should be to meet the setpoints exactly. A feasible solution to the quadratic program in Eq. 6.30 provides perfect target tracking. The output target is satisfied with an equality constraint and any input degrees of freedom are minimized from some desired input setpoint (u_{sp}).

Perfect Target Tracking

$$\min_{[u_t, x_t]^T} \Phi(u_t, u_{sp})$$

subject to:

$$\begin{aligned} \Phi(u_t, u_{sp}) &= (u_t - u_{sp})^T R_s (u_t - u_{sp}) \qquad R_s > 0 \\ \\ x_t &= A x_t + B u_t \\ y_{sp} &= \acute{C} x_t \\ H x_t &\leq h \\ D u_t &\leq d \end{aligned} \tag{6.30}$$

Theorem 6.2 (Perfect Target Tracking) *The feasible quadratic program shown in Eq. 6.30 has a unique solution for $(A, \acute{C})$ detectable.*

Proof: Consider first if there is a feasible solution. In the absence of inequality constraints the solution exists for the condition presented in Theorem 6.1. When the inequality constraints are present a linear program can determine if a feasible point exists [48]. Assume now the existence of a solution and recall Eq. 6.31 from Theorem 6.1.

$$\begin{bmatrix} 0 & -\tilde{B}_i \\ \tilde{C}_i & \tilde{C}_n (I_n - \Lambda_n)^{-1} \tilde{B}_n \end{bmatrix} \begin{bmatrix} z_i \\ u_t \end{bmatrix} = \begin{bmatrix} 0 \\ y_{sp} \end{bmatrix} \tag{6.31}$$

Eq. 6.31 is a restatement of the equality constraints from Eq. 6.30. Since R is positive definite all solutions, u, are unique determined. Therefore it is sufficient to have full column rank of $\tilde{C}_i$ in Eq. 6.31 to guarantee a unique global solution. A Hautus condition on the system given by Eq. 6.23 proves that $\tilde{C}_i$ is full column rank if the system is detectable, and the proof is complete.

$$\begin{bmatrix} (\lambda - 1)I_i & 0 \\ 0 & (\lambda I_n - \Lambda_n) \\ \tilde{C}_i & \tilde{C}_n \end{bmatrix} \tag{6.32}$$

□

Least Squares Target Tracking For more setpoints than inputs the quadratic program above has an empty feasible set. An acceptable set of targets can be calculated using the following least squares quadratic program. Since the setpoints cannot be met exactly the tuning parameters can be used to weight the relative importance of the individual setpoints. The inequality constraints are assumed to contain the origin, $h_i > 0$, $d_i > 0$. The following theorem is due to Rao [92].

Least Squares Target Tracking

$$\min_{[u_t, x_t]^T} \Phi(x_t, u_t, u_{sp}, y_{sp})$$

subject to:

$$\begin{aligned} \Phi(x_t, u_t, u_{sp}, y_{sp}) &= (y_{sp} - \acute{C}x_t)^T Q_s (y_{sp} - \acute{C}x_t) + (u_t - u_{sp})^T R_s (u_t - u_{sp}) \\ & Q_s, R_s > 0 \\ x_t &= Ax_t + Bu_t \\ Hx_t &\leq h \\ Du_t &\leq d \end{aligned} \tag{6.33}$$

Theorem 6.3 (Least Squares Target Tracking) *The quadratic program shown in Eq. 6.33 has a unique solution for $(A, \acute{C})$ detectable.*

Proof: By requiring that the inequality constraints contain the origin, a feasible solution exists. It is then sufficient to have a strictly convex quadratic program [48]. Strictly convex requires a positive definite Hessian and affine constraints.

Consider first the case in which A has no eigenvalues at 1. $(I - A)$ is invertible and the equality constraint can be used to eliminate x_t from the problem.

$$x_t = (I - A)^{-1} B u_t \tag{6.34}$$

Since R is positive definite and the constraints are affine, the strictly convex quadratic problem has a unique global solution.

Consider now the case where A has at least one eigenvalue at 1. Using the substitution $z = \acute{C} x_t - y_{sp}$ the equality constraint can be rewritten

$$\begin{bmatrix} I - A \\ \acute{C} \end{bmatrix} x_t = \begin{bmatrix} -B u_t \\ y_{sp} - z \end{bmatrix} \tag{6.35}$$

The Hautus lemma of detectability implies full rank of the following matrix.

$$\mathcal{H} = \begin{bmatrix} \lambda I - A \\ \acute{C} \end{bmatrix} \tag{6.36}$$

for all $\lambda \in \mathbb{C}$ with magnitude greater than or equal to 1. It is sufficient to only consider the Hautus matrix, $\mathcal{H}$, with $\lambda = 1$. Due to detectability $\mathcal{H}$ has full column rank. Partitioning the columns of the pseudo–inverse of $\mathcal{H}$ allows x_t to be eliminated from the constraint equations. Since the remaining constraints represent an affine relationship between u_t and z and the Hessian is positive definite, it follows that

Eq. 6.33 is strictly convex. Since the quadratic program is strictly convex for both situations, the solution exists and is unique as claimed. □

One reason the program in Eq. 6.33 is not computed before Eq. 6.30 is that the penalty on the inputs may provide an optimal answer that is not the perfect target tracking solution when perfect target tracking is, in fact, possible. Muske provides a formula for R_s that ensures perfect target tracking when possible [85].

6.4.2 Exact Soft Constraint Target Determination

It is desirable to solve only one target calculation that provides the same solution that the multistep target calculations provide. One calculation can be done with exact soft constraints. The main idea behind soft constraints is that the constraint is met when feasible. See Fletcher [48] for details on exact soft constraints. For this study it is sufficient to say that the equality constraint $y_{sp} = \acute{C}x_t$ is relaxed by redefining the constraint in terms of the slack variable ϵ.

$$\epsilon \geq y_{sp} - \acute{C}x_t \tag{6.37}$$

$$\epsilon \geq -(y_{sp} - \acute{C}x_t) \tag{6.38}$$

$$\epsilon \geq 0 \tag{6.39}$$

In addition to penalizing ϵ in the quadratic sense, a linear penalty, S, is used. S must have strictly non–negative values. Appropriate choice of the accompanying linear penalty provides the property that, if feasible, constraints are met. As before, the inequality constraints need to contain the origin. The following theorem is due to

Rao [92].

Exact Soft Constraint Target Tracking

$$\min_{[u_t, x_t]^T} \Phi(\epsilon, u_t, u_{sp})$$

subject to:

$$\begin{aligned}
\Phi(\epsilon, u_t, u_{sp}) &= \epsilon^T Q_s \epsilon + (u_t - u_{sp})^T R_s (u_t - u_{sp)} + S_s \epsilon \\
& \quad Q_s, R_s, S_s > 0 \\
\\
x_t &= A x_t + B u_t \\
y_{sp} - \acute{C} x_t &\leq \epsilon \\
-(y_{sp} - \acute{C} x_t) &\leq \epsilon \\
H x_t &\leq h \\
D u_t &\leq d
\end{aligned} \tag{6.40}$$

Theorem 6.4 (Exact Soft Constraint Target Tracking) *The feasible quadratic program in Eq. 6.40 has a unique solution for* $(A, \acute{C})$ *detectable.*

Proof: By requiring that the inequality constraints contain the origin, a feasible solution exists. It is then sufficient to have a strictly convex quadratic program [48]. Strictly convex requires a positive definite Hessian and affine constraints.

Consider first the case where A has no eigenvalues at 1. Since $(I - A)$ is invertible and the equality constraint can be used to eliminate x_t from the problem.

$$\begin{aligned}
y_{sp} - \acute{C}(I - A)^{-1} B u_t &\leq \epsilon \qquad (6.41) \\
-(y_{sp} - \acute{C}(I - A)^{-1} B u_t) &\leq \epsilon
\end{aligned}$$

Since R is positive definite and the constraints are affine, the strictly convex quadratic problem has a unique global solution.

Consider now the case where A has at least one eigenvalue at 1. The Hautus lemma of detectability implies the following rank condition

$$\mathcal{H} = \begin{bmatrix} \lambda I - A \\ \acute{C} \end{bmatrix} = n \tag{6.42}$$

for all $\lambda \in \mathbb{C}$ with magnitude greater than or equal to 1. It is sufficient to only consider the Hautus matrix, $\mathcal{H}$, with $\lambda = 1$. Due to detectability $\mathcal{H}$ has full column rank. It then follows that x_t is uniquely determined from the following equation

$$\mathcal{H}x_t = \begin{bmatrix} Bu_t \\ y_{sp} \pm (\epsilon - s) \end{bmatrix} \tag{6.43}$$

where s is a positive slack for the inequality constraints. Partitioning the columns of the pseudo–inverse of $\mathcal{H}$ allows x_t to be eliminated from the constraint equations. Since the remaining constraints represent an affine relationship between u_t and $(\epsilon - s)$, it follows that Equation 6.40 is strictly convex. Since the quadratic program is strictly convex for both situations, the solution exists and is unique as claimed. $\square$

The feasible quadratic program in Eq. 6.40 with $\epsilon = 0$ reduces to Eq. 6.30, thus the exact target tracking solution is obtained. Details involving tuning the target calculation in Eq. 6.40 are not well documented. It would be useful to have a better understanding of this potentially useful tool.

6.5 Computational Issues

While discussing the controllers above, very little was said about the ability to obtain a feasible solution in real time. In general, the chemical engineer can assume that an optimization involving a quadratic program is a solved problem. A well–posed problem has a global solution and currently available software is reliable. The control engineer, on the other hand, is concerned with real time implementation. One concern is if the solution time is prohibitive to real time implementation. In addition, the constrained quadratic program may not have a unique solution, in fact, it may not have a solution. Thus some provisions should be made to shed, soften, or violate constraints. The issue of constraint handling is truly a design issue and should be treated as such. Once infeasibility is encountered, no set formula or algorithm to regain feasibility works for all plants. The issues of speed and on line solutions and infeasibility are addressed in this section as well as some comments about the software used for the simulations in Chapter 7.

6.5.1 Infeasibility

The solution to the target tracking problem provides a new steady–state origin for the control problem. Deviation variables are used to cast the regulator problem into one of the forms given in Section 6.2. A satisfactory solution to the target QP does not guarantee that the regulator QP finds a feasible solution. For example, consider the case of a disturbance that pushes the output outside of the constraint limits. Either limited input movement or slow dynamics may make it impossible to drive the output back within constraints at the next time step. In this case, the regulator QP's feasible set is empty. Faced with infeasibility, at least two scenarios can be considered. The first situation involves an output that is violating a safety constraint. It is of interest to minimize the maximum value of the constraint violation even at the expense of a longer constraint violation time. In a second situation, violations may mean that scrap product is being produced. The maximum violation is inconsequential and the duration of the violation is critical. Thus, the constraint handling issue really is a design issue.

Many of the commercial algorithms use constraint ranking to handle infeasibility. Constraints can be removed in a predetermined manner in order to better meet the overall economic or system wide goals. Constraint shedding ultimately provides suboptimal solutions. A better solution is to use the constraint ranking system in conjunction with exact soft constraints. The same criteria used to determine the constraint ranking can be used to determine the relative weighting in the soft constraint quadratic program. The inequality constraint $Hz_j \leq h$ are relaxed by redefining the constraint in terms of the slack variable ϵ.

$$\epsilon_j \geq Hz_j - h \tag{6.44}$$

$$\epsilon_j \geq -(Hz_j - h) \tag{6.45}$$

$$\epsilon_j \geq 0 \tag{6.46}$$

In addition to penalizing ϵ in the quadratic sense, a linear penalty, S, is used. S must have strictly non–negative values. Appropriate choice of the accompanying linear penalty provides the property that, if feasible, constraints are met. As before, the inequality constraints need to contain the origin. Let σ be a vector of the ϵ_j's.

Exact Soft Constraint Regulation

$$\min_{[\pi,\sigma]} \Phi_N(x_0, \pi, \sigma)$$

subject to:

$$\begin{aligned}
\Phi_N(x_0, \pi, \sigma) &= \sum_{j=0}^{N-1} z_j^T Q z_j + v_j^T R v_j + \epsilon_j^T S_1 \epsilon_j + S_2 \epsilon_j \\
& \quad Q \geq 0, \ S_1, S_2, R > 0 \\
z_{j+1} &= A z_j + B v_j \\
z_0 &= x_0 \\
H z_j &\leq h \qquad j = 0, 1, \ldots, N-1 \\
D v_j &\leq d \qquad j = 0, 1, \ldots, N-1 \\
v_j &= -K z_j \qquad j = N, \ldots, \infty
\end{aligned} \tag{6.47}$$

Scokaert and Rawlings [103] provide some examples using exact soft constraints and compare the method to other hard and soft constraint methods. In sheet and film forming processes transient infeasibility is not a major problem. The exact soft constraint solution shows great promise, however, in other control applications. The intricacies of tuning are still an open issue. One problem large film processes do suffer from is prohibitive computing expense. The next section addresses this problem.

6.5.2 Structured Interior Point Methods

The average MPC application is easily solved with the existing software and hardware available. In fact, Qin and Badgwell report that the largest problem implemented by most commercial products is 30 inputs by 30 outputs [91]. Research in the area of plant wide control or combined economic optimization and control present problems of much greater dimension. In addition processes such as the film process provide inherently large numbers of inputs and outputs. Thus, there is motivation for faster, structured solving methods. Dave *et al.* [36] find conditions under which the QP can be reduced to an LP. Their conclusions include the outright dismissal of the QP as a feasible on–line solution technique. This conclusion is perhaps too extreme, especially in light of the available structured interior point methods.

The quadratic program that arises from a linear model predictive control problem is not an ordinary quadratic program. The optimization can be posed as a highly structured problem. Several articles are available on interior point methods applied to the MPC optimization problem [120, 122]. This method is further expanded to include exact soft output constraints and is applied to film processes in particular [93, 94]. The following discussion is intended only to introduce the idea of structured interior point methods applied to MPC. Refer to [94] for a more thorough discussion of the issues.

A straightforward way to pose the regulator optimization problem is to eliminate the states from the problem leaving the inputs as decision variables. The Hessian of the reduced problem is dense and commercial solvers such as QPSOL [51] provide reliable solutions. Now, consider leaving the states in the optimization problem. While incorporation of the states makes the dimension much greater than the standard approach, the optimization problem is sparse and highly structured. In the primal-dual interior point framework described by Wright [121], the solution to the quadratic program is obtained from a finite number of linear algebra subproblems.

Some numerical results from [93] illustrate the promise of this method when applied to film processes. Example 1 is a 50 input, 100 output polymer process. The gain matrix is based on the Toeplitz structure model generated from the 3M data (see Chapter 3). Example 2 is a 100 input, 100 output plant, and example 3 is the same plant with spatial actuator constraints. All three examples have non–zero dynamics to increase the size of the problem. A sinusoidal disturbance is rejected in these examples. Table 6.1 shows the computational times required for a single calculation of the quadratic program in (6.18) for the three examples with different control horizon lengths N. The three solving methods listed are the commercial package QPSOL (version 3.2) [51] as a representative dense optimization approach, the structured interior point method (SIP), and a sparse structured interior point method (SSIP) which exploits the structure and sparsity in the optimization data during solution of the linear algebra subproblems. The computational times were obtained using the Unix `time` command on an Alphastation 250. For the QPSOL experiments, the listed times do not account for the time required to construct the optimization data. These

Simulation Parameters	N	QPSOL	SIP	SSIP
50 inputs, 100 outputs	10	50.9	39.1	26.7
50 inputs, 100 outputs	20	410.7	75.0	54.6
100 inputs, 100 outputs	10	340.7	80.3	43.8
100 inputs, 100 outputs, Δu penalty	20	–	160.8	89.5
100 inputs, 100 outputs, Δu penalty	10	470.2	110.9	48.4

Table 6.1: Comparison of Computational Times (sec)

times were often significant, but they represent off-line calculations. These results show great promise for on–line implementation of constrained infinite horizon model predictive control.

The results show great potential for on–line infinite horizon control.

6.5.3 Software: LMPC Toolbox

The linear model predictive control toolbox (`lmpc`) was first developed by Ken Muske to run the examples presented in the February, 1993 AIChE Journal article "Model Predictive Control with Linear Models." [86]. The code was originally written in MATLAB which is a high level interactive language for numerical computations. It provided a quick and easy way to run infinite horizon model predictive control simulations. Subsequent reworking of the code provided the current user interface. The code was also ported to Octave [43] from MATLAB and is now supported for both programs. The code finds extensive use in graduate courses and industrial short courses [5]. Disturbances, state feedback, and model mismatch can be investigated with the routines. A modified version of the code was used to generate the simulations in Chapter 7. The code was modified to handle such specific problems as scanning sensors, and measured inputs. Appendix C provides more detailed information regarding the `lmpc` code.

6.6 Summary

The unconstrained infinite horizon regulator, constrained finite horizon regulator, and two forms of the constrained infinite horizon regulator are presented in this chapter. The full constrained infinite horizon controller has identical open and closed loop performance in the nominal case. In order to calculate the profile for the full constrained infinite horizon controller a QP is solved and then a series of checks determine future constraint violations. Failure to meet constraints requires increasing the control horizon and resolving the problem. The regulator was also shown to have stability when combined with a periodic estimator.

Target tracking is presented next. The conditions of detectability, stabilizability, and full row rank of the measurement matrix are required so that only sensible problems are considered. Two different target tracking strategies are presented. The first strategy is a multistep calculation in which the exact steady–state solution is calculated first. Failure to find a feasible steady–state point results in a second, least squares quadratic program. Conditions of feasibility in the unconstrained case are presented. Conditions of uniqueness are presented for any feasible solution for both of the quadratic programs. An alternative to the two target tracking problems is an exact soft constraint problem. Suitable values of the penalties ensure that constraints are met if feasible.

Among the computational issues addressed in this chapter are infeasibility in the regulator, speed of computing, and software for simulations. Infeasibility in the regulator is really a design issue. The control engineer needs to design a constraint softening or shedding policy. An exact soft constraints regulator is presented as a viable alternative, but it still needs to be tuned. In order to address the issue of speed, a brief discussion of interior point methods applied to film control problems is presented. Simulations using interior point methods provide promising results and encourages large scale on–line film control. Lastly there is a mention of the software used to generate the simulations in the next chapter.

Chapter 7

Simulations

The purpose of this chapter is to use simulations to investigate the estimation and control of sheet and film forming processes. There are four simulations in this chapter, each of which illustrate a different point. The first simulation demonstrates the use of the periodic Kalman filter. The second simulation compares three different disturbance models. The third simulation provides evidence that the measured input disturbance is useful for the film process. Finally, a simulation demonstrating the controller's robustness is presented. There are a large number of inputs and outputs making it difficult to convey results. Thus, the data is presented in several ways including two and three dimensional plots of inputs and outputs as well as plots of disturbance states.

7.1 Simulation Study I: Full Profile Scanning Versus Partial Profile Scanning

The purpose of the first simulation is three–fold. First this simulation shows how more frequent control action provides better disturbance rejection. Secondly, this simulation explains in detail many of the characteristics that are common to all of the simulations in this chapter. Lastly, the three dimensional plots are explained. A variation on this simulation was presented at the 1996 AIChE annual meeting [26].

Commercially available scanning systems ordinarily wait for the detector to complete one or possibly two scans before sending a profile to the controller. This profile may be raw data or it may be filtered. Only after a complete scan is the full profile available for use in manual, steady–state, or predictive control. Certainly the sensor vendors could provide a means for this data to be accessed more frequently. When the industrial customers demand more data then it is likely the vendors will provide a more open data delivery system.

For this simulation a distinction is made between using data at the end of a complete scan and using data immediately. The industrial method of waiting for the complete profile before taking control action is referred to as *full scan* control. The second control scheme in this example uses the scanned data immediately. Using the data as it is measured is referred to as *partial scan* control. The difference between the two data collection methods is illustrated in Figure 7.1. Although the author

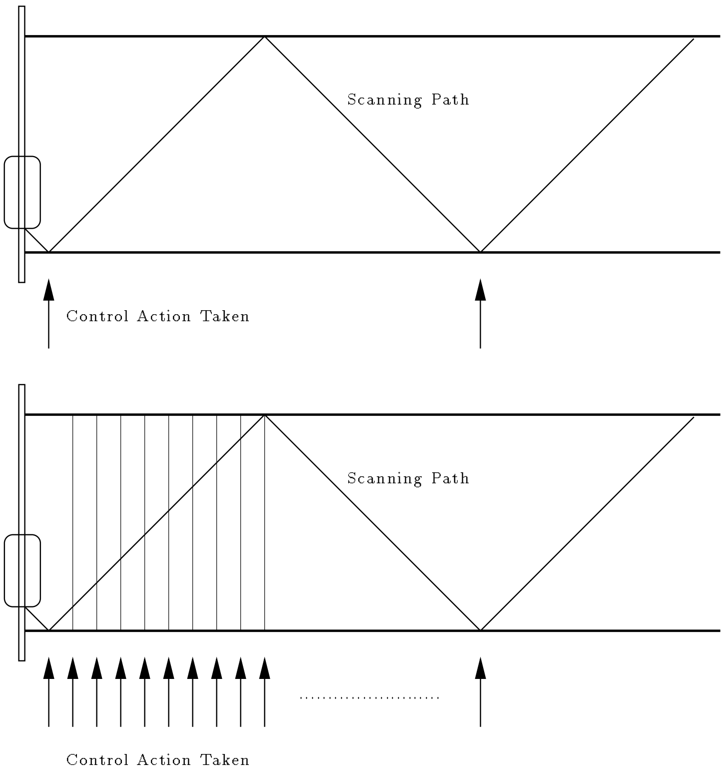

Figure 7.1: Full scan control (top) versus partial scan control (bottom).

knows of no predictive controllers in use today, both schemes use a model predictive controller in order to provide a fair comparison. The model is calculated from the 3M data (see Chapter 3). Toeplitz structure is enforced during the calculation of the model parameters. There are 12 inputs and 25 lanes measured by the scanner. A time–delay of ten sample times is assumed. The output disturbance model is chosen to remove steady–state offset. The periodic estimator is tuned such that a QDMC filter is obtained. The net effect of the tuning is that the difference between the model prediction and the measurement is assumed to be a constant bias in the output. The targets and control profiles are calculated to remove this bias. The regulator is a constrained infinite horizon model predictive regulator in which the input is set to zero after the control horizon. The controllers attempt to reject a disturbance that comes through the input. This disturbance is equivalent to an actuator bias. The simulation is summarized below and a discussion of the results follows.

Question: How does taking control action at every time step compare with taking control action after each complete scan?

Model: 3M model with Toeplitz structure enforced (see Chapter 3), 12 inputs, 25 outputs, delay of ten sample times.

Estimator: Time–varying Kalman filter based on an output disturbance model tuned as a QDMC filter.

Regulator: Constrained model predictive controller.

Simulation: Evaluate film response to disturbances coming through the input.

7.1.1 Study I: Results

Figure 7.2 shows the results of running full scan control. While the surface of the plot looks very much like a sheet, it is more precise to regard the surface as a series of snapshots specified by the sensor. For example, the data available in lanes 1–25 at time zero represent the thickness of the sheet directly underneath the sensor at time zero, although the sensor only detects the thickness in one lane. Similarly, at time one the data represents the state of the film underneath the sensor at time one.

Consider the output shown in Figure 7.2. Notice first that the disturbance is not underneath the sensor for ten time steps (the length of the time–delay). The scanner begins in lane one at time zero, traverses to lane 25 at time 24, returns to lane one at time 48. The arrows on the plot denote the scanner's position when it is in lane one or lane 25. At time 48 the full set of data is passed to the controller and input profile is calculated, and the first move in that input profile is injected into the plant. Ten time steps later the effect of the control action is seen by the sensor. Figure 7.3 shows the corresponding input profile.

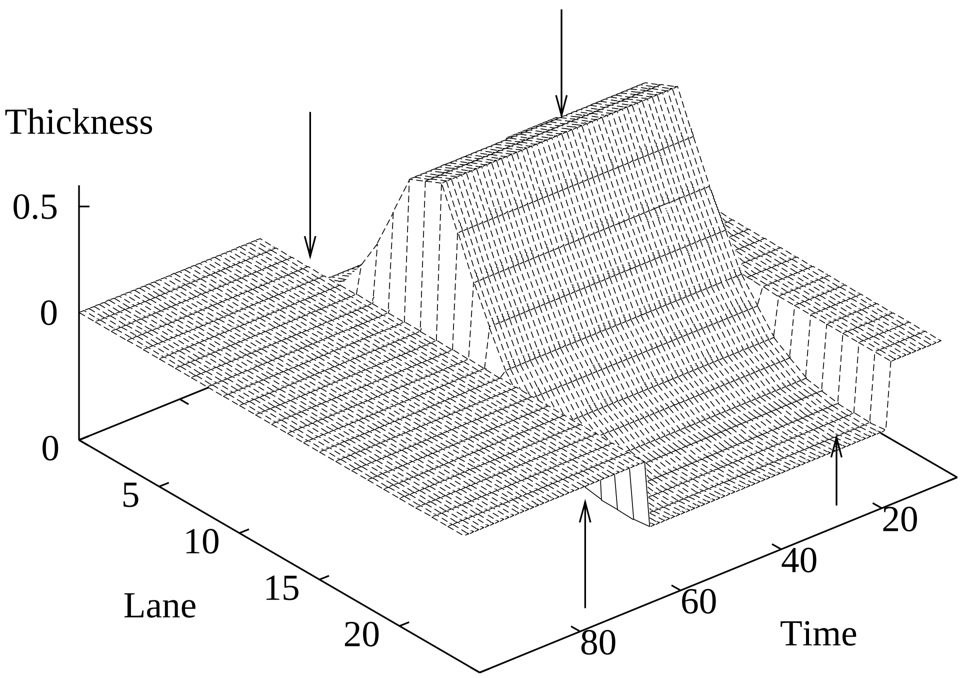

Figure 7.2: Film thickness: full scan control strategy rejecting an input disturbance (arrows denote sensor location in time).

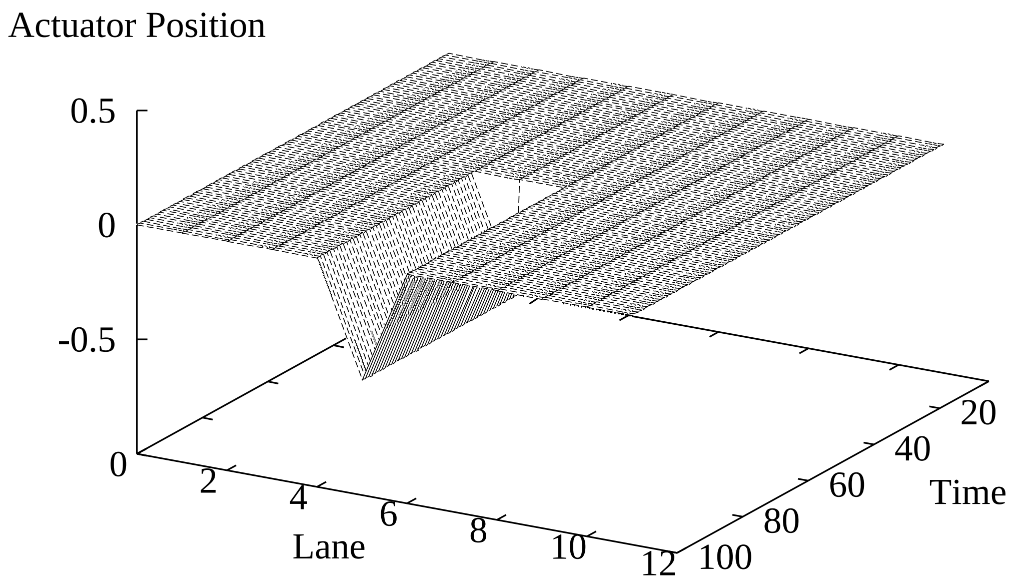

Figure 7.3: Actuator positions: full scan control strategy rejecting an input disturbance.

Figures 7.4 and 7.5 show the results of the partial scan controller. The disturbance is the same; however, the controller now calculates new control moves at each time step. As expected, the profile begins to change ten time steps after the disturbance is first detected. It takes as long as full scan control to fully reject the disturbance since the scanner must visit each lane before the true disturbance profile can be determined. The actuators are rather active and have been detuned by penalizing the rate of change of the inputs.

Figures 7.6 and 7.7 provide another way to compare the two simulations. In

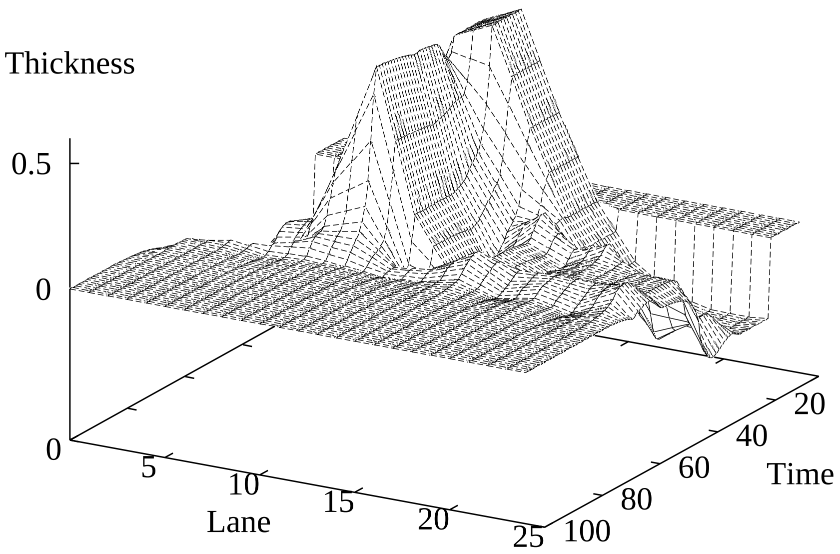

Figure 7.4: Film thickness: partial scan control strategy rejecting an input disturbance.

these figures, the mean square error of the cross directional profile and the maximum cross directional error are presented for a noisy simulation. The noise makes the three dimensional plots more difficult to interpret; thus the noise was left out in the previous simulations. While it may not be clear from the three dimensional plots, it is clear from the two dimensional plots that partial scan control is superior. The disturbance rejection is quicker as expected.

Figures 7.8 and 7.9 show an extended simulation comparing full scan control to partial scan control. The input disturbance is now varying from lane to lane every 40 time steps. This is approximately the size of the period of the scanner (48) so the controller in constantly rejecting disturbances.

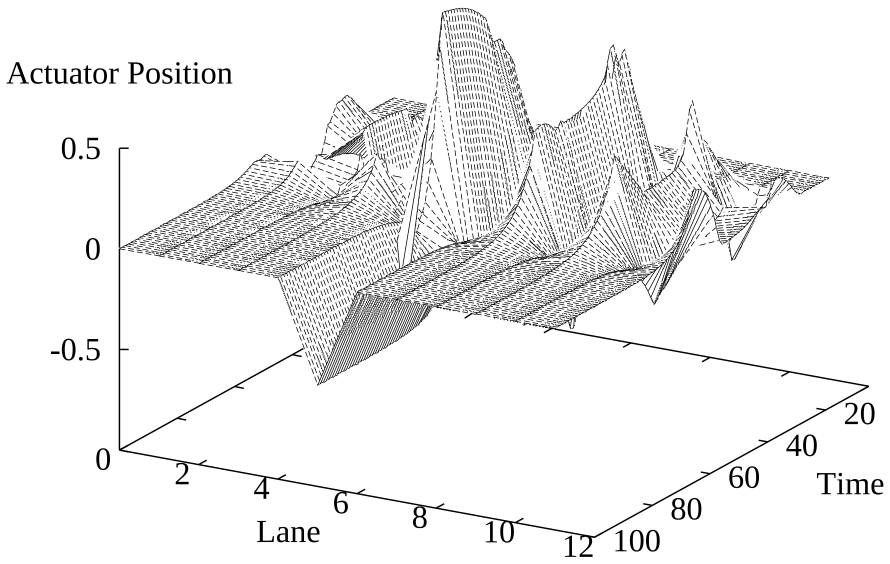

Figure 7.5: Actuator positions: partial scan control strategy rejecting an input disturbance.

Finally, Figures 7.10 and 7.11 show the mean square cross directional error and maximum cross directional error for the two longer simulations.

Study I introduces a full gage simulation in which optimal estimation and constrained linear model predictive regulation combine to provide a complete controller. A comparison was made between taking control action at every time step and taking control action at the end of each complete scan. For large paper making systems (400 outputs) it is currently impossible to compute a control move between measurements. However, with ever increasing computing speeds and more efficient quadratic program solving techniques, it is of interest to evaluate the extreme case. Some intermediate form of partial scan control could certainly be implemented. For example, control

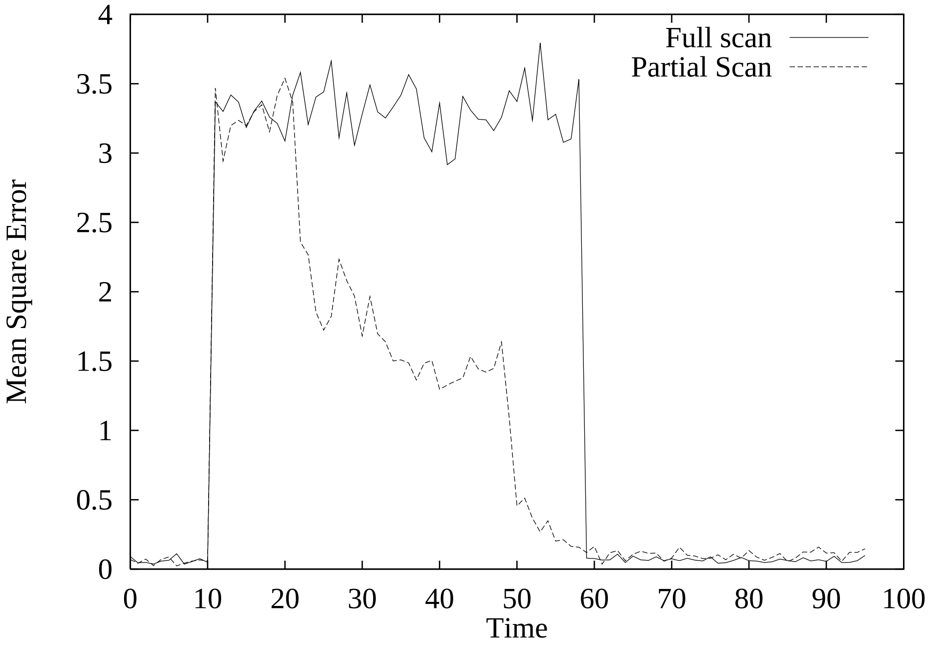

Figure 7.6: Mean square error of the cross directional profile of the film for full and partial scan control strategies.

moves could be calculated several times during a scan depending on the calculation time needed to solve the quadratic program.

It is no surprise that using scanned data as soon as possible provides better control as judged by mean square error and maximum error. It takes both controllers the same amount of time to fully remove the disturbance since the scanner must travel to each lane in order to accurately determine the disturbance profile. This filter behavior is partially due to the filter tuning. For example, no matter the scanner's location when it first detects a disturbance, QDMC tuning provides a filter that only updates the current state. There are several limitations with this tuning that are discussed in the next example.

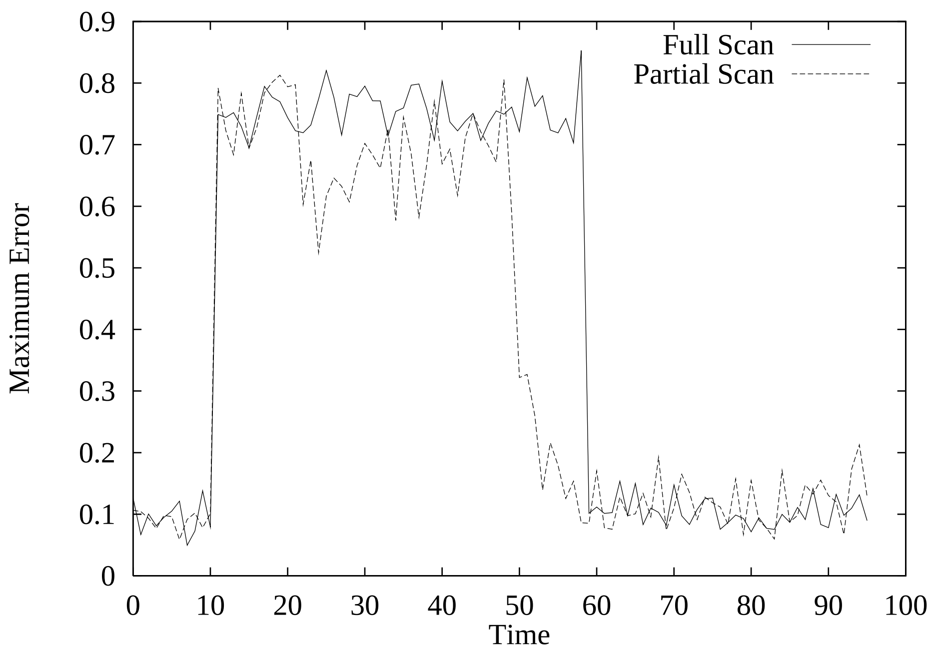

Figure 7.7: Maximum error of the cross directional profile of the film for full and partial scan control strategies.

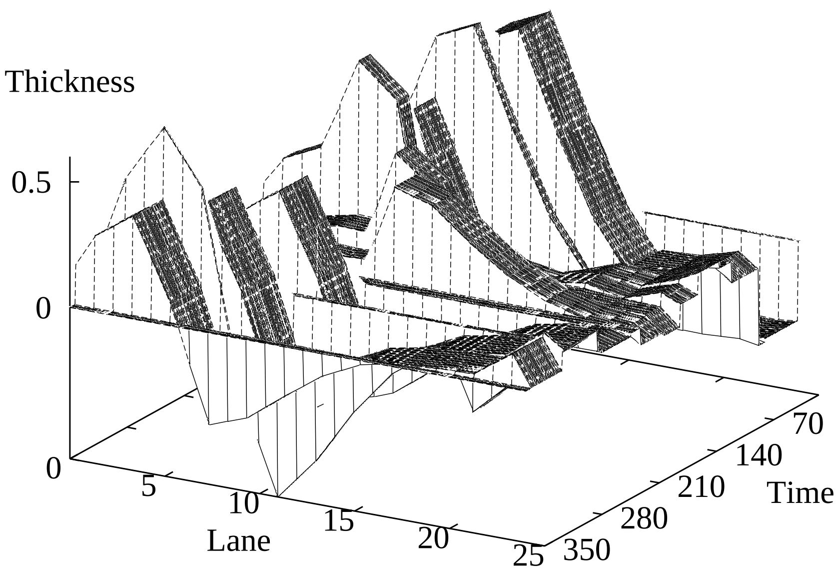

Figure 7.8: Film thickness: full scan control strategy rejecting a series of input disturbances.

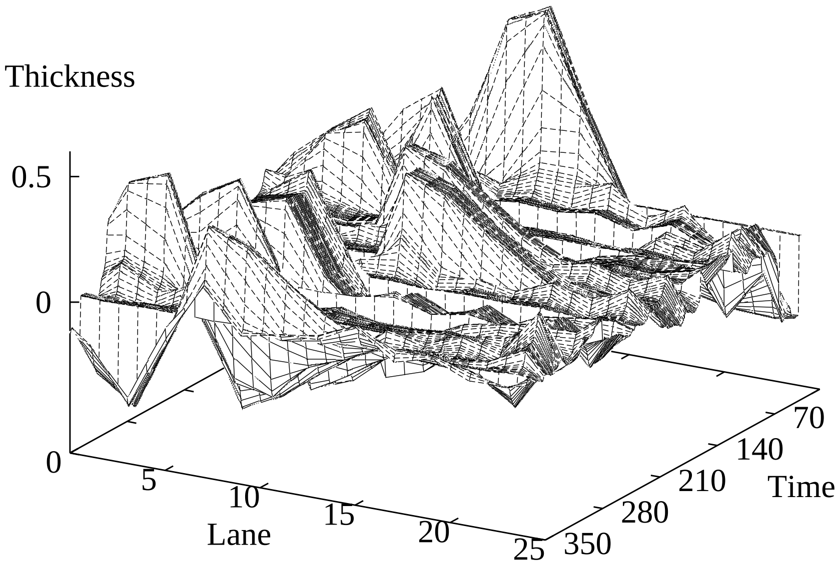

Figure 7.9: Film thickness: partial scan control strategy rejecting a series of input disturbances.

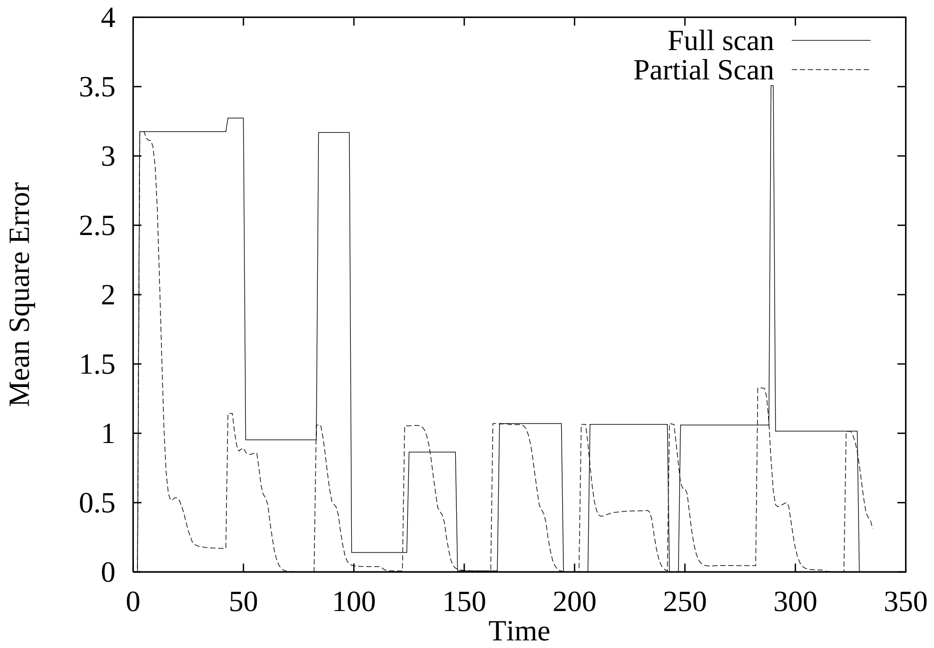

Figure 7.10: Mean square error of the cross directional profile of the film for full and partial scan control strategies.

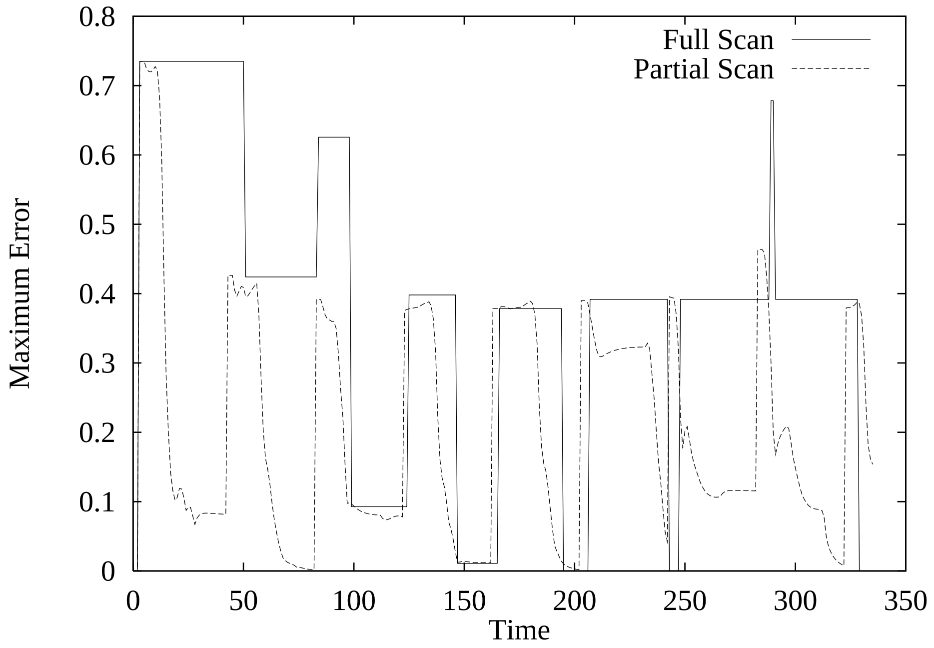

Figure 7.11: Maximum error of the cross directional profile of the film for full and partial scan control strategies.

7.2 Simulation Study II: Disturbance Models

The general disturbance model presented in Chapter 5 leaves open several questions about the actual choice of disturbance models. The purpose of this study is to evaluate several different disturbance models on a typical sheet and film forming process. The process in this example is a five input, ten output system with a time–delay of ten sample times. This relatively small example is used so the large amounts of data do not detract from the main point, which is to evaluate several simple disturbance models. The gain matrix is diagonally dominant, the first row typifies the structure of the matrix, $\begin{bmatrix} 1 & 0.61 & 0.19 & 0.012 & -0.005 \end{bmatrix}$.

The previous example pointed out the limitations of choosing QDMC tuning. This example again uses QDMC tuning, but a second set of simulations uses the input disturbance model. With the input disturbance model, the filter attributes any difference between the model prediction and the measurement to a bias in the input. A variation on this simulation is presented in [24]. The simulation is summarized below and a discussion of the results follows.

Question: Is the input disturbance model appropriate for the sheet and film forming processes?

Model: Diagonally dominant linear model with 5 inputs and ten outputs, delay of ten sample times.

Estimator: Comparison between time–varying Kalman filter based on an input disturbance model and an output disturbance model.

Regulator: Constrained model predictive controller.

Simulation: Evaluate film response to a disturbance in the input.

7.2.1 Study II: Results

Figure 7.12 shows the controller with the input disturbance model rejecting the input disturbance coming through the first actuator. As in the last example the scanner starts at time zero in lane one. Since the time–delay is ten, the scanner is in lane ten by the time the disturbance is under the sensor. When the sensor is near the lane with the disturbance, large corrections are made to the appropriate disturbance states and aggressive control action is taken; however, at the far end of the sheet (lane ten), where the disturbance has little effect, the disturbance state values are actually reduced and the control action is small. Figure 7.13 shows the same simulation for a longer period of time.

The rather slow rejection of the disturbance is due to the diagonal dominance of the gain matrix and is not related to the time–delay. When the scanner is in lane three, the error between the model and the measurement is attributed to a bias in the second actuator, and the estimate of the disturbance in lane one is greatly underestimated. Actually, the estimate of the disturbance in lane one is greatly underestimated *everywhere* except in lane one. The diagonal dominance of the gain matrix forces the input disturbance model to attribute errors to the actuators nearest the scanner's present lane. Thus, the input disturbance model is undesirably slow for this process even though the input disturbance model is specifically designed to reject this type of disturbance. Slow disturbance rejection is shown in Figure 7.14 by tracking the evolution of the five disturbance states. Given enough time they converge to the proper values of one for state one and zero for the remaining states.

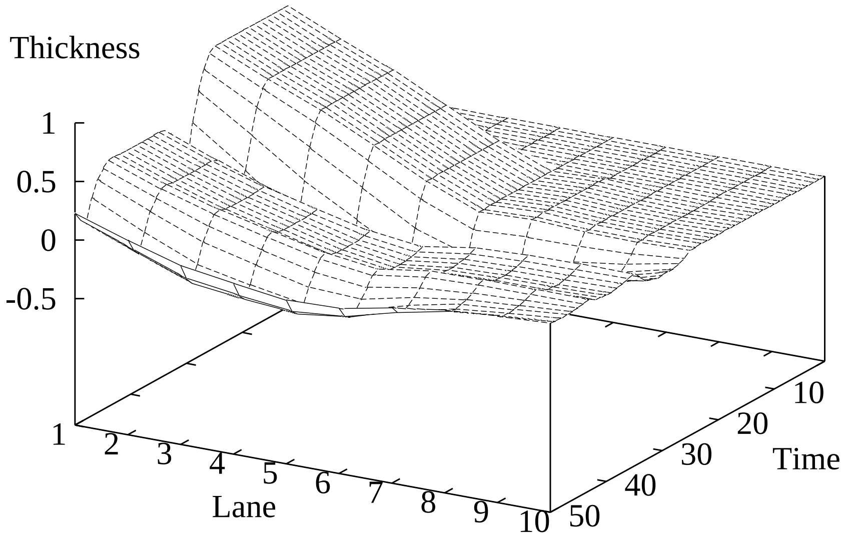

Figure 7.12: Disturbance rejection simulation using scanning sensor, time-varying Kalman filter, and model predictive controller.

Since the input disturbance model is slow and the output disturbance model does not account for correlated disturbances, a different model is needed to improve disturbance rejection. Although the output disturbance model used in Study I does not handle correlation, there is a way to shape the effect of the output disturbances. The standard formulation of the output disturbance model uses $G_p = I$ in Eq. 5.1.

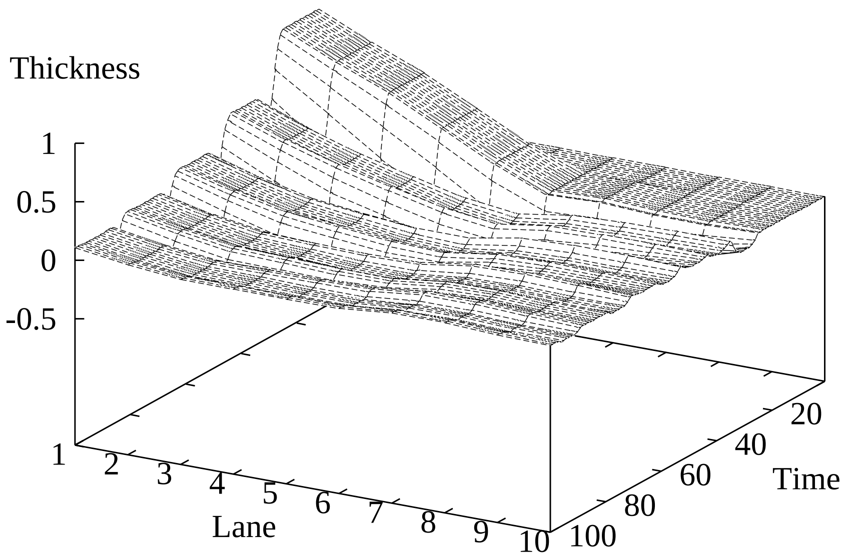

Figure 7.13: Same simulation as Figure 7.12 over a longer time period.

By choosing Gp carefully the effect of the disturbance states can be shaped. In this

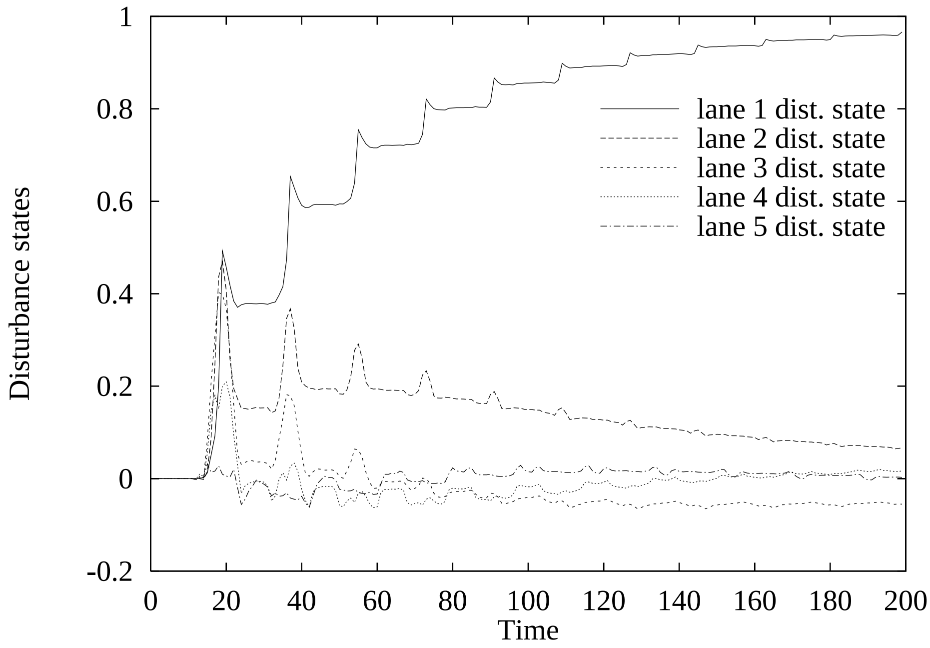

Figure 7.14: Estimated input disturbance states for simulation shown in Figures 7.12 and 7.13.

case, it is logical to pick a diagonally dominant matrix, for example

$$G_p = \begin{bmatrix} 1 & 0.6 & 0.2 & & & & & & & \\ 0.6 & 1 & 0.6 & 0.2 & & & & & & \\ 0.2 & 0.6 & 1 & 0.6 & 0.2 & & & & & \\ & 0.2 & 0.6 & 1 & 0.6 & 0.2 & & & & \\ & & 0.2 & 0.6 & 1 & 0.6 & 0.2 & & & \\ & & & 0.2 & 0.6 & 1 & 0.6 & 0.2 & & \\ & & & & 0.2 & 0.6 & 1 & 0.6 & 0.2 & \\ & & & & & 0.2 & 0.6 & 1 & 0.6 & 0.2 \\ & & & & & & 0.2 & 0.6 & 1 & 0.6 \\ & & & & & & & 0.2 & 0.6 & 1 \end{bmatrix} \tag{7.1}$$

Figure 7.15 shows the means square error of the cross directional profile of the film for the input disturbance model, standard output disturbance model ($G_p = I$), and the modified output disturbance model (G_p as given by Eq. 7.1). The modified output disturbance model appears to be a reasonable compromise between the two models. Although the disturbance is not fully rejected after one scan, the initial error does not approach the large initial errors of the other models. This example provides

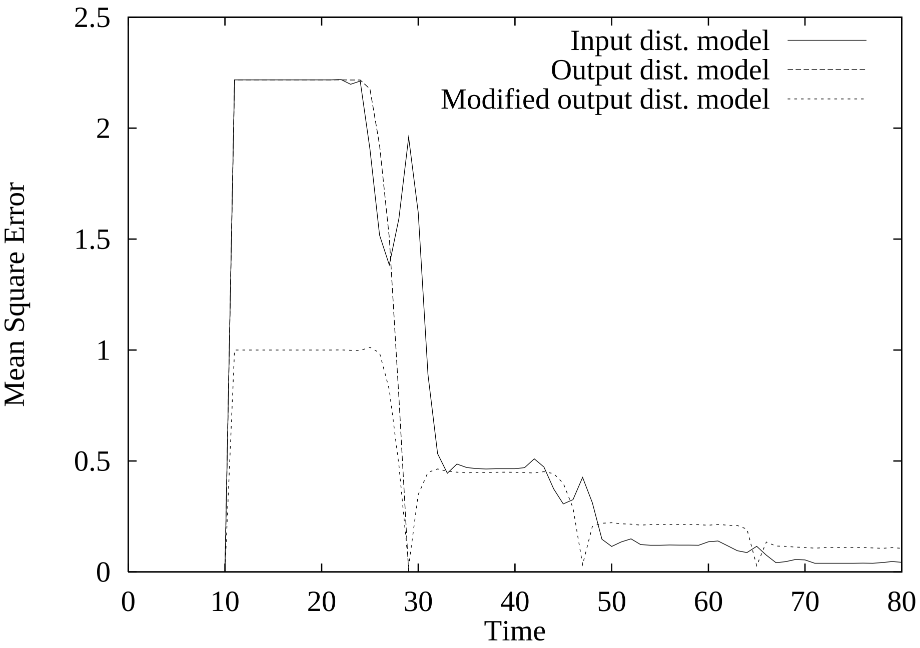

Figure 7.15: Mean square error of the cross directional profile of the film.

a comparison between several disturbance models. For the gage process, the input disturbance model does not provide adequate disturbance rejection–not even for the type of disturbance it is most suited to reject. The output disturbance model can reject an input disturbance after one full scan, but it does not take advantage of the correlation that is present in most gage processes. A reasonable alternative combines the speed of the output disturbance model with the structure on the input disturbance model.

7.3 Simulation Study III: Measured Inputs

The data set used to produce the models in Chapter 3 was lacking any information regarding the third actuator. Symmetry was used to overcome the missing data, but the question remains as to what to do about actuator failure. Actuator bias is a more common problem and this example looks at how actuator bias and actuator failure can be handled. The comparison is between the output disturbance model and a measured input model. A variation on this simulation was presented at the 1996 IFAC World Congress meeting [27].

The process has 12 measured lanes and 12 actuators. The model assumes fast actuator dynamics ($A = 0$) and a time–delay equal to the number of lanes. The centrosymmetric gain matrix from Figure 3.10 is used in the model. The simulation is summarized below and a discussion of the results follows.

Question: What is the best way for the controller to handle actuator bias or actuator failure?

Model: 3M model with centrosymmetric structure enforced (see Chapter 3), 12 inputs, 25 outputs, delay of 12 sample times.

Estimator: Time–varying Kalman filter based on an output disturbance model or a measure input disturbance model

Regulator: Constrained model predictive controller

Simulation: Monitor film response to both actuator bias and actuator failure.

7.3.1 Study III: Results

In the first simulation in this study the actuator in lane six has a bias of 0.5 (it consistently opens too far). The output disturbance model is used to compensate for the bias. The entire state of the film is shown in Figure 7.16. Keep in mind that the sensor only "sees" one lane at a time. By the time the disturbance is seen by the sensor (arbitrarily chosen at time = 14) the sensor is in lane ten and moving toward lane one. After the sensor passes through all lanes (past lane ten again to lane 12 at time = 34) it can properly reconstruct the state. After 12 more time steps the disturbance is adequately removed.

Figure 7.17 displays how the output disturbance model handles the situation in which the 6th actuator is disabled and stuck at 0.5. The controller attempts to use the 6th actuator even though it does not respond. The estimates of the error converge to a periodic set of values that provide undesirable performance. Figure 7.18 shows the mean square error of the gage profile verse time. Often, a disturbance model designed for one type of disturbance can provide adequate performance in the face of other types of disturbances. Clearly, that is not the case here, and more information is needed to solve this problem. Correct use of measurements of actuator positions can solve this problem efficiently.

The next simulation demonstrates the use of a measured input disturbance model. The actuator positions are provided to the estimator at each time step; Figure 7.19 shows the film's response. The disturbance state in lane six takes only two time steps to wind up and correctly identify the bias. If the estimator were used with an unconstrained regulate,r the regulator would continue to send larger input values to counter the bias thus making the bias larger. The estimator used with the constrained regulator, however, winds up the bias estimate and adjusts the remaining actuators to flatten the film. Note also that this simulation is using much more data (12 input measurements and 12 output measurement) to determine the state of the system than the previous two simulations. This model also works well for the case of actuators with bias.

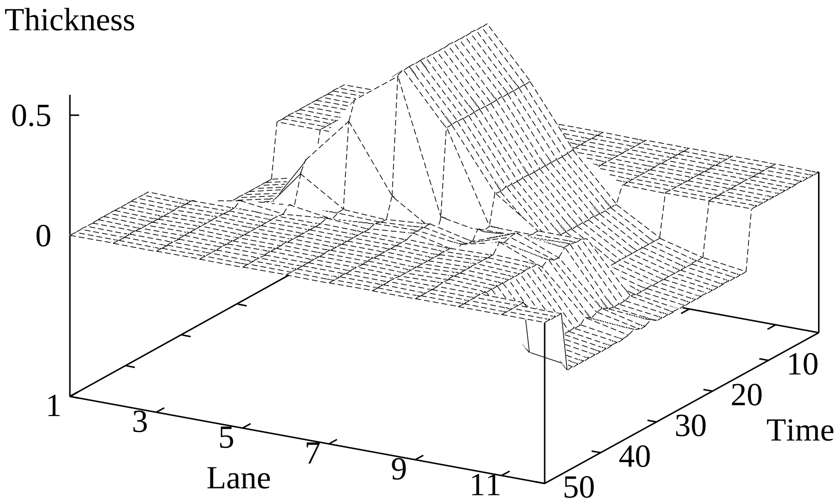

Figure 7.16: Biased actuator simulation using scanning sensor, optimal output disturbance filter, and model predictive controller.

An optimal output disturbance estimator was used in simulation with both actuator bias and actuator failure. This filter rejected the actuator bias, but needed

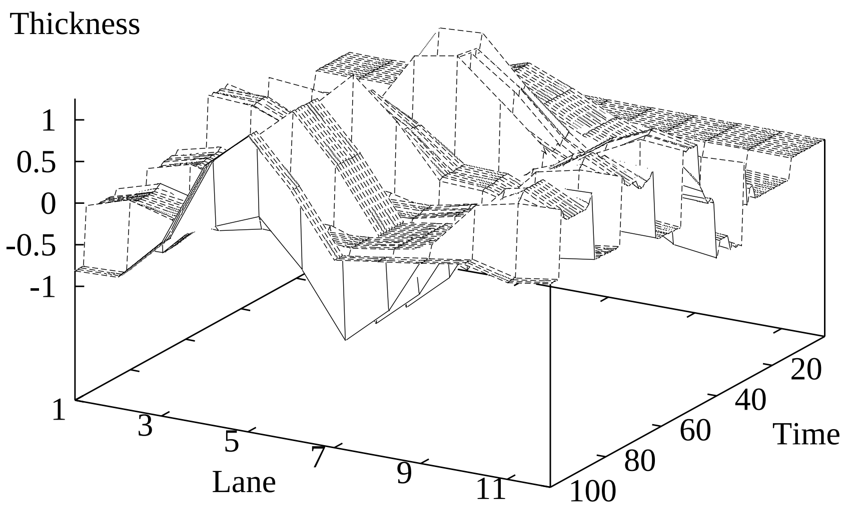

Figure 7.17: Disabled actuator simulation using scanning sensor, optimal output disturbance filter, and model predictive controller.

time to scan the entire film. The filter was unable to provide adequate performance when an actuator failed. By measuring the inputs at each time step, a measured input disturbance filter was able to remove both types of actuator problems independent of the process time–delay. This second estimator requires additional measurements. However, when actuator bias and failure is prevalent, the cost of installing these sensors may be justified by the improved performance.

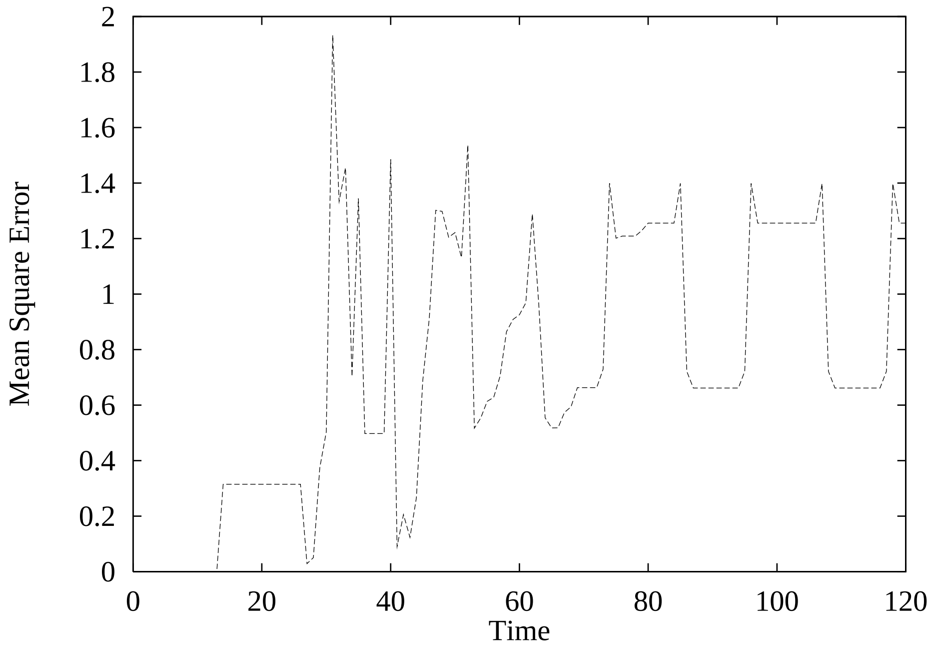

Figure 7.18: Mean square error of film thickness corresponding to disabled actuator simulation shown in Figure 7.17

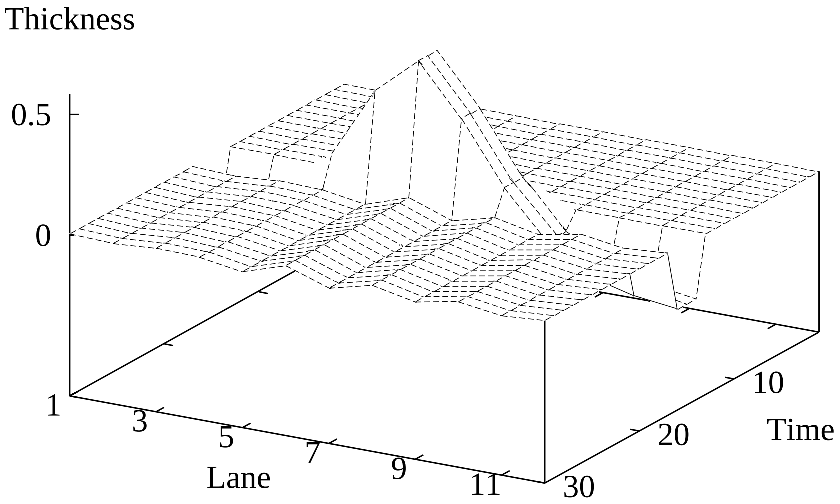

Figure 7.19: Disabled actuator simulation using scanning sensor, optimal measured input disturbance filter, and model predictive controller.

7.4 Simulation Study IV: Model Mismatch

The last simulation study provides some evidence that the combined linear estimator and non–linear regulator is robust to common types of model mismatch. The proposed control has guaranteed nominal properties, but that provides no guarantee of performance in the presence of model mismatch. With the significant time–delay, there is little chance of removing high frequency disturbances so they are ignored in this study. It is likely, however that film systems suffer from low frequency disturbances. For example, consider a model that has an incorrect estimate of the time–delay.

The simulations shown here are noise free, but the same simulations were also run with noisy sensor measurements. The noisy sensor does not change the conclusions, but it does make the plots more difficult to interpret. The simulation is summarized below and a discussion of the results follows. A variation on this simulation was presented at the Third SIAM Conference on Control and its Applications in 1995 [25].

Question: What happens if the time–delay changes?

Model: 3M model with Toeplitz structure enforced (see Chapter 3), 12 inputs, 12 outputs, delay of ten sample times.

Estimator: Time–varying Kalman filter based on an output disturbance model tuned as a QDMC filter.

Regulator: Constrained model predictive controller.

Simulation: Evaluate film response to a disturbance in the output when the plant and model have *different* time–delays.

7.4.1 Study IV: Results

Figure 7.20 illustrates the nominal case. The shape of the output disturbance is best seen by looking at the profile of the film at time zero in the 3–dimensional plot of the film thickness. The 2–dimensional plots show a reprsentave number of actuators and lane thicknesses evolving in time. As would be expected, the correct inputs are calculated once the scanner crosses the film completely. The disturbance is rejected ten time steps later. These plots are shown again in the upper left corner of Figures 7.21, 7.22, and 7.23 as a base case for comparison. In the next simulation the plant's actual time–delay is eight sample times while the model's time–delay remains at ten sample times. These simulations are shown in the upper right corner of Figures 7.21, 7.22, and 7.23. With time–delay mismatch the degradation in the output performance is only marginal. The output shown in the upper right corner of Figure 7.21 exhibits slight over compensation to the error. The inputs, however, exhibit very undesirable behavior. The regulator is tuned to be aggressive so it reacts swiftly to remove the disturbance. Unfortunately the "optimal" state estimate is not very accurate since the filter is tuned for a different type of disturbance.

The output performance is acceptable, but the input performance is not. There are methods to handle this type of undesirable performance. One way is to penalize the rate of change of the input. This is done by including a S penalty on the rate of change of the input in any of the objective functions presented in Chapter 6. The additional term is $\sum_{j=0}^{N-1} \Delta v_j^T S \Delta v_j$ (where $\Delta v_j = v_j - v_{j-1}$). Input penalties do not constrain the input from moving to its eventual steady–state value, rather the penalty shapes its path. The simulations shown in the lower left corner in Figures 7.21, 7.22, and 7.23 has a $(S/Q_\omega = 10)$. The simulations show in the lower right corner has a $(S/Q_\omega = 1000)$.

This study provides reassuring evidence that the combined linear estimator, non–linear model predictive controller can handle model mismatch. Without actually quantifying the robustness, it was shown that detuning can be used to recover desirable performance in the face of 20 percent model mismatch in the time–delay. If the time–delay drifted slowly it may be better to measure the film speed and use it as another measured output in the model. See Section 3.2.4 for more discussion on this issue. The S penalty on the rate of change of the inputs is a useful tuning parameter that provides very predictable results.

A final type of model mismatch of interest is lane drift. Each actuator is assumed to be centered about an output lane. This output lane is the primary lane for that particular actuator. If the primary lane is incorrectly identified, or if the sheet shifts during roll up, then a difficult type of model mismatch is present. Recalibration of the scanning sensor solves this problem. However, some online robustness to this problem is desirable. Zhang presents an algorithm that uses several likely models in the state estimation phase of the control calculation. The model with the lowest residual is chosen for the remaining control calculations. Since the primary lane can only shift a few lanes in either direction, there are only a small number of alternative models that need to be checked. The reader is referred to Zhang [124] for more details.

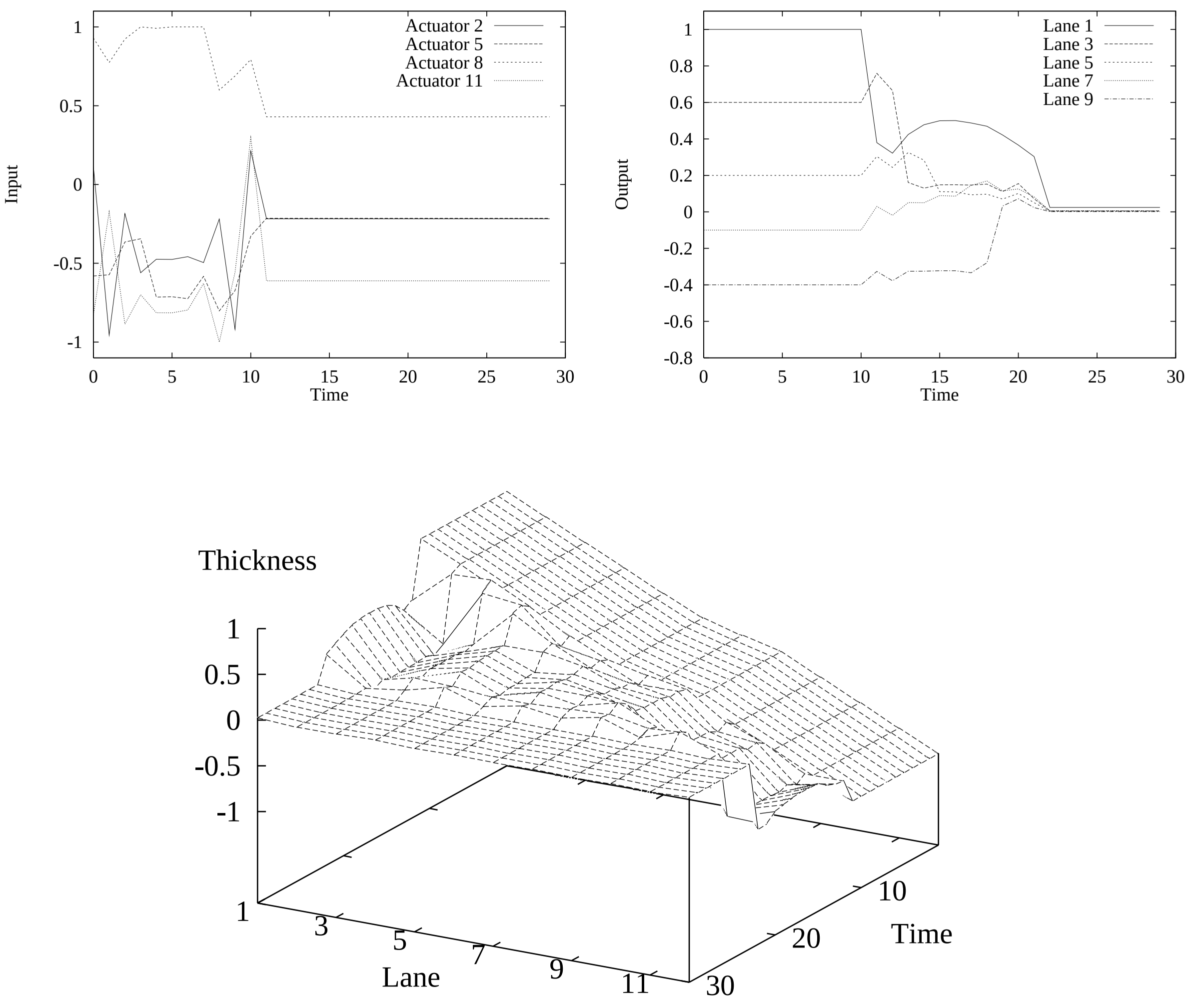

Figure 7.20: The nominal case for disturbance rejection using an output disturbance model. Shown are a representative number of the inputs (top left), the outputs (top right), and the full output profile (bottom center).

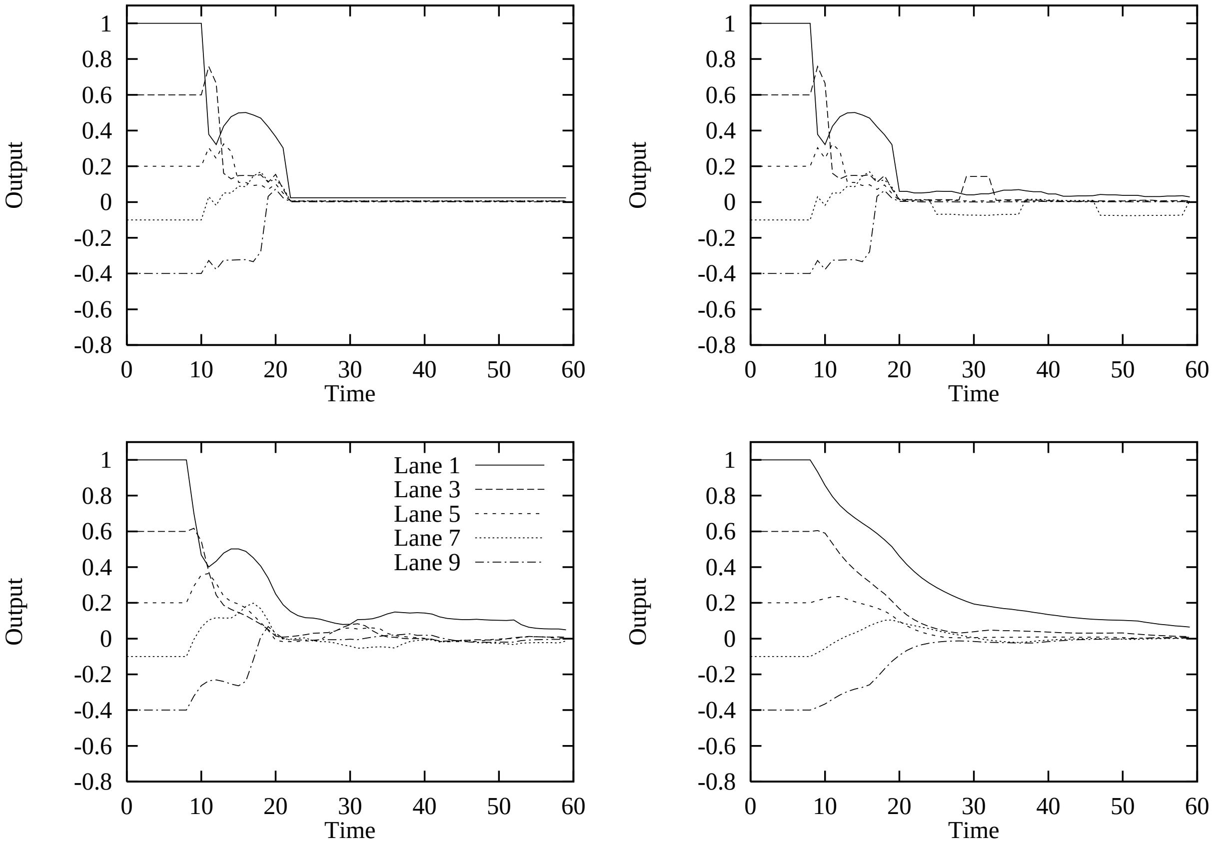

Figure 7.21: Selected outputs for no model mismatch (upper left), time–delay mismatch (upper right), time–delay mismatch with $S/Q_\omega = 10$ (lower left) and time–delay mismatch with $S/Q_\omega = 1000$ (lower right).

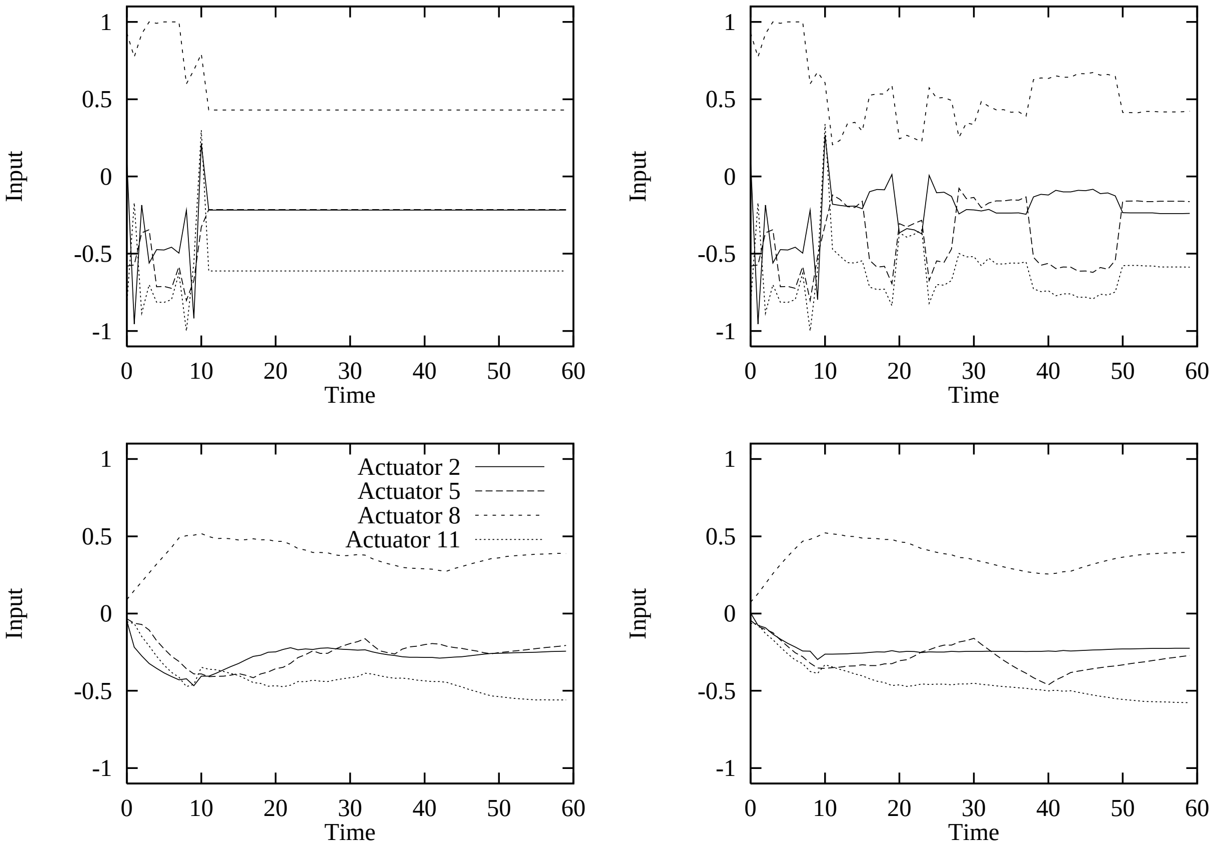

Figure 7.22: Selected outputs for no model mismatch (upper left), time–delay mismatch (upper right), time–delay mismatch with $S/Q_\omega = 10$ (lower left) and time–delay mismatch with $S/Q_\omega = 1000$ (lower right).

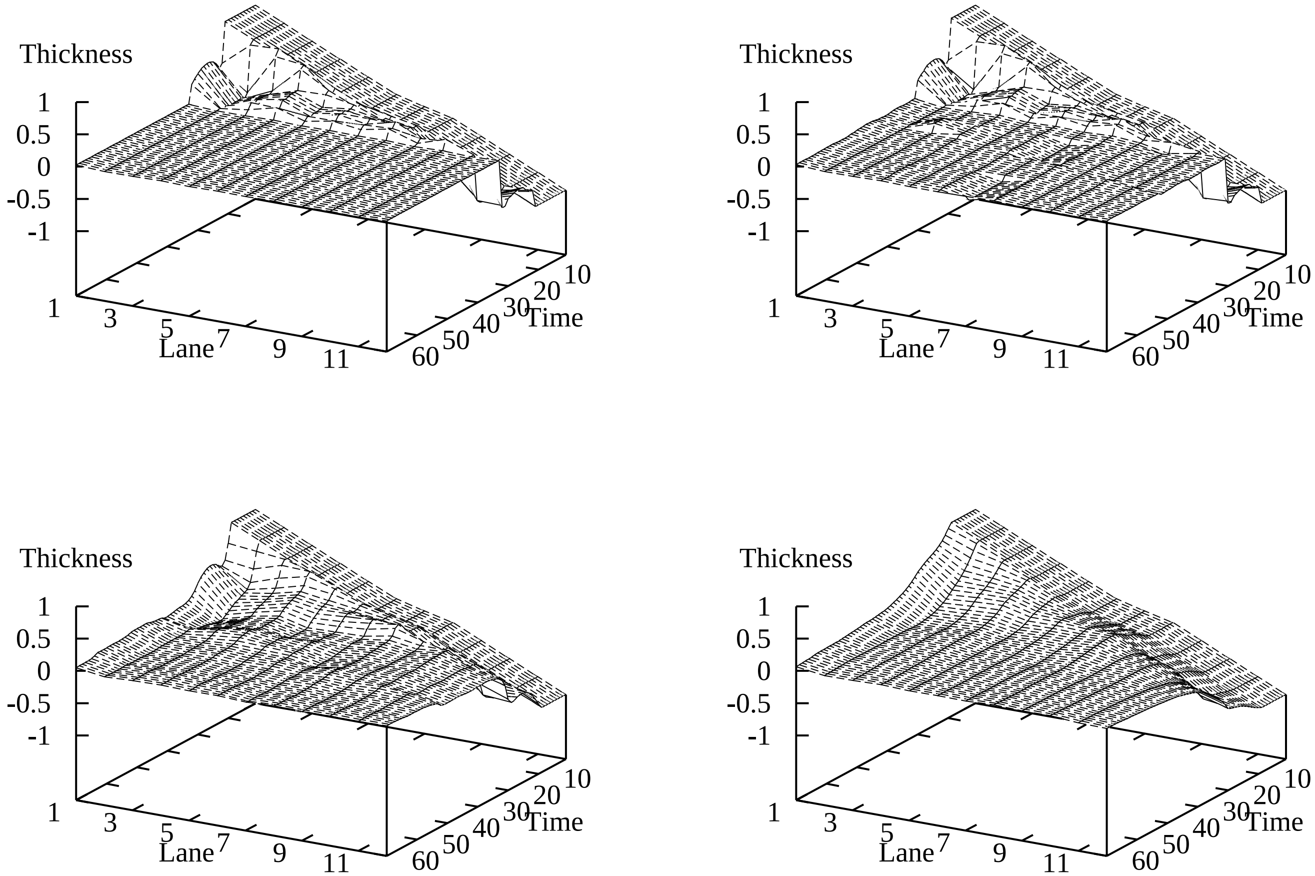

Figure 7.23: Output profiles for no model mismatch (upper left), time–delay mismatch (upper right), time–delay mismatch with $S/Q_\omega = 10$ (lower left) and time–delay mismatch with $S/Q_\omega = 1000$ (lower right).

Chapter 8

Conclusions

The purpose of this study is to recommend new and existing control technology to make superior film and paper products. Without adequate control technology, parameters such as film thickness can deviate beyond acceptable standards. The work presented in this thesis advocates a constrained infinite horizon controller. The infinite horizon controller uses a process model and an initial state estimate. Consequently the modelling and state estimation issues are of critical importance. A linear model is proposed as an adequate representation of the system. Since the model is used with a feedback controller, it need only approximate the system. Due to a periodic scanning pattern, the periodic Kalman filter is used to determine optimal estimates from the sparse data. The optimal state estimates can then be used with an infinite horizon regulator to calculate the optimal input profile. Constraints are handled in the regulator to avoid excessive bending of the die lip. The main issues covered in this thesis are summarized independently below.

Modelling

For film forming processes, negligible dynamics, a large time–delay, and a scanning sensor specify much of the model structure. Data from a 3M pilot plant was used to calculate model parameters. Process symmetry was exploited to overcome missing data and improve parameter estimates. Residuals from the data were analyzed to evaluate which type of symmetry was most appropriate. Centrosymmetry and Toeplitz symmetry were found to be relevant types of symmetry to consider for the film process. It was found that overspecifing the model by enforcing extra symmetry adversely effected the residuals. Thus, care should be taken not to enforce symmetry not present in the process. Novel modelling work presented in this thesis includes a complete linear model and the calculation of parameters from process data.

Estimation

The Kalman filter is a convenient optimal estimation tool to use for gage control. The periodic nature of the scanning sensor provides an estimation problem that can be solved via standard Kalman filtering techniques. In this work filter gains for the film process were calculated by iterating the Riccati equation to convergence and by lifting the system and solving the steady–state Riccati equation. The lifted system was shown to be an equivalent time–invariant representation of the periodic system. This equivalence is useful in the analysis of system properties such as observability and detectability. The covariances of the state estimates were analyzed by looking at the diagonal elements of the covariance matrix. These diagonal elements are related to the hyperbox that contains the hyperellipsoid of constant probability. Covariance analysis is a useful tool for evaluating sensor patterns. For example, it was shown that between sensor measurements, models with integrating states have monotonically increasing estimate covariances. Unbounded covariances imply that detectability is needed to ensure stable estimates.

Before discussing detectability, a general disturbance model was introduced. A disturbance model is useful because integrating disturbance states used in conjunction with model predictive controllers provide offset free control. Several models of interest are contained in the general model presented. Both the output disturbance model and the input disturbance model were discussed in detail. The output disturbance model is the integrating disturbance model used by QDMC controllers. The input disturbance model is used to model upstream disturbances that come through the dynamics of the system. A measured input disturbance model was presented as well. This model requires additional system measurements, but can be used to handle actuator bias and actuator failure.

In continuing with the periodic film control problem, lifted disturbance models were constructed next. The lifted systems were then used to analyze the periodic systems. Definitions of observability and detectability of both time–invariant systems and periodic systems were provided. Tests to determine observability and detectability were also provided for both time–invariant and periodic systems. Of particular use is the Hautus Lemma, which provides a simple test for observability and detectability. A periodic Hautus Lemma was presented as a useful analog for the film control system.

These definitions and tests provide a method for analyzing the disturbance models used in periodic systems. Conditions of detectability, which implies observability of the disturbance states, for the various disturbance models were presented. The number of disturbance states are limited to the number of outputs; furthermore, detectability is lost if any of the disturbance states mask any of the integrating modes of the system.

Regulation

Many of the latest theoretical advances in linear model predictive control are presented in this thesis. The constrained infinite horizon controller is the controller of choice for several reasons. Both the inputs and the states are penalized and constrained on the infinite horizon. The controller is nominally stabilizing, and, in the absence of disturbances, the open loop and closed loop trajectories are identical. Closed–loop stability of the combined periodic linear estimator and the non–linear constrained regulator was established.

Two methods of target tracking were presented. The first technique involves a multistep algorithm. If the setpoint cannot be met exactly then a second problem is solved that minimizes the square error from the desired setpoint. Conditions of feasibility were presented for the exact target problem, and detectability is shown to be sufficient to guarantee uniqueness of the solution for both problems. The second target tracking method uses exact soft constraints to solve the problem in one step. The soft constraints can be chosen such that constraints are met if feasible. Detectability is sufficient to guarantee a unique solution for the exact soft constraint target problem.

Soft constraints were also discussed in the context of the regulator. Output constraints can be relaxed in a prescribed manner by using soft constraints. As before, the constraints are met if feasible. Interior point optimization algorithms, in the context of model predictive control, were briefly discussed. Interior point methods show promise for solving large problems such as the film problem.

The simulation studies presented in Chapter 7 provide the following conclusions.

- Partial scan control, or the use of the scanned data as it is obtained, outperforms full scan control. Full scan control is the industrial standard.

- Input disturbance models are slow to reject disturbances including input disturbances, due to the diagonal dominance of the gain matrix. Output disturbance models tuned as QDMC controllers do not take into account any of the correlation present in the gain matrix. A modified output disturbance model incorporating the correlation of the input model and the speed of the output model is found to be a reasonable compromise.

- For actuator failure and actuator bias an input disturbance model can outperform other forms of disturbance models. The measured input model, however, requires measurements of the actuator positions. The measured input model has no time–delay, which provides the key advantage.

- The combined linear estimator and non–linear regulator is robust to a typical type of model mismatch (*e.g.* 20% time–delay mismatch).

8.1 Future Directions

The list of open computational and theoretical issues is certain to increase as the theory in this thesis is applied. A short discussion of further issues include the following.

- Modelling errors may be difficult to handle. This study gives no quantitative results on the size of the operation region in which the linear model is valid. If the model is inadequate, then gain scheduling or adaptive control technology may be required to update the model. Non–linear models may also be needed. In addition, many specialty film lines frequently change grades and may require many models. The design of well posed identification experiments is important to ensure high fidelity models.

- Software development paves the way for application. Packages used in higher level programming languages allow a greater number of people to examine the control algorithms. The `lmpc` package has been extensively tested through research projects, graduate courses, and industrial short courses. Surprising simulation results have produced better insight and understanding of tuning, constraints and feasibility issues. The same insight will likely come from a simulations using soft exact constraints and efficient interior point methods.

- There are currently many constraint shedding and constraint ranking methods in use. The technology used to develop past constraint handling methods may be applied to the tuning of soft constraint programs. Soft constraints techniques can avoid removing too many constraints, and may provide better performance and simpler algorithms.

This discussion only covers a few of the research topics in the area of sheet and film control. The interest in film control recently shown by academics and practitioners alike should progress in this area for the next several years.

Appendix A

A Probabilistic Derivation of the Kalman Filter

A variety of methods have been used to derive the optimal linear filter also known as the Kalman filter. Approach is often dictated by need, and thus the industrial practitioners, in this case the aerospace industry, worked out of the statistical framework that was a convenient starting place for them. Meanwhile estimation theorists found that statistical methods ignored the fundamental probabilistic structure, which allowed linear and nonlinear problems to be attacked via one method. Kalman himself used another derivation method that utilized orthogonal projections. Kalman's interest was primarily the linear problem. Perhaps the various derivations relate to the passive connotation of "observer" verses the active connotation of "estimator." Statistical methods such as recursive least squares, maximum likelihood, and the linear minimum variance estimator all assume a deterministic model. Probabilistic methods, on the other hand, compute a probability density function of the estimate in question. Choices can then be made as to which value relating to the probability density best represents the density. Median, mean and mode are all reasonable choices and some variance, or quality, of the estimate is readily available. Jazwinski attacks the full nonlinear problem via probabilistic methods, but in the linear case he also reviews orthogonal projection, recursive least squares, maximum likelihood and linear minimum variance methods [65].

It is from Jazwinski, and in the spirit of James B. Rawlings' graduate control class notes, that the following probabilistic derivation is presented. First the relevant notation and definitions are provided, after which, the linear Kalman filter is derived in the notation of this thesis.

A.1 Probability Definitions

The purpose of this section is to define some probabilistic notation and relevant properties for use in the derivation of the Kalman filter. This section is not intended as a thorough review of probability theory and the reader is encouraged to consult an appropriate statistics text [62].

Definition A.1 *The random variable ξ having a normal distribution with mean m and covariance P_ξ is denoted $\xi \sim N(m, P_\xi)$. (Note that if $\xi, m \in \mathbb{R}^n$ then $P_\xi \in \mathbb{R}^{n \times n}$ is a matrix.)*

Definition A.2 *The probability density function of a random variable ξ is denoted $p_\xi(x)$ and is shown here for a normally distributed vector valued variable.*

$$p_\xi(x) = \frac{1}{(2\pi)^{n/2}|P_\xi|^{1/2}} \exp\left[-\frac{1}{2}(x-m)^T P_\xi^{-1}(x-m)\right] \tag{A.1}$$

Definition A.3 *Two random variables are said to be jointly distributed if they are defined on the same probability space. The joint probability density function of random variables ξ and η is denoted $p_{\xi,\eta}(x,y)$ and has mean*

$$m = \begin{bmatrix} m_x \\ m_y \end{bmatrix} \tag{A.2}$$

and covariance matrix

$$P_{\xi,\eta} = \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \tag{A.3}$$

Definition A.4 *The conditional density function is defined as*

$$p_{\xi|\eta}(x \mid y) = \frac{p_{\xi,\eta}(x,y)}{p_\eta(y)} \tag{A.4}$$

Theorem A.1 (Marginal normal density) *Given that ξ and η are jointly normally distributed with mean*

$$m = \begin{bmatrix} m_x \\ m_y \end{bmatrix} \tag{A.5}$$

and covariance matrix

$$P_{\xi,\eta} = \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \tag{A.6}$$

then the marginal density of η is normal with the following parameters

$$\eta \sim N(m_y, P_y) \tag{A.7}$$

Proof: This can be shown in two ways. First x can be integrated out of the joint density function.

$$P_\eta(y) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P_{\xi,\eta}(x,y)dx_1 \ldots dx_p \tag{A.8}$$

Another, potentially simpler method, involves the following transformation, let

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{A.9}$$

where

$$w_1 = x - P_{xy}P_y^{-1}y, \qquad w_2 = y \tag{A.10}$$

Now, w is normally distributed as linear combinations of normally distributed variables are also normally distributed. Furthermore, w_1 and w_2 are uncorrelated.

$$\begin{aligned} \mathrm{cov}(w_1, w_2) &= \mathcal{E}\{(w_1 - \mathcal{E}\{w_1\})(w_2 - \mathcal{E}\{w_2\})^T\} \qquad (A.11) \\ &= \mathcal{E}\{((x - m_x) - P_{xy}P_y^{-1}(y - m_y))(y - m_y)^T\} \\ &= \mathcal{E}\{(x - m_x)(y - m_y)^T - P_{xy}P_y^{-1}(y - m_y)(y - m_y)^T\} \\ &= P_{xy} - P_{xy}P_y^{-1}P_y \quad = \quad 0 \end{aligned}$$

The variance of w_1 is calculated in a similar manner.

$$\begin{aligned} \mathrm{var}(w_1, w_1) &= \mathcal{E}\{(w_1 - \mathcal{E}\{w_1\})(w_1 - \mathcal{E}\{w_1\})^T\} \\ &= \mathcal{E}\{((x - m_x) - P_{xy}P_y^{-1}(y - m_y))((x - m_x) - P_{xy}P_y^{-1}(y - m_y))^T\} \\ &= P_x - P_{xy}P_y^{-1}P_{yx} \end{aligned} \tag{A.12}$$

so that

$$\mathcal{E}\{(w - \mathcal{E}\{w\})(w - \mathcal{E}\{w\})^T\} = \begin{bmatrix} P_x - P_{xy}P_y^{-1}P_{xy} & 0 \\ 0 & P_y \end{bmatrix} \tag{A.13}$$

Now, theorem 2.10 from Jazwinski states that if the Gaussian vectors x and y are uncorrelated, they are independent, furthermore x and y are marginally Gaussian (with appropriate parameters). Thus, w_1 and w_2 are normal and y (or η) has the desired marginal density $\eta \sim N(m_y, P_y)$. □

Theorem A.2 (Bayes' theorem) *This famous result is often referred to as Bayes' rule.*

$$p_{\xi|\eta}(x \mid y) = \frac{p_{\eta|\xi}(y \mid x)p_\xi(x)}{p_\eta(y)} \tag{A.14}$$

Proof: This follows easily from the definition of conditional probability.

$$p_{\xi,\eta}(x, y) = p_{\xi|\eta}(x|y)p_\eta(y) \tag{A.15}$$

$$p_{\eta,\xi}(y, x) = p_{\eta|\xi}(y|x)p_\xi(x) \tag{A.16}$$

Because $p_{\xi,\eta}(x, y) = p_{\eta,\xi}(y, x)$, the right hand sides of Equations A.15 and A.16 are equal thus establishing Bayes' theorem. □

Theorem A.3 (Conditional normal density) *If ξ and η are jointly normally distributed as*

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} \sim N\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix} \right) \tag{A.17}$$

then the conditional density of ξ given η is also normal

$$(\xi \mid \eta) \sim N(m, P) \tag{A.18}$$

where

$$m = m_x + P_{xy} P_y^{-1} (y - m_y) \tag{A.19}$$

$$P = P_x - P_{xy} P_y^{-1} P_{yx} \tag{A.20}$$

Proof: The definition of conditional density gives

$$p_{\xi|\eta}(x|y) = \frac{p_{\xi,\eta}(x,y)}{p_\eta(y)} \tag{A.21}$$

Because (ξ, η) is jointly normal, Theorem A.1 provides

$$p_\eta(y) = \frac{1}{(2\pi)^{n_\eta/2} |P_y|^{1/2}} \exp\left[-\frac{1}{2} (y - m_y)^T P_y^{-1} (y - m_y) \right] \tag{A.22}$$

The argument of the exponent found by combining the two probability densities in Equation A.21 is therefore

$$e = \begin{bmatrix} x - m_x \\ y - m_y \end{bmatrix}^T \begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}^{-1} \begin{bmatrix} x - m_x \\ y - m_y \end{bmatrix} - (y - m_y)^T P_y^{-1} (y - m_y) \tag{A.23}$$

P is defined as follows

$$P = P_x - P_{xy} P_y^{-1} P_{yx} \tag{A.24}$$

and, the partitioned matrix inversion formula is used to express the matrix inverse in Equation A.23 as

$$\begin{bmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{bmatrix}^{-1} = \begin{bmatrix} P^{-1} & -P^{-1} P_{xy} P_y^{-1} \\ -P_y^{-1} P_{yx} P^{-1} & P_y^{-1} + P_y^{-1} P_{yx} P^{-1} P_{xy} P_y^{-1} \end{bmatrix} \tag{A.25}$$

Substituting this result in Equation A.23 and multiplying out terms yields

$$\begin{aligned} e &= (x - m_x)^T P^{-1} (x - m_x) - 2(y - m_y)^T (P_y^{-1} P_{yx} P^{-1})(x - m_x) \\ &+ (y - m_y)^T (P_y^{-1} P_{yx} P^{-1} P_{xy} P_y^{-1})(y - m_y) \end{aligned} \tag{A.26}$$

which is the expansion of the following quadratic term

$$e = \left[(x - m_x) - P_{xy} P_y^{-1} (y - m_y)\right]^T P^{-1} \left[(x - m_x) - P_{xy} P_y^{-1} (y - m_y)\right] \tag{A.27}$$

in which it is known that $P_{xy} = P_{yx}^T$. Substituting Equation A.19 into this expression yields

$$e = (x - m)^T P^{-1} (x - m) \tag{A.28}$$

Next, the desired pre–exponential factor is a combination of the pre–exponential factor shown in Equation A.22 and the analogous pre–exponential factor of the joint probability density function (not shown) which contains the following determinant.

$$\begin{vmatrix} P_x & P_{xy} \\ P_{yx} & P_y \end{vmatrix} = \begin{vmatrix} P_x - P_{yx}P_y^{-1}P_{xy} & 0 \\ P_{yx} & P_y \end{vmatrix} = |P_x||P_y| - |P_{xy}||P_{yx}| \tag{A.29}$$

Combination of the pre–exponential factors and the exponent shown in Equation A.28 provides the following conditional density function

$$p_{\xi|\eta}(x|y) = \frac{1}{(2\pi)^{n_\eta/2}|P|^{1/2}} \exp\left[-\frac{1}{2}(x-m)^T P^{-1}(x-m)\right] \tag{A.30}$$

which is the desired result. □

Theorem A.4 (Conditional densities of several variables) *Let the joint conditional of random variables a and b given c be a normal distribution with*

$$(a, b \mid c) \sim N\left(\begin{bmatrix} m_a \\ m_b \end{bmatrix}, \begin{bmatrix} P_a & P_{ab} \\ P_{ba} & P_b \end{bmatrix}\right) \tag{A.31}$$

Then the conditional density of a given b and c is also normal

$$(a \mid b, c) \sim N(m, P) \tag{A.32}$$

where

$$m = m_a + P_{ab}P_b^{-1}(b - m_b) \tag{A.33}$$

$$P = P_a - P_{ab}P_b^{-1}P_{ba} \tag{A.34}$$

Proof: From the definition of joint density

$$p(a|b,c) = \frac{p(a,b,c)}{p(b,c)} \tag{A.35}$$

Multiplying the top and bottom of the fraction by $p(c)$ yields

$$p(a|b,c) = \frac{p(a,b,c)}{p(c)}\frac{p(c)}{p(b,c)} \tag{A.36}$$

or

$$p(a|b,c) = \frac{p(a,b|c)}{p(b|c)} \tag{A.37}$$

the marginal density, $p(b|c)$, can be calculated from the joint density, $p(a,b|c)$, as shown in Theorem A.1.

$$p(b|c) \sim N(m_b, P_b) \tag{A.38}$$

Substituting the distributions given in Equations A.31 and A.38 yields

$$(a|b,c) \sim \frac{N\left(\begin{bmatrix} m_a \\ m_b \end{bmatrix}, \begin{bmatrix} P_a & P_{ab} \\ P_{ba} & P_b \end{bmatrix}\right)}{N(m_b, P_b)} \tag{A.39}$$

And now applying the methods of Theorem A.3 this ratio of normal distributions reduces to the desired expression. □

A.2 Calculation of the Probability Density of $\mathbf{x_k}$

Assume: At time k the conditional density of x_k given the data, $\{y_0, \ldots, y_{k-1}\}$, is known to be a normal distribution (the minus sign denotes values excluding data at time k).

$$(x_k|y_0, \ldots, y_{k-1}) \sim N(\hat{x}_k^-, P_k^-) \tag{A.40}$$

Also assume that the initial conditions for the state and covariance matrix are given as $x_0 \sim N(\bar{x}_0, Q_0)$.

Goal: To determine the density of x_k given all of the data up to time k. The density is given as

$$p_{\xi_k|H_k}(x_k|y_0, \dots, y_k) \tag{A.41}$$

where $H_k = \{\eta_0, \dots, \eta_k\}$. Eq. A.41 cannot be calculated directly so the following conditional density is calculated first

$$p_{\xi_k,\eta_k|H_{k-1}}(x_k, y_k|y_0, \dots, y_{k-1}) \tag{A.42}$$

Theorem A.4 can then be used to move the y_k to the other side of the conditional symbol. Thus the following joint density is needed

$$(x_k, y_k|y_0, \dots, y_{k-1}) \sim N\left(\begin{bmatrix} \hat{x}_k^- \\ ? \end{bmatrix}, \begin{bmatrix} P_k^- & ? \\ ? & ? \end{bmatrix}\right) \tag{A.43}$$

In order to deal with y_k in Eq. A.43, recall the relationship between the output and the state.

$$y_k = Cx_k + v_k \tag{A.44}$$

Using the fact that v_k is independent of the state and the output, and $v_k \sim N(0, R)$, the following holds

$$(x_k, Cx_k + v_k|y_0, \dots, y_{k-1}) \sim N\left(\begin{bmatrix} \hat{x}_k^- \\ C\hat{x}_k^- + 0 \end{bmatrix}, \begin{bmatrix} P_k^- & P_k^- C^T \\ CP_k^- & R + CP_k^- C^T \end{bmatrix}\right) \tag{A.45}$$

substituting y_k back in yields

$$(x_k, y_k | y_0, \ldots, y_{k-1}) \sim N\left(\begin{bmatrix} \hat{x}_k^- \\ C\hat{x}_k^- \end{bmatrix}, \begin{bmatrix} P_k^- & P_k^- C^T \\ CP_k^- & R + CP_k^- C^T \end{bmatrix}\right) \tag{A.46}$$

Now applying Theorem A.4.

$$(x_k | y_0, \ldots, y_k) \sim N(\hat{x}_k, P_k) \tag{A.47}$$

in which

$$\begin{aligned} L_k &= P_k^- C^T (R + CP_k^- C^T)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + L_k (y_k - C\hat{x}_k^-) \\ P_k &= P_k^- - L_k C P_k^- \end{aligned} \tag{A.48}$$

The equations in A.48 provide the Kalman filter gain equation, update equation, and covariance matrix equation for data including y_k. By using probabilistic arguments part of the Kalman filter solution is completed. The key is to calculate the conditional densities of the estimates based on the data. Note that the mean of the estimate is the most likely value.

A.3 Calculation of the Probability Density of $\mathbf{x_{k+1}}$

The final step is to compute the density of

$$(x_{k+1} | y_0, \ldots, y_k) \sim N(\hat{x}_{k+1}^-, P_{k+1}^-) \tag{A.49}$$

We use the model for the state propagation (ignoring the inputs)

$$x_{k+1} = Ax_k + w_k \tag{A.50}$$

Using the fact that w_k is independent of the state and the output, and $w_k \sim N(0, Q)$, thus

$$(Ax_k + w_k | y_0, \ldots, y_k) \sim N(A\hat{x}_k, AP_kA^T + Q) \tag{A.51}$$

substituting x_{k+1} back in yields

$$(x_{k+1} | y_0, \ldots, y_k) \sim N(A\hat{x}_k, AP_kA^T + Q) \tag{A.52}$$

The propagation from k to $k+1$ without the next output measurement may then be summarized by

$$\begin{aligned} \hat{x}_{k+1}^- &= A\hat{x}_k \\ P_{k+1}^- &= AP_kA^T + Q \end{aligned}$$

Thus the derivation of the Kalman filter using probabilistic arguments is complete. Making the dual regulator connection clearly invites the Riccati equation, but this method emphasizes the most probable estimate, and provides a satisfying solution.

Appendix B

Additional Proofs

B.1 Proof of Theorem 3.1

Theorem 3.1 (Singular value decomposition (SVD)) *If A is a real $m \times n$ matrix with $rank(A) = r$ then there exist orthogonal matrices*

$$W = \begin{bmatrix} w_1, & \ldots, & w_m \end{bmatrix} \in \mathbb{R}^{m \times m} \qquad \textit{and} \qquad V = \begin{bmatrix} v_1, & \ldots, & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$$

such that

$$A = W^T \Sigma V$$

where

$$\Sigma = diag\{\sigma_1, \ldots, \sigma_p\} \in \mathbb{R}^{m \times n} \qquad p = \min\{m, n\}$$

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0$$

Proof: This proof follows from Davis [37]. See alternative proofs in Golub and Van Loan [52] and Horn and Johnson [63].

Let $\Sigma_1 = diag\{\sigma_1, \ldots, \sigma_r\}$ and W_1 be the $m \times r$ matrix consisting of the orthonormal eigenvectors of AA^T corresponding to the eigenvalues $\{\sigma_1^2, \ldots, \sigma_r^2\}$. Then by definition,

$$AA^T W_1 = W_1 \Sigma_1^2 \qquad \text{and} \qquad W_1^T W_1 = I_r \tag{B.1}$$

Now let W_2 be the $m \times m - r$ matrix with columns forming an orthonormal basis for the null space of A^T, then

$$A^T W_2 = 0 \qquad \text{and} \qquad W_2^T W_2 = I_{n-r}$$

Let $W = \begin{bmatrix} W_1 & W_2 \end{bmatrix}$ such that,

$$W^T W = \begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix} \begin{bmatrix} W_1 & W_2 \end{bmatrix} = \begin{bmatrix} W_1^T W_1 & W_1^T W_2 \\ W_2^T W_1 & W_2^T W_2 \end{bmatrix} \tag{B.2}$$

From Eq. B.1, $W_2^T AA^T W_1 = W_2^T W_1 \Sigma_1^2$ but $A^T W_2 = 0$ so that $W_2^T A = 0$ and therefore $W_2^T W_1 \Sigma_1^2 = 0$. Since Σ_1^2 is nonsingular $W_2^T W_1 = W_1^T W_2 = 0$ and Eq. B.2 is

$$W^T W = \begin{bmatrix} I_r & 0 \\ 0 & I_{m-r} \end{bmatrix}$$

which shows that W is unitary. Let V_1 be the $n \times r$ matrix defined by $V_1 = A^T W_1 \Sigma_1^{-1}$ and let V_2 be the $m \times n - r$ matrix with columns forming an orthonormal basis for the null space of A. Then by definition,

$$AV_2 = V_2^T A^T = 0 \qquad \text{and} \qquad V_2^T V_2 = I_{n-r}$$

Letting $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ and noting

$$V_1^T V_1 = (\Sigma_1^{-1} W_1^T A)(A^T W_1 \Sigma_1^{-1}) = \Sigma_1^{-1}\Sigma_1^2\Sigma_1^{-1} = I_r$$

and

$$V_1^T V_2 = V_2^T V_1 = V_2^T A^T W_1 \Sigma_1^{-1} = (AV_2)^T W_1 \Sigma_1^{-1} = 0$$

shows V is also unitary. Finally

$$\begin{aligned} W^T A V &= \begin{bmatrix} W_1^T \\ W_2^T \end{bmatrix} A \begin{bmatrix} V_1 & V_2 \end{bmatrix} \\ &= \begin{bmatrix} W_1^T A V_1 & W_1^T A V_2 \\ W_2^T A V_1 & W_2^T A V_2 \end{bmatrix} \\ &= \begin{bmatrix} W_1^T A A^T W_1 \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_1^2\Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

$\square$

B.2 Proof of Lemma 5.5

Lemma 5.5 (Hautus Observability Lemma) *The n–dimensional discrete–time linear system (A, C) is observable if and only if*

$$rank \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n \quad for\ all \quad \lambda \in \mathbb{C} \tag{B.3}$$

Proof: This proof is analogous to Sontag's [109] proof of the Hautus lemma for controllability. Only eigenvalues of A need be considered since $\lambda I - A$ has full rank whenever λ is not an eigenvalue of A.

Prove the forward direction first. Assume that the rank would be less than n for some $\lambda \in \lambda(A)$ then there exists some nonzero vector $v \in \mathbb{C}^n$ and some λ such that

$$\begin{bmatrix} \lambda I - A \\ C \end{bmatrix} v = 0 \tag{B.4}$$

Thus, $Av = \lambda v$ and λ is an eigenvalue of A. Since $Cv = 0$ and $CA^k v = \lambda^k Cv = 0$ for all k then

$$\mathcal{O}(A, C)v = 0 \tag{B.5}$$

which contradicts observability.

The reverse direction is proved by assuming that the system is not observable so there exists the observability canonical form given in Lemma 5.3, with $r < n$. Let λ, p be an eigenvalue/eigenvector pair of the matrix G, so that

$$(\lambda I - G)p = 0 \tag{B.6}$$

It follows that the nonzero n–vector

$$w := \begin{bmatrix} 0 \\ p \end{bmatrix} \tag{B.7}$$

is an eigenvector of $\tilde{A}$, with $\tilde{C}w = 0$. Hence $v := Tw \neq 0$ satisfies

$$\begin{bmatrix} T^{-1}(\lambda I - A) \\ C \end{bmatrix} v = 0 \tag{B.8}$$

Since the row space of the matrix in Eq. B.8 coincides with the row space of the similar matrix

$$\begin{bmatrix} (\lambda I - A) \\ C \end{bmatrix} \tag{B.9}$$

then Eq. B.4 also holds, and hence the rank of Eq. B.9 cannot be n. □

B.3 Proof of Lemma 5.6

Lemma 5.6 (Periodic Observability Lemma) *The n–dimensional periodic discrete–time linear system (A, C_k) with period q is observable if and only if*

$$rank \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C} \tag{B.10}$$

Proof: This follows easily from Theorem 5.3 and Lemma 5.5. Theorem 5.3 proves that the periodic system is observable if and only if the lifted system is observable.

Lemma 5.5 can be used to construct the Hautus condition for the lifted system.

$$\text{rank} \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C} \tag{B.11}$$

which is equivalent to Eq. B.10, therefore the periodic system is observable if and only if Eq. B.10. □

B.4 Proof of Lemma 5.7

Lemma 5.7 (Hautus Lemma of Detectability) *The n–dimensional discrete–time linear system (A, C) is asymptotically observable (detectable) if and only if*

$$rank \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C}, \quad |\lambda| \geq 1 \tag{B.12}$$

Proof: Only eigenvalues of A need be considered since $\lambda I - A$ has full rank whenever λ is not an eigenvalue of A.

First assume that there is some eigenvalue, $\bar{\lambda} \in \lambda(A)$, $|\bar{\lambda}| \geq 1$ such that

$$\text{rank} \begin{bmatrix} \bar{\lambda} I - A \\ C \end{bmatrix} < n \tag{B.13}$$

using the the similar matrices $(\tilde{A}, \tilde{C})$ from Lemma 5.3 then the following also holds

$$\text{rank}\begin{bmatrix} \bar{\lambda} I - \tilde{A} \\ \tilde{C} \end{bmatrix} < n \tag{B.14}$$

Writing the transformed matrices in block form provides the following matrix

$$\begin{bmatrix} \bar{\lambda} I - E & 0 \\ -F & \bar{\lambda} I - G \\ H & 0 \end{bmatrix} \tag{B.15}$$

The first r columns are linearly independent from Lemma 5.5 (Hautus observability condition) applied to the observable system (E, H), thus the matrix $\bar{\lambda} I - G$ must lose rank and G has an unstable eigenvalue thus detectability is violated. The reverse direction is shown by assuming the system is not detectable and therefore G has an unstable eigenvalue. The unstable eigenvalue causes the matrix in Eq. B.15 to lose rank thus contradicting the rank condition in Eq. B.13 and completing the proof. □

B.5 Proof of Lemma 5.8

Lemma 5.8 (Periodic Detectability Lemma) *The n–dimensional periodic discrete–time linear system (A, C_k) with period q is asymptotically observable (detectable)* *if and only if*

$$rank \begin{bmatrix} \lambda I - A^q \\ C_0 \\ C_1 A \\ C_2 A^2 \\ \vdots \\ C_{q-1} A^{q-1} \end{bmatrix} = n \quad \forall \quad \lambda \in \mathbb{C}, \quad |\lambda| \geq 1 \tag{B.16}$$

Proof: This result is noted throughout the literature. For example, see Bolzern and Colaneri [14] and Bittanti and Bolzern [12].

□

B.6 Output Disturbance Models and Theorem 5.5

The condition given by Eq. 5.42 requires $\mathcal{G}_p$ to have full column rank and have column dimension less than the row dimension of $\mathcal{C}$. In addition, the range of $\mathcal{G}_p$ cannot contain the output space spanned by any integrating modes of $\mathcal{A}$ (or A).

The last condition requires further explanation. Let v be an integrating mode of A and consider a nonzero $\mathcal{P}$ that satisfies $\mathcal{C}v + \mathcal{G}_p\mathcal{P} = 0$. Then, for $\lambda = 1$, there exists a nonzero vector that satisfies Eq. B.17.

$$\begin{bmatrix} I - \mathcal{A} & 0 \\ \mathcal{C} & \mathcal{G}_p \end{bmatrix} \begin{bmatrix} v \\ \mathcal{P} \end{bmatrix} = 0 \tag{B.17}$$

thus Eq. 5.42 cannot be full rank. To avoid this problem it can be specified that the range of $\mathcal{G}_p$ cannot contain the output space spanned by any integrating modes of A or $\mathcal{G}_p\mathcal{P} \neq \mathcal{C}v$. Decomposition of the lifted matrices, $\mathcal{C}$ and $\mathcal{G}_p$ and repeated use of the fact that $Av = \lambda v$ reduces the condition to $G_p p \neq C_k v$ for all $p \neq 0$ and $k \in [0, q-1]$.

B.7 Input Disturbance Models and Theorem 5.5

The condition given by Eq. 5.43 requires $\mathcal{B}$ to have full column rank and have column dimension less than the row dimension of $\mathcal{C}$. In fact $\mathcal{B}$ having full column rank implies B having full column rank (see the definition of $\mathcal{B}$ in Eq. 4.27).

In addition, the range of B cannot contain any unobservable modes of $(\mathcal{A}, \mathcal{C})$. The last condition requires further explanation. Eq. B.18 restates the rank condition

$$\begin{bmatrix} I - \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \tag{B.18}$$

partitioning of the lifted matrices $\mathcal{B}$ and $\mathcal{D}$ yields

$$\begin{bmatrix} I - \mathcal{A} & \mathcal{B}' & B \\ \mathcal{C} & \mathcal{D}' & 0 \end{bmatrix} \tag{B.19}$$

Now, with a zero block in the matrix an argument similar to that in Section B.6 can be used to show that the range of B cannot contain any observable modes of the lifted system.

Appendix C

LMPC Software

The linear model predictive control toolbox (`lmpc`) was first developed by Ken Muske to run the examples presented in the February, 1993 AIChE Journal article "Model Predictive Control with Linear Models." [86]. The code was originally written in MATLAB which is a high level interactive language for numerical computations. It provided a quick and easy way to run infinite horizon model predictive control simulations. Subsequent reworking of the code provided the current user interface. The code was also ported to Octave [43] from MATLAB and is now supported for both programs. A master set of code is used to create program specific version for Octave and MATLAB. As a consequence, few of the nifty, program specific features are used. The code has found extensive use in graduate courses and industrial short courses [5]. As would be expected, some of the ideas in MPC including infinite input parameterization, exact soft constraints, zones *etc.* are incorporated into private version of the code until adequate testing and fine tuning of user interfaces can be completed. This code is still evolving.

The Big Picture

The code is intended to take a linear plant, linear model and tuning parameters along with optional constraints, disturbances, initial conditions, and filter gains and calculate a closed loop simulation using an infinite horizon controller. As the code currently stands, the following problems are solved:

- The target is calculated by first attempting problem Eq. 6.30 then Eq. 6.33, and finally by rejecting all output constraints. This is a rather primitive way to handle the target calculation. Exact soft constraints may provide an elegant way to solve this problem. Zone variables may be included in the future as well.

- The controller currently solves the control problem presented in Eq. 6.15, although a version of the code that solves Eq. 6.18 works and will eventually be incorporated into the base code.

- Upon failure of the regulator to achieve a feasible solution, the output constraints are dropped at the current time. The new QP is calculated and another failure pushes the constraint window forward one more time step. There is an arbitrary point where the program quits.

- The program updates the states (model and disturbance) with user provided filter gains. `dlqe.m` can be used to calculate Kalman filter gains.

- A general disturbance model framework allows for disturbance models in both the output and the state. An appropriate choice of model matrices for the state disturbance model provides an input disturbance model.

The basic structure of the code is as follows: there are nine routines for the user to enter parameters, one main routine to run the simulation and seven hidden routines that are called by the main routine. The nine routines used to enter the data consist exclusively of error checking and default value assignments. A simulation provides some screen output denoting which QP's are being solved. The final output is in vector form and no plots are automatically generated.

The variable naming is consistent with the AIChE Journal article wherever possible. Nine example scripts are also available. These scripts reproduce the AIChE Journal examples. When was the last time you actually ran a simulation that you saw in a paper? There was some hope that others would modify the code to fit their needs. There has been some success with that in the graduate courses at the University of Wisconsin.

C.1 Necessary Routines

The user only calls up to ten of the routines. There are only three routines that have to be called in order to run a simulation. The three necessary routines are listed in Table C.1 and summarized below.

Necessary Routine	Brief Description
lmpc_set_plant.m	input the plant parameters
lmpc_tuning.m	input the tuning parameters
lmpc.m	runs a simulation

Table C.1: Necessary Routines

lmpc_set_plant.m This routine allows the user to enter the linear model of the plant and an optional linear plant. The plant defaults to the model.

```
% Usage: lmpc_set_plant (Am, Bm, Cm, Ap, Bp, Cp)
%        lmpc_set_plant (Am, Bm, Cm)
%
% Function to set the model of the plant
%       _          _
%       x_k+1 = Am*x_k + Bm*u_k
%                  _
%       y_k   = Cm*x_k
%
% and the plant. The plant defaults to the model.
%
%       x_k+1 = Ap*x_k + Bp*u_k
%       y_k   = Cp*x_k
%
%  Am       = state matrix for plant model
%  Bm       = input matrix for plant model
%  Cm       = output matrix for plant model
%  Ap,Bp,Cp = matrices for actual plant
```

lmpc_tuning.m This routine allows the user to enter the tuning parameters for the objective function. The parameter M was originally used to investigate constrained stabilizability. The code naturally increases j_1 (the first time at which output constraints are enforced) upon failure of the regulator. qs and rs allow for different tuning in the target QP and regulator QP.

```
% Usage:  lmpc_tuning (q, r, s, N, M, qs, rs)
%         lmpc_tuning (q, r, s, N, M)
%         lmpc_tuning (q, r, s, N)
%         lmpc_tuning (q, r, s)
%
% Function to set the tuning parameters for the lmpc routines.
%
%  q  = output weighting scalar/vector/matrix
%  r  = input weighting scalar/vector/matrix
%  s  = input velocity weighting scalar/vector/matrix
%  N  = input horizon length
%       defaults to 1
%  M  = [output constraint window length (j2-j1), start of output
%       constraint window (j1)]  defaults to [N 1]
%  qs = output target weighting scalar/vector/matrix
%       defaults to q
%  rs = input target weighting scalar/vector/matrix
%       defaults to r
```

lmpc.m This routine runs a simulation. Most of the arguments are self–explanatory although some require mention. `filter` specifies a first order reference trajectory and ranges between 0 and 1. y is only needed when lmpc is used in conjunction with Tom Badgwell's PS1 simulator. uh is the optimal control horizon computed at each simulation step.

```
% Usage:  [u,x,y,x_hat,v_hat,w_hat,uh] = lmpc (nsim, yt, filter, y)
%
% Function to simulate the plant using a model predictive
% controller.
%
%  nsim   = simulation length
%  yt     = output setpoint vector
%  filter = filter constant for reference trajectory
%  y      = current output
%
%  u     = calculated inputs
%  x     = plant states
%  y     = plant outputs
```

```
%  x_hat = estimated states
%  v_hat = estimated output disturbance state
%  w_hat = estimated disturbance state
%  uh    = control horizons
```

C.2 Optional Routines

The optional routines allow for a variety of interesting simulations. They are listed in Table C.2. lmpc_reset sets all of the global lmpc variables to empty matrices. lmpc_help provides a brief comment about each of the routines. The remaining optional routines are summarized below.

Optional Routine	Brief Description
lmpc_init_cond.m	input the initial conditions
lmpc_constraint.m	input the constraints
lmpc_dist.m	set up disturbances
lmpc_filter.m	input the filter gains
lmpc lmpc_feedfor.m	input a feedforward model
lmpc_reset.m	clear all lmpc variables
lmpc_help.m	display brief description of lmpc toolbox

Table C.2: Optional Routines

lmpc_init_cond The initial conditions for the model, plant, input and feedforward model are set with this routine. A common simulation method involves using different values for the initial states in the model and the plant.

```
% Usage: lmpc_init_cond (xo, xo_hat, uo, xfo)
%
% Function to set the initial states for the model and plant
% as well as the initial inputs.  [] designates a zero vector
```

```
% of the appropriate size.
%
%  xo     = initial states for the plant (defaults to zero)
%  xo_hat = initial states for the model (defaults to zero)
%  uo     = initial inputs (defaults to zero)
%  xfo    = initial states for the feedforward disturbance model
%           (defaults to zero)
```

lmpc_constraint The constraints include rate of change constraints on the inputs.

```
% Usage: lmpc_constraint (umin, umax, uroc, ymin, ymax)
%
% Function to set the constraints for the controller. Unwanted
% constraint variables should be specified using empty matrix
% place holders ([]).  Scalar arguments will be expanded into
% vectors of appropriate size.
%
%  umin = lower constraint on the input
%  umax = upper constraint on the input
%  uroc = rate of change constraint on the input
%  ymin = lower constraint on the output
%  ymax = upper constraint on the output
```

lmpc_dist This is by far the most confusing routine, but it allows for a large number of disturbance types. V's are associated with the output disturbance. W's are associated with the state disturbance model (defaults to the input disturbance model). v and w add white noise to the sensor and the state respectively. V and W are used to specify a variety of disturbances with a step disturbance being the most common. Care should be taken to construct the V and W matrices according to the instructions.

```
% Usage: lmpc_dist (v, w, Gvm, Gwm, Gvp, Gwp, V, W)
%
% Function to set the disturbance matrices for the plant
%
%       x_k+1 = Ap*x_k + Bp*u_k + Gwp*w_k
%       y_k   = Cp*x_k + Gvp*v_k  =======
```

```
%                        =======
% and the model.
%       -          -                  -
%       x_k+1 = Am*x_k + Bm*u_k + Gwm*w_k
%                  -          -    =======
%       y_k   = Cm*x_k + Gvm*v_k
%                        =======
% Gvp = output disturbance matrix for the plant
%       (defaults to identity matrix)
% Gwp = state disturbance matrix for the plant (defaults to Bp).
% Gvm = output disturbance matrix for the model
%       (defaults to identity matrix)
% Gwm = state disturbance matrix for the model (defaults to Bm).
% v   = standard deviation for output white noise
% w   = standard deviation for state white noise
%
% In addition, the matrices V and W can be used to specify more
% complex noise models.  If V is specified then the state vector
% v_k will be expanded to...
%
%       v_k = v_white_noise + v_other
%
%
% where v_other can be integrated white noise, a step disturbance, a
% square wave, or a user defined function. Each row of V corresponds
% to a noise state from the vector v_k.  If V does not have as many
% rows as v_k then v_other will default to zero for the extra states.
% Use the following chart to determine what values to put into the V
% matrix to specify v_other.
%
%
%  V, W = 'fun'; function to compute disturbance where the function
%         takes one argument (time) and returns one argument (the
%         disturbance),
%            or...
%  V, W = [type, value1, value2] each row represents one disturbance
%         state
%
%                  | TYPE | VALUE1(DEFAULT) | VALUE2(DEFAULT)
% -----------------|------|-----------------|----------------
% int. white noise |  1   |  std. dev.(1)   |     max.(INF)
% step disturbance |  2   |  magnitude(1)   |     start (0)
```

```
% square wave       |  3   |  magnitude(1)   |     period(2)
```

lmpc_filter The code does not automatically calculate filter gains. Use `dlqe.m` to calculate the gains. There are three gains, the first is the gain for the state (x), the second is for the output disturbance states (p), and the third is for the disturbance state (z).

```
% Usage: lmpc_filter (Lx, Lp, Lz)
%
% Function to set some Kalman filter parameters. Eventually
% will be expanded to calculate filter gains.  [] designates
% no filter.
%
%  Lx = state filter gain
%  Lp = output step disturbance filter gain
%  Lz = state step disturbance filter gain
```

lmpc_feedfor This is not a very frequently used routine. See aiche9.m for a nice example showing set point tracking in the face of a feed forward disturbance.

```
% Usage: lmpc_feedfor (Af, Bf, Cf, d)
%
% Function to set the feedforward model in order to include
% measured disturbances.  d is a constant disturbance that
% effects the output according to the following model.
%
%      xf_k+1  = Af*xf_k + Bf*d
%      yf_k    = Cf*xf_k
%
%  Af = state matrix for feedforward model
%  Bf = input matrix for feedforward model
%  Cf = output matrix for feedforward model
%   d = constant disturbance vector
```

C.3 Hidden Routines

The hidden routines made the code easier to develop and maintain. The files are listed in Table C.3 and a short description follows.

Hidden Routine	Brief Description
hmpc_check_tuning.m	checks dimensions of tuning parameters
hmpc_get_dist.m	calculates a disturbance based on the user defined inputs from lmpc_dist
hmpc_target.m	constructs the target QP and calls hmpc_qp
hmpc_mat.m	constructs the constraint matrix for the QP equality and inequality constraint
hmpc_vec.m	constructs the constraint vector for the QP equality and inequality constraint
hmpc_setup_qp.m	builds the Hessian
hmpc_qp.m	solves a generic QP

Table C.3: Hidden Routines

hmpc_check_tuning Consistency checking of the tuning parameters cannot be done in lmpc_tuning because not enough information has been entered by the user. By the time lmpc calls this program everything has been entered.

hmpc_get_dist This routine deciphers the V and W variables to decide what disturbances the user wanted and where they should enter the system. This is called whenever the linear system is advanced in time.

hmpc_target Calculates the steady–state target. The target is calculated by first attempting problem Eq. 6.30 then Eq. 6.33, and finally by rejecting all output constraints.

hmpc_mat Constructs the constraint matrix for the regulator QP. There are equality constraints associated with zeroing the unstable modes at the end of the horizon. There is a variable carried along with this matrix that tells the QP how many rows are associated with equality constraints. The remaining constraints involve the inputs and the outputs.

hmpc_vec Constructs the constraint vector for the regulator QP. There are equality constraints associated with zeroing the unstable modes at the end of the horizon. There is a variable carried along with this matrix that tells the QP how many rows are associated with equality constraints. The remaining constraints involve the inputs and the outputs.

hmpc_setup_qp Constructs the Hessian for the regulator QP. This follows the AIChE Journal paper, but that doesn't make it any easier to decipher.

hmpc_qp This merely calls the QP solver. This is kept separate because the MATLAB and Octave QP solvers are so different, thus there are two different files. Both the target calculation and the regulator calculation use this file.

C.4 Example Scripts

The files aiche*.m, where * is one of 2, 3, 5, 6, 7a, 7b, 7c, 8, 9, reproduce the AIChE Journal article examples. The number corresponds to the specific example number. These files provide the user with a working example of the code.

C.5 Known Causes of Trouble

The eigenvalue decomposition is used to determine the unstable subspace for the equality constraint. As a consequence, the code cannot handle processes with repeat eigenvalues. Use of the Schur decomposition to handle the equality constrain would solve this problem, but MATLAB does not have an ordered Schur decomposition. It is assumed that the infinite input parameterization solves this problem by removing the need for equality constraints.

Bibliography

[1] N. Amit. *Optimal Control of Multirate Digital Control Systems.* PhD thesis, Stanford University, 1980.

[2] K. J. Åström. Computer control of a paper machine – an application of linear stochastic control theory. *IBM J. Res. Dev.*, 11:389–405, 1967.

[3] Karl J. Åström. *Introduction to Stochastic Control Theory.* Academic Press, San Diego, California, 1970.

[4] James W. Atkins, Thomas E. Rodencal, and Dennis E. Vickery. Correcting a poor moisture profile improves productivity. *Tappi*, 65(2):49–52, February 1982.

[5] Thomas A. Badgwell and James B. Rawlings. Model predictive control technology, theory and applications. Course notes, Rice University, Houston, 1996.

[6] Yonathan Bard. *Nonlinear Parameter Estimation.* Academic Press, New York, 1974.

[7] A. E. Beecher and R. A. Bareiss. Theory and practice of automatic control of basis weight profiles. *Tappi*, 53(5):847–852, May 1970.

[8] Luis G. Bergh and John F. MacGregor. Spatial control of sheet and film forming processes. *The Canadian Journal of Chemical Engineering*, 65:148–155, February 1987.

[9] W. L. Bialkowski. Application of steady state kalman filters – theory with field results. In *Proc. of the Joint Auto Control Conf.*, pages 361–375, 1978.

[10] W. L. Bialkowski. Integration of paper machine operations. *Tappi Journal*, pages 65–71, September 1988.

[11] Robert R. Bitmead, Michel Gevers, and Vincent Wertz. *Adaptive Optimal Control, The Thinking Man's GPC*. Prentice-Hall, Englewood Cliffs, New Jersey, 1990.

[12] Sergio Bittanti and Paolo Bolzern. Stabilizability and detectability of linear periodic systems. *Systems & Control Letters*, 6:141–145, 1985.

[13] Sergio Bittanti and Paolo Bolzern. On the structure theory of discrete–time linear systems. *International Journal of Systems Science*, 33:33–47, 1986.

[14] Paolo Bolzern and Patrizio Colaneri. Inertia theorems for the periodic lyapunov difference equation and periodic riccati difference equation. *Linear Algebra and its Applications*, 85:249–265, 1987.

[15] Thomas J. Boyle. Control of cross-direction variations in web forming machines. *The Canadian Journal of Chemical Engineering*, 55:457–461, August 1977.

[16] Thomas J. Boyle. Practical algorithms for cross–direction control. *Tappi Journal*, 61:77–80, January 1978.

[17] Thomas J. Boyle and Horace Hamby III. Cross–direction weight control: Pro and con. *Tappi Journal*, 68(6):56–59, June 1985.

[18] Richard D. Braatz, Babatunde A. Ogunnaike, and Andrew P. Featherstone. Control of sheet and film processes – a critical review. Submitted to Automatica, February 1995.

[19] Richard D. Braatz, Matthew L. Tyler, Manfred Morari, Ferdinand R. Pranckh, and Luigi Sartor. Identification and cross-directional control of coating processes. *AIChE Journal*, 38(9):1329–1339, 1992.

[20] Donald B. Brewster. Profile control by distributed control systems: Dream or reality? *Tappi Journal*, pages 75–81, February 1989.

[21] Donald B. Brewster and Andrew K. Bjerring. Computer control in pulp and paper 1961–1969. *Proceedings of the IEEE*, 58(1):49–69, January 1970.

[22] A. E. Bryson and Y. Ho. *Applied Optimal Control*. Hemisphere Publishing, New York, 1975.

[23] J. W. Burns. The use of analysis of variance to characterize paper web nonuniformities. *Tappi Journal*, 57(12):143–146, 1974.

[24] John C. Campbell and James B. Rawlings. Estimation and control of sheet and film forming processes. In Irena Lasiecka and Blaise Morton, editors, *Control Problems in Industry*, pages 43–63, Boston, 1995. SIAM, Birkhäuser.

[25] John C. Campbell and James B. Rawlings. Gage control of film and sheet forming processes. Third SIAM Conference on Control and its Applications, April 1995.

[26] John C. Campbell and James B. Rawlings. Evaluation of model predictive control performance in sheet and film forming processes. AIChE National Meeting, Chicago, IL, November 1996.

[27] John C. Campbell and James B. Rawlings. Identification and control of sheet forming processes. In *Proceedings of the 13th IFAC World Congress*, San Francisco, California, July 1996.

[28] Torsten Cegrell and Torbjörn Hedqvist. Successful adaptive control of paper machines. *Automatica*, 11:53–59, 1975.

[29] F. Chambon. Validation of a model for the cast–film process. AIChE Annual Meeting, Chicago IL, November 1996.

[30] S. C. Chen. Kalman filtering applied to sheet measurement. In *Proceedings of the American Control Conference*, Atlanta, GA, June 1988.

[31] S. C. Chen and R. G. Wilhelm, Jr. Optimal control of cross-machine direction web profile with constraints on the control effort. In *Proceedings of the American Control Conference*, Seattle, WA, June 1986.

[32] D. Chmielewski and V. Manousiouthakis. On constrained infinite-time linear quadratic optimal control. Report ENG 95–150, UCLA School of Engineering and Applied Science (also submitted to *Systems and Control Letters*), October 1995.

[33] Edward D. Cohen. Coatings: Going below the surface. *Chemical Engineering Progress*, pages 19–23, September 1990.

[34] C. R. Cutler and B. L. Ramaker. Dynamic matrix control—a computer control algorithm. AIChE National Meeting, Houston, TX, April 1979.

[35] E. B. Dahlin. Computational methods in a dedicated computer system for measurement and control on paper machines. *Tappi*, 53(6):1100–1105, June 1970.

[36] Prashant Dave, Dennis A. Willig, Gautham K. Kudva, Joseph F Pekny, and Francis J. Doyle. LP methods in MPC of large–scale systems: Application to paper–machine CD control. *AIChE Journal*, 43(4):1016–1031, 1997.

[37] Philip J. Davis. *Circulant Matrices*. John Wiley & Sons, New York, 1979.

[38] F. M. D'Hulster, R. M. C. De Keyser, and A. R. Van Cauwenberghe. Simulations of adaptive controllers for a paper machine headbox. *Automatica*, 19(4):407–414, 1983.

[39] Guy A. Dumont. Application of advanced control methods in the pulp and paper industry – A survey. *Automatica*, 22:143–153, 1986.

[40] Guy A. Dumont, Ivar Mar Jonsson, Michael S. Davies, Fariborz T. Ordubadi, Ye Fu, K. Natarajan, Claes Lindeborg, and E. Michael Heaven. Estimation of moisture variations on paper machines. *IEEE Transactions on Control Systems Technology*, 1(2):101–113, June 1993.

[41] S.R. Duncan. Observers and controllers for cross–directional control of web processes. UMIST Control Systems Centre Report No. 806, 1994.

[42] Stephen R. Duncan. The design of robust cross–directional control systems for paper making. In *Proceedings of the American Control Conferance*, pages 1800–1805, Seattle, Washington, June 1995.

[43] John W. Eaton. Octave—a high-level interactive language for numerical computations. Technical Report CCSR-93-003, The University of Texas Center for Control and Systems Research, Austin, Texas 78712, 1993.

[44] Andrew P. Featherstone and Richard D. Braatz. Control relevant identification of sheet and film processes. In *Proceedings of the 1995 American Control Conference*, pages 2692–2696, Seattle, Washington, June 1995.

[45] V. V. Fedorov. *Theory of Optimal Experiments*. Academic Press, New York and London, 1972.

[46] Rolf Findeisen. Suboptimal nonlinear model predictive control. Master's thesis, The University of Wisconsin–Madison, 1997.

[47] Magne Fjeld. Application of modern control concepts on a kraft paper machine. *Automatica*, 14:107–117, 1978.

[48] Roger Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, 1987.

[49] Carlos E. García and A. M. Morshedi. Quadratic programming solution of dynamic matrix control (QDMC). *Chemical Engineering Communications*, 46:73–87, 1986.

[50] Elmer G. Gilbert and Kok Tin Tan. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, 36(9):1008–1020, September 1991.

[51] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright. User's guide for SOL/QPSOL: A Fortran package for quadratic programming, technical report SOL 83-12. Technical report, Systems Optimization Laboratory, Department of Operations Research, Stanford University, 1983.

[52] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins, Baltimore, 1989.

[53] O.M. Grasselli. A canonical decomposition of linear periodic discrete–time systems. *International Journal of Control*, 40:201–214, 1984.

[54] O.M. Grasselli and F. Lampariello. Stabilizability and detectability of linear periodic systems. *International Journal of Control*, 33:1091–1106, 1981.

[55] Ping Guo, Xuan Z. Ni, and Jin Zheng. Polymer extrusion production control using active recognition and adaptive control system. In *Second IEEE International Conference on Fuzzy Systems*, pages 779–784, San Francisco, California, March 1993.

[56] J. Michael Hadra. Selecting actuators for cross–machine control. *Tappi Journal*, 68(6):60–62, June 1985.

[57] Alena Halousková, Miroslav Kárný, and Ivan Nagy. Adaptive cross–direction control of paper basis weight. *Automatica*, 29:425–429, 1993.

[58] William Haverinen. Results report with a new caliper sensor and cross directional control on canadian newsprint. In *Proc. of the 3rd Int. Pulp and Paper Process Control Symp.*, pages 111–120, Vancouver B.C., Canada, 1983.

[59] Xiaohua George He. Adaptive decoupling control and application in sheet production. In *Proceedings of the American Control Conferance*, pages 2455–2456, San Francisco, California, June 1993.

[60] W. P. Heath. Orthogonal functions for cross–directional control of web forming processes. *Automatica*, 32:183–198, 1996.

[61] E. M. Heaven, T. M. Kean, I. M. Jonsson, M. A. Manness, K. M. Vu, and R. N. Vyse. Applications of system identification to paper machine model development and controller design. In *Second IEEE Conference on Control Applications*, pages 227–233, Vancouver, B.C., September 1993.

[62] Robert V. Hogg and Elliot A. Tanis. *Probability and Statistical Inference.* Macmillan Publishing Company, New York, 1988.

[63] Roger A. Horn and Charles R. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[64] A. J. Isaksson, M. Hagberg, and L.E. Jönsson. Benchmarking for paper machine MD–control: Simulation results. *Control Eng. Practice*, 3(10):1491–1498, 1995.

[65] Andrew H. Jazwinski. *Stochastic Proceses and Filtering Theory.* Academic Press, New York, 1970.

[66] Richard H. Jones. Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22(3):389–395, August 1980.

[67] Thomas Kailath. *Linear Systems.* Prentice-Hall, Englewood Cliffs, New Jersey, 1980.

[68] T. Kajiwara, S. Ninomiya, Y. Kuwano, and K. Funatsu. Numerical simulation of converging flow of polymer melts through a tapered slit die. *Journal of Non–Newtonian Fluid Mechanics*, 48(1–2):111–124, July 1993.

[69] R. E. Kalman. Contributions to the theory of optimal control. *Bull. Soc. Math. Mex.*, 5:102–119, 1960.

[70] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Engineering*, pages 35–45, March 1960.

[71] C. Eric Kan. A new caliper control actuator – evaluation process and practical results. *Tappi Journal*, pages 81–85, July 1987.

[72] Håkan Karlsson. Optimal cross–direction basis weight and moisture profile control on paper machines. In *Proc. of the 3rd Int. Pulp and Paper Process Control Symp.*, pages 139–145, Vancouver B.C., Canada, 1983.

[73] M. A. Keyes and A. Kaya. Control of paper basis weight by expert SQC/SPC algorithm using distributed controls. In *11th IFAC World Congress, Tallinn, Estonia*, pages 395–400, 1990.

[74] Dudley E. King Jr. Computer control of a combination paperboard machine. *Tappi*, 59(8):52–55, August 1976.

[75] D. L. Kleinman. An easy way to stabilize a linear constant system. *IEEE Transactions on Automatic Control*, 15(12):692, December 1970.

[76] Kristinn Kristinsson and Guy A. Dumont. Paper machine cross directional basis weight control using gram polynomials. In *Second IEEE Conference on Control Applications*, pages 234–240, Vancouver, B.C., September 1993.

[77] Huibert Kwakernaak and Raphael Sivan. *Linear Optimal Control Systems*. John Wiley and Sons, New York, 1972.

[78] W. H. Kwon and A. E. Pearson. On feedback stabilization of time-varying discrete linear systems. *IEEE Transactions on Automatic Control*, 23(3):479–481, June 1978.

[79] Daniel L. Laughlin, Manfred Morari, and Richard D. Braatz. Robust performance of cross-directional basis-weight control in paper machines. *Automatica*, 29(6):1395–1410, 1993.

[80] Joseph G. Lebrasseur. Technical advances to improve linerboard moisture. *Tappi Journal*, 67(10):60–61, October 1984.

[81] Yue Liang, Kazurou Kurihara, Kunio Saito, Hiroyasu Murakami, Kunio Kumagai, Masahiro Shshima, and Masataka Tanigaki. Profile control of plastic sheet in an industrial polymer processing process. AIChE Annual Meeting, Chicago IL, November 1996.

[82] Li-Dah Liu, Shu-Heng Wen, and Ta-Jo Liu. Extrusion die design for viscoplastic fluids. *Advances in Polymer Technology*, 13(4):283–295, 1994.

[83] H.B. Matthews. *Model Identification and Control of Batch Crystallization for an Industrial Chemical System*. PhD thesis, University of Wisconsin at Madison, 1997.

[84] Edward S. Meadows. *Stability and Continuity of Nonlinear Model Predictive Control*. PhD thesis, The University of Texas at Austin, 1994.

[85] Kenneth R. Muske. *Linear Model Predictive Control of Chemical Processes*. PhD thesis, The University of Texas at Austin, 1995.

[86] Kenneth R. Muske and James B. Rawlings. Model predictive control with linear models. *AIChE Journal*, 39(2):262–287, 1993.

[87] I. Nagy and F. Dušek. Theory and implementation of paper cross profile adaptive control. In *Advanced Methods in Adaptive Control for Industrial Applications*, pages 206–211, Prague, Czech, May 1991.

[88] Seyhan Nuyan. Requirements for the cross–machine direction basis weight control of the paper web. In *Proceedings of the 1986 American Control Conference*, pages 1416–1422, Seattle, Washington, June 1986.

[89] R. J. Pfeifer and R. G. Wilhelm, Jr. Comparing on-line gauging statistics from different paper machines. In *Proceeding of the Process and Product Quality Conference*, 1992.

[90] D. M. Prett and R. D. Gillette. Optimization and constrained multivariable control of a catalytic cracking unit. In *Proceedings of the Joint Automatic Control Conference*, 1980.

[91] S. Joe Qin and Thomas A. Badgwell. An overview of industrial model predictive control technology. Fifth International Conference on Chemical Process Control, Lake Tahoe, California, January 1996.

[92] C. V. Rao, 1997. Personal Communication.

[93] Christopher V. Rao, John C. Campbell, James B. Rawlings, and Stephen J. Wright. Efficient implementation of model predictive control for sheet and film forming processes. In *Proceedings of American Control Conference, Albuquerque, NM*, 1997.

[94] Christopher V. Rao, Stephen J. Wright, and James B. Rawlings. On the application of interior point methods to model predictive control. *Preprint MCS-P664-0597*, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., May 1997.

[95] Laxmi K. Rastogi. A flat die extruder control system using distributed microcomputers. In *Proc. of the Joint Auto Control Conf.*, pages 375–389, 1978.

[96] James B. Rawlings and I-Lung Chien. Gage control of film and sheet-forming processes. *AIChE Journal*, 42(3):753–766, March 1996.

[97] J. Richalet, A. Rault, J. L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14:413–428, 1978.

[98] Apostolos Rigopoulos, Yaman Arkun, Ferhan Kayihan, and Eric Hanczyc. Indentification of paper machine full profile disturbance modes using adaptive principal component analysis. In *Proceedings of Chemical Process Control – V*, Tahoe City, California, January 1996.

[99] Reinhold C. Roth. Electromagnetic radiation for CD moisture control. *Tappi Journal*, pages 57–61, January 1986.

[100] Michael F. Saucier and Alan D. Steinkirchner. Analysis of paper machine performance. *Tappi Journal*, 71(4):97–102, April 1988.

[101] Pierre O.M. Scokaert and James B. Rawlings. Stability of model predictive control under perturbations. In *IFAC Symposium on Nonlinear Control System Design, Tahoe City, California*, pages 21–26, 1995.

[102] Pierre O.M. Scokaert and James B. Rawlings. Infinite constraint horizons in MPC. AIChE Annual Meeting, Chicago IL, November 1996.

[103] Pierre O.M Scokaert and James B. Rawlings. Infinite horizon linear quadratic control with constraints. In *Proceedings of the 13th IFAC World Congress*, San Francisco, California, July 1996.

[104] Pierre O.M. Scokaert and James B. Rawlings. Constrained linear quadratic regulation. Accepted for publication in *IEEE Transactions on Automatic Control*, March 1997.

[105] L. E. Scriven and W. J. Suszynski. Take a closer look at coating problems. *Chemical Engineering Progress*, pages 24–29, September 1990.

[106] F. G. Shinskey. *Feedback Controllers for the Process Industries.* McGraw-Hill, New York, 1994.

[107] Steven J. Siler. Cross–machine basis–weight control: Machine considerations. *Tappi Journal*, 67(12):52–55, December 1984.

[108] Douglas E. Smith, Daniel A. Tortorellil, and Charles L. Tucker II. Sensitivity analysis and optimization of polymer sheet extrusion dies. In *Proceedings of the 1994 International Mechanical Engineering Congress and Exposition*, volume 194, pages 307–318, 1994.

[109] Eduardo D. Sontag. *Mathematical Control Theory.* Springer-Verlag, New York, 1990.

[110] H. W. Sorenson. Least–squares estimation: from Gauss to Kalman. *IEEE Spectrum*, pages 63–68, July 1970.

[111] Kathryn A. Soucy and Bradley R. Holt. Modeling, estimation and control of polymer composite processing. In *Proceedings of the American Control Conferance*, pages 1650–1655, San Diego, California, May 1990.

[112] Mario Sznaier and Mark J. Damborg. Suboptimal control of linear systems with state and control inequality constraints. In *Proceedings of the 26th Conference on Decision and Control*, pages 761–762, 1987.

[113] M. L. Tyler and M. Morari. Estimation of cross directional properties: Scanning versus stationary sensors. *AIChE Journal*, 41(4):846–854, April 1995.

[114] Barclay W. Wallace. Economic benefits offered by computerized profile control of weight, moisture, and caliper. *Tappi*, 64(7):79–83, July 1981.

[115] Michael D. Wallace. Cross–direction control: Profitability and pitfals. *Tappi Journal*, pages 72–75, June 1986.

[116] Xiaochun George Wang, Guy A. Dumont, and Michael S. Davies. Adaptive basis weight control in paper machines. In *Second IEEE Conference on Control Applications*, pages 209–216, Vancouver, B.C., September 1993.

[117] Xiaochun George Wang, Guy A. Dumont, and Michael S. Davies. Estimation in paper machine control. *IEEE Control Systems*, pages 34–43, August 1993.

[118] R. G. Wilhelm, Jr. and M. Fjeld. Control algorithms for cross directional control: the state of the art. In *Proceedings of the 5th IFAC/IMEKO Conference on Instrumentation in the Paper, Rubber, Plastic and Polymerization Industries*, Antwerp, October 11-13 1983.

[119] A. J. Wilkinson and A. Hering. A new control technique for cross machine control of basis weight. In *Proceedings of the 5th IFAC/IMEKO Conference on Instrumentation in the Paper, Rubber, Plastic and Polymerization Industries*, Antwerp, October 11-13 1983.

[120] Stephen J. Wright. Applying new optimization algorithms to model predictive control. Fifth International Conference on Chemical Process Control, Lake Tahoe, California, January 1996.

[121] Stephen J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.

[122] Steven J. Wright. Interior point methods for optimal control of discrete-time systems. *Journal of Optimization Theory and Applications*, 77:161–187, 1993.

[123] Q. Xia, M. Rao, and J. Qian. Model algorighmic control for paper machines. In *Second IEEE Conference on Control Applications*, pages 203–208, Vancouver, B.C., September 1993.

[124] Jianfeng Zhang. Model predictive control of sheet and film forming processes. Master's thesis, The University of Wisconsin–Madison, 1997.

Vita

John Carl Campbell, son of Martha Patty and Alexander James Campbell, was born in Chicago Heights, Illinois on May 4th, 1969. Just before memories could be made his family moved to Ohio. At age 10, another family move placed John in High Bridge, New Jersey. In June 1987 he graduated from nearby Voorhees High School, and in the fall of that year he enrolled in the University of Illinois at Champaign–Urbana. He received his Bachelor of Science degree in Chemical Engineering in May 1991 from the University of Illinois. In the fall of 1991 he joined the research group of Dr. James B. Rawlings at the University of Texas at Austin. He received his Masters of Science degree in Chemical Engineering from UT in December 1994. In August 1995, he moved with Dr. Rawlings from Austin to the Chemical Engineering Department at the University of Wisconsin–Madison.

Permanent Address: 12840 S. Kirkwood Dr.#522

Stafford, TX 77477

This dissertation was prepared with LaTeX 2ε[1] by the author.

[1]This particular University of Wisconsin complient style was carved from The University of Texas at Austin styles as written by Dinesh Das (LaTeX 2ε), Khe–Sing The (LaTeX), and John Eaton (LaTeX). Knives and chisels wielded by John Campbell and Rock Matthews.

Index