PERFORMANCE MONITORING AND FAULT DETECTION IN CONTROL SYSTEMS

Thesis by

Matthew L. Tyler

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1997

(Submitted November 18, 1996)

Acknowledgements

I am grateful to my thesis advisor, Prof. Manfred Morari, for having enough confidence in my abilities to grant me significant intellectual freedom during the course of this research. In addition, his insights and high standards have definitely helped to shape this work. Finally, I wish to thank him for allowing me the opportunity to present much of this research at several international conferences.

I am indebted to the many excellent teachers who instructed me, both at Caltech and MIT, and who taught me to question everything. To the entire staff and assistants at the Institut für Automatik at the Swiss Federal Institute of Technology, "merci viel mals" for contributing to a successful stay.

A highlight of my studies has been the association with the many outstanding colleagues in the Morari group: Richard, Tyler, Nikos, Alex, Simone, Mayuresh, Iftikhar, Carl, Vesna, Thomas, and Cornelius. A special thanks is due to Carl, Mayuresh, and Iftikhar for their roles as lunchmates, hiking partners, confidants, travel companions, and babysitters, and in general making Switzerland a better place to live.

For generous financial support throughout my graduate studies, I wish to acknowledge the Fannie and John Hertz Foundation. In addition, unrivaled resources and other support of the Institut für Automatik during my two year sojourn in Switzerland was greatly appreciated.

I am thankful for Jefferson, who was always willing to "play trains" or "read stories" when I needed some distraction, and who helped to remind me about the important things in life.

Above all, I would like to thank my wife Judy for her devoted friendship, unceasing support, and self-sacrificing commitment, without which this work could never have been completed. I could never hope to adequately repay the debt I owe you or to convey my indescribable gratitude.

Abstract

As the sophistication of systems used in chemical processing industries increases and demands for high quality products manufactured at low costs mount, the need for improved methods for automatic monitoring of processes arises. This is particularly true for systems operating under automatic control where the control system often acts to eliminate early warning signs of process changes. This thesis examines problems in the area of dynamic system monitoring, with emphasis on control systems. Problems in the areas of controller performance monitoring, estimation, and fault detection are considered.

In the area of controller performance monitoring, techniques for assessing performance in a minimum variance framework are developed. In contrast to previous methods, the current approach is applicable to general systems, including unstable and nonminimum-phase plants and systems with unstable controllers. Through two simple examples, it is shown that significant errors may be encountered when information on the unstable poles and non-invertible zeros of a system is not properly included in the performance evaluation techniques. An alternative approach to evaluating deterioration in performance of control systems is formulated using a framework in which acceptable performance is expressed as constraints on the closed loop transfer function impulse response coefficients. Likelihood methods are used to determine if the constraints are met. This second approach can be applied to more general performance criteria than the minimum variance based method.

The problem of constrained state estimation is pursued using Moving Horizon Estimation. It is shown that previous formulations of this estimation technique can be unstable when constraints on the innovations and estimated states are included. By expanding the constraint set and modifying the estimation objective, stability is guaranteed. The proposed algorithm can be implemented as a quadratic program.

Several approaches to fault detection are considered. First, the simultaneous

design of linear fault detection filters and controllers is considered using the four parameter controller framework. It is shown how this framework may be considered as a special case of a more general interconnection framework for which a deep synthesis theory exists. Second, using the Moving Horizon Estimation framework, a model based fault detection scheme capable of directly incorporating a class of bounded model uncertainty is developed. The proposed method is compared to other methods employing an adaptive threshold, and is demonstrated on a simulation example of a cold tandem steel mill. Finally, a statistical framework for general change detection problems is presented. This method uses a two-model approach, where signals and parameters subject to change are modeled by Brownian motion for the faulty case and by constant values in the nominal case. A detection algorithm using likelihood ratio testing is implemented through the use of recursive dynamic filtering.

The use of qualitative modeling in detection and control problems is formulated using propositional logic. By representing literals using integer variables, qualitative features can be incorporated into control and detection problems. Symptom aided detection and multiobjective performance prioritization are among the problems which can be solved in this framework using mixed integer linear and quadratic programming.

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1 Introduction

An increase in the level of automation in the chemical process industry over the past few decades has created a greater need for performing monitoring and diagnostic tasks. Under manual operation, human operators must decide to take action when unexpected problems arise, and by so doing they gather early warning signs of serious problems. By contrast, in an automated environment, control algorithms compensate for changing conditions in a manner which may not provide early warnings to the reduced staff of human operators, and greater problems might develop before being noticed. In addition, the complexity of modern control systems is increasing. Demands for higher quality products, cost efficient production, and improved reliability and safety all act to increase performance requirements. As higher requirements are placed on control system, more efficient methods are needed to monitor these automated processes.

The term monitoring can be used to describe a wide range of activities associated with automated or computer assisted process supervision. Among these activities are included performance monitoring, state and parameter estimation, and fault detection. Even though each of these pursuits has distinct characteristics, there exist common properties binding them together. In each case, process data are collected. The monitoring method processes the data with the goal of inferring additional information about the system's operation and its transient behavior. Finally, the monitoring operation is seldom the end goal but instead a means to achieving a larger objective. In the case of controller performance, this objective is to achieve the desired performance level. When the monitoring scheme indicates that system performance is not met, engineering modifications must be undertaken to achieve this goal. In state estimation, the larger goal may be controlling the system using an algorithm which requires the process state as an input. Alternatively, state estimation may be an important component of a fault detection scheme. In fault detection, the final

objective is to be able to compensate for faulty behavior. In this work, only the monitoring problems are considered and not the larger schemes of which they are vital components.

In Sections 1.1-1.3, a brief description of each of the monitoring problems which are mentioned above and which are investigated in more detail in the following chapters is given. After this introduction, the organization of this thesis will be laid out in Section 1.4.

1.1 Performance monitoring

As market place competition leads to consumer demand for higher quality products, monitoring of process performance is a key component of modern manufacturing. Traditional methods make use of statistical process control technology to monitor final and intermediate product quality [93]. In this approach, a set of product properties is measured and recorded. Bounds on acceptable ranges for these properties are determined, and statistical tests to determine if the properties violate the bounds are applied. The Shewart control chart is among the best known methods of this type.

Recent advances in this area have focused on the development of multivariate statistical control charts [81, 51]. For processes with a large number of measurements, Principle Component Analysis and Partial Least Squares are used to compress the information contained in the data trajectories into low-dimensional spaces that describe past operation. By using latent variables, univariate control charts can be constructed. Other methods from chemometrics have also been successfully applied to generate control charts [112, 64].

As with other monitoring problems, the ultimate goal of performance monitoring is not to determine if specifications have been met but rather to determine what must be done to assure that specifications will be satisfied. With this in mind, for systems operating under automated control, assessing the performance of the control loops is of vital importance. When the control system is not performing as designed, producing a quality product may be very difficult. Experience has shown that in large

production facilities, poorly operating control loops are frequently a significant cause of low quality product [52]. In modern chemical manufacturing facilities, thousands of automated control loops are employed. A single control engineer may be responsible for the maintenance of literally hundreds of controllers. Without control performance assessment tools, supervision of so many loops can be unmanageable.

1.2 State and parameter estimation

State and parameter estimation problems have been studied extensively ever since the seminal work on linear regress by Gauss [38]. Although the theory of state estimation for linear systems has been extensively developed (see for example [1]), the implementation of such estimation procedures in more general detection problems remains an active area of research. Estimation of states and parameters forms a foundation for many other monitoring schemes. For example, residual errors from Kalman filters based state estimators have been employed in detection schemes for over two decades. Recursive parameter estimation methods have been used to test for unwanted changes in the properties of system components [48]. Recently, Benveniste et al. [12] and Zhang et al. [113] have developed methods to design a change detection algorithm from any recursive parameter estimation scheme.

An important problem in estimation theory which has recently received attention involves the estimation of system states when it is known that they must satisfy certain constraints. This situation often arises in practice due to engineering knowledge on estimated variables which may be formulated as bounding constraints. Due to measurement noise and other considerations, filtered estimates obtained using traditional methods will not necessarily satisfy these conditions. By including explicitly the constraints in the estimation scheme, it may be possible to obtain more accurate estimates and reduce the sensitivity to measurement errors. Moving Horizon Estimation [92] has been proposed as an estimation strategy for this situation. An important consideration which is resolved for the first time in this thesis addresses the stability of such methods.

1.3 Fault detection

The terms "fault" and "failure" refer to any type of malfunction of a system which affect the operating conditions, such as inaccurate sensor readings, failure of actuators to respond to control signals, unexpected changes in operating conditions, poisoning of catalysts, or leaks or clogs in pipelines. Faults can be divided into two classes, depending upon the way in which they are modeled. *Additive faults* are those which can be modeled as an input signal to the system. Examples of faults which are well modeled as additive faults include sensor bias and actuator failure. Many types of system failures cannot be represented by an additive disturbance. For example, changes in location of system poles and gains affect parameters which multiply the process states and therefore cannot be accurately represented by input signals. Such faults are often referred to as *multiplicative faults.*

Significant research effort has been devoted to detection of additive faults since the seminal work of work of Beard [11] and Jones [50]. Early approaches employed Kalman filters to generate residual signals which should be zero-mean white noise when the system is operating properly [110]. Subsequently, methods based on the notion of redundancy of measured variables were developed. Physical redundancy, referring to the situation in which multiple identical sensors are used to measure the same variable, is the simplest example of this concept. In this case, deviations between the measurements indicate faults, and a voting procedure can be put to use to isolate the fault. Physical redundancy is a hardware based solution to fault detection, and will not be investigated in this work.

In addition to physical redundancy, analytical redundancy can be employed in fault detection schemes. Analytical redundancy is a result of the inherent static and dynamic relations between system inputs and outputs. For example, material and energy balances may be used to analyze the consistency of measurements obtained in a piping network. Chow and Willsky [18] have presented a discussion of the use of analytical redundancy in fault detection schemes. The simplest case of analytical redundancy, sometimes referred to as spatial or direct redundancy, involves only

static relations of variables. When the measurements from a system are linearly dependent, it is possible to find linear combinations of the measurements which are nominally zero. Even when the measurements are linearly independent, dynamic relations among the variables produce temporal redundancies. Parity space methods [18] can be used to derive detection schemes based on these relationships. A parity relation is a linear relation between the temporal process variables which, according to the dynamic model, should be zero. The parity space contains all such relations.

A detection scheme based on analytical redundancy can be implemented on an existing process without the expense of refitting sensors and actuators concomitant with physical redundancy methods. However, analytical methods require the use of process models. As mathematical models only provide an approximate description for physical systems, the reliability of analytical methods is often unsatisfactory. Model uncertainty is perhaps the largest existing obstacle in implementing analytical redundancy based detection schemes. For example, when the model for the system does not describe the system exactly, the residuals generated from parity relations may vary from zero even when no disturbances or faults are present.

Several approaches have been proposed to deal with modeling error. Chow and Willsky [18] have proposed a parity space method for generating residuals in the face of uncertainty. In their approach, a parity vector which minimizes the worst case effect of model uncertainty on the residual is sought. This method suffers from the shortcoming that it does not consider the type of fault to be detected. Thus, although the residual may be insensitive to model uncertainty, it may not be sensitive to the faults one desires to detect. Lou et al. [62] have proposed another approach to robust fault detection. They assume that the set of possible models can be parameterized by a finite number of models, $q = 1, \ldots, Q$, and then find a parity vector which minimizes the sum of residuals for each of the models. The design of such a detector reduces to finding the orthogonal complement of a matrix. Frank and Wünnenberg [35] have proposed the use of *unknown input observers* (UIO) for generating residuals which are robust to model uncertainty. The UIO is an observer which is structured in such a way that it is insensitive to certain input directions. Frank and Wünnenberg

claim that the UIO can be used to generate residuals which are robust to model uncertainty. To accomplish this, they propose that uncertainty be represented by input disturbances, and the observer be designed to be insensitive to these directions. One limitation of this method is the fact that many types of modeling errors cannot be represented by additive signals. In addition, for each of the above mentioned methods, unmodeled model perturbations are restricted to lie in a certain subspace. No bounds are associated with these perturbations. Often, this is an unrealistic representation of actual model uncertainty. For example, if there exists a small error in an input gain which affects all outputs, such a direction will not exist. A better description would provide for bounded perturbations.

Basseville [8] and Isermann [47, 48] have provided reviews of the most significant methods for detection of multiplicative faults. The majority of these methods employ recursive or some other form of on-line parameter estimation to track parameter values which change when a fault occurs. Using statistical significance tests, one determines if changes in the value of the estimated parameters warrant raising an alarm.

The methods discussed above are based upon quantitative process models. An alternative approach to fault detection which has been applied when quantitative models are unavailable uses qualitative process models. Past approaches to detection based upon qualitative modeling have used confluence equations [23, 84] and sign directed graphs [45, 71, 42, 106].

1.4 Thesis organization

This thesis is organized as follows. In Part II, the problem of controller performance monitoring is considered. Chapter 2 is concerned with performance monitoring based on a minimum variance control objective. Building on previous work by Harris [43] and Stanfelj et al. [97], tests for determining if a controller achieves this objective for general plants are developed. In Chapter 3, a new performance monitoring method which is based upon statistical hypothesis testing and is capable of dealing with a wide range of industrially important control objectives is presented.

Part III addresses problems in fault detection. Chapter 4 deals with the design of linear fault detection filters. In particular, the problem of designing an integrated control and diagnostic module is considered using the four degree of freedom controller framework introduced by Nett [78]. It is shown that this framework can be recast as a special case of a more general control theory, in which case existing results can be used to synthesize detection filters in conjunction with controllers. In Chapter 5, the problem of constrained estimation is addressed using the framework of Moving Horizon Estimation. Modifications to previous algorithms are developed which enable the designer to guarantee stability for the estimation scheme in the presence of constraints. In Chapter 6, fault detection using uncertain models is addressed using Moving Horizon Estimation. For a wide class of uncertainty descriptions, the fault estimation scheme can explicitly account for model mismatch by using constrained estimation. The approach developed in this chapter is applied to a model of an industrial cold tandem steel mill.

The methods of Chapters 4 through 6 are applicable to additive faults, i.e., those faults which can be represented as input signals. In Chapter 7, a more general fault detection method is derived which can address both additive and multiplicative failures. This approach employs statistical change detection methods along with nonlinear Bayesian filtering.

Chapter 8 addresses qualitative process modeling and performance objective formulation. In particular, qualitative features are represented using propositional logic. By using integer variables, propositions can be expressed as linear constraints and the truth values of literals can be related to process variables. This approach cannot only be applied to detection problems, but also to control problems. The formulation of constraints from logic propositions is recapitulated. Symptom based fault detection combining both a qualitative and quantitative model can be posed as a mixed integer quadratic program using this framework. In the area of control, it is shown how multiobjective performance criteria can be met using mixed integer programming methods.

Part II

Controller Performance Monitoring

Chapter 2 Minimum Variance Performance Monitoring for Unstable and Nonminimum-phase Systems

Summary

This chapter presents minimum variance control evaluation techniques and derives extensions of previous results to general systems, including unstable and nonminimum-phase plants and systems with unstable controllers. An analysis of the sensitivity of these methods to uncertainties in the location of unstable poles and non-invertible zeros is presented. An on-line implementation scheme which first divides data into segments with similar dynamical properties and then analyzes the individual segments is outlined. Through two simple examples, it is shown that significant errors may be encountered when information on the unstable poles and non-invertible zeros of a system is not properly included in the performance evaluation techniques.

2.1 Introduction

Although enormous research effort has been directed towards design and analysis of controllers, relatively little work has addressed the problem of evaluating the performance of closed loop systems. Nevertheless, automatic monitoring of control loop performance is extremely important for practical control applications, wherein changes in equipment or operating conditions may result in deterioration of a controller which originally functioned well. In a typical chemical manufacturing facility, thousands of control loops are used to track set points and reject disturbances, and manual supervision of each loop is an unwieldy task. As a single control engineer may be

responsible for over a thousand controllers, efficient tools are needed to automatically identify controllers which may need to be re-designed.

In his pioneering work in performance monitoring, Harris [43] presented a correlation test for determining whether a controller achieves minimum variance (MV). When the test fails, an estimate of the factor by which MV is exceeded can be obtained using data collected under routine closed loop conditions. Based on this estimate, Desborough and Harris [24] have introduced a normalized performance index, which expresses the fractional increase in output variance arising from not implementing an MV controller. Stanfelj et al. [97] have outlined a hierarchical method for monitoring and diagnosing performance of control loops based on these tools. Extensions to feedforward/feedback control systems have also appeared [25, 97].

Although these methods have found some application to industrial systems [31, 107], they may only be applied to systems which are stable and whose only nonminimum-phase (NMP) behavior is due to process delay. Many real systems, however, exhibit inverse response. In addition, it is known that for any continuous time system with relative degree larger than two, for sufficiently high sampling rate the sampled system will be NMP even when the underlying continuous system has no right half plane zeros [3]. Several examples where this phenomenon occurs at reasonable sampling rates are given in the cited reference.

In this chapter, extensions of the MV evaluation techniques which may be applied to general systems, including unstable and NMP systems, are derived. Section 2.2 contains some preliminaries, and Section 2.3 briefly reviews the most important past results. In Section 2.4, the main results are presented. Section 2.5 analyzes the effect of uncertainty in the plant parameters on the MV test, and issues surrounding on-line implementation are discussed in Section 2.6. Section 2.7 presents two examples which illustrate the types of errors which may be expected when NMP and unstable behavior are not considered.

2.2 Preliminaries

Consider sampled systems which can be adequately represented by discrete time models. Let q denote the delay operator (i.e., $qy_n = y_{n-1}$ or $q = z^{-1}$). Consider a process described by the following discrete dynamic model relating the output y to the manipulated input u and a disturbance d:

$$y = P(q)u + d. \tag{2.1}$$

Let $P(q)$ have θ+1 zeros at $q = 0$ (i.e., , θ is the number of integer periods of delay not including the sampling delay). Call P minimum phase modulo θ (MP/θ) if it has no other zeros within the unit disk $\{q|\ |q| < 1\}$ or, in other words, if $Pq^{-(\theta+1)}$ is MP. d is a disturbance signal generated by

$$d = W(q)e, \tag{2.2}$$

where W is a stable transfer function, with the possible exception of poles located at $q = 1$, and e is an independently and identically distributed random sequence with zero mean, constant spectrum, and variance σ_e^2 (white noise). The performance of the dynamic system described in (2.1) and (2.2) under output feedback with the input u obtained by the control law,

$$u = C(q)y, \tag{2.3}$$

will be analyzed. This system is depicted in Figure 2.1.

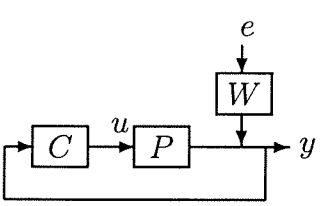

Figure 2.1: Closed loop system

Let Φ be the closed-loop transfer function from e to y:

$$\Phi(q) = \frac{W(q)}{1 - P(q)C(q)}. \tag{2.4}$$

Denote by ϕ the sequence of impulse response coefficients of Φ. The variance σ_y of the output y is given by:

$$\begin{aligned} \sigma_y^2 &= E(y_n^2) \\ &= \sum_{i=0}^{\infty} \phi_i^2 \sigma_e^2, \end{aligned}$$

where $E(\cdot)$ denotes the expectation operator. Also, let MSE(y) denote the mean square error of y,

$$\text{MSE}(y) = \frac{1}{N} \sum_{k=1}^{N} y_k^2.$$

2.3 MV control assessment: stable, MP/θ case

Minimizing the variance of y is equivalent to minimizing the l_2-norm of the sequence ϕ, or equivalently by Parseval's Theorem, the H_2-norm of transfer function $\Phi(q)$. This section reviews results aimed at assessing whether a system achieves the MV control objective. Section 2.3.1 contains a PASS/FAIL diagnostic tool, and Section 2.3.2 an approach for estimating by what factor MV is exceeded.

2.3.1 MV correlation test

The following correlation test has been proposed as a tool to assess whether MV is achieved:

Result 1 *[43] For a stable, MP/θ process P the theoretical autocorrelation of the controlled variable under MV control will be zero beyond lag θ.*

This result is derived from the fact that for stable MP/θ plants P, the MV control law results in a finite impulse response transfer function of length $\theta + 1$ [14, 4]:

$$\Phi(q) = \phi_0 + \phi_1 q + \ldots + \phi_\theta q^\theta.$$

In the language of time-series analysis, such a process is called a moving-average process of order θ (MA(θ)), and has the property that the autocorrelation function, $\rho_{yy}[p] = E(y_k y_{k-p})$, is zero for $p > \theta$.

2.3.2 Estimation of MV from data

Minimum variance is an ideal that is rarely met in practice. Frequently, this criterion may not be a suitable objective since it does not place any weighting on control action. Also, to achieve MV, reasonably accurate disturbance models must be available, whereas typical disturbances have time varying characteristics which are not well known or easily predicted. Therefore, the MV test stated above will fail in most practical cases, and when it fails it gives no information about how close the performance comes to the theoretical limit. When the output variance is determined to be suboptimal, an estimate of the extent to which the observed variance exceeds the theoretical limit serves as a useful performance measure. The following result of Harris [43] can be used to estimate the MV from closed loop data:

Result 2 *[43] For a stable MP/θ process, the minimum achievable variance can be estimated as follows:*

1. *Determine delay θ.*

2. *Fit a time-series model (e.g., AR, ARMA) $\hat{\Phi}(q)$ to the closed loop process output. Estimate variance $\hat{\sigma}_e^2$ from the model residuals.*

3. *The theoretical MV is estimated by $\hat{\sigma}_{mv}^2 = \sum_{i=0}^{i=\theta} \hat{\phi}_i^2 \hat{\sigma}_e^2$. The ratio*

$$r_{mv} = \frac{\mathrm{MSE}(y)}{\hat{\sigma}_{mv}^2}$$

gives an estimate of the factor by which MV is exceeded.

This result in based upon the fact that the first $\theta+1$ coefficients of the closed loop system are independent of control action, and when the only NMP characteristics of the plant are delays, MV corresponds to $\phi_i = 0$ for $i > \theta$.

2.4 MV control: general case

By parameterizing all stabilizing controllers, the results of Section 2.3 can be extended in a straightforward fashion. The results of Sections 2.3.1 and 2.3.2 are extended in Sections 2.4.1 and 2.4.2 respectively. Section 2.4.3 briefly outlines how a similar approach could be used for performance objectives other than minimum variance.

2.4.1 MV correlation test: general case

In this section, the correlation test is extended to the general case wherein the plant may be unstable or NMP. Let us consider anew the problem of minimizing the l_2-norm of the closed loop system given in (2.4). For SISO systems, the parameterization of all stabilizing controllers is well known [105] and given by:

$$C(q) = \frac{-U(q) + Q(q)D(q)}{V(q) + Q(q)N(q)},$$

where Q is stable, N, D, U, and V are each stable and form a coprime factorization of P, i.e.,

$$P(q) = \frac{N(q)}{D(q)}$$

$$U(q)N(q) + D(q)V(q) = 1.$$

Using this parameterization, the set of achievable closed loop maps is given by

$$\begin{aligned} \Phi(q) &= W(q)\left(D(q)V(q) + N(q)D(q)Q(q)\right) \\ &\equiv H(q) - T(q)Q(q). \end{aligned} \tag{2.5}$$

Minimization of $||\Phi(q)||_2$ subject to Q stable is well understood. The following discussion parallels that given in [22]. A stable Q cannot cancel the zeros of T which lie within the unit disk, and this fact can be used to reformulate the stability constraints as interpolation conditions. Let $\boldsymbol{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be the zeros of T within the unit circle with multiplicities $\{m_1, m_2, \ldots, m_n\}$, and let $M = \sum_i m_i$. Note that this set contains both the poles and the zeros of P which lie within the unit disk. The optimization can now be seen to be equivalent to

$$\min ||\Phi||_2^2$$

$$\text{s. t. } \left.\frac{d^k\Phi(q)}{dq^k}\right|_{q=\alpha_i} = \left.\frac{d^kH(q)}{dq^k}\right|_{q=\alpha_i} \tag{2.6}$$

$$k = 0, \ldots m_i - 1, \; i = 1, \ldots, n.$$

These constraints are linear functions of the impulse response coefficients of $\Phi(q)$, ϕ, and can be expressed using functionals. Let $V^{i,k}$ be the linear functional which takes ϕ onto $\left.\frac{d^k\Phi(q)}{dq^k}\right|_{q=\alpha_i}$, i.e., $\left.\frac{d^k\Phi(q)}{dq^k}\right|_{q=\alpha_i} = \sum_{n=0}^{\infty} V_n^{i,k}\phi_n$. $V_n^{i,k}$ is given by

$$V_n^{i,k} = \begin{cases} \frac{n!}{(n-k)!}\alpha_i^{n-k} & \text{if } n \geq k \\ 0 & \text{if } n < k \end{cases}.$$

Let V be the $\infty \times M$-dimensional matrix whose columns are composed of the functionals $V^{i,k}$, and let b be the M-dimensional vector whose elements are the corresponding right-hand side values of (2.6). The optimization problem can now be rewritten as follows:

$$\min \sum_{n=0}^{\infty} \phi_n^2 \;\; \text{s.t. } \mathrm{V}^{\mathrm{T}}\phi = \mathrm{b}$$

The solution to this problem is given by the normal equations [63]

$$\phi_{mv} = [\phi_0, \phi_1, \ldots]^T = V(V^TV)^{-1}b. \tag{2.7}$$

The matrix (V^TV) is a $M \times M$ dimensional matrix, and since $|\alpha_i| < 1$ for $i =$

$1, \ldots, n$, $(V^T V)$ is finite. ϕ_{mv} is thus a linear combination of the columns of V. The transfer function whose impulse response is given by $V_n^{i,k}$ is

$$\begin{aligned} V^{i,k}(q) &= q^k \frac{d^k}{dx^k} \frac{1}{1-x}\bigg|_{x=\alpha_i q} \\ &= q^k k! \frac{1}{(1-\alpha_i q)^{k+1}}, \end{aligned} \tag{2.8}$$

so that

$$\Phi_{mv}(q) = \sum_{i=1}^{n} \sum_{k=0}^{m_i-1} \beta^{i,k} V^{i,k}(q), \tag{2.9}$$

where $\beta^{i,k}$ corresponds to the appropriate element of the vector $(V^T V)^{-1} b$.

This construction clearly reveals that the minimum-variance control results in a closed loop FIR transfer function only when $\alpha_i = 0 \ \forall i$, that is P is MP/θ. A generalization of Result 1 which can be applied to arbitrary systems can now be stated.

Theorem 1 *Let P be a discrete-time dynamic process, and let $y = Pu + d$ where d is filtered white noise and u is obtained via a feedback control law. Let $\boldsymbol{A} = \{\alpha_1, \ldots, \alpha_n\}$ contain the poles and zeros of P which lie within the unit disk in the q-plane, and let α_i have multiplicity m_i. Let ψ be a sequence obtained by filtering the output of the closed loop system y as follows:*

$$\psi = \left[\prod_{i=1}^{n} (1 - \alpha_i q)^{m_i} \right] y \tag{2.10}$$

Then the closed loop system is MV if and only if

$$\rho_{\psi\psi}[p] = E\left(\psi_k \psi_{k-p}\right) = 0 \text{ for } p \geq M, \tag{2.11}$$

where $M = \sum_{i=1}^{n} m_i$.

Proof *From (2.9) and (2.8), the minimum variance closed loop response has the*

transfer function

$$\Phi_{mv}(q) = \sum_{i=1}^{n} \sum_{k=0}^{m_i-1} \beta^{i,k} k! q^k \frac{1}{(1-\alpha_i q)^{k+1}}. \tag{2.12}$$

Since under MV control, $y = \Phi_{mv}(q)e$, ψ *is given by*

$$\begin{aligned} \psi &= \sum_{i=1}^{n} \sum_{k=0}^{m_i-1} \beta^{i,k} k! P^{i,k}(q) e \\ P^{i,k}(q) &= q^k (1-\alpha_i q)^{m_i-k-1} \prod_{j \neq i} (1-\alpha_j q)^{m_j}. \end{aligned}$$

Each polynomial $P^{i,k}(q)$ *is degree* $M-1$, *and therefore* ψ *is a moving average process of order* $M-1$, *whose autocorrelation will be zero beyond lag* $M-1$. *This completes the proof of the necessity.*

For sufficiency, consider a signal ψ *which satisfies the autocorrelation condition.* ψ *is a moving average process of length* $M-1$, *and therefore there exist parameters* b *such that*

$$\psi = \sum_{i=0}^{M-1} b_i q_i e = B(q)e$$

By inverting (2.10), y *can be related to* ψ*:*

$$\begin{aligned} y &= \frac{1}{\prod_{i=1}^{n}(1-\alpha_i q)^{m_i}} \psi \\ &= \frac{\sum_{i=0}^{M-1} b_i q^i}{\prod_{i=1}^{n}(1-\alpha_i q)^{m_i}} e \\ &= \Phi(q)e. \end{aligned} \tag{2.13}$$

Φ *has* M *free parameters* b_i, *but (2.6) provides* M *independent equations for these parameters. The minimum-variance optimal* Φ_{mv} *given in (2.9) is also a rational transfer function with numerator of degree* M *and the same denominator as given in (2.13), satisfying the constraints in (2.6). Therefore,* $\Phi(q) = \Phi_{mv}(q)$.

Remark 1 *In the case where* $\alpha_i = 0$ *for each* i, $\psi = y$, *and* $M = \theta + 1$, *so the result reduces to Result 1.*

Remark 2 *The only process information needed to apply Result 1 is the number of*

integer periods of delay θ. By contrast, Theorem 1 requires knowledge of the location and multiplicity of all unstable poles and non-invertible zeros of the underlying process.

2.4.2 Estimation of MV from data: general case

In this section, Harris' result concerning the estimation of MV from output data is extended to the general case. Consider the parameterization of closed-loop stable transfer functions given in (2.5). Let $q = \alpha_i$, $|\alpha_i| < 1$, be a zero or pole of P with multiplicity m_i. From (2.6), Φ and its first $m_i - 1$ derivatives evaluated at α_i are independent of Q and thus independent of the controller, i.e., for any stable Φ,

$$\left.\frac{d^k\Phi(q)}{dq^k}\right|_{q=\alpha_i} = \left.\frac{d^k\Phi_{mv}(q)}{dq^k}\right|_{q=\alpha_i} \tag{2.14}$$

for $k = 0, \ldots, m_i - 1$, $i = 1, \ldots, n$. The MV response has the form (2.12), and contains M parameters. If a model of the closed loop system $\Phi(q)$ were known, the following procedure could be used to estimate the minimum achievable performance from closed loop data:

1. Determine the plant poles and zeros which lie within the unit disk in the q-domain.

2. Fit the M parameters $\beta^{i,k}$ in Φ_{mv} in (2.12) using the M equations (2.14).

Typically, the closed loop system model $\Phi(q)$ is not known a priori and must be estimated using time series methods. For any stable closed loop, the unstable poles of P and C will manifest themselves as closed loop system zeros. Therefore, whenever the plant P (or the controller C) is unstable, the closed loop will be NMP.

Standard time series identification routines, such as those found in [61] and [94], always result in an MP time series model $\Phi(q)$ with $\Phi(0) = 1$. Suppose y is a time series given by $y = He$, where e is white noise with covariance 1. Let the MP-allpass factorization of H be given by $H(q) = H_M(q)H_A(q)$, and define $H'(q)$ by

$$H'(q) = \frac{H_M(q)}{H_M(0)}.$$

Because a white noise signal has constant spectrum, without measuring e it would be impossible to distinguish between the two following scenarios from the spectrum of y:

1. y was generated by $H(q)$ and an unknown signal e_1 with unit covariance, or
2. y was generated by $H'(q)$ and an unknown signal e_2 with covariance $H_M(0)^2$ ($e_2 = H_M(0)H_A(q)e_1$).

Although $||e_1|| < ||e_2||$, identification methods which minimize the prediction error will converge to H' rather than H since calculating the prediction error for a non-invertible system is numerically unstable. With this in mind, the following method for estimating the MV from closed loop data is proposed:

Result 3 *Given the unstable poles of P and C and the zeros of P within the unit disk, the minimum achievable variance can be estimated using routine closed loop operating data by the following technique:*

1. *Fit an MP time series model Φ' to the output data, with $\Phi'(0) = 1$ and calculate σ'^2 from the covariance of the residuals.*
2. *Form $\Phi(q)$ and σ^2 by*

$$\begin{aligned} \Phi(q) &= \prod_{i=0}^{n} \left(\frac{1 - \frac{1}{p_i} q}{1 - p_i q} \right)^{m_i} \Phi'(q) \\ \sigma^2 &= \sigma'^2 \prod_{i=0}^{n} p_i^{-2m_i} \end{aligned}$$

   where $p_1, \ldots, p_n$ are the unstable poles of the plant (or controller) with multiplicity m_i.

3. *Using $\Phi(q)$ and the constraints (2.14), estimate the free parameters in $\Phi_{mv}(q)$ as in (2.12).*
4. *$\hat{\sigma}^2_{mv} = ||\Phi_{mv}||^2 \sigma^2$ is an estimate of the minimum achievable variance. The ratio*

$$r_{mv} = \frac{\mathrm{MSE}(y)}{\hat{\sigma}^2_{mv}}$$

gives an estimate of the factor by which MV is exceeded.

Remark 3 *In the case of a stable plant and controller and* $\boldsymbol{A} = \{0\}$, $m_1 = \theta + 1$, *and the equations reduce to the method used in Result 2.*

2.4.3 Other performance measures

The estimation method discussed above, involving recasting stability criteria as interpolation conditions and relating these interpolation conditions to values obtained from closed loop data, is not limited to the case of minimum variance control. For example, performance objectives other than the minimization of the l_2 norm of the closed loop could also be implemented. Although the optimization solutions will not be obtained by Euclidean projection as in (2.7), the interpolation conditions can be obtained in exactly the same way as in the MV case. For example, performance could be compared to the l_1 optimal. The problem involves solving a linear program. Let ϕ be a sequence denoting the impulse response of the closed loop. Then an infinite dimensional linear program which minimizes the l_1 norm of ϕ is given by

$$\begin{aligned} \min \quad & \sum_{i=0}^{\infty} \phi_i^+ + \phi_i^- \\ \text{s.t.} \quad & V^{i,k}\phi^+ - V^{i,k}\phi^- = b^{i,k} \\ & \phi_n^+ \geq 0 \ \forall n \\ & \phi_n^- \geq 0 \ \forall n \end{aligned} \tag{2.15}$$

It has been shown [22] that although the problem is infinite dimensional, in the optimal solution there exists an N such that $\phi_n = 0 \ \forall n > N$. This value of N can be estimated a priori. In practice, the LP can often be solved quickly be guessing N and checking the solution to assure that $\phi_n \to 0$ before $n = N$.

Note that the values of $b^{i,k}$ are obtained exactly as in the MV case.

2.5 Sensitivity analysis of MV techniques

As previously mentioned, in order to apply Theorem 1, it is necessary to know the location of unstable poles and non-invertible zeros for the underlying process P. The issue of uncertainty in the pole/zero locations is now addressed. The following result applies:

Theorem 2 *Let $\Phi_{mv}(q) = \frac{W}{1-PC_{mv}}$ where $C_{mv}(q)$ is the MV controller for the plant P and the disturbance W. Let $\boldsymbol{A} = \{\alpha_1, \ldots, \alpha_n\}$ contain the poles and zeros of P which lie within the unit disk in the q-plane, and let α_i have multiplicity m_i. Let $y = \Phi_{mv}e$, where e is white noise, and let $\hat{\psi}$ be obtained by filtering y as follows:*

$$\hat{\psi} = \prod_{i=1}^{n}(1 - \hat{\alpha}_i q)^{m_i} y$$

Then for $p \geq M$,

$$E\left(\hat{\psi}_k \hat{\psi}_{k-p}\right) = \sum_{i=1}^{n} \alpha_i^{p-M}(\alpha_i - \hat{\alpha}_i)c_i + O\|\alpha - \hat{\alpha}\|^2$$

where c_i are constants which depend on α but not $\hat{\alpha}$, and $\alpha = [\alpha_1, \ldots, \alpha_n], \hat{\alpha} = [\hat{\alpha}_i, \ldots, \hat{\alpha}_n]$.

Proof *A Taylor expansion of the filter used to obtain $\hat{\psi}$ around the point α yields*

$$\hat{\psi} = \prod_{i=1}^{n}(1 - \alpha^i q)^{m_i} y + \sum_{i=1}^{n}\left(m_i(\alpha_i - \hat{\alpha}_i)\frac{\prod_{j=1}^{n}(1 - \alpha^j q)^{m_j}}{1 - \alpha_i q} q\right) y + O\|\alpha - \hat{\alpha}\|^2$$

Let ψ be the sequence obtained through filtering y by the correct filter from (2.10):

$$\psi = \prod_{i=1}^{n}(1 - \alpha^i q)^{m_i} y,$$

and let $z^i = \frac{1}{1-\alpha^i q}\psi$. Then $\hat{\psi}$ is given by

$$\hat{\psi} = \psi + \sum_{i=1}^{n} m_i(\alpha_i - \hat{\alpha}_i) q z^i + O\|\alpha - \hat{\alpha}\|^2$$

Substituting this expression for $\hat{\psi}$ yields:

$$\begin{aligned} E\left(\hat{\psi}_k\hat{\psi}_{k-p}\right) &= E\left(\psi_k\psi_{k-p}\right) + \\ &\sum_i m_i(\alpha_i - \hat{\alpha}_i)\left[E\left(\psi_k z^i_{k-p-1}\right) + E\left(\psi_{k-p}z^i_{k-1}\right)\right] + O\|\alpha - \hat{\alpha}\|^2 \end{aligned}$$

From Theorem 1, ψ is a moving average process of order M, so that for $p \geq M$, $E\left(\psi_k\psi_{k-p}\right) = 0$. Let b_j be the moving average coefficients of ψ, i.e., $\psi_n = \sum_{j=0}^{M-1} b_j e_{n-j}$. The sequence z^i is similarly given by $z^i_n = \sum_{j=0}^{\infty} g^i_j e_{n-j}$ with g^i_j given by

$$g^i_j = \begin{cases} b_0 & j = 0 \\ b_j + \alpha_i g^i_{j-1} & 1 \leq j \leq M-1 \\ \alpha_i^{j-(M-1)} g^i_{M-1} & j \geq M \end{cases}$$

For $n \geq M$ the expectations $E\left(\psi_k z^i_{k-p-1}\right)$ and $E\left(\psi_{k-p}z^i_{k-1}\right)$ are given by

$$\begin{aligned} E\left(\psi_k z^i_{k-p-1}\right) &= \sigma_e^2 \sum_{j=0}^{M-1} b_j g^i_{j-1-p}, \\ &= 0; \\ E\left(\psi_{k-p}z^i_{k-1}\right) &= \sigma_e^2 \sum_{j=0}^{M-1} b_j g^i_{j+k-1}. \\ &= \alpha_i^{p-M} g^i_{M-1}\sigma_e^2 \sum_{j=0}^{M-1} \alpha_i^j b_j. \end{aligned}$$

Letting $c_i = g^i_{M-1}B(\alpha_i)\sigma_e^2$ gives the final result.

Remark 4 *If the plant is stable and MP/θ, underestimating θ by one delay period results in only one non-zero correlation coefficient beyond the predetermined lag. On the other hand, when the plant in non-MP/θ, the effect of errors in $\hat{\alpha}$ is manifested for all lags greater than M through the decay factor α_i.*

Remark 5 *The proportionality constants c_i can be estimated from the closed loop data by estimating the MV closed loop as in Result 3, forming the filter, and calculating the resulting coefficients b_i.*

2.6 On-line implementation of MV estimation

Recall that the minimum variance estimation involves identifying a time series model which describes the closed loop transfer function, and using this model to evaluate the interpolation conditions (2.14). In order to implement this scheme on-line, one needs efficient methods for identifying time-varying model parameters. Desborough and Harris [24] have proposed restricting the time series model structure to be autoregressive (AR), and using recursive least squares with an appropriate forgetting factor. While this approach is advantageous from a computational standpoint, it has two significant shortcomings:

1. Many closed loop systems are not well modeled by AR time series. For example, consider the simple case of an MP/θ plant driven by a discrete-time Wiener process. Under minimum variance control, the closed loop will be $\sum_{i=0}^{\theta} q^i$, which is purely moving-average (MA). Modeling the closed loop with an inappropriate model structure generally results in biased model estimates which will adversely affect the evaluation of the interpolation conditions.

2. Recursive estimation with forgetting factor essentially discards data which if used could result in more accurate minimum variance estimates. Suppose ϕ parameterizes a model, and let $f(\phi)$ be the function which calculates the minimum variance from the interpolation conditions on ϕ. If $\hat{\phi}$ is an estimate of ϕ obtained from a finite number of data, and $\bar{\phi}$ and Σ are the mean and covariance of this estimate, then under Gaussian assumptions

$$E(f(\hat{\phi})) = f(\bar{\phi}) + \frac{1}{2}\text{trace}\,(\Sigma f_{\phi\phi}) + O(||\Sigma||^2)$$

where $f_{\phi\phi}$ is the Hessian matrix of f. Standard identification techniques, such as prediction error methods, result in a covariance matrix Σ of order $1/N$ where N is the number of data points. The covariance matrix for recursive estimation behaves similarly if the forgetting factor λ is chosen as $(N-1)/(N+1)$. To minimize the error in the estimate $f(\hat{\phi})$, as many points as possible should be

used to estimate ϕ. Using recursive least squares, Σ can be made small by selecting λ nearly 1, but this subsequently increases the delay for detection. More intelligent methods are available for selecting which data should be included in the parameter estimation.

An alternative approach involves dividing the data into segments with similar dynamics. Then each of these segments can be analyzed in an off-line manner. In other words, the following steps are carried out:

1. Use a detection scheme to locate changes in the dynamics of the closed loop. Let t_{-1} denote the last time for which a change was detected, and t denote the current time.

2. Identify a model using all the data in the interval (t_{-1}, t).

By dividing the data in this fashion, the maximum amount of available data can be used to estimate the time series, resulting in a smaller covariance Σ and thus a smaller error in the MV estimate. In addition, more general model structures can be used. For example, including MA coefficients in a time series model necessitates iterative identification procedures whose approximate recursive implementations (extended Kalman filter) converge slowly, but this two-step approach can take advantage of computationally efficient detection algorithms while only requiring the use of a more expensive identification algorithm once a change has been detected. Similarly, model structure evaluation using methods such as cross-validation [98], which may not be feasibly implemented at each time step, can be carried out following a change detection. In practice, structure evaluation may be important since system changes may also produce changes in the model structure.

A few methods which are capable of addressing the problem of segmenting the data are outlined below. A more comprehensive description of these methods can be found in the individual references as well as in the monograph by Basseville and Nikiforov [9] and the survey paper by Basseville [8].

2.6.1 Two-model approach

Consider the case where the closed loop model changes from ϕ^{0*} to ϕ^1 at some time t. The output y is a random sequence with conditional probability measure functions $P_{\phi^{0*}}(y_n|y_{n-1},\ldots)$ and $P_{\phi^1}(y_n|y_{n-1},\ldots)$ before and after change respectively. To test whether the change occurs or whether a single model ϕ^0 was valid for all time up to n, the likelihood of y assuming a change is compared to that assuming a constant model over the entire horizon, yielding the following likelihood ratio:

$$\text{LR} = \frac{\prod_{k=1}^{t-1} P_{\phi^{0*}}(y_k|y_{k-1},\ldots)\prod_{k=t}^{n} P_{\phi^1}(y_k|y_{k-1},\ldots)}{\prod_{k=1}^{n} P_{\phi^0}(y_k|y_{k-1},\ldots)} \tag{2.16}$$

Since the models ϕ^0, ϕ^1, and ϕ^{0*} are generally unknown, they are replaced with their maximum likelihood estimates, yielding a so-called *Generalized Likelihood Ratio* or GLR test. Appel and Brandt [2] have shown that in the Gaussian case, the GLR may be evaluated by first dividing the data and fitting three models as follows:

1. ϕ^0 for data in the window $\{1,\ldots,n\}$
2. $\phi^{0*}(t)$ for data in the window $\{1,\ldots,t-1\}$, and
3. $\phi^1(t)$ for data in the window $\{t,\ldots,n\}$.

The log-likelihood is then given by

$$u_n(t) \;=\; n\log\hat{\sigma}_0^2 - \left[(t-1)\log\hat{\sigma}_{0*}^2(t) + (n-t+1)\log\hat{\sigma}_1^2(t)\right], \tag{2.17}$$

where $\hat{\sigma}_0^2$, $\hat{\sigma}_{0*}^2(t)$, and $\hat{\sigma}_1^2(t)$ are the variances of the innovations from the estimated models ϕ^0, $\phi^{0*}(t)$, and $\phi^1(t)$ respectively.

As the change time t is also unknown, the criterion u_n is replaced with is maximum over t

$$g_n = \max_{1\le t\le n} u_n(t),$$

and g_n is compared to a threshold to decide if a change has occurred. As the computational cost of this algorithm can be quite high due to the search over $t \in [1,n]$,

Appel and Brandt [2] propose modifying the algorithm by fitting ϕ^1 over a window of fixed length L. Equivalently, u_n is calculated only for $t = n - L$. Then if an exact change detection time is desired, when $u_n(n - L)$ exceeds a threshold, an exhaustive local search is used to find the optimal change time. This modification decreases the computational load and does not significantly affect the detection capabilities.

For efficient on-line implementation of this method, AR models should be used so that recursive identification can be done. It has already been argued that AR models may not properly describe the closed loop behavior; however, Basseville [7] has reported that in the area of speech segmentation, although AR models do not describe the speech signals well, the two-model method using AR models does segment the signals well. Since the AR model is used only to segment the data and not to evaluate the minimum variance, the interpolation conditions will not be biased by an incorrect model structure.

Other measures, such as the Kullback divergence, can also be used in the two-model method to determine if a change occurred. The interested reader is referred to [7].

2.6.2 Local approach

The local approach not only has strong theoretical justification, it is also capable of addressing more general models than AR. From a theoretical standpoint, this approach is asymptotically uniformly most powerful, which roughly means that for small changes and large data sets, no other algorithm has a higher probability of correct diagnosis. A brief outline of the local approach as applied to the present problem follows and a more complete treatise of local methods can be found in the references [12] and [26], as well as in Chapter 7.

Consider a time series model whose conditional probability measure function $P_\phi(y_t|y_{t-1}, \ldots)$ is dependent upon the parameters ϕ. Let $Y_\phi(t)$ be defined as the

derivative of $\log P$ with respect to ϕ, that is

$$Y_\phi(t) = \frac{\partial}{\partial \phi'} \log P_{\phi'}(y_t|y_{t-1}, \ldots)\Big|_{\phi'=\phi}.$$

It can be shown that, asymptotically, $Y_\phi(t)$ is a zero mean sequence, normally distributed with covariance given by the Fischer information matrix J where

$$J(\phi) = E\left(\frac{\partial^2}{\partial \phi^2} log P_\phi(y_t|y_{t-1}, \ldots)\right). \tag{2.18}$$

If a small parameter change $\Delta\phi$ occurs, then the mean of $Y_\phi(t)$ changes to $J(\phi)\Delta\phi$. Using the local approach, the problem of detecting the change in the parameters of a time series can be transformed to detecting a change in the mean of a signal, which is one of the simplest change detection problems and can be easily solved using the Cumulative Sum (CUSUM) algorithm [6].

The local approach can be easily implemented on quite general model structures. For example, for an ARMA model of the form

$$y_n + \sum_{i=1}^{N_a} a_i y_{n-i} = e_n + \sum_{i=1}^{N_b} b_i e_{n-i},$$

or equivalently $y = B(q)/A(q)e$, and for e Gaussian white noise, the conditional probability measure function is given by

$$\log P_\phi(y_n|y_{n-1}, \ldots) = -\log\left(\sqrt{2\pi}\sigma\right) - \sigma^{-2}e_n^2$$

where e is obtained by filtering y by $A(q)/B(q)$. The signal $Y_\phi(n)$ can easily be calculated by carrying out the following filtering operations:

$$\begin{aligned} w &= \frac{1}{B(q)}y \\ v &= -\frac{A(q)}{B^2(q)}y \\ \frac{\partial}{\partial a_i}\log P_\phi &= -2e_n q^i w_n \end{aligned}$$

$$\frac{\partial}{\partial b_i} \log P_\phi \;=\; -2e_n q^i v_n.$$

The matrix J can also be easily calculated from the model parameters.

2.7 Examples

This section contains two examples which demonstrate the errors which might be made when the MV techniques are applied without regard to NMP and unstable behavior. Example 1 concerns the estimation of σ^2_{mv} for a plant with a zero inside the unit disk, and Example 2 addresses the issue of NMP behavior in the closed loop due to an unstable controller. In both cases, egregious errors are made when Results 1 and 2 are directly applied.

Example 1 *Let $P(q)$ be a model for an open loop stable system which has a zero at $q = \alpha$ with $|\alpha| < 1$ as well as θ delay periods. Let $H(q)$ be a time-series model fit to the closed loop data.*

Since the process is open loop stable, the parameterization of stable closed loop transfer functions is simply given by [74]

$$\Phi = W(q)\left[1 - P(q)Q(q)\right]. \tag{2.19}$$

The MV closed loop will have the following form:

$$\Phi_{mv}(q) = \sum_{i=0}^{\theta} h_i q^i + q^{\theta+1}\frac{\beta}{1-\alpha q}, \tag{2.20}$$

where h_i are the impulse response coefficients of $H(q)$. Defining the transfer function $H_\theta(q) = \sum_{i=0}^{\theta} h_i q^i$, the stability constraints (2.14) are all met only if the constant β satisfies

$$\beta = \alpha^{-(\theta+1)}(1-\alpha^2)(H(\alpha) - H_\theta(\alpha)).$$

In the form (2.20), the norm of Φ_{mv} can be easily calculated:

$$\|\phi_{mv}\|^2 = \sum_{i=0}^{\theta} h_i^2 + \frac{\beta^2}{1-\alpha^2}. \tag{2.21}$$

If the existence of the zero at α is unknown or neglected, the MV would be incorrectly estimated by

$$\Phi_{mv} = \sum_{i=0}^{\theta} h_i q^i, \quad \|\phi_{mv}\|^2 = \sum_{i=0}^{\theta} h_i^2. \tag{2.22}$$

These values may vary significantly. For example, if $W(q) = \frac{1}{1-q}$, stability would require that $h_i = 1, i = 0, \ldots, \theta$, and also $H(\alpha) = \frac{1}{1-\alpha}$, resulting in $\beta = (1+\alpha)$. Then the correct calculation of minimum variance yields $\|\Phi_{mv}\|^2 = (\theta + 1 + \frac{1+\alpha}{1-\alpha})$, whereas neglecting the effect of the zero would erroneously yield $\|\Phi_{mv}\|^2 = \theta + 1$. For α near 1, errors arising from neglecting the zero will be much more significant than errors arising from a poor estimate of the delay θ.

Example 2 *Consider a process described by the transfer function*

$$P(q) = 20q + q^2,$$

and disturbance model $d = \frac{1}{1-q}e$, where e is a white-noise process. The MV controller is given by

$$C_{mv}(q) = -\frac{1}{(20+q)(1-q)},$$

resulting in a closed loop transfer function $\Phi_{mv}(q) = 1$. Suppose the closed loop is originally operating with the controller C_{mv}, but at time t_0, component failures result in the controller C_2 being implemented, where C_2 is given by

$$C_2(q) = \frac{-22 + 21q}{(1-q)(-21 + 419q + 21q^2)}.$$

Although C_2 is unstable, the closed loop remains internally stable, and can be described by the transfer function

$$\Phi_2(q) = \frac{1}{21}\left(21 - 419q - 21q^2\right).$$

By changing from C_{mv} to C_2, the root mean square output error increases by more than a factor of twenty. Now suppose the MV correlation test were applied to the data both before and after the change. ϕ_{mv} produces the autocorrelation function $\{1, 0, \ldots\}$, whereas ϕ_2 gives $\{1, 0, -\frac{441}{176{,}443}, 0, 0, \ldots\}$. The autocorrelation function for ϕ_2 indicates that it is not minimum-variance; however, over $640,000$ samples would be needed in order for the value $\frac{441}{176{,}443}$ to lie outside the 95% confidence limits. Thus, based solely on the autocorrelation test, one might incorrectly deduce that the increase in output variance is not due to deterioration in control, but to a larger disturbance.

Now suppose the MV ratio is also estimated. Using standard identification software [60], a MA(2) model was fit to 500 points simulated using C_2, yielding the model

$$\hat{\Phi}(q) = 1 - 0.057q - 0.038q^2, \;\; \hat{\sigma}_e^2 = 404.$$

The mean square error of simulated output data was $\mathrm{MSE}(y) = 406$. *If the method of Result 2 is used, the estimated MV ratio r_{mv} is given by*

$$r_{mv} = \frac{\mathrm{MSE}(y)}{\hat{\Phi}^2(0)\hat{\sigma}_e^2} = 1.005,$$

which would confirm the results of the correlation test. On the other hand, if the manipulated variable u is measured and a model of the form

$$u_t = \sum_{i=1}^{3} a_i u_{t-i} + b_0 y_t + b_1 y_{t-1}$$

is identified, the unstable pole in the controller at $q = 0.05$ *is easily detected. Using this information and the method of Result 3, the MV ratio estimate is*

$$\frac{\mathrm{MSE}(y)}{\hat{\Phi}^2(0)\sigma_e^2} = 402,$$

which correctly indicates that the variance exceeds the minimum by a large factor.

2.8 Conclusions

Data analysis tools which may be used to determine if a control system is achieving MV control, and to estimate the factor by which MV is exceeded, have been presented. These tools apply to general plants, including unstable and nonminimum-phase processes, and systems with unstable controllers. The examples of Section 2.7 have shown that estimates of minimum achievable variance obtained without regard to limitations imposed by NMP behavior and instabilities can be drastically incorrect.

For on-line implementation of MV control, a two-step approach has been outlined. The first step uses change detection algorithms to divide the data into segments with similar dynamical properties. Two possible algorithms for segmentation were reviewed. The second step involves analyzing the segments in an off-line fashion. This two-step approach not only allows for more general models than those which may be used effectively with recursive identification methods, but also minimizes errors in the estimated minimum variance by using the largest data segment available.

As the MV methods of this chapter follow directly from the interpolation conditions resulting from the Youla parameterization, the effective use of these methods requires knowledge of the location and multiplicity of unstable poles and non-invertible zeros. For the case where performance degradation is primarily due to changes in the disturbance spectrum (or equivalently the transfer function $W(q)$ in Figure 2.1), these methods should work well since the process parameters do not change; however, when performance degradation is due to changes in the process, including possible drifting of poles and zeros, this approach is only feasible if combined with effective

methods to estimate the value of these parameters.

Although minimum-variance is an optimal control law, in many cases it does not reflect the desired closed loop properties. For example, a near minimum-variance controller may oscillate around the set point and have a large settling time. Some practical objectives, such as minimal overshoot, can be expressed as linear programs as in Section 2.4.3; however, other objectives cannot be translated into optimization problems constrained by the interpolation condition (2.14), in which case other performance monitoring approaches are needed [103]. When implementing a scheme for controller performance monitoring, care should be taken to ensure that the monitored performance objective reflects the desired system behavior. In Chapter 3, an alternative performance monitoring method is presented which allows for more general control objectives.

Chapter 3 Performance Monitoring using Likelihood Ratio Testing

Summary

In this chapter, evaluating deterioration in performance of control systems using closed loop operating data is addressed by proposing a framework in which acceptable performance is expressed as constraints on the closed loop transfer function impulse response coefficients. Using likelihood methods, a hypothesis test is outlined to determine if control deterioration has occurred. The method is applied to a simulation example as well as data from an operational distillation column, and the results are compared to those obtained using minimum variance estimation approaches.

3.1 Introduction

In Chapter 2, performance monitoring methods based upon theoretical variance limits were presented. Even though these *"minimum variance"* methods may provide useful information about achievable performance limits, they have several shortcomings as tools for evaluating deteriorating control. For example, in many cases, minimum variance does not provide a meaningful measure of control performance. Achieving this theoretical limit may require a controller with high band width, or excessive control action which may result in the violation of robustness conditions. In addition, a shift in the ratio of actual variance to minimum variance may be due to either changes in the controller, changes in the plant, or changes in the disturbance spectrum. Whereas changes in the controller or the plant may merit retuning, changes in the disturbance spectrum may not. Finally, changes in performance may be due to changes in delay or in the location of the non-invertible zeros. Since the minimum variance methods

require these parameters to be known, the estimates obtained may be quite poor when incorrect values are used.

In this chapter, an approach to performance monitoring is developed for disturbance rejection problems applied to sampled single input/single output systems whose closed loop behavior may be accurately described by linear models. The goal of the monitoring scheme is to be able to detect violations of performance specifications from routine operating data. In other words, the method should not depend on external stimulation by test signals. Clearly if no disturbance or a constant disturbance drives the system, evaluation of the control performance may be difficult as the steady state behavior of the closed loop system will not reflect the dynamic performance. Therefore, we restrict the class of disturbances which will be considered to those which may be accurately represented by filtered white noise.

This chapter is organized as follows. A framework wherein good performance can be expressed as constraints on the impulse response coefficients is established. In Section 3.2, we show how several practical performance criteria can easily be expressed in this paradigm. Once a meaningful performance criterion is established, a hypothesis testing problem is developed to determine if the performance is being achieved. Section 3.3 shows how this test can be evaluated using a generalized likelihood ratio (GLR) approach. In Section 3.4, threshold selection criteria for the GLR test are enumerated. Extensions to systems subject to command signals are discussed in Section 3.5. Finally, Section 3.6 presents examples which compare the methods of this chapter to minimum variance methods.

3.2 Impulse response performance specifications

In order to evaluate if a given performance criterion is achieved, performance specifications must first be expressed mathematically. For controller synthesis, the performance objectives commonly used are chosen so as to result in a solvable design problem. For example, a linear quadratic objective is often used because the resulting optimal control is easy to compute. Similarly, for robust control of systems with

bounded uncertainty, a H_∞-norm performance objective leads to a solvable problem. However, many performance criteria such as settling time and overshoot are not easily translated into these objectives.

Consider systems of the form shown in Figure 3.1, with e a Gaussian white noise signal. Let ϕ denote the impulse response $[1, \phi_1, \phi_2, \ldots]$ of the closed loop transfer function from e to y. Note that ϕ depends not only on the plant P and the controller C, but also on the disturbance generator W. Also, because the magnitude of the input signal e is unknown, we can always assume $\phi_0 = 1$. Many useful performance criteria can be recast as constraints on the coefficients of the ϕ. We consider some examples here.

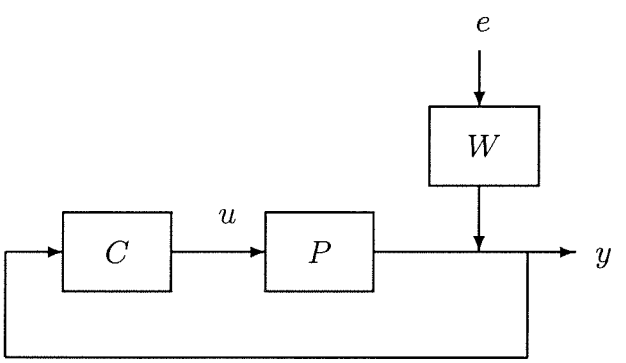

Figure 3.1: Closed loop system

1. **Closed loop settling time**. In industry, this criterion is commonly used to assess controller performance, as is evident from the following quotation from [52]:

   > *Shell control engineers typically aim for a closed loop response with a settling time close to the speed of response of the loop input transfer function when significant upsets occur.*

   Letting τ denote the response time, this performance objective could be stated as

   $$|\phi_t| < \delta\phi_0 \text{ for } t > \tau, \tag{3.1}$$

where δ is a suitably small constant.

2. **Decay rate.** Alternatively, good performance may be specified by requiring that after t_0 sampling periods, ϕ decay no slower than exponentially with time constant T. This objective translates to the following bounds on ϕ:

$$|\phi_t| < \exp\left(-\frac{t-t_0}{T}\right)\phi_0 \text{ for } t > t_0. \tag{3.2}$$

3. **Minimum variance.** For processes whose only non-invertible zeros lie at infinity, the minimum variance closed loop response is a moving average process of order d where d is the number of delays [43]. Minimum variance control could then be expressed by the impulse response constraints:

$$\phi_t = 0 \text{ for } t > d. \tag{3.3}$$

4. **Frequency domain bounds.** Often, it is desired to keep the frequency response of the closed loop ϕ small over a specified frequency range [74]. This is expressed by

$$\left|\sum_{k=0}^{\infty} e^{-jk\omega}\phi_k\right| \leq b(\omega). \tag{3.4}$$

Although this constraint is non-linear in the coefficients of ϕ, it can be approximated by linear constraints by noting that $|z| < 1$ for z complex is equivalent to $\Re(z)\cos(\theta) + \Im(z)\sin(\theta) < 1$ for $0 \leq \theta < 2\pi$. For ϕ_k real, this transforms the above constraint to the set of linear constraints

$$\sum_{k=0}^{\infty} \phi_k \cos(k\omega + \theta) \leq b(\omega) \text{ for all } \theta \in [0, 2\pi]. \tag{3.5}$$

Although in general this amounts to an infinite number of constraints at each frequency, approximating with a finite discretization of θ is usually acceptable.

5. **Filtered coefficient constraints.** Consider as an example the case where rejection of step-like disturbances is important, but due to controller bandwidth

limitations or modeling errors, high frequency oscillations must be tolerated. A meaningful performance criterion should not depend on high frequency phenomena. By constraining the low frequency components of the closed loop impulse response coefficients, this feature may be built into the performance criterion. In general, filtering introduces phase distortion. As a result, the filtered time response may look quite different from the unfiltered time response, even when the time response is band-limited by the filter cutoff frequency. When a filter with linear phase is used, the phase distortion results in a time shift of the time response coefficient for which one can easily compensate [83]. An FIR low pass filter with linear phase can be designed using the method of [65], which has been incorporated in MATLAB as the command `remez` [59]. Implementing such a linear filter on the impulse response coefficients will result in a delayed approximation of the response to low frequency disturbances. Thus, meaningful performance constraints will take the form

$$|\phi_k^F| < b_k \phi_D^F \tag{3.6}$$

for $k > D$, where $\phi_k^F = \sum_{i=0}^{k} F_{k-i}\phi_i$ are the filtered impulse response coefficients, $[F_0, \ldots, F_{2D}]$ is the FIR low pass filter, and D is the delay associated with this filter, that is the slope of the phase.

Each of the above examples involve linear equality or inequality constraints on the closed loop impulse response coefficients of the form:

$$A\phi \leq b, \tag{3.7}$$

where A is an appropriately dimensioned matrix and b is a vector. Although ϕ is an infinite dimensional vector, realistic performance criteria will restrict ϕ_t to be arbitrarily small for $t \geq t^*$, where t^* is not too large, implying that ϕ can be accurately approximated by a finite impulse response model of reasonable length when the system is performing satisfactorily.

Let us now formally state the assumptions which have been introduced in this Section.

Assumption 3.2.1 *The signal e in Figure 3.1 is a Gaussian white noise signal.*

Assumption 3.2.2 *The closed loop transfer function from e to y in Figure 3.1 can be accurately represented by a linear FIR model when good performance is achieved, and the performance objective can be expressed as constraints on the impulse response coefficients as in (3.7).*

Just as a linear quadratic objective leads to straightforward optimal control design, specifying performance criteria as bounds on ϕ results in performance evaluation methods with tractable solutions. In addition, although objectives of the form (3.7) are simple, they can represent a wide variety of realistic performance measures. In the next section, methods for detecting violations of performance criteria of the form (3.7) are developed. In applying the methods of this chapter, effort should be made so that the performance constraints are properly formulated to reflect the desired closed loop properties so that the tests are meaningful.

3.3 Generalized Likelihood Ratio test

Performance evaluation of control systems can be viewed as choosing between the two hypotheses given by:

$\mathcal{H}_0$: The closed loop behavior satisfies the performance objective,

$\mathcal{H}_1$: The closed loop behavior violates the performance objective.

For a disturbance and noise free system, the hypothesis test could be quite simple, for example noting the settling time for a set point step change. Unfortunately, many industrial control systems do not satisfy these conditions. Unmeasured disturbances which are stochastic and perhaps time-varying affect the output, and the performance criterion of interest will frequently address the capability of the control system to reject such unknown disturbances.

Throughout the ensuing discussion, the following terminology and notation will be used. Let θ denote a parameterization of a fixed model. Furthermore, let Θ be a set of such parameterizations. For example, θ may be the vector of coefficients for an n^{th} order Moving Average (MA(n)) model, and Θ the set of all MA(n) models. For convenience, we will refer to Θ as a set of models, and an element $\theta \in \Theta$ as a model. We also denote by ϕ_θ the impulse response coefficients of a model θ.

By assuming that the noise is Gaussian with zero mean and that its covariance σ^2 is part of the model parameterization, a probability density function may be associated with each model $\theta \in \Theta$. Let $\mathcal{Y}_L(t)$ be the vector containing the L most recent measurements, i.e., $\mathcal{Y}_L(t) = [y(t), y(t-1), \ldots, y(t-L+1)]$. Given a model θ, we denote the probability of $\mathcal{Y}_L(t)$ by $p_\theta(\mathcal{Y}_L(t))$.

Example Suppose $\mathcal{Y}_L(t)$ is generated by a moving average stochastic model,

$$y(t) = e(t) + \sum_{i=1}^{m} b_i e(t-i), \tag{3.8}$$

where e is Gaussian white noise with covariance σ^2. Then $\theta = [b_1, \ldots, b_m, \sigma^2]$. The probability density function p_θ is given by

$$p_\theta(\mathcal{Y}_L(t)) = \prod_{k=t-L-m}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\epsilon^2(k,\theta)}{2\sigma^2}\right), \tag{3.9}$$

with $\epsilon(k,\theta)$ obtained by filtering y according to

$$\epsilon = \frac{1}{1 + \sum_{i=1}^{m} b_i z^{-i}} y. \tag{3.10}$$

In order to calculate ϵ, the transfer function $1 + \sum_{i=1}^{m} b_i z^{-i}$ should be invertible. In practice this does not pose a problem. As the signal e is unknown, the closed loop transfer function is essentially fit to the spectrum of the output, and it is well known that for any non-invertible transfer function, there exists an invertible transfer function with the same spectrum.

We now consider the problem of evaluating the performance of a closed loop

system. To do so, first define two model sets, Θ_0 and Θ_1, corresponding to hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$:

$$\begin{aligned}\Theta_0 &= \{\theta|\text{Performance Satisfied}\},\\ \Theta_1 &= \{\theta|\text{Performance Violated}\}.\end{aligned}$$

Since a model cannot both satisfy and violate the performance specifications, the two set Θ_0 and Θ_1 will be disjoint. By defining the model sets in this fashion, the hypotheses become

$$\begin{aligned}\mathcal{H}_0 :&\ \mathcal{Y}_L(t) \text{ generated by } \theta \in \Theta_0,\\ \mathcal{H}_1 :&\ \mathcal{Y}_L(t) \text{ generated by } \theta \in \Theta_1\ .\end{aligned}$$

To address testing the hypotheses $\mathcal{H}_0$ against $\mathcal{H}_1$, consider first the idealized situation where good performance corresponds to exactly one closed loop model, θ_0, and poor performance corresponds to a different model θ_1. Such a hypothesis testing problem in which each hypothesis can be reduced to a single value in the space of probability distributions is referred to as a *simple* test. In this case, a hypothesis can be chosen using a likelihood ratio test in which the ratio of the probability density functions p_{θ_0} to p_{θ_1} is compared to a threshold λ. The decision function g_λ is given by:

$$g_\lambda(\mathcal{Y}_L(t)) = \begin{cases} \mathcal{H}_0, & \frac{p_{\theta_0}(\mathcal{Y}_L(t))}{p_{\theta_1}(\mathcal{Y}_L(t))} > \lambda, \\ \\ \mathcal{H}_1, & \frac{p_{\theta_0}(\mathcal{Y}_L(t))}{p_{\theta_1}(\mathcal{Y}_L(t))} \leq \lambda. \end{cases} \tag{3.11}$$

Two measures used to evaluate a decision function g_λ are the *size* and the *power*. The size α of a test is the probability of a false $\mathcal{H}_1$ diagnosis,

$$\alpha(\lambda) = P(g_\lambda(\mathcal{Y}_L(t)) = \mathcal{H}_1|\mathcal{H}_0), \tag{3.12}$$

and the power β is the probability of a correct $\mathcal{H}_1$ diagnosis:

$$\beta(\lambda) = P(g_\lambda(\mathcal{Y}_L(t)) = \mathcal{H}_1|\mathcal{H}_1). \tag{3.13}$$

The fundamental lemma of Neyman and Pearson [57] implies that for a simple hypothesis test, the most powerful test for a fixed size is necessarily based upon the likelihood ratio. The likelihood ratio approach has found wide use in detection problems [9] and is the basis of many common schemes, such as the cumulative sum, or CUSUM, algorithm.

In practice, the assumption of a simple hypothesis for which the likelihood ratio test may be applied will be violated. Many different closed loop models will satisfy the performance criteria, and many other models will violate them so that both Θ_0 and Θ_1 will contain many elements. In hypothesis testing terminology, this situation is called a *composite* hypothesis testing problem. A common way to deal with composite hypothesis testing problems is to formulate a likelihood ratio test with the single probability density function p_θ replaced by the density function corresponding to the most likely hypothesis in the set, i.e., $\sup_{\theta\in\Theta} p_\theta$. Replacing the densities with their maximum likelihood estimates results in the *Generalized Likelihood Ratio*, or GLR, criterion. The GLR is calculated as

$$\text{GLR} = \frac{\sup_{\theta_0\in\Theta_0} p_{\theta_0}(\mathcal{Y}_L(t))}{\sup_{\theta_1\in\Theta_1} p_{\theta_1}(\mathcal{Y}_L(t))}. \tag{3.14}$$

The decision function g_λ based on the GLR is given by:

$$g_\lambda(\mathcal{Y}_L(t)) = \begin{cases} \mathcal{H}_0, & \text{GLR} > \lambda, \\ \mathcal{H}_1, & \text{GLR} \leq \lambda. \end{cases} \tag{3.15}$$

Although the exact optimality properties of the GLR are unknown, for many special cases the GLR test is optimal [9].

In the case of Gaussian white noise, the probability measure function is given by

$$p_\theta(\mathcal{Y}_L(t)) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\sum_{k=t-L+1}^{t} \epsilon^2(k,\theta)}{2\sigma^2}\right), \tag{3.16}$$

where $\epsilon(k,\theta)$ is the model residual, which is obtained from the data and model through

the relation

$$\epsilon(k,\theta) = y(k) - E_\theta\left(y(k)|y(k-1),\ldots,y(t-L+1)\right), \tag{3.17}$$

with E_θ the conditional expectation under model θ. Since the parameter σ^2 does not affect the calculation of $\epsilon(k,\theta)$, it is possible to calculate the optimal σ^2 as a function of the other parameters. Let the model parameters be partitioned as $\theta = [\theta^*, \sigma^2]$. For any $\epsilon(\cdot,\theta^*)$, the Gaussian probability measure function is maximized with respect to σ^2 when σ^2 takes on the value $V(\mathcal{Y}_L(t),\theta^*)$ which is defined to be the mean residual square error:

$$\sigma^2 = V(\mathcal{Y}_L(t),\theta^*) \equiv \frac{1}{L}\sum_{k=t-L+1}^{t}\epsilon^2(k,\theta^*). \tag{3.18}$$

This relation can be easily verified by setting the partial derivative of p_θ with respect to σ^2 to zero. Substituting this value of σ^2 into the probability measure function, it follows that

$$\sup_\theta p_\theta(\mathcal{Y}_L(t)) = \sup_{\theta^*}\left(\frac{1}{\sqrt{2\pi V(\mathcal{Y},\theta^*)}}\right)^L \exp(-L/2). \tag{3.19}$$

Therefore, the supremum is achieved by minimizing the sum of squares of the residuals, that is

$$\begin{aligned}\hat{\theta}^* &= \arg\min_{\theta\in\Theta} V(\mathcal{Y}_L(t),\theta^*)\\ \hat{\sigma}^2 &= V(\mathcal{Y}_L(t),\hat{\theta}^*).\end{aligned} \tag{3.20}$$

The GLR then becomes

$$\text{GLR} = \left(\frac{V(\mathcal{Y}_L,\hat{\theta}_1)}{V(\mathcal{Y}_L,\hat{\theta}_0)}\right)^{\frac{L}{2}}. \tag{3.21}$$

We see that for the case of Gaussian white noise e, the GLR is calculated by minimizing the residual square error over the two sets Θ_0 and Θ_1. Assumption 3.2.1 guarantees that the driving force for the closed loop transfer function is persistently exciting so that it will be possible to identify the transfer function from routine operating data.

Now let us consider the specific case where the performance specifications are given as linear constraints on the impulse response coefficients, as discussed in Section 3.2. With the performance criterion as in (3.7), the natural structure for models in Θ_0 is that of moving average models. Although autoregressive models have the advantage that the residuals ϵ are linear functions of the model parameters, thus reducing the estimation of $\hat{\theta}_0$ in the unconstrained case to a least squares problem, they have the disadvantage that constraints of the form (3.7) are quite cumbersome to express in terms of the free model parameters. By contrast, moving average models are more difficult to estimate, but the performance criteria are easily implemented. As efficient numerical methods for estimating moving average model parameters are well known [60, 61, 94], we do not view this difficulty to be overly burdensome. Our experience has shown that including constraints in the estimation of MA coefficients only slightly increases the computational burden. We note that using models with both AR and MA terms, which in most cases gives the most parsimonious fit to a generic time series, also makes implementing constraints of the form (3.7) difficult. For these reasons, the model set Θ_0 in this chapter will be restricted to moving average models.

Although it is important that the set of models Θ_0 accurately describe the closed loop system under good performance, it is less important that some element of Θ_1 provide an unbiased fit when the performance specification is violated. Indeed, since performance violation could result in significantly different model structure which may be completely unpredictable, it will frequently be quite difficult to select a set Θ_1 which fits deteriorated performance conditions in an unbiased manner. What is important for determining performance deterioration is that when the system is not performing well, some model in Θ_1 fits the data much better than any model from the set Θ_0. When Θ_0 is the set of MA(t^*) models satisfying the constraints $A\phi_\theta \leq b$, this feature can often by achieved by allowing Θ_1 to be the complement of Θ_0 in the set of MA(t^*) models, even though this structure may not adequately describe the closed loop under deteriorated performance. Therefore, the sets Θ_0 and Θ_1 are of the

form:

$$\begin{aligned}\Theta_0 &= \{\theta \in MA(t^*) | A\phi_\theta \leq b\},\\ \Theta_1 &= \{\theta \in MA(t^*) | \max_i (A_i\phi_\theta - b_i) > 0\} = MA(t^*) \setminus \Theta_0,\end{aligned} \tag{3.22}$$

where t^* is the order of the MA process, and A_i and b_i represent the i^{th} row of the matrix A and the i^{th} element of the vector b respectively.

Evaluation of the GLR requires that the residual square error V be minimized over each of the sets Θ_0 and Θ_1. The performance criterion has been formulated so that the set Θ_0 is convex, and therefore the optimization over Θ_0 will be straightforward; however, the set Θ_1 is clearly non-convex and in general its optimal point will be difficult to calculate. Let us therefore consider another model, $\hat{\theta}_u$, the optimal unconstrained model:

$$\hat{\theta}_u = \arg \min_{\Theta_0 \cup \Theta_1} V(\mathcal{Y}_L, \theta). \tag{3.23}$$

Since the two sets Θ_0 and Θ_1 are disjoint, it follows that $\hat{\theta}_u$ lies in exactly one of these sets. Therefore, the following implications hold:

$$\hat{\theta}_u \in \Theta_0 \Rightarrow V(\mathcal{Y}, \hat{\theta}_0) < V(\mathcal{Y}, \hat{\theta}_1) \Rightarrow GLR > 1, \tag{3.24}$$

$$\hat{\theta}_u \in \Theta_1 \Rightarrow V(\mathcal{Y}, \hat{\theta}_0) > V(\mathcal{Y}, \hat{\theta}_1) \Rightarrow GLR < 1. \tag{3.25}$$

Now suppose that the decision function g_λ is as in (3.15) with $\lambda < 1$. Intuitively, the assumption $\lambda < 1$ is appealing as it implies that if the optimal unconstrained model $\hat{\theta}_u$ satisfies the performance constraints, $\mathcal{H}_0$ is accepted. Under this assumption, it is never necessary to perform the optimization over the set Θ_1. Instead, the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ defined by the model sets (3.22) are tested by implementing the following steps:

1. Calculate the unconstrained model $\hat{\theta}_u$ given by (3.23).

2. Check if $\hat{\theta}_u$ satisfies the constraints $A\phi_{\hat{\theta}_u} \leq b$. If so, accept $\mathcal{H}_0$; otherwise, proceed to step 3.

3. Find $\hat{\theta}_0$ by solving the optimization problem (3.20) for the set Θ_0, and then calculate the GLR via (3.21) with $\hat{\theta}_1 = \hat{\theta}_u$.

4. Use the decision function g_λ given by (3.15) to choose $\mathcal{H}_0$ or $\mathcal{H}_1$.

Implementation of this scheme requires the selection of a threshold λ. In Section 3.4, methods for choosing a threshold are discussed.

3.4 Threshold selection

In the previous section, a generalized likelihood ratio test for determining whether a control system meets performance constraints was presented. Typically, the threshold λ is obtained by choosing an acceptable false alarm rate, given by the size α of the test, and calculating the value of λ which yields this false alarm rate. Applying this approach to the detection problem outlined in the previous section, the false alarm rate is given by

$$\begin{aligned} \alpha(\lambda) &= P(g_\lambda(\mathcal{Y}) = \mathcal{H}_1 | \mathcal{H}_0), \\ &= P(GLR \leq \lambda | \mathcal{H}_0), \end{aligned} \tag{3.26}$$

which gives the relationship between the threshold λ and the false alarm rate of the test g_λ. Let $\bar{\theta}$ denote the actual system associated with the data $\mathcal{Y}$. To evaluate α, for each value of $\bar{\theta}$ a probability distribution function (pdf) for the GLR must be calculated from the estimation schemes used to calculate $\hat{\theta}_u$ and $\hat{\theta}_0$. Then, a pdf for $\bar{\theta}$, $\bar{p}(\bar{\theta})$ must be specified. The latter distribution determines the likelihood of various model and disturbance changes which may affect the system. Once these two pdf have been determined, the size of the test α may be calculated:

$$\alpha(\lambda) = \int_{\Theta_0} P(GLR \leq \lambda | \bar{\theta}) \bar{p}(\bar{\theta}) d\bar{\theta}. \tag{3.27}$$

Although this gives a theoretical expression for α in terms of λ, two primary difficulties render this approach impractical. First, no systematic method exists for deriving the

pdf for the GLR from the estimation schemes. Second, this approach requires that a probability distribution be assigned to the set of models $\Theta_0 \cup \Theta_1$, and in most applications, no such distribution is readily available. For these reasons, this method for selecting the threshold λ will not be pursued.

Because of the difficulties mentioned above in calculating the size of a GLR test, the asymptotic local approximation [9] is commonly used to select thresholds for GLR tests. When the probability measures associated with $\mathcal{H}_0$ and $\mathcal{H}_1$ get closer to each other as the data length L approaches infinity, for a large class of systems the distributions become locally asymptotically normal. Under these conditions, the pdf for the GLR can be well approximated by a local normal distribution. This approximation readily permits computation of the test size α and power β in terms of the threshold λ. Although the local approach is widely used to select threshold levels for GLR tests, the assumption that as L increases, the probability measures associated with the two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ approach each other is not well suited to the performance diagnosis problem considered here. The goal of the detection scheme at hand is to detect changes in the closed loop transfer function which correspond to the system changing from good performance to poor performance, which is quite different from a detection problem wherein one is interested in arbitrary small changes in the system. For the case of performance monitoring, there are many dynamical changes that may occur without violating the performance specifications, and these types of changes should not affect the monitoring scheme. Under the assumptions of the local theory, as the data length L grows, the magnitude of the changes to be detected decreases, and thresholds selected using this theory will result in detection schemes which perform poorly when these assumptions are violated.

Because the available theoretical results for threshold selection cannot be straightforwardly applied to the performance monitoring problem at hand, other approaches must be introduced to set the threshold. In the following Sections, three threshold selection methods are introduced. The first two methods consider the confidence of the parameter estimates to reduce the threshold determination to the selection of a new parameter which carries a more intuitive meaning. The third method considers

a cross-correlation criterion which produces a threshold without requiring selection of a secondary parameter.

3.4.1 Confidence limit approach

As an insightful example, consider the estimation of a MA(t_0) model using a MA(m) structure with $m > t_0$. Applying the methods of [61] or of [94] to estimating the parameters will result in an unbiased estimate. Therefore, in the limit of infinite data, the $m - t_0$ "extra" parameters in the MA(m) structure will be correctly identified as zero; however, for finite data, the variance of these estimates will be non-zero. To determine if these extra parameters could be set to zero, it is useful to compare their estimated values to their variance. More generally, the variance of the parameter estimates serve as a measure of their precision. Since the variance of the estimated parameters can also be estimated from the data, we can use this information for our criterion. In this section and the next, we discuss two possible approaches of incorporating parameter covariance estimates in the decision criterion.

Let $V(\mathcal{Y}, \hat{\theta}_u)$ be as in (3.23), with $\hat{\theta}_u \in \Theta_1$. The GLR method compares the values of the two quantities $V(\mathcal{Y}, \hat{\theta}_0)$ and $V(\mathcal{Y}, \hat{\theta}_u)$. If there were no uncertainty in the models θ_0 and θ_u, it would be reasonable to accept $\mathcal{H}_1$ whenever $V(\mathcal{Y}, \theta_u) \leq V(\mathcal{Y}, \theta_0)$; however, using finite data, the estimates $\hat{\theta}_0$ and $\hat{\theta}_u$ will always have some uncertainty associated with them. Representing this uncertainty by a probability distribution for the estimated $\hat{\theta}_u$, consider the question "Given the distribution for $\hat{\theta}_u$, is there a model θ_u which is not much less probable than $\hat{\theta}_u$ and which would give the same residual square error as $\hat{\theta}_0$ for the data $\mathcal{Y}$?" If such a θ_u exists, then accepting $\mathcal{H}_1$ may be a rash decision. This leads to the following test:

Test 3.4.1 *Is there a model θ_u such that $P(\theta_u) > r^* P(\hat{\theta}_u)$ and $V(\mathcal{Y}, \theta_u) = V(\mathcal{Y}, \hat{\theta}_0)$, where P is the probability of the parameter estimates, and r^* is a specified constant on the interval $(0, 1]$. If so, accept $\mathcal{H}_0$; otherwise, accept $\mathcal{H}_1$.*

Here, r^* determines how much less probable such a θ_u must be before one is willing to accept $\mathcal{H}_1$. Using standard identification methods, the asymptotic estimate $\hat{\theta}_u$

obtained from minimizing V will be normally distributed with mean $\bar{\theta}_u$ and covariance Σ_u. For this distribution, the relative likelihood of two models $\hat{\theta}_u$ and $\hat{\theta}_u + d\theta$ can be calculated:

$$\frac{P(\hat{\theta}_u + d\theta)}{P(\hat{\theta}_u)} = \frac{\exp(\frac{1}{2}(\hat{\theta}_u - \bar{\theta}_u)^T \Sigma_u^{-1}(\hat{\theta}_u - \bar{\theta}_u))}{\exp(\frac{1}{2}(\hat{\theta}_u + d\theta - \bar{\theta}_u)^T \Sigma_u^{-1}(\hat{\theta}_u + d\theta - \bar{\theta}_u))}. \tag{3.28}$$

By assuming $\bar{\theta}_u = \hat{\theta}_u$, the ratio (3.28) becomes:

$$\frac{P(\hat{\theta}_u + d\theta)}{P(\hat{\theta}_u)} = \exp(-\frac{1}{2} d\theta^T \Sigma_u^{-1} d\theta). \tag{3.29}$$

Choosing a value r^*, and letting $\delta^* = -2\log(r^*)$, Test 3.4.1 can be expressed as

$$\max_{d\theta^T \Sigma_u^{-1} d\theta \leq \delta^*} V(\mathcal{Y}, \hat{\theta}_u + d\theta) \gtrless V(\mathcal{Y}, \hat{\theta}_0), \tag{3.30}$$

where the poor $\mathcal{H}_1$ is accepted if $<$ holds, and $\mathcal{H}_0$ if $>$ holds. To apply this test, we could either solve the optimization problem (3.30), or we could consider an equivalent test,

$$\delta^* \gtrless \delta, \quad \delta = \min \hat{\delta} \text{ s.t. } \max_{d\theta^T \Sigma_u^{-1} d\theta \leq \hat{\delta}} V(\mathcal{Y}, \hat{\theta}_u + d\theta) = V(\mathcal{Y}, \hat{\theta}_0). \tag{3.31}$$

In general, both of these optimization problems may be very difficult to solve; however, a simple approximation can be made to the latter using that fact that $\hat{\theta}_u$ is an unconstrained optimal point of $V(\mathcal{Y}, \theta)$. Consider the second order expansion of $V(\mathcal{Y}, \theta)$ about $V(\mathcal{Y}, \hat{\theta}_u)$:

$$V(\mathcal{Y}, \hat{\theta}_u + d\theta) = V(\mathcal{Y}, \hat{\theta}_u) + V_\theta(\mathcal{Y}, \hat{\theta}_u)^T d\theta + \frac{1}{2} d\theta^T V_{\theta\theta}(\mathcal{Y}, \hat{\theta}_u) d\theta + O(||d\theta||^3), \tag{3.32}$$

where $V_\theta(\mathcal{Y}, \hat{\theta}_u)$ and $V_{\theta\theta}(\mathcal{Y}, \hat{\theta}_u)$ are the gradient and Hessian of $V(\mathcal{Y}, \theta)$ respectively evaluated at $\hat{\theta}_u$. Since the matrix $V_{\theta\theta}(\mathcal{Y}, \hat{\theta}_u)$ is usually calculated in the optimization routine to find $\hat{\theta}_u$, it is readily available, and also $V_\theta(\mathcal{Y}, \hat{\theta}_u) = 0$ since $\hat{\theta}_u$ is a minimum of V. Using this expansion, we see that for $d\theta$ small, the maximization in (3.30) can

be approximated by

$$\begin{aligned}
\max_{d\theta^T\Sigma_u^{-1}d\theta\leq\delta^*} V(\mathcal{Y},\hat{\theta}_u+d\theta)-V(\mathcal{Y},\hat{\theta}_u) &= \max_{d\theta^T\Sigma_u^{-1}d\theta\leq\delta^*}\frac{1}{2}d\theta^T V_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)d\theta+O(||d\theta||^3),\\
&= \max_{x^Tx\leq\delta^*}\frac{1}{2}x^T JV_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)Jx+O(||d\theta||^3),\\
&= \frac{1}{2}\delta^*\bar{\lambda}(JV_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)J)+O(||d\theta||^3), \qquad (3.33)
\end{aligned}$$

where $J^2=\Sigma_u$ and $\bar{\lambda}$ denotes the largest eigenvalue. The matrix J is guaranteed to exist since Σ_u is symmetric and positive definite. An approximate of the minimal δ^* as in (3.31) can easily be obtained from this solution:

$$\begin{aligned}
\delta &= 2\frac{V(\mathcal{Y},\hat{\theta}_0)-V(\mathcal{Y},\hat{\theta}_u)}{\bar{\lambda}(JV_{\theta\theta}\mathcal{Y},\hat{\theta}_uJ)}+\frac{O(||d\theta||^3)}{\bar{\lambda}(JV_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)J)},\\
&\approx 2\frac{V(\mathcal{Y},\hat{\theta}_0)-V(\mathcal{Y},\hat{\theta}_u)}{\bar{\lambda}(JV_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)J)}. \qquad (3.34)
\end{aligned}$$

If $\delta>\delta^*$, we accept hypothesis $\mathcal{H}_1$. Rearranging (3.34) results in a threshold for the GLR test,

$$\left(\frac{V(\mathcal{Y},\hat{\theta}_u)}{V(\mathcal{Y},\hat{\theta}_0)}\right)^{\frac{L}{2}} \gtrless \left(1-\frac{\delta^*\bar{\lambda}(JV_{\theta\theta}(\mathcal{Y},\hat{\theta}_u)J)}{2V(\mathcal{Y},\hat{\theta}_0)}\right)^{\frac{L}{2}}=T_1(r^*) \qquad (3.35)$$

In summary, evaluating decision function g_λ in (3.15) using the threshold $\lambda=T_1(r^*)$ consists of the following steps:

Application of Test 3.4.1:

1. Calculate $\hat{\theta}_u$. If $A\phi_{\hat{\theta}_u}\leq b$, accept $\mathcal{H}_0$; otherwise, go to step 2.
2. Estimate the covariance matrix Σ_u of $\hat{\theta}_u$. Calculate J such that $J^2=\Sigma_u$.
3. Calculate $\hat{\theta}_0=\sup_{\theta\in\Theta_0}V(\mathcal{Y},\theta)$.
4. Choose a confidence limit r^* and the associated value $\delta^*=-2\log r^*$, and calculate $\lambda=T_1(r^*)$.
5. Evaluate g_λ.

3.4.2 Constraint softening approach

In the previous section, the covariance matrix Σ_u for the unconstrained parameter estimates $\hat{\theta}_u$ was used to determine a GLR threshold λ for the decision function g_λ in (3.15). An alternative approach would be to set $\lambda = 1$, and relax the constraints until $\hat{\theta}_u \in \Theta_0$, in which case $g_1 = \mathcal{H}_0$. Then, the amount by which the constraints must be relaxed can be compared against the covariance of the parameter estimates to decide whether acceptance of $\mathcal{H}_0$ is justified. To motivate this approach, consider again the example of estimating the parameters for a MA(t_0) time series using a MA(m) model, $m > t_0$. Using the MA(m) structure, an unbiased fit can be obtained; however, for finite data, the probability that the extra $m - t_0$ coefficients are exactly zero is infinitesimally small. By contrast, the probability that these parameters lie within a standard deviation of zero will be significant.

Let $\sigma_k = \sqrt{\Sigma_u(k,k)}$. If performance constraints were all of the form $\phi_k < b_k$, softening the constraints so that $\phi_{\hat{\theta}_u,k}$ lies within $\bar{\delta}^*$ standard deviations of the constraint would result in a constraint of the form:

$$\phi_k < b_k + \bar{\delta}^* \sigma_k. \tag{3.36}$$

This constraint softening approach is depicted in Figure 3.2. The performance constraints in this figure corresponds to $\phi_k < 0.1$ for $k \geq 6$. The model $\hat{\theta}_u$ clearly violates the constraints, but each parameter ϕ_6 through ϕ_{24} lies within a standard deviation of satisfying the performance specifications.

More generally, the following test may be applied:

Test 3.4.2 *Does there exist a θ which satisfies the performance objective and is such that each parameter θ_k is within $\bar{\delta}^*$ standard deviations of the optimal estimate $\hat{\theta}_{u,k}$? If so, accept $\mathcal{H}_0$; otherwise, accept $\mathcal{H}_1$.*

For linear constraints as in (3.7) and MA(t^*) models, Test 3.4.2 may be carried out by replacing $\phi_{\hat{\theta}_u,k}$ in the constraint equations by $\phi_{\hat{\theta}_u,k} + d\phi_k \sigma_k$ and solving the feasibility

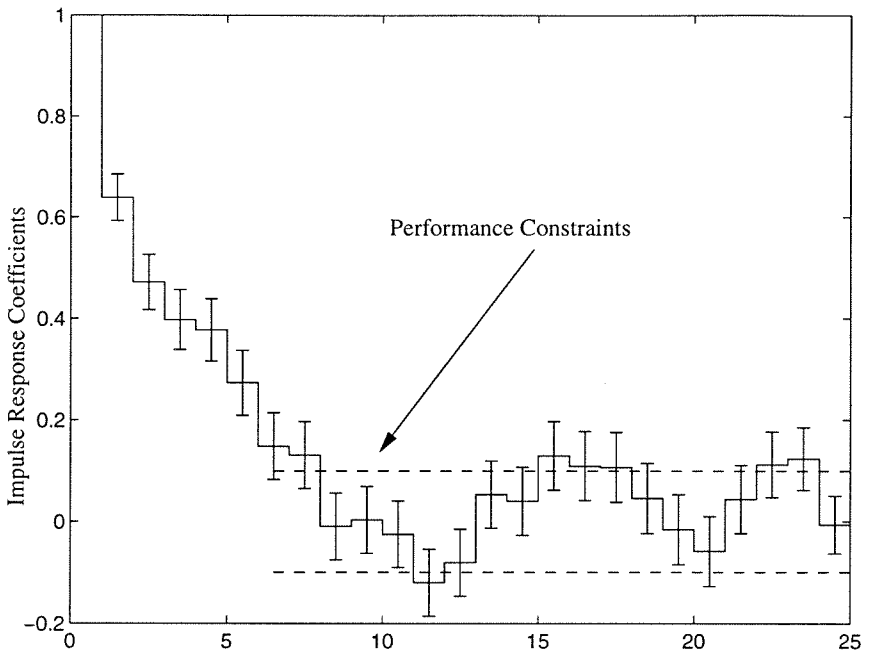

Figure 3.2: Illustration of constraint softening approach

problem: Does there exist a vector $d\phi = [d\phi_1, \ldots, d\phi_{t^*}]$ such that

$$A\phi_{\hat{\theta}_u} + Ad\phi \leq b, \ |d\phi_k| \leq \bar{\delta}^* \sigma_k. \tag{3.37}$$

This feasibility problem can easily be solved via the following linear program:

LP 1

$$\begin{aligned} \min_{d\phi} \quad & \bar{\delta}_2 \\ \text{s.t.} \quad & A(\phi + d\phi) \leq b \\ & |d\phi_k| \leq \bar{\delta}_2 \sigma_k \end{aligned}$$

The problem is feasible if and only if $\bar{\delta}_2 \leq \bar{\delta}^*$. The Test 3.4.2 can therefore be applied by carrying out the following steps:

Application of Test 3.4.2:

1. Calculate $\hat{\theta}_u$. If $A\phi_{\hat{\theta}_u} \leq b$, accept $\mathcal{H}_0$; otherwise, go to step 2.

2. Estimate the covariance matrix Σ_u of $\hat{\theta}_u$.

3. Solve the linear programming problem LP 1 for $\bar{\delta}_2$.

4. If $\bar{\delta}_2 > \bar{\delta}^*$, then choose $\mathcal{H}_1$; otherwise, choose $\mathcal{H}_0$.

We make a brief comparison between this method and the one of the previous section. Each test measures the distance in the parameter space from the unconstrained optimum of V to another point. The first method uses a weighted 2-norm, where the weight is the full covariance matrix, whereas the second method measures the distance using an ∞-norm, and only the diagonal elements of the covariance are used for weighting. When the covariance matrix is diagonal, these two measurements are equivalent in the following sense. Let $D_2\Theta(\delta)$ and $D_\infty\Theta(\delta)$ be defined in the following way:

$$\begin{aligned} D_2\Theta(\delta) &= \{d\theta \mid d\theta^T \Sigma_u^{-1} d\theta \leq \delta\}, \\ D_\infty\Theta(\delta) &= \{d\theta \mid |d\theta_k| \leq \delta\sigma_k\}, \end{aligned}$$

then for any δ_0, there exists δ_1 and δ_2 such that

$$D_2\Theta(\delta_1) \subset D_\infty\Theta(\delta_0) \subset D_2\Theta(\delta_2). \tag{3.38}$$

However, when Σ_u is not diagonal, such an equivalence does not exist, and points which are "near" $\hat{\theta}_u$ with the 2-norm measure may not be near with the ∞-norm measure. This will especially be true in the case where Σ_u is ill-conditioned. In this case, incorporating the directionality of Σ_u into the linear program may result in a smaller value of $\bar{\delta}$. Since Σ_u is positive definite and symmetric, it can be factored as $\Sigma_u = V^T D^2 V$, with V a unitary matrix and D diagonal. Using this factorization, the set $D_2\Theta$ can be expressed as

$$D_2\Theta(\delta) = \{d\theta \mid \|D^{-1} V d\theta\|_2^2 \leq \delta\}. \tag{3.39}$$

From this expression, the infinity norm measure which is equivalent to the norm used

in defining $D_2\Theta$ can easily be gleaned, $\|D^{-1}Vd\theta\|_\infty$. Minimizing this norm subject to the performance constraints results in the linear program LP 2:

LP 2

$$\begin{aligned} \min_{d\phi} \quad & \bar{\delta}_2' \\ \text{s.t.} \quad & A(\phi + d\phi) \leq b \\ & x = D^{-1}Vd\phi \\ & |x_k| \leq \bar{\delta}_2' \end{aligned}$$

3.4.3 Cross-validation approach

The GLR test (3.15) with $\lambda \leq 1$ compares the residual square error for the two models $\hat{\theta}_0$ and $\hat{\theta}_u$. Unless by chance $\hat{\theta}_u \in \Theta_0$, $\hat{\theta}_u$ will always produce a smaller residual square error than $\hat{\theta}_0$ since the optimization in the former case is performed over the larger set $\Theta_u = \Theta_0 \cup \Theta_1$. If the extra degrees of freedom of the set Θ_u actually allow a better description of the system, then the model set Θ_0 is inferior to Θ_u in describing the closed loop system. To determine whether added degrees of freedom in Θ_u are justified by the data, one could use $\hat{\theta}_u$ to calculate residuals from a different data set which was not used in the search for $\hat{\theta}_u$. This approach is commonly used for selecting model structures for process identification, and is referred to in the literature as *Cross-Validation*. As the proposed performance evaluation essentially consists of choosing between two model structures, Θ_0 and Θ_u, cross-validation tools can be used to test the hypothesis. In this section, we briefly review the general approach and develop a convenient estimation technique which is valid for model structures containing constraints. We then show how the cross-validation approach can give a threshold for the GLR test.

The idea of using cross-validation for model selection dates to [99]. The basic concept can be described as follows. Given a sequence of data and two separate model structures, first divide the data into two segments. Use one segment to obtain the best possible fit for the chosen structure, and then use the second segment to test

the model. The structure which produces the smallest error on the test sequence is chosen.

A generalization of this approach has been presented by [98]. Let $I = \{1, 2, \ldots L\}$ be an index corresponding to a data sample. Partition the data into k segments I_p such that

$$I_p = \{(p-1)m + 1, \ldots, pm\}. \tag{3.40}$$

We assume here that $L = km$. When this assumption does not hold, we simply let I_k contain more than m but less than $2m$ points. Also define the following model performance criteria:

$$V(\theta) = \frac{1}{L}\sum_{t\in I} \epsilon^2(t,\theta),\ V^p(\theta) = \frac{1}{L}\sum_{t\in I-I_p} \epsilon^2(t,\theta), \tag{3.41}$$

where for notational simplicity, we have dropped $\mathcal{Y}$ from the argument of V, although it is still implied that V depends on the data. Cross-validation involves first finding $\hat{\theta}_p \in \Theta$ which produces the smallest residual error for $t \in I - I_p$, and then using the residual errors for $t \in I_p$ which were not included in the optimization objective to evaluate the model structure Θ. When this procedure is repeated for each of the data partitions I_p, the cross-validation criterion can be mathematically expressed by:

$$C_I(\Theta) = \sum_{p=1}^{k}\sum_{t\in I_p} \epsilon^2(t,\hat{\theta}_p),\ \hat{\theta}_p = \arg\min_{\theta\in\Theta} V^p(\theta). \tag{3.42}$$

This measure can be applied to performance evaluations in a straightforward fashion by comparing the values of $C_I(\Theta_0)$ and $C_I(\Theta_u)$, resulting in the following test:

Test 3.4.3 *If $C_I(\Theta_0) < C_I(\Theta_u)$, then accept $\mathcal{H}_0$; otherwise accept $\mathcal{H}_1$.*

Although C_I could be calculated exactly, for large k this direct approach would be computationally expensive. Stoica et al. derive an asymptotic approximation for C_I which is much easier to compute for the case where the model structures do not contain constraints on the parameters. Under a relatively mild assumption, a similar approximation can be made for constrained model structures. Let us consider the

following assumption:

Assumption 3.4.1 *Let Θ be a set whose elements satisfy linear constraints of the form (3.7), and let $\hat{\theta}_p$ be as in (3.42). For each $p = 1, \ldots, k$, the active constraints of $\hat{\theta}_p$ are the same as the active constraints of $\hat{\theta}$, where $\hat{\theta}$ minimizes $V(\theta)$.*

The active constraints can easily be determined by calculating the vector $A\phi_\theta - b$ and finding the elements which are zero to within machine precision. Under the above assumption, the following theorem holds:

Theorem Let $\hat{\theta} = \arg\min_{\theta\in\Theta} V(\theta)$, and let J index the set of active constraints for $\hat{\theta}$. Then

$$\frac{1}{L}C_I(\Theta) = q_m(\Theta) + O\left(\frac{1}{k^2 m}\right), \tag{3.43}$$

where

$$\begin{aligned}
q_m(\Theta) &= V(\hat{\theta}) + \frac{4}{L^2}\sum_{p=1}^{k} w_p(\hat{\theta})^T W_J(\hat{\theta}) w_p(\hat{\theta}), \\
w_p(\hat{\theta}) &= \sum_{t\in I_p} \epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta}), \\
W_J(\hat{\theta}) &= \left(V_{\theta\theta}(\hat{\theta})^T \mathcal{N}_{A_J} V_{\theta\theta}(\hat{\theta}) + A_J^T A_J\right)^{-1} V_{\theta\theta}(\hat{\theta})^T \mathcal{N}_{A_J}, \\
\mathcal{N}_{A_J} &= (I - A_J^T (A_J A_J^T)^{-1} A_J),
\end{aligned}$$

and A_J is a matrix containing the rows of A indexed by J. When no constraints are active, $W_J = V_{\theta\theta}(\hat{\theta})^{-1}$.

Proof: (follows similar proof in [98]) First, consider the Kuhn-Tucker conditions for an optimum in $V(\theta)$ subject to linear constraints as in (3.7):

$$V_\theta(\hat{\theta}) + A_J^T \mu_J = 0. \tag{3.44}$$

This implies that $V_\theta(\hat{\theta})$ is in the range of A_J^T. As the range of A_J^T equals the null space of $\mathcal{N}_{A_J}$, this condition could be written as $\mathcal{N}_{A_J} V_\theta(\hat{\theta}) = 0$, or equivalently as $V_\theta(\hat{\theta})^T \mathcal{N}_{A_J} = 0$.

For k sufficiently large, $\hat{\theta}_p$ is close to $\hat{\theta}$, and we write

$$\epsilon^2(t,\hat{\theta}_p) = \epsilon^2(t,\hat{\theta}) + 2\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta})^T(\hat{\theta}_p - \hat{\theta}) + O(||\hat{\theta}_p - \hat{\theta}||^2). \tag{3.45}$$

Using the assumption that the active constraints for $\hat{\theta}_p$ are the same as for $\hat{\theta}$,

$$\begin{aligned}
0 &= \mathcal{N}_{A_J} V_\theta^p(\hat{\theta}_p), \\
&= \mathcal{N}_{A_J}\left(V_\theta^p(\hat{\theta}) + V_{\theta\theta}^p(\hat{\theta})(\hat{\theta}_p - \hat{\theta}) + O(||\hat{\theta}_p - \hat{\theta}||^2)\right), \\
&= \mathcal{N}_{A_J}\left(V_\theta(\hat{\theta}) - \frac{2}{L}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta})\right. \\
&\quad \left. + \left\{V_{\theta\theta}(\hat{\theta}) - \frac{m}{L}\left(\frac{\partial}{\partial\theta}\frac{2}{m}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta})\right)\Big|_{\theta=\hat{\theta}}\right\}(\hat{\theta}_p - \hat{\theta}) + O(||\hat{\theta}_p - \hat{\theta}||^2)\right), \\
&= \mathcal{N}_{A_J}\left(-\frac{2}{L}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta}) + \left[V_{\theta\theta}(\hat{\theta}) + O\left(\frac{1}{k}\right)\right](\hat{\theta}_p - \hat{\theta}) + O(||\hat{\theta}_p - \hat{\theta}||^2)\right)
\end{aligned} \tag{3.46}$$

To evaluate the order of $\mathcal{N}_{A_J}\left(\frac{1}{m}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta})\right)$, note the following:

$$\begin{aligned}
\mathcal{N}_{A_J}\left(\frac{1}{m}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta})\right) &= \mathcal{N}_{A_J}\left(E(\epsilon(\hat{\theta})\epsilon_\theta(\hat{\theta})) + O(1/m)\right), \\
&= \mathcal{N}_{A_J}\left(\frac{1}{L}\sum_{t\in I}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta}) + O(1/\sqrt{L}) + O(1/\sqrt{m})\right), \\
&= \mathcal{N}_{A_J}V_\theta(\hat{\theta}) + O(1/\sqrt{m}), \\
&= O(1/\sqrt{m}).
\end{aligned} \tag{3.47}$$

It then follows that $||\hat{\theta}_p - \hat{\theta}|| = O(1/(k\sqrt{m}))$. Therefore, we get the relation

$$\mathcal{N}_{A_J}V_{\theta\theta}(\hat{\theta})(\hat{\theta}_p - \hat{\theta}) = \mathcal{N}_{A_J}\left(\frac{2}{L}\sum_{t\in I_p}\epsilon(t,\hat{\theta})\epsilon_\theta(t,\hat{\theta}) + O(\frac{1}{k^2\sqrt{m}})\right). \tag{3.48}$$

Combined with the constraints $A_J(\hat{\theta}_p - \hat{\theta}) = 0$, we get the following system of equa-

tions for $(\hat{\theta}_p - \hat{\theta})$:

$$\begin{bmatrix} \mathcal{N}_{A_J} V_{\theta\theta}(\hat{\theta}) \\ A_J \end{bmatrix} (\hat{\theta}_p - \hat{\theta}) = \begin{bmatrix} \mathcal{N}_{A_J} \left(\frac{2}{L} w_p(\hat{\theta}) + O(\frac{1}{k^2\sqrt{m}}) \right) \\ 0 \end{bmatrix}, \tag{3.49}$$

which has the unique solution

$$\hat{\theta}_p - \hat{\theta} = \frac{2}{L} W_J(\hat{\theta}) w_p + O\left(\frac{1}{k^2\sqrt{m}}\right). \tag{3.50}$$

Next, noting that $\mathcal{N}_{A_J}(\hat{\theta}_p - \hat{\theta}) = \hat{\theta}_p - \hat{\theta}$, and substituting (3.50) into (3.45) gives

$$\begin{aligned} \frac{1}{L} C_I(\Theta) &= \frac{1}{L} \sum_{p=1}^{k} \sum_{t \in I_p} \left\{ \epsilon^2(t,\hat{\theta}) + 2\epsilon(t,\hat{\theta}) \epsilon_\theta(t,\hat{\theta})^T \mathcal{N}_{A_J} \left(W_J(\hat{\theta}) \frac{2}{L} w_p(\hat{\theta}) + O\left(\frac{1}{k^2\sqrt{m}}\right) \right) \right\} \\ &\quad + O\left(\frac{1}{k^2 m}\right), \\ &= V(\hat{\theta}) + \frac{4}{L^2} \sum_{p=1}^{k} w_p(\hat{\theta})^T W_J(\hat{\theta}) w_p(\hat{\theta}) \\ &\quad + \frac{m}{L} \sum_{p=1}^{k} \left(\mathcal{N}_{A_J} \frac{2}{m} \sum_{t \in I_p} \epsilon(t,\hat{\theta}) \epsilon_\theta(t,\hat{\theta}) \right)^T \cdot O\left(\frac{1}{k^2\sqrt{m}}\right) + O\left(\frac{1}{k^2 m}\right), \\ &= V(\hat{\theta}) + \frac{4}{L^2} \sum_{p=1}^{k} w_p(\hat{\theta})^T W_J(\hat{\theta}) w_p(\hat{\theta}) + O\left(\frac{1}{k^2 m}\right). \end{aligned} \tag{3.51}$$

By applying the approximate cross-correlation criterion $q_m(\Theta)$ to the model structures Θ_u and Θ_0, we can determine if the set Θ_u fits the data $\mathcal{Y}_L(t)$ better than Θ_0. If $q_m(\Theta_u) < q_m(\Theta_0)$, then the performance bounds imposed by Θ_0 are too restrictive. If the inequality is switched, acceptable performance is concluded. This approach can also be interpreted as the selection of a threshold for testing the GLR. The relation $q_m(\Theta_u) < q_m(\Theta_0)$ is equivalent to

$$\begin{aligned} GLR &= \left(\frac{V(\mathcal{Y}, \hat{\theta}_u)}{V(\mathcal{Y}, \hat{\theta}_0)} \right)^{\frac{L}{2}} \gtrless T_3, \\ T_3 &\equiv \left(1 - \frac{\sum_{p=1}^{k} \left(w_p(\hat{\theta}_0)^T W_J(\hat{\theta}_0) w_p(\hat{\theta}_0) - w_p(\hat{\theta}_u)^T V_{\theta\theta}(\hat{\theta}_u)^{-1} w_p(\hat{\theta}_u) \right)}{V(\mathcal{Y}, \hat{\theta}_0)} \right)^{\frac{L}{2}} \end{aligned} \tag{3.52}$$

.

In summary, a hypothesis test incorporating a cross-validation can be applied through the following steps:

Application of Test 3.4.3:

1. Calculate $\hat{\theta}_u = \sup_{\theta \in \Theta_u} V(\mathcal{Y}, \theta)$. If $\hat{\theta}_u \in \Theta_0$, accept $\mathcal{H}_0$; otherwise, go to step 2.
2. Calculate $\hat{\theta}_0 = \sup_{\theta \in \Theta_0} V(\mathcal{Y}, \theta)$ and determine the active constraints for $\hat{\theta}_0$.
3. Calculate threshold T_3 as in (3.52).
4. Evaluate the decision function g_λ in (3.15) using the threshold T_3.

3.4.4 Comparison of threshold selection methods

In this Section, the threshold selection approaches of Tests 3.4.1, 3.4.2, and 3.4.3 are compared. A brief comparison of the norms used in Tests 3.4.1 and 3.4.2 was made above. Besides the different norms used in the two approaches, another important distinction exists between the two methods. Whereas the latter measures the distance from the unconstrained optimum to the set of good performance models, the former measures the distance to the nearest point which lies on the same contour of $V(\mathcal{Y}, \theta)$ as $V(\mathcal{Y}, \hat{\theta}_0)$. Thus Test 3.4.1 gives a point $\hat{\theta}_u + d\theta$ which has the same likelihood as $\hat{\theta}_0$, but may not satisfy the constraints, and Test 3.4.2 gives a point $\hat{\theta}_u + d\theta$ which satisfies the constraints, but may be less likely than $V(\mathcal{Y}, \hat{\theta}_0)$. This is illustrated in Figure 3.3. The point A represents the unconstrained minimum $\hat{\theta}_u$ of $V(\theta)$, and B the constrained minimum $\hat{\theta}_0$. The set $\{\theta | \theta = \hat{\theta}_u + d\theta, d\theta^T \Sigma^{-1} d\theta < \delta^*\}$ is contained within the heavy ellipse, and the set $\{\theta | \theta = \hat{\theta}_u + d\theta, |d\theta_k / \sigma_k| < \delta_2\}$ is the interior of the heavy rectangle. The other curves represent the equicost curves of $V(\theta)$. The latter criterion measures the distance from A to C, the nearest feasible point, and the former criterion measures the distance from A to D, the nearest point with the same cost as the point B.

Although different philosophies lie behind the distances measured in Tests 3.4.1 and 3.4.2, it is probable that both measures will give similar diagnoses, and the examples of Section 3.6 support this claim. As no theoretical results are available,

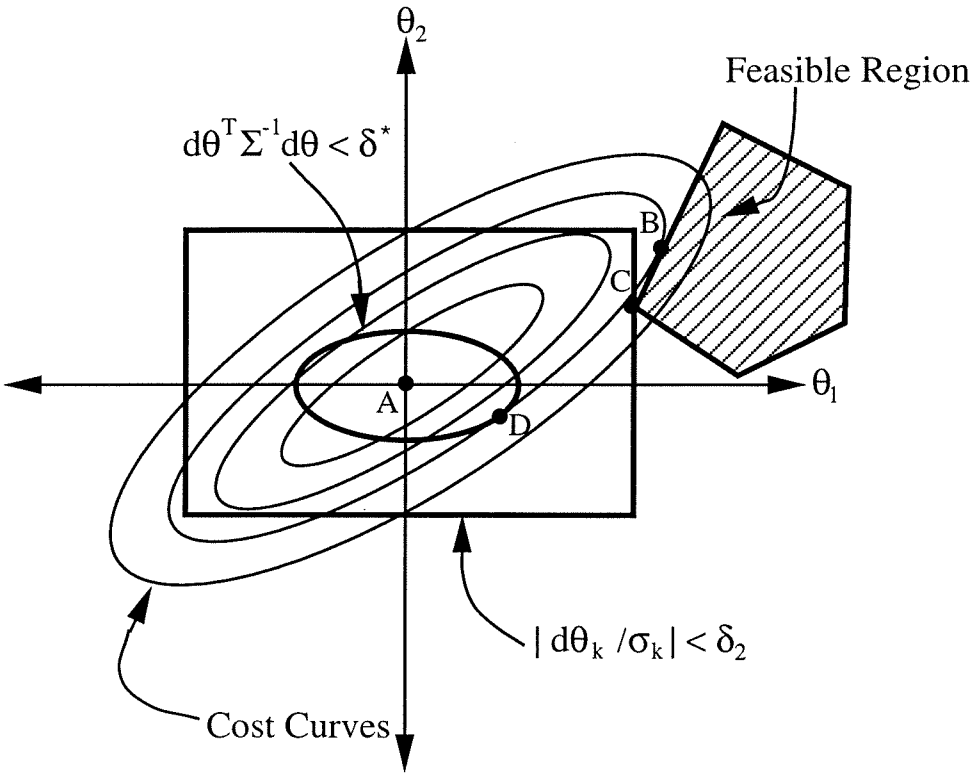

Figure 3.3: Graphic interpretation of Tests 3.4.1 and 3.4.2

further empirical comparisons between the two tests should be made with industrial data.

From a computational standpoint, a significant difference exists between the two Tests 3.4.1 and 3.4.3 and Test 3.4.2. To implement the former, both an unconstrained and a constrained optimization of $V(\mathcal{Y}, \theta)$ must be solved along with other algebraic computations, whereas for the latter, the constrained optimization of $V(\mathcal{Y}, \theta)$ is replaced by a linear program. In most cases, solution of LP 1 or LP 2 will be more efficient than the constrained optimization of V.

As noted in Section 3.3, it is desirable that given a threshold, performance criteria such as the false alarm rate can be evaluated; however, this requires knowledge of the distribution of closed loop transfer functions, which is generally unknown. For this reason, the threshold selection methods presented above were derived so as to have an intuitive meaning. For the method of Section 3.4.1, the threshold obtained depends upon the selection of a new threshold, r^*; however, for many practical applications, the

interpretation of r^* will be more straightforward as it is related to the uncertainty in the estimated model $\hat{\theta}_u$. Similarly, the parameter $\bar{\delta}^*$ required for the method Section 3.4.2 also carries an intuitive interpretation as the number of standard deviations about the estimated parameters which must be considered before the constraints are satisfied. Although the goal in deriving the criteria of Sections 3.4.1 and 3.4.2 was to provide intuitive insight into threshold selection, in many cases the parameters r^* or $\bar{\delta}^*$ will necessarily need to be tuned on line to give the desired detection properties.

The goal of performance monitoring is to identify control loops which may need to be retuned. Therefore, one approach to selecting r^* and $\bar{\delta}^*$ is to simply calculate the values they must take for the good performance hypothesis $\mathcal{H}_0$ to be accepted. In this manner, the control loops needing retuning can be prioritized. The smaller the value of r^* (or larger $\bar{\delta}^*$) which is needed to accept $\mathcal{H}_0$, the higher the priority should be set to retuning the controller.

Unlike the other two methods, using the cross-validation method of Section 3.4.3 does not require the selection of any secondary thresholds. From this standpoint, its application is more straightforward, and when selection of the levels r^* and $\bar{\delta}^*$ is not intuitive, this method will be preferred.

3.5 Extensions to systems with command signals

The methods developed in the previous sections focus on evaluating the ability of a control system to reject unknown, unmeasured disturbances. We implicitly assumed that no changes in set point occur within the data sequence being tested; however, operating data will often contain set point changes. We could apply the methods of this chapter to the signal $y - r$, where r is the reference signal, but this approach has disadvantages. For the GLR method, one essentially seeks to find the most likely unknown sequence ϵ that produces y via some model within the sets Θ_u and Θ_0. When the set point changes are large compared to the disturbances, a large value of ϵ will be needed to produce $y - r$ when no information on r is included, and this will typically bias the estimation of the underlying closed loop model.

In order to include reference signal data in the performance evaluation scheme, we consider two separate cases. The first corresponds to the case wherein the reference signal excites the system sufficiently for identification purposes, whereas in the second case, it does not. In most circumstances where routine operating data are used, the reference signal will excite the system poorly, and the second approach will be more applicable.

Consider a system as in Figure 3.1, with the modification that instead of y, the controller C is driven by $y - r$. The closed loop transfer function becomes:

$$y - r = \frac{W}{1 + PC} e - \frac{1}{1 + PC} r. \tag{3.53}$$

Previously, we have considered evaluating performance by estimating the transfer function $\frac{W}{1+PC}$ using two model structures, one of which has been constrained to correspond to acceptable performance. When the reference signal r is persistently exciting, a similar approach can be taken by specifying a set Θ_0^r which contains all transfer functions $\frac{1}{1+PC}$ which corresponds to acceptable performance levels. Similarly, a set Θ_1^r, analogous to Θ_1, is specified, and a generalized likelihood approach test may be developed using the ratio:

$$\frac{P(\mathcal{Y}_L(t)|\mathcal{R}_L(t), \theta \in \Theta_0, \theta^r \in \Theta_0^r)}{P(\mathcal{Y}_L(t)|\mathcal{R}_L(t), \theta \in \Theta_1, \theta^r \in \Theta_1^r)}, \tag{3.54}$$

where $\mathcal{R}_L(t)$ is a vector analogous to $\mathcal{Y}_L(t)$ containing the history of the reference signal. Once again, by choosing a threshold value less than one, the optimization over Θ_1^r can be replaced by a simpler unconstrained optimization over $\Theta_u^r = \Theta_0^r \cup \Theta_1^r$.

When the reference signal r is not persistently exciting, this approach will work poorly because the unconstrained structure Θ_u^r will tend to over fit the data, resulting in a small GLR. This may result in a poor performance conclusion, when the system is in fact performing satisfactorily. Consider, for example, the special case where the data $\mathcal{R}_L(t)$ contain only one step change in the reference which takes place at time $t_0 \in [t - L + 1, t]$. If Θ_1^r is an FIR(m) model, then the data can be made consistent

with $\epsilon(t^*) = 0, t_0 \leq t^* \leq t_0 + m$, whereas the constrained fit will not be able to make ϵ zero on this range. Since in most practical cases, the signal r will not be sufficiently exciting, another approach is needed.

If the plant and controller were known exactly, one could easily calculate the signal $\frac{1}{1+PC}r$, and then use the previous likelihood approach on the signal z defined by:

$$z = y - r + \frac{1}{1+PC}r = \frac{W}{1+PC}e. \tag{3.55}$$

Since we are interested in monitoring changes in the performance of the system which could have possibly been caused by changes in P or C, it is unrealistic to assume models are known, in which case z will not be independent of r. Nevertheless, when the performance has not deteriorated, the transfer function $\frac{1}{1+PC}$ will be approximately as designed, and dependence on z of r will be less significant than if the term $\frac{1}{1+PC}r$ were not included. Therefore, when the performance is satisfactory, using z as the input to the monitoring scheme should result in similar diagnosis as in the case where r is constant. On the other hand, when performance has degraded and the transfer function $\frac{1}{1+PC}$ is no longer accurate, both the estimation of $\hat{\theta}_u$ and $\hat{\theta}_0$ are affected by the changes in r, and the constraints of Θ_0 will generally result in larger residuals ϵ than will be needed for the optimal model in Θ_u. Therefore, changes in r will tend to make the GLR smaller, and thus favor a poor performance diagnosis. In the case where the reference signal is not sufficiently exciting, applying the methods of Sections 3.3 and 3.4 to the signal z in (3.55) is recommended.

3.6 Examples

In this section, we consider examples which serve to demonstrate the advantages of the methods of this chapter over minimum-variance estimation methods. Example 1 addresses disturbance rejection for a simple system in which the process parameters as well as the disturbance characteristics may vary. In Example 2, the performance of a distillation column controller is evaluated.

Example 3 Consider a simple model of a stable system, described by the transfer function

$$P = z^{-2}\frac{5}{4}\frac{1-0.6z^{-1}}{1-0.5z^{-1}}. \tag{3.56}$$

Using the IMC design method of [74], a controller is designed to reject step disturbances. The controller has the form:

$$C = \frac{Q}{1-PQ},\ Q = \frac{4}{5}\frac{(1-f)}{1-fz^{-1}}\frac{1-0.5z^{-1}}{1-0.6z^{-1}},\ 0<f<1. \tag{3.57}$$

Increasing the parameter f detunes the controller while increasing robustness to model uncertainty. In this example, $f = 0.3$, which corresponds to tight control which will be nearly minimum variance when there is no plant/model mismatch *and* the disturbances are accurately modeled as integrated white noise.

In order to use the likelihood method to determine system performance, meaningful constraints must be specified. The controller was designed to provide good rejection of step like disturbances. For the nominal plant, the closed loop step response to a unit step disturbance settles to 0.027 within four time samples, suggesting the performance specification:

$$|\phi_k| \leq 0.1 \text{ for } k \geq 4. \tag{3.58}$$

If the disturbance transfer function W were known a priori to be equal to $\frac{1}{1-z^{-1}}$, these constraints would be satisfactory. However, in practice W is not known, and may possibly have a large high frequency component. Assuming we are only interested in the rejection of low frequency disturbances, we should use the filtered coefficient constraint method of Section 3.2. A twentieth order (length 21) FIR filter was designed so that it had approximately unit amplitude in the range $0 \leq \omega \leq 0.4\pi$ and zero amplitude in the range $0.6\pi \leq \omega \leq \pi$. For systems with bandwidth less than 0.4π, filtering the impulse response coefficients with this filter will result in a sequence approximately equal to the impulse response delayed by ten samples. Because the impulse response coefficients correspond to the product of W and $\frac{1}{1+PC}$, including

Table 3.1: Nominal and Perturbed Plants

Plant	TF	OL gain	M_1	M_2	Diagnosis
P_1	$\frac{5}{4}z^{-2}\frac{1-0.6z^{-1}}{1-0.5z^{-1}}$	1	0.027	0.063	GOOD
P_2	$\frac{5}{7}z^{-2}\frac{1-0.3z^{-1}}{1-0.5z^{-1}}$	1	0.163	0.2197	POOR
P_3	$2z^{-2}\frac{1-0.75z^{-1}}{1-0.5z^{-1}}$	1	0.332	0.355	POOR
P_4	$\frac{7}{4}z^{-2}\frac{1-0.6z^{-1}}{1-0.3z^{-1}}$	1	0.226	0.235	POOR
P_5	$\frac{7}{5}z^{-2}\frac{1-0.75z^{-1}}{1-0.65z^{-1}}$	1	0.039	0.125	GOOD
P_6	$\frac{1}{2}z^{-2}\frac{1-0.6z^{-1}}{1-0.8z^{-1}}$	1	0.317	0.373	POOR
P_7	$\frac{7}{8}z^{-2}\frac{1-0.6z^{-1}}{1-0.5z^{-1}}$	0.7	0.216	0.212	POOR
P_8	$\frac{9}{8}z^{-2}\frac{1-0.6z^{-1}}{1-0.5z^{-1}}$	0.9	0.08	0.08	GOOD

the filter implies that the performance criterion depends only on the response of the system to a disturbance of the form FWe rather than We. Combining the settling time criterion (3.58) and a low pass filter gives the following performance constraints:

$$|\phi_k^F| < 0.1\ \phi_{10}^F \text{ for } k \geq 14, \tag{3.59}$$

where $\phi_k^F = F_k + \sum_{i=1}^{k} F_{k-i}\phi_i$. The set Θ_0 is given by all models of the form $\theta_0 = [1, \phi_1, \ldots, \phi_{20}]$, with ϕ satisfying the above filtered constraints.

Eight different plants were used in the simulations, as shown in Table 3.1. The table also shows the open loop steady state gain, the two measures

$$M_1 = \max_{k\geq 4} |h_k|, \quad M_2 = \max_{k\geq 4} \frac{|h_{k+10}^F|}{h_{10}^F}, \tag{3.60}$$

where h and h^F are respectively the unfiltered and filtered closed loop step response coefficients, and the proper diagnosis based on $M_1 \leq 0.1$. The plant P_1 corresponds to the nominal system. The open loop and closed loop step responses for the plants are shown in Figures 3.4 through 3.7.

For each of the plants P_1 through P_8, the closed loop system was simulated using two separate disturbances. The first disturbance was generated by $d_1 = \frac{1}{1-z^{-1}}w$, with w a white noise process. The second disturbance d_2 was obtained by adding to d_1 the sequence $\frac{1-r\cos(\omega)z^{-1}}{1-2r\cos(\omega)z^{-1}+r^2z^{-2}}w$ with $\omega = \frac{2}{3}\pi$, $r = 0.95$, and w the same sequence which

Table 3.2: Diagnostics, disturbance d_1

Plant	MV ratio	GLR	T_1	T_3	$\bar{\delta}_2$	$\bar{\delta}'_2$	κ^*
P_1	1.04	0.959	0.736	0.439	0.26	0.15	176
P_2	1.17	0.155	0.734	0.0978	1.5	1.00	153
P_3	1.07	2.46e-05	0.75	0.0965	3.8	2.8	243
P_4	1.03	0.0142	0.746	0.602	2.4	1.6	256
P_5	1.02	0.8	0.735	0.473	0.60	0.38	161
P_6	1.34	5.44e-06	0.745	0.137	3.9	2.8	220
P_7	1.15	0.0293	0.74	0.289	2.6	1.4	198
P_8	1.07	0.707	0.737	0.451	0.7	0.4	182

generated d_1. The disturbance d_2 has a significant high frequency component. Both disturbances are shown in Figure 3.8.

For each plant/disturbance pair, the GLR was computed, along with the thresholds T_1 (Eq. (3.35)) for $r^* = 0.5$ and T_3 (Eq. (3.52)) for $m = 1$. In addition, $\bar{\delta}_2$ and $\bar{\delta}'_2$ were calculated from the linear programming problems LP 1 and LP 2. For the unconstrained estimation of $\hat{\theta}_u$, standard Newton-Rapson methods such as outlined by [94] were used, and the constrained optimization of $\hat{\theta}_0$ was calculated using a quadratic penalty function [10]. Tables 3.2 and 3.3 show the decision criteria, along with the minimum variance (MV) ratio, $||\mathcal{Y}||/\sigma_{mv}$, where σ_{mv} was estimated using the method outlined in [43]. In addition, the Tables show the scaled condition number κ^* for the covariance matrix Σ_u, where κ^* is calculated by

$$\kappa^* = \min_{D_1, D_2} \kappa(D_1 \Sigma_u D_2), \tag{3.61}$$

with D_1 and D_2 diagonal matrices, and κ the usual condition number.

Although Σ_u was poorly conditioned in each example, the measure $\bar{\delta}_2$ from LP 1 gives similar diagnosis as comparing the GLR to the threshold T_1. The measure $\bar{\delta}'_2$ from LP 2 is smaller than $\bar{\delta}_2$, implying that by using the eigenvectors of Σ_u as coordinates, the perturbation which must be added to $\hat{\theta}_u$ so that the sum lies within Θ_0 is smaller than when the standard coordinate system is used. For the plants P_1, P_3, P_4, P_5, P_6, and P_7, the GLR test gives correct performance diagnosis with either

Table 3.3: Diagnostics, disturbance d_2

Plant	MV ratio	GLR	T_1	T_3	$\bar{\delta}_2$	$\bar{\delta}'_2$	κ^*
P_1	1.9	0.931	0.599	0.426	0.20	0.16	314
P_2	1.71	0.38	0.555	0.156	0.671	0.56	155
P_3	2.16	2.34e-05	0.632	0.375	2.2	2.5	773
P_4	2.22	0.00615	0.625	0.526	1.5	1.4	697
P_5	1.92	0.837	0.609	0.473	0.30	0.30	345
P_6	1.74	0.000542	0.55	0.0881	2.0	1.6	134
P_7	1.83	0.0217	0.574	0.546	1.6	1.1	235
P_8	1.87	0.602	0.59	0.438	0.55	0.41	287

threshold T_1 or T_3, and for both disturbances d_1 and d_2. For the plant P_2, a correct diagnostic conclusion will be drawn when the threshold T_1 is used, but not when the threshold T_3 is used. The performance deterioration for P_2 is the least for all of the poor performing models, so incorrect diagnosis in this case is less severe than it would be in the others. For plant P_8, all cases give correct diagnosis except when the threshold T_1 is used and the disturbance corresponds to d_1; however, the threshold T_1 depends upon the tuning parameter r^* or δ^*, and choosing a smaller value of r^* will give a correct diagnosis. That the GLR is less than the threshold T_1 approximately implies that for parameters θ to have generated $\mathcal{Y}$ with the same likelihood as as $\hat{\theta}_0$, θ must be less than half as probable as $\hat{\theta}_u$.

If the measures $\bar{\delta}_2$ or $\bar{\delta}_2$ were used to prioritize the control loops, plants P_6 and P_3 would receive high priority, followed by P_7, P_4, and with plants P_2, P_8, P_5, and P_1 receiving the lowest priority. From Figures 3.4 through 3.7, it is clear that this prioritization would be an efficient approach to re-designing the controllers.

Conclusions are more difficult to draw from the MV ratio. Consider first the case of disturbance d_1. For the plants with good performance, P_5, and P_8, the MV ratio are within 3% of the nominal case, P_1; however, two of the poor performance plants, P_3 and P_4, also have MV ratios within 3% of the nominal case, whereas the GLR clearly indicates a deterioration in performance. In addition, although P_2 has the least performance deterioration in terms of the settling criterion, its MV ratio is the largest of any of the plants. In each case, this improper diagnosis is not due to poor

estimation of the MV ratio, but rather to the fact that the MV ratio does not provide a good performance measure. In the case of the disturbance d_2, for each plant the MV ratio is significantly larger than when the simulation was carried out with d_1. Additionally, drawing conclusions about which of the models is performing worse, where performance is measured by the settling time criterion, is not possible from the MV ratio.

We would now like to consider four separate scenarios, and use the results in Tables 3.2 and 3.3 to evaluate how the GLR methods and the MV method would diagnose changes in process conditions. The scenarios will be:

1. Shift from P_1 to P_4 or P_6, constant disturbance spectrum d_1;
2. Shift from P_1 to P_4 or P_6, constant disturbance spectrum d_2;
3. Shift in disturbance spectrum from d_1 to d_2, constant plant P_1;
4. Shift in disturbance spectrum from d_2 to d_1, shift in plant from P_1 to P_4 or P_6.

Scenario 1 For a shift to P_4, the MV ratio test would indicate a 1% improvement in controller performance, incorrectly suggesting no retuning necessary, whereas for P_6, a 28% deterioration in controller performance would be noted. When the GLR test is applied, a clear deterioration for both cases P_4 and P_6 is correctly diagnosed.

Scenario 2 Here, the opposite situation occurs. Using only the MV ratio, when the plant shifts to P_4, a correct deterioration in control is concluded (increased MV ratio 1.9 to 2.22), but a shift to P_6 gives an incorrect diagnosis (decreased MV ratio 1.9 to 1.74). Again, the GLR test gives a clear poor performance diagnosis in each case.

Scenario 3 When the plant remains unchanged, but the disturbance spectrum changes from d_1 to d_2, the MV ratio indicates a substantial decrease in performance. Although the theoretical MV ratio actually does increase, the increased variance is due to the high frequency component of the disturbance d_2. Retuning the controller

would require that the model accurately describe the behavior of the system at high frequency. Because the high frequency modes of the system may be difficult to identify, retuning may not be desirable. On the other hand, the GLR test, through the filtered coefficient constraints, considers only the low frequency phenomena, and indicates that performance remains satisfactory.

Scenario 4 If the disturbance changes from d_2 to d_1, the MV ratio test would indicate substantial improvements in performance, regardless of any changes in the plant. In particular, if the plant were to shift from P_1 to either P_4 or P_6 at the same time as the disturbance dynamics change, the MV ratio diagnosis would be incorrect. On the other hand, the GLR would correctly diagnose that the performance has deteriorated.

Comparing the threshold T_1 in Tables 3.2 and 3.3, we see that lower values of the threshold are obtained in case of d_2. This is expected because although in the case of disturbance d_1, the closed loop system can be very accurately modeled by 20 impulse response coefficients, for disturbance d_2, which contains a slowly decaying, high frequency oscillation, more than 20 coefficients are needed. Since the unconstrained structure Θ_u is a MA(20) model, the true system with disturbance d_2 is not contained within Θ_u, and the estimated covariance Σ_u will be larger than in the case of d_1, resulting in the smaller threshold T_1 observed in Table 3.3.

Example 4 This example addresses the control of overhead temperature for a distillation column. The data used have been made publicly available by Shell Research Company, and a full description of the system may be found in [52], wherein it is referred to as *Column 2*. Although the complete data set contained measurements for 80,000 time samples, only the smaller segment of 1000 points shown in Figure 3.9 is analyzed in this example. From visual inspection, it is clear that a significant change in the overall process occurred after the first 500 samples.

The performance constraints were specified as in the previous example, with the settling time set to 10, which represented the open loop settling time for a step response, as indicated by the step response supplied with the data. Θ_0 and Θ_1

were respectively the set of all MA(25) models which satisfied the constraints and is complement in the set of all MA(25) models. Table 3.4 contains the results of the various diagnostic tests. The threshold T_1 was calculated for both $r^* = 0.5$ and $r^* = 0.25$, and T_3 for $m = 1$. The MV ratio indicates only a slight decrease in the performance of segment 1 compared to segment 2, suggesting that the primary source of the poor quality output is due to an increase in the energy of the disturbances. On the other hand, the GLR tests indicate that for the first segment, the performance is borderline acceptable, whereas for the second test, it is clearly below the acceptable limit.

The data also contain measurements of the feed temperature for the column input. Although this measurement was not used in calculating the control moves, it can provide some insight to the proper diagnosis. The power spectra for the feed temperature for each of the two data segments are shown in Figure 3.10. For the segment from $t \in [501, 1000]$, the energy of this disturbance is significantly higher (approximately 60 times). This accounts for much of the increased energy in the overhead temperature. In addition to the increased energy, the second segment also has a large low frequency component at approximately $\omega = 0.1\pi$. Similarly, the unconstrained model $\hat{\theta}_u$ fit to this segment has a low frequency oscillation which does not quickly decay. Since the constrained model $\hat{\theta}_0$ must be small after 10 time samples, $\hat{\theta}_u$ fits the data much better than $\hat{\theta}_0$ and the GLR is small. Thus the GLR test correctly diagnoses that the change in the signal is not only due to increased disturbance energy but also to a change in controller performance due to the disturbance spectrum. The decision whether to retune the controller should depend on several other factors, which may include the time required to retune and knowledge about the transient behavior of the disturbance. For example, if experience or other process knowledge indicated that the disturbance is likely to have the same characteristics for a considerable time, retuning the controller will be profitable.

Table 3.4: Diagnostics, Shell Column 2

Sequence	MV ratio	GLR	$T_1(0.5)/T_1(0.25)$	T_3	δ_2	$\bar{\delta}_2'$	κ^*
$1-500$	1.16	0.46	0.66/0.44	0.44	1.06	0.64	69
$501-1000$	1.30	3.6e-28	0.61/0.38	>1	5.7	5.0	1102

3.7 Conclusions

In this chapter, we have shown how many common and practical controller performance criteria can be expressed as linear constraints on the closed loop impulse response coefficients. Using this type of criterion, performance monitoring can be formulated as a generalized likelihood ratio test. Evaluating the GLR involves solving a constrained as well as an unconstrained model identification problem. In order to evaluate performance, the GLR must be compared to a threshold.

Three approaches to selecting the threshold have been discussed. The first two methods use the covariance of the estimated unconstrained model. The method of Section 3.4.1 can be interpreted as calculating the maximum relative probability that the system is described by a model with the same likelihood as the optimal constrained model, given the unconstrained estimate is distributed normally with mean $\hat{\theta}_u$ and covariance Σ_u. The method of Section 3.4.2 tests to see if a set of parameters exists, with each individual parameter within some confidence limit of its optimal estimate, which satisfies the performance bounds. Alternatively, cross-correlations between the parameter estimates can also be incorporated using a modified linear program. These methods require the selection of secondary thresholds, r^* and $\bar{\delta}^*$, which carry intuitive meanings. As an alternative to using a set value of r^* or $\bar{\delta}^*$, the maximum value of r^* or minimum value of $\bar{\delta}^*$ which would be necessary to accept $\mathcal{H}_0$ can be calculated and used to prioritize control loops targeted for retuning. Finally, in Section 3.4.3 a threshold selection with a cross-validation interpretation is derived.

When the operating data used to calculate the performance measure contain set point changes, and the reference signal excites the system sufficiently, the GLR methods can be extended directly by specifying performance bounds for reference tracking,

and parameterizing the transfer function from the reference signal to the tracking error. More frequently, changes in the set point will not produce a sufficiently exciting signal, but by subtracting from the tracking error the nominal response to the reference signal, the frequency of false poor performance diagnoses will be decreased.

The examples of Section 3.6 showed that for a meaningful performance objective, the likelihood ratio methods gave correct diagnoses whereas the minimum variance ratio of [43] did not. In particular, we demonstrated through Example 1 that the MV ratio test can be strongly influenced by changes in the high frequency component of a disturbance which may not merit retuning, and may at the same time be insensitive to model changes which do result in deteriorated controller performance and can be easily rectified by retuning the controller. On the other hand, by properly selecting the performance objective, the GLR test can be made insensitive to irrelevant changes in the disturbance dynamics, while maintaining high sensitivity to model changes. Finally, for the examples considered, the measures $\bar{\delta}_2$ or $\bar{\delta}_2'$, while easier to compute, gave diagnoses consistent with those obtained by comparing the GLR to the thresholds T_1 and T_3.

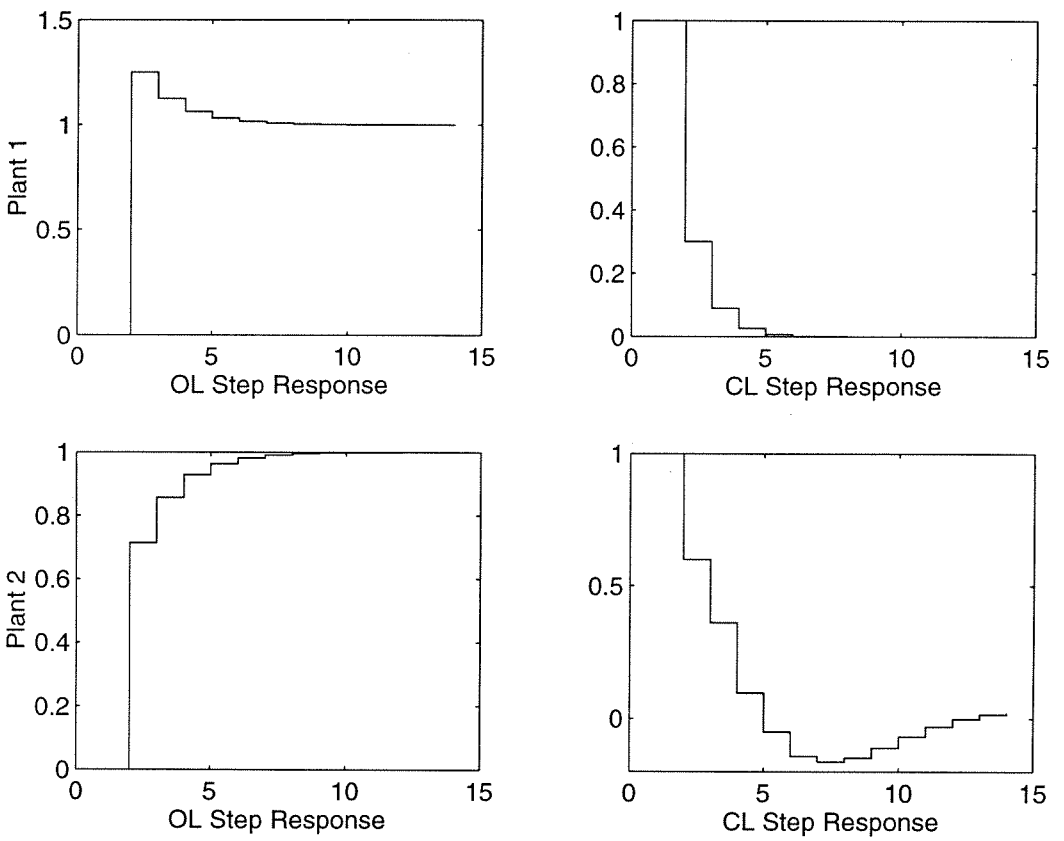

Figure 3.4: Open (OL) and closed loop (CL) step responses for Plants 1 and 2

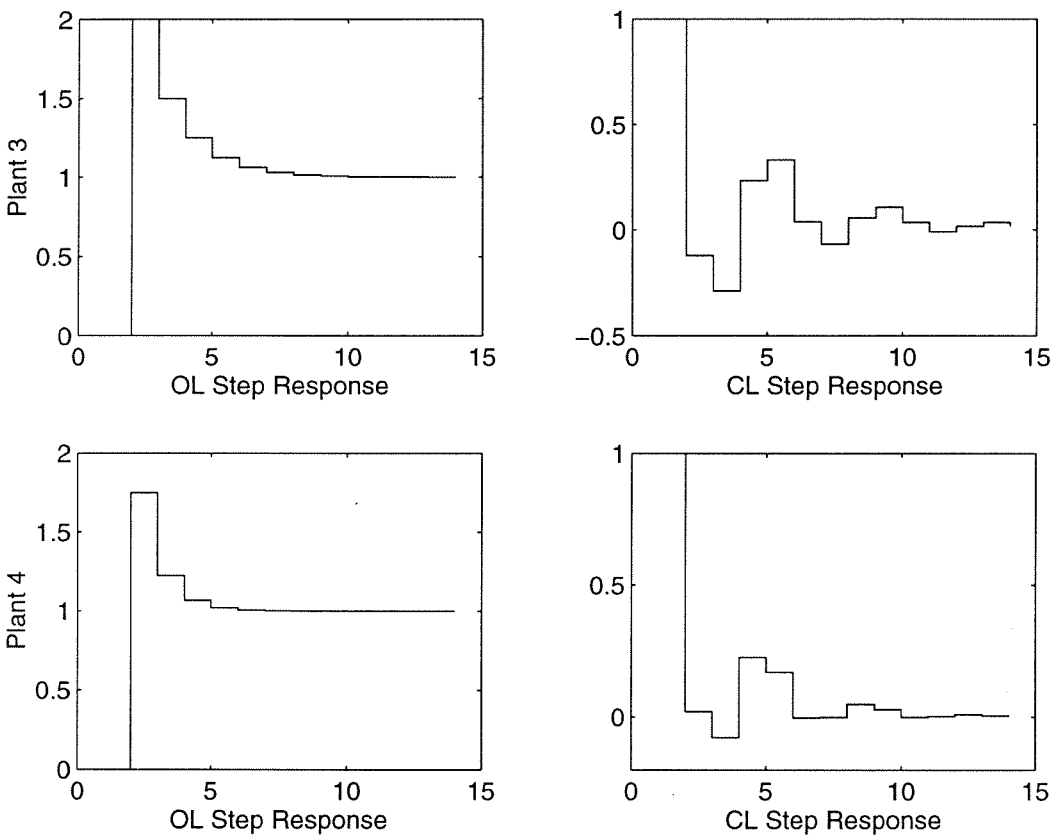

Figure 3.5: Open (OL) and closed loop (CL) step responses for Plants 3 and 4

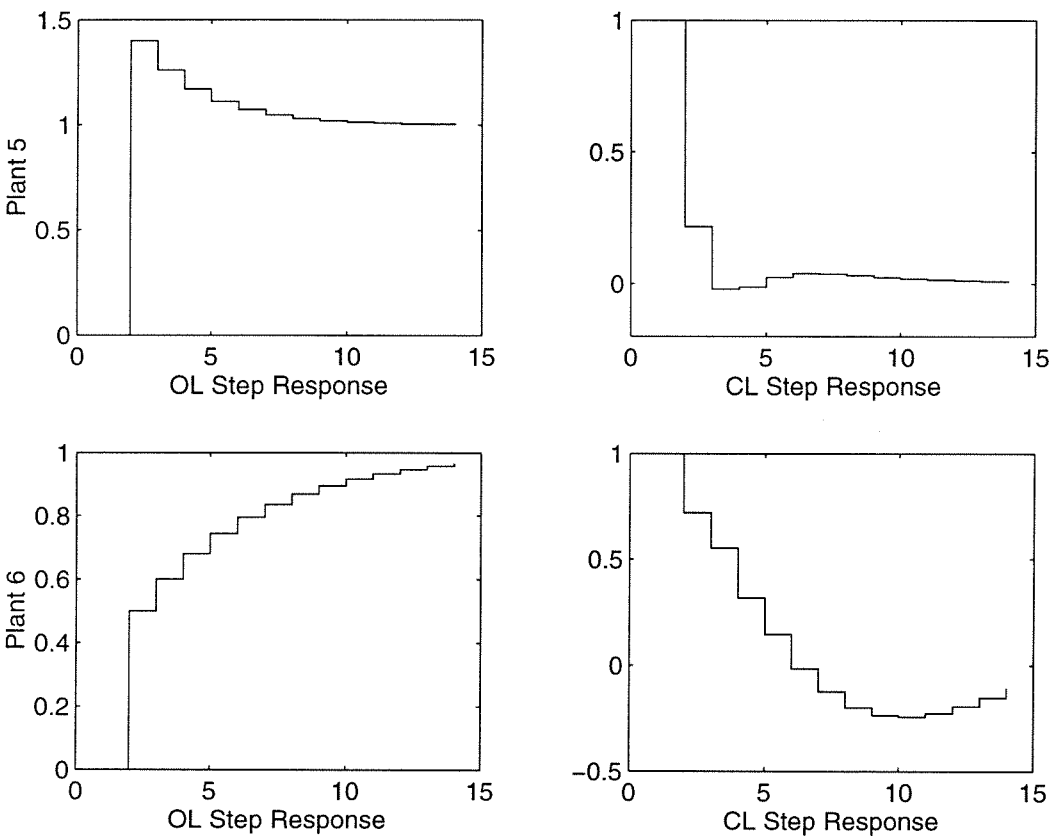

Figure 3.6: Open (OL) and closed loop (CL) step responses for Plants 5 and 6

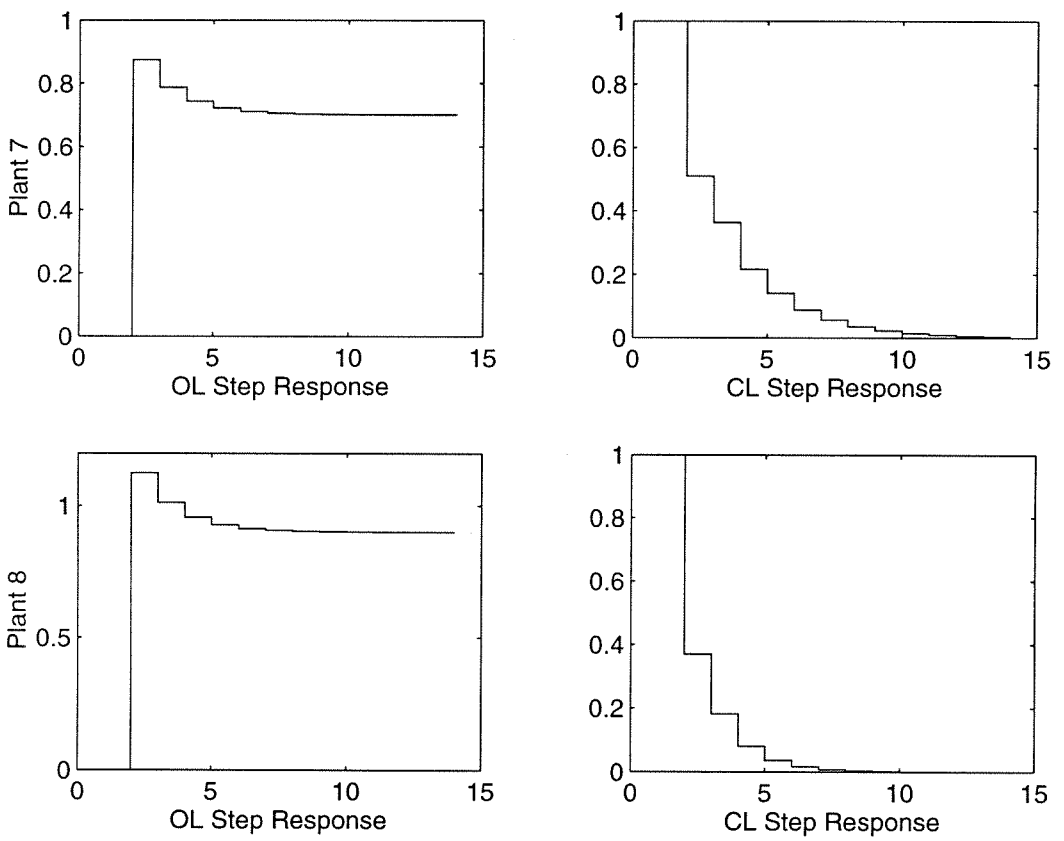

Figure 3.7: Open (OL) and closed loop (CL) step responses for Plants 7 and 8

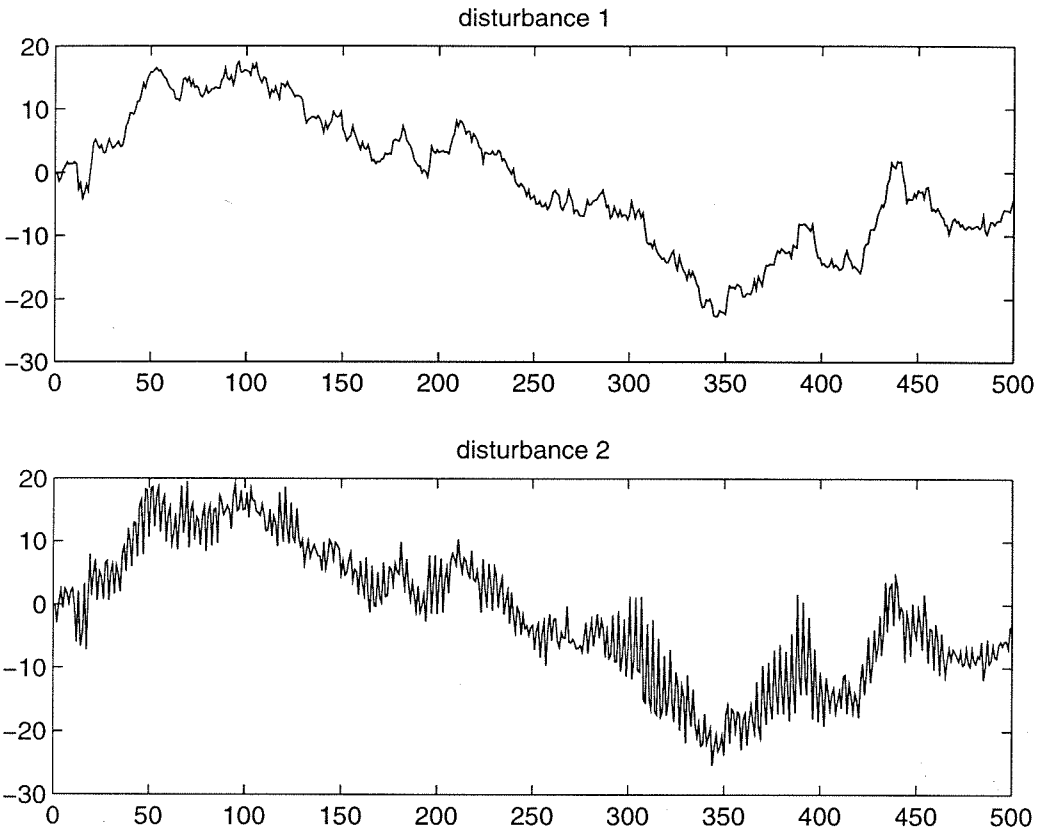

Figure 3.8: Disturbances

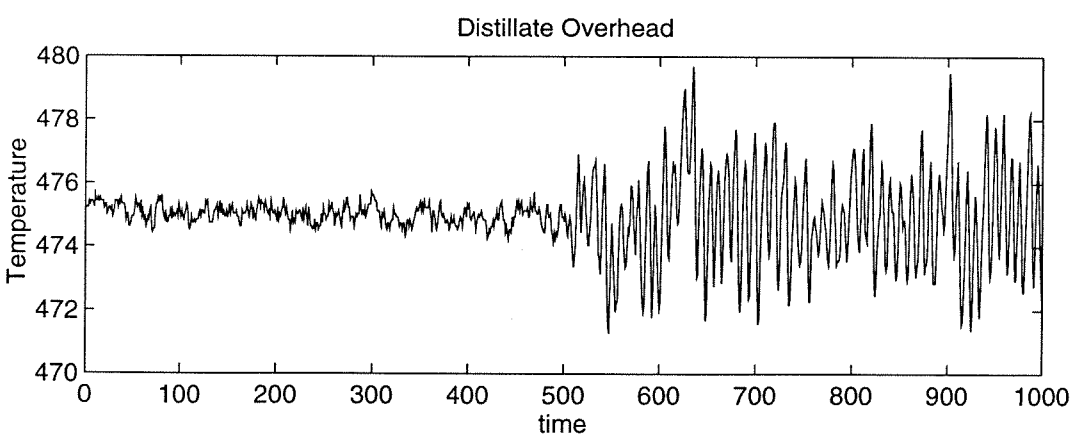

Figure 3.9: Shell Column 2, Overhead Temperature

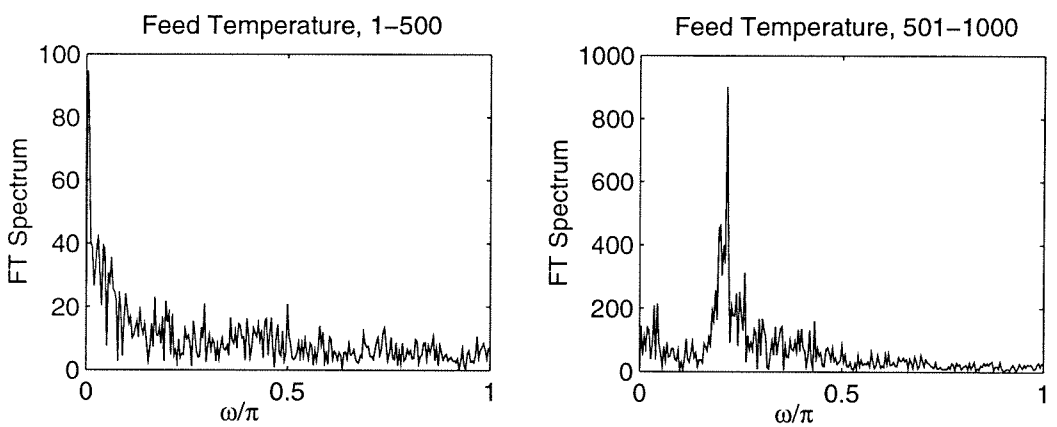

Figure 3.10: Shell Column 2, Feed Temperature Spectra

Part III

Fault Detection

Chapter 4 Integrated design of control and detection filters

Summary

The problem of designing an integrated control and diagnostic module is considered. The four degree of freedom controller is recast into a general framework wherein results from optimal and robust control theory can be easily implemented. For the case of an $\mathcal{H}_2$ objective, it is shown that the optimal control-diagnostic module involves constructing an optimal controller, closing the loop with this controller, and then designing an optimal diagnostic module for the closed loop. When uncertain plants are involved, this two-step method does not lead to reasonable diagnostics, and the control and diagnostic modules must be synthesized simultaneously. An example shows how this design can be accomplished with available methods.

4.1 Introduction

One of the earliest approaches to fault detection centered on the design of filters which when applied to system measurements would produce an output signal which deviates from zero only when the system was malfunctioning. Considerable attention has been focused on the design of such filters since the seminal work of Beard [11] and Jones [50]. The survey paper [110] provides an early summary of work in this area, and [34] provides a more recent account. Although the majority of results for model based fault detection are based solely upon nominal models, a few significant studies which incorporate uncertain models have appeared [35, 62, 18]. However, the methods developed in these references are unable to incorporate the norm bounded uncertainty descriptions commonly employed in control synthesis. In addition, for

feedback systems, these methods do not consider the interaction between detection and control, i.e., the detection algorithms consider the effect of known inputs on the outputs, but do not consider the effect of the output on the input.

Nett [78] introduced the four degree of freedom, or four parameter, controller which integrates control and diagnostics. By parameterizing all stabilizing controllers of this form, he was able to elucidate tradeoffs involving both controller performance and diagnostic performance. In addition to diagnostic tradeoffs such as between the detection of sensor and actuator faults, Nett et al. [79] have shown that when uncertain models are used, control performance must be traded off against diagnostic performance. This work suggests that both control and diagnostic modules must be designed together.

Although several years have passed since Nett parameterized the four degree of freedom controller, it has not found widespread use due to the fact that a systematic synthesis method which guarantees robust control and diagnostic performance in the face of uncertainty does not exist. Notable attempts at outlining synthesis methods can be found in [79] and [32]. The former proposes a four step procedure in which each degree of freedom is considered successively. The latter employs a method which first designs a nominal control module with robust stability in an l_∞ framework, and then designs the diagnostic module on the resultant closed loop.

In this chapter, we show how the four degree of freedom controller is simply a special case of the general interconnection structure used in modern control synthesis methods. Viewed in this framework, we employ results from $\mathcal{H}_2$-, $\mathcal{H}_\infty$-, and μ-synthesis methods to design integrated control and diagnostic systems. Through an example, we show that for uncertain systems, robust performance of both control and diagnostics requires that these two modules be designed simultaneously. The results of this chapter suggest that for control systems with inherent uncertainty and tight performance specifications, linear fault detection filters will perform poorly.

4.2 Four degree of freedom controller

Consider a general multi-input, multi-output (MIMO) system G. Partition the inputs to G into uncontrolled inputs d, and controlled inputs u. The signal d represents noises, disturbances, and faults. Also partition the outputs as those outputs used by the control algorithm (y) from those not used (z), yielding the following partition for G:

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} d \\ u \end{bmatrix}. \tag{4.1}$$

The four degree of freedom controller has access to a reference signal r as well as the measurement y, and returns not only the control moves u, but also a diagnostic alarm signal a:

$$\begin{bmatrix} a \\ u \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} r \\ y \end{bmatrix}. \tag{4.2}$$

Let the input d be described by $d = [f^T, n^T]^T$, where f represents faults, and n represents noises and disturbances which occur under normal operation. Nett et al. [79] have elucidated several algebraic tradeoffs involved with fault detection which we will not repeat, except to comment that, in general, one must tradeoff the ability to detect input or actuator faults from output or sensor faults.

Of more interest for our discussion is the tradeoff between diagnostics and control. For example, consider the block diagram shown in Figure 4.1. Suppose we desire that the alarm signal a tracks input faults f_i. The alarm signal a is given by

$$a = T_{ar} r + T_{af_o} \left(f_o + n_o + G_{22}(f_i + n_i) \right)$$

where the perturbed transfer functions T_{ar} and T_{af_o} are given by

$$T_{ar} = K_{11} + K_{12}(I - S_o G_{22} \Delta_u W_u K_{22})^{-1} S_o (G_{22} + G_{22} \Delta_u W_u) K_{21},$$

$$T_{af_o} = (I - S_o G_{22} \Delta_u W_u K_{22})^{-1} S_o,$$

and $S_o = (I - GK_{22})^{-1}$ is the output sensitivity function. For nominal systems, K can easily be designed so that $T_{ar} = 0$ [78]; however, as the above expression for this transfer function suggests, model uncertainty prohibits such a design, and the objective of making T_{ar} robustly small will compete with disturbance rejection and reference tracking objectives.

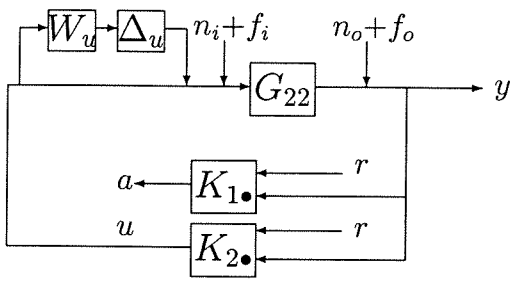

Figure 4.1: Control and diagnostic configuration for uncertain system

4.3 General framework

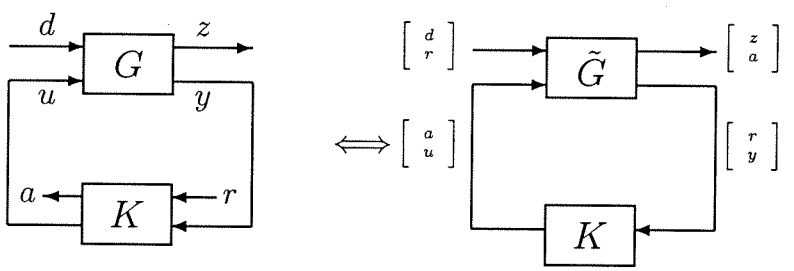

Figure 4.2: Equivalent four parameter controller configurations

Consider the plant and four parameter controller as in (4.1) and (4.2). As written, the reference signal r is not an input to G, and the output alarm a is not an output of G. We would like to obtain a reconfigured system $\tilde{G}$ so that all the inputs are inputs to $\tilde{G}$, all the outputs are outputs of $\tilde{G}$, and the closed loop systems are equivalent, as shown in Figure 4.2. This can easily be accomplished with the following augmented

plant:

$$\tilde{G} = \begin{bmatrix} G_{11} & 0 & 0 & G_{12} \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ G_{21} & 0 & 0 & G_{22} \end{bmatrix}. \tag{4.3}$$

We may therefore view the four degree of freedom controller as a special case of the more general interconnection input-output system. Several advantages arise by viewing the control-diagnostic system in this light. First, the parameterization of stabilizing controllers can be viewed as a simple application of the Youla parameterization, as shown in Section 4.4. Second, optimal control results, such as those found in [30], may be applied in a straightforward way. Finally, robust analysis and synthesis of the control-diagnostic module may be carried out in a systematic fashion using readily available tools such as those available in [5].

4.4 Four-parameter control parameterization

By using the augmented $\tilde{G}$, the parameterization of stabilizing controllers is a straightforward application of the Youla parameterization. In order to analyze stability, we need only consider the bottom right partition of $\tilde{G}$,

$$\tilde{G}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & G_{22} \end{bmatrix}. \tag{4.4}$$

Suppose a doubly coprime factorization of G_{22} is given by

$$G_{22} = N_G D_G^{-1} = \tilde{D}_G^{-1} \tilde{N}_G, \tag{4.5}$$

where

$$\begin{bmatrix} V_G & U_G \\ -\tilde{N}_G & \tilde{D}_G \end{bmatrix} = \begin{bmatrix} D_G & -\tilde{U}_G \\ N_G & \tilde{V}_G \end{bmatrix}^{-1}. \tag{4.6}$$

Then a doubly coprime factorization of $\tilde{G}_{22}$ is given by

$$M_{\tilde{G}} = \begin{bmatrix} 0 & 0 \\ 0 & M_G \end{bmatrix}, \; M \in \{N, \; U, \; \tilde{N}, \; \tilde{U}\},$$

$$M_{\tilde{G}} = \begin{bmatrix} I & 0 \\ 0 & M_G \end{bmatrix}, \; M \in \{D, \; V, \; \tilde{D}, \; \tilde{V}\}. \tag{4.7}$$

The parameterization of all stabilizing controllers can be found in [105] and is given by:

$$K = \left(V_{\tilde{G}} + Q\tilde{N}_{\tilde{G}}\right)^{-1}\left(-U_{\tilde{G}} + Q\tilde{D}_{\tilde{G}}\right), \tag{4.8}$$

where Q is any stable transfer function. By partitioning Q and substituting the coprime factorization of $\tilde{G}_{22}$ given in (4.7) into (4.8), K can be parameterized in terms of the factorization of G_{22}, resulting in

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix},$$

where

$$\begin{aligned} K_{11} &= Q_{11} - Q_{12}\tilde{N}_G\left(V_G + Q_{22}\tilde{N}_G\right)^{-1}Q_{21}, \\ K_{12} &= Q_{12}\tilde{D}_G - Q_{12}\tilde{N}_G\left(V_G + Q_{22}\tilde{N}_G\right)^{-1}(Q_{22}\tilde{D}_G - U_G), \\ K_{21} &= \left(V_G + Q_{22}\tilde{N}_G\right)^{-1}Q_{21}, \\ K_{22} &= \left(V_G + Q_{22}\tilde{N}_G\right)^{-1}(Q_{22}\tilde{D}_G - U_G). \end{aligned} \tag{4.9}$$

Using algebraic manipulations and properties of the doubly coprime factorization, one can show

$$K_{12} = Q_{12}(N_G Q_{22} + \tilde{V}_G)^{-1}, \tag{4.10}$$

in which case K has the same form as in [78].

4.5 $\mathcal{H}_2$ optimal control and diagnostics

In this section, we consider the problem of designing an $\mathcal{H}_2$ optimal controller which provides a nominally small output and also tracks faults. Let $d = [f^T, n^T, r^T]^T$, and $y = [\hat{y}^T, r^T]^T$. Let the transfer function which maps $[d^T, u^T]$ onto $[z^T, y^T]$ be given by:

$$G_c = \left[\begin{array}{c|cc} A_c & B_{c1} & B_{c2} \\ \hline C_{c1} & 0 & D_{c12} \\ C_{c2} & D_{c21} & 0 \end{array}\right] \tag{4.11}$$

The following assumptions simplify the analysis, and relaxing these assumptions do not change the nature of the results:

1. $D_{c12}^* C_{c1} = 0$.

2. $D_{c21} D_{c21}^* = I$, $D_{c12}^* D_{c12} = I$.

Assumption 1 implies that the objective function has no cross terms of the form $< x, u >$, where x represents the system states. Assumption 2 can easily be relaxed by introducing scaling matrices. Note that the common assumption $B_{c1} D_{c21}^* = 0$ is not meaningful since the signal r is directly measured.

A general diagnostic objective function is taken as the output of the following transfer function which takes d and a to z_d:

$$G_d = \left[\begin{array}{c|cc} A_d & B_{d1} & B_{d2} \\ \hline C_{d1} & 0 & D_{d12} \end{array}\right]. \tag{4.12}$$

For example, if the alarm signal should track actuator faults, then an appropriate choice of the signal z_d may be $\left[\begin{smallmatrix} W_d(a - f_i) \\ W_a(a) \end{smallmatrix}\right]$. By combining the two systems to obtain the objective output $[z^T, z_d^T]^T$, the following interconnection structure is obtained:

$$\tilde{G} = \left[\begin{array}{c|cc} \tilde{A} & \tilde{B}_1 & \tilde{B}_2 \\ \hline \tilde{C}_1 & 0 & \tilde{D}_{12} \\ \tilde{C}_2 & \tilde{D}_{21} & 0 \end{array}\right], \tag{4.13}$$

where the inputs are partitioned such that $\tilde{B}_1$ multiplies $[f^T, n^T, r^T]^T$ and $\tilde{B}_2$ multiplies $[a^T, u^T]^T$. The system matrices for $\tilde{G}$ are given by:

$$\tilde{A} = \begin{bmatrix} A_c & 0 \\ 0 & A_d \end{bmatrix}, \tilde{B}_1 = \begin{bmatrix} B_{c1} \\ B_{d1} \end{bmatrix}, \tilde{B}_2 = \begin{bmatrix} 0 & B_{c2} \\ B_{d2} & 0 \end{bmatrix},$$

$$\tilde{C}_1 = \begin{bmatrix} C_{c1} & 0 \\ 0 & C_{d1} \end{bmatrix}, \tilde{D}_{12} = \begin{bmatrix} 0 & D_{c12} \\ D_{d12} & 0 \end{bmatrix},$$

$$\tilde{C}_2 = \begin{bmatrix} C_{c2} & 0 \end{bmatrix}, \tilde{D}_{21} = D_{c21}.$$

The $\mathcal{H}_2$ optimal controller can be found in [30] and is given by

$$K = \left[\begin{array}{c|c} \tilde{A} + \tilde{B}_2 F_2 + L_2 \tilde{C}_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right], \tag{4.14}$$

where

$$-L_2 = Y\tilde{C}_2^* + \tilde{B}_1\tilde{D}_{21}^*, \; -F_2 = \tilde{B}_2^* X,$$

and X and Y are the solutions to the algebraic Riccati equations determined respectively by the Hamiltonian matrices H_2 and J_2:

$$H_2 = \begin{bmatrix} \tilde{A} & -\tilde{B}_2\tilde{B}_2{}^* \\ -\tilde{C}_1{}^*\tilde{C}_1 & -\tilde{A}^* \end{bmatrix},$$

$$J_2 = \begin{bmatrix} \tilde{A}^* - \tilde{C}_2^*\tilde{D}_{21}\tilde{B}_1^* & -\tilde{C}_2{}^*\tilde{C}_2 \\ -\tilde{B}_1(I - \tilde{D}_{21}^*\tilde{D}_{21})\tilde{B}_1{}^* & -\tilde{A} + \tilde{B}_1\tilde{D}_{21}^*\tilde{C}_2 \end{bmatrix}.$$

By partitioning X and Y in the obvious fashion and introducing the following matrices:

$$M = \tilde{A} - \tilde{B}_1\tilde{D}_{21}^*\tilde{C}_2 = \begin{bmatrix} A_c - B_{c1}D_{c21}^*C_{c2} & 0 \\ -B_{d1}D_{c21}^*C_{c2} & A_d \end{bmatrix},$$

$$R = \tilde{B}_1(I - \tilde{D}_{21}^*\tilde{D}_{21})\tilde{B}_1^* = \begin{bmatrix} B_{c1}(I - D_{c21}^*D_{c21})B_{c1}^* & B_{c1}(I - D_{c21}^*D_{c21})B_{d1}^* \\ B_{d1}(I - D_{c21}^*D_{c21})B_{c1}^* & B_{d1}(I - D_{c21}^*D_{c21})B_{d1}^* \end{bmatrix},$$

the two Riccati equations can be written as the following coupled matrix equations:

$$A_c^*X_{11} + X_{11}A_c - X_{11}B_{c2}B_{c2}^*X_{11} + C_{c1}^*C_{c1} - X_{12}B_{d2}B_{d2}^*X_{12}^* = 0,$$

$$A_d^*X_{22} + X_{22}A_d + C_{d1}^*C_{d1} - X_{22}B_{d2}B_{d2}^*X_{22} - X_{12}B_{c2}B_{c2}^*X_{12}^* = 0,$$

$$(A_c^* - X_{11}B_{c2}B_{c2}^*)X_{12} + X_{12}(A_d - B_{d2}B_{d2}^*X_{22}) = 0$$

$$M_{11}Y_{11} + Y_{11}M_{11}^* - Y_{11}C_{c2}^*C_{c2}Y_{11} + R_{11} = 0,$$

$$M_{11}Y_{12} + Y_{12}M_{22}^* + Y_{11}M_{21}^* - Y_{11}C_{c2}^*C_{c2}Y_{12} + R_{12} = 0,$$

$$M_{22}Y_{22} + M_{21}Y_{12} + Y_{22}M_{22}^* + Y_{12}^*M_{21}^* + R_{22} - Y_{12}^*C_{c2}^*C_{c2}Y_{12} = 0$$

We see that the equations for Y are coupled only in one direction, i.e., Y_{12} depends upon Y_{11}, but Y_{11} can be found independently of Y_{12}. In fact, none of the parameters for Y_{11} depend upon the diagnostic terms, so Y_{11} is the same solution that would be obtained if diagnostics were not considered. As for the X equations, the second is linear and homogeneous in X_{12}, so $X_{12} = 0$. This decouples the equations. F_2 and L_2 are then given by

$$F_2 = -\begin{bmatrix} B_{c2}^*X_{11} & 0 \\ 0 & B_{d2}^*X_{22} \end{bmatrix},$$

$$L_2 = -\begin{bmatrix} Y_{11}C_{c2}^* + B_{c1}D_{c21}^* \\ Y_{12}^*C_{c2}^* + B_{d1}D_{c21}^* \end{bmatrix}.$$

Thus, for systems whose interconnection structure can be described by (4.13), the optimal control module is independent of the optimal diagnostic module for the $\mathcal{H}_2$ solution. Thus, when considering only a nominal model, the two step procedure of designing an optimal control module, calculating the closed loop system, and then designing the optimal diagnostic module is equivalent to designing the optimal control

and diagnostic module simultaneously. However, when uncertainty is considered, we will show by example that the two modules should be considered simultaneously.

The key feature of the system (4.13) which allows for the control to be designed independently of the diagnostic is that the alarm signal a does not affect the states associated with G_c. This restriction is reasonable as its violation would imply that the overall objective included some cross-terms between the output and the alarm. For example, in a stochastic setting, assuming the form (4.13) would allow objective function terms of the form $E(y^T y)$ and $E\left((a-f)^T(a-f)\right)$, but not $E\left(y^T(a-f)\right)$, where E is the expectation operator.

4.6 Robust synthesis

Viewing the integrated control-diagnostic problem as a simple case of the general interconnection framework as shown in Figure 4.2 allows the application of robust synthesis tools such as those in [5] in a straightforward fashion. In particular, a control-diagnostic module which guarantees robust performance may be synthesized using a two block μ structure, and the DK iteration method.

Consider a system with the structure shown in Figure 4.1. We would like to design a control-diagnostic module with the following properties:

1. The output signal tracks reference commands and is insensitive to actuator faults.

2. The alarm signal is large only when a fault has occurred.

3. Properties 1 and 2 hold in the presence of a bounded uncertainty.

Let T_c be the transfer function from $[f^T, n^T, r^T]^T$ to y-r, T_d be the transfer function from $[f^T, n^T, r^T]^T$ to a-f, and T_a be the transfer function from $[f^T, n^T, r^T]^T$ to a. A mathematical statement of the design objective is formulated as follows: Find a

stabilizing K such that

$$\left\| \begin{array}{c} W_c T_c \\ W_d T_d \\ W_a T_a \end{array} \right\|_{\mathcal{H}_\infty} \leq 1 \text{ for all } \Delta_u \mid \|\Delta_u\|_{\mathcal{H}_\infty} \leq 1.$$

Including a weight on T_a is needed for the problem to be well posed. By introducing a full performance block Δ_p, this problem can be stated equivalently as: Find a stabilizing K such that

$$\mu_{\left[\begin{smallmatrix} \Delta_u & \\ & \Delta_p \end{smallmatrix}\right]} \left(F_l(\tilde{G}, K)\right) < 1,$$

where $\tilde{G}$ is an appropriate interconnection structure which contains the weighting functions, and F_l represents the lower linear fractional transformation.

Now let us consider as a specific example a second order, two-input, two-output system with actuator faults and measurement noise, i.e., $f = f_i$, $n = n_o$, and $f_o = n_i = 0$. The frequency response for the nominal model is shown in Figure 4.3, and the system matrices are given by

$$A = \begin{bmatrix} -0.5672 & -0.0629 \\ 0.7189 & -0.1175 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.0592 & 0.0014 \\ 0.1039 & -0.1132 \end{bmatrix},$$

$$C = \begin{bmatrix} -0.0916 & -0.1997 \\ 0.2251 & -0.4038 \end{bmatrix},$$

$$D = 0.$$

We assume an uncertainty of 10% in the gain of each input channel, and cover this model set by an input uncertainty, with $W_u = 0.1I$, $\Delta_u = \text{diag}(\Delta_1, \Delta_2)$, where Δ_1 and Δ_2 depict 1×1 blocks, and $\|\Delta_u\|_{\mathcal{H}_\infty} \leq 1$. The performance weights W_c, W_d, and W_a are diagonal, with the same weight on each component, and the magnitudes

of these weights are shown in Figure 4.4. Note that the performance objective is not very aggressive, namely, we desire steady state reference tracking within 10%. One would expect that this type of objective could be easily achieved.

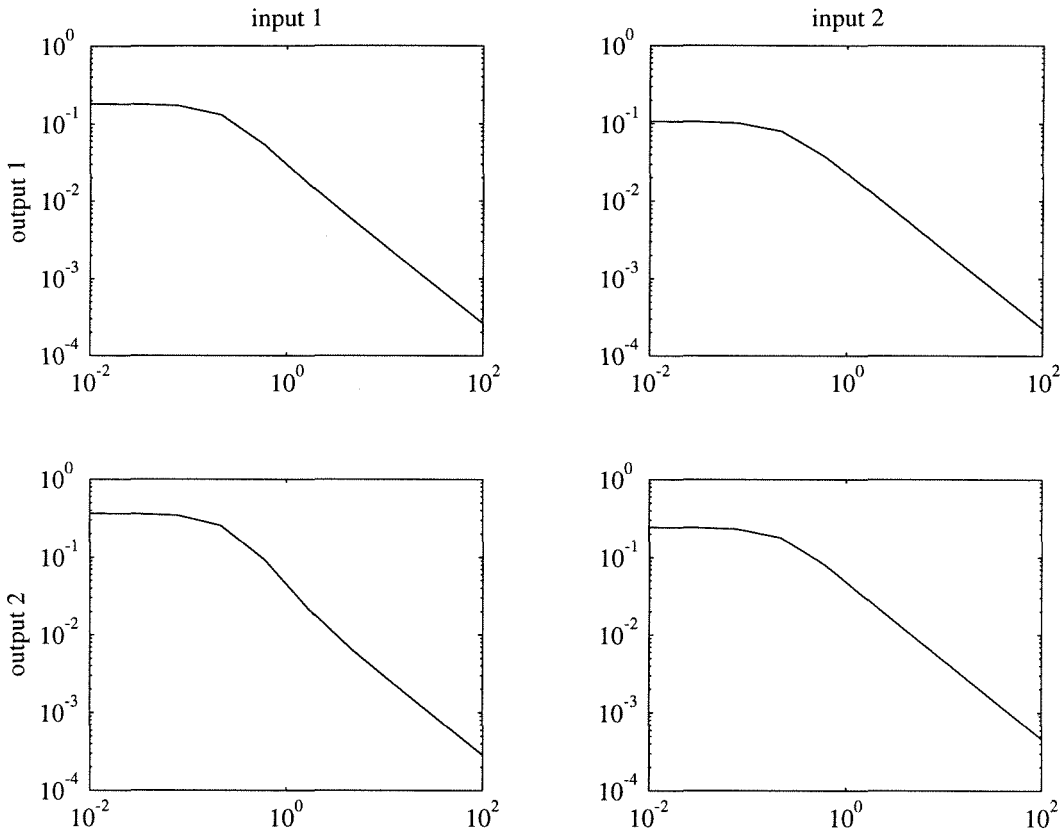

Figure 4.3: Frequency response for plant

Two controllers were synthesized. For the first controller, the DK iteration method was used to simultaneously design the control and diagnostic modules, and a final value of $\mu = 0.95$ was achieved. We will refer to this controller as the one-step controller. In the second case, a robust control module was designed without considering the diagnostic objective. The resulting controller was used to close the loop. Synthesis of an $\mathcal{H}_\infty$ optimal diagnostic led to a value of $\mu \approx 12$. Several iterations on the DK scheme resulted in $\mu \approx 6.5$. We describe the control-diagnostic module which achieves the latter μ value as the two-step controller.

For both the one-step and the two-step control-diagnostic modules, the output

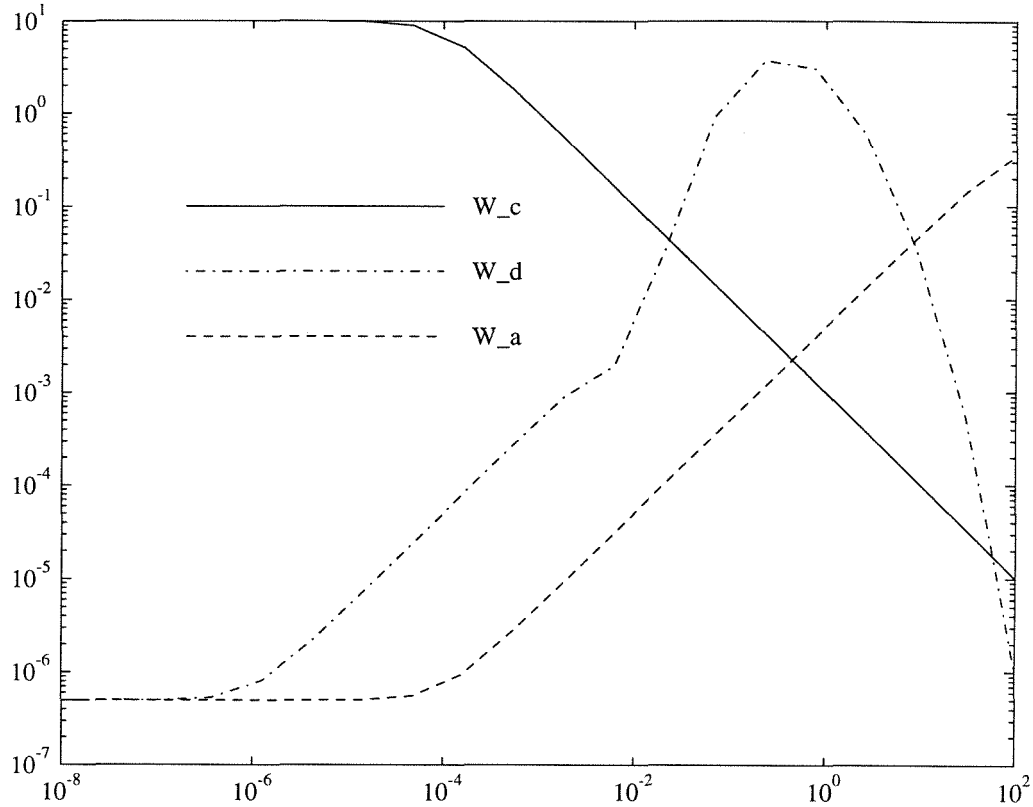

Figure 4.4: Performance weights

was simulated with Δ_u = diag(0.1,-0.1). The input consisted of a ramp actuator fault in the first channel which rose from 0 to 1 between time 100 and 200, and a set point step change of -1 for the second output at time 10, as shown in Figure 4.5. The results of the simulation are shown in Figure 4.6, which displays the plant output, and Figure 4.7, which shows the filtered alarm signal $W_d a$.

From the simulations, one can make the following observations. The two-step control-diagnostic module provides much better performance in regard to reference tracking; however, the diagnostics are inadequate. A large false alarm occurs at time 10 when the reference signal changes. The one-step module does not issue a false alarm, but experiences significantly deteriorated reference tracking performance. This suggests that robust control-diagnostic performance requires detuning of the control action to a much greater extent than would be expected when diagnostics are not

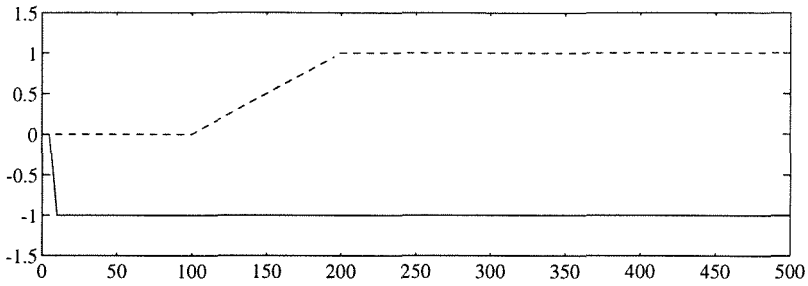

Figure 4.5: Fault in first actuator (dashed), and reference signal for second output (solid). All other input signals are zero.

considered.

For the one-step system, the alarm signal is raised when the slope of the fault changes, but not during the constant slope section of the fault ramp. This is due to the diagnostic weighting function W_d which annihilates low frequency behavior. For this example, it was not possible to achieve robust performance with a weighting with significant low frequency energy. Roughly speaking, the facts that a signal $W_u\Delta_u u$ enters the system in the same way as the fault f and that in steady state u must change to track r preclude the possibility of robust control and detection in the low frequency range.

The two-step controller achieves robust control performance, with $\mu \approx 0.1$ for the first step of this synthesis; however, the resulting control action leads to poor diagnostics. This suggests that methods for optimal or robust diagnostics which do not explicitly consider control action, such as in [28], will not work well for uncertain plants.

4.7 Conclusions

In this chapter, we have shown how systems which integrate diagnostics and control can be designed using standard methods. For nominal models, $\mathcal{H}_2$ theory justifies designing the control and diagnostic modules successively; however, uncertainty requires that the design be simultaneous since diagnostic objectives may limit achievable

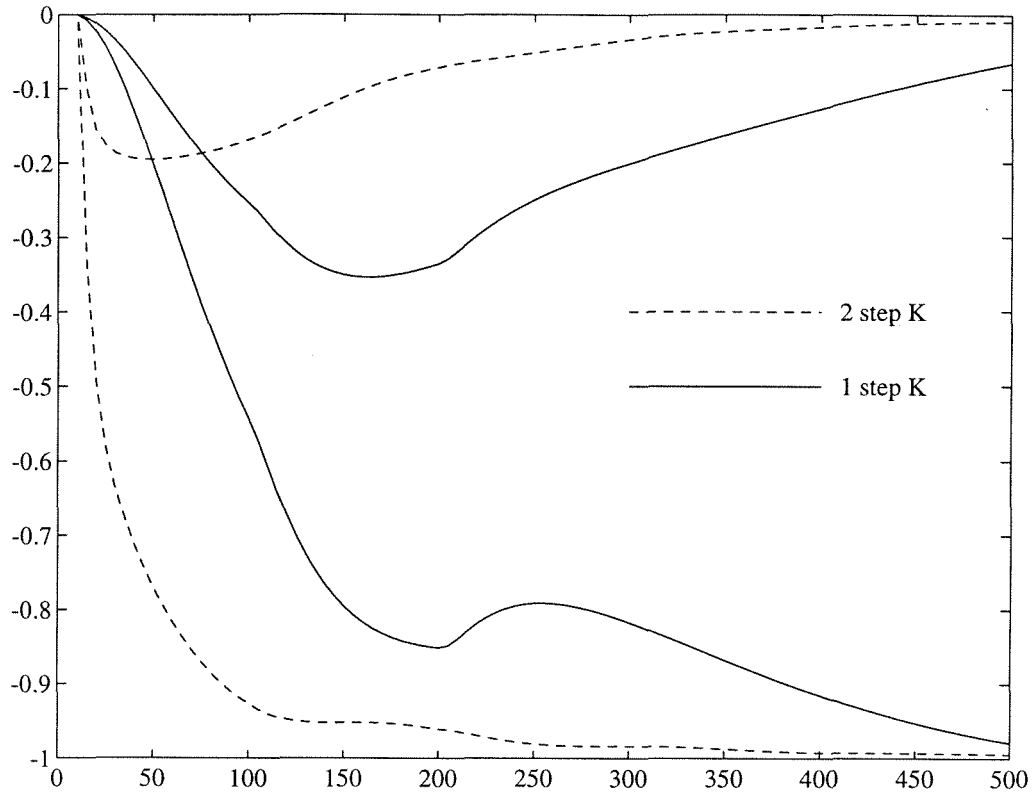

Figure 4.6: Plant output for one-step and two-step controller designs

performance.

An important question which we did not address in this study is the selection of performance measures for diagnostics. We considered the $\mathcal{H}_2$ norm for the nominal case, and the $\mathcal{H}_\infty$ norm for the robust case primarily due to the availability of methods to address these problems. In choosing a performance measure for a diagnostic system, one should consider the detection algorithm which will be used to determine when faults have occurred.

From the example in Section 4.6, two important conclusions concerning the design of linear fault detection filters can be made. First, small model uncertainties which are not problematic from a control perspective can pose severe limitations on detection performance. One explanation for this phenomenon lies in the fact that in controlling a process, the use of feedback tends to mitigate poor modeling so that

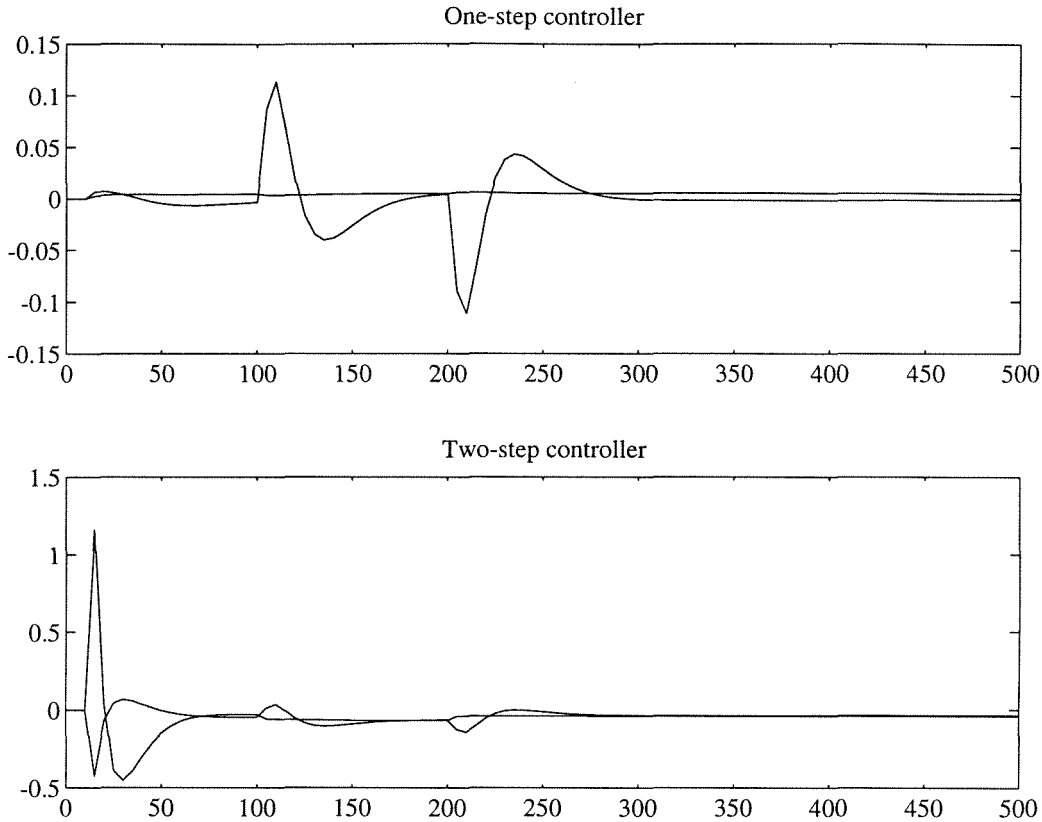

Figure 4.7: Filtered alarm signal $W_d a$ for one-step and two-step controller designs

good performance can be obtained with an uncertain model. Indeed, many classic control methods are independent of a process models. On the other hand, in the detection problem the model is fundamental as it provides the redundant relationships between the measured signals whose violation indicates a fault. When uncertainty enters the picture, these relationships based upon the model cease to hold under fault free circumstances. Second, when tight control is required for uncertain process, it may often be the case that no linear detection filter will provide satisfactory detection performance even for modest model uncertainty. In the example, in order to detect changes in the desired frequency range, it was necessary that the controller be detuned to the point that it was sluggish enough that the effect of control action propagated through the uncertainty could be distinguished from faults.

Due to the inadequacy of linear filters in model based fault detection, methods

based upon other approaches must be pursued. In the following chapters, several such methods are proposed. In Chapters 5 and 6, fault detection based upon Moving Horizon Estimation is considered. This algorithm's capacity to incorporate constraints in estimation problems can be used to reduce fault detection false alarms due to model uncertainty. In Chapter 7, a statistical approach to fault detection using nonlinear models is developed. Both of these approaches rely exclusively on quantitative models of the monitored systems. By combining qualitative and quantitative process information, improved detection may be possible. Chapter 8 investigates the incorporation of qualitative information in process engineering problems using a propositional logic formulation.

Chapter 5 Moving Horizon Estimation

Summary

A Moving Horizon Estimation method is reviewed and its stability properties are investigated. For the unconstrained case, stability is guaranteed by properly choosing an initial error weighting matrix. For the constrained case, feasibility at each step can be guaranteed by expanding the constraints to include the future trajectory. Stability in the constrained case with the expanded constraint set and a modified objective is proven. In both cases, the proposed algorithms involve solving at each time step a quadratic programming problem. Examples are used to demonstrate the improved stability properties of the modified algorithm, as well as to show how constraints can improve estimation of unknown faults and disturbances.

5.1 Introduction

In Chapter 4, the problem of designing linear filters for fault detection was considered. Filters designed for detection purposes can be considered as a special kind of estimator whose purpose is to estimate the value of unknown fault signals. For many process monitoring problems, estimation of state variables and parameters is an important component. A wide number of applications of parameter estimation in fault detection schemes have appeared in the literature as well as in review articles [47, 48]. State estimation via Kalman filters [110] has long been used to calculate residual or error signals which should be small under normal operation, but differ from zero for faulty operation.

Before collecting any data, a fair amount of information is usually known about the numerical values of parameters and states with physical significance. Often, this knowledge can be transformed into constraints on the variables. For example, a state

representing the mole fraction of a chemical species must lie between 0 and 1. If the data are processed in batch fashion, constraints can easily be incorporated within least squares estimation, and including this knowledge may result in more reliable estimates.

For estimation of dynamic states and parameters, recent attention has focused on performing on-line optimization. Zimmer [115] derived an iterative process based on Newton's method for estimating the state of a nonlinear system. Michalska and Mayne [69] presented an observer for the case of a noise and disturbance free nonlinear system which produces an estimate of the state by minimizing an L_2 cost function over a preceding interval which advances with time. Robertson et al. [92] have presented a least squares formulation for a general estimation in which noise and disturbances as well as constraints on the estimated variables are considered. A probabilistic interpretation of this approach corresponding to a modification of the Kalman filter with the innovations distributed according to a truncated Gaussian distribution rather than a pure Gaussian distribution was also developed in the cited reference. As this method updates the state and parameter estimates by solving a finite moving horizon least squares problem, it may be considered dual to the Model Predictive Control (MPC) problem [37].

As the Moving Horizon Estimation (MHE) scheme of [92] is quite similar to MPC, one may anticipate that some of the shortcomings of MPC are inherent in this scheme. For example, it is well known that MPC can result in an unstable closed loop when the horizon length is not long enough [13]. Similarly, if not properly implemented, constrained MHE may result in estimates which either fail to converge to the correct value or completely diverge. It is imperative that any estimation scheme satisfy certain stability requirements, otherwise the estimates obtained may diverge and become completely meaningless.

Muske et al. [76] have presented stability results for constrained batch estimation implemented in a recursive fashion. Although this constrained estimation scheme is stable, it suffers from the shortcoming that as more data is collecting, the dimension of the optimization problem which must be solved also increases. For on-line imple-

mentation, the computation time for this quadratic program with growing dimension soon becomes prohibitive. On the other hand, the scheme of [92] results in a constant size quadratic program, but the algorithm may result in an unstable estimate when constraints are enforced.

In this chapter, the stability properties of MHE are examined, and modifications are made to guarantee stability. Specifically, the stability of MHE for linear systems with constraints is addressed. In Section 5.3, the basic concepts of MHE are briefly reviewed, and it is shown that by properly choosing weighting matrices for the estimated variables, MHE can be made stable when no constraints are enforced. A similar result was presented by [92]. When constraints are enforced, this scheme can lead to an unstable estimate. Section 5.4 introduces a modified constrained MHE algorithm and presents stability results. Section 5.5 extends the results to the case of nonzero known input signal u. Two examples are contained in Section 5.6 to demonstrate how constraints could lead to an unstable estimate and to illustrate the proposed methods.

5.1.1 Notation

Lower case letters such as x, y, denote vector quantities. Upper case letters such as A, Σ, denote matrices. Script characters, such as $\mathcal{V}$, $\mathcal{Y}$, denote concatenations of vector variables from several consecutive sampling instances, collected into a single vector variable. The circumflex accent applied to a variable, such as $\hat{x}$, denotes an estimate of that variable.

Unless otherwise stated, $\|\cdot\|$ stands for the Euclidean norm. The notation $\|x\|_R$ is used to denote the weighted norm

$$\|x\|_R^2 = x^t R x. \tag{5.1}$$

5.2 Preliminaries

The estimation scheme presented in this chapter is applicable to sample data systems which can be accurately represented by linear time invariant models, and which are driven by both a known signal, u, and an unknown signal w. The latter signal may be interpreted as a disturbance. It is assumed that the plant transfer function mapping u and w to the output y is stable, but that the dynamics governing the evolution of the disturbance w contain a pole at 1, which is necessary to allow the disturbance to assume different steady state values. Therefore, the systems considered herein have the form:

$$\begin{aligned} x(k+1) &= Ax(k) + B_u u(k) + B_w w(k) + B_v v_x(k), \\ w(k+1) &= w(k) + v_w(k), \\ y(k) &= C_x x(k) + C_w w(k) + v_y(k). \end{aligned} \tag{5.2}$$

where the eigenvalues of A lie within the unit circle, and v_x, v_w, and v_y are deterministic innovations which are nominally zero. Special cases of this model structure include output step disturbances ($B_w = 0$, $C_w = I$) and input step disturbance ($B_w = B_u$, $C_w = 0$).

In order to simplify the discussion, it will be convenient to consider the case where u is zero, giving the simplified system

$$\begin{aligned} x(k+1) &= Ax(k) + B_w w(k) + B_v v_x(k), \\ w(k+1) &= w(k) + v_w(k), \\ y(k) &= Cx(k) + v_y(k). \end{aligned} \tag{5.3}$$

Extensions to the case where u is nonzero can be found in Section 5.5.

It will also be useful to denote by $\xi(k)$ the augmented system state $[x(k)^t, w(k)^t]^t$, and similarly by $v_\xi(k)$ the augmented innovation, $[v_x(k)^t, v_w(k)^t]^t$ and by (A_ξ, B_ξ, C_ξ)

the state space matrices for ξ:

$$A_\xi = \begin{bmatrix} A & B_w \\ 0 & I \end{bmatrix}, \; B_\xi = \begin{bmatrix} B_v & 0 \\ 0 & I \end{bmatrix}, \; C_\xi = \begin{bmatrix} C_x & C_w \end{bmatrix}. \tag{5.4}$$

In terms of these augmented variables, (5.3) becomes

$$\begin{aligned} \xi(k+1) &= A_\xi \xi(k) + B_\xi v_\xi(k), \\ y(k) &= C_\xi \xi(k) + v_y(k). \end{aligned} \tag{5.5}$$

5.3 Moving Horizon Estimation

In this section, a brief description of the Moving Horizon Estimation scheme for the system (5.3) is presented, and previous stability results are discussed.

The fundamental concept of Moving Horizon Estimation has recently appeared in several articles [76, 77, 92, 69, 68]. Although both linear and nonlinear formulations were presented in the cited references, we will focus on a linear formulation, as our goal is to guarantee stability for linear systems with constraints. Consider a horizon of length p corresponding to the p most recent time instances, $[k+1-p, k]$. Given an estimate of the system state at the beginning of the horizon, $\hat{\xi}_0(k) \approx \xi(k+1-p)$, smoothed and filtered states are obtained by solving on-line the following fixed interval smoothing problem:

$$\min_{\hat{v}_\xi(i|k), \Delta\hat{\xi}(k+1-p)} \sum_{i=k+1-p}^{k} \|\hat{v}_y(i|k)\|^2_{R^{-1}} + \sum_{i=k+1-p}^{k-1} \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}} + \|\Delta\hat{\xi}(k+1-p)\|^2_{\Sigma^{-1}}, \tag{5.6}$$

subject to

$$\hat{\xi}(k+1-p|k) = \hat{\xi}_0(k) + \Delta\hat{\xi}(k+1-p), \tag{5.7}$$

$$\hat{\xi}(i|k) = A_\xi \hat{\xi}(i-1|k) + B_\xi \hat{v}_\xi(i-1|k), \; i = k+2-p, ..., k, \tag{5.8}$$

$$\hat{v}_y(i|k) = y(i) - C_\xi \hat{\xi}(i|k). \tag{5.9}$$

Various formulations which appeared in the literature may be viewed as special cases of the above problem. For example, in the special case where $p = k + 1$, or in other words the horizon corresponds to time $[0, k]$, the problem of batch estimation is obtained. Although stability is guaranteed in this case even with the addition of constraints on the states and innovations [76], as new data become available the size of the optimization problem to be solved grows, and on-line computation becomes inhibitive. In order to obtain an optimization whose size does not increase with time, it is desirable to restrict p to be a constant. In this case, stability of the estimation scheme depends upon the procedure for updating the state at the beginning of the horizon, $\hat{\xi}_0(k)$, from the previous value $\hat{\xi}_0(k-1)$.

In order to circumvent choosing an initial state estimate, $\hat{\xi}_0(k)$, one approach is to allow $\hat{\xi}(k+1-p|k)$ to vary freely, which corresponds to $\Sigma = \infty$. Michalska and Mayne [68] presented a scheme in which the only decision variable involved in the optimization is $\hat{\xi}_0(k)$. This can be considered as a limiting case corresponding to $Q = 0$. As this formulation does not allow for non-zero values of the variables $\hat{v}_\xi$, when model errors exist or disturbances are present, large estimation errors may often result.

Muske and Rawlings [77] have noted that in the unconstrained case, allowing $\hat{\xi}_0(k)$ to vary freely guarantees stability, and that in the constrained case, by relaxing some of the constraints, stability can also be achieved. This is not surprising since a free choice of $\hat{\xi}(k+1-p|k)$ is equivalent to a time invariant FIR filter in the unconstrained case and a time varying FIR filter in the constrained case, and it is well known that FIR filters have built-in Bounded Input/Bounded Output stability (see for example [58]); however, several shortcomings exist in this approach. First, none of the information about the system which is obtained from the data on the horizon $[0, k-p]$ is used in forming the smoothed and estimated states. This is unfortunate, since typically a quite accurate estimate of $\xi(k+1-p)$ can be obtained from the data up to time $k-1$. By neglecting data on the horizon $[0, k-p]$, short horizons will lead to very poor behavior of the filtered state. For example, in the limiting case of $p = 1$, the filtered estimate $\hat{\xi}(k|k)$ will simply correspond to the projection of $y(k)$ onto

the state space: $\hat{\xi}(k|k) = \arg\min_{\xi} \|y(k) - C_{\xi}\xi\|$. Although better filtered estimates $\xi(k|k)$ will be obtained by lengthening the horizon, the smoothed states $\xi(k+i-p|k)$, especially for i near 1, will depend almost entirely on future data rather than on both past and future data. Also, in the constrained case, the smoothed states may not satisfy the *a priori* system information since the constraints may have to be relaxed.

Robertson et al. [92] suggest that the initial state $\hat{\xi}_0(k)$ be chosen as the one-step ahead predicted state, $\hat{\xi}(k+1-p|k-p) = A_{\xi}\hat{\xi}(k-p|k-p)$. In addition, the weighting matrix Σ would be updated according to the Riccati difference equation

$$\Sigma(k+1) = A_{\xi}\left(\Sigma(k) - \Sigma(k)C_{\xi}^{t}\left(C_{\xi}\Sigma(k)C_{\xi}^{t} + R\right)^{-1} C_{\xi}\Sigma(k)\right) A_{\xi}^{t} + Q. \tag{5.10}$$

This choice is motivated by a stochastic interpretation. For the case where v_{ξ} and v_{y} are white noise processes, the matrix $\Sigma(k-p)$ and vector $\hat{\xi}(k-p|k-p)$ capture all the information contained in the data on the horizon $[0, k-p]$. In the unconstrained case, the filtered state $\hat{\xi}(k|k)$ obtained from this MHE is identical to that obtained from the Kalman filter, so that in the nominal linear unconstrained case there is no advantage in using a MHE formulation. Therefore, the primary advantage of MHE over the Kalman filter lies in its ability to incorporate prior knowledge in the form of constraints; however, in the constrained case, stability is not guaranteed independent of p for the above choice of Σ and $\hat{\xi}_0$.

In the constrained case, one straightforward way to guarantee stability would be to choose $\hat{\xi}_0(k)$ as the state $\xi^S(k+1-p)$ of an unconstrained observer with known stability properties [1]. While this clearly yields a stable smoothing and filtering algorithm, many of the advantages gained by including constraints will be compromised. Since in this approach there is no guarantee that ξ^S satisfies the constraints, it may not be feasible from this initial state to obtain smoothed and filtered states which satisfy the constraints, thereby necessitating that the constraints be temporarily relaxed. More significantly, one goal of implementing constraints is to increase the performance of the smoother/filter when the model is uncertain or the measurements

[1] J. H. Lee, personal communication

are faulty. Under these circumstances, it is possible that the unconstrained estimate ξ^S deviates significantly from the true system state. By using such an $\hat{\xi}_0$ as the initial state in the optimization, the advantages gained by incorporating constraints will be compromised.

The smoothed state $\hat{\xi}(k+1-p|k-1)$ presents another choice for the state at the beginning of the horizon, $\hat{\xi}_0(k)$. Intuitively, this choice has the appeal that it represents the best estimate of the state $\xi(k+1-p)$ given all the currently available measurements. Additionally, for the constrained case stability results are more easily derived using the smoothed state since, in the MHE framework, the relationship between $\hat{\xi}(k+1-p|k-1)$ and $\hat{\xi}(k+2-p|k)$ is simpler than the relationship between $\hat{\xi}(k|k-1)$ and $\hat{\xi}(k+1|k)$. In the algorithms presented in this chapter, the smoothed state $\hat{\xi}(k+1-p|k-1)$ is therefore used as the beginning horizon state $\hat{\xi}_0(k)$.

By choosing $\hat{\xi}_0(k) = \hat{\xi}(k+1-p|k-1)$, the weighting matrix Σ obtained from the Riccati equation (5.10) may not provide stability in the unconstrained case. Since $\hat{\xi}(k+1-p|k-1)$ represents a better estimate of the true state than $\hat{\xi}(k+1-p|k-p)$ due to the additional knowledge of the system obtained from the measurements $y(k+1-p), ..., y(k-1)$, one would expect that a larger penalty on $\Delta\hat{\xi}(k+1-p)$ (smaller Σ) should be used in the objective function when $\hat{\xi}_0$ is $\hat{\xi}(k+1-p|k-1)$ than when it is $\hat{\xi}(k+1-p|k-p)$. Given the weights Q and R and the horizon length p, a constant matrix Σ which guarantees stability of the MHE in the unconstrained case can be calculated by solving an algebraic Riccati equation, as will be shown next.

For convenience define the concatenated measurement vector $\mathcal{Y}(k+1-p,k)$ as follows:

$$\mathcal{Y}(k+1-p,k) = \left[y^t(k+1-p), y^t(k+2-p), \ldots, y^t(k)\right]^t. \tag{5.11}$$

Also, define $\mathcal{V}_y(k+1-p,k)$, $\hat{\mathcal{V}}_y(k+1-p,k)$, $\mathcal{V}_\xi(k+1-p,k-1)$, and $\hat{\mathcal{V}}_\xi(k+1-p,k-1)$ similarly for the variables $v_y(\cdot)$, $\hat{v}_y(\cdot|k)$, $v_\xi(\cdot)$ and $\hat{v}_\xi(\cdot|k)$ respectively.

It can easily be shown that in the unconstrained case the optimal choice of the

decision variables is given by

$$\begin{aligned}\begin{pmatrix} \Delta\hat{\xi}(k+1-p) \\ \hat{\mathcal{V}}_\xi(k+1-p,k-1) \end{pmatrix} &= \left[\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & \bar{Q}^{-1} \end{pmatrix} + [\bar{C},\bar{D}]^t \bar{R}^{-1} [\bar{C},\bar{D}]\right]^{-1} [\bar{C},\bar{D}]^t \bar{R}^{-1}\mathcal{Y}^0(k) \\ &= \begin{pmatrix} \Sigma & 0 \\ 0 & \bar{Q} \end{pmatrix} [\bar{C},\bar{D}]^t \times \\ &= \left[[\bar{C},\bar{D}] \begin{pmatrix} \Sigma & 0 \\ 0 & \bar{Q} \end{pmatrix} [\bar{C},\bar{D}]^t + \bar{R}\right]^{-1} \mathcal{Y}^0(k) \end{aligned} \tag{5.12}$$

where $\mathcal{Y}^0(k)$ is calculated from the data and initial state:

$$\mathcal{Y}^0(k) = \mathcal{Y}(k+1-p,k) - \bar{C}\hat{\xi}(k+1-p|k-1), \tag{5.13}$$

and that this solution relates $\hat{\xi}(k+2-p|k)$ to $\hat{\xi}(k+1-p|k-1)$ through:

$$\begin{aligned}\hat{\xi}(k+2-p|k) &= A_\xi\left(\hat{\xi}(k+1-p|k-1) + \Delta\hat{\xi}(k+1-p)\right) + \bar{B}\hat{\mathcal{V}}_\xi(k-1,p), \\ &= \left(A_\xi - L\bar{C}\right)\hat{\xi}(k+1-p|k-1) + L\mathcal{Y}(k+1-p,k), \end{aligned} \tag{5.14}$$

where

$$\bar{C} = \begin{bmatrix} C_\xi \\ C_\xi A_\xi \\ \vdots \\ C_\xi A_\xi^{p-1} \end{bmatrix}, \quad \bar{D} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ C_\xi B_\xi & 0 & \dots & 0 \\ C_\xi A_\xi B_\xi & C_\xi B_\xi & \dots & 0 \\ \vdots & & \ddots & \vdots \\ C_\xi A_\xi^{p-2} B_\xi & C_\xi A_\xi^{p-3} B_\xi & \dots & C_\xi B_\xi \end{bmatrix}, \tag{5.15}$$

$$\bar{B} = [B_\xi, 0, \dots, 0], \quad \bar{R} = \operatorname{diag}[R, R, \dots, R], \quad \bar{Q} = \operatorname{diag}[Q, Q, \dots, Q]. \tag{5.16}$$

$$L = \left(A_\xi \Sigma \bar{C}^t + S\right)\left(\bar{C}\Sigma\bar{C}^t + \breve{R}\right)^{-1}, \tag{5.17}$$

$$\breve{R} = \bar{D}\bar{Q}\bar{D}^t + \bar{R}, \tag{5.18}$$

$$S = \begin{bmatrix} 0 & B_\xi Q B_\xi^t C^t & B_\xi Q B_\xi^t A^t C^t & \dots & B_\xi Q B_\xi^t (A^{p-2})^t C^t \end{bmatrix}. \tag{5.19}$$

From (5.14) one can see that the unconstrained least squares procedure results in a Luenberger observer for updating $\hat{\xi}_0$, and that the stability of this observer is determined by the eigenvalues of $A_\xi - L\bar{C}$. For any value of the weights R and Q, it is possible to choose Σ so that stability is ensured as long as the system satisfies certain detectability conditions. One such choice of Σ can be obtained by considering solutions of the following algebraic Riccati equation (ARE)

$$\Sigma = \Psi\Sigma\Psi^t - \Psi\Sigma\bar{C}^t(\bar{C}\Sigma\bar{C}^t + \breve{R})^{-1}\bar{C}\Sigma\Psi^t + B_\xi Q {B_\xi}^t - S\breve{R}^{-1}S^t, \tag{5.20}$$

with $\Psi = A_\xi - S\breve{R}^{-1}\bar{C}$. If $(\bar{C}, A_\xi)$ is detectable, then there exists a solution to the ARE such that the matrix

$$\Psi\left(I + \Sigma\bar{C}^t\breve{R}^{-1}\bar{C}\right)^{-1} = A_\xi - L\bar{C} \tag{5.21}$$

is stable [72]. Using this estimator, the state error, $\hat{\xi} - \xi$, evolves as

$$\begin{aligned} \hat{\xi}(k+2-p|k) - \xi(k+2-p) &= (A_\xi - L\bar{C})(\hat{\xi}(k+1-p|k-1) - \xi(k+1-p)) \\ &\quad + L\mathcal{V}_y(k+1-p,k) \\ &\quad + \left(L\bar{D} - \bar{B}\right)\mathcal{V}_\xi(k+1-p,k-1) \end{aligned} \tag{5.22}$$

so that for the case wherein $v_y(k)$ and $v_\xi(k)$ approach zero asymptotically, the estimation error also approaches zero asymptotically.

Let us now collect the equations describing this Moving Horizon Estimator to define the estimation scheme MHE 1:

MHE 1 *Given the estimate $\hat{\xi}(k+1-p|k-1)$, form the updated estimates according to:*

Initial Error Update:

$$\hat{\xi}(k+1-p|k) = \hat{\xi}(k+1-p|k-1) + \Delta\hat{\xi}(k+1-p), \tag{5.23}$$

Data update, *for* $i = k+2-p, ..., k$*:*

$$\hat{\xi}(i|k) = A_\xi \hat{\xi}(i-1|k) + B_\xi \hat{v}_\xi(i-1|k), \tag{5.24}$$

where $\Delta\hat{\xi}(k+1-p)$ *and* $\hat{v}_\xi(i|k)$ *minimize the following quadratic program:*

$$\min_{\Delta\hat{\xi}(k+1-p), \hat{v}_\xi(i|k)} J(k,p), \tag{5.25}$$

$$J^2(k,p) = \sum_{i=k+1-p}^{k} \|\hat{v}_y(i|k)\|^2_{R^{-1}} + \sum_{i=k+1-p}^{k} \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}} + \|\Delta\hat{\xi}(k+1-p)\|^2_{\Sigma^{-1}}, \tag{5.26}$$

and with Σ *the solution to the ARE*

$$\Sigma = \Psi\Sigma\Psi^t - \Psi\Sigma\bar{C}^t(\bar{C}\Sigma\bar{C}^t + \breve{R})^{-1}\bar{C}\Sigma\Psi^t + B_\xi Q {B_\xi}^t - S\breve{R}^{-1}S^t, \tag{5.27}$$

where $\Psi = A_\xi - S\breve{R}^{-1}\bar{C}$.

Use the estimate $\hat{\xi}(k+2-p|k)$ *in the next step.*

Remark 6 *When* Σ *is calculated as above, the dynamics of the system which generates* $\hat{\xi}(k+1-p|k-1)$ *from* $\mathcal{Y}(k+1-p,k)$ *is equivalent to the dynamics of the steady state Kalman filter for the* augmented *system*

$$\begin{aligned} \mathbf{x}(k+1) &= A_\xi \mathbf{x}(k) + \bar{v}_\xi(k), \\ \mathbf{y}(k) &= \bar{C}\mathbf{x}(k) + \bar{v}_y(k), \end{aligned} \tag{5.28}$$

with non-zero correlation between the process noise and measurement noise, that is

$$E\begin{pmatrix} \bar{v}_\xi \\ \bar{v}_y \end{pmatrix}\begin{pmatrix} \bar{v}_\xi \\ \bar{v}_y \end{pmatrix}^t = \begin{bmatrix} B_\xi Q B_\xi^t & S \\ S^t & \breve{R} \end{bmatrix}. \tag{5.29}$$

The Kalman filter for this system is given by [1]:

$$\begin{aligned} \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \Sigma\bar{C}^t\left(\bar{C}\Sigma\bar{C}^t + \breve{R}\right)^{-1}\left(\mathbf{y}(k) - \bar{C}\hat{\mathbf{x}}(k|k-1)\right), \\ \hat{\mathbf{x}}(k+1|k) &= A_\xi\hat{\mathbf{x}}(k|k) + S\breve{R}^{-1}\left(\mathbf{y}(k) - \bar{C}\hat{\mathbf{x}}(k|k)\right). \end{aligned} \tag{5.30}$$

5.4 Constrained Moving Horizon Estimation

5.4.1 Constrained formulation

Now consider the case where the state estimates and innovations variables must satisfy certain linear constraints of the form:

$$\begin{aligned} G_\xi \hat{\xi}(k+i-p|k) &\leq h_\xi, \ i \geq 1, \\ G_{\Delta\hat{\xi}} \Delta\hat{\xi}(k+1-p) &\leq h_{\Delta\hat{\xi}}, \\ G_v \hat{v}_\xi(k+i-p|k) &\leq h_v, \ 1 \leq i \leq p-1. \end{aligned} \tag{5.31}$$

The constraints on $\hat{\xi}$ require that not only the smoothed and filtered states, but also the predicted state trajectory satisfy the a priori system information. We shall see that by enforcing the constraints on the predicted state trajectory, it will always be possible to find a feasible solution to the optimization problem. As the predicted state trajectory constraints form an infinite set, we first show how to obtain a finite number of constraints which if satisfied guarantee that the infinite predicted trajectory lies within the feasible region.

The predicted state trajectory is obtained from the filtered state $\hat{\xi}(k|k)$ according to:

$$\hat{\xi}(k+i|k) = A_\xi^i \hat{\xi}(k|k). \tag{5.32}$$

Asymptotically, $\hat{\xi}(k+i|k)$ converges to a constant value $\xi^* = \lim_{j\to\infty} A_\xi^j \hat{\xi}(k|k)$, or equivalently

$$\xi^*(k) = \begin{bmatrix} x^*(k) \\ w^*(k) \end{bmatrix} = \begin{bmatrix} (I-A)^{-1} B_w \hat{w}(k|k) \\ \hat{w}(k|k) \end{bmatrix}. \tag{5.33}$$

Using the asymptotic state $\xi^*(k)$, the predicted state trajectory may be expressed as

$$\hat{\xi}(k+i|k) = \xi^*(k) + A_\xi^i \left(\hat{\xi}(k|k) - \xi^*(k) \right) \tag{5.34}$$

$$= \begin{bmatrix} x^*(k) \\ w^*(k) \end{bmatrix} + \begin{bmatrix} A^i \left(\hat{x}(k|k) - x^*(k) \right) \\ 0 \end{bmatrix}. \tag{5.35}$$

If a linear observer of the form

$$\hat{\xi}(k+1|k+1) = A_\xi \hat{\xi}(k|k) + L\left(y(k+1) - C_\xi A_\xi \hat{\xi}(k|k)\right) \tag{5.36}$$

were used to estimate the state $\hat{\xi}(k|k)$, then $\xi^*(k)$ corresponds to the value that the observer state would converge to under the condition that the predicted output error were identically zero for all time after time k.

Lemmas 1 and 2 use the asymptotic state ξ^* to show that under mild assumptions, only a finite number of constraints for $\hat{\xi}$ need to be considered. In addition, a method is established for determining off-line this finite set of constraints which will guarantee that all constraints are met over the infinite trajectory.

Lemma 1 *Let $G_\xi = [G_x,\ G_w]$. Consider the following maximization problem:*

$$\begin{aligned} \max_{\zeta} \quad & G_{x,j} A^n \zeta \\ \text{s.t.} \quad & G_{x,j} A^l \zeta \leq 1,\ l = 0, \ldots, n-1, \end{aligned} \tag{5.37}$$

where $G_{x,j}$ denotes the j^{th} row of G_x. If there exists an $n > 0$ such that the maximum is less than or equal to 1, and $\xi^*(k)$ satisfies $G_{\xi,j}\xi^*(k) \leq h_{\xi,j}$, then the constraints

$$G_{\xi,j}\hat{\xi}(k+i|k) \leq h_{\xi,j} \tag{5.38}$$

are satisfied for all i if they are satisfied for $i = 0, \ldots, n-1$.

Proof See appendix.

Lemma 2 *Suppose, in addition to the constraints in (5.37), it is required that $\|\zeta\|$ be bounded. Then there exists a finite n such that with the additional constraint on $\|\zeta\|$ the maximization (5.37) is bounded above by one.*

Proof See appendix.

Remark 7 *When $\|\zeta\|$ is not bounded, it may not be possible to find an n such that the condition of Lemma 1 is satisfied. Consider the following example:*

$$G_j = \begin{bmatrix} 1 & -1 \end{bmatrix}, \; A = \mathrm{diag}(\lambda_1, \lambda_2), \tag{5.39}$$

with $1 > \lambda_1 > \lambda_2$. Then

$$G_j A^n \zeta = \lambda_1^n \zeta_1 - \lambda_2^n \zeta_2. \tag{5.40}$$

If $\zeta_2 > \zeta_1 > 0$, then the right-hand side is non-positive whenever n satisfies

$$n \leq \frac{\log(\lambda_1/\lambda_2)}{\log(\zeta_2/\zeta_1)}, \tag{5.41}$$

and strictly positive otherwise. Clearly, for any n it is possible to find ζ_1 and ζ_2 such that $G_j A^m \zeta \leq 0$ for $m < n$, but $G_j A^n \zeta$ is arbitrarily large.

Remark 8 *The condition that the constraints impose a bound on the $\|\zeta\|$ are not very restrictive in practice. Recall from the proof of Lemma 1 that ζ represents $\hat{x}(k|k) - x^*(k)$. Suppose bounds of the form $l \leq C_\xi \xi \leq u$ are imposed, where (A_ξ, C_ξ) form an observable pair. Then it follows that $\|\hat{x}(k|k) - x^*(k)\|$ must be bounded. If not, there exists an arbitrarily large vector $z = [(\hat{x}(k|k) - x^*(k))^t, 0]^t$ such that*

$$\begin{bmatrix} l' \\ l' \\ \vdots \\ l' \end{bmatrix} \leq \begin{bmatrix} C_\xi \\ C_\xi A_\xi \\ \vdots \\ C_\xi A_\xi^{n-1} \end{bmatrix} z \leq \begin{bmatrix} u' \\ u' \\ \vdots \\ u' \end{bmatrix}, \tag{5.42}$$

where $u' = u - C_\xi \xi^*(k)$ and $l' = l - C_\xi \xi^*(k)$. But this can only hold if the matrix has linearly dependent columns, which contradicts the observability assumption. Therefore, in this case, the constraints (5.31) impose a bound on $\hat{x}(k|k) - x^*(k)$ for any finite $\xi^*(k)$.

Define $\mathcal{F}$ to be the set of feasible steady state measurements,

$$\mathcal{F} = \{y | y = C_\xi \xi^*, A_\xi \xi^* = \xi^*, G_\xi \xi^* \leq h_\xi\}, \tag{5.43}$$

or in other words, $y(k) \in \mathcal{F}$ if there exists a corresponding ξ^* which satisfies the constraints (5.31) and constitutes an equilibrium point. If a measured $y(k)$ does not lie within $\mathcal{F}$, then one of several possibilities exist:

1. The constraints are too restrictive for the physical behavior of the system.
2. The dynamic system is in transient.
3. The measurement $y(k)$ is too inaccurate.

For the first case, it would be necessary to gain further insight into the system, and use this new information to reformulate the constraints so they are physically meaningful. In this chapter, we assume that the constraints are valid, so that when $y(k)$ lies outside the set $\mathcal{F}$ asymptotically, it is due to inaccurate measurements. One of the primary motivations for including constraints in an estimation scheme is to improve performance when inaccurate measurements are obtained.

5.4.2 Stability

Now that we have shown that by enforcing a precalculated finite number of constraints, all future predictions can be guaranteed to satisfy these constraints, let us consider the stability properties of constrained MHE. In Section 5.6.1, it is demonstrated through an example that the formulation MHE 1 can result in an unstable estimator when constraints are imposed. Therefore, this algorithm must be modified to ensure stability.

From results in MPC, one can get ideas on how to stabilize constrained MHE. In MPC, one uses control errors in the objective function and extends the prediction horizon to the infinite future [91]. This makes the objective function a Lyapunov function and allows one to prove stability. The dual problem in MHE would be to

consider the state estimation error in the objective function and to extend the horizon to the infinite past. Two difficulties arise in this approach. First, the state estimation error is not known. If it were known, the state would also be known, and therefore the need for an estimator would be eliminated. Second, by extending the problem to the infinite past, the problem size grows at each step and soon becomes intractable.

Instead of the infinite past, consider modifying the MHE objective by extending the output error terms to the infinite future, resulting in an objective of the form:

$$\begin{aligned}\Phi^2(k,p) &= \sum_{i=k+1-p}^{\infty} \|\hat{v}_y(i|k)\|^2_{R^{-1}} \\ &\quad + \sum_{i=k+1-p}^{\infty} \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}} + \|\Delta\hat{\xi}(k+1-p)\|^2_{\Sigma}. \end{aligned} \tag{5.44}$$

To evaluate the terms of this objective function, one must assume the future behavior of the system. Assuming that disturbances remain constant in the future gives $\hat{v}_\xi(i|k) = 0$ for $i \geq k$. Next, consider the terms $\hat{v}_y(i|k)$ which contain measurements $y(i)$ which for $i > k$ have not yet been made. One possibility is to replace these values with the predicted values; however, in this case $\hat{v}_y(i|k) = 0$, and the objective function remains unchanged. We assume instead that the future measurements remain constant, that is $y(i) = y(k)$ for $i > k$, or equivalently

$$\hat{v}_y(i|k) = y(k) - C_\xi\hat{\xi}(i|k) \text{ for } i > k. \tag{5.45}$$

With this modification, the objective function remains bounded if and only if there exists an $\xi^*(k)$ satisfying (5.35) as well as the relation

$$y^*(k) = y(k),\ y^*(k) \equiv C_\xi\xi^*(k). \tag{5.46}$$

In this case, the first term in the objective function becomes

$$\sum_{i=k+1-p}^{\infty} \|\hat{v}_y(i|k)\|^2_{R^{-1}} = \sum_{i=k+1-p}^{k-1} \|y^*(i) - C_\xi\hat{\xi}(i|k)\|^2_{R^{-1}} + \|x^*(k) - \hat{x}(k|k)\|^2_P, \tag{5.47}$$

where P is given by

$$P = \sum_{i=0}^{\infty} A^t C^t R^{-1} C A = A^t P A + C^t R^{-1} C. \tag{5.48}$$

Because the measurement $y(k)$ may not lie within the set $\mathcal{F}$, the constraint (5.46) may not be feasible. This constraint can be relaxed in order to guarantee feasibility of the optimization problem by removing the constraint and penalizing the objective function when this equality is violated by including a term of the form

$$\|y(k) - C_\xi \xi^*(k)\|_\Lambda^2, \tag{5.49}$$

where Λ is any positive definite matrix [114].

By modifying the objective function as described above, the following algorithm results.

MHE 2 *Given the estimate* $\hat{\xi}(k+1-p|k-1)$ *, form the updated estimates according to:*

Initial Error Update:

$$\hat{\xi}(k+1-p|k) = \hat{\xi}(k+1-p|k-1) + \Delta\hat{\xi}(k+1-p), \tag{5.50}$$

Data update, *for* $i = k+2-p, ..., k$*:*

$$\hat{\xi}(i|k) = A_\xi \hat{\xi}(i-1|k) + B_\xi \hat{v}_\xi(i-1|k). \tag{5.51}$$

where $\Delta\hat{\xi}(k+1-p)$ *,* $\hat{v}_\xi(i|k)$*, and* $x^*(k)$ *solve the following quadratic program:*

$$\min_{\Delta\hat{\xi}(k+1-p), \hat{v}_\xi(i|k), x^*(k)} \Phi(k,p), \tag{5.52}$$

$$\Phi^2(k,p) = \sum_{i=k+1-p}^{k-1} \|\hat{v}_y^*(i|k)\|_{R^{-1}}^2 + \|x^*(k) - \hat{x}(k|k)\|_P^2 + \|y(k) - C_\xi \xi^*(k)\|_\Lambda^2$$

$$+ \sum_{i=k+1-p}^{k-1} \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}} + \|\Delta\hat{\xi}(k+1-p)\|^2_{\Sigma^{-1}} \tag{5.53}$$

subject to the constraints (5.31) as well as

$$x^*(k) = Ax^*(k) + B_w\hat{w}(k|k), \tag{5.54}$$

$$\hat{v}_y^*(i|k) = y^*(i) - C_\xi\hat{\xi}(i|k). \tag{5.55}$$

Use the estimate $\hat{\xi}(k+2-p|k)$ in the next step.

In the following, it is shown that MHE 2 with P chosen as the solution to the Lyapunov inequality:

$$A^tPA - P + C^tR^{-1}C \leq 0, \tag{5.56}$$

has desirable convergence properties. Lemma 3 establishes that for asymptotically constant signals y, the estimate $\hat{\xi}(k+1-p|k)$ converges to values determined by ξ^*. Lemmas 4 and 5 establish similar convergence results for the estimates $\hat{\xi}(k+i-p|k)$ for $i \geq 1$. Finally, these preliminary results are used in Theorem 3 to establish the convergence of $\xi^*(k)$.

Lemma 3 *Assume that the signal $y(k)$ converges to a constant value y_∞ such that*

$$\sum_{i=0}^{\infty} \|y(i) - y_\infty\| < m_y. \tag{5.57}$$

Then for any positive definite matrices R, Q, and Σ, and for any horizon length p and positive constant Λ the estimator MHE 2 with P satisfying (5.56) has the following asymptotic convergence properties as $k \to \infty$:

$$C_\xi\hat{\xi}(k+1-p|k) \to y^*(k+1-p) \tag{5.58}$$

$$\Delta\hat{\xi}(k+1-p) \to 0 \tag{5.59}$$

$$\hat{v}_\xi(k+1-p|k) \to 0 \tag{5.60}$$

$$y^*(k+1-p) - y^*(k-p) \to 0. \tag{5.61}$$

Proof See appendix.

Lemma 4 *Suppose that the pair* (C_ξ, A_ξ) *is observable, and that the convergence results (5.58)-(5.61) hold. Then*

$$\hat{\xi}(k+2-p|k) \to \hat{\xi}(k+1-p|k-1). \tag{5.62}$$

Proof See appendix.

Lemma 5 *Under the assumptions of Lemmas 3 and 4,*

$$\hat{\xi}(k+i-p|k) \to \hat{\xi}(k+1-p|k-1),\ i \geq 1. \tag{5.63}$$

Proof See appendix.

Theorem 3 *For* y *as in (5.57), the estimates* $\hat{\xi}(i|k)$ *for* $i \in [k+1-p, k]$ *obtained using MHE 2 with the matrix* P *satisfying (5.56) converge asymptotically to a state* $\xi^*(i)$ *which satisfies*

$$\xi^*(i) \quad = \quad \arg \min_{C_\xi \xi \in \mathcal{F}} \|C_\xi \xi - y(i)\|_\Lambda^2. \tag{5.64}$$

Proof Lemma 3 through 5 imply that $\hat{\xi}(i|k)$ converges to some value $\xi^*(i)$ such that $C_\xi \xi^* \in \mathcal{F}$. It remains to show that ξ^* satisfies (5.64). Lemma 5 implies that every term in the objective Φ approaches zero asymptotically, with the possible exception of the term $\|y(k) - C_\xi \xi^*(k)\|_\Lambda^2$. Let $e^2(k)$ denote this term. Letting $\mathbf{z}$ denote the optimization variables:

$$\mathbf{z} = [\Delta\hat{\xi}^t(k+1-p),\ \hat{\mathcal{V}}_\xi^t(k+1-p, k-1)]^t, \tag{5.65}$$

and denoting $\Delta x^* = x^*(k) - x^*(k-1)$, the objective Φ can be written as

$$\Phi^2 \quad = \quad [\mathbf{z}, \Delta x^*]^t M [\mathbf{z}, \Delta x^*] + [\mathbf{z}, \Delta x^*]^t f + e^2(k), \tag{5.66}$$

where M is a positive definite matrix and f converges to zero. The constraints may be abbreviated as

$$A_{1z}\mathbf{z} + A_{1x}\Delta x^* \leq b_1, \tag{5.67}$$

$$A_{2z}\mathbf{z} + A_{2x}\Delta x^* = 0. \tag{5.68}$$

with $b_1 \geq 0$. One feasible solution to the above problem is $\mathbf{z} = 0$, $\Delta x^* = 0$. Suppose this is the optimal feasible solution. Then the constraints do not change for the next step, and it follows that $\xi^*(k+i) = \xi^*(k)$ for all $i > 0$. From Lemma 5 it follows that $C_\xi \hat{\xi}(k+i-p|k)$ converges to $C_\xi \xi^*(k)$ for all $i > 0$, so that each of the constraints on $\hat{\xi}(k+i-p|k)$ are satisfied asymptotically if and only if $C_\xi \xi^*(k)$ lies in the set $\mathcal{F}$. Therefore, $\xi^*(k)$ is an optimal feasible solution to the problem (5.64).

Suppose the solution $[\mathbf{z}, \Delta x^*] = 0$ is not optimal. For this suboptimal solution, the objective function has the value

$$\Phi^2 = \|y(k) - C_\xi \xi^*(k-1)\|_\Lambda^2. \tag{5.69}$$

Since this solution is a suboptimal, there exists a $\hat{\mathbf{z}}$, $\Delta\hat{x}^*$ such that

$$[\hat{\mathbf{z}}, \Delta\hat{x}^*]^t M [\hat{\mathbf{z}}, \Delta\hat{x}^*] + [\hat{\mathbf{z}}, \Delta\hat{x}^*]^t f + \|y(k) - C_\xi \xi^*(k)\|_\Lambda^2 \leq \|y(k) - C_\xi \xi^*(k-1)\|_\Lambda^2 \tag{5.70}$$

The left-hand side is bounded below by the term

$$\|[\hat{\mathbf{z}}, \Delta\hat{x}^*]\|^2 \underline{\sigma}(M) - \|[\hat{\mathbf{z}}, \Delta\hat{x}^*]\| \|f\| + \|y(k) - C_\xi \xi^*(k)\|_\Lambda^2. \tag{5.71}$$

Since $\|f\| \to 0$, for sufficiently large k, the first two terms are greater than zero. This implies that

$$\|y(k) - C_\xi \xi^*(k)\|_\Lambda^2 < \|y(k) - C_\xi \xi^*(k-1)\|_\Lambda^2. \tag{5.72}$$

Together with the convergence of y, this implies that $\|y(k) - C_\xi \xi^*(k)\|_\Lambda^2$ approaches a limit and therefore $\Delta\hat{\xi}^*(k) = [\Delta x^*(k)^t, v_w(k|k)^t]^t \to 0$, which in turn implies $\mathbf{z} \to 0$,

so that the above argument can be used to show that ξ^* converges to the specified value. ■

5.5 Extension to nonzero input u

The algorithms developed can easily be extended to the case where u is non zero by transforming the system so that $\Delta u(k) = u(k) - u(k-1)$ is considered as the input:

$$\begin{aligned}
\xi(k+1) &= \begin{bmatrix} A & B_w & B_u \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \xi(k) + \begin{bmatrix} B_x & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix} v_\xi(k) + \begin{bmatrix} B_u \\ 0 \\ I \end{bmatrix} \Delta u(k), \\
\begin{bmatrix} y(k) \\ u(k) \end{bmatrix} &= \begin{bmatrix} C_x & C_w & 0 \\ 0 & 0 & I \end{bmatrix} \xi(k) + \begin{bmatrix} 0 \\ I \end{bmatrix} \Delta u(k) + \begin{bmatrix} v_y \\ 0 \end{bmatrix}.
\end{aligned} \tag{5.73}$$

Redefining A_ξ, B_ξ, and B_{ξ_u} as the matrices multiplying $\xi(k)$, v_ξ, and Δu, the update equations become

$$\hat{\xi}(k+1-p|k) = \hat{\xi}(k+1-p|k-1) + \begin{bmatrix} I \\ 0 \end{bmatrix} \Delta\hat{\xi}(k+1-p) \tag{5.74}$$

$$\begin{aligned}
\hat{\xi}(i|k) &= A_\xi \hat{\xi}(i-1|k) + B_\xi \hat{v}_\xi(i-1|k) \\
&\quad + B_{\xi_u} \Delta u(i-1),\ i = k+2-p, ..., k
\end{aligned} \tag{5.75}$$

$$\hat{\xi}(k+i|k) = A_\xi^i \hat{\xi}(k|k),\ i > 0. \tag{5.76}$$

In choosing $\hat{\xi}^*(k)$, it is again required that this auxiliary state vector correspond to an equilibrium position of the systems, and it can easily be seen that $\hat{\xi}^*(k) = \lim_{i\to\infty} A_\xi^i \hat{\xi}(k|k)$ corresponds to the following conditions:

$$\xi^*(k) = \begin{bmatrix} (I-A)^{-1}\left(B_w \hat{w}(k|k) + B_u u(k)\right) \\ \hat{w}(k|k) \\ u(k) \end{bmatrix}. \tag{5.77}$$

5.6 Examples

5.6.1 Instability via constraints

This example is designed to demonstrate that the estimation scheme MHE 1 can become unstable if constraints are enforced in the minimization of the objective $J(k,p)$. Consider a dynamic system consisting of a linear plant G and a controller C operating in closed loop. The plant and controller are given by:

$$G = \frac{10z^3 - 11}{10z^4 - 8},\ C = \frac{Q}{1 - PQ},\ Q = \frac{10z^3 - 8z^2}{-11z^3 + 10}\frac{z}{2z - 1}. \tag{5.78}$$

A disturbance w is introduced at the plant input. Three separate estimation techniques are used to estimate the system states and input disturbance. The first method corresponds to the constrained MHE with stability guarantee MHE 2, the second method to MHE 1 with the addition of constraints, and the third method to an unconstrained linear Kalman filter. The MHE methods both used a horizon of $p = 5$, and weights $R = 0.01$, $Q = 1$, $\Lambda = 1$, and Σ as in (5.20). The constraints enforced were as follows:

$$0 \leq \hat{w} \leq 1,\ |\hat{y}| \leq 0.5,\ |\Delta\hat{x}| \leq 0.05. \tag{5.79}$$

The disturbance w, as well as the estimated disturbances for each of the three methods, is shown in Figure 5.1. Because the constraints will not allow the estimate $\hat{w}$ to fall below zero, the constrained method with stability guarantee, MHE 2, reflects the disturbance more accurately than the Kalman filter. When constraints are added to MHE 1, the disturbance estimate does not converge asymptotically, although it does not diverge due to the constraints on $\hat{w}$. By relaxing these constraints while leaving the others unchanged, the estimate can be made to diverge.

5.6.2 Chemical reactor

This example demonstrates how including constraints can increase the robustness of fault detection using linear models. The detection of faulty operation for a chemical

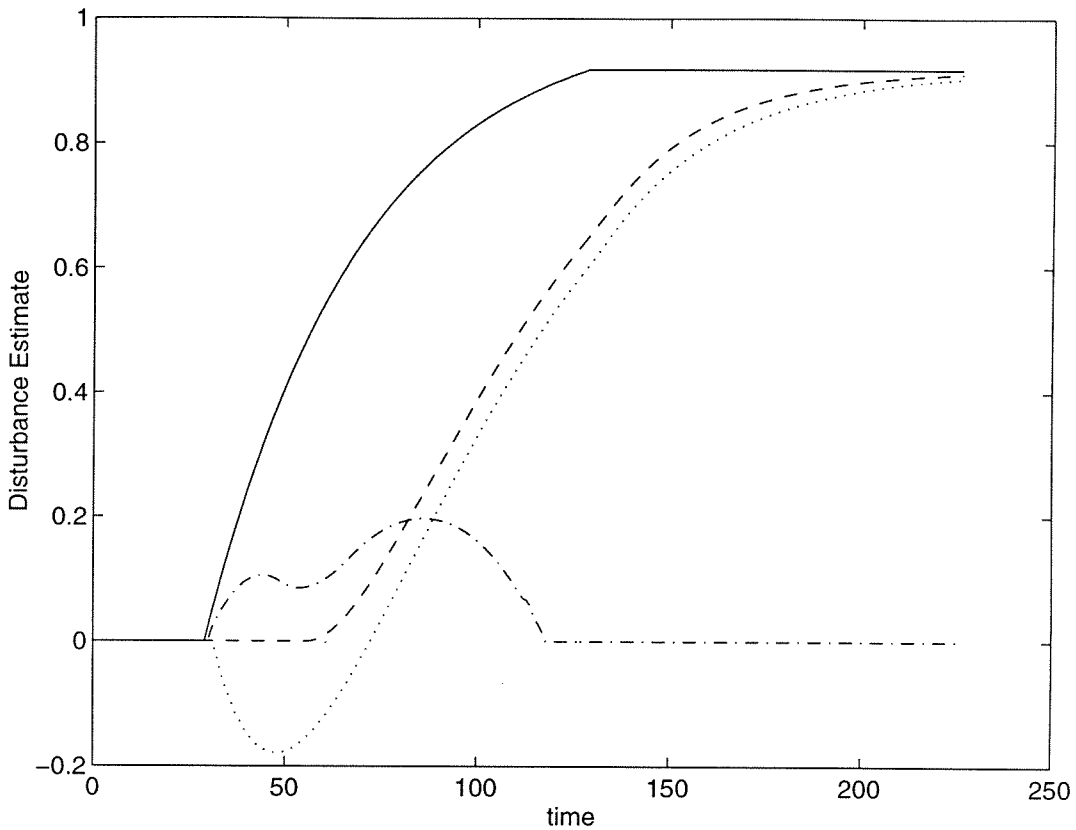

Figure 5.1: Disturbance estimates. Solid line – actual disturbance w, dashed line – MHE 2 (guaranteed stability), dash-dot line – MHE 1 (no stability guarantee), dotted line – unconstrained Kalman filter

reactor is considered. Since the reactor is nonlinear, one expects that by using a nonlinear model rather than a linear model, better estimates would be obtained; however, in contrast to this simple example, frequently a nonlinear model based on first principles is not available, and one must rely upon a linear input/output model obtained from identification experiments. In this example, a linear model obtained from linearization about the nominal operating point will be used in the estimation schemes. It will be shown that for this example, constrained estimation using a linear model is significantly more robust to model uncertainties due to nonlinearities than the linear Kalman filter.

Consider a continuous stirred tank reactor (CSTR) in which chemical species A is converted to species B through an irreversible first order liquid phase reaction

Parameter	Value
$\frac{F_{in}}{V}$	0.5
T_{in}	298
T_c	373
$\frac{-\Delta HC}{\rho C_p}$	10
$\frac{UA}{\rho C_p V}$	0.92
A_1	2.4e13
A_2	1.0e3,

Table 5.1: Parameter Numerical Values

as shown in Figure 5.2. The reactor temperature is regulated by passing a heat transfer fluid at temperature T_c through a jacket. Assume that the reactor volume V is maintained constant by appropriately adjusting the outlet, and that the physical properties (molar concentration C, heat capacity C_ρ, density ρ) of the inlet and outlet streams are the same. An energy and material balance leads to the following governing differential and algebraic equations:

$$\begin{aligned}
\frac{dT}{dt} &= \left(\frac{F_{in}}{V}\right)(T_{in} - T) + \left(\frac{-\Delta HC}{\rho C_p}\right)k(T)x_A + \left(\frac{UA}{\rho C_p V}\right)(T_c - T), \\
\frac{dx_A}{dt} &= \frac{F_{in}}{V}(1 - x_A) - k(T)x_A, \\
k(T) &= A_1 \exp(\frac{-A_2}{T}),
\end{aligned} \tag{5.80}$$

where T is the reactor temperature, x_A the mole fraction of A in the reactor, $k(T)$ the Arrhenius reaction constant, F_{in} the inlet volumetric flow rate, ΔH the heat of reaction, and UA the heat transfer coefficient times area. The parameter values used for this example are given in Table 5.1, which results in the steady state operating condition $T = 350$, and $x_A = 0.051$.

Imagine that two types of faulty performance frequently affect the system. Fault 1, f_1, corresponds to the clogging of the inlet pipe so that the inlet flow rate F_{in} decreases from its nominal value of $0.5V$. Fault 2, f_2, is related to the heat transfer fluid. This fluid is normally maintained at constant temperature $T_c = 373K$ by being passed through an external heater. When the heater is not functioning properly, the

temperature of the heat transfer fluid to the reactor jacket drops, that is $T_c < 373$. Each of the two faults may be modeled by *decreases* in one of the system variables F_{in} or T_c.

By linearizing the system equations about the nominal operating point, and defining the states $x = [\Delta T, \Delta x_A]^t$, $w = [\Delta \frac{F_{in}}{V}, \Delta T_c]^t$, the following linear system is obtained:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -1.0325 & 93.2043 & -51.9334 & 0.9200 \\ 0.0388 & -9.8204 & 0.9491 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{5.81}$$

By discretizing this model, a system in the form of (5.3) with $x = [\Delta T, \Delta x_A]$ and $w = [\Delta \frac{F_{in}}{V}, \Delta T_c]$ is obtained. The constraints which should be enforced include

$$w_1 \leq 0, \; w_2 \leq 0, \; -w_1 \leq 0.5, \tag{5.82}$$

$$-x_2 \leq 0.051, \; x_2 \leq (1 - 0.051). \tag{5.83}$$

The first two constraints imply that F_{in} and T_c can only decrease from their nominal values. The third constraint implies that F_{in} must be positive, and the final two express the requirement that the mole fraction of A is in the interval $[0, 1]$.

For the objective function, the outputs $y = [\Delta T, \; \Delta x_A]$ and the innovations v must be scaled. One approach to choosing the scaling consists of finding diagonal matrices D_1 and D_2 such that the steady state values of $\|D_1 y\|$ and $\|D_2 w\|$ are close to the same when a step in w (or an impulse in v) occurs. Let P_0 be the steady state value of the impulse response matrix for the system (5.81). Then at steady state,

$$D_1 y = \left(D_1 P_0 D_2^{-1}\right) D_2 w, \tag{5.84}$$

which gives the bounds

$$\underline{\sigma}\left(D_1 P_0 D_2^{-1}\right) \|D_2 w\| \leq \|D_1 y\| \leq \bar{\sigma}\left(D_1 P_0 D_2^{-1}\right) \|D_2 w\|. \tag{5.85}$$

The gap between the upper and the lower bound is as small as possible when the condition number of $D_1 P_0 D_2^{-1}$ is minimized. The condition number is independent of constant scaling, so it is possible to further specify $\|D_1 P_0 D_2^{-1}\| = 1$. This minimization is easily calculated [15], and for this example results in

$$D_1 = \text{diag}(1, 189), D_2 = \text{diag}(0.0139, 0.8719). \tag{5.86}$$

This scaling is implemented in the objective function by letting $R = D_1^{-2}$, $Q = D_2^{-2}$.

For each disturbance, three detection schemes were implemented: Constrained MHE 2 using a linear model, with horizon $p = 5$ and weighting matrix $\Lambda = 5I$; Unconstrained linear Kalman filter based on the model linearization at the fault free operating point; and the Unconstrained Extended Kalman filter. For the schemes requiring a linear model, the model (5.81) was transformed to a discrete difference equation using a zero order hold transformation.

The results for a step fault in the input flow rate (f_1) are shown in Figures 5.3 and 5.4. Two different size steps were simulated using the full nonlinear model. The simulated output was sampled at a rate of 0.1 sec^{-1}. For the step of magnitude 0.01 (a 2% decrease in the flow), each detection scheme correctly estimates Fault 1; however, the Kalman and Extended Kalman filters have larger errors in the estimation of Fault 2 than the MHE estimator. Of particular concern is the fact that the Kalman filter results in a non-zero asymptotic estimate of Fault 2. For the larger fault of magnitude 0.25 (a 50% decrease in the flow), both the MHE 2 and the Extended Kalman filter adequately diagnose the failure. The Extended Kalman filter has a larger transient error for Fault 2 than the MHE 2, but the MHE 2 has steady state offset. This offset is due to inaccuracies in the model used in the MHE which may be ascribed to nonlinearities, but does not affect the proper diagnosis of Fault 1. The Kalman

filter is also affected by the same modeling errors, but because the constraints are not enforced, the estimated faults result in a poor diagnosis, incorrectly assigning the majority of the poor performance to Fault 2. Since the EKF uses the full nonlinear model, there are no modeling errors for this case. For this example, it is clear that MHE 2 with constraints is more robust to modeling errors than is the Kalman filter.

Figures 5.5 and 5.6 show the results of applying the same detection schemes to step faults in the heating temperature T_c (f_2) of 0.5 and 10 degrees respectively. The results are similar to those for Fault 1.

From the basis of steady state offset, only the Extended Kalman filter gives error free estimation of the disturbance for large disturbances; however, if the cause of the fault is of more interest than its magnitude, one may argue that the performance of MHE 2 is comparable. In addition, the method MHE 2 achieves this level of performance while using a simpler, local linear model, which would be more easily identified from process data.

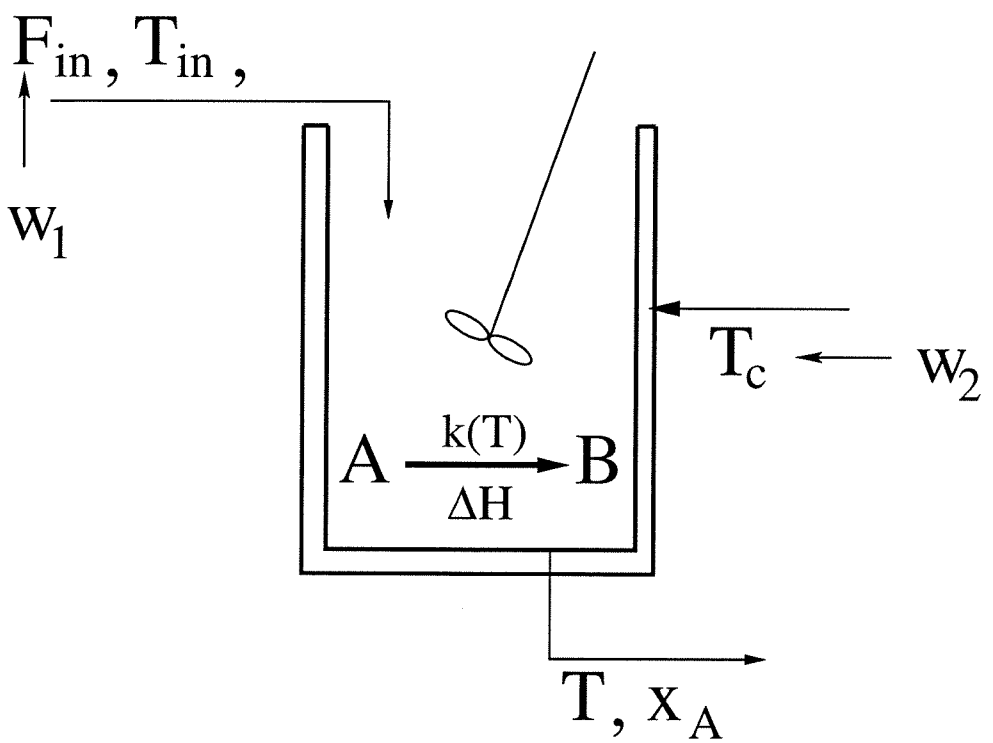

Figure 5.2: Continuously stirred tank reactor

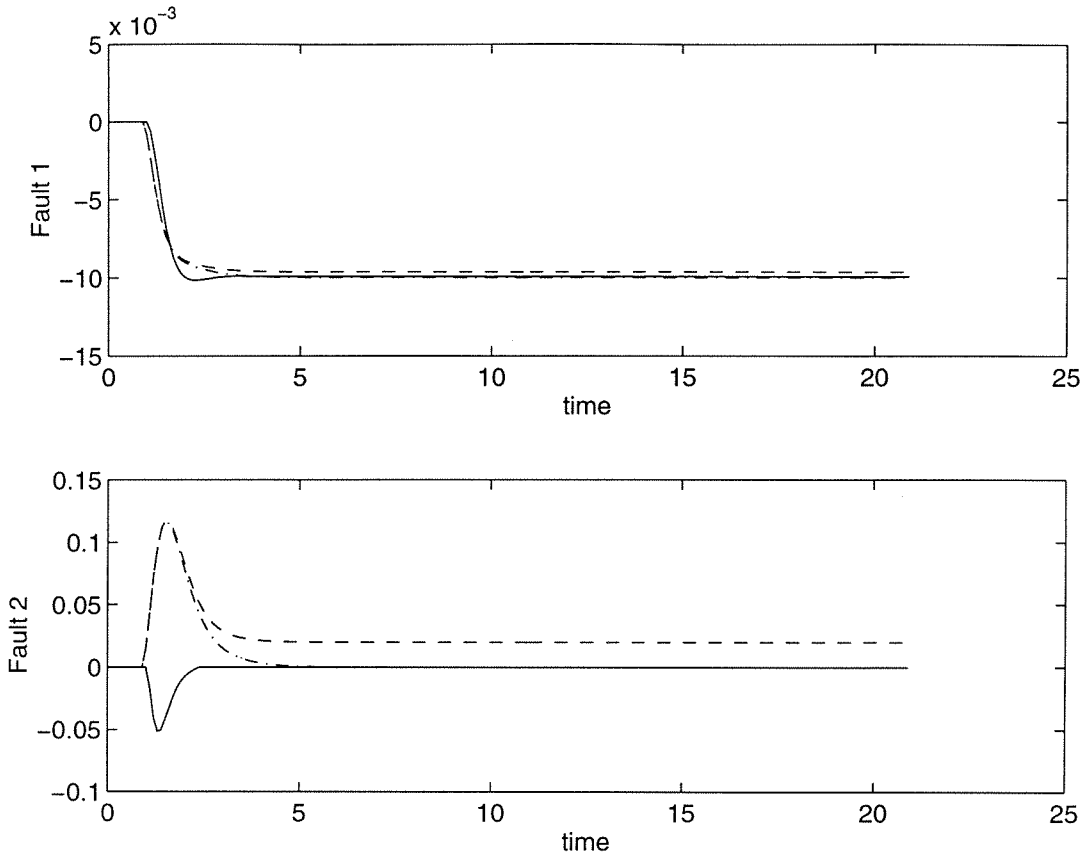

Figure 5.3: Fault estimation for $f = -0.01 f_1$. Solid – MHE 2 scheme, dashed – unconstrained linear Kalman filter, dash-dot – unconstrained extended Kalman filter

5.7 Conclusions

Moving Horizon Estimation (MHE) provides a framework for dynamic estimation of system states and parameters wherein constraints on disturbances and states can be implemented. Including process information in the form of constraints can improve the robustness to model error for estimation schemes. Although under the assumptions which render the Kalman filter optimal the MHE scheme is suboptimal, MHE may outperform "optimal" estimation schemes when the assumptions such as Gaussian distributed white noise do not hold by constraining the estimates to lie within a given region. For the unconstrained MHE, stability can be guaranteed by properly selecting the initial state error weighting matrix Σ; however, the addition of constraints can lead to an unstable estimator.

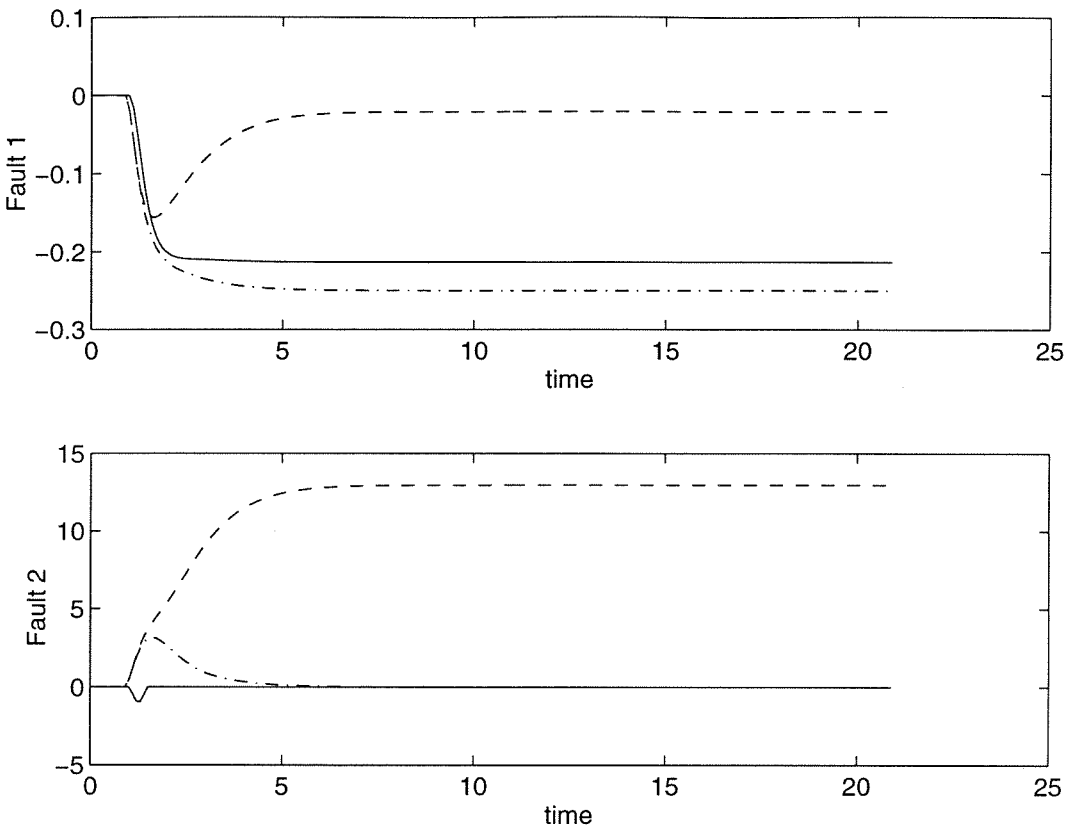

Figure 5.4: Fault estimation for $f = -0.25f_1$. Solid – MHE 2 scheme, dashed – unconstrained linear Kalman filter, dash-dot – unconstrained extended Kalman filter

For the case of constrained MHE, several modifications to the estimation scheme were needed to guarantee stability. First, the set of infinite state constraints must be replaced by a finite set which covers the same space. Under fairly mild assumption on the form of the state constraints, this substitution will be possible as was shown in Lemmas 1 and 2. With the constraints extended over the infinite prediction horizon, it was shown that stability can be guaranteed by including in the MHE objective function an additional penalty on the difference between the final estimated state and an auxiliary state ξ^*, and it was shown how ξ^* can be chosen to guarantee asymptotic convergence of the output estimation error.

Two examples demonstrated the use of the new MHE algorithm. The first example showed that the algorithm MHE 2 maintains a stable estimate on a sample system for which previous versions of constrained MHE diverge. A second example illus-

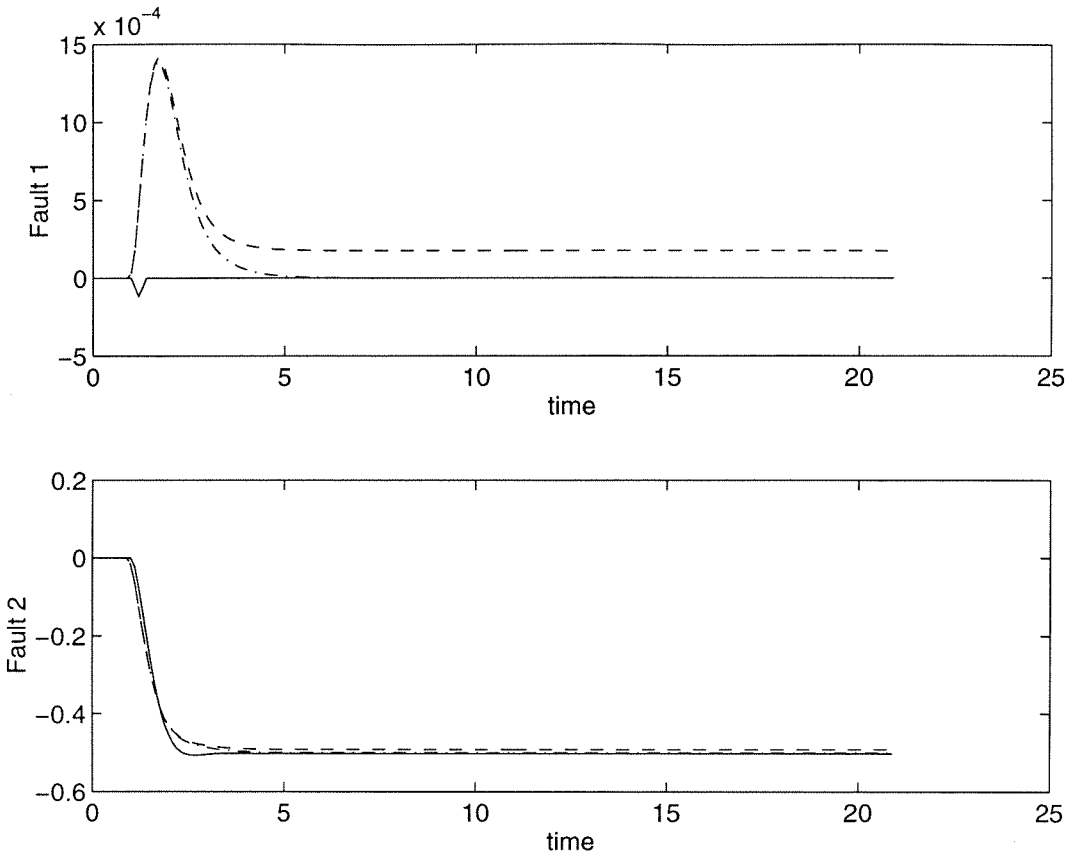

Figure 5.5: Fault estimation for $f = -0.5f_2$. Solid – MHE 2 scheme, dashed – unconstrained linear Kalman filter, dash-dot – unconstrained extended Kalman filter

trated how incorporating constraints on estimated variables can improve robustness to modeling errors.

In Chapter 6, it will be shown how constrained estimation can be used to explicitly account for model uncertainty in fault detection problems and thereby provide further improvements in detection performance.

Appendix: Proof of Lemmas

Proof *Lemma 1* The proof is by induction. Suppose $G_{x,j}A^n\zeta \leq 1$ for all ζ satisfying the first n constraints. Assume the claim holds for all $i \leq m$ where $m > n$, and

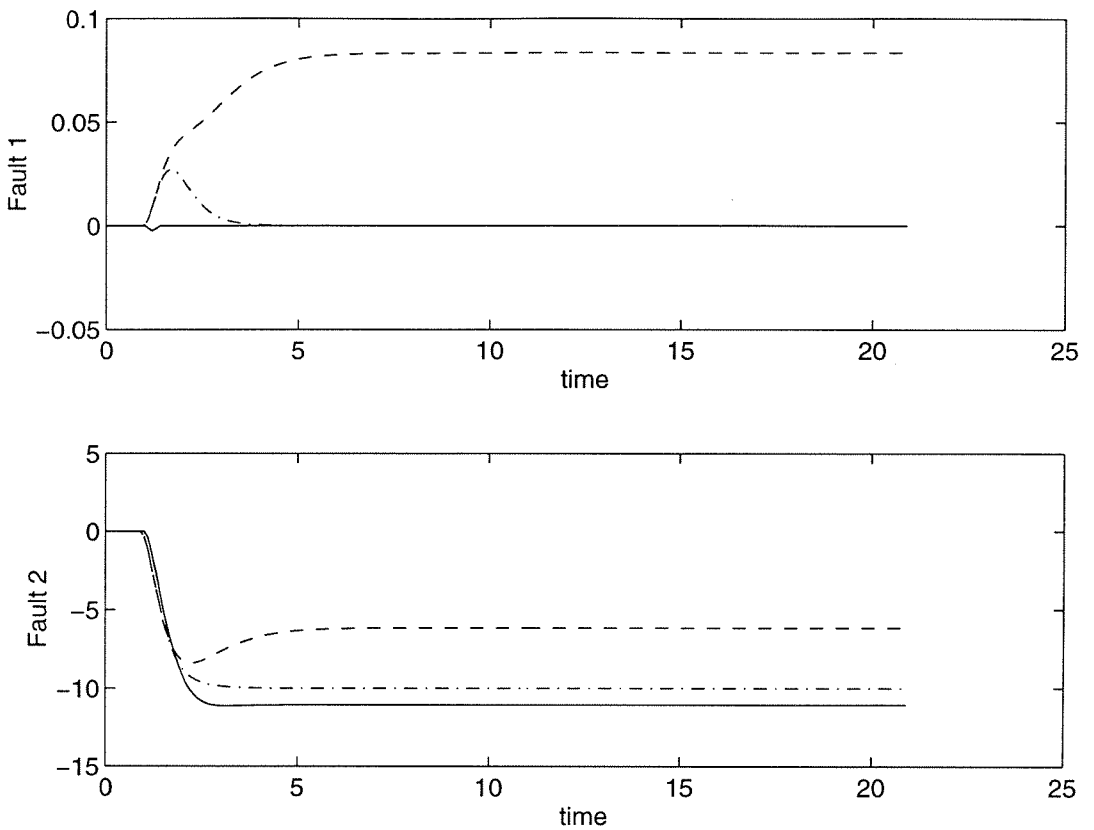

Figure 5.6: Fault estimation for $f = -10f_2$. Solid – MHE 2 scheme, dashed – unconstrained linear Kalman filter, dash-dot – unconstrained extended Kalman filter

consider the maximization

$$\begin{aligned} \max_{\zeta} \quad & G_{x,j}A^{m+1}\zeta \\ \text{s.t.} \quad & G_{x,j}A^{l}\zeta \leq 1,\ l = 0, \ldots, m. \end{aligned} \tag{5.87}$$

By removing the constraint corresponding to $l = 0$, an upper bound may be achieved. Substituting $\eta = A\zeta$, the bounding problem becomes

$$\begin{aligned} \max_{\eta} \quad & G_{x,j}A^{m}\eta \\ \text{s.t.} \quad & G_{x,j}A^{l}\eta \leq 1,\ l = 0, \ldots, m-1 \\ & \eta \in \text{ Range } A, \end{aligned} \tag{5.88}$$

which is bounded above by the problem obtained by removing the constraint that η lie in the range of A. This is exactly the linear program corresponding to $i = m$, which is less than one by induction. It follows that $G_{x,j}A^{m+1}\zeta \leq 1$.

Now suppose the bound 1 is replaced by h'_j, where $h'_j \geq 0$. Then it follows that $G_{x,j}A^i x \leq h'_j$ for all i given that it holds for $i < n$. This holds for any $h'_j \geq 0$, and in particular for h'_j given by

$$h'_j = h_{\xi,j} - G_{x,j}x^* - G_{w,j}w^*.$$

Finally, by using (5.35), note that

$$G_{\xi,j}\hat{\xi}(k+i|k) = G_{x,j}x^*(k) + G_{w,j}w^*(k) + G_{x,j}A^i\left(\hat{x}(k|k) - x^*(k)\right), \tag{5.89}$$

which completes the proof. ■

Proof *Lemma 2* Since A is stable, there exists a positive definite X such that

$$A^tXA - X + I = 0,\ X \geq I. \tag{5.90}$$

Using this X, it follows that

$$\zeta^tA^tXA\zeta = \zeta^tX\zeta - \zeta^t\zeta \leq \left(1 - \frac{1}{\bar{\sigma}(X)}\right)\zeta^tX\zeta. \tag{5.91}$$

Defining $\lambda^2 = \left(1 - \frac{1}{\bar{\sigma}(X)}\right)$ and repeating the above procedure, it follows that

$$\left[A^n\zeta\right]^t X \left[A^n\zeta\right] \leq \lambda^{2n}\zeta^tX\zeta. \tag{5.92}$$

Now by using the fact that

$$\underline{\sigma}(X)\|\zeta\|^2 \leq \zeta^tX\zeta \leq \bar{\sigma}(X)\|\zeta\|^2,$$

it follows that

$$\|A^n\zeta\| \leq \lambda^n\sqrt{\kappa(X)}\|\zeta\|, \tag{5.93}$$

where $\kappa(X)$ is the condition number $\bar{\sigma}(X)/\underline{\sigma}(X)$ and λ is strictly less than 1. Now consider the following inequalities:

$$\begin{aligned} G_{x,j}A^n\zeta &\leq |G_{x,j}A^n\zeta|, && (5.94)\\ &\leq \bar{\sigma}(G_{x,j})\|A^n\zeta\|, && (5.95)\\ &\leq \bar{\sigma}(G_{x,j})\sqrt{\kappa(X)}\lambda^n\|\zeta\|. && (5.96) \end{aligned}$$

Since $\lambda < 1$, there exists an n such that the right-hand term is less than one as long as $\|\zeta\|$ is bounded. ■

Proof *Lemma 3* Suppose at time k, the objective function is minimized by

$$\Delta\hat{\xi}(k+1-p), \hat{v}_\xi(k+1-p|k), \ldots, \hat{v}_\xi(k-1|k), x^*(k) \tag{5.97}$$

and that this solution is used to update $\hat{\xi}(k+1-p|k-1)$ to $\hat{\xi}(k+2-p|k)$. Consider the following feasible solution at time $k+1$

$$\begin{aligned} \Delta\hat{\xi}(k+2-p) &= 0 && (5.98)\\ \hat{v}_\xi(i|k+1) &= \hat{v}_\xi(i|k),\ i = k+2-p, \ldots, k-1 && (5.99)\\ \hat{v}_\xi(k|k+1) &= 0, && (5.100)\\ x^*(k+1) &= x^*(k). && (5.101) \end{aligned}$$

The minimum value $\Phi(k+1,p)$, $\tilde{\Phi}(k+1,p)$, is bounded above by the value obtained from this feasible solution:

$$\begin{aligned} \tilde{\Phi}^2(k+1,p) \leq & \sum_{i=k+2-p}^{k} \|\hat{v}_y(i|k)\|^2_{R^{-1}} + \|x^*(k) - \hat{x}(k+1|k)\|^2_P \\ & + \|y(k) - C_\xi\xi^*(k) + y(k+1) - y(k)\|^2_\Lambda \\ & + \sum_{i=k+2-p}^{k-1} \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}} \end{aligned} \tag{5.102}$$

$$\leq \tilde{\Phi}^2(k,p) - \Psi'(k+1-p|k)$$
$$+2\|y(k) - C_\xi \xi^*(k)\|_\Lambda \|y(k+1) - y(k)\|_\Lambda$$
$$+\|y(k+1) - y(k)\|_\Lambda^2 \tag{5.103}$$

where $\Psi'(i|k)$ is given by

$$\Psi'(i|k) = \|\Delta\hat{\xi}(i)\|^2_{\Sigma^{-1}} + \|\hat{v}_y(i|k)\|^2_{R^{-1}} + \|\hat{v}_\xi(i|k)\|^2_{Q^{-1}}. \tag{5.104}$$

In going from (5.102) to (5.103), the facts that $\hat{x}(k+1|k) - x^*(k) = A(\hat{x}(k|k) - x^*(k))$ and $A^tPA + C^tR^{-1}C \leq P$ were used. By defining the nonnegative quantity $\Psi^*(k)$

$$\Psi^*(k) = \tilde{\Phi}(k,p) - \left(\tilde{\Phi}^2(k,p) - \Psi'(k+1-p|k)\right)^{1/2}, \tag{5.105}$$

the triangle inequality together with the fact that

$$\|y(k) - C_\xi \xi^*(k)\|_\Lambda \leq \left(\tilde{\Phi}^2(k,p) - \Psi'(k+1-p|k)\right)^{1/2}$$

imply that

$$\tilde{\Phi}(k+1,p) \leq \tilde{\Phi}(k,p) - \Psi^*(k) + \|y(k+1) - y(k)\|_\Lambda. \tag{5.106}$$

Now denote by $\tilde{\Phi}'$ the quantity

$$\tilde{\Phi}' = \tilde{\Phi} + \sum_{j=0}^{\infty} \|y(k+j+1) - y(k+j)\|_\Lambda \tag{5.107}$$

which is guaranteed to be finite by the convergence properties of y. It then follows that

$$\tilde{\Phi}'(k+1,p) + \Psi^*(k) \leq \tilde{\Phi}'(k,p), \tag{5.108}$$

which implies that $\tilde{\Phi}'$ is a non-increasing sequence bounded below, and thus must converge to a limit. Therefore,

$$\Psi^*(k) \to 0, \tag{5.109}$$

but this can only be the case when (5.58)-(5.60) hold. Since (5.59)-(5.60) imply that

$\hat{\xi}(k+1-p|k)$ converges to a constant value, $\Psi^*(k)$ can only approach zero if $y^*(k)$ also approaches a constant value, which establishes the limit (5.61) and completes the proof. ■

Proof *Lemma 4* Since $\hat{\xi}(k+1-p|k) = \hat{\xi}(k+1-p|k-1) + \Delta\hat{\xi}(k+1-p)$, (5.59) implies that for any ϵ there exists a K such that for all $k > K$

$$\hat{\xi}(k+1-p|k) = \hat{\xi}(k+1-p|k-1) + O(\epsilon). \tag{5.110}$$

Combining (5.110) with (5.60) and (5.61) implies that

$$C_\xi\hat{\xi}(k+2-p|k) = C_\xi\hat{\xi}(k+1-p|k-1) + O(\epsilon) \tag{5.111}$$

which in turn implies that

$$\hat{\xi}(k+2-p|k) = \hat{\xi}(k+1-p|k-1) + O(\epsilon) + N(k), \tag{5.112}$$

where $C_\xi N(k) = 0$. It now suffices to show that $N(k)$ must converge to zero. The convergence (5.59) and (5.60) along with (5.111) imply that for large enough $k > K$,

$$\begin{aligned}
\hat{\xi}(k+3-p|k+1) &= A_\xi\hat{\xi}(k+2-p|k) + O(\epsilon), && (5.113)\\
&= A_\xi\hat{x}(k+1-p|k-1) + A_\xi N(k) + O(\epsilon) && (5.114)\\
&= \hat{x}(k+2-p|k) + A_\xi N(k) + O(\epsilon). && (5.115)
\end{aligned}$$

By evaluating (5.111) at time $k+1$ and substituting (5.115) for $\hat{x}(k+3-p|k+1)$ on the left-hand side, it follows that

$$C_\xi\hat{x}(k+2-p|k) + C_\xi A_\xi N(k) = C_\xi\hat{x}(k+2-p|k) + O(\epsilon), \tag{5.116}$$

which implies that $C_\xi A_\xi N(k) \approx O(\epsilon)$. Repeating this procedure,

$$\begin{bmatrix} C_\xi \\ C_\xi A_\xi \\ \vdots \\ C_\xi A_\xi^{n-1} \end{bmatrix} N(k) \approx O(\epsilon). \tag{5.117}$$

Observability of (C_ξ, A_ξ) implies the matrix on the left has linearly independent columns so that $N(k) \to 0$. ■

Proof *Lemma 5* The proof is by induction. In the previous proof, it was established that

$$\hat{\xi}(k+1-p|k) \rightarrow \hat{\xi}(k+1-p|k-1), \tag{5.118}$$

$$\hat{\xi}(k+2-p|k) \rightarrow \hat{\xi}(k+1-p|k-1). \tag{5.119}$$

In addition, the limit

$$\hat{\xi}(k+2-p|k) \rightarrow \hat{\xi}(k+2-p|k-1) \tag{5.120}$$

follows directly from the fact that

$$\begin{aligned} \hat{\xi}(k+2-p|k) &= A_\xi \hat{\xi}(k+1-p|k-1) + A_\xi \Delta\hat{\xi}(k+1-p) + B_\xi v_\xi(k+1-p|k) \\ &= \hat{\xi}(k+2-p|k-1) + O(\epsilon), \end{aligned} \tag{5.121}$$

which implies that

$$\hat{\xi}(k+i-p|k) \rightarrow \hat{\xi}(k+i-p|k-1), \; i \leq 2. \tag{5.122}$$

Now assuming (5.63) and (5.122) hold for $i \leq j$, and substituting $k+1$ for k into

these two relations yields

$$\hat{\xi}(k+1+j-p|k+1) \rightarrow \hat{\xi}(k+2-p|k) \tag{5.123}$$

$$\hat{\xi}(k+1+j-p|k+1) \rightarrow \hat{\xi}(k+1+j-p|k) \tag{5.124}$$

which together with (5.119) imply that

$$\hat{\xi}(k+(j+1)-p|k) \rightarrow \hat{\xi}(k+1-p|k-1), \tag{5.125}$$

so that the claim holds for $j+1$. ■

Chapter 6 Application of Moving Horizon Estimation to Robust Fault Detection

Summary

For a class of model uncertainty descriptions, plant/model mismatch can be directly incorporated into a model based fault detection scheme using Moving Horizon Estimation. The model uncertainty is represented by a set of bounded parameters which can be used to alter the dynamics of the model through injection of the measured output as well as inputs. The uncertainty class includes gain uncertainty as well as uncertainty in pole locations. Using a bank of filters, detection of exclusive fault scenarios can be accomplished. The proposed method is compared to other methods employing an adaptive threshold, and is demonstrated on a simulation example of a cold tandem steel mill.

6.1 Introduction

One approach to fault detection and isolation in dynamical systems consists of using linear observers or Kalman filters to estimate the system state [110]. Information about the normal operating behavior of the system can be used along with the state estimate to deduce faulty operation. Methods for deriving linear observers assume an accurate representation of the system, including models for nominal disturbances. Although detection schemes using such observers work very well when the model exactly describes the system, small inaccuracies severely limit the performance. For example, in Chapter 4 it was shown that for some small model perturbations which

would have little effect on control performance, linear observer based detection methods will perform poorly. As mathematical descriptions of physical processes rarely provide a precise description of the system behavior, the use of linear observers in fault detection has its limits.

Traditionally, fault detection is divided into a two step approach. The first step involves generating residual signals which are small when the process is operating normally. These residual signals are often obtained from the output prediction error derived from an observer as described above. Then a decision algorithm based on sequential probability ratio tests, such as CUSUM, is used to determine whether the residual is different enough from zero to indicate a fault. Attempts at robust fault detection have usually focused on either one or the other of these steps, as is evident from the review articles [85]. [35], [62], and [18] address the issue of robust residual generation by choosing linear combinations of the output variables which are insensitive to modeling uncertainties. Recently, an optimal residual decoupling method has appeared [41]. One significant drawback to these approaches lies in the fact that frequently a reduction in the sensitivity to model uncertainty can only be accomplished with a high loss in sensitivity to faults.

In the decision phase, the methods of robustness which have received the most attention are based on the use of adaptive thresholds [33, 29, 19, 44, 46, 108]. In this approach, the threshold becomes a function of a measurable quantity, usually either a reference signal or control moves. The threshold for each residual is varied according to the measured variables.

In Chapter 5, estimation of dynamic states and parameters using Moving Horizon Estimation was discussed. In particular, modifications to previous algorithms were proposed in order to guarantee stability of the estimation scheme. In this chapter, an alternative approach to achieving robustness in fault detection is pursued using the MHE framework. Following a traditional approach, a potential fault is modeled by including auxiliary states which represent possible faults. Then process measurements are used to estimate the states by minimizing a weighted sum of the fault vectors and errors between estimated and measured outputs using Moving Horizon

Estimation. Unlike other state observers, the Moving Horizon Estimator has the advantage of being able to incorporate constraints. This feature is used to enhance the robustness of the fault detection scheme by defining the model sets via constraints, and thereby minimizing the Moving Horizon Estimation objective over the entire set of models. For a class of uncertainty descriptions, the resulting optimization problem is a quadratic program. Although this approach could be interpreted as an adaptive threshold selection technique, in contrast to previously proposed methods, the fault estimates are not generated independently of the threshold selection process as the on line optimization couples the two problems.

The proposed robust fault detection scheme is illustrated through a simulation example on a cold tandem steel mill.

6.2 Robust fault detection

In this Section, model based fault detection for uncertain process models is considered. The general approach can be summarized as follows. Using a process model to represent a dynamical system, an estimator is used to form estimates of the system state and possible faults from output measurements. The estimator attempts to keep output errors small by introducing when necessary non-zero fault signals. In the case of a single, precise model describing the system and no measurement noise, in the no-fault case the estimator should be able to exactly match the measured output without introducing any non-zero fault signals. On the other hand, when a fault which affects the output enters the system, the estimator can no longer match the output without introducing a fault signal. By using a single nominal model in the case wherein plant/model mismatch occurs, even in the no-fault case the estimator will often introduce non-zero fault estimates in order to account for discrepancies between the modeled output and the measured output.

In order for the estimation scheme to be able to account for mismatch between the process and the model, this approach must be modified. Rather than using a single process model, let us assume that the physical process can be well modeled

by some model lying within a bounded set of models Π. Then, given the initial state of the system x_0 and measurements up to some time $T-1$, an estimator is used to minimize the output error. As before, the estimator can use non-zero fault signals to reduce the output error, but unlike the single model case, the estimator performs the minimization not only over possible fault signals but also over the set of possible models. Letting $g_k(x_0, u, w, P)$ denote the function which maps the initial state, inputs u, and fault w through the model P to the output $y(k)$, the estimator performs the following minimization:

$$\begin{array}{l} \min \\ P \in \Pi \\ w \in W \end{array} \sum_{i=0}^{T-1} \|y(i) - g_i(x_0, u_0^i, w_0^i, P)\|_V^2 + \|v_w(i)\|_Q^2. \tag{6.1}$$

where $\|\cdot\|_V$ denotes the weighted Euclidean norm

$$\|x\|_V^2 = x^T V x, \tag{6.2}$$

$v_w(i)$ denotes changes in w, $w(i)-w(i-1)$, $u_{t_1}^{t_2}$ denotes the sequence of input variables

$$u_{t_1}^{t_2} = [u(t_1)^T, ..., u(t_2)^T]^T, \tag{6.3}$$

and $w_{t_1}^{t_2}$ is defined analogously to $u_{t_1}^{t_2}$.

This approach is similar to the time domain robust model validation techniques introduced by Poola et al.[86]. In their work, tests are developed to determine if some model within a given set of models is consistent with the input/output data. These tests are based on determining the smallest perturbation from a nominal model which would account for the observed behavior. In the current approach, the set of models is fixed, and the detection scheme attempts to minimize a weighted sum of the output error and fault value by choosing some model within the model set.

Although, in general, the above optimization will be difficult to carry out, for a certain class of uncertainty structures it can be recast as a quadratic programming

problem. Consider the case where the set of models can be represented in state space by the following equations:

$$x(k+1) = Ax(k) + B_u u(k) + B_w w(k) + \sum_{i=1}^{n_\theta} \theta_i \left(F_i u(k) + G_i y(k)\right) \quad (6.4)$$

$$y(k) = Cx(k) + D_u u(k) + D_w w(k), \quad (6.5)$$

where θ_i are scalar variables satisfying the bounds

$$|\theta_i| \leq 1. \quad (6.6)$$

When the system can be described by a model of this form, the mapping g is bilinear in the variables θ_i and the pair $[u, y]$. Since at any given time, all past values of u and y are known, the mapping g is a time varying linear function in the decision variables θ, and therefore the optimization (6.1) has a quadratic objective.

The uncertainty description (6.4) can be considered a special case of the more general linear fractional transformation uncertainty which has been widely used in control design:

$$x(k+1) = Ax(k) + Bu(k) + Gp(k) \quad (6.7)$$

$$y(k) = Cx(k) \quad (6.8)$$

$$q(k) = C_q x(k) + D_q u(k) \quad (6.9)$$

$$\|p(k)\| \leq \|q(k)\| \quad (6.10)$$

with the restrictions that the rows of C_q are linear combinations of the rows of C, and that the relation between p and q are structured according to the structure imposed by the θ variables. Although the uncertainty description (6.4) is less general, it is still useful for a wide range of problems. For example, this uncertainty structure can account for gain uncertainty through the terms involving u, as well as uncertainty in the poles through the terms involving y.

To implement the above algorithm in real time, the size of the optimization prob-

lem grows as the number of measurements increase, and soon becomes computationally intractable. To circumvent this difficulty, a moving horizon estimation (MHE) scheme is employed [92]. In the moving horizon scheme, the horizon length p does not change, but rather as a new measurement becomes available, the starting point of the horizon moves forward one step. At time k, the state at the start of the horizon can be obtained from a smoothed state $\hat{x}(k-p+1|k-1)$. Since this smoothed state is not precise, it is advantageous to modify the optimization to allow for deviations from this value by introducing an additional decision variable $\Delta\hat{x}$.

Let us consider the overall structure of the model for the system driven by both known signals u and y and unknown faults. Since the system is linear, let us divide it into two components with states x_u and x_w. The state x_u represents the position of the state due to the inputs u and the model uncertainties θ, whereas the state x_w represents the position of the state due to faults and errors in the initial state estimate. Then there exist mappings g^u and g^w such that

$$y(j) = g_j^u(x_u(k-p+1), u_{k-p+1}^j, y_{k-p+1}^{j-1}, \theta) + g_j^w(x_w(k-p+1), w_{k-p+1}^j). \tag{6.11}$$

Here, $y_{t_1}^{t_2}$ is defined in an analogous fashion to $u_{t_1}^{t_2}$. Since x_u and x_w are generally unknown, they must be estimated. In the MHE framework, the estimated states $\hat{x}_u$ and $\hat{x}_w$ are updated according to

$$\begin{aligned} \hat{x}_u(k-p+1) &= A\hat{x}_u(k-p) + B_u u(k-p) \\ &\quad + \sum_{i=1}^{n_\theta} \theta_i \left(F_i u(k-p) + G_i y(k-p)\right) \qquad (6.12) \\ \hat{x}_w(k-p+1|k) &= A\hat{x}_w(k-p|k-1) + A\Delta\hat{x} + B_w\hat{w}(k-p|k), \qquad (6.13) \end{aligned}$$

and the smoothed fault estimate obeys the update relation

$$\hat{w}(k-p+1|k) = \hat{w}(k-p|k-1) + \hat{v}_w(k-p). \tag{6.14}$$

The following optimization problem is then solved:

$$\min \sum_{j=k-p+1}^{k} \left(\|y(j) - g_j^u(\theta) - g_j^w(\Delta\hat{x}, \hat{v}_w)\|_V^2 + \|\hat{v}_w(j)\|_Q^2 \right) + \|\Delta\hat{x}\|_P^2 \tag{6.15}$$

subject to the constraints

$$|\theta_i| \leq 1. \tag{6.16}$$

Here, for simplicity in notation, the dependence of g on the non-decision variables $\hat{x}_u(k-p+1)$, $\hat{x}_w(k-p+1|k-1)$, $\hat{w}(k-p+1|k-1)$, y_{k-p+1}^k, and u_{k-p+1}^k has been neglected.

For stable systems, since the input signals u and the measured outputs y are bounded, the state estimate $\hat{x}_u$ remains bounded for any bounded values of θ. Therefore, the overall stability of the state updates depends on the stability for updating $\hat{x}_w$. In the case wherein no constraints are applied to the decision variables $\Delta\hat{x}$ or $\hat{v}_w$, this can be accomplished by properly selecting the weighting matrix P after specifying V and Q. If constraints are enforced on these decision variables, the objective function should be modified as described in Chapter 5. Although it is convenient to visualize the process by breaking down the state into x_u and x_w, the algorithm may easily be implemented using only the combined state $x = (x_u + x_w)$.

6.3 Fault detection using a bank of robust MHE filters

In the case where there is zero initial state error, i.e., $x(k-p+1) = \hat{x}_u(k-p+1) + \hat{x}_w(k-p+1|k-1)$, no disturbances, no measurement noise, and the physical system is exactly modeled by some model in the set Π, one of two situations arises. If there is no-fault, then the objective function used in the MHE problem can be made identically zero by properly choosing the parameters θ while keeping the decision variables corresponding to faults zero. When a fault is present in the system, a non-zero value of the fault variables will result in a smaller objective. Therefore, in this

idealized case, faults are successfully diagnosed.

In reality, the idealized conditions described above do not hold. Measurement noises will be present, random disturbances will affect the system, and the smoothed state estimate $\hat{x}_w$ may be inaccurate. In this situation, rather than determining if some non-zero fault reduces the output error, one wishes to determine if the amount which the residual would be decreased by such a fault is significant considering the measurement noise and the uncertainty in the initial state. To address this problem, an alternative approach is taken using a bank of MHE filters.

Banks of filters have been employed for detection of faults for over two decades. Willsky [110] provides a description of the general approach, as well as an early survey of applications of Kalman filter banks in fault detection. Generally, the bank of filters consists of a filter for each fault type, including the no-fault case, and for each possible failure time. Therefore, for a system with r possible faults and T measurements, there would be $(r+1)^T$ filters, each one corresponding to a particular failure occurring at a particular time. Since this general formulation leads to an exponential growth in the number of filters, a number of approximate techniques have been proposed. One of the most common involves assuming that shifts between the various faults occurs only once every N steps. At the end of the N steps, one of the fault scenarios is accepted, typically using sequential probability ratio tests (SPRT), and the state of each of the $r+1$ filters is set to the state of the filter corresponding to the accepted fault scenario. Frank [34] discusses more recent filter bank approaches, including the *dedicated observer scheme* and the *generalized observer scheme*. Recent methods and applications using banks of filters have also appeared [20, 21, 36, 82].

The detection procedure using a bank of MHE filters follows. Minimally, there must be two filters, one corresponding to the no-fault case, and one corresponding to the faulty case. For the no-fault case, a filter is used which allows for changes in the initial state $\hat{x}_w(k-p+1)$ but not for changes in the fault from its starting value of $\hat{w}(k-p+1)$. Therefore, the no-fault case does not actually correspond to *no-fault*, but rather to *no change* in the fault. The no-fault MHE filter uses an objective function as in (6.15), but in this case the minimization is performed with the additional equality

constraints $\hat{v}_w = 0$ so that only errors in the initial state estimate are considered.

In the two filter case, the second filter corresponds to the original MHE fault detection filter. If a bank of several filters is used, then the individual filters correspond to the optimization problem (6.15) with a subset of the variables $\hat{v}_w$ constrained to be zero. For example, if a separate filter is used for each fault type, then the filter corresponding to fault i would solve the optimization problem subject to the constraint

$$\hat{v}_{wj} = 0, j \neq i, \tag{6.17}$$

while $\hat{v}_{wi}$ would be free to vary.

To implement the bank of filters, at each time step, an optimization of the form (6.15) is performed for each of the individual filters. If r filters are used, this results in r possible update equations of the from (6.12) through (6.14). In order to decide which update procedure to implement, the output errors, i.e., the first term in (6.15) will be used to determine how well each filter fits the system. Some filters may use more decision variables than others. In particular, the no-fault filter will always use fewer parameters than any of the filters which allows for faults. Because of the added degrees of freedom, the faulty filters will always give a smaller value of the overall objective and therefore a smaller value of the output error norm. Therefore, rather than directly comparing the value of the output variance for each filter, the performance of each filter should be determined using an information based criterion which adds penalties for extra parameters, such as Akaike's information criterion (AIC) [94]. This criterion can be considered to be an asymptotic approximation to a cross-validation test (see for example [98]), and is given by

$$AIC = N \log \hat{V} + 2n_p, \tag{6.18}$$

where $N = n_y p$ is the number of scalar output variables in the estimation horizon (n_y is the number of measurements made at each sampling time), $\hat{V}$ is the output error

term

$$\hat{V} = \sum_{j=k-p+1}^{k} \|y(j) - g_j^u(\theta) - g_j^w(\Delta\hat{x}, \hat{v}_w)\|_V^2 \tag{6.19}$$

and n_p is the number of free variables (those not constrained to zero) over which the objective in (6.15) is minimized. This approach is closely related to model structure determination. Given the data within the considered horizon, the no-fault case can be considered as a data model which only includes parameters for model uncertainty (θ) and initial state error ($\Delta\hat{x}$). On the other hand, the other filters in the bank include these parameters as well as additional parameters to describe changes in the value of faults ($\hat{v}_w$). A change in the fault structure is concluded only when the difference in the residuals is large enough to warrant using the additional parameters.

By implementing a bank of MHE filters, one optimization problem must be solved at each time step for each of the filters. Although this increases the computational cost, each of the resulting optimization problems is smaller than one corresponding to a MHE filter which includes each of the faults. Because each of the optimizations is independent of the others, this procedure can easily be implemented in a parallel architecture, with one processor for each sub-filter. In this fashion, although more hardware is required, the computational time for a bank of MHE filters will be less than for a single MHE with more possible faults.

In summary, in the presence of noises, disturbances, and uncertain initial conditions, a fault detection procedure based on robust MHE consisting of the following steps is proposed:

1. Divide the faults into classes. Perform the optimization step of MHE for each class of faults, as well as for the no-fault case.

2. Compute AIC for each of the MHE filters in the bank.

3. Update the initial conditions according to (6.12) - (6.14) using the decision variables from the optimization which produces the smallest AIC.

4. Move the horizon forward one step.

6.4 Comparison to adaptive threshold methods

The use of an adaptive threshold for fault detection residuals has been proposed by several researchers. Clark [19] employed an empirical adaptive law. Horak [44] used a bounded parameter uncertainty description to calculate at every time instant the extremal values of the parameters which would maximize and minimize the residual. Using these values, a range of outputs was determined such that the residual can be completely explained by parameter uncertainty only if it lies within the determined range. Similarly, the threshold selection method of Emami-Naeini et al. [33] characterized the set of detectable faults as those whose smallest possible effect on the residuals exceeds the largest possible effects due to noise and uncertainty. Isaksson [46] and Ding and Frank [29] used similar concepts based on frequency domain uncertainty descriptions. A time domain approach using a convolution kernel whose frequency domain counterpart bounds model uncertainty was proposed by Weiss [108].

In each of these approaches, an adaptive threshold which varies with the control action must be calculated for each time step. In most of the approaches, this threshold calculation is carried out by performing an optimization over the model uncertainty. In this sense, the threshold selection is similar to the proposed MHE filter scheme, which also performs an optimization over the model uncertainty. However, unlike the other schemes, the fault signal generated by MHE is directly affected by the uncertainty.

One way to view the MHE scheme is as a multi-hypothesis testing problem. Each hypothesis can be stated as follows:

> The data y_{k-p+1}^{k} were generated from the initial condition $\hat{x}_u + \hat{x}_w$ and the input u_{k-p+1}^{k} by some model $P \in \Pi$ and some non-zero value of fault i.

Each MHE filter in the bank corresponds to a different value of i, including the no-fault case, the AIC criterion is employed to decide which hypothesis is accepted, and the initial state $x_w + x_u$ is updated accordingly. On the other hand, the previous approaches using adaptive thresholds have the following interpretation. Suppose there exists a relationship among the measured variables which is nominally zero. Given

that this relationship is non-zero, can this possibly be due to noise and plant/model mismatch? A fault is detected whenever this question cannot be answered in the affirmative.

A significant shortcoming of previous adaptive threshold methods lies in the fact that although the threshold level is computed for the residual, the residual is usually processed via a SPRT to decide if a fault has occurred. For example, if the residual comes from a Kalman filter, it should ideally be zero-mean white noise. To determine if a fault is present, a SPRT which tests for changes in the mean and autocorrelation of the residual should be used. This approach is much more reliable then simply considering bounds on the residual. However, the adaptive threshold approaches merely compute bounds which when exceeded indicate the residual cannot be described by noise and model uncertainty alone. By contrast, the multi-bank scheme proposed in this chapter produces an estimated fault by incorporating the model uncertainty within an information based criterion, AIC. The fault is non-zero only when the information measure calculated by including the model uncertainty so indicates.

6.5 Tandem cold rolling mill example

Cold tandem rolling is an important process in steel manufacturing. In a cold tandem rolling mill, a series of rollers is used to decrease the thickness of a sheet of steel, as depicted in Figure 6.1. In the simulation example, a mill with five stands is used as shown in the figure. Fundamental equations describing the tandem roller can be found in the reference [70]. Variable definitions are summarized in Table 6.1. For each variable, a subscript identifies the stand number so that, for example, h_i denotes the exit thickness from the i^{th} stand. The sheet thickness at the exit is determined by the roll gap and the rolling force:

$$h_i = S_i + \frac{P_i}{K}. \tag{6.20}$$

The rolling force is given by the implicit relationship

$$F_p(P_i, H_1, H_i, h_i, t_{bi}, t_{fi}, \mu_i) = 0. \tag{6.21}$$

The strip velocity is determined by the roll velocity and the forward slip coefficient:

$$v_{fi} = (1 + f_i)V_{Ri}. \tag{6.22}$$

The tension is governed by differential equation

$$\dot{t}_{fi} = \frac{E}{L}\left(v_{b(i+1)} - v_{fi}\right) \tag{6.23}$$

where E is Young's modulus, L is the distance between stands, and v_b is related to v_f through the material balance

$$v_{bi} = \frac{h_i}{H_i} v_{fi}. \tag{6.24}$$

The implicit rolling force function F_p is determined by the following set of equations

$$F_p = P_i - W\kappa_i \bar{k}_i \sqrt{R'_i(H_i - h_i)} D_{pi} \tag{6.25}$$

$$\kappa_i = 1 - \frac{(a-1)t_{bi} + t_{fi}}{a\bar{k}_i} \tag{6.26}$$

$$\bar{k}_i = \alpha\left(\bar{r}_i + \beta\right)^{\gamma} \tag{6.27}$$

$$D_{pi} = 1.08 + 1.79 r_i \sqrt{1 - r_i}\,\mu_i \sqrt{\frac{R'_i}{h_i}} - 1.02 r_i \tag{6.28}$$

$$r_i = \frac{H_i - h_i}{H_i} \tag{6.29}$$

$$r_{bi} = 1 - \frac{H_i}{H_1} \tag{6.30}$$

$$r_{fi} = 1 - \frac{h_i}{H_1} \tag{6.31}$$

$$R'_i = R_i\left(1 + \frac{cP_i}{W(H_i - h_i)}\right) \tag{6.32}$$

Symbol	Physical Meaning
h	Exit thickness
H	Entry thickness
S	Roll gap
P	Rolling force
K	Elastic constant
t_b	Back tension
t_f	Forward tension
μ	Coefficient of friction
$\bar{k}$	Mean resistance to deformation
W	Strip width
κ	Defined by (6.26)
R'	Deformed roll radius
R	Roll radius
D_p	Defined by (6.28)
r	Reduction in thickness
r_b	Total reduction of in-going strip
r_f	Total reduction of outgoing strip
f	Forward slip
ϕ_n	Neutral angle
s_f	Yield stress at entry
s_b	Yield stress at exit
v_f	Strip velocity at exit
v_b	Strip velocity at entry
V_R	Peripheral roll velocity

Table 6.1: Meaning of Variables

$$\bar{r}_i = \begin{cases} r_{fi} & \text{for } i = 1, \\ 0.4r_{bi} + 0.6r_{fi} & \text{for } i > 1 \end{cases} \tag{6.33}$$

The forward slip is governed by

$$f_i = \phi_{ni}^2 \frac{R'_i}{h_i} \tag{6.34}$$

where

$$\phi_{ni} = \sqrt{\frac{h_i}{R'_i}} \tan\left(\sqrt{\frac{h_i}{R'_i}} \frac{H_{ni}}{2}\right) \tag{6.35}$$

$$H_{ni} = \frac{H_{bi}}{2} - \frac{1}{2\mu_i} \log\left(\frac{H_i}{h_i} \frac{1 - \frac{t_{fi}}{s_{fi}}}{1 - \frac{t_{bi}}{s_{bi}}}\right) \tag{6.36}$$

$$s_{fi} = \alpha\,(r_{fi} + \beta)^{\gamma} \tag{6.37}$$

$$s_{bi} = \alpha\,(r_{bi} + \beta)^{\gamma} \tag{6.38}$$

$$H_{bi} = 2\sqrt{\frac{R_i'}{h_i}} \arctan\left(\sqrt{\frac{r_i}{1 - r_i}}\right). \tag{6.39}$$

In the above, α, β, γ, and a are material dependent parameters.

6.5.1 Steady state solution

At steady state, the equation governing tension becomes

$$h_i v_{fi} - h_{i+1} v_{f(i+1)} = 0, \tag{6.40}$$

which can be viewed as a material balance. In addition, the entry and exit strip thickness are related by

$$H_{i+1} = h_i, \tag{6.41}$$

whereas the forward and back tension are related as follows:

$$t_{b(i+1)} = t_{fi}. \tag{6.42}$$

Using the physical parameters found in [70] for strip width, roll radius, elastic constant, and friction coefficient, a rolling schedule with positive gap width at each stand was found as follows. Using (6.41) and (6.42), H_i and t_{bi} are eliminated for $i = 2, ..., 5$. Next the input and output thicknesses, H_1 and h_5, the back tension for the first stand, t_{b1}, and the velocity from the last stand v_{f5} are specified. Then a solution $h_i, S_i, P_i, , v_{fi}, t_{fi}$ to the fundamental equations (6.20), (6.21), and (6.40), which satisfies the constraints $S_i \geq 0$ and $h_{i+1} \leq h_i$ is sought. Since this system is under-determined, sequential quadratic programming was used to find the solution

Parameter	Stand 1	Stand 2	Stand 3	Stand 4	Stand 5
H_i (mm)	3.20	2.50	1.79	1.50	1.33
h_i (mm)	2.50	1.79	1.50	1.33	1.20
t_{bi} (kg/mm^2)	0	26.2	26.2	26.2	26.2
t_{fi} (kg/mm^2)	26.2	26.2	26.2	26.2	26.2
R_i (mm)	273	273	292	292	292
μ_i (-)	0.0714	0.0714	0.0714	0.0714	0.0714
K $(10^3\ kg/mm)$	470	470	470	470	470
S_i (mm)	0.89	0.46	0.36	0.35	0.24
V_{Ri} $(10^4\ mm/sec)$	0.684	1.07	1.27	1.44	1.60
V_{fi} $(10^4\ mm/sec)$	0.784	1.10	1.31	1.47	1.63
P_i/K (mm)	1.61	1.33	1.14	0.98	0.96

Table 6.2: Steady State Operating Point for Tandem Mill

which also minimized the following objective:

$$J = \sum_{i=2}^{5}(P_i - P_{i-1})^2 + \sum_{i=2}^{4}(t_{fi} - t_{f(i-1)})^2 + \sum_{i=2}^{5} t_{fi}^2 + 100\sum_{i=1}^{5}(S_i - 0.5h_i)^2 \qquad (6.43)$$

Then the roll velocity V_{Ri} is found from (6.22) and (6.34).

Data for the steady state operating point is found in the Table 6.2. In addition, the values of the physical parameters were taken from [70] as $\alpha = 84.6$, $\beta = 0.00817$, $\gamma = 0.3$, and $a = 3.0$, which reportedly corresponds to low-carbon steel (C = 0.08%), and the strip width and spacing between stand were taken as $W = 930\ mm$ and $L = 4600\ mm$ respectively. The value $E = 21000\ kg/mm^2$ is used for Young's modulus.

6.5.2 Dynamic solution

The following dynamic relations exist between the variables. The entry thickness at a given stand is related to the exit thickness at the previous stand through the transport delay:

$$H_{i+1}(t) = h_i(t - D_{i,i+1}) \qquad (6.44)$$

where $D_{i,i+1}$ is the time required for the strip to travel the distance between the two stands:

$$D_{i,i+1} = \frac{L}{v_{fi}}. \tag{6.45}$$

Assume that the transport delay for the tension is negligible, so that the steady state relation (6.42) between forward and back tension holds. Finally, assume that the roll velocity V_{Ri} and the roll gap S_i can be manipulated by inputs u_{si} and u_{vi} according to the following dynamics:

$$\dot{S}_i = -\frac{1}{T_s}(S_i - U_{si}) \tag{6.46}$$

$$\dot{U}_{si} = u_{si} \tag{6.47}$$

$$\dot{V}_{Ri} = -\frac{1}{T_v}(V_{Ri} - U_{vi}) \tag{6.48}$$

$$\dot{U}_{vi} = u_{vi} \tag{6.49}$$

where the nominal time constants T_s and T_v used in the simulation were specified as 0.3 and 0.6 seconds respectively. These equations assume that the response of S_i is independent of the rolling pressure. In reality, as the rolling force increases, a larger force will be needed to make changes in S_i. One way to account for this phenomenon will be to consider it as a source of model uncertainty, and will be represented by multiplying u_{si} by a gain $(1 + \theta_i)$. Variations in T_s from its nominal value will also be considered as a source of plant/model mismatch. It will be assumed that V_{R5} is used to control the strip velocity v_{f5} at its setpoint, and is therefore not used to reject disturbances. For this reason, from this point on V_{R5} will not be considered as a free variable.

Because the tension responds much faster to changes than do S_i and V_{Ri}, the dynamic model can be simplified by assuming that the forward tension is always in equilibrium. In this situation, the relation (6.40) can be used. Now the state of the system is completely determined by the controlled variables S_i and V_{Ri}, and by the values of H_i which are easily related to past values of h_i. Given these quantities, h_i, P_i, v_{fi}, and t_{fi} can be determined from Equations (6.20), (6.21), (6.22), and (6.40)

given the boundary values t_{b0}, t_{f5}, v_{f5}, and H_1.

6.5.3 Linearization

A local linearization of the nonlinear tandem rolling model was used to carry out the fault detection and to design a feedback controller. This linearization was based on the linear state equation model (6.46)-(6.49) and a linear approximation to the output equations. In addition to the command signals u_{vi} and u_{si}, the friction coefficients μ_i were considered as input variables in the linearization in order to be able to detect faults corresponding to changes in these parameters. In addition, the signals H_i were considered as inputs to form the linearization.

Using (6.20) and (6.22) to eliminate h_i and v_{fi}, an implicit function for t_{fi} and P_i can be obtained from (6.21), (6.34), and (6.40). Letting $\mathbf{y} = [t_{f1}, ..., t_{f4}, P_1, ...P_5]$, $\mathbf{x} = [V_{R1}, ..., V_{R4}, S_1, ..., S_5]$, $\mathbf{u} = [u_{s1}, ..., u_{s5}, u_{v1}, ..., u_{v4}, H_1, ..., H_5]$, and $\mathbf{w} = [\mu_1, ..., \mu_5]$, the implicit function has the form

$$\mathcal{F}(\mathbf{y}, \mathbf{x}, \mathbf{u}, \mathbf{w}) = 0. \tag{6.50}$$

Then, using the implicit function theorem, the linearization about the steady state operating point is given by

$$\Delta\mathbf{y} = C\Delta\mathbf{x} + D_v\Delta\mathbf{u} + D_w\Delta\mathbf{w}, \tag{6.51}$$

where

$$C = -\left(\frac{\partial\mathcal{F}}{\partial\mathbf{y}}\right)^{-1}\left(\frac{\partial\mathcal{F}}{\partial\mathbf{x}}\right) \tag{6.52}$$

$$D_v = -\left(\frac{\partial\mathcal{F}}{\partial\mathbf{y}}\right)^{-1}\left(\frac{\partial\mathcal{F}}{\partial\mathbf{u}}\right) \tag{6.53}$$

$$D_w = -\left(\frac{\partial\mathcal{F}}{\partial\mathbf{y}}\right)^{-1}\left(\frac{\partial\mathcal{F}}{\partial\mathbf{w}}\right), \tag{6.54}$$

Δ denotes change from steady state values, and $\partial\mathcal{F}/\partial z$ represents the Jacobian matrix

of $\mathcal{F}$ with respect to the variables z evaluated at the steady state operating point. The linearization for h_i is then given by

$$\Delta h_i = \Delta S_i + \frac{1}{K}\left(C_{P_i}\Delta\mathbf{x} + D_{v,P_i}\Delta\mathbf{u} + D_{w,P_i}\Delta\mathbf{w}\right), \tag{6.55}$$

where C_{P_i} is the row of C corresponding to the output P_i, etc.

To simulate closed loop behavior of the system, a feedback controller was designed using $\mathcal{H}_\infty$ design methods. A third order Padé approximation was used for the transport delays relating H_{i+1} to h_i in the dynamic model used for the controller design. The controller used measurement of t_{fi} and h_i to calculate values of u_{si} and u_{vi} which would regulate t_{fi} and h_i to their nominal values.

In order to account for changes in $\mathbf{w}$, corresponding integrating states were added to the linear state model (6.46). Letting A and B be the state space matrices corresponding to the the system (6.46), the augmented linear system can now be described in state space as:

$$\left[\begin{array}{cc|cc} A & 0 & B & 0 \\ 0 & 0 & 0 & I \\ \hline C & D_w & D_v & 0 \end{array}\right]. \tag{6.56}$$

For small changes in μ_i, this local linear model will be valid. Note that the dynamic state equations (6.46) are globally valid since they are linear.

The linear model was transformed using the Tustin bilinear approximation with a sampling time of 0.1 seconds to obtain a discrete time description of the system. Denote the discrete version of (A, B) by (a, b).

6.5.4 Robust fault detection

For the cold tandem steel rolling process, consider faults which are manifest by changes in value of the frictional coefficients μ_i. This type of fault could represent either wear in the roll or contamination. In five separate simulations, a fault corresponding to a 10% decrease in each coefficient μ_i was introduced at time $t = 1$ which corresponded to sample $k = 10$. In each case, the data was generated using the

nonlinear model and the Runge-Kutta fifth order method. Measurement noise was added, resulting in a signal to noise ratio of approximately 10.

Using the procedure described in Section 6.3, a bank of two MHE filters was implemented with a horizon length 10. The first filter corresponded to the no-fault case, whereas the second allowed for any fault. A second detector was implemented using a bank of six MHE filters, one for each fault type as well as the no-fault case. For both filter bank architectures, the measured variables were taken to be $t_{f1}, ..., t_{f4}$ and $h_1, ..., h_5$. Delayed versions of h_i were used as the inputs H_{i+1}. A diagonal matrix V was used in each MHE filter. Since a fault of the type considered in the simulations introduced a considerably smaller magnitude change in h_i than t_{fi}, elements of V weighting h_i were set to 100, whereas elements weighting t_{fi} were set to 1. The matrix Q was set to the identity, and P was calculated using the appropriate Riccati equation to guarantee stability [102]. The fault detection results are shown for in Figures 6.2 and 6.3.

To examine the false alarm robustness properties, the system was simulated in open loop using the control moves which would be obtained from a step decrease in μ_1, but without changing μ_1. This situation simulates the case of no-fault, but changing operating conditions. Plant/model mismatch was introduced in two separate fashions. First, an uncertainty in the gain of a subset of manipulated variables was introduced. The uncertainty in the model gain was incorporated in the detection scheme by introducing five parameters θ_i, one corresponding to each of the manipulated variables u_{si}. Letting a and b denote the discrete time state space matrices for the linearized process, and denoting by b_{si} and b_{vi} the columns of b corresponding to u_{si} and u_{vi} respectively, the uncertain dynamic model becomes

$$\mathbf{x}(k+1) = a\mathbf{x}(k) + \sum_{i=1}^{5}(1+\theta_i)b_{si}u_{si}(k) + \sum_{i=1}^{4} b_{vi}u_{vi}(k). \tag{6.57}$$

In the simulations, the uncertainty parameters assumed the values $\theta_1 = 0.2$, $\theta_2 = -0.2$, and $\theta_i = 0$ for $i > 2$. Robust fault detection was carried out using the above algorithm with bounds $|\theta_i| \leq 0.2$, and the results were compared to the nominal case

in which no parameter uncertainties were used in the fault detection scheme.

In a separate simulation, uncertainty in the time constant T_s was considered. The discrete time difference equation for the gap width is given by

$$S_i(k+1) = a_{ii}S_i(k) + b_{si}U_{si}(k), \tag{6.58}$$

where $a_{ii} = \exp(-\frac{1}{10T_s})$. Uncertainty in the first order time constant can be built in to the MHE fault detection filter by introducing factors θ_i:

$$S_i(k+1) = a_{ii}S_i(k) + b_{si}U_{si}(k) + a_{ii}\theta_i S_i(k). \tag{6.59}$$

Although S_i is not directly measured, the readout matrix C corresponding to measurements of t_{fi} and h_i has full row rank. Therefore, when no fault is present it is possible to express S_i as a linear combination of the measured quantities and obtain a difference equation of the form (6.4). Since the measurements do depend directly on the fault through the term D_w, S_i cannot be obtained explicitly from known quantities when a fault occurs. Therefore, if the estimated fault is non-zero, one obtains an estimate of S_i from the known quantities, t_f and h_i, and the estimate of the fault μ_i. Using the same control moves as above, the no-fault case was simulated with plant model mismatch by changing a_{11} and a_{22} from their nominal values of 0.7165 to 0.85 and 0.6, respectively. These values corresponded to $\theta_1 = 0.186$ and $\theta_2 = -0.1626$. The bound $|\theta_i| < 0.2$ was used in the MHE filter. The results for both the gain uncertainty and the pole uncertainty cases are shown in Figure 6.4.

6.5.5 Discussion

From Figures 6.2 and 6.3, the reader sees that both the two bank MHE filter and the six bank MHE filter give correct diagnosis in each fault case, although there is a small offset in the magnitude of the fault. Since the algorithm will attempt to compensate as much as possible using the uncertainty parameters θ, one should anticipate such an offset. However, when the primary goal is isolation of the fault rather than estimation

of its magnitude, this offset error will not be important. For the two bank MHE filter, the detection scheme indicated small but non-zero values of the absent faults. This is to be expected because the two bank scheme attempts to find the best combination of faults using a quadratic objective which includes penalties for changes in the fault estimate. On the other hand, as the six bank filter considers the presence of various faults independently and then chooses only one non-zero fault change at each time step in the MHE algorithm, estimates for absent faults are zero.

In the no-fault but significant model uncertainty cases shown in Figure 6.4, the reader sees that when the MHE optimization does not consider uncertainty parameters θ, the two bank MHE algorithm mistakenly ascribe a fault to the model uncertainty. However, including θ parameters in the MHE calculations results in correct no-fault diagnosis. Therefore, by explicitly accounting for model uncertainty through the parameters θ, the false alarm rate can be decreased. Similar results were obtained using a six bank MHE algorithm.

6.6 Conclusions

A new methodology for robust fault detection for a class of bounded uncertainties has been presented. This method uses Moving Horizon Estimation with constraints on θ-variables which are used to represent the model uncertainty. Due to the presence of constraints, the MHE filters are nonlinear. Because the MHE filter explicitly accounts for the model uncertainty at each step, a fault detection scheme which is less prone to false alarms caused by model mismatch can be formulated.

The robustness improvements attainable with a MHE fault detection filter are not cost free. Implementation of the MHE scheme requires on-line solution of quadratic programming problems, whereas Kalman filter schemes can be implemented with simple linear filters. In the case where a bank of MHE filters is employed, computation time can be decreased by using a parallel architecture.

Using a model of a complicated physical system, the cold tandem steel mill, the capabilities of the MHE detection scheme have been demonstrated. In particular, the

case study showed improved robustness to modeling uncertainty using the proposed MHE scheme with uncertainty parameters.

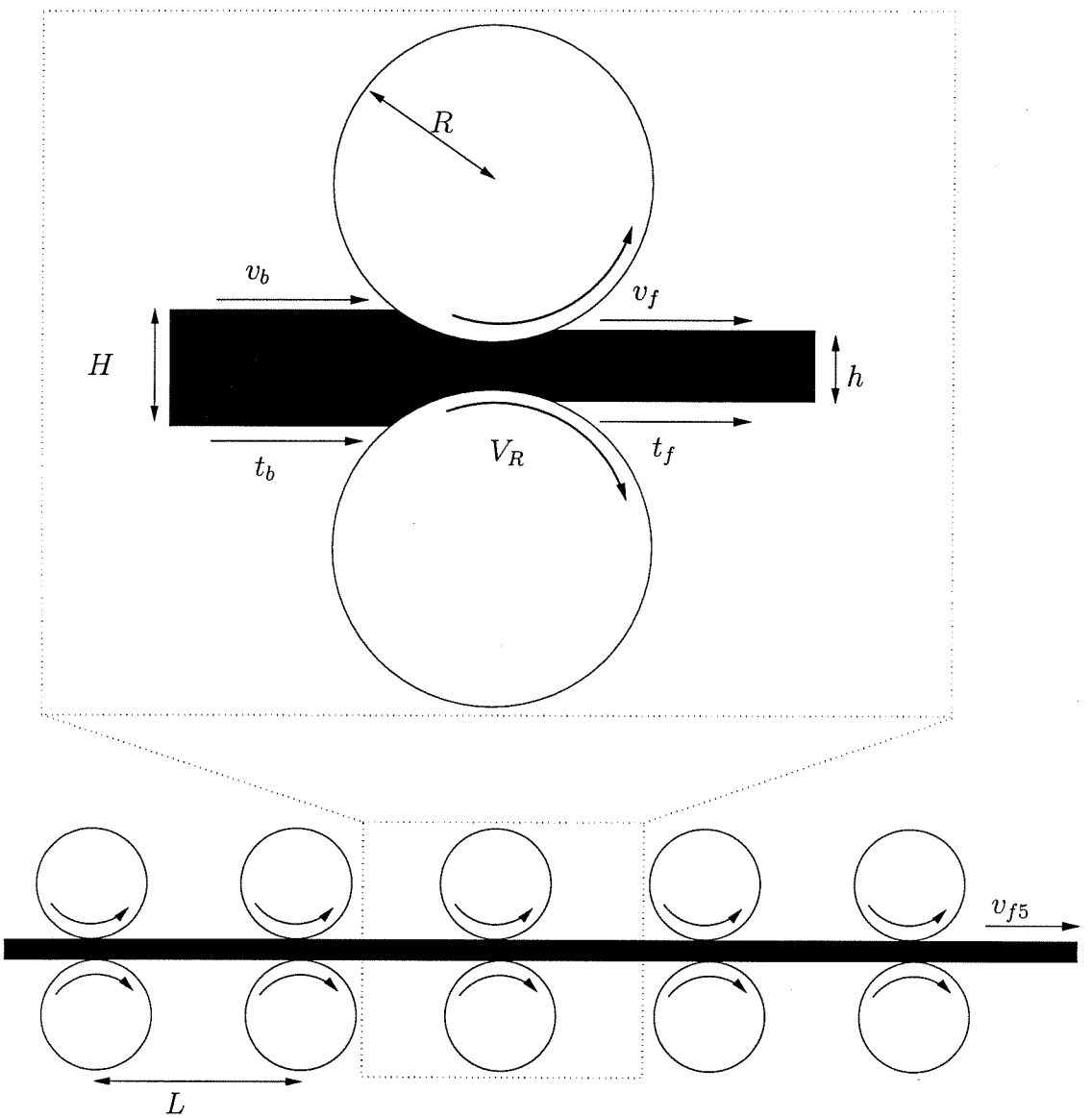

Figure 6.1: Cold tandem steel mill

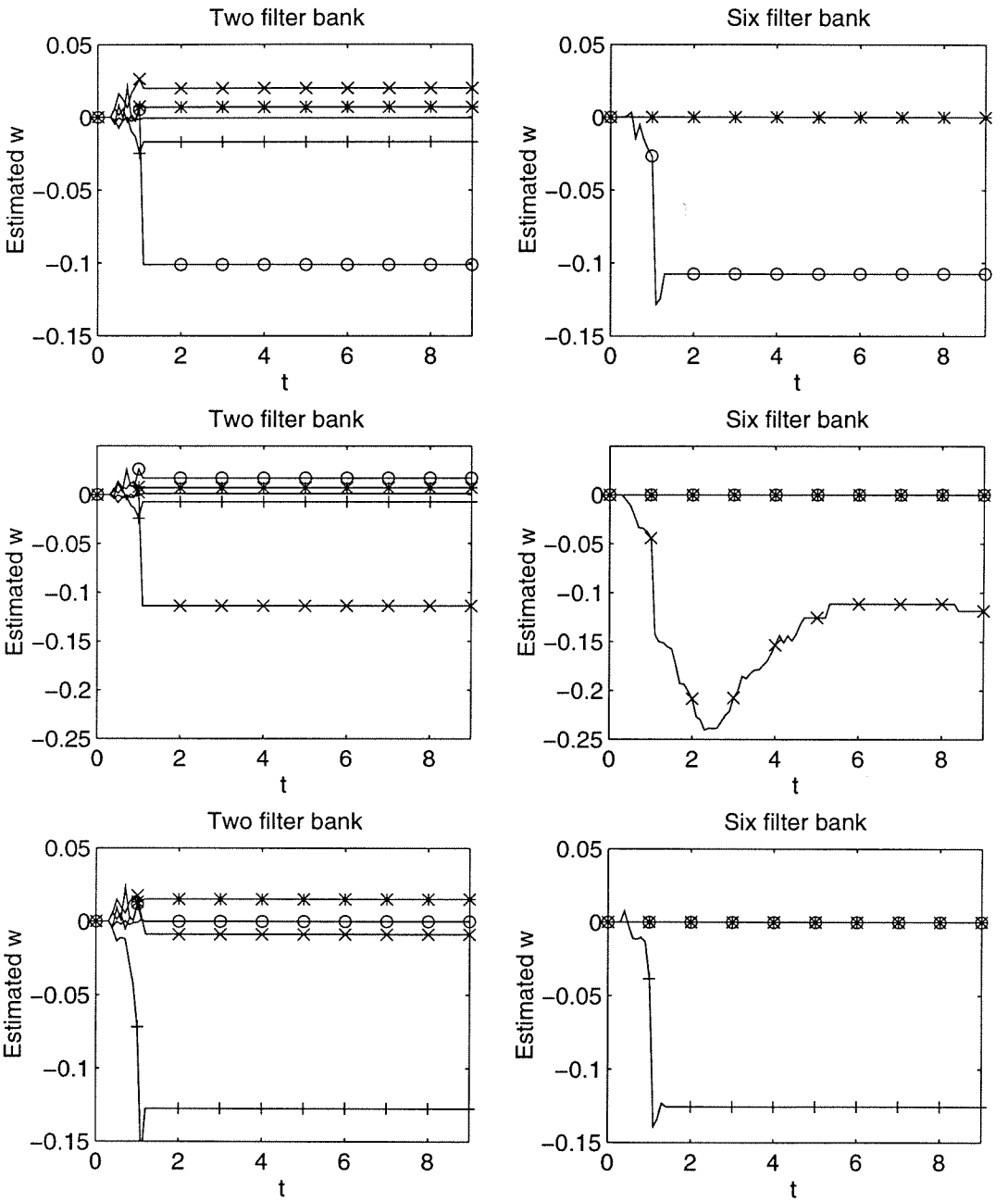

Figure 6.2: Estimated faults as measured in fractional change from nominal values of μ_i. Top: fault in μ_1. Middle: fault in μ_2. Bottom: fault in μ_3. Legend: `-o-` : μ_1, `-x-` : μ_2, `-+-` : μ_3, `-*-` : μ_4, `-` : μ_5.

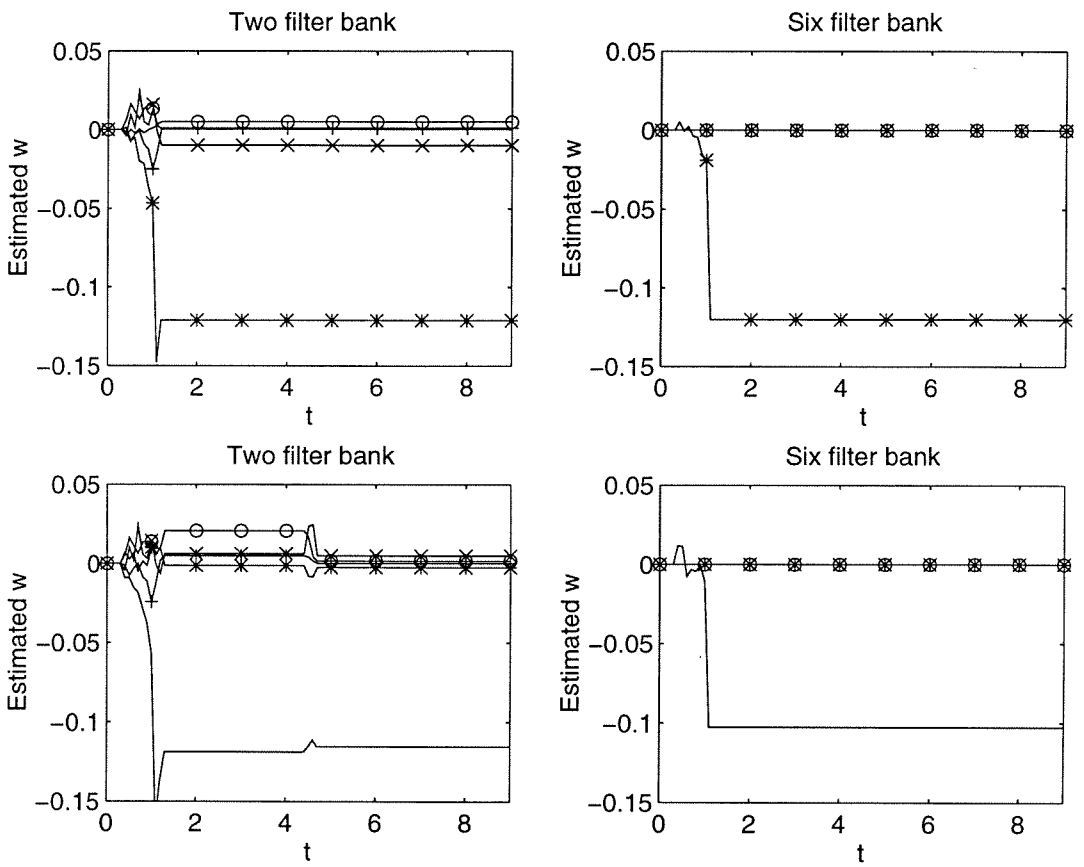

Figure 6.3: Estimated faults as measured in fractional change from nominal values of μ_i. Top: fault in μ_4. Bottom: fault in μ_5. Legend: `-o-` : μ_1, `-x-` : μ_2, `-+-` : μ_3, `-*-` : μ_4, `-` : μ_5.

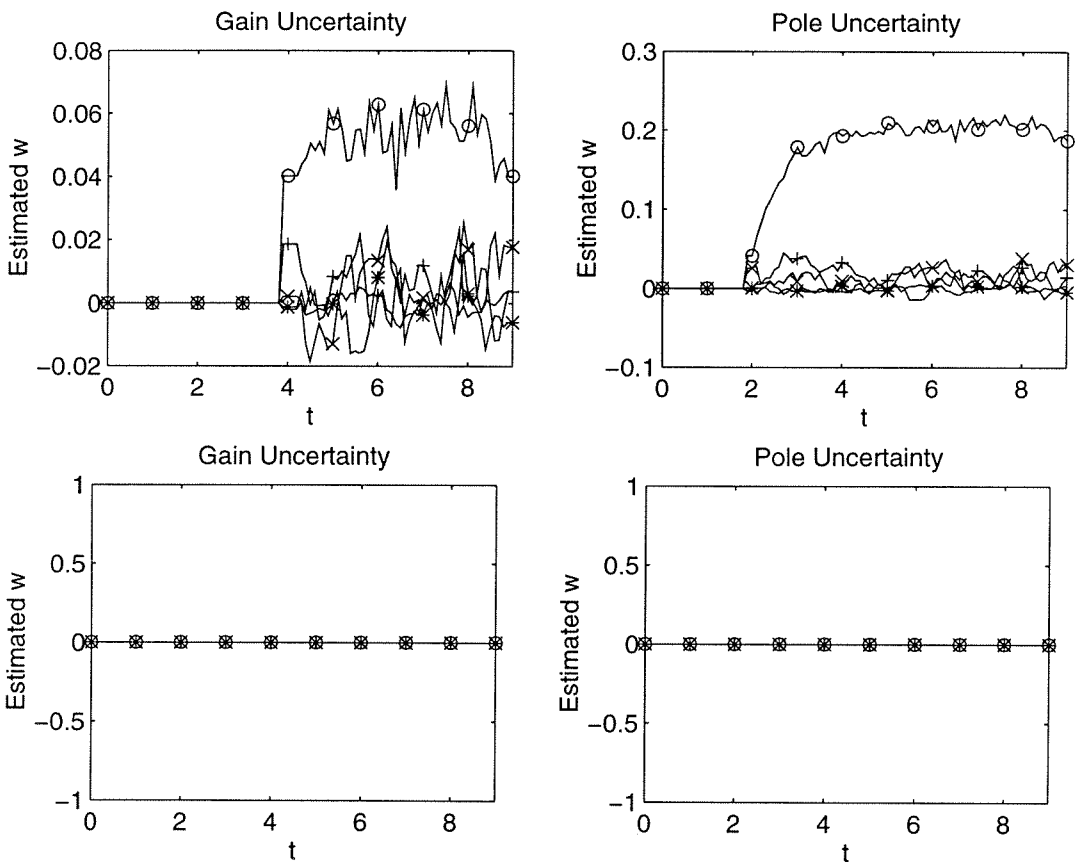

Figure 6.4: Estimated faults as measured in fractional change from nominal values of μ_i using two bank MHE filter. No-fault case. Top: no-fault, MHE without model uncertainty parameters. Bottom: no-fault, MHE with model uncertainty parameters. Legend: `-o-` : μ_1, `-x-` : μ_2, `-+-` : μ_3, `-*-` : μ_4, `-` : μ_5.

Chapter 7 Likelihood Ratio Detection using Nonlinear Filtering

Summary

This chapter presents a statistical framework for general change detection problems. A two-model approach is used, wherein signals and parameters subject to change are modeled by Brownian motion for the faulty case and by constant values in the nominal case. A detection algorithm using likelihood ratio testing is implemented through the use of recursive dynamic filtering. In the case of change in mean of a Gaussian sequence, a detailed analysis of the detection scheme reveals that for fixed error rates, there exist optimal filtering parameters which minimize the detection rate. For non-linear and non-Gaussian change detection, approximate filtering algorithms based on Bayes' law can be employed in the present framework. A computational filtering algorithm based on Bayes' law, probability grid filtering, is reviewed. The proposed framework combined with probability grid filtering is compared to the local asymptotic approach through an example containing non-linear dynamics. The proposed method's performance is vastly superior to the latter's.

7.1 Introduction

In Chapter 4, the design of filters which are sensitive to certain fault classes was considered. It was seen that the use of linear filters for this purpose can yield unsatisfactory false alarm rates, especially when the detection scheme is coupled with a controlled process. In Chapters 5 and 6, constrained Moving Horizon Estimation was used to directly account for a class of model uncertainties.

An alternative approach to detection problems which has been used extensively

consists of filtering observations through a whitening filter and testing the filter innovations for deviations from the white noise hypothesis. The whitening filter can either be known, such as in the cases of tests based on banks of Kalman filters [111] and on local tests [80], or identified as in the Generalized Likelihood Ratio test. A second approach uses a two-model structure. Traditional two-model approaches identify two dynamic models using different subsets of the signal and then compare the identified models using conventional distance measures [7].

In this chapter, we present a general statistical approach to the detection of changes in dynamic systems which uses a two-model structure and combines likelihood ratio testing with non-linear Bayesian filtering. This framework is capable of encompassing parameter change detection [48] as well as sensor and actuator failure detection [34] for a wide class of systems. As compared to other two-model approaches such as found in [7], the proposed method contains a significant difference: the model corresponding to faulty operating conditions represents by Brownian motion those parameters and signals whose changes correspond to faults.

This chapter is organized as follows. In Section 7.2, the general framework is outlined, and the fundamental result of Neyman-Pearson is reviewed. Next, computational methods for non-linear filtering, an important component of the proposed method, are discussed in Section 7.3. Section 7.4 contains an in-depth analysis of detection properties for the proposed method in the basic case of change in mean of a Gaussian sequence. In Section 7.5, an alternative approach for detecting changes in non-linear systems, the local asymptotic approach, is reviewed, and a case study is presented in Section 7.6 to compare the two methods.

7.2 Framework for detection

In this work, we will only discuss problems of detection; however, the methods developed could easily be applied to model validation problems as well. One approach which has proven to be efficacious for solving detection and diagnosis problems consists of casting them in a hypothesis testing framework, in which the null hypothesis

$\mathcal{H}_0$ corresponds to the situation in which no fault has occurred. The fault detection problem is then the problem of deciding whether to accept or reject $\mathcal{H}_0$. For diagnosis problems, a multiple hypothesis testing approach is used where from a set of hypotheses $\{\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_n\}$ one is chosen.

In this chapter we shall consider processes which can be described as controlled semi-Markov processes. A sequence (y_k) is a controlled semi-Markov process with control parameter θ_* if (y_k) if it can be expressed in the form

$$\begin{aligned} P\{\xi_k \in G | \xi_{k-1}, \xi_{k-2}, \ldots\} &= \int_G \pi_{\theta_*}(\xi_{k-1}, dx), \\ y_k &= h(\xi_k), \end{aligned} \tag{7.1}$$

where $\pi_{\theta_*}(\xi, dx)$ is the transition probability of a Markov chain ξ_k depending on the parameter θ_*. Furthermore, assume that the changes to be detected correspond to changes in the value of the control parameter θ_*. Suppose that a change occurs in the parameter θ_* at time r. In other words, there exists a time r such that (y_k) is controlled by the parameter $\theta_* = \theta_0$ for $k < r$ and by the parameter $\theta_* = \theta_1$ for $k \geq r$. The detection problem consists of determining from a record $y_0, \ldots, y_N$ whether a change has occurred, i.e., whether $r \leq N$.

In the sequel, we will consider the case where the probability law (7.1) can be described as the solution to a stochastic difference equation with parameter θ_*. Let us partition the Markov parameter ξ as $\xi = [x^T, u^T]$, where the parameter u is known precisely (or delta distributed). In practical applications, u will represent a known process variable, such as a manipulated input. The process (y_n) will be described as the output of the stochastic difference equation

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, \theta_*) + w_k, \\ y_k &= h(x_k, p_k) + v_k, \end{aligned} \tag{7.2}$$

where w_k and v_k are random variables with distributions p_w and p_v. By specifying the probability distribution functions p_w and p_v along with an initial distribution

for x_0, the probability of a sequence $Y_0^N = \{y_0, y_1, \ldots, y_N\}$ is determined for a fixed value of the control parameter θ_*. This model is quite general, and encompasses two important classes, additive faults and multiplicative faults in linear systems. For a linear system of the form

$$x_{k+1} = Ax_k + Bw_k + Fp_k, \tag{7.3}$$

$$y_k = Cx_k + Dv_k + Hp_k, \tag{7.4}$$

additive faults correspond to $\theta_* = p$, whereas multiplicative faults refer to the case where $\theta_* = \{A, C\}$. This nomenclature is derived from the fact that for additive faults the parameter subject to changes is added to the state, whereas in the case of multiplicative faults, the parameter multiplies the state.

Consider the problem of determining a change in the control parameter θ_* from θ_0 to θ_1, and let $\mathcal{H}_0$ and $\mathcal{H}_1$ be the hypotheses corresponding to each of these parameter values. Let g be a decision function with domain Y_0^N and range $\{\mathcal{H}_0, \mathcal{H}_1\}$. Two properties which are used to quantify the performance of a test g are the *size* α, which measures the probability of a false detection, and the *power* β, which measures the probability of correctly diagnosing a change. These quantities are formally defined as follows:

$$\alpha = P\left(g(Y_0^N) = \mathcal{H}_1 | \mathcal{H}_0\right), \tag{7.5}$$

$$\begin{aligned} \beta &= P\left(g(Y_0^N) = \mathcal{H}_1 | \mathcal{H}_1\right), \\ &= 1 - P\left(g(Y_0^N) = \mathcal{H}_0 | \mathcal{H}_1\right). \end{aligned} \tag{7.6}$$

An effective test g should have a small size, and a power near one; however, there exists inherent tradeoffs between these measures. By increasing the test power, the size is usually also increased. A fundamental result, the Neyman-Pearson lemma, states that a test is optimal in the sense of maximizing the power for a fixed size if and only if it is based upon the likelihood ratio. In this case, the decision function g

is given by:

$$g(Y_0^N) = \begin{cases} \mathcal{H}_0, & \text{when } \frac{p_1(Y_0^N)}{p_0(Y_0^N)} < \lambda_\alpha \\ \mathcal{H}_1, & \text{when } \frac{p_1(Y_0^N)}{p_0(Y_0^N)} \geq \lambda_\alpha \end{cases}, \tag{7.7}$$

where p_i is the probability distribution function obtained from the difference equation (7.2) with $\theta_* = \theta_i$. A test satisfying the optimality criterion of the Neyman-Pearson lemma is said to be the *Most Powerful* test with size α, where α is defined by

$$\alpha = P\left(\frac{p_1(Y_0^N)}{p_0(Y_0^N)} \geq \lambda_\alpha | \mathcal{H}_0\right). \tag{7.8}$$

The problem of devising an "optimal" decision function g therefore reduces to one of evaluating the probability functions $p_i(Y_0^N)$.

Often, for computational purposes it is convenient to replace the probability ratio with its logarithm. Let us define the S_N as the log likelihood ratio,

$$S_N = \ln \frac{p_1(Y_0^N)}{p_0(Y_0^N)}. \tag{7.9}$$

To evaluate S_N, we first expand the probability function in terms of conditional probabilities:

$$\begin{aligned} p_i(Y_0^N) &= p_i(y_N|Y_0^{N-1})p_i(Y_0^{N-1}), \\ &= p_i(y_N|Y_0^{N-1})p_i(y_{N-1}|Y_0^{N-2})p_i(Y_0^{N-2}), \\ &\vdots \\ &= \left(\prod_{k=1}^{N} p_i(y_k|Y_0^{k-1})\right) p_i(y_0). \end{aligned} \tag{7.10}$$

Now, by taking the logarithm, the product is transformed to a sum so that S_N can be calculated by

$$\begin{aligned} S_N &= \ln\left(p_1(Y_0^N)/p_0(Y_0^N)\right), \\ &= S_{N-1} + s_N, \end{aligned} \tag{7.11}$$

where s_k is given by

$$s_k = \ln \frac{p_1(y_k|Y_0^{k-1})}{p_0(y_k|Y_0^{k-1})}. \tag{7.12}$$

Therefore, to recursively calculate S_N, a probability distribution for y_k conditioned on past values Y_0^{k-1} must be available for each value of θ.

For on-line detection of changes, a sequential probability ratio test (SPRT) is practical. An SPRT is a decision function together with a stopping time T of the form

$$g(Y_0^T) = \begin{cases} \mathcal{H}_0 & \text{when } S_T \leq -a \\ \mathcal{H}_1 & \text{when } S_T \geq \lambda \end{cases}, \tag{7.13}$$

$$T = \min_{S_k \geq \lambda \bigcup S_k \leq -a} k. \tag{7.14}$$

Although one can with a reasonable degree of certainty assume that the parameter value θ_0 before change is known (it can always be estimated from a set of test data), more often than not, the value of θ_1 after a change is completely unknown. For simple detection problems, several methods exist to account for an unknown parameter value after change. In the case of change in mean of a scalar Gaussian process, one can specify a minimum change magnitude ν, and use two parallel tests, one corresponding to a change of $+\nu$ and the other to a change of $-\nu$, yielding the so called "two-sided" test [6]. An alternative approach is to replace the posterior parameter with its maximum likelihood estimate, resulting in the Generalized Likelihood Ratio (GLR) approach. For multi-dimensional θ_*, the two sided approach cannot easily be extended as there exists a continuum of directions in which the change could occur. On the other hand, the GLR approach can be extended in a straightforward fashion to multivariate problems. For problems which can be reduced to changes in the mean of Gaussian processes, the GLR approach can be applied with an explicit optimization for estimating θ_1. For more complicated processes, optimization over θ_1 becomes impractical, and the local asymptotic approach, which essentially approximates the problem at hand by a simpler change in mean problem, is often the only alternative. A brief description of this method is contained in Section 7.5.

In this work, a new approach to specifying the parameter after the change is pur-

sued. As before, the no-fault case will be modeled by a system with fixed parameter value θ_0; however, rather than modeling the failure case by using the same model with a different parameter value θ_1, a different stochastic difference equation will be used. In this second model, the parameter subject to change will be modeled as a Brownian motion. Thus, whereas (7.2) with $\theta_* = \theta_0$ will be used to generate the distribution p_0 corresponding to no-fault, the distribution p_1 will be generated by the difference equations

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, \theta_k) + w_k, \\ \theta_{k+1} &= \theta_k + q_k, \\ y_k &= h(x_k, p_k) + v_k, \end{aligned} \tag{7.15}$$

where q_k is a zero mean white noise process with given covariance. In this framework, rather than representing faults by abrupt changes in the parameter, the faulty situation is modeled by a drift of the parameter with the rate of the drift determined by the variance of q. Because the model (7.15) allows the parameter to drift, the resulting distribution function p_1 will be more spread out than the distribution p_0. Consequently, when the parameter θ_0 accurately describes the system dynamics, the expected value of s_k will be non-positive. On the other hand, when a change has occurred in the dynamics, the model (7.15) will better describe the process behavior since it allows θ_k to change, whereas θ_0 remains fixed in (7.2). In this case, s_k is expected to be positive.

The framework is now laid for the general detection problem. A two model approach is used, with the "no-fault" model and its accompanying probability distribution function p_0 determined by the stochastic difference equation (7.2), and the "faulty" model and its distribution p_1 determined by (7.15). For both models, the problem remains of solving the conditional probabilities needed to calculate the log likelihood terms s_k in (7.12). For the case of linear dynamics with additive faults and Gaussian distributions for p_w, p_q, and p_v, this is accomplished by using the appropriate Kalman filter [1]. In this case, the probability measure $p_i(y_k|Y_0^{k-1})$ corresponds

to a normal distribution with mean $C\hat{x}^i_{k|k-1}$ and covariance $C\Sigma^i_k C^T + R$, where $\hat{x}^i_{k|k-1}$ is the Kalman filter state prediction and Σ^i_k is the variance of $x_k - \hat{x}^i_{k|k-1}$ which can be calculated from an appropriate Riccati equation for Model i. To calculate s_k, two Kalman filters should be run in parallel, and the results are used to evaluate the SPRT (7.13). At first glance, this approach seems to be related to the likelihood ratio test where the parameter value θ_1 after the change is replaced by its Kalman filter estimate; however, this is not the case. Even in the simplest case where y is a Gaussian white noise sequence whose mean is subject to change, the distributions p_0 and p_1 do not differ only in mean; p_1 has a larger variance. This case is examined in more detail in Section 7.4.

For more general problems, no closed form solution exists for calculating the conditional probabilities needed in (7.12), and one must resort to approximate methods. A very useful approach based upon Bayesian estimation is discussed in the next section.

7.3 Nonlinear probability grid filters

Given a method for calculating $p_i(y_k|Y_0^{k-1})$, the conditional probability distribution function associated with the difference equations (7.2) and (7.15), an efficient SPRT can be obtained using the result of Neyman-Pearson. As mentioned in the previous section, in the case of linear dynamics and Gaussian disturbances the resulting distribution is normal and can be calculated recursively using a Kalman filter. When these restrictions do not hold, alternative methods must be used. In this section, the Bayesian estimation approach is reviewed along with an approximate implementation, probability grid filtering (PGF).

An in-depth discussion of the Bayesian approach to nonlinear filtering can be found in [49]. The results will be briefly summarized here. The Bayesian approach provides a recursive procedure for updating the prediction density. Given the conditional distribution $p(x_k|Y_0^{k-1})$ and a new measurement y_k, the distribution is updated to

obtain $p(x_k|Y_0^k)$ by using Bayes' law

$$p(x_k|Y_0^k) = \frac{p(y_k|x_k)p(x_k|Y_0^{k-1})}{p(y_k|Y_0^{k-1})}, \tag{7.16}$$

where

$$p(y_k|Y_0^{k-1}) = \int p(y_k|x)p(x|Y_0^{k-1})dx. \tag{7.17}$$

and the distribution $p(y_k|x)$ is given by $p_v(y_k - h(x))$. This step is commonly referred to as the *measurement update* step. Next, in the *dynamic update* step the prediction density is calculated from the updated distribution and the state transition probability:

$$p(x_{k+1}|Y_0^k) = \int p(x|Y_0^k)p(x_{k+1}|x)dx, \tag{7.18}$$

where the transition density $p(x_{k+1}|x_k)$ is obtained from the dynamics and p_w. The equations (7.16) and (7.18) give the complete solution to the filtering problem. When the dynamics are linear and the distributions p_w and p_v are Gaussian, the Bayesian approach yields the Kalman filter [1]. For the general case, a closed form recursion cannot be obtained and approximate numerical methods must be used.

Applied to problems of estimation and filtering, a noted weakness of the Bayesian approach is that rather than producing a finite-dimensional description of the variables that are estimated, it produces a function, the posterior distribution. Although this is considered a weakness in filtering and estimation, for change detection problems it becomes a strength as the Neyman-Pearson test must evaluate the distribution. By applying the Bayesian recursion formula to the data, the conditional probability $p(y_k|Y_0^{k-1})$ can be calculated at each time step.

A large number of methods for the numerical evaluation of the Bayesian recursion relations have been proposed. The interested reader is referred to the works [95], [53], [54], [96], and references found therein. In this work, we will not attempt to provide a comprehensive discussion of these methods, but rather will present a short description of one such algorithm with the intent of demonstrating the fundamental concept. The various methods have the common characteristic of defining a grid of points in the

probability space on which the approximations are based. Because a finite number of points are used, the grid is restricted to a finite region of state space. When choosing the grid, it is important to assure that the probability mass is non-significant outside this region. Proper grid selection is an important consideration in the algorithm, but will not be discussed here.

The extended Kalman filter (EKF) uses a single grid point and linearizes the functions f and h around this grid point. The posterior density obtained from the EKF is Gaussian. The EKF has found wide use in nonlinear filtering problems because of its simplicity and efficiency, especially for processes which operate near a steady state and thereby mitigate linearization errors. However, when considering dynamic change detection, the situation arises where the density may not be well approximated by a Gaussian distribution which would assign negligible probability mass to points corresponding to changing dynamics. The EKF can give very poor results because the probability density will necessarily be small away from the grid point. Extensions to the EKF which retain second-order and higher terms in the expansion of the system functions f and h exist [49]. The EKF and its extensions can be considered as local methods as the resulting distributions are most accurate within a neighborhood of the single grid point. In all local methods, the use of a single point to form the approximation presents a disadvantage.

Due to the recognized weaknesses of local methods, in the late 1960's, global approximate methods were introduced to deal with nonlinear filtering problems. Bucy [16] proposed a general approach, the point mass method, but due to the computational limits of the day, few applications were forthcoming. Although since that time a variety of global methods have appeared, all specific global methods must provide solutions to the following sub-problems as noted by Sorenson [96]:

1. An initial grid must be defined.

2. A procedure must be given for defining the grid at each subsequent sampling time.

3. A method for carrying out the Bayes' rule calculations on the grid must be

specified.

The specific method discussed below is the "p-Vector" approach [96], which derives its name from the fact that the algorithm calculates at each step a vector of probability values $p_{k,i}$ corresponding to grid points $\eta_{k,i}$.

Assume that the one-step prediction density is known. For the initial step, this density must be specified by the user, and for subsequent steps, it is available as the output of the previous step. Let $\eta_{k,i}$ denote the location of i^{th} grid point during the k^{th} step of the algorithm. Similarly, let $p_{k,i}$ denote the value of the conditional distribution $p(x_k|Y_0^{k-1})$ evaluated at the i^{th} grid point, that is $x_k = \eta_{k,i}$. The measurement update is given by

$$p(\eta_{k,i}|Y_0^k) = \frac{1}{c_k} p(\eta_{k,i}|Y_0^{k-1}) p(y_k|\eta_{k,i}) \tag{7.19}$$

$$= \frac{1}{c_k} p_{k,i} p_v\left(y_k - h(\eta_{k,i})\right) \tag{7.20}$$

where $c_k = p(y_k|Y_0^{k-1})$ is given by (7.17) and can be approximated numerically using the grid points $\eta_{k,i}$ and the densities $p_{k,i}$:

$$c_k \approx \sum_i \alpha_i p_{k,i} p_v\left(y_k - h(\eta_{k,i})\right), \tag{7.21}$$

and α_i are constants depending upon the numerical integration routine. The normalization factor c_k gives the conditional probability value needed to calculate s_k for the likelihood ratio test. Once the measurement update has been performed, the dynamic update equations are used to obtain the one-step prediction densities.

Given the grid points $\eta_{k,i}$, define new grid points as the one-step dynamic evolution of these points:

$$\eta'_{k,i} = f(\eta_{k,i}, u_k, \theta). \tag{7.22}$$

Next, let J denote the Jacobian of the inverse of f, $J = \partial f^{-1}/\partial \eta'$. Then $p(\eta'_{k,i}|Y_0^k)$ can be calculated [49]:

$$p(\eta'_{k,i}|Y_0^k) = p(\eta_{k,i}|Y_0^k) \det\left(J(\eta'_{k,i})\right). \tag{7.23}$$

If there is no system noise ($w_k = 0$), then this relation determines the prediction density at the grid points η'. In the more general case of system noise, the prediction density is determined by convolving the update density with the noise density as in (7.18). Numerically, these values may be calculated using the grid points η':

$$p(\eta_{k+1,i}|Y_0^k) = \sum_{j=1}^{n_k} p(\eta'_{k,j}|Y_0^k) p_w\left(\eta_{k+1,i} - \eta'_{k,j}\right) \tag{7.24}$$

Since the noise density p_w is known, the density can be evaluated for any grid location η_{k+1}. This convolution presents the largest computational burden. In the event that the grid points are equally spaced, fast Fourier transform (FFT) can be utilized; however, the dynamic update (7.22) will destroy constant spacing. Kramer and Sorenson [54] have proposed an alternative algorithm using piece-wise constant approximations which has the property that constant spacing can be preserved; however, their algorithm contains other computationally expensive steps. Which algorithm performs better will often depend on the specific application. The purpose of this section is to provide the reader the fundamental concept of probability grid filtering and not to furnish a critical review of the various implementations. The reader interested in comparing implementations of probability grid filters is referred to the citations above and the references found therein.

For practical implementation, an important consideration is the choice of new grid points $\eta_{k+1,i}$. From (7.24) it is clear that if $\eta_{k+1,i}$ is much different from $\eta'_{k,i}$, where difference is measured by the probability p_w, then $p_{k+1,i}$ may become negligibly small due to the approximations used in the calculation. Evidently there is a strong relation between the density p_w and the choice of new grid points η_{k+1}.

7.4 Unknown change in mean of Gaussian sequence

Let us analyze the properties of the change detection scheme for the special case of a change in the mean of a random Gaussian sequence. Using the method outlined above, we consider the likelihood ratio for each of the following two models:

Model 0

$$y_k = \mu_0 + v_k,\ E[v_k^2] = R. \tag{7.25}$$

Model 1

$$\mu_{k+1} = \mu_k + w_k,\ E[w_k^2] = Q, \tag{7.26}$$

$$y_k = \mu_k + v_k,\ E[v_k^2] = R. \tag{7.27}$$

The probability distribution functions for each of the models are given by:

$$p_0(y_k|Y_0^{k-1}) = \frac{1}{\sqrt{2\pi R}} \exp\left(-\frac{(y_k-\mu_0)^2}{2R}\right), \tag{7.28}$$

$$p_1(y_k|Y_0^{k-1}) = \frac{1}{\sqrt{2\pi(\Sigma+R)}} \exp\left(-\frac{(y_k-\hat{\mu}_k)^2}{2(\Sigma+R)}\right), \tag{7.29}$$

where $\hat{\mu}_k$ is the Kalman filter prediction based upon the data Y_0^{k-1} and satisfies the following recursion relation

$$\hat{\mu}_{k+1} = \hat{\mu}_k + K\,(y_k - \hat{\mu}_k) \tag{7.30}$$

where K is the Kalman filter gain, $K = \Sigma(\Sigma+R)^{-1}$, and Σ is related to Q by $Q = \Sigma^2(\Sigma+R)^{-1}$. The log-likelihood ratio therefore satisfies the relationship

$$\begin{aligned} s_k &= \ln\left(p_1(y_k|Y_0^{k-1})\right) - \ln\left(p_0(y_k|Y_0^{k-1})\right), \\ &= \frac{1}{2}\left(\ln(R) - \ln(\Sigma+R) + \frac{(y_k-\mu_0)^2}{R} - \frac{(y_k-\hat{\mu}_k)^2}{\Sigma+R}\right). \end{aligned} \tag{7.31}$$

Now let us consider the properties of this change detection algorithm. In particular, we will consider the mean delay for detection T_λ and the mean time between false alarms L for the case of a step change in the mean from μ_0 to μ_1. For efficient detection, L should be large and T_λ small.

Consider an SPRT as in (7.13) and (7.14) with lower limit $a = 0$. Let P_0 be the

probability that the SPRT algorithm stops by hitting its lower boundary of 0 given that it starts as zero, and let T be the stopping time associated with this SPRT. The expectation $E_i(T)$, $i \in \{0,1\}$ denotes the average sample number (ASN) before the algorithm terminates at either the lower boundary 0 or the upper boundary λ, with the subscript i indicating which model is used to carry out the expectation. With this notation, the mean time between alarms is given by

$$L = \frac{E_0(T)}{1 - P_0}. \tag{7.32}$$

In the case of stationary increments s_k and negligible excess over the boundary (i.e., $E(S_T | S_T \geq \lambda) = \lambda$), Wald's approximation for L can be used (see [9]):

$$L \approx \frac{1}{E_0(s_k)} \left(h + \frac{e^{-w_0 \lambda} - 1}{w_0} \right), \tag{7.33}$$

where w_0 is unique non-zero root of the equation $E_0(e^{-w_0 s_k}) = 1$. In order to calculate w_0, we remark that y_k and $\hat{\mu}_k$ are independent Gaussian sequences with mean μ_0 and variance R and $\frac{KR}{2-K}$ respectively. Then $E_0(e^{-w_0 s_k})$ can be calculated by evaluating the integral

$$\begin{aligned} E_0(e^{-w_0 s_k}) &= \frac{1}{2\pi} \sqrt{\frac{2-K}{K}} \int dy \int d\hat{\mu} \times \\ &\quad \exp\left(-w_0 s_k - \frac{(y-\mu_1)^2}{2R} - \frac{(\hat{\mu}-\mu_1)^2}{2\frac{RK}{2-K}} \right), \qquad (7.34) \\ &= \left(\frac{(X+2)(X+1)^{(w_0+1)}}{(w_0+1)X^2 + (3+w_0-w_0^2)X + 2} \right)^{1/2}, \qquad (7.35) \end{aligned}$$

where $X = \Sigma R^{-1}$. The equation $E_0(e^{-w_0 s_k}) = 1$ is satisfied for $w_0 \in \{0, -1\}$. Evaluating a similar integral, $E_0(s_k)$ can also be calculated:

$$E_0(s_k) = \frac{1}{2} \left(1 - \ln(1+X) - \frac{2}{X+2} \right). \tag{7.36}$$

Substituting $w_0 = 1$ and $E_0(s_k)$ into Wald's approximation for L yields the final

expression:

$$L = 2\frac{e^{\lambda} - \lambda - 1}{\frac{2}{X+2} + \ln(1+X) - 1}. \tag{7.37}$$

We see that the mean time between false alarms L depends not only upon the threshold λ but also upon the parameter X. The denominator of L is a monotonically increasing function of X. Therefore, to avoid a high false alarm rate, it is desirable that X should be small; however, the smaller X, the slower $\hat{\mu}$ will respond to changes in the mean of y. Intuitively, this suggests that for smaller X the delay for detection will be increased. Similarly, as λ increases, L will increase as will the delay for detection.

Now let us consider the delay for detection. In the case of a step change in μ from μ_0 to μ_1 at time zero, the random variable $\hat{\mu}_k$ will again be Gaussian with constant variance $KR(2-K)^{-1}$ and with time varying mean $\bar{\mu}_k$ given by

$$\bar{\mu}_k = \mu_1 + (1-K)^k(\mu_0 - \mu_1). \tag{7.38}$$

Using this distribution and defining the signal to noise ratio $snr = (\mu_0 - \mu_1)^2 R^{-1}$, the expectation $E_1(s_k)$ is given by the time varying expression

$$E_1(s_k) = \frac{1}{2}\left[-\ln(1+X) + \frac{X}{X+2} + snr\right] - \frac{snr}{X+1}\gamma^k, \tag{7.39}$$

where $\gamma = (1-K)^2 = \left(\frac{X}{X+1}\right)^2$. Since $0 < \gamma < 1$, for large k, $E_1(s_k)$ approaches a constant value given by the first term on the right-hand side of (7.39). If this term is negative, then the expectation satisfies $E_1(s_k) < 0$ for all k after the change, and the probability that S_T reaches λ before it reaches 0 becomes minutely small. Therefore, changes in the mean will be correctly diagnosed with high probability only when the first term is positive or, in other words, when the following relation between snr and X holds:

$$snr \geq \ln(1+X) - \frac{X}{X+2}. \tag{7.40}$$

Even when this relation is satisfied, $E_1(s_k)$ may be negative for small values of k.

Since $E_1(s_k)$ is a monotonically increasing function of k, this will occur only when $E_1(s_0)$ is negative, or equivalently:

$$snr \leq \frac{X+1}{X}\ln(1+X) - \frac{X+1}{X+2}. \tag{7.41}$$

Since $E_1(s_k)$ is time varying sequence, Wald's approximation cannot be employed to estimate the mean delay for detection. Instead, we will use a different approximation. Assume the SPRT terminates at time T'. If we also assume that with probability 1 the SPRT terminates by crossing the lower boundary 0 whenever $E_1(s_{T'}) \leq 0$ and by crossing the upper boundary whenever $E_1(s_{T'}) > 0$, then T_λ can be estimated by solving $E_1(S_{T_\lambda}) = \lambda$, where $E_1(S_T)$ is given by

$$E_1(S_T) = \sum_{k=0}^{T} \max\{0, E_1(s_k)\}. \tag{7.42}$$

This assumption will be valid in the limit of large threshold λ. Let $k = T_0$ be the last time at which $E_2(s_k) \leq 0$. T_0 is non-zero only when (7.41) is satisfied, in which case it is given by

$$T_0 = \frac{1}{\ln\gamma}\left[\ln(X+1) - \ln(snr) + \ln\left(\frac{X}{X+2} - \ln(1+X) + snr\right)\right]. \tag{7.43}$$

By substituting the expression (7.39) for $E_1(s_k)$ into (7.42), $E_1(S_T)$ can be related to X and snr:

$$E_1(S_T) = \begin{cases} \frac{T}{2}\left[\frac{X}{X+2} - \ln(1+X) + snr\right] + \frac{1}{2}\frac{\gamma}{1-\gamma}\frac{snr}{X+1}\left(\gamma^T - 1\right), \\ \qquad (7.41) \text{ violated}, \\ \frac{1}{2}\left[\frac{X}{X+2} - \ln(1+X) + snr\right]\left[(T-T_0) + \frac{\gamma}{1-\gamma}\left(\gamma^{(T-T_0)} - 1\right)\right], \\ \qquad (7.41) \text{ satisfied}. \end{cases} \tag{7.44}$$

Letting $F(T, X, snr)$ denote the right-hand side of (7.44), the approximate mean delay for detection T_λ as a function of X and snr is determined by equating λ and

$E_1(S_{T_\lambda})$:

$$\lambda = F(T_\lambda, X, snr). \tag{7.45}$$

In Figure 7.1, curves corresponding to constant mean delay for detection are shown as a function of X and snr for a fixed value of $L = 1000$. Since the mean time between false alarms L is held constant, different threshold levels λ apply for each value of X. For large values of X, smaller changes result in $E_1(s_k) < 0$ and the algorithm fails to detect the changes. As the magnitude of the parameter change increases, the delay for detection decreases as expected. The shape of the curves indicates that for a fixed value of snr, there exists a value of X which minimizes the delay for detection for constant L.

In Figure 7.2, the dependence of mean delay for detection T_λ on the tuning parameter X is depicted for a fixed signal to noise ratio $snr = 1$. As this figure shows, for a given false alarm rate L, there exists a value of X which minimizes the detection delay. This can be explained by noting that for small values of X, the filter gain K is close to zero. With a small gain K, $\hat{\mu}_{k+1}$ obtained from (7.30) responds slowly to changes in the mean of y_k. Therefore, many samples are needed before the estimated $\hat{\mu}$ changes significantly from μ_0 to provide a more accurate description of the distribution. On the other hand, for X large, the variance of distribution function p_1 in (7.29) is large and the estimate $\hat{\mu}$ is strongly affected by the measurement y_k. In this situation, a larger threshold λ is required to keep L large, and a larger value of λ results in a longer delay for detection T_λ. Since the mean delay for detection increases in the limit of small X as well as large X, there exists an intermediate value which optimizes the algorithm's performance for a fixed signal to noise ratio.

7.5 Local asymptotic approach

To gain a deeper understanding of the advantages of the proposed change detection algorithm consisting of likelihood ratio testing together with probability grid filtering, let us compare the proposed algorithm to an alternative method, the local asymptotic

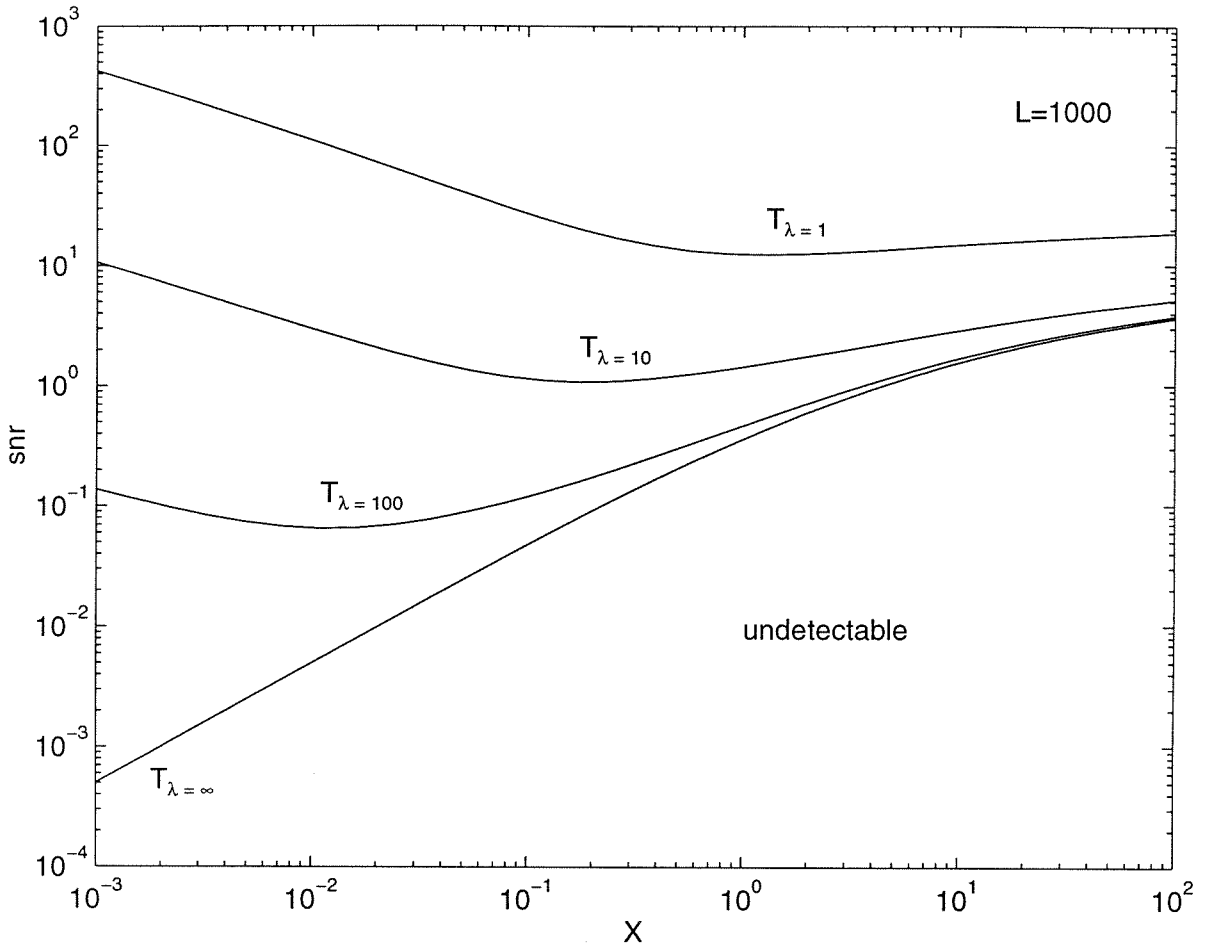

Figure 7.1: Delay for detection T_λ as a function of signal to noise ratio *snr* and tuning parameter X for fixed mean time between false alarms $L = 1000$

approach. In this section, a brief review of the local asymptotic approach based upon recursive identification algorithms is presented following the references [80], [12], [113], and [9]. The summary given here is meant to give the reader the necessary background to understand the comparisons which follow, and the reader interested in applying the local approach should refer to the cited references. An off-line description of the local approach is first presented, followed by modifications for on-line implementation.

Consider semi-Markov process as in (7.1). Suppose θ is identified via an adaptive algorithm of the form

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \delta_k Z(\hat{\theta}_{k-1}; y_k). \tag{7.46}$$

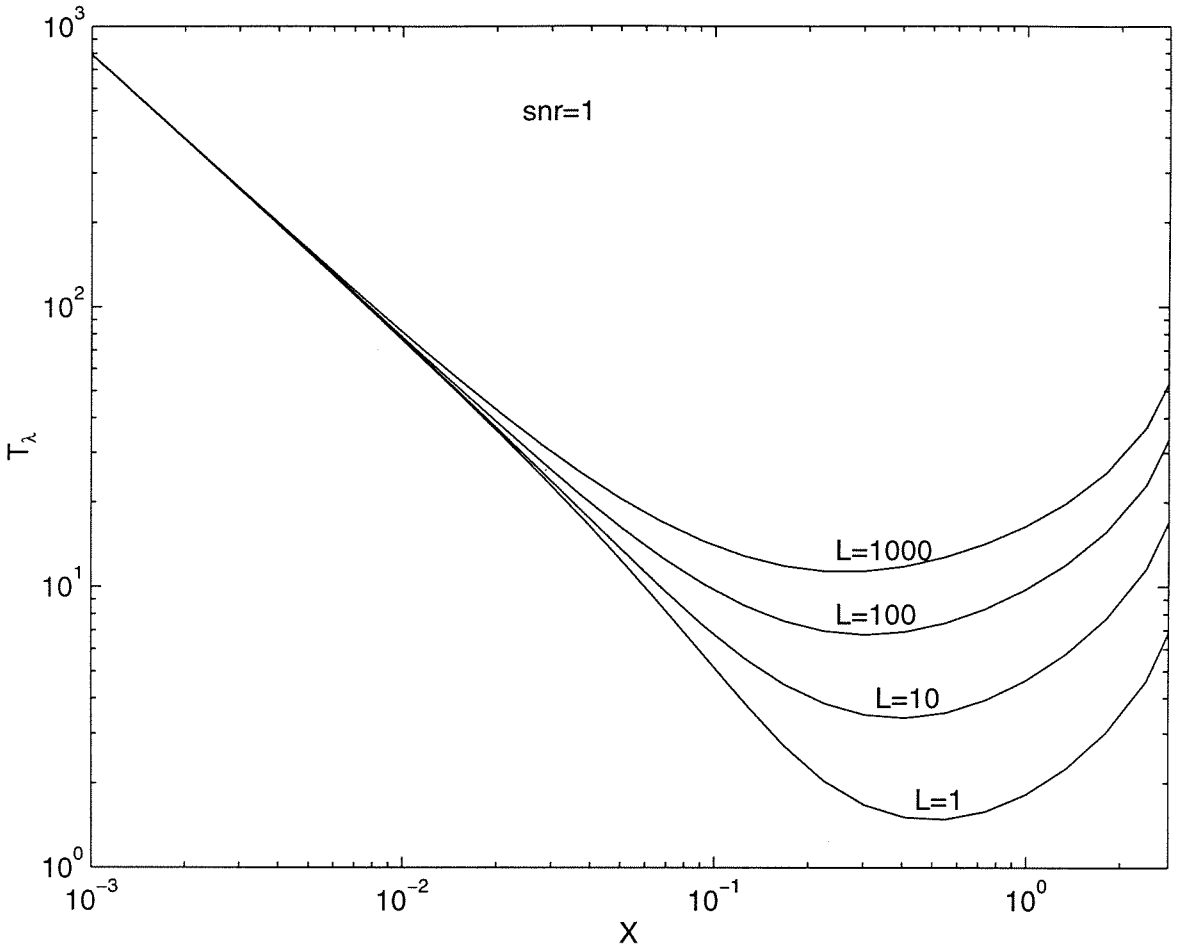

Figure 7.2: Delay for detection T_λ as a function of mean time between false alarms L and tuning parameter X for fixed signal to noise ratio $snr = 1$

Also, assume there exists a nominal model $\theta_* = \theta_0$. In practice, this nominal model is chosen by the user. Also, assume that the true system is controlled by $\theta_* = z$, where z is related to the nominal model through the relation

$$z = \theta_0 + \frac{d\theta}{\sqrt{N}}. \tag{7.47}$$

In other words, as the data record grows, the deviation between the nominal model and the true system converges to zero at a rate which is inversely proportional to the square root of the record length. The nomenclature "local asymptotic approach"

is derived from this assumption as it states that in the asymptotic limit of infinite data record, the change becomes infinitely small [26]. This behavior permits the distribution of the cumulative sum to be specified asymptotically. The following two results describe this behavior for the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1'$:

- $\mathcal{H}_0$: $d\theta = 0$.
- $\mathcal{H}_1'$: there exist $\tau \in [0,1]$ such that:

$$\begin{aligned} d\theta &= 0 \text{ for } k < \tau N, \\ d\theta &\neq 0 \text{ for } \tau N \leq k \leq N. \end{aligned}$$

Theorem 4 *[12]* **Behavior under the hypothesis $\mathcal{H}_0$.** *Define the cumulative sum $D_{N,m}$ as follows:*

$$D_{N,m}(\theta_0, d\theta) = \frac{1}{\sqrt{N}} \sum_{k=1}^{m} Z(\theta_0; y_k).$$

- *In the limit $N \longrightarrow \infty$ the cumulative sum $D_{N,m}$ is normally distributed with mean 0 and covariance $R(\theta_0)$, where $R(\theta)$ is given by*

$$R(\theta) = \sum_{k=-\infty}^{+\infty} \mathrm{cov}_{\theta,\theta} \left[Z(\theta; y_k), Z(\theta; y_0) \right]. \tag{7.48}$$

  where $\mathrm{cov}_{\theta,\theta}$ denotes the covariance with respect to the probability law obtained when the true system and the nominal model are both equal to θ.

- *For $t \in [0,1]$, set*

$$D_{N,t}(\theta_0, d\theta) = D_{N,m}(\theta_0, d\theta) \text{ where } m = [Nt]. \tag{7.49}$$

  Then the process $\{D_{N,t}(\theta_0, 0)\}_{0 \leq t \leq 1}$ converges weakly to the process $R^{1/2}(\theta_0)(B_t)_{0 \leq t \leq 1}$, where (B_t) is a Brownian motion.

Theorem 5 *[12]* **Behavior under the hypothesis $\mathcal{H}_1'$.** *Let $\tau \in [0,1]$ be given. Let*

$$D_{N,t}(\theta_0, d\theta, \tau) = \frac{1}{\sqrt{N}} \sum_{k=1}^{m} Z(\theta_0; y_k), \; m = [Nt]$$

where (y_k) *is a controlled semi-Markov process with control parameter* θ_* *given by*

$$\begin{aligned} \theta_* &= \theta_0 \text{ for } k \leq \min(m, [N\tau]), \\ \theta_* &= \theta_0 + N^{-1/2} d\theta \text{ for } k > \min(m, [N\tau]). \end{aligned} \tag{7.50}$$

Then when N *tends to infinity, the process* $\{D_{N,t}(\theta_0, d\theta, \tau)\}_{0 \leq t \leq 1}$ *converges weakly towards the process* D_t *which is the solution to the linear stochastic differential equation*

$$dD_t = -1_{t \geq \tau} \cdot (h_\theta d\theta)\, dt + R^{1/2} dB_t, \tag{7.51}$$

where $R = R(\theta_0)$ *and* $h_\theta = \lim_{N \to \infty} \frac{\partial E_\theta[Z(\theta_0; y_N)]}{\partial \theta}$.

These two theorems can be interpreted as follows. Consider the random variables $\zeta_k(\theta_0)$ defined by $\zeta_k(\theta_0) = Z(\theta_0; y_k)$. If a change of magnitude $N^{-1/2}$ occurs at time r, and for large enough N, then if the random variables ζ_k were independent and distributed according to the following law:

$$\begin{aligned} \zeta_k(\theta_0) &\approx \mathcal{N}(0, R(\theta_0)) \text{ for } k < r, \\ \zeta_k(\theta_0) &\approx \mathcal{N}(-h_\theta(\theta_0) d\theta, R(\theta_0)) \text{ for } k \geq r, \end{aligned} \tag{7.52}$$

where $\mathcal{N}(x, \Sigma)$ denotes a normal distribution with mean x and variance Σ, then the asymptotic behavior described in Theorem 5 would result. The fundamental idea in the asymptotic approach is therefore to replace the original testing problem of $\mathcal{H}_0$ against $\mathcal{H}_1'$ by the asymptotically equivalent problem of detecting a change in the mean of independent Gaussian ζ_k as in (7.52). In this case, the log likelihood ratio between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1'$ is given by

$$\begin{aligned} S_r^N(d\theta) &= \sum_{k=r}^{N} \zeta_k^T R^{-1} \zeta_k - \sum_{k=r}^{N} (\zeta_k - (h_\theta d\theta))^T R^{-1} (\zeta_k - (h_\theta d\theta)), \\ &= 2 \sum_{k=r}^{N} \zeta_k^T R^{-1} (h_\theta d\theta) - (N - r + 1)(h_\theta d\theta)^T R^{-1} (h_\theta d\theta). \end{aligned} \tag{7.53}$$

Since the change parameter $d\theta$ is unknown both in magnitude and direction, a GLR

approach is used by replacing $d\theta$ by its most likely value:

$$S_r^N = \max_{d\theta} S_r^N(d\theta) = (\Delta_r^N)^T R^{-1} (\Delta_r^N), \tag{7.54}$$

where

$$\Delta_r^N = (N - r + 1)^{-1/2} \sum_{k=r}^{N} \zeta_k. \tag{7.55}$$

To summarize, given an adaptive identification algorithm, the asymptotic local approach for detection of change in the identified parameters at time r consists of first calculating the random variables $\zeta_k = Z(\theta_0; y_k)$ for each observation $y_0, \ldots, y_N$. Note that the value of ζ_k calculated does not correspond to the increments for the identification algorithm since the nominal model θ_0 is always used in the calculation rather than the updated estimate $\hat{\theta}_{k-1}$. Next, the variables ζ_k are used to calculate the GLR statistic S_r^N. Since the change time is also unknown, the stopping rule is obtained by maximizing S_r^N over r to obtain the statistic G_N defined below:

$$G_N = \max_r S_r^N. \tag{7.56}$$

In the asymptotic limit, the sequence S_r^N is distributed according to a central $\mathcal{X}^2$ law with d degrees of freedom, where $\theta \in \Re^d$. Therefore, $\mathcal{X}^2$ tables can be used to choose a threshold with a specified false alarm rate.

In the algorithm outlined above, it is implicitly assumed that the true system can be parameterized by some model within the model set, that the adaptive identification procedure is unbiased, and that the nominal model θ_0 is identified using the same identification algorithm on a training set. When these assumptions do not hold, it is still possible to apply the local asymptotic approach, but slight modifications to the algorithm are needed. For a detailed discussion of this approach, the interested reader is referred to [113].

The formulas (7.55) and (7.56) define the local test and their practical application does not depend on the asymptotic approximation. Therefore, although the asymptotic assumption was used in deriving the test, and the test has the theoretical

optimal property of being uniformly most powerful when the asymptotic assumption holds, the implementation is independent of asymptotic considerations. However, as the length of the data record increases, detection of a given change will become easier.

Although the above exposition has been presented in an off-line framework, it can easily be implemented on-line by calculating G_k for each time k; however, in an on-line implementation, performing the optimization in (7.56) over all $r \in [1, k]$ becomes prohibitive as more data become available. To address this problem, Willsky [111] suggested using a sliding window of possible change times, in which optimization over r is carried out over a fixed length subinterval M of the range $[1, k]$, for example $M_k = [k - n_1, k - n_2]$ with $n_1 > n_2$, yielding the statistic:

$$G'_k = \max_{r \in M_k} S_r^k. \tag{7.57}$$

The constants n_1 and n_2 should be chosen in a manner such that n_2 time steps after a change, enough data is available that a decision may be made with reasonable accuracy, and that after n_1 time steps the decision accuracy is scarcely improved by more data. For high accuracy, both n_1 and n_2 should be large, but the facts that larger n_2 leads to larger detection delay and that larger n_1 increases the computational burden present a tradeoff which frequently can only be addressed through trial and error.

7.6 Batch reactor example

As a case study, consider the application of the proposed detection method consisting of Likelihood Ratio plus PGF (LR+PGF) testing to a batch chemical reactor in which feed stock A is converted to desired product B. This example is taken from [100]. The decomposition of A to B is governed by second order dynamics. In addition, B decomposes to an inert species via a first order reaction. The measurement quantity depends quadratically on both component A and B. Due to change in the catalyst, the rate constant governing the decomposition of A to B is prone to change during

the course of the reaction. Therefore, the goal of the detection scheme is to detect from the available measurement changes in this reaction rate. The system dynamics and detection objective are stated below:

- System Dynamics:

$$\begin{aligned} \frac{d[A]}{dt} &= -\kappa_1[A]^2, \\ \frac{d[B]}{dt} &= \kappa_1[A]^2 - \kappa_2[B], \\ y &= 5[A]^2 + 10[B]^2 + v. \end{aligned} \tag{7.58}$$

- Detect changes in rate constant κ_1.

The method based upon likelihood ratio testing using PGF (LR+PGF) was compared to an alternative algorithm implementing the local asymptotic approach based, using an extended Kalman filter (LA+EKF) for recursive estimation of the combined state and parameter κ_1 to provide the statistics ζ_k. This was carried out first by obtaining an approximate discrete time model of the system using an Euler approximation for the system dynamics. Denoting the state vector $x = ([A], [B], \kappa_1)^T$, the discretized dynamics are given by

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \\ x_{3,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} - x_{3,k}x_{1,k}^2 dt \\ x_{2,k} - (x_{3,k}x_{1,k}^2 - \kappa_2 x_{2,k})dt \\ x_{3,k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} q_k, \tag{7.59}$$

$$= f(x_k) + Gq_k, \tag{7.60}$$

where q_k was modeled by a zero-mean, Gaussian noise with variance $\sigma^2 = 2.5e{-}5$, and $dt = 1.0$. To apply the local asymptotic method, the statistic ζ_k should be calculated based upon the nominal model. Therefore, although the EKF for the combined state and parameter estimation problem was used to update the state estimates of $[A]$ and $[B]$, the estimate of κ_1 was held constant at its nominal value of 0.02. The equations

governing this modified version of the EKF are given by

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k\left(y_k - h(\hat{x}_{k|k})\right), && (7.61)\\
\hat{x}_{3,k|k} &= 0.02, && (7.62)\\
\hat{x}_{k+1|k} &= f(\hat{x}_{k|k}), && (7.63)\\
K_k &= \Sigma_k H_k^T\left[H_k\Sigma_k H_k^T + R_k\right]^{-1}, && (7.64)\\
\Sigma_k &= \Phi_k\left(\Sigma_k - K_k H_k \Sigma_k\right)\Phi_k^T + GG^T\sigma^2, && (7.65)\\
\Phi_k &= \left.\frac{\partial f}{\partial x}\right|_{x=\hat{x}_{k|k}}, && (7.66)\\
H_k &= \left.\frac{\partial h}{\partial x}\right|_{x=\hat{x}_{k|k}}. && (7.67)\\
& && (7.68)
\end{aligned}$$

For both the PGF and the EKF it is necessary to specify an initial distribution. For the EKF, this distribution is restricted to be normal, and is specified by the initial mean $\hat{x}_0$ and covariance Σ_0. For each of the simulations, this initial distribution was specified as:

$$\mathcal{N}\left(\begin{bmatrix} 0.999 \\ 0.001 \\ 0.02 \end{bmatrix}, \begin{bmatrix} 1e-2 & 0 & 0 \\ 0 & 2.5e-3 & 0 \\ 0 & 0 & 1e-4 \end{bmatrix}\right).$$

For ease of comparison, the identical initial distribution was used for the PGF.

For implementing the PGF, the p-vector approach of Sorenson [96] was used. An initial grid consisting of 20 points for each state for a total of 8000 points was employed to cover the Cartesian product $x_i \in [\hat{x}_{i,0} - 3\sigma_i, \hat{x}_{i,0} + 3\sigma_i]$ for $i = 1, 2$ and $x_3 \in [0, 0.1]$ where $\sigma_i = \sqrt{\Sigma_{ii,0}}$. The initial grid points were equally space. In the p-vector approach, the grid evolves according to the system dynamics, so the uniform spacing is subsequently lost except for the variable x_3. However, as the absence of process noise affecting x_1 and x_2 results in the fact that the convolution step only involves x_3, the convolution can be carried out using FFT methods.

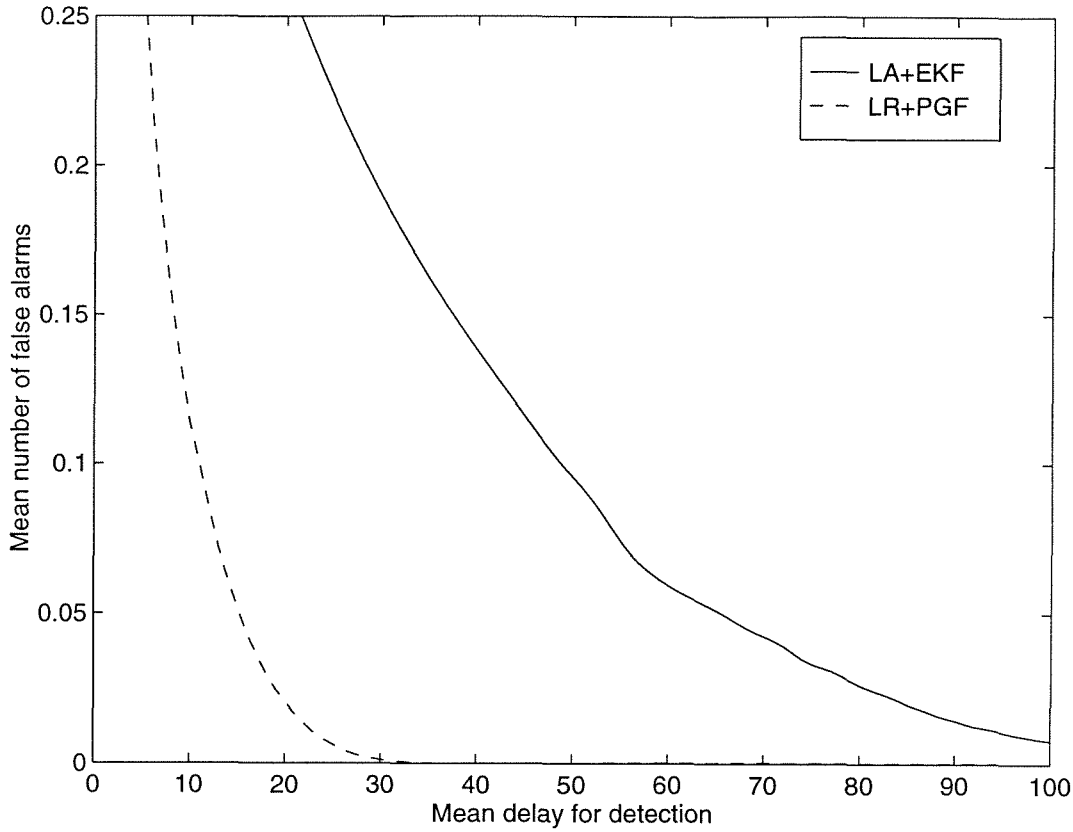

Figure 7.3: False alarms as a function of detection delay, LA+EKF and LR+PGF methods

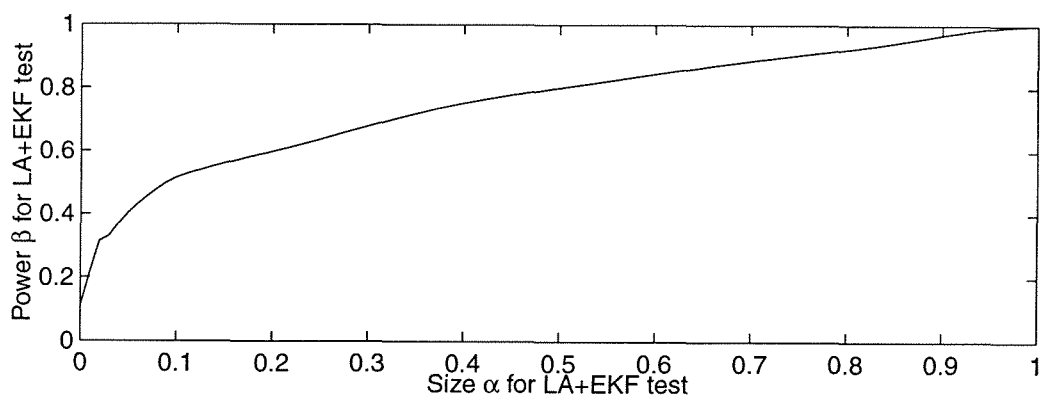

Figure 7.4: Test power β as a function of size α for LA+EKF method

For both the LR+PGF and LA+EKF methods, detection performance criteria were evaluated using Monte Carlo simulation. In each simulation the initial conditions were $[A] = 1, [B] = 0$, the terminal time for the batch was $t = 250$, and measurement noise with variance 0.04 was added to the simulation output. For sequential detection, the relevant performance criteria are mean delay for detection, T_λ, and mean time between false alarms. However, because the batch operation dictates a finite run length for the detection scheme, these criteria must be modified slightly. In the case of the former criterion, if the threshold is significantly large it is possible that the change is not detected prior to the end of the batch run. Let t_r denote the batch run time, t_c the time at which the change occurs, t_d the time at which the change is detected, and P_d the probability that a detection occurs prior to the end of the batch ($P_d = P(t_d < t_r|\mathcal{H}_1')$ where $\mathcal{H}_1'$ is changing value hypothesis). Then the mean delay for detection is given by

$$T_\lambda = E(t_d - t_c|t_d < t_r)P_d + E(t_d - t_c|t_d > t_r)(1 - P_d), \tag{7.69}$$

$$\geq E(t_d - t_c|t_d < t_r)P_d + (t_r - t_c)(1 - P_d). \tag{7.70}$$

When $P_d \neq 1$, T_λ cannot be estimated directly from the results of the Monte Carlo simulations since no estimate of $E(t_d - t_c|t_d > t_r)$ is available; however, the lower bound and T_λ can always be estimated from the simulation runs. In the following, this lower bound will be used when the Monte Carlo simulations indicate that $P_d < 1$. For the latter criterion, difficulties arise estimating the mean time between false alarms in the case where fewer than two false alarms occur during a batch run. For this reason the mean number of false alarms per unit time will be calculated instead. These two measures are related as the mean time between false alarms can be considered as the reciprocal of the mean number of false alarms per unit time.

In order to evaluate the delay for detection T_λ, simulation runs incorporated a change in κ_1 from its nominal value of 0.02 to 0.05 at time $t_c = 100$. Simulation runs used in determining the mean time between false alarms maintained a constant value of $\kappa_1 = 0.02$ throughout the run. In the case of the LA+EKF algorithm, G_k' as in

(7.57) was calculated using a growing data window $M = [1, k - 20]$. In Figure 7.3 the mean number of false alarms per unit time is shown as a function of the mean delay for detection (or its lower bound). The estimation error in T_λ for LR+PGF method is 1.0%, for LA+EKF 1.3%. As is clearly evident, for a fixed false alarm rate, the detection performance of the proposed method LR+PGF gives substantially improved detection performance. For practical considerations the frequency of false alarms must be kept small. Otherwise, the operating personnel will fail to heed the warnings generated by the alarm. (This situation is very similar to that of the classic fairy tale "The Boy Who Cried Wolf.") If the goal is to maintain the average number of false alarms per batch below 1 (1 false alarm per 250 samples), using the LR+PGF method this can be accomplished with a mean detection delay of approximately 25, and 100% detection prior to batch termination. On the other hand, imposing a similar performance criterion on the LA+EKF method results in 63% of changes going undetected prior to batch termination, and a lower bound on the mean detection delay of approximately 110.

The test LA+EKF can also be viewed as an off-line test, in which case the relevant performance measures are the power β and the size α. In Figure 7.4, power is plotted as a function of size for this test. The values for the power and size used in this plot were estimated from the results of the Monte Carlo simulation runs. As can be seen from this figure, using this test a high likelihood of detection can only be achieved with a substantial false alarm rate. This confirms the on-line results discussed above.

7.7 Summary and conclusions

A new approach for change detection problems which encompasses a wide class of problems, including additive and multiplicative faults as well as general parameter changes, has been presented. The approach uses two dynamic models. In the first model, system parameters remain fixed at their predetermined, nominal values, whereas in the second model they are modeled by Brownian motion. When the nominal parameter values accurately describe the system, the first model will result in

a higher likelihood than the second due to the diffusive nature of Brownian motion; however, when a change occurs, the second model will provide a higher likelihood. For each of the two models, the probability of each new measurement conditioned on previous measurements and the model is calculated and used to form a Neyman-Pearson test based on likelihood ratios.

The proposed two model structure allows for arbitrary dynamics as well as system changes of arbitrary magnitude and direction in the parameter space. By contrast, using a GLR approach in the case of unknown parameter value following the change, the optimization of post-change parameter values is usually impractical except in the case of linear dynamics in which case the optimization can be carried out explicitly. This feature presents a significant advantage for the proposed framework over the latter method.

For linear problems including additive faults on linear systems, the method can be applied by running two Kalman filters, one for each model type. In Section 7.4 the properties of the proposed method were evaluated for the problem of change in mean of a Gaussian sequence. For this basic problem, the performance depends on two parameters, the threshold level λ and the ratio $X = \Sigma/R$. Choosing the threshold λ so as to result in a constant mean time between false alarms L, there exists an optimal value of X which minimizes the mean delay for detection for a fixed change size. This suggests the following approach for tuning the algorithm. First, determine an acceptable level for the mean time between false alarms L. Next, specify the magnitude snr_* of the minimum change for which detection is desired. Using these settings, determine the value of X which minimizes the mean delay for detection. For system changes with magnitude larger than snr_*, the mean delay for detection will be decreased without changing the mean time between false alarms. Therefore, this approach will result in an algorithm which minimizes the *worst case* mean delay for detection where worst case is taken with respect to all changes such that $snr \geq snr_*$.

Although it is not possible to carry out an in-depth analysis of the performance properties for a general non-linear change detection problem, the analysis for the simplified scenario provides insight into tuning the algorithm for the general case. The

parameter X depends upon the system to measurement noise ratio Q/R, suggesting that in the general case, the system noise distribution law p_w will affect the detection performance in a similar fashion. Although in some cases it may be possible to determine p_w from historical process data, in many instances it will be necessary to view p_w as a tuning parameter. Restricting p_w to be a zero mean Gaussian distribution, one would expect that the covariance matrix Q_w associated with p_w will play a similar role as in the change in mean detection problem. Initially increasing the elements Q_w from small values will decrease the detection delay, but further increases will result in longer detection delays and ultimately undetectability. Although in general it is doubtful analytical relations expressing the effect of Q_w on detection performance can be derived, often Monte Carlo simulations can be employed during the tuning phase to obtain good values.

Since analytical expressions for recursively updating the conditional probabilities which are needed to evaluate the likelihood ratio test are not available for non-linear problems or problems with non-Gaussian stochastic dynamics, approximate methods must be used. The Bayesian approach, which is consistent with Kalman filtering for the linear case, can be used to obtain computational procedures for evaluating the necessary distributions. The class of algorithms known as probability grid filtering can be used to obtain a global approximation to the conditional probability density. Although these algorithms carry a high computational cost, as computational power continues to increase, implementation for larger systems will become feasible.

The case study in Section 7.6 reveals that the LR+PGF approach to change detection can provide significant improvement over methods based on the local asymptotic approach. Considering the result of Neyman-Pearson, this is not surprising. The most powerful test is based upon a likelihood ratio test, and when the approximations associated with the non-linear filtering algorithm are valid and changes are accurately modeled by Brownian motion, the proposed test should provide an excellent approximation to the likelihood ratio. By contrast, the assumptions used to derive the local asymptotic test are more restrictive, requiring a large data record and a small change. On the other hand, the local asymptotic approach provides a theoretical foundation

for selecting detection thresholds, whereas in the proposed method, threshold selection must be made through simulation studies or on line tuning; however, for problems in which a limited data record is available or rapid detection is necessary, the inferior performance of the local asymptotic method outweighs any potential benefit derived from simpler threshold selection.

Chapter 8 Propositional logic in control and monitoring problems

Summary

By using linear constraints on integer variables, logical inference can be incorporated within Model Predictive Control and model based detection problems resulting in mixed integer quadratic or mixed integer linear programming problems. Qualitative knowledge can be used to prioritize control objectives, or to improve performance of detection schemes. This chapter demonstrates some capabilities of combining logic using integer variables with quantitative models.

8.1 Introduction

The vast majority of modern control methods implement in one fashion or another a model of a system. Some methods, such as model predictive control (MPC), use the model directly to calculate control moves from measured data. Other methods, such as LQG, $\mathcal{H}_\infty$, and μ-based control use the model to synthesize a linear time invariant control law. It has been widely recognized by the control community that an inherent limitation in model based designs lies within the problem of robustness to model uncertainty. Namely, due to differences between the mathematical model and the real physical system, methods based upon an imprecise model may perform unsatisfactorily. In the area of monitoring and detection, this is especially true. For example, in the area of fault detection, there exists no satisfactory method to date to design model based detection schemes in the presence of model uncertainty.

In addition to the quantitative models used to describe a system, often other information is available about the behavior of the system. Such qualitative information

may exist in the form of relationships which are implied when other relationships hold. In addition, practical control objectives are often qualitative. For example, an operator may desire that the feed to a chemical reactor remain near a set-point when the reactor pressure is low, but when the reactor pressure is high, the throughput should be used to reduce the pressure. Current control methods are inadequate at incorporating such qualitative features.

In the area of process monitoring, qualitative information can often be useful in diagnosing failures. Initial approaches attempting to incorporate qualitative and quantitative models have typically focused on using a quantitative model to generate a residual signal which should be nominally zero. Then qualitative methods are used to diagnose the residual. See for example [56], [55], and [104]. Although this approach reportedly helps improve robustness to modeling errors, it is lacking in the fact that the qualitative knowledge cannot be used to generate the residuals.

In this chapter, a systematic method for incorporating qualitative information in control and monitoring problems is presented. The method consists of first stating qualitative information in terms of propositional logic. Then, using integer variables, the propositions are translated into linear constraints. For control problems, the qualitative knowledge may easily be incorporated within the Model Predictive Control framework by appending the constraints to the control calculation problem. Similarly, qualitative knowledge can be used in fault detection problems by estimating faults while enforcing the logical constraints.

In Section 8.2, previous results are outlined which demonstrate how to convert logical propositions into linear constraints and how to pose logical inference as an integer programming problem. In Section 8.3, alternative approaches to translating propositions to constraints which may affect the complexity of the ensuing optimization problem are discussed. In Section 8.4, control and monitoring problems using combined qualitative and quantitative information are presented. Optimization techniques for the solution of the type of mixed integer programming problems resulting from these control and monitoring problems are outlined in Section 8.5.

8.2 Representation of logic

Cavalier et al. [17] and Post [87] have shown how propositional logic can be represented by formulating linear constraints on integer variables. In this framework, reasoning or logical inference can be shown to be equivalent to solving an integer programming (IP) problem. More recently, Raman and Grossman [90] have used this framework in process synthesis to combine qualitative reasoning and quantitative models to obtain mixed integer mathematical programs. In the area of design, this approach is very useful as qualitative approaches have been successful for preliminary design considerations, whereas quantitative approaches are better suited for analyzing interactions between variables and arriving at an optimal configuration.

In this section, we will briefly review the representation of propositional logic using linear constraints. The building block of propositional logic is the literal, a variable which can assume the value of true or false. With each literal $\mathcal{L}$, we can associate another literal, $\neg\mathcal{L}$, the negation of $\mathcal{L}$, such that the proposition $\neg\mathcal{L}\vee\mathcal{L}$ is tautological. A proposition is formed by combining literals with the operators AND ($\wedge$), OR ($\vee$), and IMPLIES ($\Rightarrow$). An important result in propositional logic is that any proposition can be reduced to conjunctive normal form, consisting of a conjunction of clauses [67]:

$$C_1 \wedge C_2 \wedge \ldots \wedge C_n \tag{8.1}$$

where each clause C_i is a disjunction of literals of the form

$$\mathcal{L}_1 \vee \mathcal{L}_2 \vee \ldots \vee \mathcal{L}_m. \tag{8.2}$$

The conjunctive normal form can be obtained systematically by applying the following rules:

1. Replace implications with the equivalent disjunction:

$$(\mathcal{L}_1 \Rightarrow \mathcal{L}_2) \text{ becomes } (\neg\mathcal{L}_1 \vee \mathcal{L}_2) \tag{8.3}$$

Table 8.1: Propositional Logic Constraints

Logical Relation	Expression	Linear Inequalities
OR	$\mathcal{L}_1 \vee \mathcal{L}_2 \vee \ldots \vee \mathcal{L}_n$	$l_1 + l_2 + \ldots + l_n \geq 1$
AND	$\mathcal{L}_1 \wedge \mathcal{L}_2 \wedge \ldots \wedge \mathcal{L}_n$	$l_1 \geq 1;\ l_2 \geq 1;\ \ldots; l_n \geq 1$
$\Rightarrow$	$\mathcal{L}_1 \Rightarrow \mathcal{L}_2$	$l_1 - l_2 \leq 0$
$\Leftrightarrow$	$\mathcal{L}_1 \Leftrightarrow \mathcal{L}_2$	$l_1 - l_2 = 0$
XOR	$\mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \ldots \oplus \mathcal{L}_n$	$l_1 + l_2 + \ldots + l_n = 1$

2. Use DeMorgan's Theorem to distribute negation:

$$\neg(\mathcal{L}_1 \wedge \mathcal{L}_2) \quad \text{becomes} \quad (\neg\mathcal{L}_1 \vee \neg\mathcal{L}_2) \tag{8.4}$$

$$\neg(\mathcal{L}_1 \vee \mathcal{L}_2) \quad \text{becomes} \quad (\neg\mathcal{L}_1 \wedge \neg\mathcal{L}_2) \tag{8.5}$$

3. Distribute OR over AND using the tautologies

$$[(\mathcal{L}_1 \wedge \mathcal{L}_2) \vee \mathcal{L}_3] \text{ becomes } [(\mathcal{L}_1 \vee \mathcal{L}_3) \wedge (\mathcal{L}_2 \vee \mathcal{L}_3)]\,. \tag{8.6}$$

To convert propositions into mathematical expressions, we first define an integer variable $l_j \in \{0, 1\}$ associated with each literal $\mathcal{L}_j$. If $\mathcal{L}_j$ is true, then $l_j = 1$, otherwise $l_j = 0$. Clearly, negation can be represented by $(1 - l_j)$. Additionally, disjunction of several literals is equivalent to constraining the sum of the corresponding integers to exceed or equal 1. For example, $\mathcal{L}_1 \vee \mathcal{L}_2$ is equivalent to $l_1 + l_2 \geq 1$. Finally, conjunction can be expressed as a set of inequality constraints, i.e., $\mathcal{L}_1 \wedge \mathcal{L}_2$ becomes $l_1 \geq 1,\ l_2 \geq 1$. By associating integer variables with each literal and combining the rules for negation, disjunction, and conjunction, any proposition in conjunctive normal form can be expressed as a set of linear inequality constraints. Table 8.1 summarizes linear inequality relations for the most important logical relations.

The framework outlined above is suitable for describing hard logical facts about a system. However, frequently the system knowledge is in the form of heuristics or rules of thumb. In this case, it is possible that the knowledge is violated, but our experience tells us that it usually is not. Post [87] allows for uncertain knowledge of this type by

introducing a new variable which represents violation of the heuristic. Therefore, if C_i is a logical clause corresponding to a heuristic, a new literal V_i is introduced which corresponds to violation of the heuristic, leading to the new proposition:

$$C_i \vee V_i. \tag{8.7}$$

Letting v_i denote the integer variable associated with V_i, corresponding linear inequalities can be established. In order to discriminate between strong rules and heuristic rules, penalties are associated with the violation of each heuristic. The penalty w_i associated with v_i should be a nonnegative number whose magnitude reflects the level of uncertainty associated with the heuristic. Thus, small values of w_i indicate that the corresponding heuristic is relatively weak, whereas large values correspond to a stronger hypothesis. Clearly, the selection of penalties is an important design consideration.

Using the framework described above, logical inference with uncertain knowledge can be formulated as an IP problem whose objective is to find a feasible point which minimizes the violation of heuristics. Formally, the IP is given by

$$\begin{aligned} \min \quad & \sum_i w_i v_i \\ \text{s.t.} \quad & G_1 l + v \geq H_1; \text{ (heuristic knowledge)} \\ & G_2 l \geq H_2; \text{ (hard knowledge)} \\ & l \in \{0,1\}^n, \ v \geq 0, \end{aligned} \tag{8.8}$$

where G_i and H_i are appropriately dimensioned matrices and vectors respectively. In applications in control and monitoring, the situation will arise in which the value of process variables and measurements, represented by continuous variables, will determine the value of a literal. For example, consider the case where distillation column pressure is monitored. When the pressure exceeds a certain limit, action must be taken to relieve it. Let P denote the measured pressure, P_T the upper threshold, $\mathcal{L}_P$ the literal corresponding to "high pressure," and l_P the corresponding integer

variable. We would like to implement a constraint such that the following relation holds:

$$(P \geq P_T) \Rightarrow \mathcal{L}_P. \tag{8.9}$$

In the case where a priori upper and lower bounds for P are known, which will be the case for any physical quantity, this relation can be implemented by the following constraint:

$$(P - P_T) - Ul_P \leq 0, \tag{8.10}$$

where U is an upper bound for $(P - P_T)$. Raman and Grossman [90] have suggested that any constraint of the form $f(x) \leq 0$, where x varies continuously and where $f(x) \in [L_f, U_f]$, can be associated with a binary variable l_f through the following inequality:

$$L_f l_f + \epsilon \leq f(x) \leq U_f(1 - l_f), \tag{8.11}$$

where ϵ is a small positive tolerance. Thus, if $f(x) \leq 0$ is satisfied, the variable l_f assumes the value of 1. Conversely, when $f(x) \geq \epsilon$, $l_f = 0$. Since a constraint of the form $f(x) = 0$ can be considered as a combination of $f(x) \leq 0$ and $-f(x) \leq 0$, equality constraints can also be cast into the present framework. This is done by introducing one binary variable l_- corresponding to the former inequality, another variable l_+ corresponding to the latter, and a third binary variable $l_=$ which are related through the following proposition:

$$l_= \Leftrightarrow l_+ \wedge l_-. \tag{8.12}$$

Integer variables can also be used to represent functions which may be defined in a piecewise manner. For example, suppose that a function $\phi(x)$ is given by

$$\phi(x) = \begin{cases} f_1(x); & x \in X_1 \\ f_2(x); & x \in X_2 \\ \vdots & \\ f_n(x); & x \in X_n \end{cases}, \tag{8.13}$$

where X_i are disjoint polyhedra. Each region X_i can be described by a set of linear constraints $G^i x \leq H^i$, and we associate a binary variable l_i with each region X_i through the proposition

$$x \in X_i \Leftrightarrow l_i = 1. \tag{8.14}$$

Since only one region may be valid, the variables l_i also must satisfy the following XOR relation:

$$l_1 + l_2 + \ldots + l_n = 1. \tag{8.15}$$

Let G^i_j and H^i_j represent the j^{th} row of G^i and H^i respectively. One way to formulate the relation (8.14) would be as follows:

$$(G^i_j x - H^i_j) - M(1 - l_i) \leq 0, \forall i, j \tag{8.16}$$

where M is a suitably large positive constant. These constraints combined with (8.15) will guarantee that $l_i = 1$ if and only if $x \in X_i$.

The function $\phi(x)$ can now be represented by the following set of constraints:

$$L_i(1 - l_i) \leq \phi - f_i(x) \leq U_i(1 - l_i) \tag{8.17}$$

where L_i and U_i are lower and upper bounds on the $\phi - f_i(x)$. For practical problems, such bounds will always exist. If $l_i = 1$, then the above constraint requires $\phi = f_i(x)$, whereas if $l_i = 0$, the constraint is automatically satisfied. By relating continuous quantities such as ϕ in (8.17) and P in (8.10) to the value of integer variables representing logical propositions, mixed integer programming problems are obtained.

8.3 Constraint formulation and complexity

An important issue affecting the complexity of the integer programming problem which must be solved involves the formulation of the logical constraints. More often than not, there exist many different ways to formulate equivalent conditions. As an example, consider the proposition $\mathcal{L}_1 \Leftrightarrow \mathcal{L}_2 \wedge \mathcal{L}_3$. Two equivalent formulations are

given by

$$\begin{aligned} l_2 + l_3 - l_1 &\leq 1 \\ l_2 + l_3 - 2l_1 &\geq 0 \end{aligned} \tag{8.18}$$

and

$$\begin{aligned} l_2 + l_3 - l_1 &\leq 1 \\ l_2 - l_1 &\geq 0 \\ l_3 - l_1 &\geq 0. \end{aligned} \tag{8.19}$$

Rather than using conjunctive normal form, Williams [109] has suggested a substitution method for forming sets of constraints corresponding to logical propositions. In this approach, new variables are introduced to represent conjunction and disjunction. Thus, in the above, the variable $\mathcal{L}_1$ would be substituted for the conjunction $\mathcal{L}_2 \wedge \mathcal{L}_3$ via the constraint (8.19). For disjunctions, $\mathcal{L}_1 \Leftrightarrow (\mathcal{L}_2 \vee \mathcal{L}_3)$ is enforced via the constraints on the corresponding binary variables:

$$\begin{aligned} l_2 + l_3 - l_1 &\geq 0 \\ l_2 - l_1 &\leq 0 \\ l_3 - l_1 &\leq 0. \end{aligned} \tag{8.20}$$

The constraint sets (8.19) and (8.20) can then be used as building blocks to form more elaborate propositions.

As noted by Cavalier et al. [17], whether conjunctive normal form is more advantageous than the substitution method depends upon the type of knowledge considered. For example, production rule type logic statements of the form

$$(\mathcal{L}_1 \vee \mathcal{L}_2 \vee \ldots \vee \mathcal{L}_n) \Rightarrow \mathcal{L}_{n+1} \tag{8.21}$$

can be expressed in conjunctive normal form using $n+1$ variables and n constraints. By contrast, the substitution method leads to a set of constraints requiring $2n$ variables and $3n-2$ constraints. On the other hand, a clause of the form

$$(\mathcal{L}_1 \wedge \mathcal{L}_2) \vee (\mathcal{L}_3 \wedge \mathcal{L}_4) \vee \ldots \vee (\mathcal{L}_{n-1} \wedge \mathcal{L}_n) \tag{8.22}$$

requires $3n/2$ variables and $3n/2+1$ constraints using the substitution method whereas the conjunctive normal form requires n variables and $2^{n/2}$ constraints. From these examples, we can conclude that the better modeling approach depends upon the structure of the clause.

Cavalier et al. [17] also addressed the issue of preprocessing the constraints in order to remove redundant constraints and obtain a more concise and efficient formulation of the problem. The following results are based on conjunctive normal form. Every constraint in conjunctive normal form can be expressed in binary variables in the form

$$\sum_{i \in I} l_i - \sum_{j \in J} l_j \geq 1 - |J| \tag{8.23}$$

where $|J|$ denotes the cardinality of the index set J. Using this expression, the following result may be used to reduce the complexity of constraints posed in conjunctive normal form.

Theorem 6 *The constraint (8.23) dominates any constraint of the form*

$$\sum_{i \in M} l_i - \sum_{j \in N} l_j \geq 1 - |N| \tag{8.24}$$

for any sets $M \supseteq I$ and $N \supseteq J$. In particular, the constraint

$$l_i \geq l_j \tag{8.25}$$

dominates any constraint of the form (8.24) when M and N are disjoint and $i \in M$ and $j \in N$.

8.4 Application to control and detection

In the previous sections, it was shown how integer variables and linear constraints can be used to represent any relationship which can be expressed in propositional logic. In addition, it was shown how propositions which depend on continuous variables can be related to these variables through linear constraints. In the remainder of this chapter, we will demonstrate several applications of this method for incorporating qualitative information in problems in the area of control and detection. In particular, we consider incorporating logical constraints within the Model Predictive Control (MPC) [73] and Moving Horizon Estimation (MHE) [92] frameworks.

Many formulations of MPC algorithms exist. To review the relevant literature in this area is beyond the scope of the present work; however, all MPC algorithms have in common the feature that control moves are calculated by solving on-line an optimization problem. This optimization consists of minimizing the predicted output trajectory subject to constraints on the control inputs and predicted output. In this work, we will consider the following optimization problem which can be related to an MPC problem:

$$\min_{u \in \mathcal{U}(\hat{x}(k|k))} \sum_{i=1}^{P} [\hat{x}(k+i|k) - x^*(k+i)]' R [\hat{x}(k+i|k) - x^*(k+i)] + u(k+i-1)' Q u(k+i-1). \tag{8.26}$$

The variable x^* is a known reference state. The region $\mathcal{U}$ is determined by linear constraints of the form:

$$G_u u_k^{k+P} + G_x \hat{x}(k|k) \leq H, \tag{8.27}$$

and $u_k^{k+P} = [u(k), \ldots, u(k+P)]$, and $\hat{x}(k+i|k)$ satisfies the dynamic constraints:

$$\hat{x}(k+j|k) = A\hat{x}(k+j-1|k) + Bu(k+j-1), \; j \geq 1. \tag{8.28}$$

It is possible to include constraints on the control increments Δu, but for our discussion, the above formulation suffices. In MHE, dynamic state estimates are obtained

by solving a similar optimization problem:

$$\min_{w \in \mathcal{W}} \quad \Delta\hat{x}(k-p+1)'\Sigma^{-1}\Delta\hat{x}(k-p+1) + \sum_{i=k-p+1}^{k} [y(i)-\hat{y}(i|k)]' R^{-1} [y(i)-\hat{y}(i|k)] + w(i|k)'Q^{-1}w(i|k) \quad (8.29)$$

subject to the constraints

$$\hat{x}(k-p+1|k) = \hat{x}(k-p+1|k-1) + \Delta\hat{x}(k-p+1) \quad (8.30)$$

$$\hat{x}(i+1|k) = A\hat{x}(i|k) + B_w\hat{w}(i|k),\ i = k-p+1, ..., k-1 \quad (8.31)$$

$$\hat{y}(i|k) = C\hat{x}(i|k),\ i = k-p+1, ..., k-1, \quad (8.32)$$

where $\hat{w}$ is interpreted as a disturbance estimate and $\Delta\hat{x}$ as the error in the initial state estimate. Since the objective is quadratic and the constraints are linear for both of these problems, the optimization corresponds to a standard quadratic programming problem. Alternative implementations employ an l_1 or l_∞ norm in the objective functions and result in linear programming problems.

By formulating logic propositions as linear integer constraints, it is possible to incorporate logic based control decisions within the MPC and MHE frameworks. The resulting optimization problem belongs to the class of mixed integer quadratic programming (MIQP), or mixed integer linear programming (MILP), depending on the type of objective function used. Many commercial optimization packages are capable of solving mixed integer optimization problems [75]. One algorithm for solving MIQP problems by iterating between quadratic programming (QP) and MILP subproblems is described in Appendix 8.5. Ding and Sargent [27] have recently developed an efficient branch and bound algorithm for MIQP's. Therefore, the class of optimization problems resulting from incorporating logical constraints can be solved by a variety of readily available methods.

With the addition of logic based constraints, the class of schemes which can be formulated as MPC and MHE problems is enlarged. In the following sections, several examples of the capabilities of this approach are enumerated. In Section 8.4.1, it is

shown how problems in which the control objective depends on the operating region can be formulated. In Section 8.4.2, the prioritization of multiobjective control problems is addressed. In Section 8.4.3, fault detection using combined qualitative and quantitative information is formulated in the current framework.

8.4.1 Logic dependent objective

In many practical control applications, the objective that the controller should address is dependent upon the operating characteristics. This situation can be handled by formulating a problem in which the objective that the controller seeks to meet depends upon the satisfaction of certain logical conditions. The following instructive example serves to clarify this application.

Consider a semi-batch reactor equipped with a cooling jacket. In order to maximize conversion, the feed addition rate F should be maintained at the set point $F = 1$. In addition, for performance and safety considerations the reactor temperature should be regulated at $T_r = T^*$. The steady state reactor temperature may be related to the feed rate F and the coolant flow rate Q through the linearized relations:

$$T = \frac{\partial T}{\partial F}F + \frac{\partial T}{\partial Q}Q + T_0, \tag{8.33}$$

where $T = T_r - T^*$. To simplify the discussion, we will consider the case wherein this linear relationship holds, the partial derivatives are constant, and the variables are scaled such that $\frac{\partial T}{\partial F} = 1$ and $\frac{\partial T}{\partial Q} = -1$. The manipulated variables F and Q are each constrained to lie within the range $[0, 1]$. The value T_0 can be viewed as a disturbance load.

With the manipulated variables F and Q constrained to lie between 0 and 1, the reactor can operate at its optimal feed rate and temperature only when the load T_0 lies within the range $[-1, 0]$. For smaller values of T_0, the reactor will necessarily run cold. For larger values of T_0, the reactor can operate at the optimal temperature only by decreasing the feed rate below its optimal point. For $T_0 > 1$, the reactor will run hot even when the feed rate is reduced to 0.

Table 8.2: Objective $J = (1-F)^2 + \alpha T^2$

T_0	Q	F	T	J
$T_0 \leq -1$	0	1	$1+T_0$	$\alpha(1+T_0)^2$
$-1 < T_0 \leq 0$	T_0+1	1	0	0
$0 < T_0 \leq 1+\frac{1}{\alpha}$	1	$1-\frac{\alpha}{1+\alpha}T_0$	$\frac{1}{1+\alpha}T_0$	$\frac{\alpha}{1+\alpha}T_0^2$
$1 < T_0$	1	0	T_0-1	$1+\alpha(T_0-1)^2$

If this system were controlled with a conventional MPC algorithm, the control moves dF and dQ would be calculated by minimizing an objective containing terms of the form

$$\sum_{j=1}^{P} J(k+j) = \sum_{j=1}^{P} (1-F(k+j))^2 + \alpha T(k+j|k)^2. \tag{8.34}$$

In steady state, the values of F and Q would then be determined by the solution to the quadratic program

$$\min_{F\in[0,1],Q\in[0,1]} (1-F)^2 + \alpha T^2. \tag{8.35}$$

Substituting $T = F - Q + T_0$, the steady state solution can be determined using the Kuhn-Tucker criteria. This solution is shown in Table 8.2. When the load $T_0 \leq 0$, the steady state behavior is as desired for any choice of α; however, when the load is positive, T is small in steady state only when α is large. When Q is not saturated, a large value of α results in F deviating significantly from 1 for extended time. Therefore, using a traditional objective one can guarantee small steady state temperature offset when Q saturates only at the expense of poor response in F for all times.

By including simple logic into the controller, this problem can be remedied. The logic should reflect the fact that when Q saturates, the manipulated variable F should be used to control the temperature. Otherwise, F should be kept near its optimal value of 1 and Q should control the temperature. We will therefore consider a logic based objective of the following form:

1. If Q is not saturated, minimize $(1-F)^2 + \alpha T^2$.

2. If Q saturates, minimize βT^2.

Using integer variables, this objective can be implemented through the following constraints:

$$L(1-I_q)-\epsilon \leq X_1(k) \leq U(1-I_q)$$

$$L(1-I_q)-\epsilon \leq X_2(k)-\sqrt{\beta}T(k) \leq U(1-I_q)$$

$$LI_q-\epsilon \leq X_1(k)-(1-F(k)) \leq UI_q$$

$$LI_q-\epsilon \leq X_2(k)-\sqrt{\alpha}T(k) \leq UI_q$$

$$Q_{sat}(I_q-1) \leq Q(k)-Q_{sat}+\epsilon_q \leq \epsilon_q I_q$$

The last constraints require that $I_q = 1$ if and only if $Q(k)$ is within ϵ_q of the saturation level Q_{sat}. The first four constraints require that X_1 and X_2 assume the following values:

$$X_1(k) = \begin{cases} 0, & I_q = 1 \\ (1-F(k)), & I_q = 0 \end{cases} \tag{8.36}$$

$$X_2(k) = \begin{cases} \sqrt{\beta}T(k), & I_q = 1 \\ \sqrt{\alpha}T(k), & I_q = 0 \end{cases} \tag{8.37}$$

By replacing the objective term (8.34) with terms of the form $J = X_1^2 + X_2^2$, and incorporating the above constraints involving I_q, a MIQP is obtained with the desired logic based objective. The solution satisfying these constraints can be calculated by using Kuhn-Tucker criteria for each of the two cases $I_q = 0$ and $I_q = 1$. The solution for $I_q = 0$ is the same as in Table 8.2. For $I_q = 1$, the optimal solution is shown in Table 8.3. The MIQP optimum is shown in Table 8.4.

A dynamic simulation with a control and prediction horizon of 5 was conducted for both the MIQP based control and the traditional MPC controller. First order temperature dynamics are used:

$$T = \frac{0.5z}{z-0.5}F - \frac{0.2z}{z-0.8}Q + T_0. \tag{8.38}$$

Table 8.3: Objective $J = \beta T^2$, $Q = 1$

T_0	Q	F	T	J
$T_0 \leq 0$	1	1	T_0	βT_0^2
$0 < T_0 \leq 1$	1	$1 - T_0$	0	0
$1 < T_0$	1	0	$T_0 - 1$	$\beta(T_0 - 1)^2$

Table 8.4: MIQP, $\alpha < \beta < \alpha(1 + \alpha)$

T_0	Q	F	T	J
$T_0 \leq -1$	0	1	$1 + T_0$	$\alpha(1 + T_0)^2$
$-1 < T_0 \leq 0$	$T_0 + 1$	1	0	0
$0 < T_0 \leq 1$	1	$1 - T_0$	0	0
$0 < T_0 \leq 1 + \frac{1}{\alpha}$	1	0	0	$\beta(T_0 - 1)^2$
$1 + \frac{1}{\alpha} < T_0$	1	0	$T_0 - 1$	$\beta(T_0 - 1)^2$

For both controllers, the control action was rate limited by $|dF| < 0.1$ and $|dQ| < 0.1$. For the MIQP controller, $\alpha = 0.1$ and $\beta = 0.105$, whereas for the traditional MPC controller $\alpha = 10$ which is necessary for small steady state temperature offset when Q saturates. The simulation results are shown in Figures 8.1 and 8.2.

For the MIQP based controller, the flow rate remains close to $F = 1$ even when significant changes occur in the load as long as Q does not saturate. At time 100 the load T_0 exceeds 0 and F must decrease in order to provide zero steady state offset in T.

For the traditional MPC controller, each change in the load T_0 results in a significant deviation in F from its nominal value. This is due to the large weight on T in the objective function and the fact that T responds more rapidly to F than to Q. When the load exceeds 0 after time 100, the feed rate F decreases, but there remains a small steady state offset in the temperature T. This offset can be reduced by increasing α at the expense of deteriorated flow control.

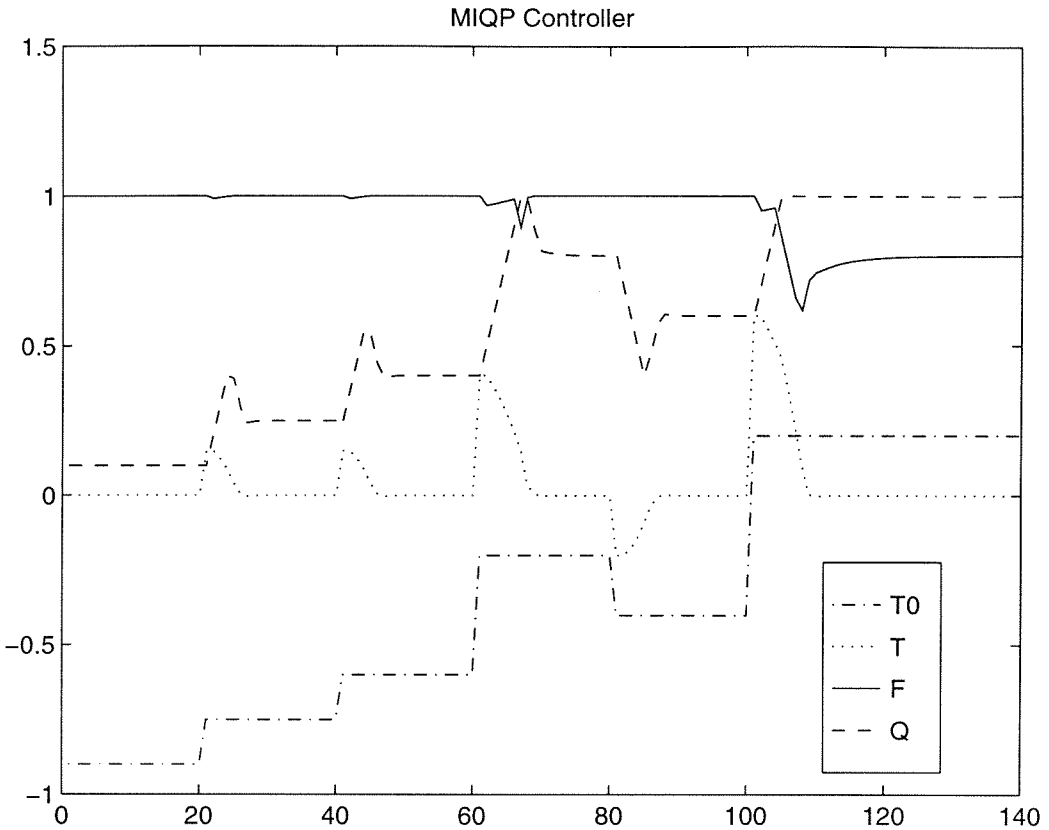

Figure 8.1: Simulation for MIQP based controller

8.4.2 Multiobjective prioritization

An engineer designing a multivariate control system is essentially faced with a multiobjective optimization problem [89]. The control system should achieve process safety by maintaining all necessary variables within safe limits, optimize economic considerations such as production levels and product quality, and ensure environmental regulations are met. In general, the various objectives of the control system do not have equal priority. For example, the most significant objective may be to maintain the process within safe operating limits, at the expense of profitability if necessary. If large fines are imposed when environmental regulations are violated, satisfying regulatory demands will take preference over productivity.

In order to quantify process control requirements, they are commonly stated as either *optimization objectives* or as *constraints*. An example of the former would be to

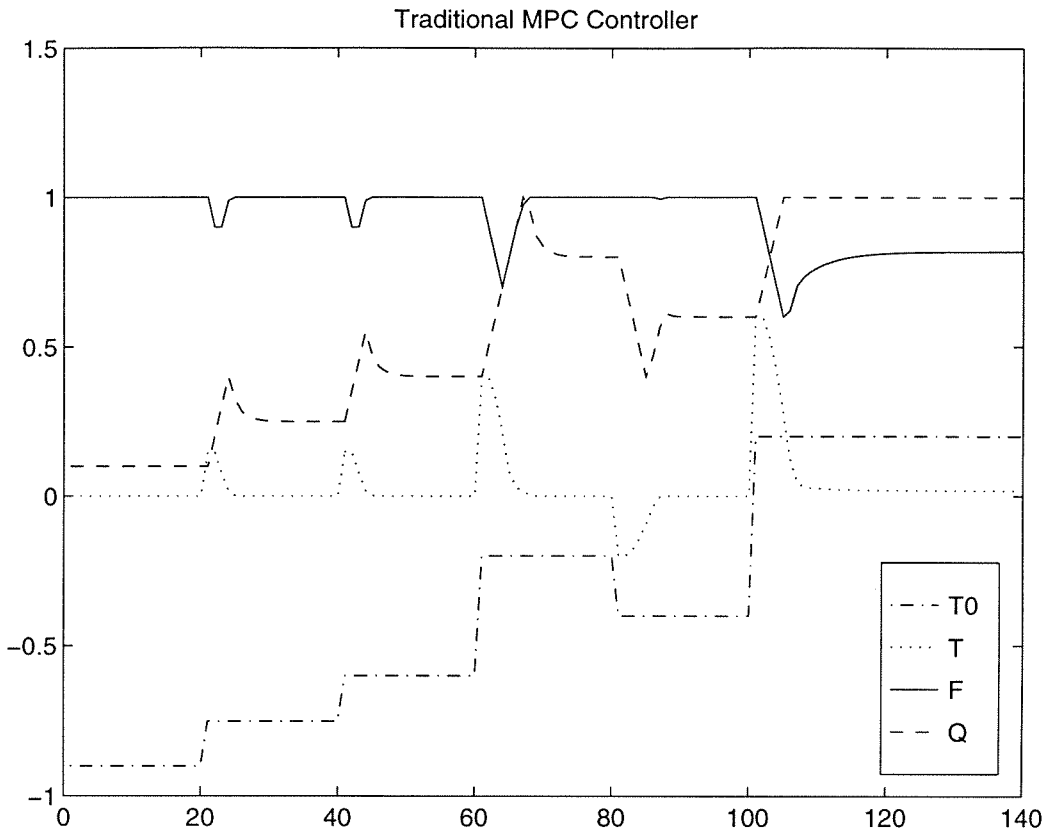

Figure 8.2: Simulation for traditional MPC controller

minimize the integral square error of a measured variable from its set-point, whereas the latter would correspond to maintaining the variable within a specified range. Constraints can be considered to be either *hard* or *soft* [114]. Hard constraints refer to those constraints which cannot be violated and usually correspond to physical limitations of the system. For example, hard constraints can be used to represent limitations in the capacity or in the rate of change of an actuator. Soft constraints are those which can be temporarily relaxed if operating conditions so dictate and are most frequently used to describe objectives on output variables. Imposing a soft upper bound on a temperature, for example, can be interpreted as requiring that the temperature remain below the bound if at all possible, where the realm of possibility is defined by the hard constraints.

The vast majority of the previous literature for multivariate, multiobjective con-

trol problems has relied upon constructing a single utility function which seeks to incorporate each of the objectives. The most common approach has been to set the utility function equal to a weighted sum of the multiple objectives [89, 74, 88]. Soft constraints are transformed to hard constraints through the introduction of slack variables v, for example the soft constraint $x \leq x_u$ becomes $x \leq x_u + v$, and a weighted norm on the slack variables is included in the utility function. The weights of the various objectives comprising the overall utility function must be chosen by trial and error, through numerical simulations, or via other ad hoc approaches in order that more important objectives are given higher priority. The design of weights is an arbitrary procedure, guided by the heuristic that larger weights should be used for more important objectives, and there is no guarantee that a choice of weights will result in the desired prioritization of objectives for all operating conditions. Although the use of a single utility function has led to a useful framework for theoretical analysis of control systems as well as a plethora of successful industrial applications, it lacks the ability to address in a systematic and effective manner the design tradeoffs among conflicting objectives.

Let us formalize the criteria that a designer wishes to satisfy for a multiobjective control algorithm. We will assume that there exist N objectives, each of which can be expressed via constraints on the output variables. The i^{th} constraint is expressed as

$$G^i y \leq H^i, \tag{8.39}$$

where y is the vector of process measurements, G^i is an appropriately dimensioned real vector, and H^i is a real number. Set point tracking objectives can be formulated in this framework using constraints of the form $|y_j - r_j| \leq \epsilon$, where ϵ is a small tolerance. The objectives are ordered according to their relative priorities, with objective i receiving a higher priority than objective $i+1$. The goal of the controller is to satisfy as many objectives as possible while requiring that objectives with higher priority are met first. If all objectives can be satisfied, the goal is achieved; otherwise, the controller should minimize the amount by which the highest non satisfied objective is violated.

Meadowcroft et al. [66] introduced the Modular Multivariate Controller to address the problem of multivariate multiobjective control. This approach incorporated *lexicographic goal programming.* The fundamental concept of this approach is to solve sequentially a set of optimization problems, with each problem corresponding to minimizing the slack variable associated with violating a performance objective. The highest priority objective is solved first, and the optimal value of the slack variable associated with this objective is used to form a bound on acceptable performance for the first objective. A constraint corresponding to this bound is then enforced while attempting to minimize the second objective. The procedure is repeated until as many objectives as possible have been satisfied. In each step of the procedure, a primary manipulated variable is specified, and one attempts to meet the performance criteria using only this variable while constraining the other variables to predetermined values. If the objective cannot be met using the primary manipulated variable, a second manipulated variable is added to the search space, and so on. At each step of the algorithm, all primary variables for higher priority objectives must be recalculated.

Although the Modular Multivariate Controller addressed the issue of prioritization of objectives, the algorithm used is confusing and can be quite complicated for large dimensioned systems with a high number of objectives. In the following, it is shown that the constraint prioritization problem posed above can easily be solved through the use of integer variables. First with each objective, associate an integer variable l_i, $i = 1, ..., N$. When $l_i = 1$, the constraint corresponding to the concomitant objective must be satisfied; otherwise, the constraint may be violated. This condition can be expressed in the form of the following modified constraint:

$$G^i y \leq H^i + M^i(1 - l_i), \tag{8.40}$$

where M^i is a (conservative) upper bound on $G^i y - H^i$ which ensures that the constraint is always satisfied when $l_i = 0$. The term $M^i(1 - l_i)$ can be thought of as a slack when $l_i = 0$. Often, a single integer variable l_i can be used for more than one constraint. This will be true when the constraints involved cannot be simultaneously

violated. For example, in the case where the constraints form upper and lower bounds (u_j and ν_j) on an output y_j, at least one of the upper and lower bounds will be met all the time. The modified constraints can be written as

$$y_j \leq u_j + m^i(1 - l_i), \quad -y_j \leq -\nu_j + m^i(1 - l_i). \tag{8.41}$$

If G^i is viewed as a 2 row matrix and H^i by a 2-vector, and M^i a 2-vector with equal elements, these constraints can be expressed in the form of (8.40). For generality, we consider priority constraints of this form (8.40) where G^i is a matrix, H^i a vector, and M^i a vector with equal elements, and the structure of G^i and H^i are such that only one row of the inequality can be violated at a time.

By specifying constraints of the form (8.40), the designer ensures that objective i is met whenever $l_i = 1$. We now address the issue of prioritization. Prioritization of objectives implies that objectives 1 through i must be satisfied before attempting to satisfy objective $i + 1$. In other words, if constraint $i + 1$ is enforced with zero slack, then constraint i must also be enforced with zero slack. This condition can be easily expressed through the following constraints on the integer variables l:

$$-l_{i+1} + l_i \geq 0, \ i = 1, ..., N - 1. \tag{8.42}$$

The following mixed integer programming problem can then be used to satisfy the multiobjective criteria:

$$\min \left(-\sum_{i=1}^{N} l_i \right) \text{ subject to (8.40) and (8.42).} \tag{8.43}$$

Now suppose that due to the operating regime it is not possible to meet each performance objective simultaneously. Let i_f denote the index of the first priority which failed. In this situation, we desire that the controller come as close as possible to meeting objective i_f. This can be done by introducing a slack variable ϕ satisfying

the following system of constraints:

$$G^i y \leq H^i + \phi + M^i \left[(i-1) + l_i - \sum_{j=1}^{i-1} l_j \right], \; i = 1, ..., N. \tag{8.44}$$

Consider the following objective function:

$$\min \left(-M_\phi \sum_{i=1}^{N} l_i + f(\phi) \right) \text{ subject to (8.40) and (8.42) and (8.44)} \tag{8.45}$$

where f is a positive strictly increasing function and M_ϕ denotes a strict upper bound on f. Since M_ϕ forms an upper bound on f, the objective can always be made smaller by increasing the number of non-zeros l's than by decreasing ϕ. Once as many l's are one as possible, ϕ will be minimized. Since $l_i = 1$ for each $i < i_f$, the constraint (8.44) will be automatically satisfied for $i < i_f$ for any value of ϕ. For $i = i_f$, the term multiplying M^i in (8.44) is zero, whereas for $i > i_f$, this term is greater than 1. Therefore, the only active constraint involving ϕ corresponds to i_f, and the minimal ϕ is equal to the slack in objective i_f, and the solution of this problem satisfies the multiobjective criterion.

The nature of the above optimization problem depends upon the relationship between the manipulated variables u and the output y involved in the constraints, as well as the function f. For the case of linear input-output systems, the constraints will all be linear in the manipulated variables u. Therefore, if f is a linear function of ϕ, a MILP is obtained. For f a quadratic function in ϕ, a MIQP results. Although it is possible to extend the approach above to the case where ϕ is not a scalar, when ϕ is a positive scalar, the identical solution is obtained whether $f(\phi) = \phi$ or $f(\phi) = \phi^2$. Due to the fact that MILP's can typically be solved more efficiently than MIQP, the former criterion is preferred.

In some cases, particularly for large systems, the situation may arise where, for example, objectives $1, ..., i_f - 1$ may be met, objective i_f cannot be satisfied, but after minimizing the slack associated with i_f, it is possible to meet objectives with lower priority than i_f. In such a situation, it is clearly desirable that the control algorithm

yield such a solution. In this case, the optimal multiobjective solution can be found by repeated application of the above optimization. After each iteration, the bound H^{i_f} is increased by ϕ and the minimum is recomputed. By increasing H^{i_f}, the subsequent solution is guaranteed to satisfy $l_{i_f} = 1$. The iteration may be terminated when there are no degrees of freedom left in the manipulated variables, or when $i_f = N - 1$.

In order to apply the above procedure to MPC, two alternatives exist. In the first alternative, a separate value $l_i(k+j)$ is used for each time in the prediction horizon. In addition, each of the constraints (8.40), (8.42), and (8.44) are formulated for each time in the horizon using the forecast model for y. Such a method results in a total of $N * P$ integer variables and a large number of constraints, and for large problems this method may become computationally intractable for on-line implementation. By using blocking strategies, the computational demand may be decreased, but for large enough systems computational limits will again be reached.

The second alternative uses only steady state relations between the inputs and the outputs in formulating the priorities. Using a cascade structure, the multiobjective optimization problem is then solved for a steady state solution in the outer loop, and the resulting values of the outputs and inputs are then passed as reference signals to the inner loop which calculates the dynamic control output for the desired reference.

To illustrate this approach consider the system with two inputs and three outputs whose step response is shown in Figures 8.3 and 8.4. The steady state response of this system is given by the following equation:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}, \tag{8.46}$$

where the vector d corresponds to a disturbance which is either measured or estimated. The performance criteria in order of decreasing importance are given as follows:

1. $\max_i |y_i| \leq 2$

2. $|y_2 - r_2| \leq \epsilon$

3. $|y_3| \leq 1$

4. $|y_1 - r_1| \leq \epsilon$

The first objective states that all of the outputs should be kept within 2 measurement units of 0. When this can be satisfied, y_2 should then be controlled to its set-point r_2. Once this goal has been met, it is desirable to keep the magnitude of y_3 below 1. Finally, y_1 should be controlled to its set-point r_1. In the following, we will assume throughout that $r_1 = r_2 = 0$. In addition to the output objectives, the hard constraints $|u_i| \leq 1$ must be enforced at all times. Table 8.5 expresses the constraints which need to be enforced in the current framework to achieve the multiobjective optimal solution.

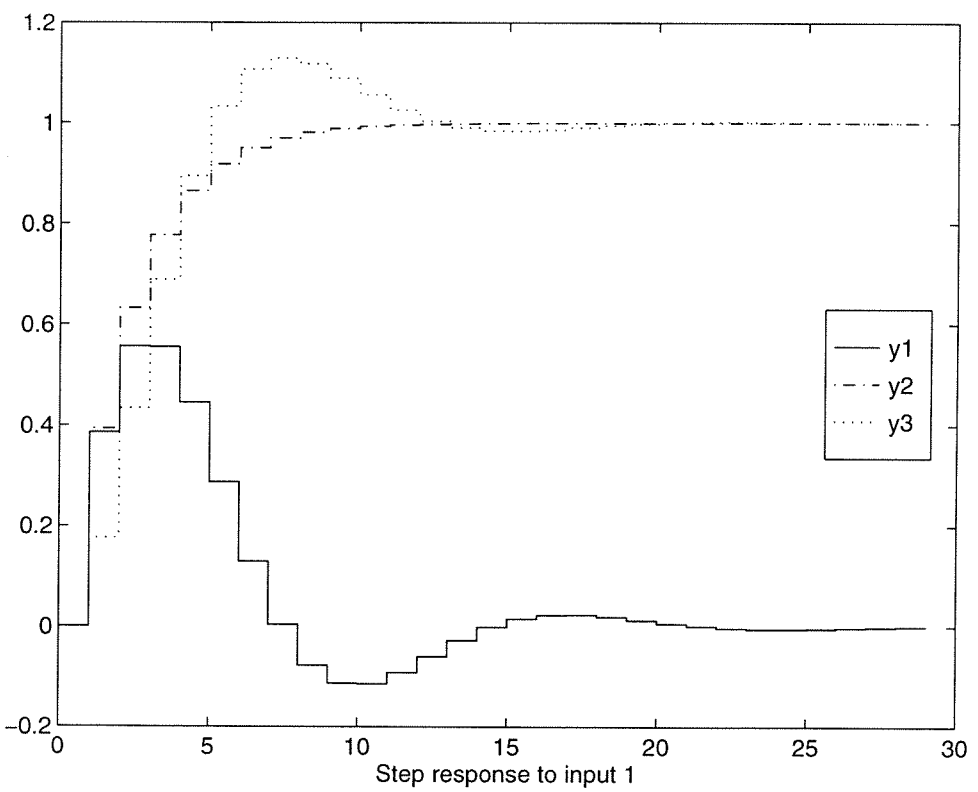

Figure 8.3: Step response to input 1 for dynamic system

The proposed multiobjective method was compared to a traditional approach using a single utility function consisting of a weighted sum of squares of the individual

Table 8.5: Constraints for Multiobjective Controller

Objective 1	$\lvert y_1\rvert \leq 2$	$u_2 + d_1 \leq 2 + 10(1 - l_1)$ $-u_2 - d_1 \leq 2 + 10(1 - l_1)$
	$\lvert y_2\rvert \leq 2$	$u_1 + d_2 \leq 2 + 10(1 - l_1)$ $-u_1 - d_2 \leq 2 + 10(1 - l_1)$
	$\lvert y_3\rvert \leq 2$	$u_1 + u_2 + d_3 \leq 2 + 10(1 - l_1)$ $-u_1 - u_2 - d_3 \leq 2 + 10(1 - l_1)$
Objective 2	$\lvert y_2\rvert \leq \epsilon$	$u_1 + d_2 \leq \epsilon + 10(1 - l_2)$ $-u_1 - d_2 \leq \epsilon + 10(1 - l_2)$
Objective 3	$\lvert y_3\rvert \leq 1$	$u_1 + u_2 + d_3 \leq 1 + 10(1 - l_3)$ $-u_1 - u_2 - d_3 \leq 1 + 10(1 - l_3)$
Objective 4	$\lvert y_1\rvert \leq \epsilon$	$u_2 + d_1 \leq \epsilon + 10(1 - l_4)$ $-u_2 - d_1 \leq \epsilon + 10(1 - l_4)$
$\min \phi$	$u_2 + d_1 \leq 2 + \phi + 10l_1$ $-u_2 - d_1 \leq 2 + \phi + 10l_1$ $u_1 + d_2 \leq 2 + \phi + 10l_1$ $-u_1 - d_2 \leq 2 + \phi + 10l_1$ $u_1 + u_2 + d_3 \leq 2 + \phi + 10l_1$ $-u_1 - u_2 - d_3 \leq 2 + \phi + 10l_1$ $u_1 + d_2 \leq \epsilon + \phi + 10(1 - l_1 + l_2)$ $-u_1 - d_2 \leq \epsilon + \phi + 10(1 - l_1 + l_2)$ $u_1 + u_2 + d_3 \leq 1 + \phi + 10(2 - l_1 - l_2 + l_3)$ $-u_1 - u_2 - d_3 \leq 1 + \phi + 10(2 - l_1 - l_2 + l_3)$ $u_2 + d_1 \leq \epsilon + \phi + 10(3 - l_1 - l_2 - l_3 + l_4)$ $-u_2 - d_1 \leq \epsilon + \phi + 10(3 - l_1 - l_2 - l_3 + l_4)$	
Integer Constraints	$l_i \in \{0, 1\},\ i = 1, ..., 4$	
Prioritize Objectives	$-l_2 + l_1 \geq 0$ $-l_3 + l_2 \geq 0$ $-l_4 + l_3 \geq 0$	

objectives and slack variables. First, for the disturbance $d = [-0.5, 1, 2.5]$ weights for the traditional method were obtained through trial and error in order to yield a response which was close to the optimal multiobjective method. The results for both methods are shown in Figures 8.5 and 8.6. The optimal steady state solution corresponds to $u = [-0.5, -1]$ and $y = [-1, 0, 1]$. If the same weights are used for a second disturbance $d = [-0.5, -0.5, 2.5]$, the responses shown in Figures 8.7 and 8.8 result. Here the optimal multiobjective solution is given by $u = [0.5, -1]$ and $y = [-1.5, 0, 2]$. Although the integer variable method gives the proper response, the

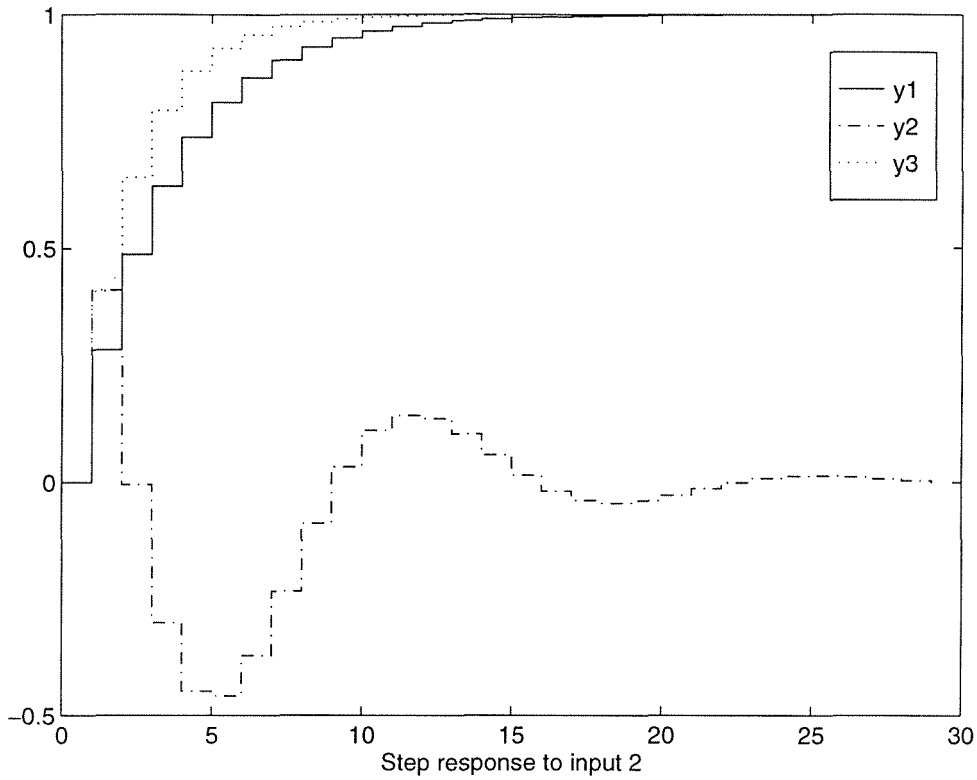

Figure 8.4: Step response to input 2 for dynamic system

traditional approach does not work well using the previous weights. Because of the large weight associated with the goal $|y_3| \leq 1$, reference tracking of y_2 is sacrificed. Choosing weights for the traditional method such that the performance is satisfactory in both cases is an unwieldy task. For the first disturbance, the first two objectives assist each other. By driving y_2 to the origin, the magnitude of y_3 is also reduced. By contrast, for the second disturbance, y_2 can only be driven to zero at the expense of increasing the magnitude of y_3.

We note that the semi-batch reactor control problem posed in Section 8.4.1 could also be formulated as a multiobjective problem. In this case, the top priority would be to maintain the temperature at its set-point, with the second priority to maximize the flow rate.

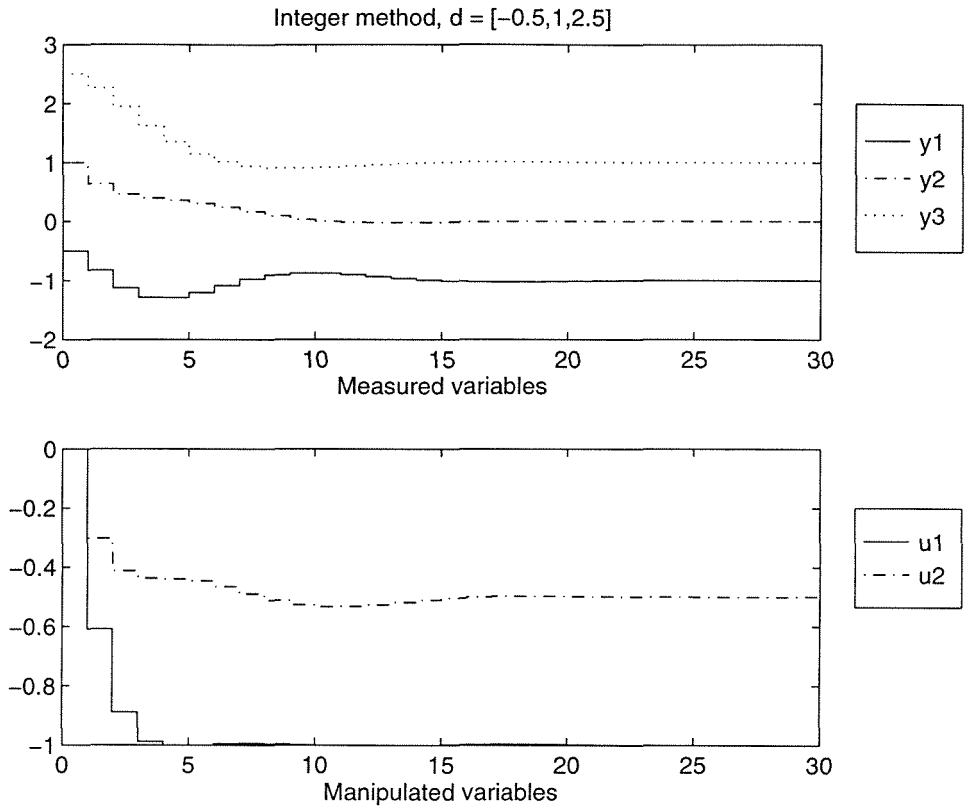

Figure 8.5: Integer method controller. Response to $d = [-0.5, 1, 2.5]$.

8.4.3 Symptom aided detection

In the area of process monitoring, the goal is to determine why a process is not behaving normally. Although the model based approach works well when there are minimal modeling errors or when the process is not operating in closed loop, when feedback control is used to rectify failures and the model between the control action and the measurement is uncertain, model based approaches are susceptible to high false alarm rates. An alternative approach is based on symptoms. For example, in an exothermic reactor coupled with a cooling jacket, excessive temperature may be symptomatic of failure in the cooling jacket. Symptoms may be represented by binary variables, and combined with models to yield a detection scheme which is both symptom and model driven.

This approach is perhaps best illustrated through an example. We consider here

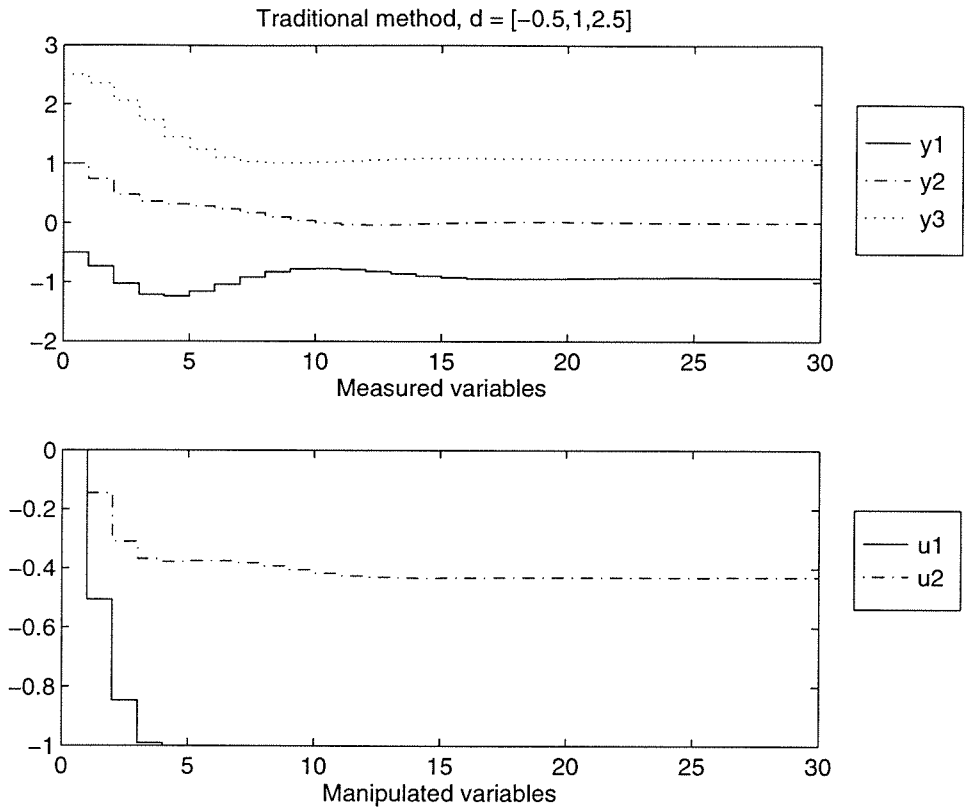

Figure 8.6: Single object MPC controller. Response to $d = [-0.5, 1, 2.5]$.

fault detection in a cold tandem steel mill. A detailed description of the process can be found in [101]. In this process, nine measurements are made, corresponding to four tensions $t_{f1}, \ldots, t_{f4}$ and five thicknesses $h_1, \ldots, h_5$. The process failures considered manifest themselves as changes in the friction coefficient μ of one of five different rollers. Let F_i^+ and F_i^- denote respectively the literals corresponding to an increase and a decrease in value of the friction coefficient μ_i associated with roller i, and f_i^+ and f_i^- the corresponding integer variables. With each of the five faults, there exists a pattern of changes in the measured variables, as shown in Table 8.6. In this table, $\uparrow$ denotes an increase in the variable of interest, whereas $\downarrow$ denotes a decrease.

Here, we assume that it is improbable for multiple faults to occur simultaneously. The rules state that if a fault F_i has occurred, a certain pattern must follow in the measurements. Let us define binary variables τ_i and η_i to correspond to the

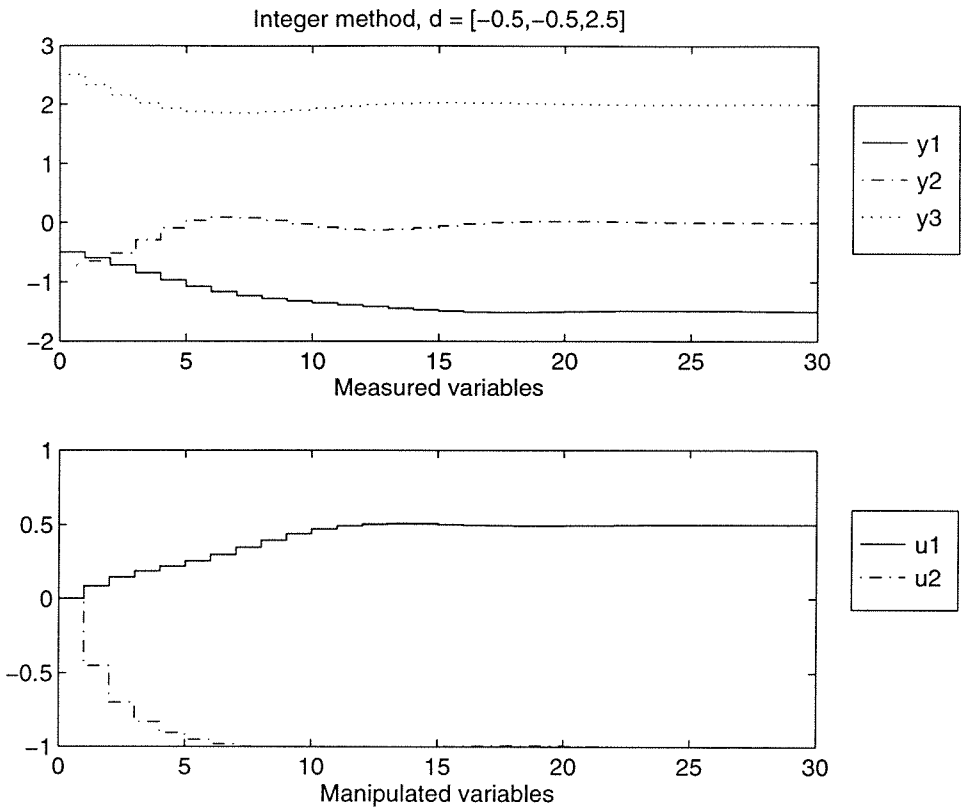

Figure 8.7: Integer method controller. Response to $d = [-0.5, -0.5, 2.5]$.

literals $t_{fi} \downarrow$ and $h_i \downarrow$. The symptoms $t_{fi} \uparrow$ and $h_i \uparrow$ then correspond to $(1 - \tau_i)$ and $(1 - \eta_i)$. The value of τ_i and η_i are determined by the sign of the mean of the respective variables over the optimization horizon. Then the rules in Table 8.6 may be expressed in terms of linear constraints on the integer variables. For example, the rule involving F_1^- may be written as

$$-f_1^- + \tau_1 \geq 0 \tag{8.47}$$

$$-f_1^- + \eta_1 \geq 0 \tag{8.48}$$

$$-f_1^- - \eta_2 \geq -1. \tag{8.49}$$

In the case where the fault patterns represent heuristic knowledge which may be violated, an integer variable v_1^- should be added to the left-hand side of each of

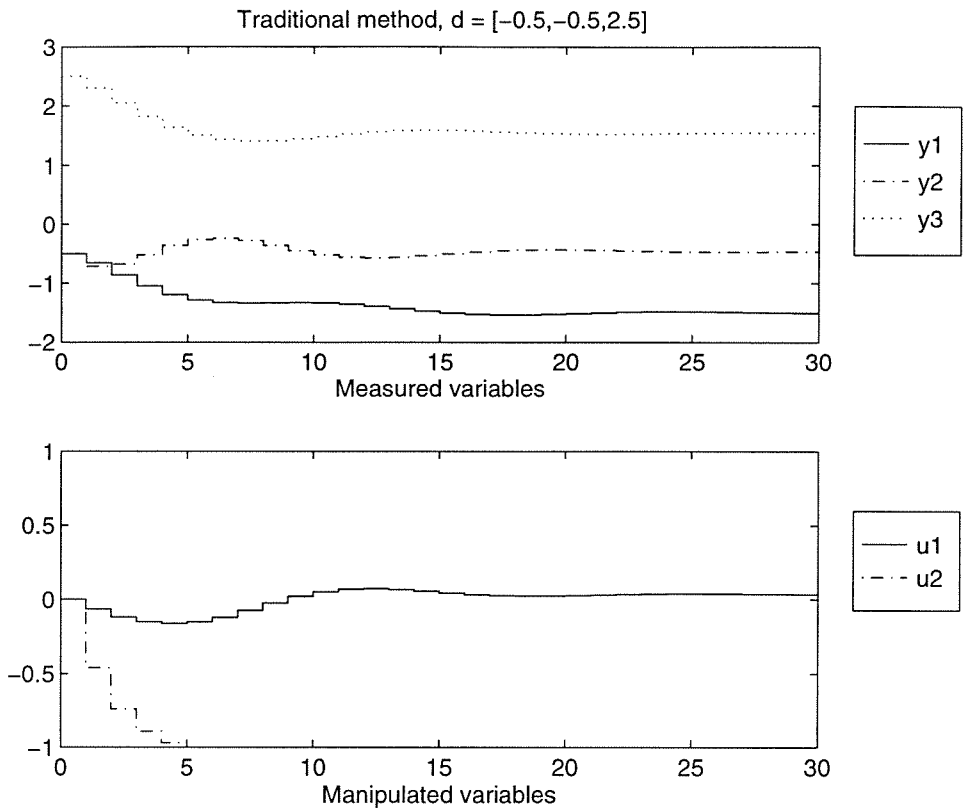

Figure 8.8: Single objective MPC controller. Response to $d = [-0.5, -0.5, 2.5]$.

the three constraints. Then, if the fault F_1^- occurs, either the associated pattern is observed, or the heuristic is violated ($v_1^- = 1$).

We would now like to combine the knowledge representation of the system with a model in order to detect faults. Let us assume that the relationship between the measurements y, the manipulated variables u, and the faults μ can be expressed via a dynamical input/output system of the form

$$y = P_u * u + P_\mu * \mu. \tag{8.50}$$

Thus, we assume that the effect of the manipulated variables u and the faults μ are uncoupled. The knowledge system described above is based upon the model P_μ. Consider the case wherein the system operates in closed loop. In this situation, whenever a fault occurs the control system generates a signal u which attempts to

Table 8.6: Fault Patterns

Fault	Pattern	Fault	Pattern
F_1^-	$(t_{f1}\downarrow, h_1\downarrow, h_2\uparrow)$	F_1^+	$(t_{f1}\uparrow, h_1\uparrow, h_2\downarrow)$
F_2^-	$(t_{f1}\downarrow, h_1\uparrow, h_2\downarrow)$	F_2^+	$(t_{f1}\uparrow, h_1\downarrow, h_2\uparrow)$
F_3^-	$(t_{f2}\downarrow, h_2\uparrow, h_3\downarrow)$	F_3^+	$(t_{f2}\uparrow, h_2\downarrow, h_3\uparrow)$
F_4^-	$(t_{f3}\downarrow, h_3\uparrow, h_4\downarrow)$	F_4^+	$(t_{f3}\uparrow, h_3\downarrow, h_4\uparrow)$
F_5^-	$(t_{f4}\downarrow, h_4\uparrow, h_5\downarrow)$	F_5^+	$(t_{f4}\uparrow, h_4\downarrow, h_5\uparrow)$

compensate for the change in the friction coefficient. Thus, the closed loop response typically will not conform to the rules based on open loop performance. However, when a model $\hat{P}_u$ of the system P_u is available, we can consider the signal Δy given by

$$\begin{aligned} \Delta y &= y - \hat{P}_u * u \\ &= P_\mu * \mu + \left(P_u * u - \hat{P}_u * u\right). \end{aligned} \tag{8.51}$$

The term in brackets represents modeling error. In the case of perfect modeling, the signal Δy will exactly correspond to the response from the faults. However, as modeling error is inevitable, Δy will be corrupted whenever u is not identically zero. One goal of including information on the fault pattern is to be able to distinguish between the faults and modeling errors. We assume that the modeling errors do not result in consistent patterns in the outputs; if they did, we could use this information to obtain a better model. Thus, if Δy is significantly different than zero and matches one of the fault patterns, then we postulate that the non-zero value of Δy is due to a fault, but if the pattern is inconsistent with any of the faults, model errors are to blame.

Let us now consider fault detection using the quantitative model. In particular, we consider the model $\Delta y = P_\mu * \mu$. One approach involves modeling μ as a random walk and using moving horizon estimation to estimate the value of μ from the measurements y. Denoting w_i as the increments for the variables μ_i, we estimate μ by

solving the following optimization:

$$\min_{w} \sum_{j=k-p+1}^{k} \|\Delta y(j) - \hat{\Delta y}(j)\|^2 + \|w(j)\|^2, \tag{8.52}$$

where $\Delta y(j)$ represents the measurements minus the model output, $\hat{\Delta y}(j)$ represents the output from the model $P_\mu * \mu$, and $\|\cdot\|$ a is vector norm. For the case wherein the norm is a weighted l_1 norm, i.e., $\|x\| = \sum_{i=j}^{n} r_i|x_j|,$, $r_i > 0$, the combined qualitative/quantitative model based detection problem becomes a mixed integer linear program. When a Euclidean norm is employed, a mixed integer quadratic program is obtained.

To integrate the qualitative and quantitative approach, one needs to relate the binary fault variables f_i to the increments w_i. Since F_i^+ corresponds to an increase in the value of μ_i, we know that if w_i is positive, then F_i^+ must be true. On the other hand, if F_i^+ is true, then w_i must be positive. Similarly, F_i^- should be true if and only if w_i is negative. These relationships can be incorporated through the following set of constraints:

$$-Uf_i^- - w_i(j) \geq -U \tag{8.53}$$

$$Lf_i^+ + w_i(j) \geq L \tag{8.54}$$

$$b_i - w_i(j) \geq 0 \tag{8.55}$$

$$b_i + w_i(j) \geq 0 \tag{8.56}$$

$$Uf_i^- + Uf_i^+ - b_i \geq 0, \tag{8.57}$$

where L is a lower bound for w_i and U is an upper bound on the magnitude of w_i. The first constraint requires that $w_i(j)$ is negative for all j whenever F_i^- is true. The second constraint requires that $w_i(j)$ is positive whenever F_i^+ is true. Together these two constraints prohibit f_i^+ and f_i^- from both being true at the same time. The third and fourth constraints state that b_i is an upper bound for the magnitude of $w_i(j)$ over the horizon. The final constraints state that if b_i is positive, then either F_i^+ or F_i^- must hold. The constraints do not prevent b_i from attaining a non-zero value even when

$w_i = 0$. By including b_i multiplied by a small positive weight in the objective function, we can be certain that this scenario does not happen since for any w_i the objective will be minimized by choosing b_i to be the smallest upper bound. Therefore, if b_i is non-zero, w_i is non-zero, and F_i must be true. The mixed quantitative/qualitative fault detection problem corresponds to solving the mixed integer program (either MILP or MIQP) consisting of minimizing the objective (8.52) subject to the logic constraints implied by the patterns in Table 8.6 and the constraints (8.53) thru (8.57).

Let us now consider a numeric example consisting of a step decrease of 10% is the second friction coefficient at time 10. The linear model described in [101] is used to carry out the detection performance. Model uncertainty was introduced by using +20% and −20% errors in the first and second input gains respectively. In addition, the full nonlinear model was used in the simulation whereas the local linearization was employed for the detection computations. A quadratic error objective was used, resulting in a MIQP. The optimization was run on a window corresponding to the first 10 samples following the introduction of the fault. In the case where no heuristic violations were allowed, the resulting solution for the logic variables was $f_2^- = 1$, $f_4^+ = 1$ with all other $F's$ zero. The quantitative estimates for the changes in the friction coefficients were $\Delta\mu_2 = -0.10$ and $\Delta\mu_4 = 0.0006$, yielding an optimal objective value of 0.27.

To consider the effects of allowing heuristic violation, the MIQP was solved repeatedly while forcing heuristic violation for each of the zero valued logical fault variables, one at a time. The resulting decrease in the objective function indicates the maximum magnitude by which the violation variable V could be weighted. The results are shown in Table 8.7. For example, by allowing the rule corresponding to F_3^- to be violated, the objective would decrease by −0.0007. If the weight on the heuristic violation variable v_3^- were larger than 0.0007, the MIQP would be minimized by enforcing the heuristic. By choosing a weight smaller than 0.0007, the designer effective places very little confidence in this heuristic.

When neither sufficient data or quantitative models are available to generate fault symptoms, qualitative simulation can be used. A significant literature exists on the

Table 8.7: Change in Objective Value from Allowing Heuristic Violation

Fault Type	Δ Objective
F_1^-	0.00
F_3^-	-0.0007
F_4^-	-0.0006
F_5^-	-0.0006
F_1^+	-0.0013
F_2^+	1.28
F_3^+	0.00
F_5^+	0.00

use of qualitative modeling in detection problems. In qualitative simulation, the effects of faults are causally propagated from one variable to another, ultimately satisfying steady-state process constraints. The qualitative representation of these relations can be expressed using confluence equations [23]. Oyeleye and Kramer [84] have used confluence equations in developing a qualitative model based detection scheme. Sign directed graphs [45, 71, 42, 106] have also been proposed as useful tools for qualitative process modeling. Confluence equations as well as sign directed graphs can be easily converted to a series of if-then statements, and therefore recast into the mixed integer optimization framework developed in this paper. Future research will address the efficacy of this approach.

8.5 Optimization methods

In this section, solution methods for mixed integer quadratic programming are considered. In Section 8.5.1 a general procedure for solving mixed integer problems, the generalized Benders decomposition, is reviewed. Section 8.5.2 follows with the specific application of this decomposition to quadratic programming problems. The resulting MIQP algorithm consists of iterations between MILP and QP problems.

8.5.1 Generalized Benders decomposition

Mixed integer programming problems can be efficiently solved using the generalized Benders decomposition (GBD). In this section, a brief description of the method and the underlying theory will be presented.

The GBD can be used to solve mathematical programming problems of the form

$$\min_{x\in X, y\in Y} f(x,y) \text{ subject to } G(x,y) \leq 0, \tag{8.58}$$

where y is a vector of complicating variables in the sense that the above optimization is a much easier problem in x when y is considered fixed, and $f(x,y)$ and $G(x,y)$ are convex on X for fixed values of y. The key idea of the GBD is to exploit computational efficiencies which can be achieved by considering the problem in y-space rather than in xy-space. This is done by projecting the original optimization onto y-space as follows:

$$\min_{y} v(y) \text{ subject to } y \in Y \cap V, \tag{8.59}$$

where

$$v(y) = \inf_{x\in X} f(x,y) \text{ subject to } G(x,y) \leq 0 \tag{8.60}$$

and

$$V = \{y : G(x,y) \leq 0 \text{ for some } x \in X\}. \tag{8.61}$$

One can easily see that $v(y)$ is the optimal value of the original problem (8.58) for fixed y. For the GBD to be an efficient solution algorithm, evaluating $v(y)$ should be much easier than solving (8.58) itself. The set V constitutes those values of y for which the right-hand side of (8.60) is feasible.

Geoffrion [39] has shown that problem (8.58) and problem (8.59)-(8.61) are equivalent in the sense that (8.58) is infeasible or has an unbounded optimal value if and only if the same is true for (8.59)-(8.61), and if (x^*, y^*) is optimal in (8.58) then y^* is optimal in (8.59)-(8.61) and x^* achieves the infimum in (8.60). Therefore, the optimal xy solution can be found by focusing on the projection onto y. In order to

solve this problem in y space, duality theory is employed. First, the region V can be characterized by its dual representation in terms of the intersection of a collection of regions that contain it. Geoffrion [40] has shown that a point $\bar{y} \in Y$ is also in the set V if and only if $\bar{y}$ satisfies the infinite system of constraints

$$\left[\inf_{x \in X} \lambda^t G(x, y)\right] \leq 0, \forall \lambda \in \Lambda \tag{8.62}$$

where Λ is the set defined by

$$\Lambda = \left\{\lambda : \lambda_i \geq 0, \sum_i \lambda_i = 1\right\}. \tag{8.63}$$

Next, under the assumption that whenever $v(\bar{y})$ is finite, an optimal multiplier vector u exists for the infimum in (8.60), the optimal value equals that of its dual on $Y \cap V$:

$$v(y) = \sup_{u \geq 0} \left[\inf_{x \in X} f(x, y) + u^t G(x, y)\right]. \tag{8.64}$$

By defining y_0 to be the least upper bound implied by the supremum operation, these manipulations can be used to formulate the following problem which is equivalent to (8.59)-(8.61)

$$\min_{y \in Y, y_0} y_0 \tag{8.65}$$

subject to the infinite set of constraints

$$y_0 \geq \inf_{x \in X} \left[f(x, y) + u^t G(x, y)\right], \text{ all } u \geq 0, \tag{8.66}$$

$$\inf_{x \in X} \left[\lambda^t G(x, y)\right] \leq 0, \text{ all } \lambda \in \Lambda. \tag{8.67}$$

Geoffrion [40] has developed an efficient routine for the problem (8.65)-(8.67) based upon relaxation. This is done by first solving (8.65) ignoring all but a few of the constraints (8.66)-(8.67). If the resulting solution satisfies all of the ignored constraints, then an optimal solution has been found. Otherwise, one or more violated constraints should be generated and added to the relaxed problem. This procedure is

repeated until all the constraints are satisfied or until a solution of acceptable accuracy has been found. Geoffrion shows how the solution of the subproblem associated with the infimum on the right-hand side of (8.60) can be used to test for feasibility and to generate an index of violated constraints using the multipliers. If $(\hat{y}, \hat{y}_0)$ is a solution to the relaxed problem, a violated constraint index is a vector $\hat{u} \geq 0$ such that

$$\hat{y}_0 < \inf_{x \in X} \left[f(x, \hat{y}) + \hat{u}^t G(x, \hat{y}) \right] \tag{8.68}$$

if (8.66) is violated or a vector $\hat{\lambda} \in \Lambda$ such that

$$\inf_{x \in X} \left[\hat{\lambda}^t G(x, \hat{y}) \right] > 0 \tag{8.69}$$

if (8.67) is violated. Geoffrion notes that virtually all modern algorithms yield such a $\hat{\lambda}$ when the problem is infeasible, or an optimal multiplier $\hat{u}$ when a feasible solution exists.

The solution by the relaxation method can now be stated. First, for reference define the following functions

$$L^*(y, u) = \inf_{x \in X} \left[f(x, y) + u^t G(x, y) \right] \tag{8.70}$$

$$L_*(y, \lambda) = \inf_{x \in X} \left[\lambda^t G(x, y) \right] \tag{8.71}$$

The Generalized Benders Decomposition consists of the following steps:

1. Let $\bar{y}$ in $Y \cap V$ be a known point. Calculate $v(\bar{y})$ as in (8.60) and obtain an optimal multiplier u^1 as well as the function $L^*(y, u^1)$. Put $p = 1$, $q = 0$, $UBD = v(\bar{y})$. Choose a convergence tolerance parameter $\epsilon > 0$.

2. Solve the current relaxed master problem

$$\min_{y \in Y, y_0} y_0 \text{ subject to } \begin{array}{ll} y_0 \geq L^*(y, u^j) & j = 1, ..., p \\ L_*(y, \lambda^j) \leq 0, & j = 1, ..., q \end{array} \tag{8.72}$$

by any applicable algorithm. Let $(\hat{y}, \hat{y}_0)$ be the optimal solution; $\hat{y}_0$ is a lower

bound on the optimal value in (8.58) since it is a lower bound to the dual problem (8.65)-(8.67), and the dual forms a lower bound on the primal (8.58). If $\hat{y}_0 \geq UBD - \epsilon$, terminate.

3. Solve the revised problem associated with $v(\hat{y})$. One of the following cases will occur:

   (a) The quantity $v(\hat{y})$ is finite. If $\hat{y}_0 \geq v(\hat{y}) - \epsilon$, terminate. Otherwise, determine an optimal multiplier vector $\hat{u}$ and the function $L^*(y, \hat{u})$. Increase p by 1 and put $u^p = \hat{u}$. If $v(\hat{y}) < UBD$, put $UBD = v(\hat{y})$. UBD is an upper bound on the optimal value of (8.58). Return to Step 2.

   (b) The problem associated with evaluating $v(\hat{y})$ is infeasible. Determine $\hat{\lambda} \in \Lambda$ satisfying (8.69) and the function $L_*(y, \hat{\lambda})$. Increase q by 1 and put $\lambda^q = \hat{\lambda}$. Return to Step 2.

Geoffrion [40] has shown that when Y is a finite discrete set, X is a non-empty convex set, G is convex on X for each fixed $y \in Y$, and that $v(y)$ possesses an optimal multiplier vector whenever $v(y)$ is finite, the generalized Benders decomposition terminates in a finite number of steps for any given $\epsilon \geq 0$ including $\epsilon = 0$. Therefore, for mixed integer quadratic programs, convergence is guaranteed in a finite number of steps.

8.5.2 MIQP algorithm

In this section, the implementation of the generalized Benders decomposition to MIQP problems will be set forth. In particular, problems of the form

$$\min \frac{1}{2} x^t Q x + g_x^t x + g_y^t y \tag{8.73}$$

subject to

$$G_x x + G_y y - H \leq 0, \; y \in \{0,1\}^{n_2}, \; x \in \Re^{n_1} \tag{8.74}$$

will be considered. The symbol Q denotes an $n_1 \times n_1$ real positive semi-definite matrix; g_x, g_y, and H denote real vectors of dimension n_1, n_2, and m respectively; G_x and G_y denote real matrices of dimensions $m \times n_1$ and $m \times n_2$ respectively. In this form, the objective function can depend quadratically on the variables x but only linearly on y. This is consistent with the interpretation of y as logic variables which may correspond to heuristics. In addition, the computation of $v(y)$ in (8.60) can be easily accomplished using quadratic programming. By using a dual method, the optimal multiplier vector can easily be obtained.

Now let us consider the function $L^*(y,u)$. By substituting (8.73) for $f(x,y)$ and (8.74) for $G(x,y)$ into the (8.70), the function $L^*(y,u)$ becomes

$$L^*(y,u) = \inf_x \left[\frac{1}{2}x^tQx + g_x^tx + g_y^ty + u^t\left(G_xx + G_yy - H\right)\right] \tag{8.75}$$

$$= \left(u^tG_y + g_y^t\right)y + \inf_x \left[\frac{1}{2}x^tQx + \left(g_x^t + u^tG_x\right)x - u^tH\right]. \tag{8.76}$$

The first term on the right-hand side is linear in y, whereas the second term is a constant whose value is obtained from the evaluation of $v(\hat{y})$. Denote this value by $C(x,u)$. Then the first constraint in (8.72) assumes the form

$$y_0 - \left((u^j)^tG_y + g_y^t\right)y \geq C(x^j,u^j). \tag{8.77}$$

Now consider the function $L^*(y,\lambda)$, which is obtained by

$$L_*(y,\lambda) = \inf_x \left[\lambda^t(G_xx + G_yy - H)\right] \tag{8.78}$$

$$= \lambda^tG_yy - \lambda^tH + \inf_x \lambda^tG_xx. \tag{8.79}$$

The infimum on the right-hand side must be bounded below for if it were not, then the condition (8.69) could not hold. The minimal value of λ^tG_xx can easily be obtained as the solution to a linear program. Denoting this value by $D(\lambda)$, the second equation

in (8.72) can be expressed as

$$(\lambda^j)^t G_y y \leq (\lambda^j)^t H - D(\lambda^j). \tag{8.80}$$

For the special case of the MIQP problem (8.73)-(8.74), the constraints in (8.72) are linear and are given by the equations (8.77) and (8.80). Therefore, the relaxed master problem from Step 2 can be solved using a MILP algorithm. With the integer variables fixed at $\hat{y}$, $v(\hat{y})$ can be evaluated by solving a QP problem. Therefore, in the special case of an MIQP as in (8.73)-(8.74), the generalized Benders decomposition presented in Section 8.5.1 results in an algorithm with iteration between a MILP in Step 2 and a QP in Step 3.

8.6 Conclusions

This chapter has investigated the capabilities of integer variables and linear constraints to represent heuristic process knowledge. Any relationship which can be expressed as propositional logic can be translated into this framework. In particular, we see that many possible applications of this approach exist in the area of control and detection.

In the area of control, by including integer variables representing logic propositions, it is possible to combine logic based control decisions within the MPC framework. This allows innovative control strategies which are capable of prioritizing constraints as well as altering the control objective depending upon the positions of control inputs. By implementing such a strategy, controller performance can be improved. The examples in Section 8.4.1 and 8.4.2 showed that for multivariable systems wherein saturation of one of the manipulated variables prevents all objectives from being met, integer constraints can be used to improve performance and prioritize the objectives.

Integer variables can be used in detection problems to represent the occurrence of symptoms which are indicative of classes of failures. In applications where uncertain

models must be used, false alarms due to uncertainty can be reduced by combining quantitative fault estimation with symptom based fault estimation. When residuals are primarily due to modeling uncertainty, the use of logic variables corresponding to symptoms will prevent erroneous fault alarms.

This chapter has demonstrated how qualitative and quantitative models can be combined using integer variables and linear constraints. Future work in this area should study the properties of such control and detection strategies and determine methods for tuning this type of control algorithm. For example, in the system considered in Section 8.4.1, the controller had the desired properties only for $\alpha \leq \beta \leq \alpha(1+\alpha)$. In more general situations, determining the proper relationship between weights may not be as straightforward and more general methods are needed. By contrast, the only design parameters needed for the prioritization problem of Section 8.4.2 are bounds on the variables of interest.

Part IV

Conclusion

Chapter 9 Summary

9.1 Summary of contributions

In this thesis, a variety of problems in system monitoring have been investigated, including controller performance monitoring, fault detection, and estimation. In this chapter, original contributions to this field which are contained in the present work are summarized.

Controller performance monitoring Previous results for testing whether a controller achieves minimum variance control and for estimating this theoretical limit using routine closed loop data were restricted to stable systems whose only non-invertible behavior consisted of process delays. As many industrial systems exhibit inverse response, the results using the previous theory could be arbitrarily conservative. The minimum variance monitoring philosophy was extended to provide a general theory which could be applied to arbitrary linear systems, including unstable and nonminimum-phase systems. As the new theory indicated that such a performance evaluation requires knowledge of the location of all non-invertible process zeros, an analysis of the effect of errors in these values on the estimated variance limit was provided.

Because minimum variance control is seldom an industrial control objective, a new controller performance monitoring technique was developed. This technique is capable of addressing any performance objective which can be specified in terms of constraints on the impulse response coefficients. Using likelihood ratio testing, an algorithm for verifying that the performance objective is achieved was developed. This algorithm can be implemented by fitting two time series models to the closed loop operating data. In the first model, the closed loop behavior is freely estimated within a class of models, whereas in the second model, the closed loop is constrained

to satisfy the performance objective.

Estimation For the first time, an implementation of constrained Moving Horizon Estimation which guarantees stability of the estimator was presented. Previous implementations of Moving Horizon Estimation can be destabilized by including constraints on estimated variables and innovations. As the primary motivations for implementing such an estimation scheme for linear systems lies in its ability to incorporate constraints, the contribution in this area is fundamental.

Fault detection For systems operating in closed loop, it had been previously recognized that there exists strong interactions linking the design of control and diagnostic components. Although a framework for studying such problems, the four degree of freedom controller, had been previously proposed, no efficient method for designing such integrated control and detection systems existed for processes with uncertain linear models. In this work, it was shown how the previously proposed four parameter control framework could be reconfigured as a special case of a general interconnection structure for which a rich synthesis theory exists. Despite the fact that designing a control and diagnostic system in this framework allows for optimal tradeoffs between control and monitoring objectives, in many cases small model errors will require the control performance to be sacrificed in favor of satisfactory diagnostics. Therefore, the effective use of linear filters for detection purposes is limited.

Two alternative approaches to fault detection were developed. In the first method, a constrained estimation scheme was used to allow for certain types of model uncertainty to be directly incorporated within the detection scheme. The proposed detection method can be implemented using Moving Horizon Estimation in order to keep the size of the problem from growing as more data become available. By using a bank of such estimators with one for each fault type considered, fault diagnosis can be accomplished. A case study involving an important industrial process, the cold tandem steel mill, demonstrated the benefits in detection robustness which can be obtained using this method.

In the second method, a likelihood ratio based detection algorithm was used. According to the Neyman-Pearson lemma, algorithms based on such tests satisfy certain optimality criteria. Unlike the other detection methods presented in this thesis, this method is capable of addressing multiplicative faults as well as additive faults. A new framework for detection of changes in parameters and signals was introduced in which two models are used. In the first model, parameters and signals subject to change are modeled using nominal values, whereas in the second model, Brownian motion is used to describe these terms. Approximate nonlinear global filtering is proposed to estimate the likelihood of the observations under each of the two models. Substantial improvement over detection methods based on local approximations were shown.

Qualitative modeling using propositional logic For many systems, much of the known process information can be expressed in terms of propositional logic. Previous researchers have shown how propositional logic can be translated into linear constraints on zero-one integer variables. In this work, it has been shown how this concept can be applied to problems in monitoring as well as control. In the area of monitoring, symptoms can be included in model based detection schemes through the use of integer variables. Multiobjective and region dependent performance criteria which are difficult to implement using traditional approaches can be easily formulated using constraints on logic variables. Detection and control algorithms can be implemented using this framework by solving mixed integer linear and quadratic programming problems.

Bibliography

[1] B. Anderson and J. B. Moore. *Optimal Filtering.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1979.

[2] U. Appel and A. V. Brandt. Adaptive sequential segmentation of piecewise stationary time series. *Inf. Sci.*, 29:27–56, 1983.

[3] K. J. Åström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20:31–38, 1984.

[4] K. J. Åström and B. Wittenmark. *Computer Controlled Systems Theory and Design.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1984.

[5] G. J. Balas, J. C. Doyle, K. Glover, A. K. Packard, and R. S. R. Smith. *μ-Analysis and Synthesis Toolbox (μ-Tools) : Matlab Functions for the Analysis and Design of Robust Control Systems.* The Mathworks, Inc., Natick, MA, 1991. Computer Software.

[6] M. Basseville. On-line detection of jumps in mean. In M. Basseville and A. Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, number 77 in Lecture Notes in Control and Information Sciences. Springer, Berlin, 1986.

[7] M. Basseville. The two-models approach for the on-line detection of changes in AR processes. In M. Basseville and A. Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, number 77 in Lecture Notes in Control and Information Sciences. Springer, Berlin, 1986.

[8] M. Basseville. Detecting changes in signals and systems – a survey. *Automatica*, 24(3):309–326, 1988.

[9] M. Basseville and I. Nikiforov. *Detection of Abrupt Changes: Theory and Applications.* Prentice-Hall Information and Systems Series. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.

[10] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming, Theory and Algorithms.* John Wiley & Sons, second edition, 1993.

[11] R. V. Beard. *Failure Accommodation in Linear Systems through Self-reorganization.* PhD thesis, MIT, Cambridge, MA, 1971.

[12] A. Benveniste, M. Basseville, and G. V. Moustakides. The asymptotic local approach to change detection and model validation. *IEEE Trans. Aut. Control,* 32(7):583–592, July 1987.

[13] R. R. Bitmead, M. Gevers, and V. Wertz. *Adaptive Optimal Control.* Prentice Hall, Englewood Cliffs, N.J., 1990.

[14] G. E. P. Box and G. M. Jenkins. *Time Series Analysis, forecasting and control.* Holden-Day, San Francisco, second edition, 1976.

[15] R. D. Braatz and M. Morari. Minimizing the Euclidean condition number. *SIAM Journal of Control and Optimization,* 32(6):1763–1768, November 1994.

[16] R. S. Bucy. Bayes' theorem and digital realization for nonlinear filters. *J. Astronaut. Sci.*, 17:80–94, 1969.

[17] T. M. Cavalier, P. M. Pardalos, and A. L. Soyster. Modeling and integer programming techniques applied to propositional calculus. *Computers Opns. Res.*, 17(6):561–570, 1990.

[18] E. Y. Chow and A. S. Willsky. Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. Aut. Control,* 29:603–614, July 1984.

[19] R. N. Clark. State estimation schemes for instrument fault detection. In R. Patton, P. Frank, and R. Clark, editors, *Fault Diagnosis in Dynamic Systems, Theory and Applications.* Prentice Hall, New York, New York, 1989.

[20] R. Da and C. F. Lin. Failure-detection and isolation structure for global positioning system with autonomous integrity monitoring. *J. of Guidance, Control, and Dynamics*, 18(2):291–297, 1995.

[21] R. Da and C. F. Lin. A new failure-detection approach and its application to GPS autonomous integrity monitoring. *IEEE Trans. Aero. Elec. Sys.*, 31(1):499–506, 1995.

[22] M. A. Dahleh and I. J. Diaz-Bobillo. *Control of Uncertain Systems: A Linear Programming approach.* Prentice-Hall, 1995.

[23] J. de Kleer and J. S. Brown. A qualitative physics based on confluences. *Artificial Intelligence*, 24(1-3):7–83, 1984.

[24] L. Desborough and T. Harris. Performance assessment measures for univariate feedback control. *Canadian Journal of Chemical Engineering*, 70:1186–1197, Dec. 1992.

[25] L. Desborough and T. Harris. Performance assessment measures for univariate feedforward/feedback control. *Canadian Journal of Chemical Engineering*, 71:605–616, Aug. 1993.

[26] J. Deshayes and D. Picard. Off-line statistical analysis of change-point models using non parametric and likelihood methods. In M. Basseville and A. Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, number 77 in Lecture Notes in Control and Information Sciences. Springer, Berlin, 1986.

[27] M. Ding and R. W. H. Sargent. *MISQP: A mixed Integer Nonlinear Programming Solver.* Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College, London, 1995.

[28] X. Ding and P. M. Frank. Fault identification filter design via $\mathcal{H}_\infty$-optimization techniques. In *Proc. IFAC Symp. on Id. & Sys. Param. Est.*, pages 219–224, Budapest, 1991.

[29] X. Ding and P. M. Frank. Frequency domain approach and threshold selector for robust model-based fault detection and isolation. In *Fault Detection, Supervision, and Safety for Technical Processes*, pages 271–276, Baden-Baden, Germany, 1991. IFAC.

[30] J. C. Doyle, K. Glover, P. Khargonekar, and B. Francis. State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Trans. Auto. Cont.*, 34(8):831–847, Aug. 1989.

[31] G. A. Dumont. Fifteen years in the life of an adaptive controller. In *IFAC Symposium on Adaptive Systems in Control and Signal Processing*, pages 261–272, Grenoble, France, July 1992. IFAC.

[32] A. Elias-Juarez, A. Ajbar, and J. C. Kantor. Multiobjective l_∞ design with integrated diagnostics. In *Proceedings of the 1991 American Control Conference*, pages 1671–1672, 1991.

[33] A. Emami-Naeini, M. M. Akhter, and S. M. Rock. Effect of model uncertainty on failure detection. *IEEE Trans. Aut. Control*, 33:1106–1115, 1988.

[34] P. M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – a survey and some new results. *Automatica*, 26(3):459–474, 1990.

[35] P. M. Frank and J. Wünnenberg. Robust fault diagnosis using unknown input observer schemes. In R. Patton, P. Frank, and R. Clark, editors, *Fault Diagnosis in Dynamic Systems, Theory and Applications*, New York, New York, 1989. Prentice Hall.

[36] A. W. Fuxjaeger and R. A. Iltis. Adaptive parameter-estimations using parallel Kalman filtering for spread-spectrum code and Doppler tracking. *IEEE Trans. Comm.*, 42(6):2227–2230, 1994.

[37] C. E. García, D. M. Prett, and M. Morari. Model predictive control: Theory and practice — A survey. *Automatica*, 25(3):335–348, May 1989.

[38] K. F. Gauss. *Theoria motus corporum celestium.* English Translation: *Theory of the Motion of the Heavenly Bodies.* Dover (1963), 1809.

[39] A. M. Geoffrion. Elements of large-scale mathematical programming. *Management Science*, 16(11), 1970.

[40] A. M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.

[41] J. J. Gertler and M. M. Kunwer. Optimal residual decoupling for robust fault detection. *Int. J. Control*, 61(2):395–421, 1995.

[42] C. Han, R. Shih, and L. Lee. Quantifying signed directed graphs with the fuzzy set for fault diagnosis resolution improvement. *Ind. Eng. Chem. Res.*, 33(8):1943–1954, 1994.

[43] T. J. Harris. Assessment of control loop performance. *Canadian Journal of Chemical Engineering*, 67:856–861, 1989.

[44] D. T. Horak. Failure detection in dynamic systems with modeling errors. *J. Guidance, Control and Dynamics*, 11(6):508–516, 1988.

[45] M. Iri, K. Aoki, E. O'Shima, and H. Matsuyama. An algorithm for diagnosis of system failures in the chemical process. *Comp. and Chem. Eng.*, 3:489–493, 1979.

[46] A. J. Isaksson. An on-line threshold selector for failure detection. In *TOOLDIAG*, pages 628–634. CERT-ONERA, 1993.

[47] R. Isermann. Process fault detection based on modeling and estimation methods - a survey. *Automatica*, 20(4):387–404, 1984.

[48] R. Isermann. Fault diagnosis of machines via parameter estimation and knowledge processing – tutorial paper. *Automatica*, 29(4):815–835, 1993.

[49] A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 1970.

[50] H. L. Jones. *Failure Detection in Linear Systems.* PhD thesis, MIT, Cambridge, MA, 1973.

[51] T. Kourti, P. Nomikos, and J. F. MacGregor. Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *J. of Process Control*, 5(4):277–284, 1996.

[52] D. J. Kozub and C. García. Monitoring & diagnosis of automated controllers in the chemical processing industries. In *AIChE Annual Meeting*, St. Louis, MO, 1993.

[53] S. C. Kramer and H. W. Sorenson. Bayesian parameter estimation. *IEEE Trans. Aut. Control*, 33(2):217–222, 1988.

[54] S. C. Kramer and H. W. Sorenson. Recursive Bayesian estimation using piecewise constant approximations. *Automatica*, 24(6):789–801, 1988.

[55] N. Kuipel and P. Frank. Fuzzy supervision and application to lean production. *Int. J. Systems SCI*, 24:1935–1944, 1993.

[56] N. Kuipel, S. Koppen-Seliger, H. Kellinghaus, and P. Frank. Fuzzy residual evaluation concept. In *Proc. IEEE/SMC Conference*, pages 13–18, Vancouver, CANADA, 1995.

[57] E. L. Lehmann. *Testing Statistical Hypotheses.* John Wiley & Sons, 1986.

[58] J. S. Lim. Fundamental of digital signal processing. In T. Kailath, editor, *Modern Signal Processing*, pages 1–58. Hemisphere Publishing Co., 1985.

[59] J. N. Little and L. Shure. *Signal Processing Toolbox.* The Mathworks, Inc., Natick, MA, 1992. Computer Software.

[60] L. Ljung. *MATLAB System Identification Toolbox User's Guide.* The Mathworks, Inc., Sherborn, Massachusetts, 1986.

[61] L. Ljung. *System Identification: Theory for the User.* Prentice-Hall Information and System Sciences Series. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.

[62] X. C. Lou, A. S. Willsky, and G. C. Verghese. Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica*, 22(3):333–344, 1986.

[63] D. G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons, Inc., New York, New York, 1968.

[64] E. B. Martin and A. J. Morris. Non-parametric confidence bounds for process performance monitoring charts. *Journal of Process Control*, 6(6):349–358, 1996.

[65] J. H. McClellan and T. W. Parks. A unified approach to the design of optimum FIR linear phase digital filters. *IEEE Transactions on Circuits Theory*, CT-20:697–701, 1973.

[66] T. A. Meadowcroft, G. Stephanopoulos, and C. Brosilow. The modular multivariate controller: Part I: Steady-state properties. *AIChE Journal*, 38(8):1254–1278, 1992.

[67] E. Mendelson. *Introduction to mathematical logic.* Van Nostrand, Princeton, N.J., 1964.

[68] H. Michalska and D. Q. Mayne. Moving horizon observers. In M. Fliess, editor, *IFAC Nonlinear Control Systems Design Symposium (NOLCOS '92)*, pages 576–581, Bordeaux, France, 1992.

[69] H. Michalska and D. Q. Mayne. Moving horizon observers and observer-based control. *IEEE Trans. Aut. Control*, 40(6):995–1006, June 1995.

[70] Y. Misaka. Control equations for cold tandem mills. *Transactions Iron and Steel Inst. of Japan*, 8:86–96, 1968.

[71] S. Mohindra and P. A. Clark. A distributed fault diagnosis method based on digraph models: steady-state analysis. *Comp. and Chem. Eng.*, 17(2):193–209, 1993.

[72] B. P. Molinara. The stabilizing solution of the discrete algebraic riccati equation. *IEEE Trans. Aut. Control*, 20:396–399, 1975.

[73] M. Morari, C. E. Garcia, J. H. Lee, and D. M. Prett. *Model Predictive Control.* Prentice-Hall Inc., Englewood Cliffs, N.J. (in preparation), 1997.

[74] M. Morari and E. Zafiriou. *Robust Process Control.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1989.

[75] J. J. Moré and S. J. Wright. *Optimization Software Guide.* SIAM, 1993.

[76] K. Muske, J. B. Rawlings, and J. H. Lee. Receding horizon recursive state estimation. In *American Control Conf.*, pages 900–904, San Francisco, CA, 1993.

[77] K. R. Muske and J. B. Rawlings. Nonlinear moving horizon state estimation. In R. Berber, editor, *Methods of Model Based Process Control*, pages 349–365. Kluwer Academic Publications, 1995.

[78] C. N. Nett. Algebraic aspects of linear control system stability. *IEEE Trans. Aut. Control*, 31:941–949, 1986.

[79] C. N. Nett, C. A. Jacobson, and A. T. Miller. An integrated approach to controls and diagnostics: The 4-parameter controller. In *American Control Conf.*, pages 824–835, 1988.

[80] I. V. Nikiforov. Sequential detection of changes in stochastic systems. In M. Basseville and A. Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, number 77 in Lecture Notes in Control and Information Sciences. Springer, Berlin, 1986.

[81] P. Nomikos and J. F. MacGregor. Multivariate SPC charts for monitoring batch processes. *Technometrics*, 37(1):41–59, 1995.

[82] D. Y. Oh, H. C. No, and S. H. Kim. Instrument failure-detection and reconstruction methodology in space-time nuclear-reactor dynamic-systems with fixed in-core detectors. *IEEE Trans. Nucl. Sci.*, 38(5):1026–1034, 1991.

[83] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing.* Prentice-Hall, 1989.

[84] O. O. Oyeleye and M. A. Kramer. Qualitative simulation of chemical process systems: steady-state analysis. *AIChE Journal*, 34(9):1441–1454, 1988.

[85] R. J. Patton. Robustness in model-based fault diagnosis: The 1995 situation. In *IFAC workshop on on-line fault detection and supervision in the chemical process industries*, Newcastle-upon-Tyne, England, June 1995. Preprints pp. 55-77.

[86] K. Poola, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal. A time-domain approach to model validation. In *American Control Conf.*, pages 313–317, 1992.

[87] S. Post. Reasoning with incomplete and uncertain knowledge as an integer linear program. In *Expert Systems and Their Applications*, volume 2, pages 1361–1377, Avignon, France, 1987.

[88] D. M. Prett, C. Garcia, and B. L. Ramaker, editors. *Shell Process Control Workshop.* Buttwerworths, Stoneham, MA, 1990.

[89] D. M. Prett and C. E. Garcia. *Fundamental Process Control.* Butterworths, Stoneham, MA, 1988.

[90] R. Raman and I. E. Grossman. Integration of logic and heuristic knowledge in minlp optimization for process synthesis. *Comp. and Chem. Eng.*, 16(3):155–171, 1992.

[91] J. B. Rawlings and K. R. Muske. The stability of constrained receding horizon control. *IEEE Trans. Aut. Control*, 38(10):1512–1516, October 1993.

[92] D. G. Robertson, J. H. Lee, and J. B. Rawlings. A moving horizon-based approach to least squares estimation. *AIChE Journal*, 42(8):2209–2224, 1996.

[93] G. Smith. *Statistical process control and quality improvement.* Prentice Hall, Englewood Cliffs, N.J., 1995.

[94] T. Söderström and P. Stoica. *System Identification.* Prentice Hall International Series in Systems and Control Engineering. Prentice Hall, New York, 1989.

[95] H. W. Sorenson. On the development of practical nonlinear filters. *Information Sciences*, 7:253–270, 1974.

[96] H. W. Sorenson. Recursive estimation for nonlinear dynamic systems. In J. C. Spall, editor, *Bayesian Analysis of Time Series and Dynamic Models*, pages 127–165. Marcel Dekker, 1988.

[97] N. Stanfelj, T. E. Marlin, and J. F. MacGregor. Monitoring and diagnosing process control performance: The single-loop case. *Ind. Eng. Chem. Res.*, 32:301–314, 1993.

[98] P. Stoica, P. Eykhoff, P. Janssen, and T. Söderström. Model-structure selection by cross validation. *Int. J. Control*, 4(6):1841–1878, 1986.

[99] M. Stone. Cross-validatory choice and assessment of statistical predictions. *J. Roy. Statist. Soc. Ser. B*, 36:111–147, 1974.

[100] P. Terwiesch and M. Agarwal. A discretized nonlinear state estimator for batch processes. *Comp. and Chem. Eng.*, 19(2):155–169, 1995.

[101] M. Tyler, K. Asano, and M. Morari. Application of moving horizon estimation based fault detection to cold tandem steel mill. Technical Report AUT96-06, Automatic Control Laboratory, Swiss Federal Institute of Technology, 1996.

[102] M. Tyler and M. Morari. Stability of constrained moving horizon estimation schemes. *Automatica*, 1996. (submitted).

[103] M. L. Tyler and M. Morari. Performance monitoring of control systems using likelihood methods. *Automatica*, 1996. in press.

[104] M. Ulieru. A fuzzy logic based computer assisted fault diagnosis system. In *TOOLDIAG'93*, pages 689–698, Toulouse, France, 1993.

[105] M. Vidyasagar. *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, Massachusetts, 1985.

[106] X. Z. Wang, S. A. Yang, E. Veloso, M. Lu, and C. McGreavy. Qualitative process modeling – a fuzzy signed directed graph method. *Comp. and Chem. Eng.*, 19(Suppl):S735–S740, 1995.

[107] B. Weinstein. A sequential approach to the evaluation and optimization of control system performance. In *American Control Conf.*, pages 2354–2358, Chicago, June 1992.

[108] J. L. Weiss. Threshold computations for detection of failures in SISO systems with transfer function errors. In *American Control Conf.*, pages 2213–2218, 1988.

[109] H. P. Williams. Linear and integer programming applied to the propositional calculus. *Syst. Res. Inf. Sci.*, 2:81–100, 1987.

[110] A. S. Willsky. A survey of design methods for failure management in dynamical systems. *Automatica*, 12:601–611, 1976.

[111] A. S. Willsky. Detection of abrupt changes in dynamic systems. In M. Basseville and A. Benveniste, editors, *Detection of Abrupt Changes in Signals and Dynamical Systems*, chapter 2, pages 27–49. Springer-Verlag, 1986.

[112] B. M. Wise and N. B. Gallagher. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6(6):329–348, 1996.

[113] Q. Zhang, M. Basseville, and A. Benveniste. Early warning of slight changes in systems. *Automatica*, 30(1):95–113, 1994.

[114] Z. Q. Zheng and M. Morari. Stability of model predictive control with mixed constraints. *IEEE Trans. Aut. Control*, 40(10):1818–1823, October 1995.

[115] G. Zimmer. State observation by on-line minimization. *Int. J. Control*, 60(4):595–606, 1994.