Measurement System Design for Chemical Processes

by

Michael Joseph Liba

A thesis submitted to the Department of Chemical Engineering
in conformity with the requirements for the degree of Master of Applied Science

Queen's University
Kingston, Ontario, Canada
June 2011

Abstract

The problem of measurement system design for stochastic linear systems, a popular modeling strategy for chemical processes, is addressed. A multi-objective optimization approach is used. The design metrics are the capital cost of measurement equipment and a weighted uncertainty in terms of the Kalman filter state estimation error covariance matrix. The Pareto optimal set of measurement systems is identified by any of the following applicable techniques: an exhaustive combinatorial search, sequential sensor addition/removal, and branch-and-bound with semi-definite programs (BnB/SDP) solved at each node. The decision-making process involves the use of simulation experiments as a means to map Pareto optimal measurement systems to a dollar cost of operation. The closed-loop performance of Pareto optimal measurement systems are then simulated under a joint Kalman filter and robust profit-maximizing model predictive control strategy.

The design methodology is applied to two example problems. The first involves a low-dimensionality fluid handling network where a number of stream flow rates, a tank level, and a leak stream describe the dynamics. The design variables are the process outputs and the precision with which they are measured. It is observed that the sequential and BnB/SDP techniques are able to approximate the true Pareto optimal set very well, with improved performance in the latter case attainable through trial and error. The second example problem involves a high-dimensionality thermal network model of a one-floor office building. The impact of zero and nonzero state noise covariance structures on the results of the proposed design procedure is investigated. It is shown that measurement importance is placed on the disturbance variables in the deterministic case, whereas importance is placed on the controlled variables when model uncertainty is assumed. Closed-loop simulations incorporating MPC and Kalman filtering are then performed to generate expected operational cost data. The measurement system that minimizes the overall cost of capital investment and operation over the expected lifespan of the measurement is chosen as the final design. It is shown that the combination of measured variables which minimizes the overall cost is those of the three largest bodies of air that are to be controlled.

Acknowledgements

I would like to first thank my supervisors, Professor Tom Harris and Professor Martin Guay, whose encyclopedic knowledge of a variety of topics helped craft this thesis. In particular, thank you Tom for finding the time every week, and thank you Martin for providing a comfortable and fruitful working environment. My deepest thanks also goes to Sorin, Keunmo, Veronica, and Satish at the United Technologies Research Center for their efforts to bring me into a friendly and high-performance workplace, and to Professor Allan Gilbert at Lakehead University, who helped spark my interest in process control.

Thanks are also in order for my officemates in G37. To Saman, whom I've known since day one, and with whom I have shared the ups and downs of the learning process. To Nick, whose character, sense of humour, and friendship is missed. To Scott, for the many fruitful conversations whose topics span an enormous spectrum, and for taking me into your home over this last week. Thanks also to Kai, Y.J., Devon, and Derek for making the office such a comfortable place to work and learn. I would be lucky to have such courteous coworkers as you in the future.

I would also like to express my gratitude to a number of people whom have only recently entered my life. First and foremost, to my dear friend Nancy: it was a pleasure to share my experience at Queen's with you. I wish you all the best in the future. I wish to ackhowledge Jonas and Allison Gerson for their kind hospitality and friendship. To Nadim, Jay, and Jeremy in Connecticut: I couldn't have asked for a better group to share my experience abroad with (especially you, Jay, who may very well have saved my life!). Thanks also to Tanya, Jordan, Jeff, Dan, my courteous roommate Joe, and the many others too numerous to list here that made Kingston an enjoyable place to work and live.

Finally, I thank my family and friends in my hometown for their undying support. To my parents, who have worked so hard to provide me with a comfortable life and the necessary tools for success: I hope I have made you proud. To my friends, thank you for giving me something to look forward to during the tough times while away from home. `OATH!`

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation for Measurement System Design

In any control system, the availability and quality of measurements will inevitably affect closed-loop performance. For many systems, however, it is unreasonable to measure all candidate variables due to the unavailability of sensors or ceilings on funds available for capital investment. It is natural then to ask: *"which variables are the most important to measure?"*

The answer to this question has proven to be elusive, even when attention is restricted to linear systems, whose properties have been investigated extensively (see, for example [Brockett, 1970; Kwakernaak and Sivan, 1972; Dullerud and Paganini, 2000]). Among the many possible reasons for this elusiveness is the general inability of researchers to pose an optimality criterion on measurement systems which completely captures the quality of their design. This is not surprising, since deciding on which variables to measure without knowledge of exactly how information will be manipulated and used by the control system clearly introduces subjectivity into the design procedure. This non-consideration of subsequent information usage prompted researchers to initially pose measurement system design as a maximization

of certain scalar values related to a linear system's observability gramian [Müller and Weber, 1972]. The idea with these analyses being that "more observable" systems should, in some general sense, perform better in closed-loop than those which are "less observable". In the field of process control, the influence of measurement system design on process economics is of particular interest due to the bottom-line oriented nature of profit-maximizing firms and their various technical support subdivisions. As such, it is desirable to consider the selection of measured variables from a process economics point of view. To do so, logical connections between: (1) the selection of measured variables and their measurement devices; (2) data processing (i.e. state estimation/filtering); and (3) usage by the control system are required to fully analyze the performance of measurement system design on closed-loop performance. Life cycle analysis via simulations can then be performed in order to justify the allocation of capital towards the purchase and installation of measurement equipment.

1.2 Organization of the Thesis

CHAPTER 2: Chapter 2 is divided into two parts. First, the technical preliminaries required to develop the measurement system design methodology proposed in Chapter 3 are introduced. The topics include observability/detectability, Kalman filtering, and model predictive control. The second section contains a review of the early and more recent works in the field of measurement system design for deterministic and stochastic linear systems.

CHAPTER 3: In this chapter, measurement system design is posed as a multi-objective optimization problem where the quality of all candidate measurement systems are quantified based on their Kalman filtering state estimation uncertainty and purchased and installed cost of measurement equipment. A number of potential approximate solution strategies are developed for a range of problems whose state space dimensionalities and number of available sensors types are variable. A standardized decision-making process based on model

predictive control is then used to decide on a final measurement system design.

CHAPTER 4: A low state space dimensionality fluid handling network is considered in Chapter 4 as a means to showcase the effectiveness of the multi-objective optimization solution techniques presented in Chapter 3. All three techniques are applied to investigate their relative performances. Controller design and simulation experiments are not considered, since they are the focus of the next chapter.

CHAPTER 5: The second example problem, which is characterized by a large state space dimensionality, is discussed in detail. A full measurement system design is performed by first approximating the solution to the multi-objective optimization problem and then by employing simulation experiments as a means to map Pareto optimal designs to a dollar cost of operation. Chapter 5 concludes with a measurement equipment lifecycle analysis that selects a single measurement system design from the 10^{23} theoretically possible sensor configurations.

CHAPTER 6: A summary of the design procedure developed in Chapter 3 is provided followed by a discussion concerning its strengths and weaknesses. The design procedure's contingency on the applicability of the proposed control, state estimation, and optimization techniques is addressed, and suggestions for future work is provided.

Chapter 2

Literature Review

The measurement system design methodology developed in this work is largely synthetic, relying on concepts from the fields of linear systems, state estimation, model predictive control, and optimization. In the first part of this chapter, these concepts are introduced, and their relevance is identified. Detailed discussion regarding the interrelationships between components is left to the next chapter. The second part of this chapter summarizes the early and more recent works by researchers active in measurement system design and identifies where the approach considered here fits in.

2.1 Technical Preliminaries

2.1.1 Observability & Detectability

Consider a discrete-time linear system of the form

$$x_{k+1} = Ax_k + Bu_k, \tag{2.1a}$$

$$y_k = Cx_k, \tag{2.1b}$$

where $x \in \mathbb{R}^{n_x}$ is the state vector, $u \in \mathbb{R}^{n_u}$ is the control input, $y \in \mathbb{R}^{n_y}$ are the system outputs, and matrices A, B, and C are of appropriate dimension. Observability and detectability are properties of such a dynamical system, first introduced by Kalman [1960]. These properties are meant to express the availability of measurement data with respect to one's ability to reconstruct or make inferences regarding the values of unmeasured state variables.

Observability

Definition 2.1.1. *[Simon, 2006]: A linear discrete-time system given by Eq. (2.1a) and (2.1b) is "observable" (or (C, A) is an "observable pair") if for any initial state x_0 and some final time t the initial state x_0 can be uniquely determined by knowledge of the inputs u_k and outputs y_k for all $k \in [0, t]$.*

If a system is observable, then its initial state can be determined. If the initial state is known, then values of the state at any time $k \in [0, t]$ can be calculated. Hence, observability implies that values of the state at all times $k \in [0, t]$ are fully reconstructible as long as the inputs and outputs are known exactly.

Observability can be checked by a matrix rank test performed on either the system's *observability gramian* or *observability matrix*:

Theorem 2.1.1. *The discrete LTI system (2.1) is observable if and only if the observability gramian defined by*

$$\mathcal{M}_O \triangleq \sum_{k=0}^{t} \left(A^T\right)^k C^T C A^k$$

has rank n_x for some $t \in (0, \infty)$.

Theorem 2.1.2. *The discrete LTI system (2.1) is observable if and only if the observability matrix defined by*

$$\mathcal{O}(C, A) \triangleq \left[C^T, (CA)^T, \cdots, (CA^{n_x-1})^T\right]^T$$

is of rank n_x.

A LTI system can also be said to be observable if it contains no *unobservable modes*. A system's modes can be identified by applying the similarity transformation $z = Vx$ to Eq. (2.1)

$$z_{k+1} = V^{-1}AVz_k + V^{-1}Bu_k,$$
$$y_k = CVz_k,$$

that results in the matrix $(V^{-1}AV)$ having Jordan normal form. The i^{th} mode is then defined as the state (or group of states) that correspond to the i^{th} Jordan block of $(V^{-1}AV)$. The i^{th} mode of a transformed system is said to be unobservable if the i^{th} column of V is orthogonal to every row of C. If this is the case, then the i^{th} column of the matrix CV will contain only zeros. Consequently, the i^{th} column of the observability matrix defined by $\mathcal{O}(CV, V^{-1}AV)$ will also only contain zeros, which implies that rank $\left[\mathcal{O}(CV, V^{-1}AV)\right] < n_x$. Since observability is unaffected by similarity transformations, the presence of an unobservable mode is indicative of system unobservability.

Detectability

Definition 2.1.2. *[Dullerud and Paganini, 2000]: A linear system is "detectable" if for a non-observable system, after a state transformation T such that*

$$TAT^{-1} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}, \quad CT^{-1} = \begin{bmatrix} \tilde{C}_1, 0 \end{bmatrix}$$

where $\left(\tilde{C}_1, \tilde{A}_{11}\right)$ is an observable pair; the eigenvalues of $\tilde{A}_{22}$ all have magnitude < 1.

Note that due to the structure of CT^{-1} the states corresponding to $\tilde{A}_{22}$ are all completely unobservable. However, since the eigenvalues of the matrix $\tilde{A}_{22}$ all have magnitude < 1,

then under no control input, convergence of these states to zero is guaranteed.

Similarly, a LTI system can also be said to be detectable if and only if all of its unobservable modes are stable (i.e. $\|\lambda_i\| < 1$ where λ_i is the i^{th} eigenvalue of A that corresponds to an unobservable mode).

The concepts of observability and detectability are central to the design of *state observers* and *state estimators*, which are discussed in the next section.

2.1.2 State Observers and Kalman Filtering

In control theory, a state observer is a dynamical system whose outputs are the estimates of the state variables of the system [Dullerud and Paganini, 2000]. The main criterion that observers must satisfy is that the estimation error $e_k = (x_k - \hat{x}_k)$ tends to zero as $k \to \infty$ where $\hat{x}_k$ is the estimate of the state x_k at sample time k. If the dynamics of the plant give rise to a linear time-invariant system, then there exists an estimator of the form

$$\hat{x}_{k+1} = A\hat{x}_k + L(y_k - \hat{y}_k) + Bu_k, \tag{2.2a}$$

$$\hat{y}_k = C\hat{x}_k + Du_k, \tag{2.2b}$$

which guarantees convergence of the state estimation error to zero, provided that the plant is detectable. The observer given by Eqs. (2.2a) and (2.2b) is referred to as a *Luenberger observer* [Dullerud and Paganini, 2000]. The matrix L is designed so that the eigenvalues of $(A - LC)$ all have magnitude < 1, which ensures the stability of the observer's error dynamics.

The Kalman Filter for Stochastic Linear Systems

In this work we consider a Kalman filter (KF) that is designed to act as a state observer for linear systems whose state evolution and outputs are corrupted by noise:

$$x_{k+1} = A_k x_k + B_k u_k + w_{k+1},$$
$$y_k = C_k x_k + v_k,$$

where $w_k \sim \mathcal{N}(0, Q_k)$ and $v_k \sim \mathcal{N}(0, R_k)$ are Gaussian white noise sequences.

Assuming that the linear system is also time-invariant, the KF estimate of the state given all information up to and including time k (referred to as the *a posteriori* estimate), $\hat{x}_{k|k}$ is given by

$$\hat{x}_{k|k} = (I - K_k C)(A\hat{x}_{k-1|k-1} + Bu_{k-1}) + K_k y_k, \tag{2.3}$$

where K_k is referred to as the "Kalman gain matrix". Note that the most recent estimate $\hat{x}_{k|k}$ in Eq. (2.3) depends on the previous *a posteriori* estimate $\hat{x}_{k-1|k-1}$. As such, the Kalman filter is a *recursive* state estimation strategy where new information in the form of measurements is combined with knowledge of the system's dynamics and statistical parameters to produce an optimal estimate. Optimality is defined based on the KF's minimization of the objective measure

$$J_{KF} = \mathrm{tr}\left(\mathbf{E}\{(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T\}\right), \tag{2.4}$$

where $\mathrm{tr}(\cdot)$ and $\mathbf{E}\{\cdot\}$ denote the matrix trace and expected value operators, respectively. That is, the Kalman filter is designed to minimize the total variance of the state estimates at all sampling instants.

Assuming a full characterization of the process' dynamics and statistical parameters are

available, the recursive KF equations are:

$$\hat{x}_{k|k} = (I - K_k C)(A\hat{x}_{k-1|k-1} + Bu_{k-1}) + K_k y_k, \tag{2.5a}$$

$$K_k = (P_k^- C^T)(CP_k^- C^T + R)^{-1}, \tag{2.5b}$$

$$P_k^- = AP_{k-1}^- A^T - A(P_{k-1}^- C^T)(CP_{k-1}^- C^T + R)^{-1}(CP_{k-1}^-)A^T + Q, \tag{2.5c}$$

$$P_k^+ = (I - K_k C)P_k^-(I - K_k C)^T + K_k R K_k^T, \tag{2.5d}$$

where P^- and P^+ are the *a priori* and *a posteriori* estimation error covariance matrices:

$$P_k^- \triangleq \mathbf{E}\left\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\right\}, \quad P_k^+ \triangleq \mathbf{E}\left\{(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T\right\},$$

where $\hat{x}_{k|k-1}$ is the estimate of x_k given all information up to and including sampling instant $k-1$. To ensure acceptable performance of the Kalman filter, it may be necessary to occasionally estimate the entries of the process and output noise covariance matrices Q and R, respectively. The recent techniques which attempt to achieve this rely on the use of a set of historic process data in conjunction with the minimization of a least-squares objective function based on predicted and observed process and output noise autocorrelation [Odelson et al., 2006]. Other recent works include the addition of positive definiteness constraints on the Q and R matrices in the least-squares optimization problem formulation, leading to a semi-definite program [Rajamani and Rawlings, 2009]. Most recently, a covariance estimation technique for systems described by nonlinear dynamics was proposed by Lima and Rawlings [2011] as one of the components of a full state estimator design procedure.

Some state estimation applications may not require the recursive calculations of the Kalman filter gain and covariance matrices. Instead, it may only be necessary to consider their steady-state values during on-line operation, which would reduce the computational burden. Such a state estimation strategy is referred to as "steady-state filtering", and is possible contingent on the existence of the converged solution to the discrete algebraic riccati

equation (DARE) Eq. (2.5c). The following are the necessary and sufficient conditions that ensure convergence of the Kalman filter covariance matrices and gain matrix to steady-state values.

Theorem 2.1.3. *[Simon, 2006]: The DARE has at least one positive definite solution P_∞^- if and only if both of the following conditions hold.*

1. *The pair (A, C) is detectable.*
2. *The pair (A, B) is controllable.**

Furthermore, exactly one of the positive definite DARE solutions results in a stable steady-state Kalman filter.

When the controllability condition is relaxed, the following result is obtained.

Theorem 2.1.4. *[Simon, 2006]: The discrete algebraic riccati equation (DARE) (3.8c) has a unique positive semidefinite solution P_∞^- if and only if both of the following conditions hold.*

1. *The pair (A, C) is detectable.*
2. *The pair (A, B) is stabilizable.*

Furthermore, the corresponding steady-state Kalman filter is stable. That is, the eigenvalues of $[(I - K_\infty C) A]$ have magnitude < 1.

These results clearly specify the conditions on which the design of measurement systems rely. These theorems will be used to justify the subsequent measurement system design procedure in the next chapter.

*The duals of observability and detectability: controllability and stabilizability, are omitted from this chapter. See any linear systems text for a detailed discussion.

2.1.3 Model Predictive Control

Model predictive control (MPC), known as "receding horizon control" (RHC), is a control strategy that explicitly utilizes a process model to generate a sequence of control inputs which minimizes some user-specified objective function [Camacho and Bordons, 2007].

One of the main advantages of MPC over other "optimal control" strategies such as the linear-quadratic regulator (LQR) is its ability to enforce constraints on both the state variables and control inputs. This differs from so-called *unconstrained* optimal control strategies. The ability of MPC to enforce state and input constraints is crucial for real application since it is not difficult to imagine cases where states of physical importance (temperatures or pressures, for example) must be maintained within prescribed limits for safety reasons. Constraints on inputs typically arise from the physical limitations of actuators (for example, control valves cannot be opened more than 100%).

In general, the optimization problem solved at each controller execution is of the form

$$\begin{aligned} &\min_{u_{N_{pr}}} J_{MPC} = \sum_{i=0}^{N_{pr,1}} J_i^c(x_{k+i}, u_{k+i}) + J^f(x_{k+N_{pr,1}}), \\ &\text{subject to:} \begin{cases} x_{k+i} \in \mathbb{X}_i & \forall\ i = \{1, \cdots, N_{pr,1}\}, \\ u_{k+i} \in \mathbb{U}_i & \forall\ i = \{0, \cdots, N_{pr,2}\}, \\ x_{k+1} = Ax_k + Bu_k + w_{k+1}, \end{cases} \end{aligned} \tag{2.6}$$

where $u_{N_{pr}} = [u_k^T, u_{k+1}^T, \cdots, u_{k+N_{pr,2}}^T]^T$ is a vector that contains the control inputs for the current sampling instant and the next $N_{pr,2}$ number of samples into the future. The cost function J_{MPC}, is composed of stage costs J_i^c and a terminal cost J^f. The sets, $\mathbb{X}$ and $\mathbb{U}$, are the admissible sets for the state and control input, respectively. The scalars $N_{pr,1}$ and $N_{pr,2}$ are referred to as the *prediction* and *control input* horizons, respectively, where (in general) $N_{pr,1} \geq N_{pr,2}$. The final defining feature of MPC is that only the first calculated control input u_k is actually used. The entire optimization problem is re-solved at the next

controller execution using any new information that has become available. A schematic of MPC's operation is given in Figure 2.1, where a particular control sequence $u_{N_{pr}}$ is considered, but will be abandoned due to a predicted violation of the state constraints.

One of the earliest MPC formulations for single-input/single-output (SISO) systems, known as Dynamic Matrix Control (DMC), considers a step response model in conjunction with an objective function representing the sum of future squared output errors [Bequette, 2003]. Under this unconstrained MPC formulation, an analytical solution for the minimizing control input sequence exists, and computational burden is placed mostly on matrix inversion. It is possible, however, that the combination of objective function and constraints may result in a more complex optimization problem such as a linear/quadratic program (LP/QP), semi-definite program (SDP), etc. Under these scenarios, computational power may very well be the limiting factor for the applicability of MPC's on-line use. In distributed control systems where older processing units are still being used, even LPs and QPs may lead to computational intractability when a long control input horizon is necessary to achieve acceptable performance [Morari and Lee, 1999].

Model predictive control offers the control system engineer a flexible and powerful tool by which to achieve a variety of control objectives. Surprisingly, though, theoretical results regarding stability, the primary control objective for any control system, were only obtained after years of practical implementation of MPC.

Stability analysis is performed by treating the MPC objective function J_{MPC} as a Lyapunov function. Conditions under which the cost functional of the MPC controller becomes a Lyapunov function for the closed-loop have been derived by many authors. Of these, the so-called *dual mode configuration* is of particular interest (it is applied in the example problem of Chapter 5). Under this formulation, closed-loop stability can be theoretically guaranteed by imposing the constraint that x must be contained within a *terminal set* $\mathcal{X}_f$ at the end of the prediction horizon. The control inputs following $u_{k+N_{pr,2}}$ are assumed to be produced by a separate, *terminal* control law $u^f(x)$ which is designed to ensure that

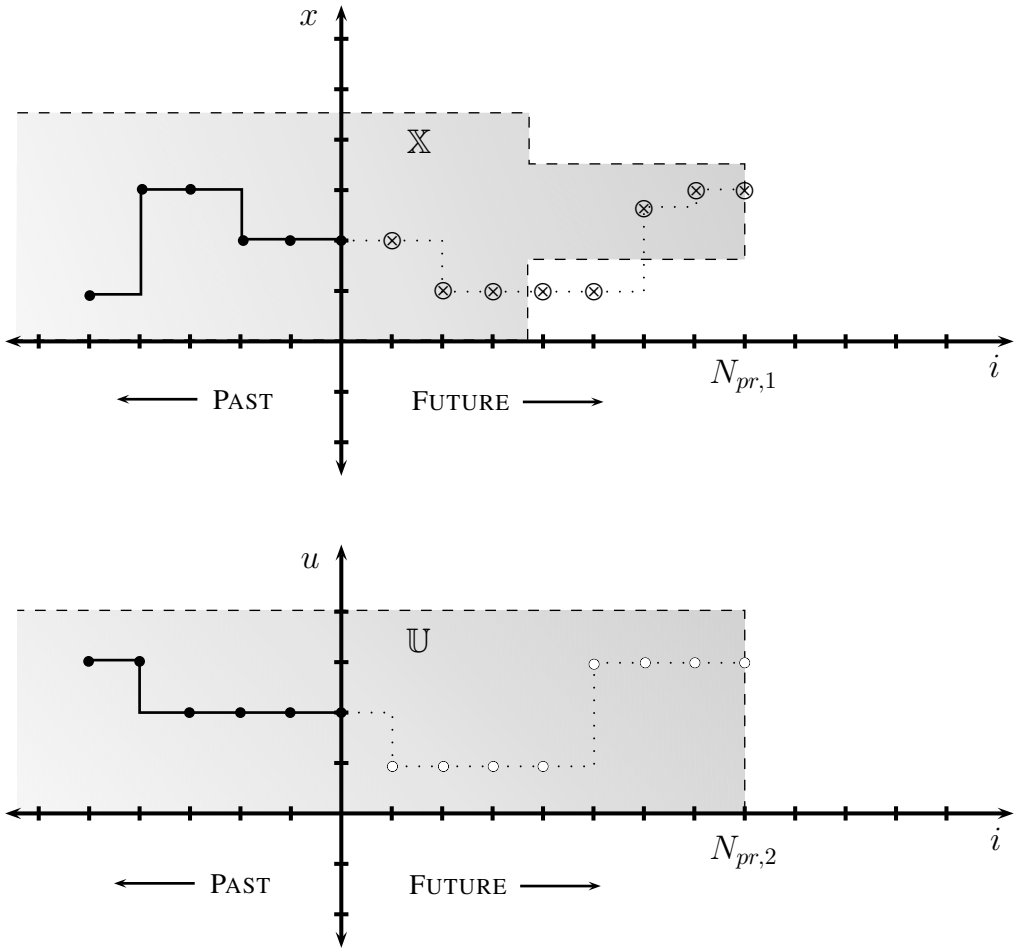

Figure 2.1: A schematic of a single-state system under MPC with $N_{pr,1} = N_{pr,2}$. The past and current values of x and u ($\bullet$) are known, and the future state ($\otimes$) under a candidate control sequence ($\circ$) is prediced. At the current execution, a potential control sequence $u_{N_{pr}}$ is considered, but is predicted to result in 2 state constraint violations.

$x_{N_{pr,1}+i} \in \mathcal{X}_f \ \forall i = \{1, \cdots, \infty\}$ (i.e. $\mathcal{X}_f$ is *positively invariant* under u^f).

When stability analysis is performed by regarding J_{MPC} as a Lyapunov function, each technique specifies $\mathcal{X}_f$, J^c, J^f, and u^f to satisfy the (sufficient) conditions for closed-loop stability which are conveniently summarized in the survey paper by Mayne et al. [2000]:

C1: $\mathcal{X}_f \subset \mathbb{X}$, $\mathcal{X}_f$ closed, $0 \in \mathcal{X}_f$ (state constraints are satisfied in the terminal region $\mathcal{X}_f$),

C2: $u^f(x) \in \mathbb{U} \quad \forall x \in \mathcal{X}_f$ (terminal control inputs are admissible),

C3: $\left[Ax + Bu^f(x)\right] \in \mathcal{X}_f \quad \forall x \in \mathcal{X}_f$ ($\mathcal{X}_f$ is positively invariant under $u^f(\cdot)$),

C4: $(J_f(x^+) - J_f(x)) + J_s(x, u^f(x)) \leq 0 \quad \forall x \in \mathcal{X}_f$ where $x^+ = Ax + Bu^f(x)$ ($J_f(\cdot)$ is a local Lyapunov function),

along with the implicit condition that the optimization problem Eq. (2.6) be feasible. In cases where the terminal set $\mathcal{X}_f$ is small, it may be necessary to increase the prediction and control horizons $N_{pr,1}$ and $N_{pr,2}$ to ensure feasibility. However, such a reformulation may lead to on-line computational intractability. Though this is a critical concern, a general discussion regarding practical implementation issues for MPC is lengthy. The reader is directed to [Camacho and Bordons, 2007, Ch. 7] and references therein where a number of techniques and modifications to ensure feasibility in a variety of MPC contexts is discussed.

2.2 Measurement System Design

When a process model is known to be reliable, it may be possible to make inferences regarding the values of several variables without directly measuring them by using state observers as discussed in section 2.1.2. When these so-called observer-based strategies are employed, the choice of measured variables becomes a fundamental design consideration. If attention is restricted to either stochastic or deterministic linear systems, a logical starting point is to consider the connection between the choice of measured variables and their effect on various scalar measures linked to observability, detectability and state observer design.

Early attempts at quantifying the observability of linear systems was performed by Müller and Weber [1972], who considered the maximum eigenvalue, trace, and determinant of the system's inverse observability gramian. Since a system is either observable or not, these measures could be interpreted as the "degree" of observability. Given observability's physically motivated definition, it is natural to expect that "more observable" systems should in some broad sense outperform those whose observability measures are less. Similar measures were considered in the control of a tubular reactor by Waldraff et al. [1998], and for nonlinear systems by Singh and Hahn [2004] who considered *empirical gramians* [Lall et al., 2002].

Uncertainty metrics have also been used to design measurement systems. In particular, the more recent works in this field rely on some scalar measure of the state estimation error covariance matrix to guide the placement of sensors. In [Harris et al., 1980], the authors consider the trace and determinant of the state prediction error covariance matrix for LQG control of a tubular reactor. In [Musulin et al., 2005], optimal sensor placement in the context of system parameter identification via Kalman filtering is discussed.

The problem of *economically optimal* measurement system design was considered by Muske and Georgakis [2003], who explicitly took the capital cost of sensors into consideration. A multi-objective optimization problem [Steuer, 1986; Branke et al., 2008] was formulated where measurement systems were categorized solely by their cost of measurement equipment and uncertainty in terms of a weighted state estimation error covariance matrix. Muske and Georgakis [2003] then constructed the Pareto optimal set of measured variables by considering every possible combination of sensor arrangements that resulted in system observability, but omitted the *decision-making* process from their analysis. Though a fully exhaustive search may be useful for illustration purposes, such a computationally expensive method of locating the Pareto optimal measurement system configurations is realistic only when the number of potential measurement system configurations is low.

One possible alternative to this exhaustive search technique is to apply the branch-and-bound (BnB) algorithm. This is possible through the work by Chmielewski et al. [2002], who have expressed measurement system design as a binary integer program (BIP) subject to linear matrix inequality (LMI) constraints. A BIP of this type can then be solved via BnB, with semi-definite programs (SDPs) being solved at each node. The key feature of this design technique is that the construction of the Pareto optimal set for high-dimensional systems may in fact be tractable since the BnB algorithm either explicitly or implicitly examines all possible measurement system configurations. As such, it is possible that a significant number of these configurations can be ruled out of consideration by the bounding of the objective function that occurs at each node [Edgar et al., 2001].

High state-space dimensionality is a potential consequence of the discretization of heat transfer models, such as in the building system temperature control example problem of Chapter 5. Discretized heat transfer models, known as "thermal networks" have proven effective in building system control research due to their relative simplicity and ability to accurately describe temperature dynamics [Wang and Xu, 2006; Zhang and Hanby, 2006; Xu and Wang, 2008; Lee and Braun, 2008; O'Neill et al., 2010]. Under this modeling strategy, entities of physical significance, called "nodes" (walls, masses of air, etc.) are assumed to have uniform temperature, and heat transfer between nodes is modeled as due to relevant heat transfer phenomena [Çengel, 2006].

For thermal network models which consider several individual rooms rather than macro-scale zones (entire floors, whole buildings), the total number of walls, floor surfaces, and air masses available for temperature measurement can become very large. If the temperature control of individual rooms (for occupant comfort reasons) is the primary concern, measuring the temperature at each node is impractical and intuitively unnecessary since, for example, heat transferred from small surfaces will have little impact on the temperature dynamics of large masses of air. Hence, the location of the most important measurements for the optimal control of such systems is a challenging problem whose results may find use in practical

application.

A brief history and some of the current active research related to measurement system design has been discussed. The approach to measurement system design in this work can be thought of as a combined application of the previous work by Muske and Georgakis [2003] and Chmielewski et al. [2002]. Furthermore, a decision-making process involving model predictive control as a means to map candidate measurement systems to a cost of operation is introduced to complete the multi-objective optimization procedure initiated by Muske and Georgakis [2003].

Chapter 3

Measurement System Design Methodology

In this chapter, the various topics introduced in Chapter 2 are assembled under a multi-objective optimization framework to produce the measurement system design procedure. A number of strategies by which to approximate or exactly solve the multi-objective optimization problem are presented to ensure that the procedure is applicable to wide range of systems.

3.1 System Representation & Basic Components

3.1.1 System Representation

We consider a stochastic linear discrete-time systems of the form

$$x_{k+1} = Ax_k + Bu_k + Fd_k + a_{k+1}, \tag{3.1}$$

where $x \in \mathbb{R}^{n_x}$ is the state vector of the plant, $u \in \mathbb{R}^{n_u}$ is the control input, $d \in \mathbb{R}^{n_d}$ is the vector of disturbances, and $a \sim \mathcal{N}(0, Q_a)$ is a vector of zero-mean normally distributed

white noise sequences with covariance matrix Q_a. Matrices A, B, and F are of appropriate dimension, and assumed to be time-invariant. The disturbance variables are modeled as autoregressive time series

$$d_{k+1} = \Phi d_k + v_{k+1}, \tag{3.2}$$

where Φ is the disturbance state transition matrix, and $v \sim \mathcal{N}(0, Q_v)$ is the vector of driving white noise sequences. Combining Eqs. (3.1) and (3.2), we obtain

$$\begin{bmatrix} x_{k+1} \\ d_{k+1} \end{bmatrix} = \begin{bmatrix} A & F \\ 0 & \Phi \end{bmatrix} \begin{bmatrix} x_k \\ d_k \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} I_x & 0 \\ 0 & I_d \end{bmatrix} \begin{bmatrix} a_{k+1} \\ v_{k+1} \end{bmatrix}, \tag{3.3}$$

where I_x and I_d are $(n_x \times n_x)$ and $(n_d \times n_d)$ identity matrices, and 0 are zero-matrices of appropriate dimensions. The output variables used for feedback control are

$$y_k = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \begin{bmatrix} x_k \\ d_k \end{bmatrix} + w_k, \tag{3.4}$$

where $y \in \mathbb{R}^{n_y}$, C_1 and C_2 are matrices that specify which of the variables belonging to the plant and disturbance model, respectively, are to be measured, and $w \sim \mathcal{N}(0, R)$ are normally distributed additive output measurement noise not correlated to a. To simplify notation, we consider only system (3.3) with state variable $z_k = \left[x_k^T, d_k^T\right]^T$, driving white noise sequence $\nu_k = \left[a_k^T, v_k^T\right]^T$, and redefine the output matrix in Eq. (3.4) as $\mathbf{H}$ to read

$$\begin{aligned} z_{k+1} &= \mathbf{A} z_k + \mathbf{B} u_k + \nu_{k+1}, \\ y_k &= \mathbf{H} z_k + w_k, \end{aligned} \tag{3.5}$$

where matrices **A** and **B** follow from Eq. (3.3):

$$\mathbf{A} = \begin{bmatrix} A & F \\ 0 & \Phi \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}. \tag{3.6}$$

For inputs u which are known to be bounded by upper and lower limits, the guarantee of stability for *open-loop unstable* systems can not made. Since these bounds are usually present in the constraints of the MPC formulation, the guarantee for stability is restricted to systems whose matrix **A** has all of its eigenvalues inside the unit circle. Given the fact that stability is the primary concern for any control system, we state the following technical assumption to make clear the applicability of the design procedure in this thesis:

Assumption 3.1.1. *The eigenvalues of the combined state transition matrix* **A** *in Eq. (3.3) are all contained inside the unit circle.*

3.1.2 Control System Components

To simplify notation, we define a measurement system $\mathcal{M}$ as a pair $\mathcal{M} = (\mathbf{H}, R)$. The matrix **H** specifies the variables that are measured while R indicates the *level of precision* with which they are measured. Note that for any practical problem, the number of rows and non-zero entries of **H** along with the allowable variances in R must be known. Restrictions on **H** may result from physical limitations or accessibility hindrances, while the diagonal entries of R may be limited by the availability of sensors. For example, thermistors, resistive thermal devices (RTDs), and thermocouples can all measure a temperature state, but the precision with which temperature measurements are made varies amongst the three sensor types (as does the capital cost).

The objective is to find the economically optimal measurement system configuration for stochastic linear systems. From dynamic programming and Bellman's principle of optimality, the optimization of a stage-wise process is initiated by optimizing the final stage n over

all possible inputs (outputs from stage $n-1$). Once stage n has been optimized, stage $n-1$ is treated as the new final stage, and the optimization is performed again until the entire process has been optimized. This idea, in the context of the present work, is depicted in Figure 3.1 where the progression from measured variable selection to closed loop performance over the expected lifespan of the measurement equipment is sketched. In this thesis, we design the first stage (measured variable selection and sensor specification) based on a filtering and control strategy which are known in advance, and are designed to be optimal during their operation.

- Control Strategy: The control strategy is chosen to be model predictive control (MPC) due to its optimization-based nature and ability to enforce constraints on both the state x and control inputs u. With the proper selection of internal objective function and constraints formulation, it may be possible to design the control algorithm to produce economically optimal control inputs at each execution, subject to the quality of the state estimate supplied.

- State Estimator: The proposed state estimator is the Kalman filter (KF) due not only to its popularity and flexibility, but more importantly due to its inherently optimal design. The KF equations are constructed in such a way that a minimum variance estimate of the state is produced for any measurement system $\mathcal{M}$ (subject to structural conditions on the system such as detectability).

This central idea is very similar to the problem considered by Mellefont and Sargent [1977] and Harris et al. [1980], where measurement system design was considered in a pre-specified Kalman filter/linear-quadratic-Gaussian control context. In what follows, the details necessary to synthesize a full connection between the selection of measured variables and process economics are developed.

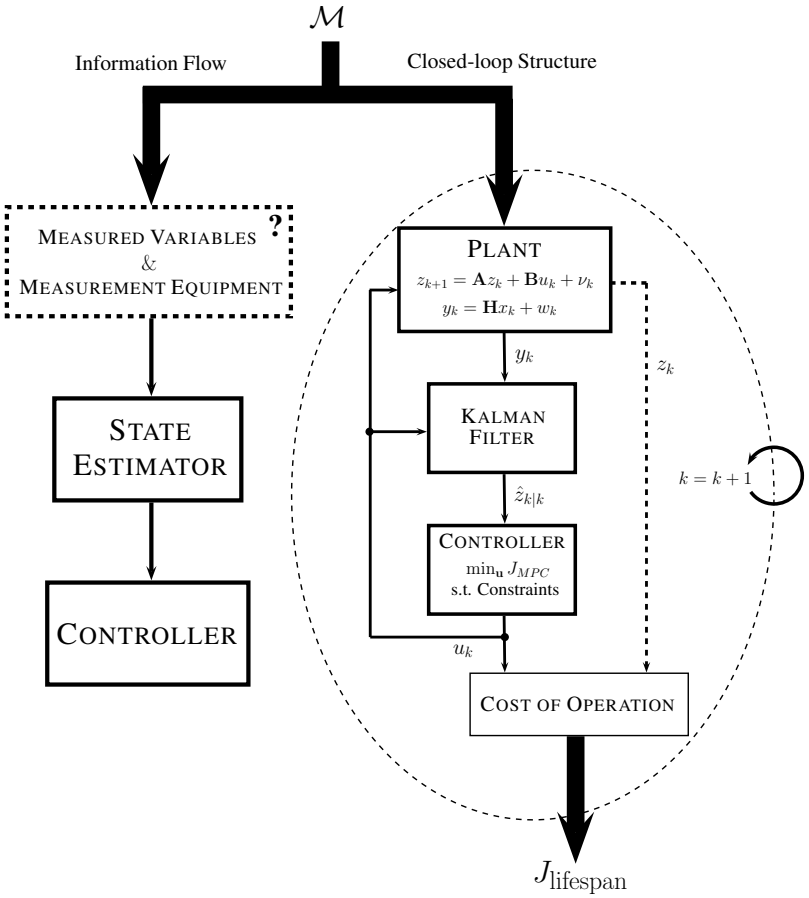

Figure 3.1: A conceptualization of measurement system design's affect on life cycle cost. The measurement system $\mathcal{M}$ (specified once) affects economics through the internal connection between the plant, state estimator (Kalman filter), and controller.

3.2 Controller Synthesis

3.2.1 General Form

Since process economics is a major concern, the control system is designed such that at each execution, the control input administered to the plant results in an optimal economic behaviour. We also wish to ensure that the state be contained in the admissible region $\mathbb{X}$, which may be time-dependent. Constraints on the inputs must also be included in the formulation due to the physical limitations that the actuators may possess.

For simplicity, we will restrict ourselves to the case where $N_{pr,1} = N_{pr,2} = N_{pr}$. The most general form of our model predictive controller is as follows:

$$\max_{u_{N_{pr}}} J_{MPC},$$
$$\text{subject to:} \begin{cases} \Pr(\underline{X}_i \leq x_{k+i} \leq \bar{X}_i) \geq 1 - \alpha_i \quad \forall\ i = \{1, \cdots, N_{pr}\}, \\ \underline{U}_i \leq u_{k+i} \leq \bar{U}_i \quad \forall\ i = \{0, \cdots, N_{pr} - 1\}, \\ x_{k+1} = Ax_k + Bu_k + Fd_k + w_{k+1}, \\ d_{k+1} = \Phi d_k + v_{k+1}, \end{cases} \tag{3.7}$$

where J_{MPC} is the expected revenue generated by operation over the prediction horizon, $u_{N_{pr}} = \left[u_k^T, u_{k+1}^T, \cdots, u_{k+N_{pr}-1}^T\right]^T$ contains each control input for the next N_{pr} number of samples into the future, $\underline{X}_i$ and $\bar{X}_i$ are the lower and upper state constraints for the next i-number of sampling instants into the future, and $\underline{U}_i$ and $\bar{U}_i$ are the lower and upper constraints on the input vector for the next i-number of sampling instants into the future. The inequalities in Eq. (3.7) are understood to apply element-wise. Under this formulation, the admissible sets $\mathbb{X}_i$ introduced in Eq. (2.6) are orthotopes*, which we denote by $\mathcal{O}_i$. Note that the exact form of the objective function J_{MPC} is yet to be specified, since it depends on the particular application. Also, constraints on the state have necessarily been posed probabilistically, since the evolution of the state is subject to the influence of random variables.

It is desirable to convert the probabilistic constraints on the state x_{k+i} in Eq. (3.7) into deterministic ones which can be written solely in terms of the decision variable $u_{N_{pr}}$. If this can be done, it is possible that the optimization problem might be able to be rewritten in some standard form (such as a linear program (LP), quadratic program (QP), etc.). This conversion is the subject of the next section.

*An *orthotope* is the generalization of a rectangle for higher dimensions. In the present context, an orthotope defines the set of interval constraints on the elements of the state vector.

3.2.2 Conversion to a Deterministic Problem

In this section, the probabilistic constraint on the state vector x in Eq. (3.7) is converted into a deterministic constraint expressed solely in terms of the MPC decision variable $u_{N_{pr}}$. The conversion follows directly from the work by Yan and Bitmead [2005]; van Hessem and Bosgra [2002] and references therein. As will be shown, the conversion leads to a linear constraint on the decision variable. This is desirable, since LPs and QPs, among other optimization problem types, are both characterized by linear inequality constraints.

At each controller execution, an estimate of the combined state vector given all information up to and including time k, $\hat{z}_{k|k}$, is required to predict the future values of the constrained state x. A Kalman filter (KF) is used to supply this state estimate:

$$\hat{z}_{k|k} = (I - K_\infty \mathbf{H})(\mathbf{A}\hat{z}_{k-1|k-1} + \mathbf{B}u_{k-1}) + K_\infty y_k, \tag{3.8a}$$

$$K_\infty = (P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}, \tag{3.8b}$$

$$P_\infty^- = -\mathbf{A}(P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}(\mathbf{H}P_\infty^-)\mathbf{A}^T + \mathbf{A}P_\infty^- \mathbf{A}^T + \mathbf{Q}, \tag{3.8c}$$

$$P_\infty^+ = (I - K_\infty \mathbf{H})P_\infty^-(I - K_\infty \mathbf{H})^T + K_\infty R K_\infty^T, \tag{3.8d}$$

where P_∞^- and P_∞^+ are the converged *a priori* and *a posteriori* estimation error covariance matrices, respectively, K_∞ is the converged KF gain matrix, and $\mathbf{Q}$ has the form:

$$\mathbf{Q} = \begin{bmatrix} Q_a & 0 \\ 0 & Q_v \end{bmatrix}. \tag{3.9}$$

The KF equations can be modified to handle so-called "coloured" noise processes, which exhibit autocorrelation. For simplicity, however, we consider only white noise processes. A comprehensive discussion on the KF, and many references, can be found in [Simon, 2006]. Even if the state variable z is subject to normally distributed random inputs, under a constrained MPC control law the control u depends on $\hat{z}$ and z in a nonlinear fashion, and

so the state variables themselves are not normally distributed. However, it can be said that the state estimation error $\tilde{z}_{k|k} = (z_k - \hat{z}_{k|k})$ is zero-mean normally distributed with covariance matrix P_∞^+ (see [Yan and Bitmead, 2005] and references therein). The i^{th}-step-ahead prediction error covariance matrix is given by

$$P_i = \mathbf{A}^i P_\infty^+ \mathbf{A}^{iT} + \sum_{j=0}^{i-1} \mathbf{A}^j \mathbf{Q} \mathbf{A}^{jT}, \tag{3.10}$$

and the probabilistic constraint $\Pr(x_{k+i} \in \mathcal{O}_i) \geq 1 - \alpha_i \ \forall i$ can be rewritten as

$$\frac{1}{\sqrt{(2\pi)^{n_x} \det(P_i)}} \int_{\mathcal{O}_i} e^{-\frac{1}{2}(x_{k+i} - \hat{x}_{k+i|k})^T P_i^{-1} (x_{k+i} - \hat{x}_{k+i|k})} dx_{k+i} \geq 1 - \alpha_i, \tag{3.11}$$

which expresses the constraint on x in terms of $\hat{x}$, whose future values can be determined exactly. For the purposes of this investigation, the ellipsoidal approximation of the probability constraint in question, as suggested by Yan and Bitmead [2005] is employed. The ellipsoidal approximation for constraint enforcement involves defining confidence ellipsoids $\mathcal{E}_i$ around the prediction of the state $\hat{x}_{k+i|k}$ such that $\Pr(x_{k+i} \in \mathcal{E}_i) \geq 1 - \alpha_i$:

$$\mathcal{E}_i = \{x_{k+i} : \left(x_{k+i} - \hat{x}_{k+i|k}\right)^T (P_i^{1/2} P_i^{1/2})^{-1} \left(x_{k+i} - \hat{x}_{k+i|k}\right) \leq r_i^2\}, \tag{3.12}$$

where $P_i^{1/2} P_i^{1/2} = P_i$ and r_i is chosen so that

$$1 - \alpha_i = \frac{1}{2^{n_x} \Gamma(n_x/2)} \int_0^{r_i^2} \chi^{\frac{n_x}{2} - 1} e^{-\frac{\chi}{2}} d\chi, \tag{3.13}$$

since the probability that $x_{k+i} \in \mathcal{E}_i$ has χ^2-distribution with n_x degrees of freedom. Because the prediction error is zero-mean, the ellipsoid in Eq. (3.12) is centered at $\hat{x}_{k+i|k}$, with size and shape determined by entries in P_i and r_i^2 (through α_i), as depicted in Figure 3.2.

It is unlikely that the allowed probability of violation α_i can vary over the course of the prediction horizon, and so the subscript "i" from α_i, r_i^2, $\mathcal{E}_i$, and $\mathcal{O}_i$ is dropped. Finally, the

condition that $\mathcal{E} \subset \mathcal{O}$, and hence that $\Pr(x_{k+i} \in \mathcal{O}) \geq 1 - \alpha$ is guaranteed as long as the following constraints are satisfied (the derivation of this is given in Appendix A):

$$r\sqrt{e_j^T P_i e_j} + e_j^T \hat{x}_{k+i|k} \leq e_j^T \bar{X}_i, \tag{3.14}$$

and

$$r\sqrt{e_j^T P_i e_j} - e_j^T \hat{x}_{k+i|k} \geq e_j^T \underline{X}_i, \tag{3.15}$$

for all $j = \{1, \cdots, n_x\}$ where e_j are the standard basis vectors. The probability constraints in the MPC formulation Eq. (3.7) can therefore be written as

$$e_j^T \underline{X}_i + r\sqrt{e_j^T P_i e_j} \leq e_j^T \hat{x}_{k+i|k} \leq e_j^T \bar{X}_i - r\sqrt{e_j^T P_i e_j}, \tag{3.16}$$

which must hold for all $j = \{1, \cdots, n_x\}$ and $i = \{1, \cdots, N_{pr}\}$.

Replacing the probabilistic constraints in Eq. (3.7) with those derived above, the MPC

optimization problem can be written as

$$\max_{u_{N_{pr}}} J_{MPC},$$

$$\text{subject to:} \begin{cases} e_j^T \hat{x}_{k+i|k} \leq e_j^T \bar{X}_i - r\sqrt{e_j^T P_i e_j}, \\ e_j^T \hat{x}_{k+i|k} \geq e_j^T \underline{X}_i + r\sqrt{e_j^T P_i e_j}, \\ \underline{U}_i \leq u_{k+i} \leq \bar{U}_i, \\ \hat{z}_{k|k} = (I - K_\infty \mathbf{H})(\mathbf{A}\hat{z}_{k-1|k-1} + \mathbf{B}u_{k-1}) + K_\infty y_k, \\ K_\infty = (P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}, \\ P_\infty^- = -\mathbf{A}(P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}(\mathbf{H}P_\infty^-)\mathbf{A}^T \\ \qquad + \mathbf{A}P_\infty^- \mathbf{A}^T + \mathbf{Q}, \\ P_i = \mathbf{A}^i P_\infty^+ \mathbf{A}^{iT} + \sum_{j=0}^{i-1} \mathbf{A}^j \mathbf{Q} \mathbf{A}^{jT}, \\ P_\infty^+ = (I - K_\infty \mathbf{H})P_\infty^-(I - K_\infty \mathbf{H})^T + K_\infty R K_\infty^T, \end{cases} \tag{3.17}$$

In the derivation of Eq. (3.17), the state constraints have been defined with respect to the *entire state vector*. Such a constraint formulation only makes sense when, for example, the constraint violation of a single element of x is as undesirable as the simultaneous constraint violations of *several* elements of x. Formulation of performance constraints on linear combinations of one or several sub-vectors of the plant's state follows in a straightforward manner (see Appendix A), in which case additional probabilistic constraints in Eq. (3.7) will be required.

Note that the choice of measurement system $\mathcal{M} = (\mathbf{H}, R)$ affects the cost of operation first through the Kalman filter equations, then through multi-step-ahead state prediction error covariance matrix, and finally affects the max/min constraints on the states (Eqs. (3.8), (3.10), and (3.16), respectively). The loss of information (availability or quality) results in a reduced search region during MPC's operation, which limits the process's performance.

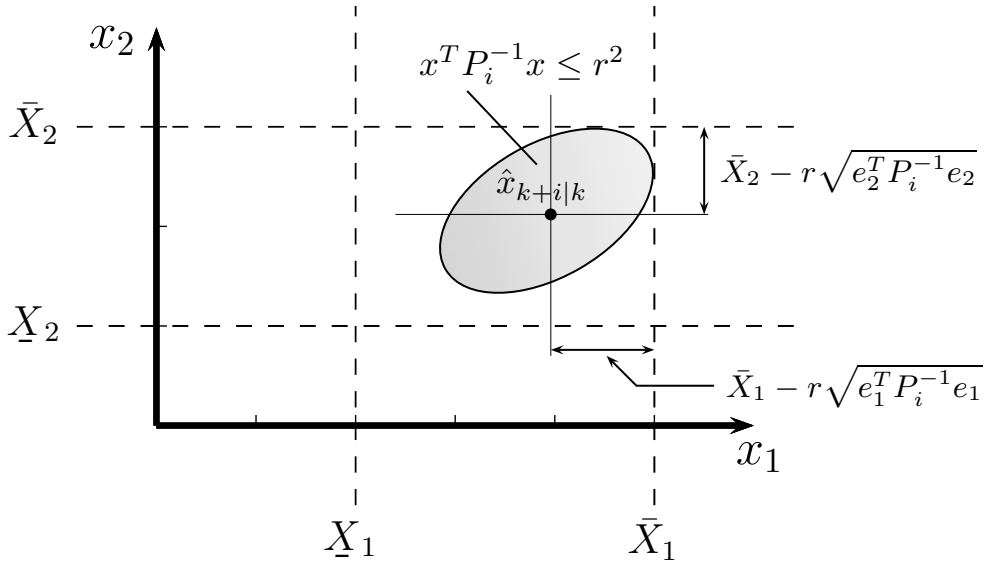

Figure 3.2: A 2-D sketch showing a feasible region defined by the upper and lower constraints $\bar{X}_j$ and $\underline{X}_j$, $j = \{1, 2\}$ for variables x_1 and x_2. Also shown is a confidence ellipsoid centered at $\hat{x}_{k+i|k}$.

Feasibility of the Optimization Problem

As is discussed by Camacho and Bordons [2007], it is possible that as the prediction horizon increases, the growing uncertainty region around the expected values of future states might grow so large that they preclude the existence of a solution to the optimization problem Eq. (3.17). Even if the problem remains feasible, excessively conservative performance might be encountered. The problem stems from the fact that when the controller is executed, a prediction is made for the entire prediction horizon without taking into account in any way the fact that additional information in the form of measurements will be available in the future. The concept of "closed-loop covariances" introduced and discussed by Yan and Bitmead [2005] is employed here to prevent infeasibility problems associated with these growing uncertainty regions around state trajectories. The closed-loop covariance strategy involves only the one-step-ahead prediction error covariance matrix to build a confidence interval around all state trajectories for each of the N_{pr} number of steps into the future. Only the one-step-ahead prediction error covariance matrix is used since the implemented

version of MPC only relies on one-step-ahead prediction error uncertainties because only the first control input calculated by the optimization routine is actually used at each sampling time.

Complete Model Predictive Control Problem

Using the closed-loop covariance strategy, the MPC optimization problem is rewritten as:

$$\begin{aligned} &\max_{u_{N_{pr}}} \; J_{MPC}, \\ &\text{subject to:} \begin{cases} e_j^T \hat{x}_{k+i|k} \leq e_j^T \bar{X}_i - r\sqrt{e_j^T P_1 e_j}, \\ e_j^T \hat{x}_{k+i|k} \geq e_j^T \underline{X}_i + r\sqrt{e_j^T P_1 e_j}, \\ \underline{U}_i \leq u_{k+i} \leq \bar{U}_i, \\ \hat{z}_{k|k} = (I - K_\infty \mathbf{H})(\mathbf{A}\hat{z}_{k-1|k-1} + \mathbf{B}u_{k-1}) + K_\infty y_k, \\ K_\infty = (P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}, \\ P_\infty^- = -\mathbf{A}(P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}(\mathbf{H}P_\infty^-)\mathbf{A}^T \\ \qquad + \mathbf{A}P_\infty^- \mathbf{A}^T + \mathbf{Q}, \\ P_1 = \mathbf{A}P_\infty^+ \mathbf{A}^T + \mathbf{Q}, \\ P_\infty^+ = (I - K_\infty \mathbf{H})P_\infty^-(I - K_\infty \mathbf{H})^T + K_\infty R K_\infty^T, \end{cases} \end{aligned} \tag{3.18}$$

for all $i = \{1, \cdots, N_{pr}\}$ for constraints on $\hat{x}_{k+i|k}$, for all $i = \{0, \cdots, N_{pr} - 1\}$ for constraints on u_{k+i}, and for all $j = \{1, \cdots, n_x\}$. Specification of the objective function J_{MPC} is the final component required to fully construct the model predictive controller.

In process control, revenues generated through operation are typically described by "price-times-quantity" terms, that can be written as a linear combination of state and input variables:

$$J_{MPC} = \beta(k)^T \hat{x}_{N_{pr}} + \delta(k)^T u_{N_{pr}}, \tag{3.19}$$

where $\hat{x}_{N_{pr}} = \left[\hat{x}_{k+1}^T, \cdots, \hat{x}_{k+N_{pr}}^T\right]^T$ is a vector which contains the expected values of the state x for all future sampling instants $i = \{1, \cdots, N_{pr}\}$. The vectors β and δ contain price factors which convert the entries of $\hat{x}_{N_{pr}}$ and $u_{N_{pr}}$ to units of currency. The time argument on β and δ indicates that the entries of these vectors may depend on the "real" time of operation. An example of time varying prices are those imposed by electric utility companies in order to reflect the change in demand during on- and off-peak hours.

If J_{MPC} can be written as in Eq. (3.19), then Eq. (3.18) becomes a linear program (LP). From an implementation point of view, this is desirable since LPs are convex problems whose globally optimal solutions can be calculated via the simplex or interior-point algorithms in a relatively low amount of time [Nash and Sofer, 1996]. On-line implementation can become an issue when other algorithms are required to solve more difficult optimization problems. Under the proposed control strategy, only the two parameters α and N_{pr} are required to tune the controller. In addition to the low number of parameters, another attractive feature of this MPC formulation is the fact that the two parameters are both physically meaningful. As such, *a priori* knowledge of the statistical parameters ($\mathbf{Q}$ and R) and process heuristics can be used to design α and N_{pr}. Increasingly conservative performance can be achieved by lowering the value of α (or increasing r), which tightens the admissible region for $\hat{x}$ in Eq. (3.18). Economic improvements to operation can be achieved by increasing N_{pr}, which might allow the controller to exploit time-dependent prices or costs ($\beta(k)$ and $\delta(k)$, respectively). Of course, the disadvantage with manipulating the parameters in these directions is the possibility of an infeasible or computationally demanding optimization problem.

3.3 Multi-objective Optimization Procedure

It is clear from the controller's formulation that uncertainty in the future value of the state will affect economics in an undesirable way, though exactly how severe this degradation in

performance will be remains to be estimated. In this study, various measures of uncertainty and capital cost are assumed to be the primary measures of quality for any candidate measurement system. The structural condition of observability is not assumed to be necessary *a priori*.

An appropriate capital cost of the measurement system J_{meas} is given by

$$J_{\text{meas}} \triangleq \sum_i \sum_j c_j q_{i,j}, \tag{3.20}$$

where c_j is the expected purchased and installed cost of sensor type j, and $q_{i,j} = \{0, 1\}$ specifies if variable i is measured by sensor type j.

Additionally, consider the single amalgamated uncertainty measure

$$J_{\text{uncert}} \triangleq \text{tr}\{WP_1\}, \tag{3.21}$$

where W is a weighting matrix whose entries are assigned based on engineering judgement. The above measure is motivated by the closed-loop covariances discussion and final MPC formulation. Other measures of uncertainty could be used, such as those considered by Harris et al. [1980]; Muske and Georgakis [2003] and Musulin et al. [2005], however the measure presented above is more representative of the uncertainty considering MPC's explicit dependency on the entries of the one-step-ahead prediction error covariance matrix. The weighting matrix W allows one to assign relative importance to the prediction quality of the individual states x_i. Therefore, in this thesis J_{uncert} alone is used to describe the uncertainty aspect of all candidate measurement systems.

Through the design of a measurement system $\mathcal{M}$, we wish to minimize both entries of the multi-objective vector

$$\mathbf{J} = \left[J_{\text{meas}}, J_{\text{uncert}}\right]^T, \tag{3.22}$$

though to do so is clearly not possible. For example, consider a given measurement system

$\mathcal{M} = (\mathbf{H}, R)$. If an additional measurement is added, we expect the value of J_{uncert} to decrease. However, the capital cost of the measurement system J_{meas} has increased due to the additional equipment that must be purchased.

Consider the equivalent measurement system description $\bar{\mathcal{M}} = (\mathbf{H}^{\text{full}}, \bar{R})$, where $\mathbf{H}^{\text{full}}$ corresponds to an output matrix $\mathbf{H}$ which measures every possible entry of the combined state vector z. Under this reformulation, all sensors are assumed to be physically present, though the structure of R is modified so that its diagonal elements may take on extremely large values (in order to mimic the absence of a sensor). The $(i,i)^{th}$ entry of the modified measurement noise covariance matrix $\bar{R}$ is therefore parameterized as

$$\bar{R}_{i,i} = \left(\sigma^2_{i,1} \cdot q_{i,1} + \sigma^2_{i,2} \cdot q_{i,2} + \cdots + \sigma^2_{i,N_{s,i}} \cdot q_{i,N_{s,i}} + \sigma^2_{\text{large}} \cdot q_{i,N_{s,i}+1}\right), \tag{3.23}$$

where $\sigma^2_{i,j}$ is the measurement variance of sensor type j on variable i, $N_{s,i}$ is the number of sensor types available for the measurement of variable i, σ^2_{large} is a relatively large number used to mimic the absence of a sensor, and the binary variable $q_{i,j}$ is the same as in Eq. (3.20). For a practical design, it is likely that only one of $\{q_{i,1}, q_{i,2}, \cdots, q_{i,N_{s,i}}\}$ for any i can be non-zero.

With this structural modification, a measurement system's design can be expressed solely in terms of the binary decision variables $q_{i,j}$. A multi-objective optimization problem can

therefore be formulated:

$$\min_{q_{i,j}} \mathbf{J} = [J_{\text{meas}}, J_{\text{uncert}}]^T,$$

$$\text{subject to:} \begin{cases} J_{\text{meas}} = \sum_i \sum_j c_j q_{i,j}, \\ J_{\text{uncert}} = \text{tr}\{W P_1\}, \\ P_1 = \mathbf{A} P_\infty^+ \mathbf{A}^T + \mathbf{Q}, \\ P_\infty^+ = \left(I - K_\infty \mathbf{H}^{\text{full}}\right) P_\infty^- \left(I - K_\infty \mathbf{H}^{\text{full}}\right)^T + K_\infty \bar{R} K_\infty^T, \\ K_\infty = (P_\infty^- (\mathbf{H}^{\text{full}})^T)(\mathbf{H}^{\text{full}} P_\infty^- (\mathbf{H}^{\text{full}})^T + \bar{R})^{-1}, \\ P_\infty^- = -\mathbf{A}(P_\infty^- (\mathbf{H}^{\text{full}})^T)(\mathbf{H}^{\text{full}} P_\infty^- (\mathbf{H}^{\text{full}})^T + \bar{R})^{-1}(\mathbf{H}^{\text{full}} P_\infty^-)\mathbf{A}^T \\ \qquad + \mathbf{A} P_\infty^- \mathbf{A}^T + \mathbf{Q}, \\ \bar{R}_{i,i} = \left(\sigma_{i,1}^2 \cdot q_{i,1} + \sigma_{i,2}^2 \cdot q_{i,2} + \cdots + \sigma_{i,N_{s,i}}^2 \cdot q_{i,N_{s,i}} + \sigma_{\text{large}}^2 \cdot q_{i,N_{s,i}+1}\right), \\ q_{i,j} = \{0, 1\}, \\ c_{N_{s,i}} = 0, \end{cases} \tag{3.24}$$

for all $i = \{1, \cdots, n_{mv}\}$ where n_{mv} is the number of variables available for measurement (which is not necessarily equal to $n_x + n_d$), and where the final constraint $c_{N_{s,i}} = 0$ reflects the fact that not measuring variable i results in no increase of the capital cost J_{meas}. Note that if observability is a requirement, it is possible to obtain a theoretical lower bound on the number of sensors that must be installed. The discussion on observability in Chapter 2 indicates that the minimum number of sensors required to ensure observability is equal to the number of independent modes of the system. In this thesis, however, the condition of observability, which corresponds to full state reconstructability, is not assumed to be necessary. Instead, the weighting matrix W is introduced as a design parameter that allows the designer to directly place relative estimation importance on the individual states.

The solution to a multi-objective optimization problem whose decision variables are binary is a finite set of points termed the *Pareto optimal set*, also referred to as the *Pareto frontier* [Branke et al., 2008; Steuer, 1986]. Three potential strategies by which to exactly solve or approximate the multi-objective optimization problem in Eq. (3.24) are presented next.

3.3.1 Identifying the Pareto Optimal Set

In [Muske and Georgakis, 2003] a Pareto optimal set was constructed for a continuously stirred tank reactor (CSTR) process by considering every possible combination of nine state variables whose measurements resulted in an observable system. If a state space formulation is small enough to accommodate this approach, construction of the Pareto optimal set is tractable. However, if the problem contains many potential measurements, an approximation is required due to the combinatorial explosion associated with the increasing number of variables available for measurement. For example, a system with 30 variables available for measurement by one type of sensor results in $2^{30} \approx 10^9$ potential measured variable combinations.

The construction of the Pareto optimal set corresponds to a significant reduction in the number of measurement system configurations that need to be considered. This is because non-Pareto optimal points can be completely ruled out from the design procedure, since we know that there exists a measurement system with the same capital cost and has a lower measure of uncertainty. This idea is depicted visually in Figure 3.3.

How one chooses to identify the Pareto optimal set depends on several factors including the dimensionality of the state space representation, the number of viable sensor types, and the computing power available. In the discussion that follows, three possible methods by which to approximate or exactly determine the Pareto optimal set are presented.

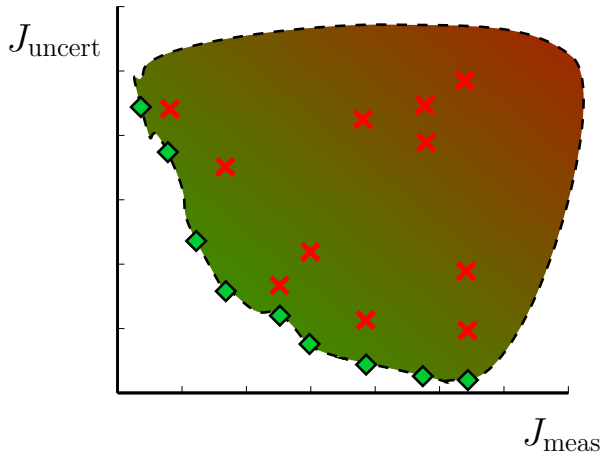

Figure 3.3: A sketch of a Pareto optimal set which considers $\mathbf{J} = [J_{\text{meas}}, J_{\text{uncert}}]^T$. Pareto optimal measurement systems are deonted by ($\diamond$) and sub-optimal points are denoted by ($\times$).

Branch-and-bound with Semi-definite Programming (BnB/SDP)

The work by Chmielewski et al. [2002] forms a basis for a means by which to construct the Pareto optimal set via binary integer programming. The idea behind this approach is that an optimization problem with J_{meas} as the objective function subject to constraints on individual state estimation error variances can be posed as a binary integer program (BIP) subject to linear matrix inequality (LMI) constraints. In what follows, the important results reported by Chmielewski et al. [2002] are first presented and then modified slightly to suit our specific needs.

Theorem 3.3.1. *[Chmielewski et al., 2002]: Let $\bar{R} > 0$, $\mathbf{Q} > 0$, $\mathbf{A}^{-1}$ exist, and the pair $(\mathbf{A}, \mathbf{Q}^{1/2})$ be stabilizable. Then there exists $P_\infty^+ > 0$ such that $\text{tr}\{U^T P_\infty^+ U\} < \gamma^2$ and Eq. (3.8d) is satisfied if and only if there exists $X > 0$ such that*

$$\text{tr}\{U^T(\mathbf{A}^T X \mathbf{A})^{-1} U\} < \gamma^2, \tag{3.25}$$

and

$$\begin{bmatrix} X - \mathbf{A}^T X \mathbf{A} + (\mathbf{H}^{\text{full}})^T \bar{R}^{-1} \mathbf{H}^{\text{full}} & X \\ X & X + \mathbf{Q}^{-1} \end{bmatrix} > 0. \tag{3.26}$$

The inequality on a generic matrix Z means $Z > 0 \Leftrightarrow v^T Z v > 0 \; \forall v \neq 0$, i.e. Z is *positive definite*. If individual state estimate precision constraints are considered ($U = \mathbf{h}_i^T$, where $\mathbf{h}_i$ is the i^{th} row of $\mathbf{H}^{\text{full}}$), application of the Schur complement theorem on Eq. (3.25) leads to the positive definiteness requirement

$$\begin{bmatrix} \gamma_i^2 & \mathbf{h}_i \\ \mathbf{h}_i^T & \mathbf{A}^T X \mathbf{A} \end{bmatrix} > 0. \tag{3.27}$$

Using the results of Theorem 3.3.1, the identification of Pareto optimal measurement systems can be performed by solving

$$\min_{q_{i,j}, X} J_{\text{meas}} = \sum_i \sum_j c_j q_{i,j},$$

$$\text{subject to:} \begin{cases} \begin{bmatrix} X - \mathbf{A}^T X \mathbf{A} + (\mathbf{H}^{\text{full}})^T \bar{R}^{-1} \mathbf{H}^{\text{full}} & X \\ X & X + \mathbf{Q}^{-1} \end{bmatrix} > 0, \\ \begin{bmatrix} \gamma_i^2 & \mathbf{h}_i \\ \mathbf{h}_i^T & \mathbf{A}^T X \mathbf{A} \end{bmatrix} > 0, \\ X > 0, \\ \mathbf{A}^{-1} \text{ exists}, \\ \bar{R}_{i,i} = \left(\sigma_{i,1}^2 \cdot q_{i,1} + \cdots + \sigma_{i,N_{s,i}}^2 \cdot q_{i,N_{s,i}} + \sigma_{\text{large}}^2 \cdot q_{i,N_{s,i}+1} \right), \\ q_{i,j} = \{0, 1\}. \end{cases} \tag{3.28}$$

An optimization problem such as Eq. (3.28) can be solved via the branch-and-bound algorithm, where at each node a semi-definite program (SDP) is solved (as opposed to the usual linear programs for mixed integer linear programming problems). Details on the branch-and-bound algorithm can be found in [Edgar et al., 2001].

Note that the precision constraints considered by Chmielewski et al. [2002] are defined in terms of the 0-step ahead prediction error covariances. Due to the dependence of J_{uncert} on 1-step ahead prediction error covariances, a reformulation to reflect this dependence is desired. Conversion of the precision constraints in Eq. (3.28) to constraints expressed in terms of the 1-step ahead prediction error covariances is given in the following proposition.

Proposition 3.3.1. *Individual precision constraints γ_i in Eq. (3.28) can be expressed in terms of the* 1*-step ahead prediction uncertainties by replacing the positive definiteness constraint given by Eq. (3.27) with*

$$\begin{bmatrix} \tilde{\gamma}_i^2 & \mathbf{h}_i \\ \mathbf{h}_i^T & X \end{bmatrix} > 0. \tag{3.29}$$

To show this, consider the precision constraint in terms of the 1-step ahead prediction uncertainty

$$\text{tr}\{U^T P_1 U\} < \gamma^2, \tag{3.30}$$

keeping in mind that U will be set to $U = \mathbf{h}_i^T$. Eq. (3.30) is equivalent to the inequality

$$\text{tr}\{U^T(\mathbf{A}P_0\mathbf{A}^T + \mathbf{Q})U\} < \gamma^2,$$

which implies

$$\text{tr}\{U^T\mathbf{A}P_0\mathbf{A}^TU\} + \text{tr}\{U^T\mathbf{Q}U\} < \gamma^2,$$

or

$$\text{tr}\{U^T\mathbf{A}P_0\mathbf{A}^TU\} < \tilde{\gamma}^2,$$

where $\tilde{\gamma}^2 = \left(\gamma^2 - \text{tr}\{U^T\mathbf{Q}U\}\right)$.

Since the matrix U is arbitrary, define $\bar{U}^T = U^T\mathbf{A}$ to obtain

$$\text{tr}\{U^T P_1 U\} < \gamma^2 \Leftrightarrow \text{tr}\{\bar{U}^T P_0 \bar{U}\} < \tilde{\gamma}^2.$$

By Theorem 3.3.1, the constraint $\text{tr}\{\bar{U}^T P_0 \bar{U}\} < \tilde{\gamma}^2$ is satisfied if and only if there exists $X > 0$ such that

$$\text{tr}\{\bar{U}^T(\mathbf{A}^T X \mathbf{A})^{-1}\bar{U}\} = \text{tr}\{U^T X^{-1} U\} < \tilde{\gamma}^2,$$

where $\mathbf{A}$ is assumed invertible. Substituting $U^T = \mathbf{h}_i$ to represent individual precision constraints and applying the Schur complement theorem, we arrive at the expression given by Eq. (3.29).

The equivalent optimization problem in terms of 1-step ahead prediction uncertainty is therefore given by:

$$\min_{q_{i,j},X} J_{\text{meas}} = \sum_i \sum_j z_j q_{i,j},$$

$$\text{subject to:} \begin{cases} \begin{bmatrix} X - \mathbf{A}^T X \mathbf{A} + (\mathbf{H}^{\text{full}})^T \bar{R}^{-1} \mathbf{H}^{\text{full}} & X \\ X & X + \mathbf{Q}^{-1} \end{bmatrix} > 0, \\ \begin{bmatrix} \gamma_i^2 - \mathbf{h}_i \mathbf{Q} \mathbf{h}_i^T & \mathbf{h}_i \\ \mathbf{h}_i^T & X \end{bmatrix} > 0, \\ X > 0, \\ \mathbf{A}^{-1} \text{ exists}, \\ \bar{R}_{i,i} = \left(\sigma_{i,1}^2 \cdot q_{i,1} + \cdots + \sigma_{i,N_{s,i}}^2 \cdot q_{i,N_{s,i}} + \sigma_{\text{large}}^2 \cdot q_{i,N_{s,i}+1}\right), \\ q_{i,j} = \{0,1\}. \end{cases} \tag{3.31}$$

Exact Pareto optimal points can be found by adjusting values of γ_i^2 near a point which

corresponds to a measurement system known in advance to be nearly Pareto optimal. Such approximately optimal $\mathcal{M}$ may be located by the following approximation technique.

Sequential Sensor Addition and Removal

This strategy is a heuristic way of locating Pareto optimal measurement systems which works by turning "on" or "off" the decision variables $q_{i,j}$ in an order that is determined by a series of local evaluations of the two objective measures J_{meas} and J_{uncert}. The process is analogous to a gradient-based optimization procedure for functions of continuous variables. The proposed approximation strategy via sequential sensor removal is as follows (the addition procedure follows naturally):

1. Construction of the Pareto optimal set is initialized by the full-state measurement and highest precision sensors point $\mathcal{M}^\star = (\mathbf{H}^{\text{full}}, R^{\text{high}})$, which corresponds to the Pareto optimum characterized by both the highest measurement cost J_{meas} and lowest uncertainty J_{uncert}.

2. A high-quality sensor is replaced by its next lowest quality option (which might correspond to sensor absence) based on which replacement has the least impact on the scalar measure
$$J_r = \frac{J_{\text{uncert}}(\mathcal{M}_{\text{next}}) - J_{\text{uncert}}(\mathcal{M}^\star)}{J_{\text{meas}}(\mathcal{M}^\star) - J_{\text{meas}}(\mathcal{M}_{\text{next}})}, \tag{3.32}$$
where $\mathcal{M}_{\text{next}}$ is the measurement system similar to $\mathcal{M}^\star$ but with one sensor of lower quality (or absent). The incremental measure J_r is used since at each step it is desirable to have a small uncertainty increase and large cost decrease.

3. Step 2 is repeated until all sensors have been removed.

Exhaustive Search

The third and final technique, known as "exhaustive search", is the simplest to perform (when possible). The exhaustive search technique involves evaluating J_{meas} and J_{uncert} for every combination of measured variables and sensors types available. This technique is only feasible when the total number of possible $\mathcal{M}$ is low. The advantage, of course, is that all Pareto optimal points can be identified exactly.

It should be noted that unlike the work by Muske and Georgakis [2003], we do not limit our search to observable systems, since for controllable or stabilizable systems only the weaker condition of detectability is required for the existence of the steady-state solutions to the KF equations (see Chapter 2).

3.3.2 Decision-making Process

The final step of the measurement system design procedure requires a *decision maker* [Branke et al., 2008; Steuer, 1986] to select one from the potentially several Pareto optimal measurement system configurations. In this thesis, simulation experiments are the proposed means to generate data (process economics in particular) over a sample-length duration of typical operation. Only the exact or best approximate Pareto optimal sensor configurations located by the above strategies need to be considered. Process economics data can then be scaled up to the expected lifespan of the measurement equipment and combined with the purchased and installed cost of measurement equipment to generate a final total measurement system cost:

$$J_{\text{total}} = J_{\text{meas}} + t_{\text{scale}} \cdot J_{op} \tag{3.33}$$

where J_{total} is the total cost incurred by the choice of measurement system $\mathcal{M}$, and t_{scale} is a dimensionless scaling which converts the cost of a sample-length operation J_{op} to the expected lifespan of the measurement equipment. The measurement system which minimizes

the total cost is then chosen as the overall optimum.

For the simulation experiments, the MPC strategy given by Eq. (3.18) is used to design the controller since its formulation provides a direct way of maximizing revenue while satisfying safety and product quality constraints. If a different control strategy is known to be necessary *a priori*, it can be used for decision-making purposes, but closed-loop performance results are likely to be plagued by trade-offs between constraint violations and economics.

3.3.3 Summary

In this chapter, measurement system design for stochastic linear systems has been posed as a multi-objective optimization problem in order to preserve the characteristic objective measures of a measurement system design's quality. Minimization of the two objective measures – capital cost and uncertainty – is desired, but is not possible due to their conflicting nature. Three methods have been proposed by which the Pareto optimal set of measurement systems can be identified, either exactly or approximately, with the applicability of each strategy contingent on the number of decision variables involved. For decision-making purposes, model predictive control is the control strategy of choice. A controller whose operation maximizes process revenue subject to constraints on the inputs and state variables has been designed. Data from closed-loop simulation experiments can then be used to map approximate or exact Pareto optimal measurement systems to a cost of operation over an expected lifespan of the measurement equipment. The measurement system which minimizes the combined cost of capital investment and operation over the predicted lifespan is then chosen as the optimum design.

The next two chapters contain illustrative examples which showcase the various components of the full measurement system design procedure presented in this chapter. The first example problem involves a low-dimensionality fluid handling network, on which the three multi-objective optimization solution techniques are applied in order to investigate their relative performances. The second example problem considers a high-dimensionality

thermal-network description of a one-floor office building whose air temperatures are to be controlled at minimum cost. The focus of this example problem is the performance of the model predictive controller and subsequent decision-making strategy.

Chapter 4

Application to a Fluid Handling Network

The purpose of this chapter is to investigate the performance of each of the three multi-objective optimization solution techniques on a low-dimensionality state space model. The application of MPC as a decision-making aide is the focus of the next chapter. The problem is taken from [Chmielewski et al., 2002], and is of low enough dimensionality to include the exhaustive search technique as a potential Pareto optimal set identification strategy. Clearly, if the exhaustive search is possible, then the Pareto optimal set can be identified exactly, and BnB/SDP and sequential techniques are not required. The latter techniques are included in this chapter to assess their performance in comparison to the true Pareto optimal set identified by an exhaustive search method.

4.1 System Description

The flow diagram for this example problem is shown in Figure 4.1. Node 2 is a storage tank whose mass at time k is denoted by $m(k)$. The tank has a leak stream, denoted by $l(k)$, that cannot be measured directly. A disturbance model is not known, and so an augmented

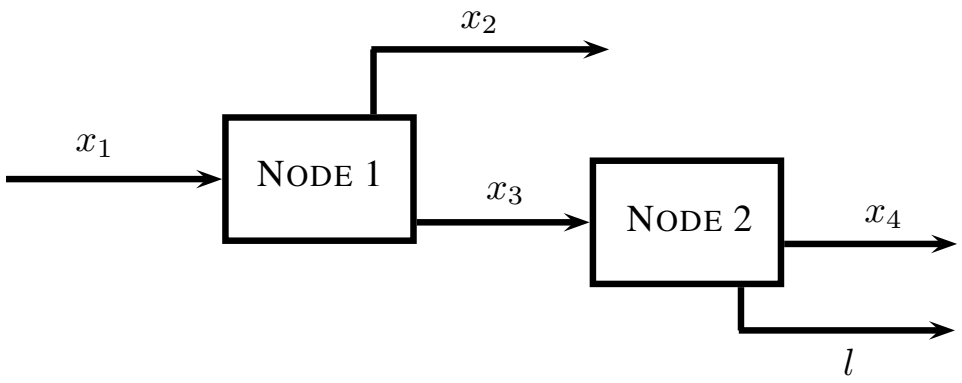

Figure 4.1: Flow diagram of the example problem given in [Chmielewski et al., 2002]. The flow rates x_i can be measured, but the leak stream l cannot.

state-space representation such as Eq. (3.5) is not possible. The mass balances equations for this system are

$$x_1(k) = x_2(k) + x_3(k),$$
$$m(k+1) = m(k) + x_3(k) - x_4(k) - l(k),$$

where $x_i(k)$ is the mass of fluid passed through stream i between sampling instants. Recall that the applicability of the BnB/SDP strategy is contingent on the invertibility of the discrete dynamics' A matrix. To ensure that A^{-1} exists, one considers a set of *primary* variables $x_p(k) = [x_2(k) \;\; x_3(k) \;\; x_4(k) \;\; m(k) \;\; l(k)]^T$, whose dynamics are given by $x_p(k+1) = Ax_p(k)$. For analysis purposes, measurement systems of the type $\bar{\mathcal{M}} = (\mathbf{H}^{\text{full}}, \bar{R})$ are considered, as defined in Chapter 3, and whose outputs $y(k) = Cx_p(k)$ correspond to the entire state. In this case, $\mathbf{H}^{\text{full}} = C$, with the A and C matrices given by

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.1}$$

Sensor Type	Precision 1 ($\sigma_{1,j}$)	Precision 2 ($\sigma_{2,j}$)	Precision 3 ($\sigma_{3,j}$)	Precision 4 ($\sigma_{4,j}$)	Precision 5 ($\sigma_{5,j}$)	Cost ($)
1	28.5	9.9	18.5	18.5	–	2,500
2	57.0	19.8	37.1	37.1	–	1,500
3	85.4	29.8	55.7	55.7	–	800
4	–	–	–	–	16	2,600
5	–	–	–	–	64	500

Table 4.1: Precision and costs for the five sensor types available for measurement system design in the current example problem.

It is also assumed that flow during a given time period is subject to noise of a known covariance structure:

$$x_p(k+1) = Ax_p(k) + Fw(k), \tag{4.2}$$

where $w(k)$ is normally distributed white noise with covariance matrix Q, and F is given as

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.3}$$

It is assumed that the amount of correlation between $x_i(k)$ and $x_i(k+1)$, $i = \{2, 3, 4\}$ is small when Δt is large. Therefore the covariance matrix Q is defined via $Q^{-1} = \text{diag}\{[10^{-6} \;\; 10^{-6} \;\; 10^{-6} \;\; 1/(2.15)^2]\}$*, where the entries 10^{-6} indicate a large variance on the first three components of $w(k)$.

A number of different measurement devices exist for the various state variables, each with their own capital cost and measurement precision. The precision (in terms of standard deviation values) as well as purchased and installed costs of the sensors are given in Table 4.1.

*In [Chmielewski et al., 2002], Q is defined via $Q^{-1} = \text{diag}\{[0 \;\; 0 \;\; 0 \;\; 1/(2.15)^2]\}$, which, as the authors argue, is sufficient given that Q does not need to be evaluated at any point during the BnB/SDP analysis.

4.2 Procedure

4.2.1 Sequential Addition and Removal

The sequential addition and removal techniques as outlined in section 3.3.1 are applied to generate an approximate Pareto optimal set. A slightly modified version of the uncertainty given by Eq. (3.21) is used. It is given by:

$$J_{\text{uncert}} = \text{tr}\{WP_{\infty}^{+}\}, \tag{4.4}$$

since no model predictive control is to be applied to this problem, and hence 1-step ahead prediction error covariances do not need to be considered. The above uncertainty measure is calculated by weighting the individual estimation error uncertainties of states 1 (inlet flow), 4 (tank mass), and 6 (leak stream) a factor of 100,000 times higher than those of the remaining states. This is chosen following suggestions reported by Chmielewski et al. [2002] that these states are the most crucial for estimation. Specifically, diagonal entries $W(1,1) = W(4,4) = W(6,6) = 1000$ while remaining diagonal entries were set to 0.01. For each $\mathcal{M}$ considered during the sequential techniques, the estimation error covariance matrix P_{∞}^{+} is calculated recursively using the following form of the *a posteriori* error covariance matrix:

$$P_{\infty}^{+} = \left[\left(AP_{\infty}^{+}A^{T} + FQF^{T}\right)^{-1} + C^{T}\bar{R}^{-1}C\right]^{-1} \tag{4.5}$$

where $\bar{R}$ has the form given by Eq. (3.23). The recursion is initialized at $P_{\infty}^{+}(0) = 1000 \cdot I$ where $I \in \mathbb{R}^{5\times 5}$ is the identity matrix. Recursion is terminated after the following termination criterion is reached:

$$\|P_{\infty}^{+}(i) - P_{\infty}^{+}(i-1)\|_{F} < \epsilon = 1 \times 10^{-6} \tag{4.6}$$

where $P_{\infty}^{+}(i)$ is the i^{th} iteration of P_{∞}^{+} and $\|\cdot\|_{F}$ denotes the matrix Frobenius norm.

4.2.2 BnB with Semi-definite Programming

A number of approximate Pareto optimal measurement systems can be generated by following the procedure outlined in the previous section. Following this, a select number of the resulting estimation variances $P_\infty^+(1,1)$, $P_\infty^+(4,4)$, and $P_\infty^+(6,6)$ can then be relaxed slightly to form upper bound precision constraints γ_i^2 in Eq. (3.28). In this thesis, the MATLAB^{TM} optimization package YALMIP [Löfberg, 2004] is used to solve binary integer programming problems and semi-definite programs. By defining a total of 19 binary decision variables (one for each of the 4 sensor options on states 1-4 and 3 sensor options on state 5) plus positive-definiteness constraints on the matrices in Eq. (3.28), the branch-and-bound algorithm can then be executed. The branch-and-bound algorithm used is contained in the file `bnb.m` of the YALMIP optimization package. The semi-definite programs at each node are solved using the semi-definite programming solver named "SeDuMi" [Pólik, 2009]. The feasibility tolerance of the branch-and-bound algorithm is set to 0 to ensure that the precision constraints γ_i^2 are not violated.

4.2.3 Exhaustive Search

In this problem, 5 states available for direct measurement with the possibility of 4 different types of sensors on states 1-4, and 3 different sensor types on state 5. The total number of measurement system configurations is 768. Consequently, the individual calculations of J_{meas} and J_{uncert} as given by Eqs. (3.20) and (4.4), respectively, can be reasonably performed. These calculations can be performed by executing the MATLAB^{TM} file `Chmielewski_2b_Exhaustive.m`, which calculates every J_{meas} and J_{uncert} for the 768 possible measurement system designs in a straightforward manner.

4.3 Results & Discussion

The results obtained by the sequential addition/removal strategies and exhaustive search technique are shown in Figure 4.2. The vertical axis displays the natural logarithm of J_{uncert}. The results demonstrate that both the sequential addition and removal techniques locate measurement systems characterized by comparable uncertainty and capital costs. Some differences in the quality of the approximate Pareto optimal points are seen in the interior region of Figure 4.2, and especially at the low number of sensors region. This is due to the nature of the sequential techniques, which are merely a heuristic way of locating approximately Pareto optimal points. As their operations progress, a further departure from true Pareto optimal points is increasingly more likely to occur because at stage n a removal/addition decision is made based on decisions that were made at all stages $\{n-1, \cdots, 1\}$, where the optimality of stage n was not of concern. Once the sequential removal technique has begun approximating Pareto optimal points at the low number of sensors region, this effect of sub-optimal optimization has been felt to a greater extent than by the sequential addition's operation at the same region, which has only just started to locate Pareto optimal points. These differences in approximation quality at the high and low number of sensor regions suggests that one should use the sequential techniques in conjunction – rather than one or the other exclusively – to approximate the Pareto optimal set. Figure 4.3 displays nearly the same results as Figure 4.2, the difference being that the results of the exhaustive search have been removed, and the results of the BnB/SDP strategy have been colour-coded. Examination of the interior region of Figure 4.3 reveals the result that the BnB/SDP strategy is able to locate points that found by the sequential techniques (for example, the green star near \$9000), as well as points that were not. The quality of the unique points that BnB/SDP is able to find is variable. Table 4.2 shows two particular cases: one where BnB/SDP outperforms the sequential techniques, and one where the BnB/SDP strategy falls short. The upper half of Table 4.2 shows the results of

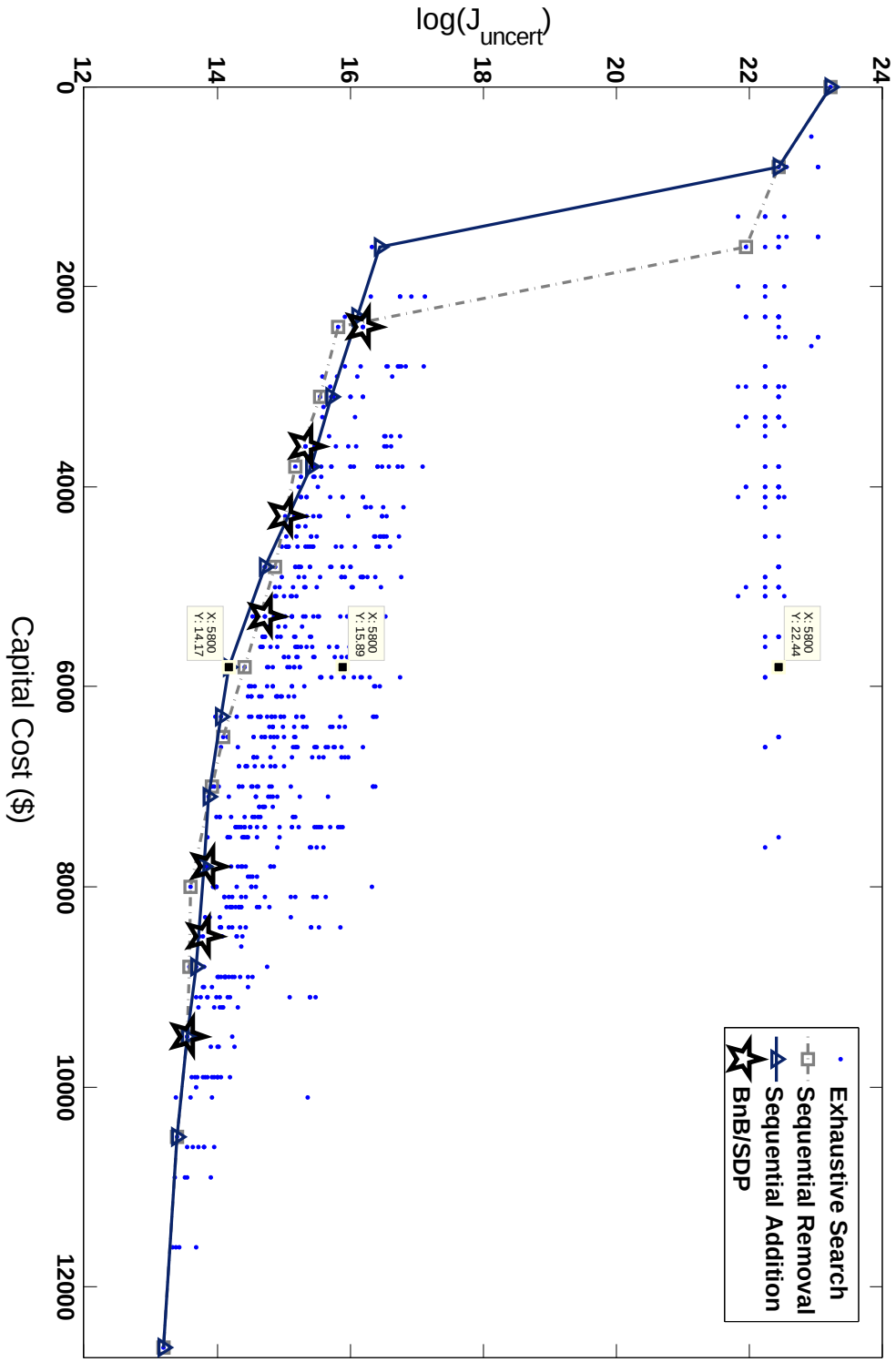

Figure 4.2: The results of the sequential addition and removal strategies (△ and □, respectively), BnB/SDP (★), and exhaustive search (·) superimposed for visual comparison. Three sample measurement system configurations are shown to emphasize that many *poor* combinations of measured variables and sensor types exist.

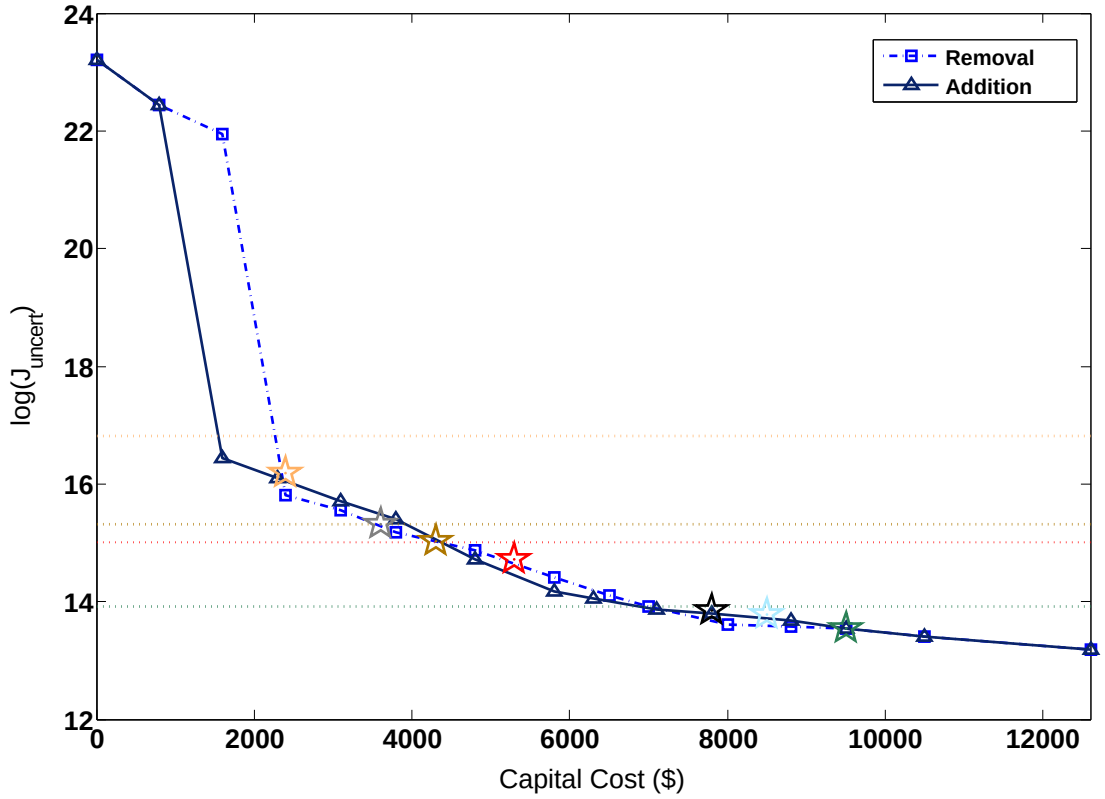

Figure 4.3: Pareto frontier generated via sequential removal (□) and sequential addition (△). Also shown are the Pareto optimal points generated by semi-definite programming (★). Dashed horizontal lines correspond to maximum uncertainty values used to construct stars of corresponding colour (for cleanliness, only a select 4 are shown).

a particular BnB/SDP run which produced a measurement system with the same capital cost as one located by the sequential addition technique ($J_{\text{meas}} = \$7800$). In this case, the sequential technique was able to find a better approximate Pareto optimal point than the BnB/SDP approach, as seen by the two differing values of the uncertainty measure J_{uncert}. The opposite result is also possible. An example of this is given in the lower half of Table 4.2, where BnB/SDP outperforms the sequential addition strategy by the same arguments as above.

The reason for the identification of sub-optimal Pareto points by BnB/SDP stems from the slight incompatibility between the uncertainty (precision) constraints formulation of

Location Method	$P_\infty^+(1,1)$	$P_\infty^+(4,4)$	$P_\infty^+(6,6)$	$\log_e(J_{\text{uncert}})$	J_{meas}
BnB/SDP	652.48	325	61.75	13.85	\$7,800
Upper Limit γ_i^2	700	550	80	–	–
Sequential Addition	580	326.2	75.5	13.80 ✓	\$7,800
BnB/SDP	812.1	1376.2	266.98	14.7153 ✓	\$4,800
Upper Limit γ_i^2	820	1400	290	–	–
Sequential Addition	812.1	1376.2	267.08	14.7201	\$4,800

Table 4.2: Precision and costs for the five sensor types available for measurement system design in the current example problem.

Eq. (3.31) and the Pareto uncertainty measure used to categorize measurement systems given by Eq. (4.4). The remedy for this incompatibility is to simply pose the first stage of the multi-objective optimization procedure as a multi-objective optimization problem characterized by several independent objective function uncertainty measures (one for each state estimate uncertainty). The advantage of this is a more direct method of locating Pareto optimal points. However, it is likely that many more Pareto optimal measurement systems may then need to be individually considered during the subsequent decision-making process.

Despite the possibility that BnB/SDP can converge to sub-optimal Pareto points, it is clear from Figure 4.2, which contains the results from the exhaustive search technique, that BnB/SDP is successful at finding very good measurement system configurations (their proximity to the Pareto optimal set is very low). It is easily seen that the sequential techniques are also quite successful at finding nearly or exactly Pareto optimal points, though a number of exact Pareto optimal points are missed along the way.

4.4 Summary

In this chapter, the three multi-objective optimization solution strategies presented in Chapter 3 were applied to a low-dimensionality fluid handling network. The results suggest that

the combined use of the sequential and BnB/SDP strategies can result in a high quality approximation of the true Pareto optimal set. This can be achieved since: 1) the sequential addition and removal strategies can identify high quality approximate Pareto optimal points at the low and high number of sensors region, respectively, and 2) the BnB/SDP is capable of finding high quality Pareto optimal points in the interior region by adjusting the precision constraints of the SDP.

Chapter 5

Application to a Thermal Network

In this chapter, the measurement system design procedure of Chapter 3 is applied to a thermal network description of a building system model. Due to the high dimensionality of the state space model, only the sequential addition/removal techniques are used to approximate the Pareto optimal set. Decision-making via simulation experiments is performed using the steady-state Kalman filtering and profit maximizing MPC strategies.

5.1 Example Problem Description

We consider the temperature control of a one-floor office building represented by a thermal network model, a popular approach for modeling heat transfer in the building systems literature [Wang and Xu, 2006; Zhang and Hanby, 2006; Xu and Wang, 2008; Lee and Braun, 2008; O'Neill et al., 2010]. The purpose of this example problem is to apply the techniques developed in Chapter 3 to design an economically optimal measurement system whose expected lifespan is 10 years.

The building system, whose floor plan is shown in Figure 5.1, contains seven zones that are defined based on the availability of local actuators. In this example, it is assumed that the office is subject to a warm climate. As a result, only air conditioning (cooling) is required

to maintain a comfortable atmosphere for the building's occupants. Air conditioning in each zone is accomplished via fan coil units (FCUs), which supply cool air to the zones. The white, unlabeled spaces in Figure 5.1, are uncontrolled areas whose temperatures are considered to be disturbances. Surrounding the building is a corridor whose temperature is also uncontrolled. Further disturbances include the ground temperature, ceiling temperature, and internal load (which is due to occupants, lights, etc.). Under the thermal network formulation, each surface (wall, floor, ceiling) and mass of air is assigned a uniform temperature whose dynamic behaviour is modeled as due to conductive or convective heat transfer [Çengel, 2006]. Hence, for a reasonably complex model, many temperatures are available for measurement.

The objective of indoor temperature control is to keep the temperatures of each zone within their user-specified upper and lower limits, while minimizing the cost of operation. This problem, also in the context of model predictive control is considered in [Lute and van Paassen, 1994], though the MPC objective function considered is of a mixed type, including terms which correspond to dollar-valued operation cost and occupant comfort. Model predictive controllers whose objective functions include terms corresponding to various performance measures have been successfully applied in industry (such as quadratic dynamic matrix control (QDMC) [Bequette, 2003]). However, a drawback with such control strategies is that even if the controller parameters can be tuned to achieve acceptable performance, the analysis of closed-loop operation is likely to be complicated by tradeoffs involving the multiple performance criteria.

This problem is well suited to illustrate the application of the techniques proposed in this thesis since:

- a large number (> 100) of physically interpretable variables are candidates for measurement,
- the controlled variables are subject to several stochastic disturbance inputs,

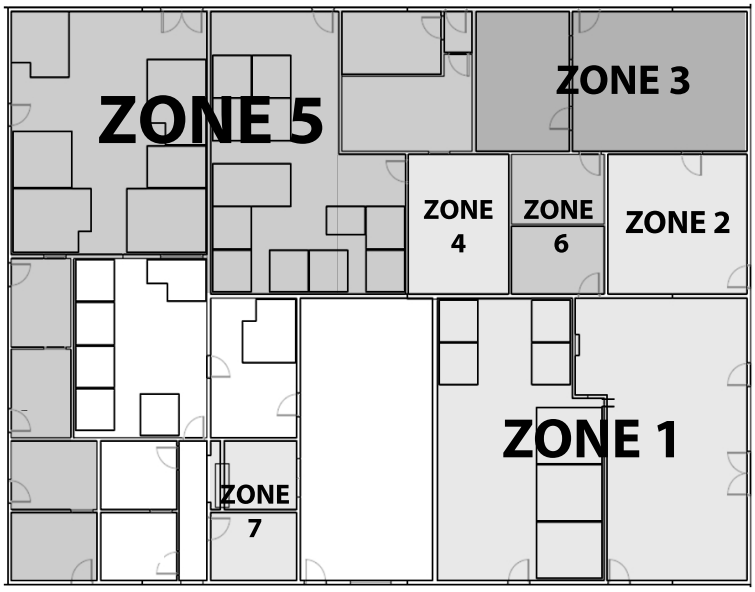

Figure 5.1: A floor plan of the one-floor office building considered in this chapter. Zones are defined based on the availability of an actuator. White spaces and the area surrounding the figure are uncontrolled areas whose temperature is considered a disturbance.

- comfort constraints (max/min room temperatures) are crucial to enforce.

5.2 Procedure

5.2.1 Modeling Indoor Temperature and Disturbance Dynamics

Indoor Temperature Dynamics

The reduced order model of the temperature dynamics is based on mass and energy balances around the entire building envelope and around each zone. Under the thermal-network framework, an effective modeling technique is the "3R2C" approximation. The "3R" component refers to the three resistances to heat transfer between adjacent zones. Resistances 1) and 2) are due to the combined effects of two convective heat transfer stages between surfaces and air, and resistance 3) originates from the conductive heat transfer through solid surfaces. The term "2C" refers to the discretization of walls and other surfaces into

two separate entities, each with its own uniform temperature and thermal capacitance. Neglecting water evaporation, exogenous air infiltration, and interzone air mixing, an energy balance around zone j is given by:

$$M_{air,j}C_{p,air}\frac{dT_{zone,j}}{dt} = \dot{m}_{sa}C_{p,air}(T_{sa} - T_{zone,j}) + Q_{int,j} + \sum_{i=1}^{N_{surf,j}} Q_{surf,i}, \tag{5.1}$$

where $T_{zone,j}$ (K) and $M_{air,j}$ (kg) are zone j's air temperature and mass of air, $C_{p,air}$ $(\frac{J}{kg \cdot K})$ is the specific heat capacity of air (assumed constant), $\dot{m}_{sa}$ (kg/s) and T_{sa} (K) are the mass flow rate and temperature of air supplied by the FCU, $Q_{int,j}$ (W) is zone j's internal load, $Q_{surf,i}$ (W) is the heat exchanged to/from the room via the various surfaces (walls, ceiling, ground), and $N_{surf,j}$ is the number of surfaces adjacent to zone j's air mass. For a surface adjacent to zone j, the rate of change of its temperature is given by

$$C_{s,i}\frac{dT^{in}_{surf,i}}{dt} = h_j A_{s,i}(T_{zone,j} - T^{in}_{surf,i}) + \frac{T^{out}_{surf,i} - T^{in}_{surf,i}}{R_{con,i}}, \tag{5.2}$$

where $T^{in}_{surf,i}$ (K) is the i^{th} surface adjacent to zone j's mass of air. The temperature $T^{out}_{surf,i}$ (K), is that of the surface adjacent to the i^{th} surface, h_j $\left(\frac{W}{m^2 \cdot K}\right)$ is the convective heat transfer coefficient between the surface and zone j's air (assumed constant), $R_{con,i}$ (K/W) is the conductive resistance between two adjacent surfaces, and $A_{s,i}$ (m^2) is the area of the surface. The thermal resistance and capacitance are defined by

$$R_{con,i} = \frac{w_i}{k_i A_{s,i}}, \quad C_{s,i} = \frac{\rho_i C_{p,i} w_i A_{s,i}}{2}, \tag{5.3}$$

where w_i (m), and k_i $\left(\frac{W}{m \cdot K}\right)$ are the thickness and thermal conductivity of surface i respectively, ρ_i (kg/m^3) is the density of the surface's material, and $C_{p,i}$ $\left(\frac{J}{kg \cdot K}\right)$ is the specific heat capacity of the surface's material. The thermal-network parameters used are based on those determined during on-line estimation by the control systems research group at the United

Technologies Research Center (UTRC) for the particular building that is considered in this chapter, all of which are contained at the top level of the MATLAB^{TM} file `makeSS_D.m` The FCU is assumed to have local level controls that regulate both the air flowrate $\dot{m}_{sa}$ and temperature of air delivered to the zone, T_{sa}. If the method of feedback linearization [Khalil, 2002] is employed, all seven $\dot{m}_{sa}C_{p,air}(T_{sa} - T_{zone,j})$ terms in Eq. (5.1) can be grouped into one control input vector, u_k, whose entries at time k represent the power removed (for cooling) from each room. Note that $T_{zone,j}$ is a state variable whose value at any sampling instant is not known exactly. If the zone temperature $T_{zone,j}$ in the proposed expression for u is replaced by an estimate, the result is a stochastic control variable, which renders the applicability of the Kalman filter in Chapter 3 invalid. To ensure that u_k can be known exactly at all future sampling instants, it is assumed that measurement of the mass flow rate and temperatures at the inlet and outlet of each FCU's cool water stream are available to the local level controllers. A straightforward energy balance over time-span dt can then be performed to calculate $T_{zone,j}$, since it is assumed that all heat removed from a zone's air is gained by the cool water stream passing through the FCU's cooling coil. It is unknown if FCUs with a hardware arrangement such as this currently exist. As such, it is recognized that this assumption may be a critical factor governing the applicability of the MPC control strategy. Nevertheless, in this thesis the standard technical assumption that the control input is known exactly at all sampling instants is made to keep the focus on the topics at hand.

Disturbances

The disturbances that affect the indoor temperature of the office building consist of:

- a corridor temperature, T_{corr} (surrounding the building and in the white space of Figure 5.1),
- a ground temperature, T_{gr},

- a ceiling temperature, T_{ceil},
- internal loads $Q_{int,j}, \quad j = \{1, \cdots, 7\}$.

Since the disturbances, D^i are periodic in nature, they can be modeled as the discrete-time version of an under-damped second order continuous-time system with very low damping ratio ($\zeta = 0.005$)

$$D^i_{k+1} = \phi^i_1 D^i_k + \phi^i_2 D^i_{k-1} + v^i_{k+1}. \tag{5.4}$$

The under-damped second order systems of Eq. (5.4) are zero mean, and so the introduction of additional components is required. The *actual* disturbance variables which affect the plant include the periodic component given by Eq. (5.4) combined with an additional noise term $\bar{v}^i$ (perfectly correlated with v^i). A bias b_i is also added to introduce additional uncertainty phenomenon and move the expected value to a user-defined average daily temperature. As a result, the model takes the form:

$$\bar{D}^i_{k+1} = D^i_k + \bar{v}^i_{k+1} + b^i, \tag{5.5}$$

where $\bar{D}^i$, $i = \{1, \cdots, 4\}$ are the actual values of the disturbances that affect the plant, $\bar{v}^i \sim \mathcal{N}(0, \sigma^2_{\bar{v},i})$ are additive white noise sequences and b^i are constant bias terms. In matrix notation, the i^{th} disturbance variable is modeled as follows:

$$\begin{bmatrix} \bar{D}^i_{k+1} \\ D^i_{k+1} \\ D^i_k \\ b^i_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & \phi^i_1 & \phi^i_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{D}^i_k \\ D^i_k \\ D^i_{k-1} \\ b^i_k \end{bmatrix} + \begin{bmatrix} \sigma_{\bar{v},i} \\ \sigma_{v,i} \\ 0 \\ 0 \end{bmatrix} v^i_{k+1}, \tag{5.6}$$

which can be conveniently rewritten as:

$$d^i_{k+1} = \Phi^i d^i_k + v^i_{k+1}, \tag{5.7}$$

where $d^i_{k+1} = \begin{bmatrix} \bar{D}^i_{k+1} & D^i_{k+1} & D^i_k & b^i_{k+1} \end{bmatrix}^T$, and

$$\Phi^i = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & \phi_1^i & \phi_2^i & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.8}$$

For this example problem, the entire disturbance vector d as given by Eq. (3.2) has the form $d_k = \left[{d_k^1}^T, {d_k^2}^T, {d_k^3}^T, {d_k^4}^T\right]^T$. The dynamics of the disturbance variables are therefore modeled as

$$d_{k+1} = \Phi d_k + v_{k+1}, \tag{5.9}$$

where $v_k = \left[{v_k^1}^T, {v_k^2}^T, {v_k^3}^T, {v_k^4}^T\right]^T$. The block diagonal disturbance state transition matrix Φ is given by

$$\Phi = \begin{bmatrix} \Phi^1 & 0 & 0 & 0 \\ 0 & \Phi^2 & 0 & 0 \\ 0 & 0 & \Phi^3 & 0 \\ 0 & 0 & 0 & \Phi^4 \end{bmatrix}, \tag{5.10}$$

where Φ^i are defined by Eq. (5.8) and 0 are zero matrices of appropriate dimension. The maximum daily temperature/internal load variation (i.e. amplitude) is then adjustable by appropriately specifying the initial condition d_0.

Full State-space Description

A full description of the indoor temperature dynamics results in a system of 99 continuous-time deterministic linear ordinary differential equations:

$$\dot{x} = A_c x + B_c u + F_c d, \tag{5.11}$$

with the dynamics governing d to be included later. The subscript "c" on the A, B, and F matrices in Eq. (5.11) indicates that these matrices describe the continuous-time dynamics. The approach taken to integrate Eq. (5.11) in this thesis is by invoking the MATLAB$^{\text{TM}}$ command `c2d` with the zero-order hold option on a state space object defined by A_c and B_c

in the MATLAB$^{\text{TM}}$ environment. The output is the formation of a new *discrete* state space object containing the matrices A, B, and F of Eq. (3.1). The dynamics of the discrete state space model are therefore described by the linear difference equations

$$x_{k+1} = Ax_k + Bu_k + Fd_k,$$

to which a stochastic component used to simulate plant/model mismatch $a \sim \mathcal{N}(0, Q_a)$ is added. The resulting model of the building system's state evolution is therefore given by

$$x_{k+1} = Ax_k + Bu_k + Fd_k + a_{k+1}. \tag{5.12}$$

Combining Eqs. (5.12) and (5.9), we arrive at the augmented description of the dynamics with state variable $z_k = \left[x_k^T, d_k^T\right]^T$ and driving white noise sequence $\nu_k = \left[a_k^T, v_k^T\right]^T \sim \mathcal{N}(0, \mathbf{Q})$:

$$\begin{aligned} z_{k+1} &= \mathbf{A}z_k + \mathbf{B}u_k + \nu_{k+1}, \\ y_k &= \mathbf{H}z_k + w_k, \end{aligned} \tag{5.13}$$

where $w \sim \mathcal{N}(0, R)$, and matrices $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\mathbf{A} = \begin{bmatrix} A & F \\ 0 & \Phi \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \tag{5.14}$$

where "0" denote zero matrices of appropriate dimensions. The output matrix $\mathbf{H}$ takes the block diagonal form

$$\mathbf{H} = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}, \tag{5.15}$$

where matrices C_1 and C_2 specify which of the plant and disturbance variables, respectively, are to be measured, and "0" again denote zero matrices of appropriate dimensions. Note that the notation used to represent the dynamics of the combined plant/disturbance model

is exactly the same as the notation used in Chapter 3.

A full characterization of the heat transfer between all internal temperature and disturbance states results in a system of 115 difference equations (99 internal states, 16 disturbance states) with 7 control inputs. For measurement system design, the output matrix $\mathbf{H}^{\text{full}}$ as defined in Chapter 3 corresponds to the measurement of all 99 internal states plus all $\bar{D}^i$ and all b^i. It is understood, however, that the internal load disturbance state cannot actually be measured since it is an amalgamation of many variables. It is included to determine its relative importance in the description of the building system's dynamics.

5.2.2 Controller Design

In this section, a suitable model predictive controller is designed to minimize the cost of temperature regulation subject to user-specified max/min zone temperature constraints. Numerical implementation details and tuning parameter selection is left for the later section which discusses the simulation experiments in detail.

The control objective is to maintain a comfortable atmosphere for the building's occupants while incurring as little operational cost as possible. As such, the model predictive controller is designed in such a way that its operation minimizes the cost of electricity usage subject to input constraints (due to limits on actuator capabilities) and constraints on the states corresponding to zone air temperatures.

For this control problem, the expected cost of operation is given by price-times-input terms where the input u in Watts is multiplied by time-dependent price values that reflect the differences in demand during peak- and off-peak hours. The cost of control is performed assuming a 1:1 direct conversion between cooling power and electricity consumption. For more detailed analysis, the conversion can be accomplished via the Air Conditioning, Heating, and Refrigeration Institute's "Seasonal Energy Efficiency Ratio", or SEER.*

*A cooling unit's SEER rating is defined as the ratio between the cooling power delivered by the appliance and its electrical power input [Air-Conditioning, Heating, & Refrigeration Institute, 2008].

Maximum and minimum constraints on the zone air temperatures are assumed to be time-dependent, with the tighter constraints imposed during occupied office hours (8:00 am - 6:00 pm). This operating policy allows off-peak hour "pre-cooling", where thermal storage through the use of the building's thermal mass is exploited. Since the temperature control of individual zones is the primary concern in this problem, it is more appropriate to pose the constraints in the MPC formulation on a zone-by-zone basis, specifying the allowed probability of violation for each zone (as opposed to the entire state vector). As such, r in Eq. (3.18) is replaced by Z_j, the standard normal value corresponding to a maximum violation probability of $1-\alpha_j$ for zones $j = \{1, \cdots, 7\}$. The optimization problem solved at each controller execution is given by

$$\min_{u_{N_{pr}}} J_{MPC} = \delta^T(k) u_{N_{pr}}$$

$$\text{subject to:} \begin{cases} e_j^T \hat{x}_{k+i|k} \leq e_j^T \bar{X}_i - Z_j \sqrt{e_j^T P_1 e_j}, \\ e_j^T \hat{x}_{k+i|k} \geq e_j^T \underline{X}_i + Z_j \sqrt{e_j^T P_1 e_j}, \\ \underline{U}_i \leq u_{k+i} \leq \bar{U}_i, \\ \hat{z}_{k|k} = (I - K_\infty \mathbf{H})(\mathbf{A}\hat{z}_{k-1|k-1} + \mathbf{B}u_{k-1}) + K_\infty y_k, \\ K_\infty = (P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}, \\ P_\infty^- = -\mathbf{A}(P_\infty^- \mathbf{H}^T)(\mathbf{H}P_\infty^- \mathbf{H}^T + R)^{-1}(\mathbf{H}P_\infty^-)\mathbf{A}^T \\ \qquad + \mathbf{A}P_\infty^- \mathbf{A}^T + \mathbf{Q}, \\ P_1 = \mathbf{A}P_\infty^+ \mathbf{A}^T + \mathbf{Q}, \\ P_\infty^+ = (I - K_\infty \mathbf{H})P_\infty^-(I - K_\infty \mathbf{H})^T + K_\infty R K_\infty^T, \end{cases} \tag{5.16}$$

for all $i = \{1, \cdots, N_{pr}\}$ and $i = \{0, \cdots, N_{pr} - 1\}$ for constraints on $\hat{x}_{k+i|k}$ and u_{k+i}, respectively and for all j indexed appropriately to enforce constraints on the states which represent the seven zone air temperatures. The entries of vector $\delta(k)$ are time-dependent power-to-dollar conversion coefficients.

Since an output reference signal is not involved in the MPC formulation Eq. (5.16), closed-loop stability under MPC can be ensured by directly following the dual mode MPC configuration guidelines presented in Chapter 2. Because the augmented system is open-loop stable, by assigning the terminal region $\mathcal{X}_f = \mathbb{R}^{n_x+n_d}$, terminal control policy $u^f = 0$, and stage costs following the prediction horizon $J^s(x, 0) = 0 \; \forall x \in \mathcal{X}_f$, we ensure that conditions **C1** through **C4** are met (**C4** being met with equality). Of course, the terminal control policy $u^f = 0$ is not actually be implemented, since convergence of the plant's state x to the origin is not one of the control objectives.

5.2.3 Measurement System Design

As stated previously, the objective of this example problem is to design the most economically optimal measurement system by completing the two-stage multi-objective optimization procedure of Chapter 3. The first stage involves the determination of the Pareto optimal set, and the second stage relies on simulation experiments as a means to map Pareto optimal measurement configurations to a cost of operation. The cost values generated by the simulation experiments designed to emulate typical working conditions are then scaled up to the expected lifespan of the measurement equipment. The measurement system that minimizes the combined cost of capital investment and scaled-up operation is chosen as the optimum.

Construction of the Pareto Optimal Set

Given the dimensionality of the state-space model used to describe the temperature dynamics, an exhaustive search was not possible. Therefore, only the sequential addition/removal and branch-and-bound/semi-definite programming techniques were attempted.

For the sequential techniques, two cases are examined. The first assumes that the effect of unmodeled dynamics on the plant were negligible ($Q_a = 0$), while the second case acknowledges the effect of unmodeled dynamics whose covariance structure ($Q_a \neq 0$) is

known exactly. It is reasonable to assume that some amount of model inaccuracy exists since the thermal network model assumes that a number of parameters (heat transfer coefficients, heat capacities, resistances, etc.) are temperature-independent. In reality, these parameters do exhibit a slight temperature dependence. In addition to parameter-based simplifications, other unmodeled mass and heat transfer phenomena cause a discrepancy between the control model's predictions and observed behaviour.

To simulate the effect of these unmodeled phenomena, Q_a is assumed to be diagonal with the i^{th} diagonal entry $Q_a(i,i) = \sigma_a^2 = 0.05$. Variances on the random component of the measurements are modeled by assigning R as a diagonal matrix with $R(i,i) = \sigma_R^2 = 0.005$. That is, all available temperature measurements are assumed to be reliable, with most of the prediction uncertainty originating from the unmodeled dynamics of the plant. The approximate Pareto optimal sets are constructed by following the sequential sensor addition/removal procedures of section 3.3.1. Numerically, the sequential techniques can be performed by executing the MATLAB$^{\text{TM}}$ files `RemoveSensors1.m` and `RemoveSensors1_reverse.m` (corresponding to the sequential removal and addition strategies, respectively).

In all cases, the measure of uncertainty from Eq. (3.21) is used, since the one-step-ahead prediction error covariances are used by the model predictive controller. To assign representative weighting on the importance of the prediction of various states, the weighting matrix W is made diagonal with zero weights on non-zone temperature states and weights of $500 \cdot V_j$ on zone temperature states where V_j is the volume of zone j.

The branch-and-bound/semi-definite programming strategy is applied on the case where $Q_a \neq 0$ by again using the optimization package YALMIP with MATLAB$^{\text{TM}}$. Both the SeDuMi semi-definite programming solver (already introduced in Chapter 4) as well as the MATLAB$^{\text{TM}}$ Robust Control Toolbox's "LMILAB" solver are employed to solve the relaxed problems at each node. Before the branch-and-bound algorithm is executed, feasibility tests for a combination of precision constraints and measured variables, known to lead to a feasible solution performed to obtain estimates of the time-to-completion of the branch-and-bound

algorithm. The BnB/SDP strategies can be initiated by executing the MATLAB^{TM} files `ParetoExact.m` and `LMISetup.m`. The YALMIP command `solvesdp` can then be executed to begin the branch-and-bound procedure.

Simulation-based Experiments

Only the measured variable combinations resulting from sequential and BnB/SDP strategies which were closest to Pareto optimal are considered. Furthermore, only the unmodeled dynamics case $Q_a \neq 0$ is considered, since it is the most representative of real operation. Reliable measurements, as given by R defined previously are used. For these measurement systems, three days of typical operation are simulated in the MATLAB^{TM}/Simulink environment as a means to generate operational cost data for use in the final decision-making process.

The response of the plant (along with the Kalman filter and controller) is simulated in Simulink using the "Discrete State Space" block. This block collects the components of a "discrete state space" object constructed through the conversion of the continuous-time ordinary differential equations of Eq. (5.2) to discrete-time form through the MATLAB^{TM} command `c2d`. Conversion of the ODEs to matrix form takes place in the file named `makeSS_D.m`, while the disturbance transition matrices of Eq. (5.6) are augmented afterwards in the file `RunSim.m` to produce the **A**, **B**, and **H** matrices of Eq. (3.5).

The update/sampling interval for the plant is set to 15 minutes, a reasonable length of time since the relatively slow temperature dynamics of building systems results in appreciable temperature changes occurring on the order of minutes (as opposed to seconds or milliseconds). To standardize the results, all noise added to the plant's states is generated in advance via the MATLAB^{TM} command `randn` and stored in the MAT file `state_noise.mat` which is called upon by Simulink during simulation via a "From File" block. Noise associated with the disturbance variables originate through "Random Number" blocks whose constant seed value of 0 ensures reproducibility of the results.

The Kalman filter operates in nearly the same manner as the Plant. Since the KF is simply a discrete-time dynamical system itself, its operation can be simulated by another "Discrete State Space" block (in conjunction with other math operation blocks) whose arguments follow from the KF equations (3.8a) and (3.8b). Because conditions that represent typical operation are to be simulated, the steady-state filtering strategy is used over the entire course of the 3 day simulations. This experimental design detail is equivalent to the assumption that convergence of P_k^-, P_k^+, and K_k to P_∞^-, P_∞^+, and K_∞, respectively, occurs quickly relative to the expected lifespan of the measurement equipment. Numerically, this is accomplished by using only the converged matrices in Eq. (3.8) and initializing $\hat{z}_{0|0} = z_0$. The solution to the DARE in Eq. (3.8c) can be calculated by invoking the MATLAB^{TM} command `dare`.[†] Further numerical implementation details regarding the KF's operation are available in the subsystem `KALMAN FILTER` contained in Simulink file `BuildingSim.mdl`.

The optimization problem given by Eq. (5.16) is a linear program that lends itself to solution by either the interior-point or simplex algorithms. Preliminary investigations showed that the simplex algorithm that comes standard with the MATLAB^{TM} optimization toolbox (executed by invoking the command `linprog`) is more reliable than the interior-point algorithm supplied in the same software package. In a small number of instances where the interior-point algorithm was used to solve Eq. (5.16), optimization was terminated prematurely for no clearly identifiable reason. For this reason, the simplex algorithm was chosen to solve Eq. (5.16) at each controller execution for all simulation experiments.

Interaction between `linprog` and the remaining simulation components is made possible through the introduction of the "MATLAB-S-Function" block named `MPC_LinProg.m`. Conversion of the MPC equations in Eq. (5.16) into the form required by `linprog` occurs in this file during a call to an internal function named `ieq_output_constraint_gen`. The controller is set to execute every 30 minutes to calculate a new $u_{N_{pr}}$ whose first entry is supplied under a zero-order hold to the plant. The prediction horizon N_{pr} is set to 12 hours

[†]Execution of this command requires that the MATLAB^{TM} Control System Toolbox be installed.

to allow enough time for the controller to utilize the relatively slow dynamics of the plant to store cooling energy in the building's walls, floor, ceiling, and bodies of air. In addition to returning control inputs, `MPC_LinProg.m` also returns a diagnostic variable called `exitflag`, whose values at each controller execution indicate whether or not convergence to a feasible optimal solution is achieved. Knowledge of this diagnostic is critical to ensure that all simulation experiments are equivalent in terms of the controller's successful operation. For example, faulty linear programming operation can lead to excessive constraint violations and unrepresentative cost function values if control inputs corresponding to zero cooling energy are erroneously being administered to the plant.

A number of "To Workspace" blocks are utilized in the Simulink model file whose outputs are first stored in the MATLAB^{TM} workspace and then saved to the hard disk as a `MAT` file under a representative file name. When multiple simulations are run, the file `RunSims.m` is executed. This file automatically changes the parameters of interest between successive simulation experiments, and automatically stores data upon completion.

Post-simulation processing occurs through execution of the file `Performance.m` once the desired simulation results (stored in a `MAT` file) are loaded to the MATLAB^{TM} workspace. This file plots the response of one or all of the seven zone air temperatures on one panel and the control inputs at each sample time on the lower panel. This file also calculates the percentage of simulation time where all seven zone temperatures violated upper and lower temperature constraint values.

A number of operating conditions used in the simulations are given in Table 5.1. Other parameters which are less critical for this discussion are clearly marked at the top level of the source code file `RunSim.m`.

Final Design

The operating cost data generated by the three day long simulations was scaled up to ten years, which was assumed to be the life expectancy of the measurement equipment. Each

Parameter Type	Variable		
	Description	Name	Value (units)
Building Properties	Wall thermal resistance	R_{con}	0.43 ($^\circ C$/W)
	Wall thermal capacitance	C_s	42000 (J/$^\circ C$)
	Occupied hours	t_{oc}	7:00 am - 6:00 pm
	Zone Temp. Max. (all hours)	$\bar{X}$	25 ($^\circ C$)
	Zone Temp. Min. (occupied hrs.)	$\underline{X}_o$	22 ($^\circ C$)
	Zone Temp. Min. (unoccupied hrs.)	$\underline{X}_u$	17 ($^\circ C$)
Disturbances	Ground temperature (avg.)	$\bar{T}_{gr}$	18 ($^\circ C$)
	Ground temperature (amplitude)	$T_{gr,A}$	1 ($^\circ C$)
	Ceiling temperature (avg.)	$\bar{T}_{ce}$	30 ($^\circ C$)
	Ceiling temperature (amplitude)	$T_{ce,A}$	3 ($^\circ C$)
	Corridor temperature (avg.)	$\bar{T}_{co}$	25 ($^\circ C$)
	Corridor temperature (amplitude)	$T_{co,A}$	2 ($^\circ C$)
MPC Parameters	Controller execution period	T_c	30 (min)
	Prediction horizon	N_{pr}	48 (samples)
	Probability of constraint violation for a zone	$1-\alpha$	0.05
Other Parameters	Electricity high-price	e_{high}	0.099 ($/$kWh$)
	Electricity low-price	e_{low}	0.051 ($/$kWh$)
	Electricity high-price time	$t_{e,h}$	7:00 am - 6:00 pm

Table 5.1: Some of the parameters used in closed-loop MPC simulations of the one-floor office building described by a thermal network.

sensor was assigned a price tag of $200, which is understood to be low, but is useful for illustration purposes. The final cost of the measurement system is then calculated by

$$J_{\text{total}} = J_{\text{meas}} + t_{\text{scale}} \cdot J_{op}, \tag{5.17}$$

where J_{meas} is the capital cost of measurement equipment (defined in Eq. (3.20)), t_{scale} is a dimensionless scaling factor that converts the cost of three days of simulation to the 10 year expected lifespan of the measurement equipment, and J_{op} is the cost of operation relative to the full state measurement case:

$$J_{op} = \bar{\delta}^T \left(\mathbf{u}(\mathcal{M}) - \mathbf{u}(\mathcal{M}^\star) \right), \tag{5.18}$$

where the vector $\bar{\delta}$ contains power-to-dollar conversion coefficients for the entire length of a three day simulation, $\mathbf{u}(\mathcal{M})$ is a vector containing the control inputs over the course of the simulation run with measurement system $\mathcal{M}$, and $\mathcal{M}^\star = \left(\mathbf{H}^{\text{full}}, R^{\text{high}} \right)$. As such, J_{op} represents the *opportunity cost* of not measuring every possible state variable with the highest precision.

The exact or approximate Pareto optimal measurement system, $\mathcal{M}$ that minimizes Eq. (5.17) is chosen as the overall optimum.

5.3 Results and Discussion

The results of the investigations described above are presented in this section. Additionally, an account of the attempts at applying the BnB/SDP technique are provided. The BnB/SDP technique is not considered as a candidate Pareto optimal set identification strategy due primarily to the computational infeasibility of the semi-definite programs when a large state space model is considered. An account of several attempts at applying the BnB/SDP strategy is provided.

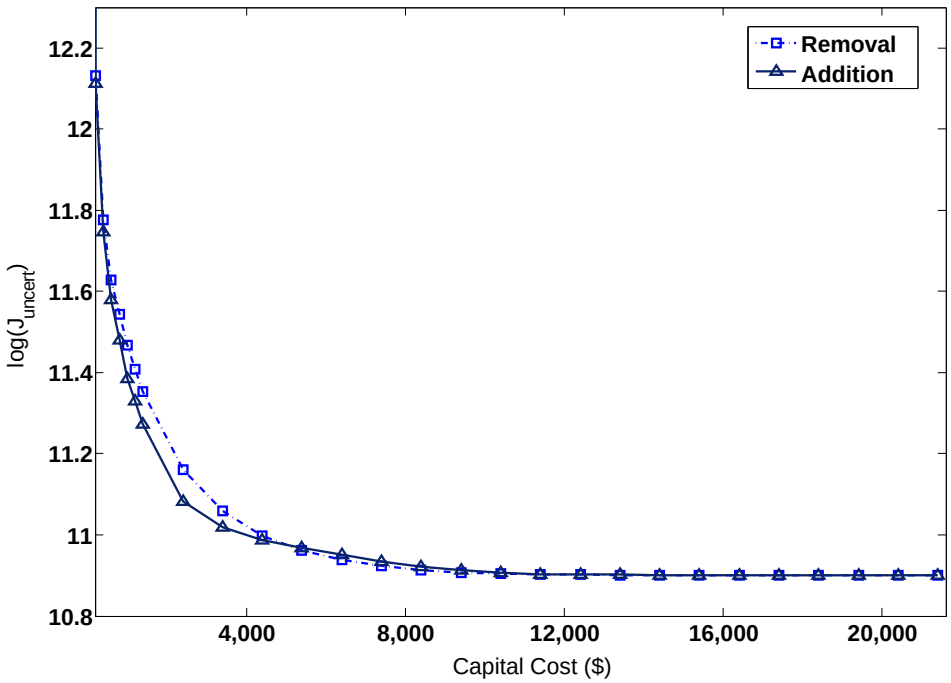

Figure 5.2: Approximate Pareto optimal sets generated via sequential addition (△) and removal (□) strategies for $Q_a \neq 0$.

5.3.1 Approximate Pareto Optimal Set

Sequential Addition/Removal

Figure 5.2 shows the approximate Pareto optimal set produced by the sequential addition/removal strategies for the case where $Q_a \neq 0$. Figure 5.3 shows the same information but at the low number of sensors range (which is the region of interest) for the case where $Q_a = 0$. In both cases, a dramatic decrease in uncertainty is observed after a relatively small number of sensors have been included. Further purchase and installation of measurement equipment results in very modest reductions in uncertainty. It is also observed the sequential addition technique is able to locate measured variable configurations closer to the true Pareto optimal than the sequential removal technique at the low number of sensors region. The explanation for the sequential addition strategy's superior performance in this region is the same as the one already given in Chapter 4.

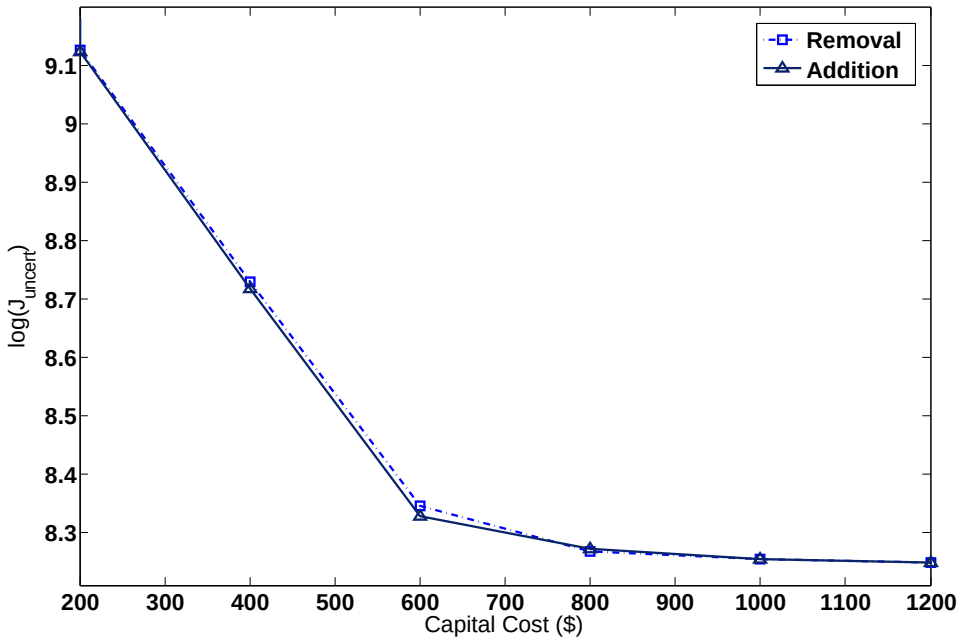

Figure 5.3: Approximate Pareto optimal sets generated via sequential addition (△) and removal (□) strategies for $Q_a = 0$. Only the low number of sensors region is shown.

The states of the augmented model whose measurements result in the dramatic uncertainty decrease at the low number of sensors range are given in Table 5.2 (sequential removal) and Table 5.3 (sequential addition). The clear difference between the two unmodeled dynamics cases is that when $Q_a \neq 0$, greater importance is placed on states whose heat transfer mechanism is in direct contact with the constrained highly-weighted states (recall that the weighting matrix W in Eq. (3.21) resulted in heavy weighting of the zone states) or the constrained states themselves. This is in opposition to importance being placed on the disturbances (which is the case when $Q_a = 0$), whose future effect on air temperature states are relatively less certain due to the assumed model uncertainty. Similar physically intuitive results, though in a gramian-based observability context, have been reported by Brewer et al. [2007] for a distillation process.

Sequential Sensor Removal		
	No Unmodeled Dynamics	Unmodeled Dynamics
State Rank	State Description	State Description
1	Zone 1 air temperature	Zone 5 inner floor temperature
2	Disturbance (ceiling temp.)	Zone 1 inner floor temperature
3	Disturbance (internal load)	Zone 3 inner floor temperature
4	Disturbance (corridor temp.)	Zone 4 inner floor temperature
5	Zone 5 air temperature	Zone 2 inner floor temperature
6	Zone 3 air temperature	Disturbance (ceiling temp.)
7	Disturbance (ground temp.)	Zone 5 inner wall temperature

Table 5.2: The final 7 measurements removed during the sequential removal process.

Sequential Sensor Addition		
	No Unmodeled Dynamics	Unmodeled Dynamics
Variable Rank	Variable Description	Variable Description
1	Zone 5 air temperature	Zone 5 air temperature
2	Disturbance (ceiling temp.)	Zone 1 air temperature
3	Disturbance (internal load)	Zone 3 air temperature
4	Disturbance (corridor temp.)	Zone 4 air temperature
5	Zone 1 air temperature	Zone 2 air temperature
6	Zone 3 air temperature	Zone 7 air temperature
7	Disturbance (ground temp.)	Zone 6 air temperature

Table 5.3: The first 7 measurements added during the sequential addition process.

Branch-and-bound/Semi-definite Programming

The results of the branch-and-bound/semi-definite programming strategy as applied to this example problem are unavailable due to the following complications. Firstly, re-examination of Theorem 3.3.1 shows that existence of $\mathbf{A}^{-1}$ is necessary. It may not always be a simple task to ensure the invertibility of $\mathbf{A}$ by expressing the evolution of states in terms of an invertible basis, as was performed in the low-dimensional example of Chapter 4. Luckily, the non-invertibility issue arises not from the dynamical equations describing the evolution of the internal states, but rather from the ad-hoc modeling of the temperature dynamics. It is clear from Eqs. (5.6) and (5.10), that the disturbance state transition matrix Φ is non-invertible. Re-modeling (purely for BnB/SDP purposes, not closed-loop simulation) is considered by adding a small constant to the 4^{th} row/1^{st} column element of the four Φ^i matrices given by Eq. (5.8) to render the block matrix Φ invertible. This simple procedure effectively improves the condition number of Φ, though the exact physical motivation for this is dubious.

Despite this, the dimensionality of the state-space description remains a critical factor. The necessary calculations performed at one node is prohibitively slow (taking on the order of hours to complete). This is unacceptable in this case, since several hundred nodes would be required to fully explore all of the nodes corresponding to the large number of measured variable combinations (either explicitly or implicitly).

To reduce the size of the state space, model reduction is performed by combining the two isothermal halves of each wall, floor, and ceiling, into one isothermal entity. The resulting ODEs are available in the file `makeSS_D_red2.m`. Simulation experiments are performed where the outputs of the original full-state model and newly reduced model subject to a variety of inputs are compared. A noticeable discrepancy in the outputs of the two models is observed, with a maximum temperature difference of approximately 1.0 - 1.5 $^{\circ}C$ occurring during mid-day. The length of time required for a one-node evaluation was significantly

decreased to approximately 12 minutes.

Despite this successful reduction, it was decided that the new model, with reduced number of states and modified disturbance transition matrix, is in many ways overly simplistic. Any BnB/SDP analysis arising from such a model may be of limited use in practice. Additionally, the time associated with a total enumeration via branch-and-bound to produce just *one* approximate Pareto optimal point is uncertain and potentially prohibitive for a complete analysis of the system. For these reasons, the BnB/SDP strategy is not considered as a Pareto optimum identification strategy for this example problem.

5.3.2 Simulation Experiments

Closed-loop Response of the Building System

The top portion of Figure 5.4 shows the closed-loop behaviour of zone number two's air temperature over the course of three days of operation when unmodeled dynamics are present. The controlled response under both full-state-measurement (dashed lines) and the three most important sensors as determined during the sequential addition construction of the approximate Pareto frontier (solid line) are shown. Also included are the constant upper temperature and time-varying lower temperature limits. The lower panel displays the control input at each sample time for the three-sensor case.

By design, the controller searches for the most cost effective means of maintaining the zone temperatures between their upper and lower limits. The off-peak electricity price is taken advantage of by pre-cooling the building in anticipation of higher priced mid-day electricity and peak disturbance temperatures. The model predictive controller effectively solves a thermal storage allocation problem, such as in [Zhou et al., 2005], although in a slightly different context. During mid-day, the economic incentive is to operate as close to the upper limit as possible, without violating the constraints. The benefit of additional information in the full-state-measurement case is physically realized by the controller's

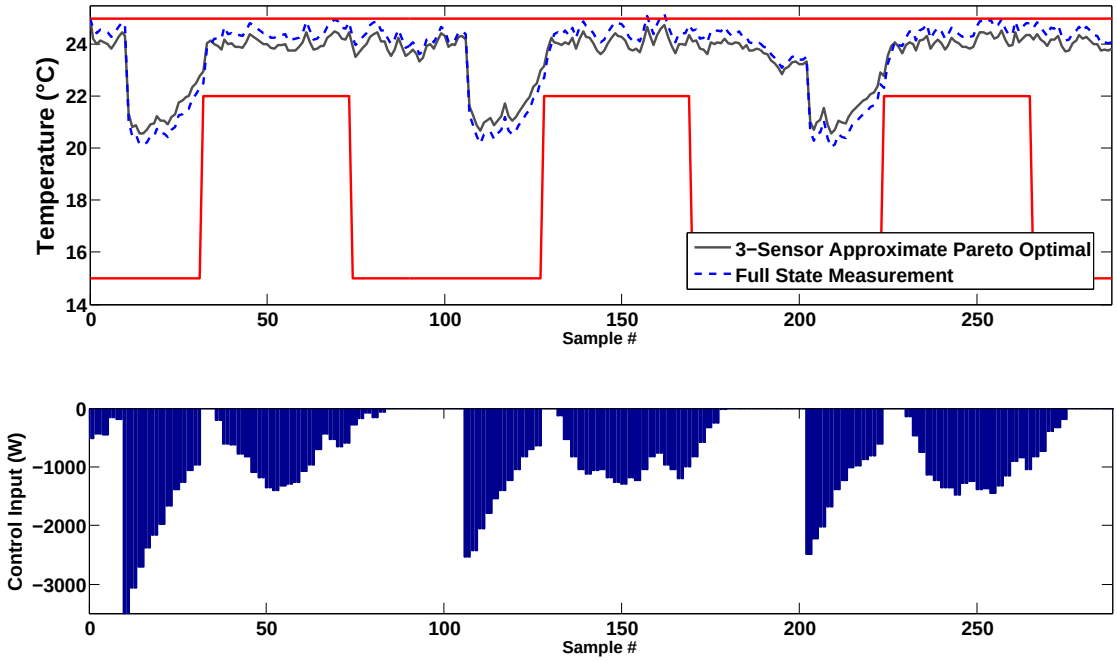

Figure 5.4: UPPER PANEL: zone number two's temperature over the course of a three day long simulation subject to $\mathcal{M}^\star$ (dashed line) and the three-sensor approximate Pareto optimal measurement system (solid line). LOWER PANEL: control inputs at each controller execution for the three-sensor approximate Pareto configuration.

ability to 1) operate closer to the upper constraint limit during mid-day and 2) achieve a lower pre-cooling temperature during off-peak hours. Hence, the availability of additional information leads to a reduction in operating cost, but this benefit is heavily out-weighed by the large capital investment associated with the measurement of every temperature node.
In all cases, the violation of the constraints on the zone temperatures remains within the allowable frequency. Inspection of the elements of the linear programming diagnostic `exitflag` shows that only a small number of simulations experienced optimization issues during execution. The cases where errors are reported correspond to the sub-optimal measurement systems in Figure 5.2 whose results are not required for the remainder of the design procedure.

Economic Analysis

Figure 5.5 shows values of J_{op} as defined by Eq. (5.18) for the best measured variable combinations found during the sequential addition/removal techniques. The decrease in operating cost follows a similar trend as the reduction in uncertainty at the low number of sensors range. This trend is reasonable given the strong relationship between information availability and closed-loop control.
The additional cost of capital investment is then added to J_{meas} to produce the total dollar costs J_{total} displayed in Figure 5.6. The point marked by "★" is the overall optimum. For this example problem, the optimum number of measured variables is three; the specific states are the air temperatures of zones 5, 1, and 3. They are, in that order, the three largest in terms of volume.

5.3.3 Summary

The design procedure of Chapter 3 clearly identified one measurement system design out of a theoretical maximum $2^{107} \approx 10^{23}$ configurations. Of course, there is no guarantee that the three-sensor configuration is the true optimum since only an approximation of the Pareto

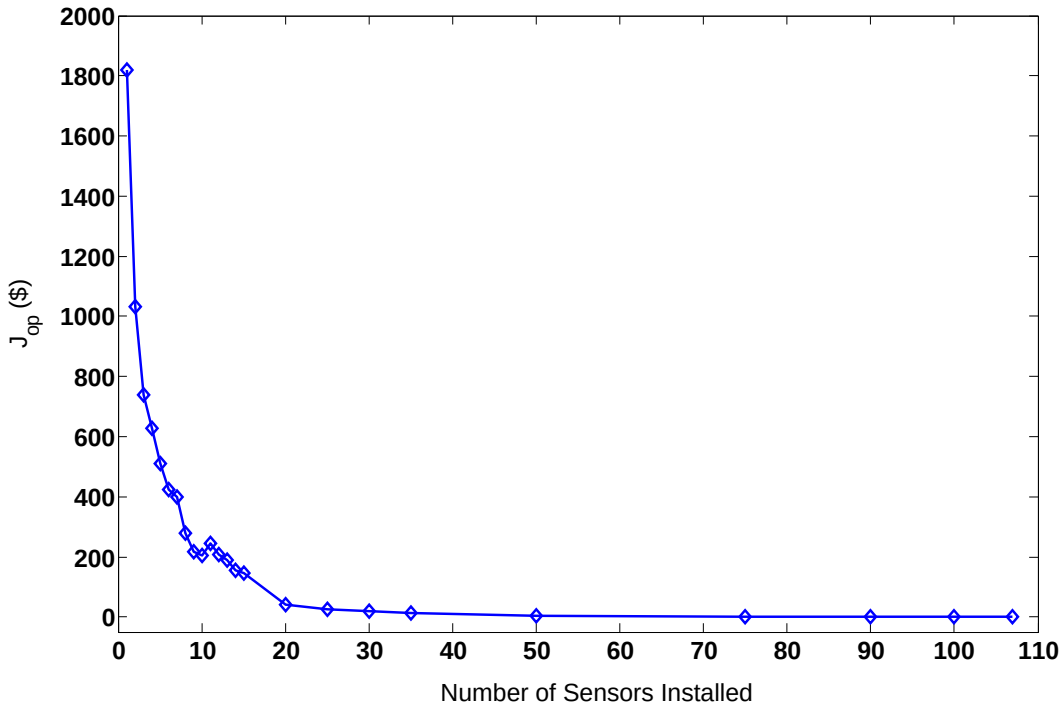

Figure 5.5: The opportunity cost of not using a full state/high precision measurement system over the course of 10 years of operation.

optimal set was possible. Despite this, the likelihood of a good approximation of the Pareto optimal set at the low number of sensors region, combined with engineering judgement gives us confidence that the three-sensor design is at least *nearly* optimal, if not *exactly*.

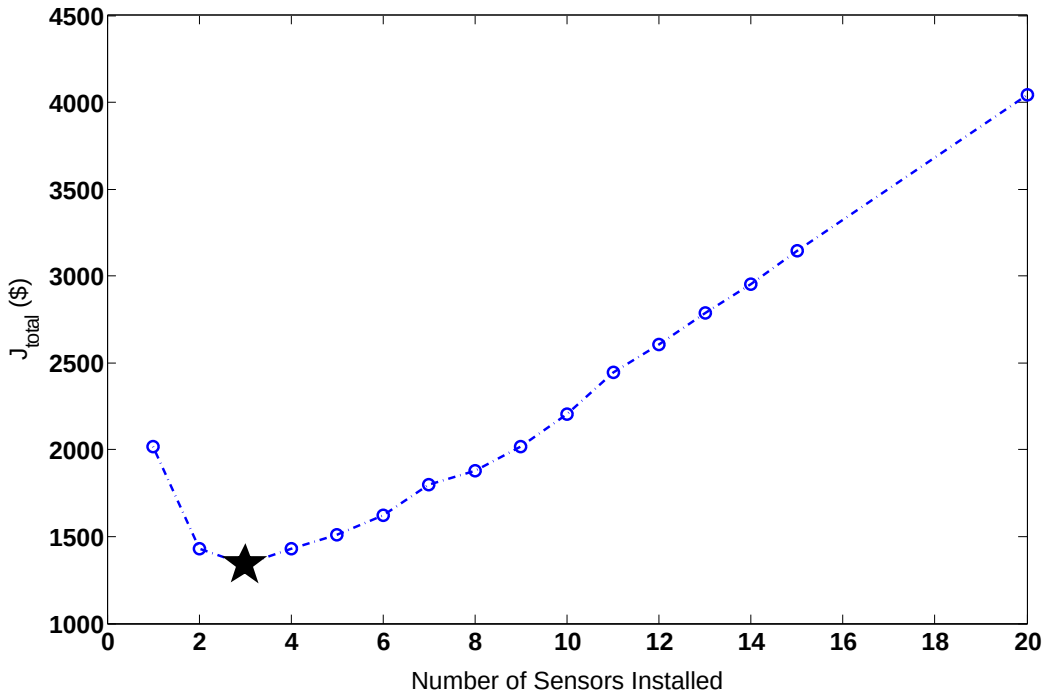

Figure 5.6: The expected total cost of various measurement systems (∘) at the low number of sensors range. The predicted optimum number of measured variables is three (★).

Chapter 6

Conclusions and Future Work

A measurement system design strategy for stochastic linear systems under the framework of a multi-objective optimization problem has been developed. The first stage – identification of the Pareto optimal set – is performed by using any of three applicable techniques: 1) exhaustive search, 2) sequential sensor addition/removal, and 3) branch-and-bound with semi-definite programming. The second stage – decision-making – utilized closed-loop simulations with a pre-specified state estimation strategy (Kalman filter) and optimal control policy (model predictive control) whose objective function and constraints formulation is designed to achieve economically optimal operation subject to user-specified constraints on the state and input variables.

The design methodology is applied to two simulation example problems: a low-dimensional fluid flow network problem and a high-dimensional thermal network model of a building. In the low-dimensional problem, all three Pareto optimal set identification strategies are applied, since the exhaustive search technique was applicable and hence a discussion regarding the quality of the sequential and BnB/SDP strategies was possible. For this example problem, the approximating sequential techniques are shown to identify measurement systems that are nearly Pareto optimal. One drawback of these strategies are their inability to identify a detailed (containing many elements) approximate Pareto optimal set. The

BnB/SDP procedure for the low-dimensional system also performs reasonably well, even though the semi-definite programming problem formulation is not exactly compatible with measurement system design's multi-objective optimization problem formulation.

The high-dimensional building system example problem limits the number of applicable strategies to determine the Pareto optimal set. Since the theoretical number of possible measurement system configurations is high (2^{107}), only the sequential and BnB/SDP strategies are believed to be applicable. The latter strategy results in severe numerical complications stemming from the size of the underlying process model. As such, its applicability to specific problems depends primarily on the computational power available. The sequential techniques yield measurement systems whose configurations corresponded to measurement importance being placed on state variables in a physically intuitive hierarchy. Since direct comparison with a known Pareto optimal set was not possible, heuristics and engineering judgement were the only tools available for commenting on the sequential strategy's effectiveness. When the process model is assumed to be reliable, measurement of the disturbance variables is assigned high importance, since the outdoor environment and building's occupants completely dictate the thermal dynamics of the interior. Conversely, when the process model is assumed to be unreliable, measurement importance is placed on the variables that are to be kept within their upper and lower limits. These logical designs indicate that the sequential techniques may be an effective means to solve the multi-objective optimization problem.

Model predictive control is assumed to be an appropriate control strategy for the high-dimensional example problem based on the results of a number of simulation-based experiments. The controller's operation results in economically optimal performance that successfully enforces constraints on the state and input variables. Another attractive feature of the model predictive controller is the minimal tuning required to ensure acceptable closed-loop performance. This relative ease is due to the low number of tuning parameters which are all physically meaningful in this case. It is recognized, however, that real *on-line*

feasibility is variable, and depends on many control system design and implementation factors. Specifically, when air handling actuators in building systems are considered, it may be difficult to identify an objective function whose minimization ensures economically optimal performance if certain hardware conditions are not satisfied. Several simulation experiments on a sample of approximate Pareto optimal measurement system configurations are performed to map the design of measurement systems to a cost of operation, which is then scaled up to the expected ten year lifespan of the hardware.

The particular set of economic design parameters used are somewhat unrealistic, but useful for illustration purposes. When economic analysis is performed with these parameters, the optimal measurement system called for the measurements of three states whose estimation error variances were deemed important to minimize. These three states correspond to the temperatures of the three largest bodies of air. Given the nature of the problem, it is safe to conclude that the resulting measurement system design is reasonable from an engineering point of view.

Economics, in terms of capital investment and future pay-back, has been the underlying theme throughout the course of the design procedure. However, there may be other factors to consider during the design of a measurement system, such as measurement redundancy in the case of unexpected sensor failures. Incorporation of these considerations is crucial to develop a robust design strategy. Nevertheless, the procedure in this thesis provides a good starting point.

An extension to nonlinear systems is also possible as long as the scalar uncertainty measures associated with relevant state estimation techniques are applicable as design parameters. One likely complication may be that the internal optimization procedures of the resulting cost-minimizing MPC strategy results in significant increases in computational demand, putting on-line and even simulation-based experimental feasibility into question.

For any real control system, the selection of measured variables and the degree of precision with which they are measured is a key design parameter. This decision is typically made

on a case-by-case basis, however, in order for firms to remain cost-competitive, process economics should be considered at every stage during a process's design phase, including the design of the measurement system. It makes sense, then, to consider the purchase of measurement equipment as an investment that provides some form of return. Measurement system design in the context of linear systems has been the starting point for this type of design philosophy, likely due to the success of linearized dynamical equations as a modeling strategy and the tractability of solutions to problems relying on their properties. However, a concrete design procedure for measurement systems where capital investment and process economics are primary concerns is still far from complete. Their design – as a field of research – remains wide open.

Bibliography

B. Wayne Bequette. *Process Control: Modeling, Design, and Simulation.* Prentice Hall International Series in the Physical and Chemical Engineering Sciences. Prentice Hall, New Jersey, 2003.

Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Słowiński. *Multiobjective Optimization: Interactive and Evolutionary Approaches.* Lecture Notes in Computer Science. Springer, Berlin, 2008.

Jeremy Brewer, Zuyi Huang, Abhay K. Singh, Manish Misra, and Juergen Hahn. Sensor Network Design via Observability Analysis and Principal Component Analyais. *Industrial & Engineering Chemistry Research*, 46:8026–8032, 2007.

Roger W. Brockett. *Finite Dimensional Linear Systems.* Decision and Control. John Wiley & Sons, Inc., New York, 1^{st} edition, 1970.

Eduardo F. Camacho and Carlos Bordons. *Model Predictive Control.* Advanced Textbooks in Control and Signal Processing. Springer, London, 2^{nd} edition, 2007.

Yunus A. Çengel. *Heat and Mass Transfer: A Practical Approach.* McGraw-Hill, 3^{rd} edition, 2006.

Donald J. Chmielewski, Tasha Palmer, and Vasilios Manousiouthakis. On the theory of optimal sensor placement. *AIChE Journal*, 48:1001–1012, 2002.

Geir E. Dullerud and Fernando Paganini. *A Course in Robust Control Theory.* Number 36 in Texts in Applied Mathematics. Springer, 1^{st} edition, 2000.

Thomas F. Edgar, David M. Himmelblau, and Leon Lasdon. *Optimization of Chemical Processes.* Chemical Engineering Series. McGraw-Hill, New York, 2^{nd} edition, 2001.

T. J. Harris, J. F. MacGregor, and J. D. Wright. Optimal sensor location with an application to a packed bed tubular reactor. *AIChE Journal*, 26:910–916, 1980.

Rudolf E. Kalman. On the general theory of control systems. In *Proceedings of the* 1^{st} *IFAC Congress*, pages 481–492, London, 1960. Butterworth's.

Hassan K. Khalil. *Nonlinear Systems.* Prentice Hall, New Jersey, 3^{rd} edition, 2002.

Huibert Kwakernaak and Raphael Sivan. *Linear Optimal Control Systems.* Wiley-Interscience, 1972.

Sanjay Lall, Jerrold E. Marsden, and Sonja Glavaški. A subspace approach to balanced truncation for model reduction of non-linear control systems. *International Journal of Robust and Nonlinear Control*, 12:519–535, 2002.

Kyoungho Lee and James E. Braun. Model-based demand-limiting control of building thermal mass. *Building and Environment*, 43:1633–1646, 2008.

Fernando V. Lima and James B. Rawlings. Nonlinear Stochastic Modeling to Improve State Estimation in Process Monitoring and Control. *AIChE Journal*, 57:996–1007, 2011.

Johan Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *IEEE International Symposium on Computer Aided Control Systems Design*, pages 284–289, Taipei, Taiwan, September 2004.

Peter J. Lute and Dolf van Paassen. Optimal indoor temperature control with the use

of a predictor. In *Proceedings of the Third IEEE Conference on Control Applications*, volume 3, pages 1843 – 1848, 1994.

D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000. Survey Paper.

D. J. Mellefont and R. W. H. Sargent. Optimal measurement policies for control purposes. *International Journal of Control*, 26:595–602, 1977.

Manfred Morari and Jay H. Lee. Model predictive control: past, present and future. *Computers & Chemical Engineering*, 23:667–682, 1999.

P. C. Müller and H. I. Weber. Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems. *Automatica*, 8:237–246, 1972.

Kenneth R. Muske and Christos Georgakis. Optimal measurement system design for chemical processes. *AIChE Journal*, 49:1488–1494, 2003.

Estanislao Musulin, Chouaib Benqlilou, Miguel J. Bagajewicz, and Luis Puigjaner. Instrumentation design based on optimal Kalman filtering. *Journal of Process Control*, 15: 629–638, 2005.

Stephen G. Nash and Ariela Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York, 1996.

Brian J. Odelson, Murali R. Rajamani, and James B. Rawlings. A new autocovariance least-squares method for estimating noise covariances. *Automatica*, 42:303–308, 2006.

Zheng O'Neill, Satish Narayanan, and Rohini Brahme. Model-based thermal load estimation in buildings. In *Proceedings of SimBuild, 4th National Conference of IBPSA-USA*, pages 474 – 481, New York, August 2010.

Imre Pólik. Addendum to the SeDuMi user guide: Version 1.1. Online, 2009. URL `http://sedumi.ie.lehigh.edu/`.

Murali R. Rajamani and James B. Rawlings. Estimation of the disturbance structure from data using semidefinite programming and optimal weighting. *Automatica*, 45:142–148, 2009.

Dan Simon. *Optimal State Estimation*. John Wiley & Sons, Inc., New Jersey, 2006.

Abhay K. Singh and Jurgen Hahn. Optimal sensor location for nonlinear dynamic systems via empirical gramians. In *Proceedings of DYCOPS*, Boston, MA, 2004.

Ralph E. Steuer. *Multiple Criteria Optimization: Theory, Computation, and Application*. Wiley Series in Probability and Mathematical Statistics - Applied. John Wiley & Sons, Inc., New York, 1986.

D. H. van Hessem and O. H. Bosgra. Closed-loop stochastic dynamic process optimization under input and state constraints. In *Proceedings of the American Control Conference*, pages 2023 – 2028, Anchorage, AK, May 2002.

Walter Waldraff, Denis Dochain, Sylvie Bourrel, and Alphonse Magnus. On the use of observability measures for sensor location in tubular reactor. *Journal of Process Control*, 8:497–505, 1998.

Shengwei Wang and Xinhua Xu. Simplified building model for transient thermal performance estimation using GA-based parameter identification. *International Journal of Thermal Sciences*, 45:419–432, 2006.

Xinhua Xu and Shengwei Wang. A simplified dynamic model for existing buildings using CTF and thermal network models. *International Journal of Thermal Sciences*, 47:1249–1262, 2008.

Jun Yan and Robert R. Bitmead. Incorporating state estimation into model predictive control and its application to network traffic control. *Automatica*, 41:595–604, 2005.

Yi Zhang and Vic I. Hanby. Model-Based Control of Renewable Energy Systems in Buildings. *HVAC&R Research*, 12(3a):739–760, 2006.

Jijun Zhou, Song Deng, Guanghua Wei, David E. Claridge, W. Dan Turner, and Oscar Contreras. Control Optimization for a Chilled Water Thermal Storage Sysem Under a Complicated Time-of-Use Electricity Rate Schedule. *ASHRAE Transactions: Research*, pages 4759–4771, 2005.

Appendix A

Derivation of Eqs. (3.14) and (3.15)

In what follows, we consider constraints on the one-step ahead value of the state x_{k+1} and a more general (polytopic) representation of the admissible set $\mathbb{X}$. Constraints on multi-step ahead values of x for an orthotopic admissible set follow in a straightforward way.

The polytope $\mathcal{P}$ defines the allowable region for future state $x_{k+1} \in \mathbb{R}^{n_x}$,

$$\mathcal{P} = \{x_{k+1} : Ax_{k+1} \leq g\}. \tag{A.1}$$

Our requirement is that the probability that $x_{k+1} \in \mathcal{P}$ must be larger than the user-specified value $1 - \alpha$

$$\Pr(x_{k+1} \in \mathcal{P}) \geq 1 - \alpha. \tag{A.2}$$

Eq. (A.2) is equivalent to the following integral constraint:

$$\frac{1}{\sqrt{(2\pi)^{n_x}\det(Z)}} \int_{\mathcal{P}} e^{-\frac{1}{2}(x_{k+1}-\hat{x}_{k+1|k})^T Z^{-1}(x_{k+1}-\hat{x}_{k+1|k})} dx_{k+1} \geq 1 - \alpha. \tag{A.3}$$

where $\hat{x}_{k+1|k}$ is the conditional expectation of x_{k+1} given all available information up to and including sample time k.
Define the ellipsoid $\mathcal{E}_r$:

$$\mathcal{E}_r = \{x_{k+1} : \tilde{x}_{k+1|k}^T Z^{-1} \tilde{x}_{k+1|k} \leq r^2\} \tag{A.4}$$

where $\tilde{x}_{k+1|k} = \left(x_{k+1} - \hat{x}_{k+1|k}\right)$. Since the scalar variable $(\tilde{x}_{k+1|k}^T Z^{-1} \tilde{x}_{k+1|k})$ has χ^2 distribution with n_x d.o.f., the constraint

$$\Pr(\tilde{x}_{k+1|k}^T Z^{-1} \tilde{x}_{k+1|k} \leq r^2) = 1 - \alpha, \tag{A.5}$$

is equivalent to

$$\frac{1}{2^{n_x/2}\Gamma(n_x/2)} \int_0^{r^2} \chi^{n_x/2-1} e^{-\chi/2} d\chi = 1 - \alpha. \tag{A.6}$$

Hence, through proper specification of r^2 it is possible to ensure that $\Pr\left(x_{k+1} \in \mathcal{E}_r\right) = 1 - \alpha$. With the size and shape of $\mathcal{E}_r$ now known, we wish to establish conditions on the ellipsoid's center $\hat{x}_{k+1|k}$ to guarantee that $\mathcal{E}_r \subset \mathcal{P}$ and therefore that inequality (A.2) holds.
Note that the ellipsoid $\mathcal{E}_r$ can also be defined by

$$\mathcal{E}_r = \{x_{k+1} : x_{k+1} = \hat{x}_{k+1|k} + r \cdot Z^{1/2} s,\ \|s\| \leq 1\}, \tag{A.7}$$

with $\|s\| = 1$ defining the edge. From the definition of the constraints (A.1), the ellipsoid $\mathcal{E}_r$ is contained entirely in $\mathcal{P}$ if and only if:

$$\sup_{\|s\|=1} \left\{ a_i^T (\hat{x}_{k+1|k} + r \cdot Z^{1/2} s) \right\} \leq g_i \tag{A.8}$$

for all i where a_i^T is the i^{th} row of A, and g_i is the i^{th} entry of g. To find this supremal element, we use the method of Lagrange multipliers:

$$\begin{aligned} \max f(s) &= r \cdot a_i^T Z^{1/2} s \\ \text{s.t. } g(s) &= c \end{aligned}$$

where $g(s) = s^T s = 1$ since the solution will occur on the edge of the ellipsoid.
Define the Lagrange function

$$L(s, \lambda) = f(s) + \lambda \cdot (g(s) - c).$$

Solving $\nabla_{s,\lambda} L(s, \lambda) = 0$ we obtain:

$$\frac{\partial L}{\partial s} = r \cdot Z^{1/2} a_i + 2\lambda s, \qquad \frac{\partial L}{\partial \lambda} = s^T s - 1.$$

We then have

$$s = \frac{-r}{2\lambda} Z^{1/2} a_i, \qquad s^T s = 1, \qquad \lambda = r \cdot \frac{\pm \|Z^{1/2} a_i\|}{2},$$

for which

$$s = \frac{Z^{1/2} a_i}{\|Z^{1/2} a_i\|}$$

is the maximizer. Substituting the above expression into Eq. (A.8), we arrive at the

following condition on $\hat{x}_{k+1|k}$:

$$a_i^T \hat{x}_{k+1|k} + \frac{r \cdot a_i^T Z^{1/2} Z^{1/2} a_i}{\|Z^{1/2} a_i\|} \leq g_i,$$

which is equivalent to the result reported by van Hessem and Bosgra [2002]:

$$a_i^T \hat{x}_{k+1|k} + r \cdot \sqrt{a_i^T Z a_i} \leq g_i.$$