PARAMETER ROBUST REDUCED-ORDER CONTROL OF FLEXIBLE STRUCTURES

by

Stephen H. Jones

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Electrical Engineering

APPROVED:

William T. Baumann, Chairman

Hugh F. VanLandingham

Harry H. Robertshaw

John S. Bay

Harley H. Cudney

August, 1991

Blacksburg, Virginia

PARAMETER ROBUST REDUCED-ORDER CONTROL OF FLEXIBLE STRUCTURES

by

Stephen H. Jones

Department of Electrical Engineering

William T. Baumann, Chairman

(ABSTRACT)

This thesis generalizes the concept of internal feedback loop modeling, due to Tahk and Speyer, to arrive at two new LQG-based methods of parameter robust control. One component of the robustness procedure, common to both methods, is the application of an auxiliary cost functional penalty to desensitize the system to variations in selected parameters of the state-space model. The other component consists of the formulation of a fictitious noise model to accommodate the effect of these parameter variations.

The "frequency-domain method" utilizes knowledge of the system dynamics to create a frequency-shaped noise model with a power spectrum that approximates the frequency content of unknown error signals in the system due to parameter uncertainties. This design method requires augmentation of additional dynamics to the plant, which results in higher-dimensional full-order controllers. However, the controller design computations are identical to those of a standard LQG problem.

The "time-domain method" emulates the same error signals by means of a multiplicative white noise model which reflects the time-domain behavior of those signals. The resulting robust controller is of the same order as the standard LQG controller, although the design involves a more complex computational algorithm. The application of multiplicative white noise to the system model requires the solution of a system of four coupled equations — two modified Riccati equations and two modified Lyapunov equations.

In addition, the optimal projection equations are applied to both robustness methods to reduce the controller order with minimal loss in performance.

Comparisons are drawn between these and related robust control methods, and it is shown that the relative effectiveness of such methods is problem dependent. Parameter sensitivity analysis is carried out on a simply supported

plate model subject to external disturbances. The appropriate robust controller is selected, and it is found to stabilize the plate with little sacrifice in performance.

Acknowledgements

A few people contributed to make this work much more manageable than it otherwise would have been.

Dr. William T. Baumann gave direction to my research and contributed a number of ideas. He consistently provided valuable technical assistance, sometimes producing methods of attacking a difficult problem before I could finish posing the question.

Dr. Hugh F. VanLandingham secured funding throughout the course of this study. The commencement of his funding was particularly timely, as it occurred shortly after receiving word from the graduate school that I would be denied consideration for all of their scholarships due to my race.

Dr. Harry H. Robertshaw was generous in granting me access to his laboratory, where the principles studied in this thesis could be tested and verified on a hardware experiment.

Loretta Estes, my graduate counselor, repeatedly found creative ways through the massive bureaucratic jungle we have come to know as a "university" and always did so with a smile. Without her encyclopedic knowledge of the many rules and regulations, her kindness, and her sanity in this academic environment, the task of pursuing a Ph.D. would be fraught with despair.

Deserving a special mention is Dr. Layne T. Watson, without whose help I was entirely able to design homotopy algorithms on my own.

Table of Contents

List of Figures

List of Tables

0. Introduction

0.1 Purpose and Importance of This Work

The purpose of this research is to explore two new and related methods of designing compensators for the active control of uncertain systems. Both methods are designed to improve the robustness of a control system's stability and performance with respect to uncertainty in selected parameters of the state-space model of the plant. They accomplish this by modeling the effect of uncertainties as accurately as possible by fictitious noise sources while simultaneously reducing sensitivity by means of additional penalties in the cost functional. The first method constructs the noise model in the frequency domain, whereas the second formulates noise in the time domain. In addition, both incorporate optimal order reduction directly into the design. These methods are expected to have general application to plants with uncertain state-space parameters; however, they will be applied in this work to the problems involved in the suppression of disturbances in flexible structures, which suffer particularly from the difficulties of large plant uncertainty and large compensator order.

Modeling and identification of flexible structures, themselves, constitute a difficult problem and an active area of research (e.g., Balas and Doyle 1990). The prevalence of this type of research is an indicator that linear, finite-dimensional, time-invariant (LFDTI) models of flexible structures are examples of highly uncertain plants, and it points to the need for robust controllers for such systems.

Also, LFDTI models of flexible structures tend to be of very high order. Firstly, since flexible structures are infinite-dimensional, it is desirable to consider as many vibrational modes as possible in order to reduce the effect of unmodeled dynamics on the control system. Secondly, it is frequently necessary to augment the plant with additional dynamics in order to model disturbances more accurately or to meet more precise performance specifications. Disturbance modeling (e.g., Kwakernaak and Sivan 1972, sec. 1.11.4) is desirable when the disturbances are correlated and something is known about their frequency content. This allows the exogenous signals to be modeled by white noise processes (in the case of LQG-based designs), provided the disturbance dynamics are appended to the plant at the disturbance inputs. Likewise, frequency-shaped cost functionals (Gupta 1980) are used when the state and control weighting matrices are functions of frequency. This is accomplished in the LQG framework by

appending dynamics to the plant at the controlled-variable outputs. It has been shown (Sievers and von Flotow 1989) that these two extensions to LQG theory are duals and, in the single-input single-output (SISO) case, equivalents of one another. Also, the same two methods are used for analogous purposes in $\mathcal{H}_\infty$-based design (e.g., Doyle 1984).

In either case, an LQG or $\mathcal{H}_\infty$-optimal compensator, being of the same order as the augmented plant, must also be of high order. As a result, full-order controllers tend to be of very high order and therefore place a great computational burden on the real-time processor hardware.

0.2 Research Objectives

The proposed parameter robust reduced-order control methods are described in Chapter 1 and developed in Chapters 2 through 5. Then, the resulting designs are applied to a continuous-time FDLTI model of a simply supported rectangular plate (in Chapter 6), as well as to the actual hardware (in Chapter 7). The control system has one control input actuator, one disturbance input actuator, and twelve accelerometer sensors. The state-space model of the plate has been derived from modal frequency and modeshape data obtained by an identification procedure. This model is augmented by the dynamics associated with a contol signal smoothing filter and a noise shaping filter designed to reflect the characteristics of a colored noise disturbance. Analysis has shown the stability of the closed-loop system to be most susceptible to errors in the natural frequencies and control input modeshape vector, and experimentation has confirmed this. Therefore, emphasis is placed on making the system less sensitive to errors in the parameters of the state-space model corresponding to these quantities.

There are three main objectives to this research. The first objective is to evaluate the efficiency of the performance/stability robustness tradeoff for the two proposed design methods. This is done by comparing compensators designed by the frequency-domain and time-domain noise modeling techniques (of Chapters 2 and 3, respectively) with those designed by two existing methods discussed in Chapter 1: LQG/PRE and multiplicative white noise modeling (without the auxiliary output modeling feature used in Chapter 3). First, a parameter range is specified for which the system must be stable. This parameter range is larger than the stability range obtained

by the standard LQG design. Second, full-order compensators are designed by all four techniques to just meet this stability robustness specification, using the same performance criterion and assumed exogenous noise covariances used by the LQG design. Finally, it is determined how much performance was sacrificed to attain this level of robustness by measuring the performance of the designs (i.e., the quadratic cost) over the stability range of interest. This method is applied to a simple 1-mode model of the plate in Chapter 6 to illustrate the design considerations involved. Uncertainties in both the control input modeshape vector (i.e., the controller actuator location) and the natural frequency are considered. In Chapter 7, a 4-mode model of the plate is analyzed to determine the cause of the poor stability properties of the actual hardware. Then the most suitable robust controller is selected and implemented on the plate.

The second objective is to investigate the extent to which the order of a standard LQG compensator can be reduced for FDLTI models of the rectangular plate. When one adds the states of a second-order smoothing filter and a noise shaping filter to the two states per vibrational mode of the plate, the high order of the plant becomes apparent, as does the need for a reduced-order compensator. The procedure is as follows. First, a full-order LQG compensator is designed to provide satisfactory nominal performance. Then reduced-order compensators are designed using the optimal projection equations, and it is determined how far the order can be reduced without significant degradation in performance. This is carried out in Chapter 6, and the results of implementation in hardware are described in Chapter 7.

The third and final objective is to combine the best modified LQG/PRE design method with the optimal projection equations to determine the overall merit of the new robust, minimal order design. This work is carried out in Chapter 7.

1. Motivation for This Study

The ideas developed in this work were motivated by the shortcomings of existing $\mathcal{H}_2$- and $\mathcal{H}_\infty$-based techniques for rejecting disturbances in uncertain systems. In this chapter, standards are defined for the practical merit of a controller, and two approaches are proposed to better meet those standards than do existing methods. Then the specific contributions of this work are outlined.

1.1 Criteria for the Merit of Candidate Designs

The standards that will be used to judge and compare the effectiveness of various compensator designs are: 1) nominal disturbance rejection performance; 2) "efficiency" of tradeoff between performance and stability robustness; 3) compensator order; and 4) ability to reject both steady state and transient disturbances. What follows is a discussion of these criteria and the means by which they will be met.

A. Nominal Performance

The *performance* of a compensator will be measured by its ability to minimize the effect of unwanted exogenous signals on the prescribed controlled-variable vector, whose elements consist of weighted linear combinations of system states and controls. *Nominal performance* refers to the performance of a control system with all plant parameters equal to their assumed values. Two competing classes of performance optimality have been considered, LQG and $\mathcal{H}_\infty$. These two definitions of optimality differ in the way they model the exogenous signals and in the way they measure the magnitude of the resulting disturbance in the controlled variables.

In the traditional LQG formulation of the disturbance rejection problem, optimal performance is defined to be that which minimizes the cost functional

$$J = \lim_{T\to\infty} \frac{1}{2T} \mathbb{E} \left\{ \int_{-T}^{T} \left[x^{\mathrm{T}}(t) Q x(t) + u^{\mathrm{T}}(t) R u(t) \right] dt \right\} \tag{1.1}$$

in the presence of unit intensity, uncorrelated white process and sensor noise (v and n, respectively). For the purpose of easier comparison with $\mathcal{H}_\infty$-based methods, an equivalent formulation will be made in the frequency domain. Define the noise vector, w, and the controlled-variable vector, z, as follows:

$$w = \begin{bmatrix} v \\ n \end{bmatrix}, \quad z = \begin{bmatrix} Q^{1/2}x \\ R^{1/2}u \end{bmatrix} \tag{1.2}$$

and denote the closed-loop transfer function matrix from w to z by $H(s)$. Then the $\mathcal{H}_2$-optimal compensator is that which minimizes the $\mathcal{H}_2$ norm of $H(s)$ (i.e., $\|H(s)\|_2$). The fact that the LQG and $\mathcal{H}_2$ definitions of optimal performance are equivalent are an immediate consequence of the following theorem:

Theorem. $J = \|H(s)\|_2^2$.

Proof. Using the subscript "T", define time-truncated functions as follows:

$$\begin{aligned} x_T(t) &\triangleq x(t), \quad -T \le t \le T; \\ &\triangleq 0, \qquad \text{otherwise} \end{aligned}$$

Then

$$\begin{aligned} J &\triangleq \lim_{T\to\infty} \tfrac{1}{2T}\, \mathbf{E}\left\{ \int_{-T}^{T} \left[x^{\mathrm{T}}(t)Qx(t) + u^{\mathrm{T}}(t)Ru(t)\right] dt \right\} \\ &= \lim_{T\to\infty} \tfrac{1}{2T}\, \mathbf{E}\left\{ \int_{-T}^{T} z^{\mathrm{T}}(t)z(t)dt \right\} \\ &= \lim_{T\to\infty} \tfrac{1}{2T}\, \mathbf{E}\left\{ \int_{-\infty}^{\infty} {z_T}^{\mathrm{T}}(t)z_T(t)dt \right\} \\ &= \lim_{T\to\infty} \tfrac{1}{2T}\, \mathbf{E}\left\{ \tfrac{1}{2\pi} \int_{-\infty}^{\infty} {z_T}^*(j\omega)z_T(j\omega)d\omega \right\} \\ &= \lim_{T\to\infty} \tfrac{1}{2T}\, \mathbf{E}\left\{ \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{trace}\left[{w_T}^*(j\omega)H^*(j\omega)H(j\omega)w_T(j\omega)\right] d\omega \right\} \\ &= \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{trace}\left[H(j\omega) \lim_{T\to\infty} \left\{ \frac{\mathbf{E}\left[w_T(j\omega){w_T}^*(j\omega)\right]}{2T} \right\} H^*(j\omega) \right] d\omega \\ &= \tfrac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{trace}\left[H^*(j\omega)H(j\omega)\right] d\omega \\ &= \|H(s)\|_2^2. \end{aligned}$$

In the proof of the theorem, use was made of the definition of the $\mathcal{H}_2$ operator norm (e.g., Dailey 1990), Parseval's Theorem, and the fact that the covariance of the

assumed disturbance vector, w, is the identity matrix.

In $\mathcal{H}_\infty$ controller design, the same closed-loop transfer function is considered, and optimality of performance is achieved when its $\mathcal{H}_\infty$ norm (i.e., $\|H(s)\|_\infty$) is minimized. However, the disturbances, w, and the controlled variables, z, are viewed differently. The $\mathcal{H}_\infty$ operator norm is the induced L_2 norm, meaning

$$\|H(s)\|_\infty = \sup_{\|v(t)\|_2 = 1} \|z(t)\|_2 \tag{1.3}$$

Therefore, w and z are considered to be L_2 (i.e., square integrable) signals. This expression for the $\mathcal{H}_\infty$ norm reveals a possible drawback to $\mathcal{H}_\infty$-based control methods for disturbance rejection. Since the $\mathcal{H}_\infty$ controller is designed for the worst case unity-norm L_2 disturbance, no consideration is given to the relative intensities of multiple noise sources. The worst case disturbance may be one whose entire power is delivered to a single process noise or sensor noise port. LQG-based methods, on the other hand, model the noise intensity at each port separately.

Another expression for the $\mathcal{H}_\infty$ norm is given by its definition,

$$\|H(s)\|_\infty \triangleq \sup_\omega \bar{\sigma}[H(j\omega)] \tag{1.4}$$

where $\bar{\sigma}$ denotes the maximum singular value. This expression shows that the $\mathcal{H}_\infty$-optimal compensator is the one which minimizes the *worst case* disturbance rejection over all frequencies. In contrast, the LQG compensator provides better disturbance rejection *averaged* over all frequencies. This principle is illustrated in the singular value plot of a closed-loop vibrational system in Figure 1.1. Using the same state and control weightings, Q and R, an LQG and an $\mathcal{H}_\infty$ compensator were designed to reject disturbances in an eight-mode model of a flexible structure. The two curves indicate the maximum singular value of the transfer function matrix, $H(s)$, for the two different compensator designs. The $\mathcal{H}_\infty$ compensator provided slightly better worst case performance, but at the cost of significantly worse performance averaged over the entire frequency range.

Because of these two potential problems associated with disturbance rejection in $\mathcal{H}_\infty$-based designs (i.e., less precise noise modeling and less practical definition of

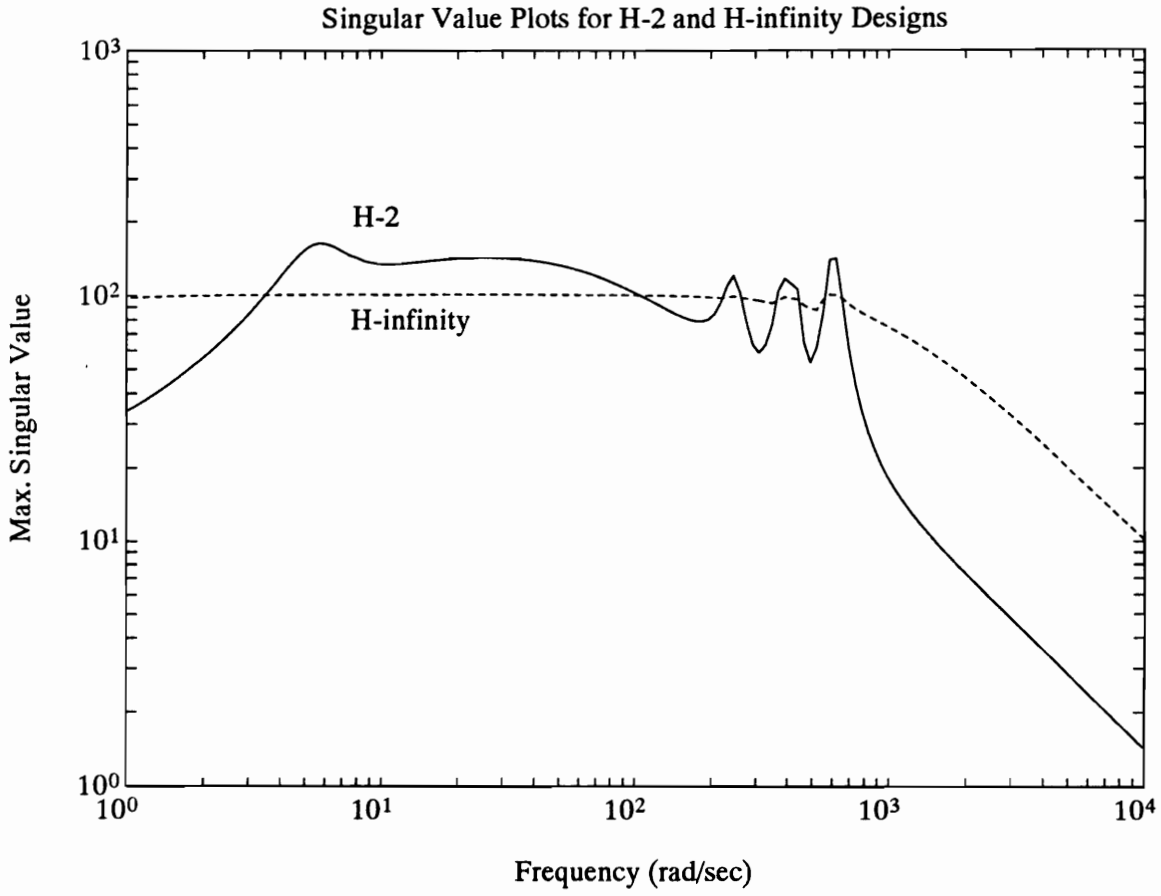

Figure 1.1: Nominal Performance for LQG vs. $\mathcal{H}_\infty$-Optimal Control

optimality) preferential consideration was given to improving LQG-based design methods in the expectation that they would, by conventional measures, provide significantly better nominal performance.

B. Performance/Stability Robustness Tradeoff

For uncertain systems, $\mathcal{H}_\infty$-based methods, such as μ-synthesis (Doyle 1985), have the advantage that plant uncertainties with singular value bounds can be incorporated directly into the design, without the need for fictitious noise sources to model the uncertainties. However, these methods also have serious limitations, as illustrated in their recent application to a flight control problem (Doyle, Lenz, and Packard 1987). Firstly, $\mathcal{H}_\infty$ controllers do not provide optimal performance and stability robustness simultaneously. For example, an $\mathcal{H}_\infty$ controller designed for performance may yield very poor stability margins. Secondly, the μ-synthesis technique corrects this problem by attempting an optimal tradeoff between performance and stability robustness, but the resulting controllers tend to be of very high order. Therefore, ad hoc controller and/or model reduction schemes are required to reduce the controller order to a practical size.

An LQG-based method can avoid these problems, provided that its noise models adequately represent the uncertainties without greatly increasing the order of the compensator. For this reason, the main thrust of this work is to improve on an existing LQG-based robust control method by more accurately modeling the plant uncertainties by means of more suitable fictitious noise sources. The purpose of doing so is to give the resulting compensators the same degree of stability robustness with a smaller sacrifice in performance. The degree to which this occurs will be referred to as the "efficiency" of the performance/stability robustness tradeoff.

Tahk and Speyer (1987) introduced a modified LQG synthesis procedure, later called parameter robust LQG (PRLQG), in which perturbations in the elements of the A-matrix are modeled by a fictitious internal feedback loop. Given a nominal system matrix, A, and the perturbed matrix, $\hat{A}$, define the perturbation matrix by $\Delta A = \hat{A} - A$. Then factor it as follows: $\Delta A = MLN$, where M and N are of full rank (to minimize their dimensions) and L is without loss of generality a diagonal matrix, whose diagonal elements reflect the (unknown) magnitude of the perturbations. An accurate model of the perturbed system results after the addition of this internal

feedback loop, as shown in Figure 1.2. Now, given the nominal system,

$$\begin{aligned} \dot{x} &= Ax + Bu + G_1 w \\ y &= Cx + G_2 w \end{aligned} \tag{1.5}$$

the perturbed system is described by

$$\begin{aligned} \dot{x} &= Ax + Bu + G_1 w + M w_a \\ y &= Cx + G_2 w \\ z_a &= Nx \\ w_a &= L z_a \end{aligned} \tag{1.6}$$

where w_a and z_a are auxiliary input and output variables, respectively. Intuitively, one may suspect that the robustness of the system to the prescribed parameter variations would be enhanced if w_a were modeled as white noise, thereby adding $\mu M M^{\mathrm{T}}$ (for some scalar μ) to the process noise covariance. Actually, Tahk and Speyer showed that under certain minimum phase and similarity conditions, the robustness of an LQ regulator is recovered asymptotically as $\mu \to \infty$. Note that this robustness recovery technique is identical to LQG/LTR for uncertainties at the input (Doyle and Stein 1981), provided $M = B$. However, the PRLQG method has the advantage that M contains partial information as to the structure of the perturbation, ΔA. In fact, Tahk and Speyer showed that LQG/LTR fails to asymptotically desensitize the estimator (when PRLQG succeeds) for certain structures of ΔA, since B does not necessarily pass a crucial similarity condition with respect to the matrix ΔA, namely that $\mathrm{span}\{M\} \subset \mathrm{span}\{B\}$ for some MLN-factorization of ΔA. Vibration control examples have demonstrated that in such cases LQG/LTR may provide relatively poor robustness to parameter uncertainties (Tahk and Speyer 1989). PRLQG also includes the dual of the above procedure — analogous to LQG/LTR for uncertainties at the output. In the dual procedure, z_a is treated as an auxiliary controlled variable, and the term $\rho N^{\mathrm{T}} N$ (for some scalar ρ) is added to the state weighting matrix in the regulator design. This procedure also makes sense intuitively, as it seems reasonable to penalize the auxiliary output variable in the cost functional in order to minimize the magnitude of the auxiliary inputs that result. Combining the two procedures asymptotically yields "absolute robustness" (i.e., stability robustness to parameter uncertainties of arbitrary

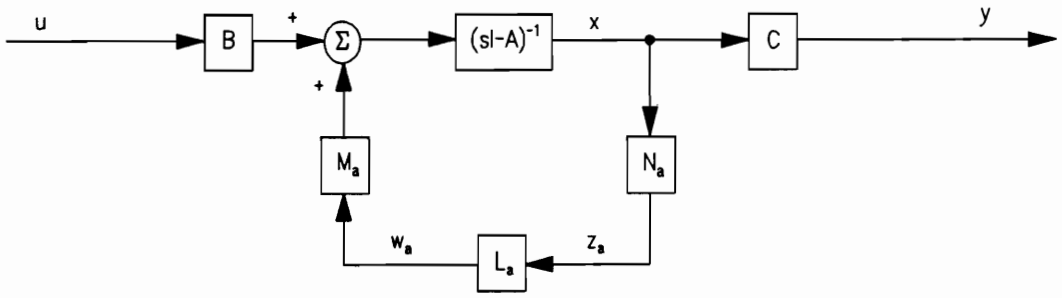

Figure 1.2: PRLQG Error Model

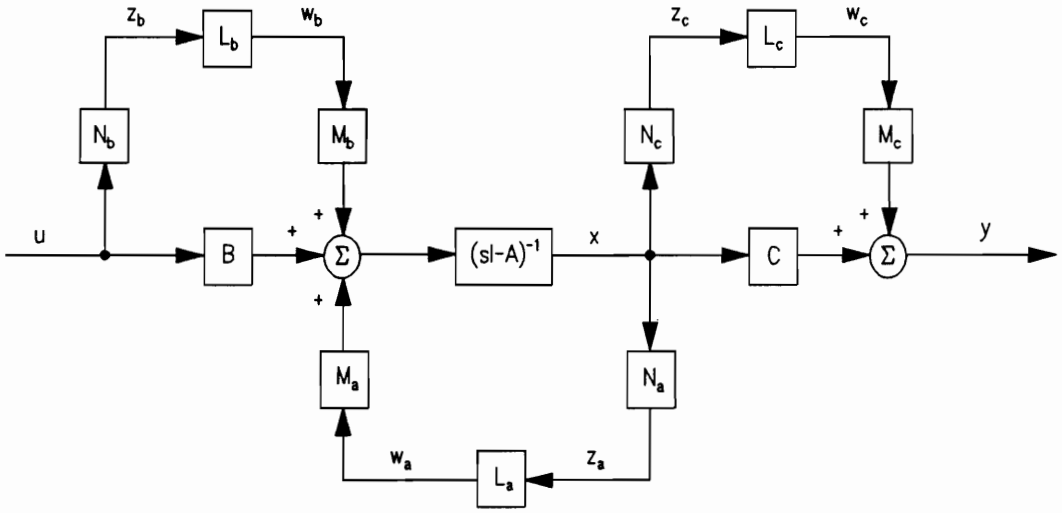

Figure 1.3: LQG/PRE Error Model

magnitude) as μ, $\rho \to \infty$.

Lin (1989) studied the combined effect of both procedures in more detail and also extended the method to include parameter variations in the B- and C-matrices in a way that is less cumbersome than that suggested by Tahk and Speyer. This extended method was termed LQG/PRE (standing for Linear Quadratic Gaussian design with Parameter Robustness Enhancement). The resulting model of the plant, with perturbations in A, B, and C is shown in Figure 1.3. The system equations are:

$$\begin{aligned} \dot{x} &= Ax + Bu + G_1 w + M_a w_a + M_b w_b \\ y &= Cx + G_2 w + M_c w_c \\ z_a &= N_a x, \quad z_b = N_b u, \quad z_c = N_c x \\ w_a &= L_a z_a, \quad w_b = L_b z_b, \quad w_c = L_c z_c \end{aligned} \tag{1.7}$$

where

$$\Delta A \triangleq M_a L_a N_a, \quad \Delta B \triangleq M_b L_b N_b, \quad \Delta C \triangleq M_c L_c N_c \tag{1.8}$$

Given the nominal regulator weighting matrices, Q and R, and the nominal Kalman filter noise covariances, Q_f and R_f, this leads to the following modified LQG design matrices (denoted with carats):

$$\begin{aligned} \hat{Q} &= Q + \rho_a {N_a}^{\mathrm{T}} N_a + \rho_c {N_c}^{\mathrm{T}} N_c \\ \hat{R} &= R + \rho_b {N_b}^{\mathrm{T}} N_b \\ \hat{Q}_f &= Q_f + \mu_a M_a {M_a}^{\mathrm{T}} + \mu_b M_b {M_b}^{\mathrm{T}} \\ \hat{R}_f &= R_f + \mu_c M_c {M_c}^{\mathrm{T}} \end{aligned} \tag{1.9}$$

The main weakness of the LQG/PRE scheme seems to be the fact that it models the auxiliary input variables as white noise. In order to reflect reality, these fictitious input signals should generally be modeled quite differently. For example, suppose there is a single uncertain, but constant, parameter in A (i.e., w_a is scalar, and ΔA is constant). Then a true model of the perturbed system would require w_a to be proportional to one of the states, say x_i (through $w_a = L_a N_a x$). Since the sign and magnitude of the perturbation is unknown (i.e., L_a unknown), w_a cannot be modeled faithfully; however, intuitively, it seems its dynamics should more closely resemble the dynamic behavior of x_i rather than that of a constant-intensity white noise variable.

Considerable research has already been conducted that will be applied to constructing a more meaningful model of the auxiliary inputs. This research was done

outside the LQG/PRE-framework, and did not include the related problem of penalizing auxiliary outputs. Using stochastic differential equation theory, Wonham (1967, 1968) developed modified Riccati equations for both LQ regulator and Kalman filter design for systems with "state-dependent noise". State-dependent noise arises in systems with state equations of the form

$$\dot{x} = Ax + Bu + G(x)w \tag{1.10}$$

where $G(x)$ is some function of the state, and w is a (vector) white noise process. In the above example, we would model w_a as state-dependent noise, which would make it white noise with time-varying intensity proportional to x_i. Equivalently, one could conceptualize ΔA as a matrix whose elements are white noise processes.

Hyland (Hyland and Madiwale 1981, Hyland 1982), motivated by the maximum entropy principle and the concept of ΔA as a matrix of white noise processes, modeled uncertainties in the A-matrix by adding state-dependent noise to the plant model. In doing so, he developed the equations for the full LQG compensator for systems with state-dependent noise. For such systems, the separation principle no longer holds. The regulator and filter design require the solution not of two uncoupled Riccati equations, but of four coupled equations — two modified Riccati equations and two modified Lyapunov equations. Later, Bernstein and Hyland (1988a) extended this result to apply to systems with state-, control-, and measurement-dependent noise, effectively allowing parameter perturbations in all three system matrices (ΔA, ΔB, and ΔC, respectively) to be modeled as white noise processes. The design process is computationally more difficult than that of simple LQG in that it requires the iterative solution of four coupled matrix equations, but notably the optimal compensator is of the same order as the plant. That is, adding state-, control-, and measurement-dependent noise (henceforth referred to collectively as *multiplicative white noise*) to a system involves no increase in compensator order.

As stated at the beginning of this section, this research involves modeling fictitious noise sources more accurately in an LQG-type problem. Specifically, the LQG/PRE approach will be adopted, with modifications that are intended to improve the efficiency of the performance/stability robustness tradeoff by remodeling the auxiliary input signals such that they more closely reflect reality. Two different fictitious noise models will be investigated, one motivated by consideration of the time-

domain knowledge of the auxiliary input signals, and the other by frequency-domain knowledge.

The time-domain method consists of using multiplicative white noise in place of simple white noise to model the auxiliary input signals. Recalling the simple example mentioned above, the motivation for this approach is illustrated in Figure 1.4. These plots were derived from a two-mode simulation of the rectangular plate apparatus described in section 1.3. A true constant parameter uncertainty in the A-matrix would be compensated for optimally if w_a were modeled perfectly (i.e., $w_a = kx_i$, where the constant k is determined by the sign and magnitude of the parameter variation), as shown in Figure 1.4*a*. However, since k is unknown, this model is impossible to implement. The LQG/PRE method uses constant-intensity white noise to model w_a (say, $w_a = w$), as in Figure 1.4*b*. But the multiplicative white noise approach will use state-dependent noise to take advantage of the fluctuations in x_i (i.e., $w_a = x_i w$), as shown in Figure 1.4*c*. Presumably, the extent to which this noise model improves the LQG/PRE design will depend on the degree to which the amplitude of x_i varies over time. In particular, the multiplicative noise model should provide a greater improvement for transient noise rejection than for steady-state (i.e., constant-intensity) noise rejection. This modification of LQG/PRE requires solution of the coupled Riccati and Lyapunov equations, but does not result in an increase in compensator order.

The frequency-domain method involves an attempt to match the true auxiliary inputs in the frequency domain. Time-domain matching suffers from the fact that the sign of the parameter perturbations is unknown. As a result, a very broad band signal is used in an attempt to cover both possibilities at every time instant. However, our knowledge of the *frequency* content of the auxiliary inputs is limited only by our knowledge of the disturbance's frequency content and of the closed-loop system dynamics. So, given an assumed disturbance model, the true auxiliary inputs can be modeled fairly accurately in the frequency domain by closing the loop with an LQG compensator which gives the desired nominal performance and measuring the frequency response of w_a, w_b, and w_c (see Figure 1.3). Therefore, the frequency-domain behavior of the auxiliary inputs will be approximated by passing fictitious white noise signals through suitable frequency-shaping filters. The effect is demonstrated in the frequency spectrum plots of Figure 1.5, generated from the same data used in Figure 1.4. The frequency-shaped noise model, produced by means of a second-order filter, provides a much better frequency-domain replica of the true signal ($w_a = kx_i$) than does a white

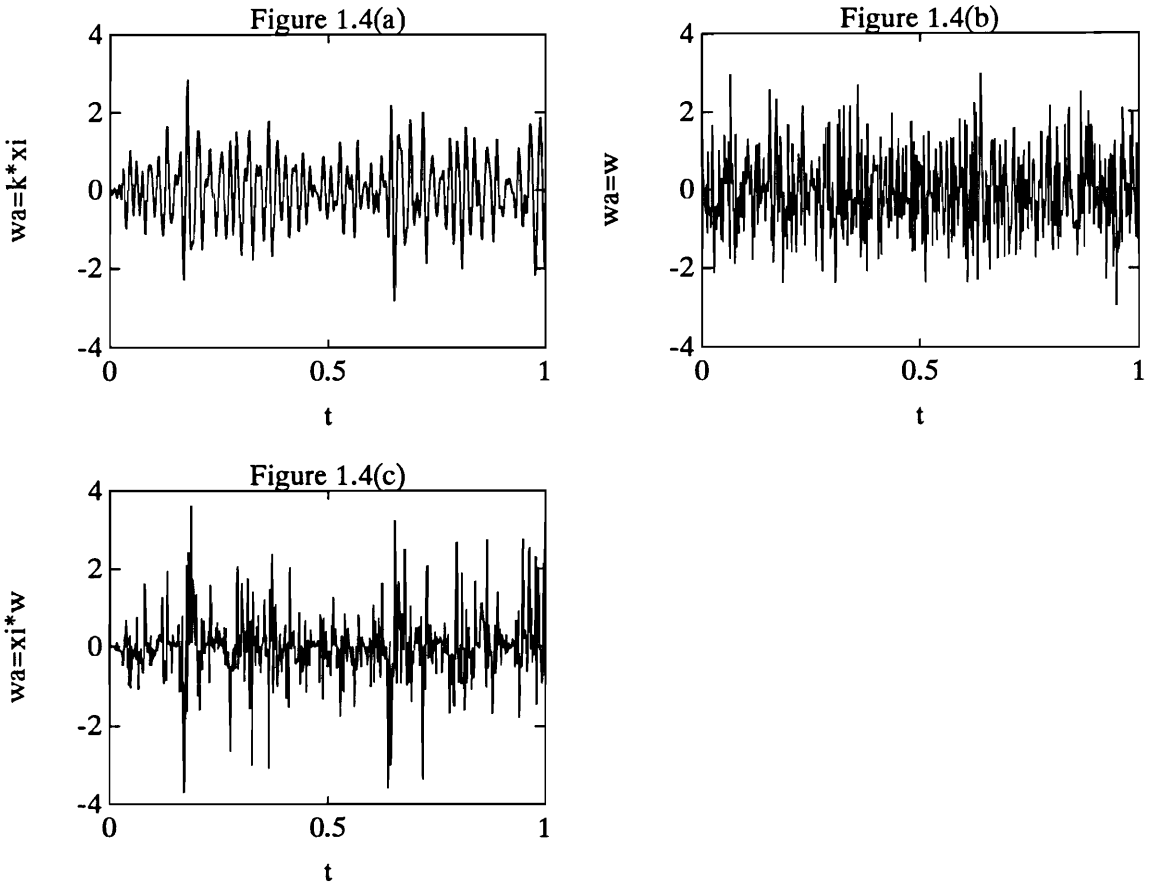

Figure 1.4: Time-Domain Auxiliary Input Models

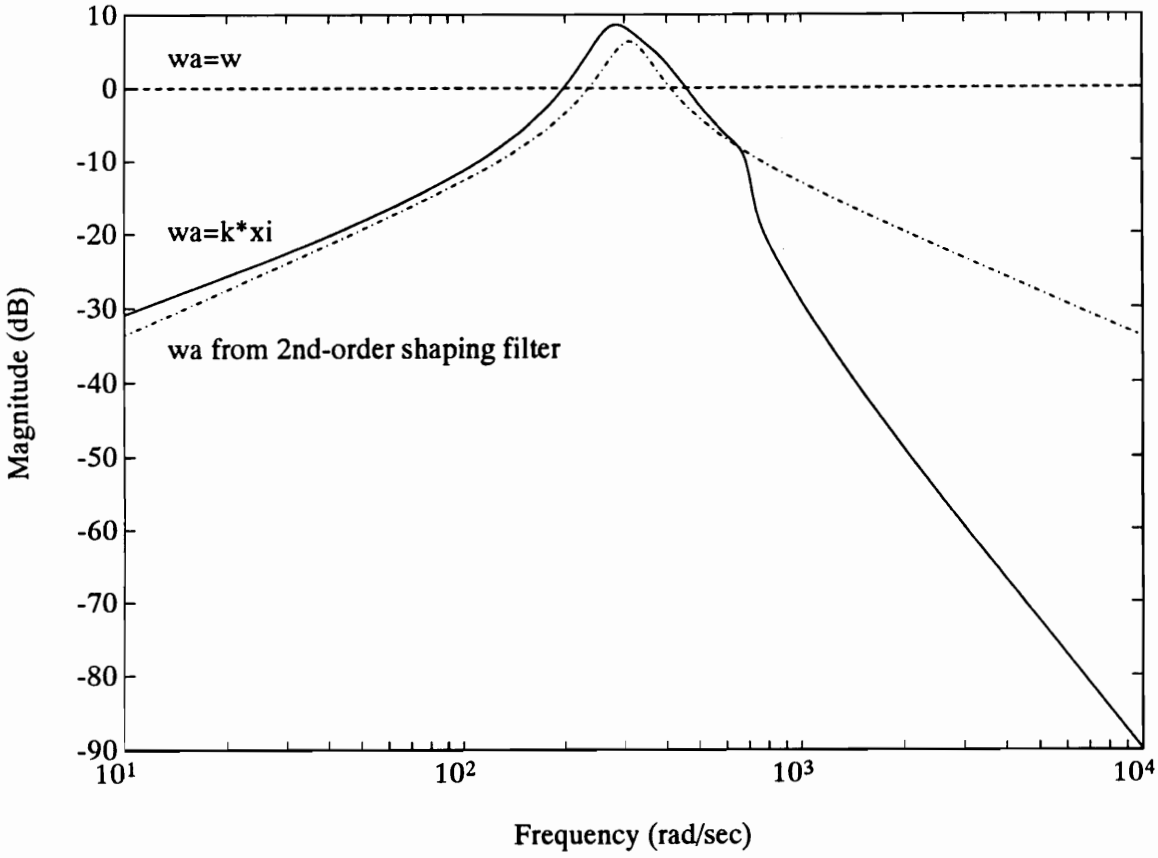

Figure 1.5: Frequency-Domain Auxiliary Input Models

noise model ($w_a = w$). As with time-domain matching, the correct magnitude of the auxiliary inputs cannot be duplicated, but the intensity of the noise signals used to generate the noise models may be adjusted as design parameters. This modification of LQG/PRE does not require solution of the coupled Riccati and Lyapunov equations; however, it effectively increases the size of a full-order compensator, since it involves augmentation of the plant model with additional dynamics.

C. Compensator Order

Many methods are available to reduce compensator order, but they fall into three basic categories — *model reduction, controller reduction,* and *direct design* (Anderson and Liu 1989). Model and controller reduction may be accomplished by such methods as balanced truncation (Moore 1981) or Hankel norm approximation (Glover 1984), but the optimality of the compensator (for a given controller order) — and sometimes even stability — is preserved only by designing a reduced-order controller *directly* from a full-order plant model. The equations for the direct design of LQG reduced-order controllers were developed by Kwakernaak and Sivan (1972, sec. 5.7), among others. However, the solution of these equations requires a gradient search on the many free parameters of the controller state-space model. A far more practical method was developed by Hyland and Bernstein (1984), involving solution of the *optimal projection equations* — a coupled system of two modified Riccati and two modified Lyapunov equations, similar to those mentioned above in the discussion on stochastic control with multiplicative white noise. The Lyapunov-type equations are analogous to the Lyapunov equations required in solving the balanced truncation of the plant or controller, but they are coupled with the modified Riccati equations by means of an optimal projection matrix when the controller is less than full-order. This coupling demonstrates the fact that balancing and controller design cannot be performed separately to obtain an optimal reduced-order controller.

The suggested modifications to LQG/PRE easily fit into the framework of optimal reduced-order compensators, with minor modifications to the optimal projection equations. This is an important feature, since the degree to which controller reduction is possible may very well determine whether modern control applied to flexible structures can be implemented.

D. Steady State, Transient Disturbance Rejection

When sensors are available to measure incoming disturbances, feedforward control becomes an alternative to feedback control. The main advantage of feedforward control is that the compensator does not need to, in effect, model the plant dynamics. Therefore, feedforward control appears to be an attractive alternative when small compensator order is important. However, experience has shown that the advantage in compensator order can be significantly reduced if the feedback controller is designed by optimal projection. Also, feedback controllers can provide added damping to the plant and therefore provide the ability to reject transient disturbances. For these reasons, only feedback control techniques are studied here.

1.2 Contributions

This research has a number of contributions. Firstly, the internal feedback loop (IFL) modeling principle of Tahk and Speyer is generalized and fully exploited to maximize its potential. Secondly, the concept of multiplicative white noise is given broader application by means of a new interpretation, and a complete derivation is given of the controller design equations from first principles. Thirdly, implementation concerns are thoroughly discussed, and explicit algorithms are developed that have application beyond this study. Lastly, the validity of the controller design techniques is demonstrated by their implementation on flexible structure hardware.

The original IFL modeling technique, used by Tahk and Speyer for PRLQG and by Lin for LQG/PRE, was restricted to a white noise model for the auxiliary inputs and resulted only in full-order controllers. In this research, it is shown that a reliance on white noise models may result in relatively poor performance. Frequency-shaped noise and multiplicative white noise models are developed to improve the flexibility of this LQG-based design method. Also, optimal order reduction is incorporated into this framework for the first time. The tremendous savings in controller duty cycle that result are demonstrated, suggesting that the addition of this feature may be critical to the ability of these parameter-robust controllers to be implemented.

The controller design equations for plant models with multiplicative white noise have already been derived for application to a less general and differently motivated robustness problem. However, derivations published up to now have been incomplete,

frequently presuming knowledge of stochastic differential equations and omitting nontrivial steps and clarifications which fully explain the applicability of the design equations. This work attempts to fill that void by supplying an entire and unbroken derivation, complete with surrounding discussion and explicit references to readily available sources.

In the implementation phase of the design, a number of decisions must be made which affect the performance of the control system. The procedures that were used to resolve these problems are discussed in detail, and explicit algorithms are developed for the solution of the coupled Riccati and Lyapunov equations of Chapters 3 and 4.

The application of the parameter-robust and reduced-order controller designs to the control of an actual flexible structure demonstrate their ability to solve real problems and allow us to quantify the performance of these designs. Sacrifices in performance due to robustness enhancement and controller reduction are measured and limitations are discovered as to the maximum amount of controller reduction possible.

2. Frequency-Shaped Noise

2.1 Problem Statement

The frequency-domain method of auxiliary input modeling attempts to replicate the frequency-domain behavior of the (unknown) auxiliary input signals by means of a frequency-shaped noise model. Rather than using white noise (as does LQG/PRE) to model the auxiliary input signals w_a, w_b, and w_c (see Figure 1.3), a much more precise determination is made of the actual power spectral density of these signals. The resulting auxiliary noise model is then combined with a cost functional penalty on the auxiliary output signals — z_a, z_b, and z_c — to form a new method of parameter-robust controller design. The noise modeling phase of this design method consists of 1) finding an approximation for the power spectra of the auxiliary input signals, 2) designing a noise shaping filter to warp the power spectrum of white noise into this shape, and 3) appending the shaping filter dynamics to the plant.

Referring to the LQG/PRE model of the plant in equations (1.7) and Figure 1.3, we see that the actual power spectra of the auxiliary inputs cannot be found, because these signals are a function of the unknown diagonal matrices L_a, L_b, and L_c, and of the compensator matrix triple, $\{A_c,B_c,C_c\}$, which is yet to be designed. However, a very good approximation to the power spectra may still be found. For relatively small perturbations in the A-, B-, and C-matrices of the plant, compared to the nominal values, the feedback and feedforward loops containing the auxiliary signals have little effect on the frequency content of the states and controls, upon which the auxiliary inputs depend. Therefore, we may safely break these loops at the auxiliary inputs (see Figure 2.1) for the purpose of approximating the power spectra. Now, the unknown diagonal L-matrices only determine the sign and magnitude of the auxiliary inputs, not their frequency content, and the elements of the vectors w_a, w_b, and w_c are proportional to those of z_a, z_b, and z_c, respectively. Therefore, the task of finding the power spectral densities of the auxiliary input signals is equivalent to finding the power spectra of the auxiliary outputs, modulo some proportionality constants. The values of those constants are not important, because the relative noise intensities of the auxiliary inputs must later be adjusted in order to provide the desired amount parameter robustness. The problem of not knowing the compensator matrices a priori can be treated by designing a standard LQG compensator based on the nominal parameter values of the plant and using it to calculate power spectra. Experience has shown that the LQG

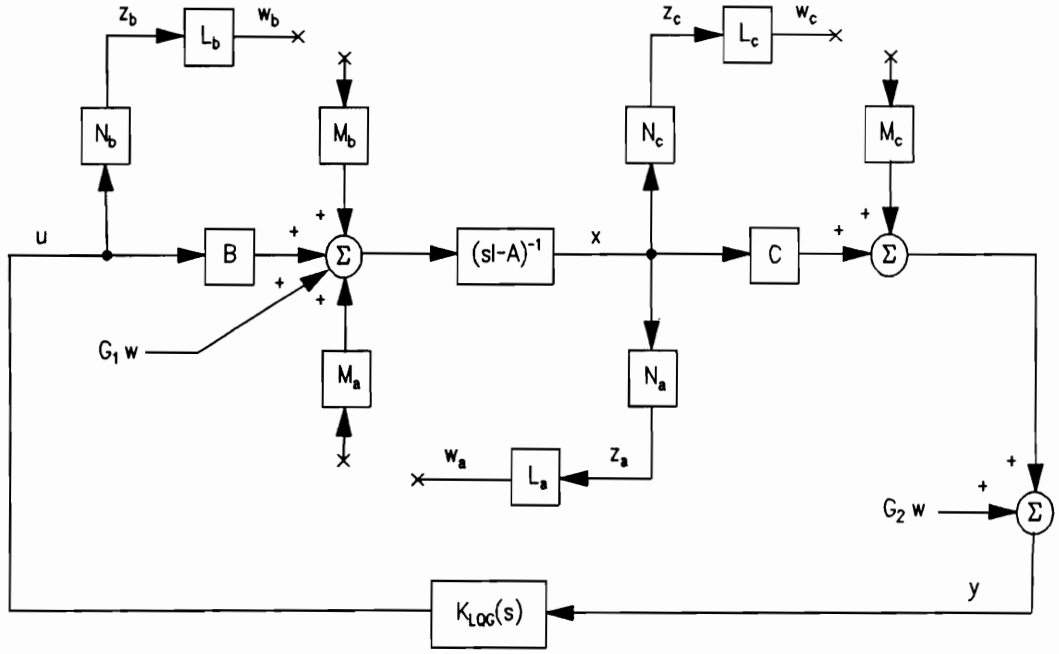

Figure 2.1: Model Used to Approximate Power Spectral Densities of Auxiliary Inputs

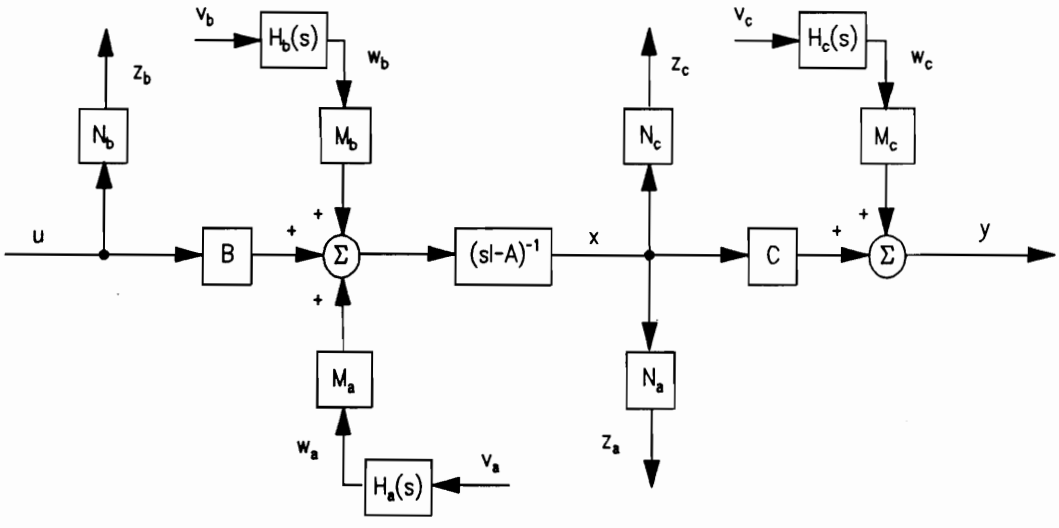

Figure 2.2: Plant Augmented with Noise Shaping Filters

compensator can provide an excellent substitute for this purpose. If greater accuracy were required, one could use an iterative process of alternately designing a compensator and approximating power spectra until a satisfactory noise model were found.

2.2 Noise Shaping Filter Design

Figure 2.1 illustrates the resulting model used to obtain approximations of the power spectral densities of the auxiliary input signals. The corresponding state-space model of the n-th order plant, n-th order compensator, and auxiliary inputs (assuming the L-matrices are all identity matrices) is given by

$$\dot{x}(t) = Ax(t) + Bu(t) + G_1 w(t) \tag{2.1}$$

$$y(t) = Cx(t) + G_2 w(t) \tag{2.2}$$

$$x_c(t) = A_c x_c(t) + B_c y(t) \tag{2.3}$$

$$u(t) = C_c x_c(t) \tag{2.4}$$

$$w_a(t) = N_a x(t), \quad w_b(t) = N_b u(t), \quad w_c(t) = N_c x(t) \tag{2.5}$$

It is assumed without loss of generality that the exogenous noise signal, $w(t)$, is a vector of mutually independent, unit intensity white noise processes. Included in $w(t)$ are both the process noise and sensor noise. The closed-loop system with input w and outputs w_a, w_b, and w_c is described by

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{G}w(t) \tag{2.6}$$

$$w_a(t) = [N_a \quad 0]\,\tilde{x}(t) \tag{2.7}$$

$$w_b(t) = [0 \quad N_b C_c]\,\tilde{x}(t) \tag{2.8}$$

$$w_c(t) = [N_c \quad 0]\,\tilde{x}(t) \tag{2.9}$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \quad \tilde{A} \triangleq \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \quad \tilde{G} \triangleq \begin{bmatrix} G_1 \\ B_c G_2 \end{bmatrix} \tag{2.10}$$

From this closed-loop model and the assumptions on w, the power spectral densities of the elements of w_a, w_b, and w_c may be derived one at a time. Let us say, arbitrarily, that we are interested in w_a and that it is a scalar. The following procedure is easily repeated for all other existing auxiliary inputs. Then we have the frequency-domain description of w_a,

$$w_a = [N_a \quad 0](sI - \tilde{A})^{-1}\tilde{G}w \tag{2.11}$$

$$\triangleq H(s)w$$

In general, then, $H(s)$ is a $2n$-th order, multi-input, single-output transfer function matrix. If the plant is of sufficiently low order, we could stop here and use $H(s)$ as the noise shaping filter to be appended to the plant. Normally, however, this would not be practical, so a lower-order filter will be found, which is single-input, single-output. One might suspect that a low-order balanced truncation approximation to $H(s)$ would provide a good substitute, but examination of their comparative frequency responses for particular examples has shown that there is a much more effective method of designing an accurate noise shaping filter.

First we find an expression for the power spectral density, $S_{w_a w_a}(\omega)$, of w_a. For a vector exogenous noise process w, the definition of $H(s)$ in (2.11) leads to (Maciejowski 1989, p. 98):

$$S_{w_a w_a}(\omega) = H(\omega)S_{ww}(\omega)H^*(\omega) \tag{2.13}$$

$$= H(\omega)H^*(\omega)$$

where $H^*(\omega)$ is the complex conjugate transpose of the transfer function matrix from w to w_a. A suitable shaping filter designed to yield this power spectrum may be found by deriving its magnitude frequency response from the information in (2.13) and curve fitting a transfer function to give that response. The desired frequency response of our shaping filter transfer function, $H_{sf}(\omega)$, is

$$\left|H_{sf}(\omega)\right| = \sqrt{H(\omega)H^*(\omega)} \tag{2.14}$$

In order to hold down the controller order, the magnitude response in (2.14) should be fit using a shaping filter of as low order as possible.

2.3 Controller Design Summary

Once the shaping filters to produce auxiliary inputs w_a, w_b, and w_c have been found — say, $H_a(\omega)$, $H_b(\omega)$, and $H_c(\omega)$, respectively — they may be appended to the plant, as shown in Figure 2.2. Define the state-space models of the shaping filters by the following identities,

$$H_a(\omega) = C_{wa}(j\omega I - A_{wa})^{-1}B_{wa} + D_{wa} \tag{2.15}$$

$$H_b(\omega) = C_{wb}(j\omega I - A_{wb})^{-1}B_{wb} + D_{wb}$$

$$H_c(\omega) = C_{wc}(j\omega I - A_{wc})^{-1}B_{wc} + D_{wc}$$

Then the state-space model of the plant, (2.1) and (2.2), is augmented so that we have,

$$\frac{d}{dt}\begin{bmatrix} x \\ x_a \\ x_b \\ x_c \end{bmatrix} = \begin{bmatrix} A & M_aC_{wa} & M_bC_{wb} & 0 \\ 0 & A_{wa} & 0 & 0 \\ 0 & 0 & A_{wb} & 0 \\ 0 & 0 & 0 & A_{wc} \end{bmatrix}\begin{bmatrix} x \\ x_a \\ x_b \\ x_c \end{bmatrix} + \begin{bmatrix} B \\ 0 \\ 0 \\ 0 \end{bmatrix}u + \begin{bmatrix} G_1 & M_aD_{wa} & M_bD_{wb} & 0 \\ 0 & B_{wa} & 0 & 0 \\ 0 & 0 & B_{wb} & 0 \\ 0 & 0 & 0 & B_{wc} \end{bmatrix}\begin{bmatrix} w \\ v_a \\ v_b \\ v_c \end{bmatrix}$$

$$y = \begin{bmatrix} C & 0 & 0 & M_cC_{wc} \end{bmatrix}\begin{bmatrix} x \\ x_a \\ x_b \\ x_c \end{bmatrix} + Du + \begin{bmatrix} G_2 & 0 & 0 & M_cD_{wc} \end{bmatrix}\begin{bmatrix} w \\ v_a \\ v_b \\ v_c \end{bmatrix} \tag{2.16}$$

The supplementary independent white noise variables v_a, v_b, and v_c are treated as exogenous noise sources, and there are as many of these additional noise sources in the augmented model as there are independent parameter uncertainties in the plant.

After augmenting the plant with the noise shaping filter dynamics as above, the rest of the frequency-domain method is the same as LQG/PRE. The intensities of the auxiliary inputs and the weights on the auxiliary outputs are adjusted to provide the desired amount of parameter robustness. Then a standard LQG compensator is designed based on the augmented model with modified noise intensities and cost functional.

3. Multiplicative White Noise

3.1 Problem Statement

The time-domain method of auxiliary input modeling uses artificial noise of time-varying intensity to account for the time-varying amplitude of the actual auxiliary input signals. Since the amplitude of the true signals is proportional to a known linear combination of the states or controls, a multiplicative white noise model provides the desired proportionate noise intensity as a function of time. As in Chapter 2, the auxiliary input model is supplemented with a cost functional penalty on the auxiliary outputs. The time-domain method without the added flexibility of auxiliary output penalties is equivalent to the compensator design method of Bernstein and Hyland (e.g., 1988a) for multiplicative white noise for uncertain systems, although the motivation here is different. Since the optimal linear quadratic compensator for the resulting plant model cannot be found by standard LQG techniques, a more lengthy development is required here than was necessary for the frequency-domain method of Chapter 2.

Modeling of the auxiliary input signals by multiplicative white noise requires modification of the state and/or output equations of the plant. For the general case of uncertainties in the A-, B-, and C-matrices, we have the n-th order system,

$$dx_t = Ax_t dt + \sum_{i=1}^{p} \gamma_i A_i x_t dv_{it} + Bu_t dt + \sum_{i=1}^{p} \gamma_i B_i u_t dv_{it} + G_1 d\beta_t \tag{3.1}$$

$$dy_t = Cx_t dt + \sum_{i=1}^{p} \gamma_i C_i x_t dv_{it} + G_2 d\beta_t \tag{3.2}$$

These equations are written in differential form, because the development of the LQ optimal compensator for systems with multiplicative white noise requires the use of a more rigorous form of stochastic differential equation theory. The theory makes subtle distinctions in the interpretation of noise processes, which will be reviewed briefly in the next section. Such precise definitions are inconsistent with the use of the concept of Gaussian white noise, which, strictly speaking, is not well defined in the more rigorous framework. Therefore, all of the noise variables in this chapter will be written as differentials of *Wiener processes,* also known as *Brownian motion processes* and loosely referred to as the "integral of Gaussian white noise". Note also that variables which are a function of time are identified as such by a "t" in the subscript. This

compact notation is used throughout the chapter in order to prevent the equations from becoming too cumbersome.

The model of equations (3.1) and (3.2) represents a system with p independent auxiliary input signals, arising from p independent parameter uncertainties. The v_{it} ($i = 1,2,\ldots,p$) are scalar independent Wiener processes, with $\mathbb{E}\{dv_{it}{}^2\} = \gamma_i dt$, and β_t is a vector Wiener process with $\mathbb{E}\{d\beta_t d\beta_t{}^{\mathrm{T}}\} = V dt$. As an illustrative example, consider the following 1-mode damped oscillator with acceleration output and two independent auxiliary input signals.

$$dx_t = \begin{bmatrix} 0 & 1 \\ -\omega^2 & -2\zeta\omega \end{bmatrix} x_t dt + \begin{bmatrix} 0 & 0 \\ 0 & \delta_1 \end{bmatrix} x_t dv_{1t} + \begin{bmatrix} 0 \\ b \end{bmatrix} u_t + \begin{bmatrix} 0 \\ \delta_2 \end{bmatrix} u_t dv_{2t} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} d\beta_t \tag{3.3}$$

$$dy_t = \begin{bmatrix} -\omega^2 & -2\zeta\omega \end{bmatrix} x_t dt + \begin{bmatrix} 0 & \delta_1 \end{bmatrix} x_t dv_{1t} + \begin{bmatrix} 0 & 1 \end{bmatrix} d\beta_t \tag{3.4}$$

An uncertainty in the damping parameter, $-2\zeta\omega$, must occur in both the A- and C-matrices. Since these parameter uncertainties are necessarily the same, they comprise a single independent uncertainty and utilize the same noise process, dv_{1t}. The resulting auxiliary inputs consist of (1) *state-dependent noise* (noise with intensity proportional to the states and entering at the process noise port) and (2) *measurement-dependent noise* (noise with intensity proportional to the states and entering at the sensor noise port). An uncertainty in the B-matrix will generally be independent of the damping parameter uncertainty and is provided for by an independent noise process, v_{2t}, yielding an auxiliary input which is *control-dependent noise* (noise with intensity proportional to the input and entering at the process noise port). The positive scalar constants δ_1 and δ_2 are design parameters which effectively adjust the intensities of the auxiliary input noise signals according to the magnitude of the uncertainties present.

Given the n_c-th order state-space model of the compensator,

$$dx_{ct} = A_c x_{ct} dt + B_c dy_t \tag{3.5}$$

$$u_t = C_c x_{ct} \tag{3.6}$$

where $n_c = n$ for a full-order controller, we can represent the closed-loop system as follows.

$$d\tilde{x}_t = \tilde{A}\tilde{x}_t dt + \sum_{i=1}^{p} \tilde{A}_i \tilde{x}_t dv_{it} + \tilde{G} d\beta_t \tag{3.7a}$$

$$= \tilde{A}\tilde{x}_t dt + \begin{bmatrix} \tilde{A}_1\tilde{x}_t & \tilde{A}_2\tilde{x}_t & \cdots & \tilde{A}_p\tilde{x}_t & \tilde{G} \end{bmatrix} \begin{bmatrix} dv_t \\ d\beta_t \end{bmatrix} \tag{3.7b}$$

In equation (3.7) we have defined the vector Wiener process

$$dv_t \triangleq \begin{bmatrix} dv_{1t} & dv_{2t} & \cdots & dv_{pt} \end{bmatrix}^{\mathrm{T}} \tag{3.8}$$

and the closed-loop quantities, denoted by tildas, as follows:

$$\tilde{n} \triangleq n + n_c, \qquad \tilde{x}_t \triangleq \begin{bmatrix} x_t \\ x_{ct} \end{bmatrix}, \qquad \tilde{G} \triangleq \begin{bmatrix} G_1 \\ B_c G_2 \end{bmatrix},$$

$$\tilde{A} \triangleq \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \qquad \tilde{A}_i \triangleq \begin{bmatrix} A_i & B_i C_c \\ B_c C_i & 0 \end{bmatrix} \tag{3.9}$$

Now the compensator design objective may be stated in terms of an LQ optimization problem. Define the standard cost functional,

$$J(A_c, B_c, C_c) \triangleq \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left\{ \int_{t_0}^{T} \left[x_t^{\mathrm{T}} R_1 x_t + 2x_t^{\mathrm{T}} R_{12} u_t + u_t^{\mathrm{T}} R_2 u_t \right] dt \right\} \tag{3.10}$$

In order to simplify the problem, the cost functional will be rewritten in an equivalent form without the integral,

$$J(A_c, B_c, C_c) = \lim_{t\to\infty} \mathbb{E}\{ x_t^{\mathrm{T}} R_1 x_t + 2x_t^{\mathrm{T}} R_{12} u_t + u_t^{\mathrm{T}} R_2 u_t \, dt \} \tag{3.11}$$

[Kwakernaak and Sivan 1972, p. 394, Theorem 5.4]. In fact, the limit may be eliminated as well, due to the assumed stationarity of the exogenous noise processes, provided the initial time is infinitely far in the past (i.e., for $t_0 \to -\infty$).

In light of the performance objective (3.11), the LQ optimal compensator design problem for systems with multiplicative white noise may be stated as follows. Given a plant (1.1)-(1.2) with p independent auxiliary inputs, find a matrix triple, $\{A_c, B_c, C_c\}$, to realize a compensator (3.5)-(3.6) which minimizes the cost functional (3.11).

3.2 Preliminaries of Stochastic Differential Equation Theory

Before deriving the necessary conditions for the optimization problem just stated, a brief overview will be given of those elements of stochastic differential equation theory which will be used in the derivation. For more information on matters where specific references are not given, see Jazwinski (1970) and references therein.

Consider the stochastic differential equation,

$$dx_t = f(x_t, t)dt + G(x_t, t)d\beta_t\,, \quad t \geq t_0 \tag{3.12}$$

where β_t is a vector Wiener process with $\mathbb{E}\{d\beta_t d\beta_t^{\mathrm{T}}\} = Q(t)dt$. Associated with this equation is the integral equation,

$$x_t - x_{t_0} = \int_{t_0}^{t} f(x_\tau, \tau)d\tau + \int_{t_0}^{t} G(x_\tau, \tau)d\beta_\tau \tag{3.13}$$

The first integral in (3.13) can be defined as a Riemann integral for sample functions of x, or as a *mean square Riemann integral* (Jazwinski 1970, p. 66) for x a stochastic process. If G were a function of time only, as would be the case for ordinary state equations with *additive* white noise, the second integral would be known as a *Wiener integral.* However, added complications arise when *multiplicative* white noise is present (G a function of x). In this more general case, the second integral may also be defined in a mean square sense, but such a definition will not be unique. Itô (1944) was the first to define this type of stochastic integral. In doing so he modeled the Riemann sum after a *forward difference equation*, effectively sampling the integrand at the beginning of each partition, and then proceeded with a mean square limit to define the integral. When the integral is interpreted in this sense, it is called an *Itô stochastic integral*, and the associated differential equation (3.12) is called an *Itô stochastic differential equation*, or *Itô equation.* The rules of calculus which result from this interpretation are called *Itô stochastic calculus.* Note that the closed-loop state equation, (3.7), can be interpreted as an Itô equation.

What is not immediately apparent is that the second integral in (3.13) has different interpretations according to where the partitions of the Riemann sum are sampled. For example, another important interpretation, called the *Stratonovich stochastic integral* (Stratonovich 1966), arises when the differential equation is modeled

as a *central difference equation* before taking the limit (i.e., partitions sampled in the center). Actually, the differences between the Itô and Stratonovich stochastic integrals are somewhat more profound than that, but the mathematical details which distinguish the two are beyond the scope of this work. The purpose of the foregoing discussion is to inform the reader that the stochastic differential equation (3.12) which appears in state-space models with multiplicative white noise has no meaning without assigning the associated integrals a particular interpretation (e.g., Itô or Stratonovich integrals).

The Itô stochastic integral is defined over a much broader class of functions than the Stratonovich integral and is used for most theoretical work in stability and control. Therefore, Itô stochastic calculus will be applied in the development of the next section. The Stratonovich integral does have a number of nice properties, but a simple transformation exists between differential equations of the two different interpretations (Jazwinski 1970, pp. 118–20, 131). It is not clear which interpretation is more "correct" for the application studied in this work. However, Wong and Zakai (1965) discovered a property of the Stratonovich noise model that gives it intuitive appeal for applications to physical systems:

Theorem 3.1: Let x_t be the solution to the stochastic differential equation (3.12). Now, replace the Wiener process β_t with a sequence of continuous piecewise linear approximations, $\beta_t^{(n)}$, such that $\beta_t^{(n)}$ converges to β_t as $n \to \infty$. Then the solutions, $x_t^{(n)}$, to the resulting sequence of ordinary differential equations converge to x_t if β_t in (3.12) is interpreted in the Stratonovich sense. This property does not hold in general for β_t interpreted in the Itô sense.

Because of Theorem 3.1 and the fact that Bernstein (e.g., 1987, Bernstein and Hyland 1988b) argues in favor of the Stratonovich interpretation for applications to lightly damped flexible structures, a comparison will be made in Chapter 6 of the performance of the two different controllers that result from the two stochastic integral interpretations when applied to a particular example.

An important result of Itô stochastic calculus, called *Itô's integration formula*, will be used in the next section. It is derived directly from *Itô's Theorem* below.

Theorem 3.2 (Itô's Theorem or *Itô's Chain Rule)* (e.g., Jazwinski 1970, p. 112, Lemma 4.2) Also, for a more formal proof, see Gikhman and Skorokhod (1969, pp. 387–91):

Let $\phi(x_t, t)$ be a real, scalar linear functional with continuous partial derivatives,

$$\phi_t \triangleq \frac{\partial \phi}{\partial t}, \qquad \phi_x \triangleq \frac{\partial \phi}{\partial x}, \qquad \phi_{xx} \triangleq \frac{\partial^2 \phi}{\partial x^2} \tag{3.14}$$

where x_t is the unique solution to the vector Itô stochastic differential equation, (3.12), and β_t is a vector Wiener process with $\mathbb{E}\{d\beta_t d\beta_t^{\mathrm{T}}\} = Q(t)dt$. Then the stochastic differential $d\phi$ of ϕ is

$$d\phi = \phi_t dt + \phi_x{}^{\mathrm{T}} dx_t + \tfrac{1}{2} \operatorname{trace}\{GQG^{\mathrm{T}}\phi_{xx}\}dt \tag{3.15}$$

The existence of the third term in (3.15) establishes the fact that the usual chain rule does not hold in Itô stochastic calculus, although it does hold, incidentally, in Stratonovich stochastic calculus.

Integrating (3.15) and taking the expected value conditioned on $x_{t_0} = x_0$, *Itô's integration formula* results (Wonham 1970, p. 137):

$$\mathbb{E}\Big\{\phi(x_t, t) \mid x_{t_0} = x_0\Big\}$$

$$= \phi(x_0, t_0) + \mathbb{E}\left\{ \int_{t_0}^{t} \Big[\phi_\tau + \phi_x{}^{\mathrm{T}} f + \tfrac{1}{2} \operatorname{trace}\{GQG^{\mathrm{T}}\phi_{xx}\}\Big]\, d\tau \mid x_{t_0} = x_0 \right\} \tag{3.16}$$

The arguments, (x_τ, τ), have been omitted for functions in the integrand in order to conserve space. The term in (3.15) involving $d\beta_t$ (by way of dx_t) does not appear in (3.16) due to the Martingale property of a stochastic integral w.r.t. a Wiener process. With this formula, we are now ready to find an expression for $\tilde{Q}_t$.

3.3 Conversion to Deterministic Minimization Problem

The first step in deriving the necessary conditions will be to simplify the optimization problem stated at the end of section 3.1 by converting it into a *deterministic* minimization problem. Define the closed-loop state covariance,

$$\tilde{Q}_t \triangleq \mathbb{E}\{\tilde{x}_t \tilde{x}_t{}^{\mathrm{T}}\}, \qquad \tilde{Q} \triangleq \lim_{t \to \infty} \tilde{Q}_t \tag{3.17}$$

and the closed-loop state weighting matrix,

$$\tilde{R} \triangleq \begin{bmatrix} R_1 & R_{12}C_c \\ C_c{}^{\mathrm{T}}R_{12}{}^{\mathrm{T}} & C_c{}^{\mathrm{T}}R_2C_c \end{bmatrix} \tag{3.18}$$

Then the cost functional of (3.11) becomes,

$$J(A_c, B_c, C_c) = \lim_{t\to\infty} \mathbb{E}\{\tilde{x}_t{}^{\mathrm{T}}\tilde{R}\tilde{x}_t\} \tag{3.19a}$$

$$= \mathrm{trace}\{\tilde{Q}\tilde{R}\} \tag{3.19b}$$

The objective then is to minimize the cost as described by (3.19b). The quantity $\tilde{R}$ is known, because it is specified by the cost criterion in the problem statement. It remains, however, to find an expression for the closed-loop state covariance, $\tilde{Q}$.

Define the functional, $\phi_{ij}(\tilde{x}_t) = \tilde{x}_{ti}\tilde{x}_{tj}$, where $\tilde{x}_{ti}$ denotes the i^{th} element of the solution vector $\tilde{x}_t$ to the closed-loop state equation (3.7). Note that $\tilde{Q}_t(i,j) = \mathbb{E}\{\phi_{ij}(\tilde{x}_t)\}$. The f, G, and Q quantities of (3.7) as defined by the generic Itô equation, (3.12), are given by,

$$f(\tilde{x}_t) = \tilde{A}\tilde{x}_t, \qquad G(\tilde{x}_t) = \begin{bmatrix} \tilde{A}_1\tilde{x}_t & \tilde{A}_2\tilde{x}_t & \dots & \tilde{A}_p\tilde{x}_t & \tilde{G} \end{bmatrix},$$

$$Q = \mathrm{diag}\{\gamma_1, \gamma_2, \dots, \gamma_p, V\} \tag{3.20}$$

The partial derivatives needed for Itô's integration formula are computed as follows.

$$\phi_t = 0, \tag{3.21}$$

$$\phi_{\tilde{x}}{}^{\mathrm{T}} = \begin{cases} [0 \dots 0 \underset{(i)}{\tilde{x}_{tj}} 0 \dots 0 \underset{(j)}{\tilde{x}_{ti}} 0 \dots 0], & i \neq j \\ [0 \dots 0 \; 2\tilde{x}_{ti} \; 0 \dots 0], & i = j \end{cases} \tag{3.22a}$$

$$= \tilde{x}^{\mathrm{T}}(E_{ij} + E_{ji}) \tag{3.22b}$$

where E_{ij} denotes the elementary matrix with 1 in the $(i,j)^{\mathrm{th}}$ position and 0's elsewhere.

$$\phi_{\tilde{x}\,\tilde{x}} = \begin{cases} \begin{matrix} & & (i) & & (j) & \\ (i) \\ (j) \end{matrix} \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & 1 & \vdots \\ \vdots & 1 & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, & i \neq j \\ \begin{matrix} & & (i) & & \\ (i) \end{matrix} \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & 2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, & i = j \end{cases} \tag{3.23a}$$

$$= (E_{ij} + E_{ji}) \tag{3.23b}$$

Therefore, the second and third terms of the integrand of Itô's integration formula are

$$\begin{aligned} \phi_{\tilde{x}}{}^{\mathrm{T}} f &= \tilde{x}_{t i}(\tilde{A}\tilde{x}_t)_j + \tilde{x}_{t j}(\tilde{A}\tilde{x}_t)_i \\ &= (\tilde{A}\tilde{x}_t \tilde{x}_t{}^{\mathrm{T}})_{ji} + (\tilde{A}\tilde{x}_t \tilde{x}_t{}^{\mathrm{T}})_{ij} \\ &= (\tilde{x}_t \tilde{x}_t{}^{\mathrm{T}} \tilde{A}^{\mathrm{T}})_{ij} + (\tilde{A}\tilde{x}_t \tilde{x}_t{}^{\mathrm{T}})_{ij} \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} \operatorname{trace}\{GQG^{\mathrm{T}}\phi_{\tilde{x}\,\tilde{x}}\} &= \operatorname{trace}\Big\{\Big[\tilde{A}_1\tilde{x}_t \quad \tilde{A}_2\tilde{x}_t \quad \dots \quad \tilde{A}_p\tilde{x}_t \quad \tilde{G}\Big] \operatorname{diag}\{\gamma_1, \gamma_2, \dots, \gamma_p, V\} \\ &\qquad \Big[\tilde{A}_1\tilde{x}_t \quad \tilde{A}_2\tilde{x}_t \quad \dots \quad \tilde{A}_p\tilde{x}_t \quad \tilde{G}\Big]^{\mathrm{T}} (E_{ij} + E_{ji})\Big\} \\ &= \operatorname{trace}\Big\{\Big[\sum_{k=1}^{p} \gamma_k \tilde{A}_k \tilde{x}_t \tilde{x}_t{}^{\mathrm{T}} \tilde{A}_k{}^{\mathrm{T}} + \tilde{G}V\tilde{G}^{\mathrm{T}}\Big] (E_{ij} + E_{ji})\Big\} \\ &= 2\Big[\sum_{k=1}^{p} (\gamma_k \tilde{A}_k \tilde{x}_t \tilde{x}_t{}^{\mathrm{T}} \tilde{A}_k{}^{\mathrm{T}})_{ij} + (\tilde{V})_{ij}\Big] \end{aligned} \tag{3.25}$$

where $\tilde{V} \triangleq \tilde{G}V\tilde{G}^{\mathrm{T}}$. In the last step of (3.25), the identity

$$\text{trace}\{A_{n\times m}B_{m\times n}\} = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}B_{ji} \tag{3.26}$$

was used, along with the symmetry of the expression enclosed in brackets.

Now we may apply Itô's integration formula to find an expression for the closed-loop state covariance, $\tilde{Q}_t$.

$$\mathbb{E}\{\tilde{x}_{ti}\tilde{x}_{tj} \mid \tilde{x}_{t_0} = \tilde{x}_0\} = \tilde{x}_{t_0 i}\tilde{x}_{t_0 j} + \mathbb{E}\Big\{\int_{t_0}^{t}[(\tilde{x}_\tau \tilde{x}_\tau^{\mathrm{T}}\tilde{A}^{\mathrm{T}})_{ij} + (\tilde{A}\tilde{x}_\tau \tilde{x}_\tau^{\mathrm{T}})_{ij} + \sum_{k=1}^{p}(\gamma_k \tilde{A}_k \tilde{x}_\tau \tilde{x}_\tau^{\mathrm{T}}\tilde{A}_k^{\mathrm{T}})_{ij} + (\tilde{V})_{ij}]d\tau \mid \tilde{x}_{t_0} = \tilde{x}_0\Big\} \tag{3.27}$$

$$\tilde{Q}_t(i,j) = \tilde{Q}_{t_0}(i,j) + \int_{t_0}^{t}[(\tilde{Q}_\tau \tilde{A}^{\mathrm{T}})_{ij} + (\tilde{A}\tilde{Q}_\tau)_{ij} + \sum_{k=1}^{p}(\gamma_k \tilde{A}_k \tilde{Q}_\tau \tilde{A}_k^{\mathrm{T}})_{ij} + (\tilde{V})_{ij}]d\tau \tag{3.28}$$

$$\tilde{Q}_t = \tilde{Q}_{t_0} + \int_{t_0}^{t}[\tilde{Q}_\tau \tilde{A}^{\mathrm{T}} + \tilde{A}\tilde{Q}_\tau + \sum_{i=1}^{p}\gamma_i \tilde{A}_i \tilde{Q}_\tau \tilde{A}_i^{\mathrm{T}} + \tilde{V}]d\tau \tag{3.29}$$

By the Fundamental Theorem of Integral Calculus,

$$\dot{\tilde{Q}}_t = \tilde{A}\tilde{Q}_t + \tilde{Q}_t\tilde{A}^{\mathrm{T}} + \sum_{i=1}^{p}\gamma_i \tilde{A}_i \tilde{Q}_t \tilde{A}_i^{\mathrm{T}} + \tilde{V} \tag{3.30}$$

The cost functional, (3.19b), requires the steady state solution to (3.30), but first we must determine under what conditions such a solution exists. To that end, we will consider an alternate expression for $\tilde{Q}_t$ which is useful for examining the *stochastic stability* of the system. The following lemma will be used repeatedly.

Lemma 3.1: Given matrices A, B, and C of compatible dimensions,

$$\text{vec}(ABC) = (C^{\mathrm{T}} \otimes A)\text{vec}(B) \tag{3.31}$$

Proof: See Graham (1981, p. 25, Property VIII).

In Lemma 3.1, the "vec" operator merely organizes the columns of its matrix argument into a single column vector with the first column on top and each subsequent column

beneath the previous one. The symbol "$\otimes$" represents the Kronecker product. Now define

$$\mathcal{A} \triangleq \tilde{A} \oplus \tilde{A} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i \otimes \tilde{A}_i \tag{3.32}$$

where "$\oplus$" is the Kronecker sum. Then taking the "vec" operation of (3.30) and applying Lemma 3.1 to the first three terms on the right hand side, we obtain the alternate expression,

$$\mathrm{vec}(\dot{\tilde{Q}}_t) = \mathcal{A}\ \mathrm{vec}(\tilde{Q}_t) + \mathrm{vec}(\tilde{V}) \tag{3.33}$$

From this equation we see that when $\mathcal{A}$ is a stable matrix, a steady state solution, $\tilde{Q}$, exists. Under this condition (i.e., stable $\mathcal{A}$) the closed-loop system (3.7) is said to be *second-moment stable* (Arnold 1974), and $\tilde{Q}$ is the solution to the algebraic equation,

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i \tilde{Q} \tilde{A}_i^{\mathrm{T}} + \tilde{V} \tag{3.34}$$

which is a modified Lyapunov equation. Note that when the plant uncertainties are zero (i.e., $\gamma_i = 0$ or $\tilde{A}_i = 0$, for all $i = 1,\ldots,p$) the third term in (3.34) vanishes and the standard Lyapunov equation results for the steady-state second-moment matrix of the state vector for a system driven by white noise [see e.g., Kwakernaak and Sivan (1972, p.101, Theorem 1.52)]. In order to assure that the cost functional, J, is finite and independent of the system initial values, only second-moment stabilizing compensators will be admissible.

Therefore the original stochastic optimization problem has been reduced to the following deterministic one:

minimize:	$\mathrm{trace}(\tilde{Q}\tilde{R})$	(3.35)
over the set:	$\{(A_c, B_c, C_c)$: $\mathcal{A}$ is stable$\}$	
subject to:	$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i \tilde{Q} \tilde{A}_i^{\mathrm{T}} + \tilde{V}$	

3.4 Derivation of the First-Order Necessary Conditions

The first-order necessary conditions that a controller triple, $\{A_c, B_c, C_c\}$, must meet to solve this type of problem have been derived by at least two different methods — the Lagrange multiplier method (Bernstein and Haddad, 1989), used to solve an $\mathcal{H}_\infty$ problem, and the calculus of variations method (Bernstein and Hyland, 1988a). The Lagrange multiplier solution is somewhat easier to follow but could not be found in complete form in the literature for this particular problem. Therefore, the bulk of it will be derived here, leading up to the simple but more tedious algebraic manipulations, which will be referenced to a paper where they are found in their entirety.

The Lagrangian corresponding to the optimization problem (3.35), can be written as,

$$\mathcal{L}(A_c, B_c, C_c, \tilde{Q}, \tilde{P}, \lambda) = \mathrm{tr}\{\lambda \tilde{Q}\tilde{R} + [\tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i \tilde{Q} \tilde{A}_i^{\mathrm{T}} + \tilde{V}]\tilde{P}\} \qquad (3.36)$$

where $\tilde{P}$ is an $\tilde{n} \times \tilde{n}$ matrix of Lagrange multipliers and λ is the scalar supplementary Lagrange multiplier, which is without loss of generality equal to 1 if the problem is *normal* (i.e., $\lambda = 0 \Rightarrow \tilde{P} = 0$) — see, for example, Ewing (1985, sec. 5.5). Since the conventional Lagrange multiplier problem involves a *vector* of constraints and hence a *vector* of Lagrange multipliers, the use of a Lagrange multiplier *matrix* requires some justification. The conventional problem with equality constraints takes the form,

$$\begin{aligned} &\text{minimize:} && f(x,y) \\ &\text{subject to:} && g(x) = 0 \end{aligned} \qquad (3.37)$$

where f is a scalar function of the vector variables x and y, and g is a vector function of x. Then the Lagrangian to be minimized becomes,

$$\mathcal{L}(x, y, p, \lambda) = \lambda f + p^{\mathrm{T}} g \qquad (3.38)$$

with p a column vector of Lagrange multipliers with dimensions equal to those of x. The matrix equality constraint in (3.35) may be rewritten in the form (3.38) by applying the "vec" operator to the equation. Define the matrix expression on the right hand side of the equality constraint in (3.35) as X. Then an equivalent to the constraint $X = 0$ is $\mathrm{vec}(X) = 0$, which yields the Lagrangian,

$$\mathcal{L}(A_c, B_c, C_c, \tilde{Q}, p, \lambda) = \lambda \ \mathrm{tr}\{\tilde{Q}\tilde{R}\} + p^{\mathrm{T}}\mathrm{vec}(X) \tag{3.39}$$

where p represents an $\tilde{n}^2 \times 1$ Lagrange multiplier vector. Of course, (3.39) is equivalent to (3.36). Applying the identity, $(\mathrm{vec}A^{\mathrm{T}})^{\mathrm{T}}\mathrm{vec}(B) = \mathrm{tr}(AB)$ (Graham 1981, p. 18, Example 1.4), to the term involving p, we see that (3.39) may be written as

$$\mathcal{L}(A_c, B_c, C_c, \tilde{Q}, \tilde{P}, \lambda) = \lambda \ \mathrm{tr}\{\tilde{Q}\tilde{R}\} + \mathrm{tr}\{\tilde{P}^{\mathrm{T}}X\} \tag{3.40}$$

where the $\tilde{n} \times \tilde{n}$ Lagrange multiplier matrix $\tilde{P}$ is defined by $p = \mathrm{vec}(\tilde{P})$. Clearly, (3.40) is equivalent to (3.36). Note that the symmetry of the expression defined as X means that we may assume without loss of generality that $\tilde{P}$ is symmetric.

In taking the partial derivatives of the Lagrangian, the following properties will be used (Graham 1981, pp. 76–78, Examples 5.4–5.6),

$$\frac{\partial \mathrm{tr}(AX)}{\partial X} = A^{\mathrm{T}}, \quad \frac{\partial \mathrm{tr}(AX^{\mathrm{T}})}{\partial X} = A, \quad \frac{\partial \mathrm{tr}(X^{\mathrm{T}}AXB)}{\partial X} = AXB + A^{\mathrm{T}}XB^{\mathrm{T}} \tag{3.41}$$

Taking the partial derivatives of the Lagrangian, (3.36), with respect to each of its arguments and setting each equal to zero:

1 $\frac{\partial \mathcal{L}}{\partial \tilde{Q}} = 0$:

$$\lambda\tilde{R} + \tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{P}\tilde{A} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i^{\mathrm{T}}\tilde{P}\tilde{A}_i = 0 \tag{3.42}$$

In order to test whether the optimization problem is normal, we set $\lambda = 0$ in (3.42). Taking the "vec" operation of the equation, as well, the result is

$$\left(\tilde{A}^{\mathrm{T}} \oplus \tilde{A}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i^{\mathrm{T}} \otimes \tilde{A}_i^{\mathrm{T}}\right) \mathrm{vec}\tilde{P} = 0 \tag{3.43}$$

$$\mathcal{A}^{\mathrm{T}} \ \mathrm{vec}\tilde{P} = 0$$

Since $\mathcal{A}$ is stable by assumption, $\mathcal{A}^{\mathrm{T}}$ has no nullspace, and therefore $\tilde{P} = 0$. As stated above, $\lambda = 0 \Rightarrow \tilde{P} = 0$ means that the problem is normal, and we can take $\lambda = 1$, without loss of generality. Therefore, with $\lambda = 1$,

$$\boxed{\tilde{R} + \tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{P}\tilde{A} + \sum_{i=1}^{p} \gamma_i \tilde{A}_i^{\mathrm{T}} \tilde{P} \tilde{A}_i = 0} \tag{3.44}$$

Before proceeding, partition the symmetric matrices, $\tilde{V}$, $\tilde{Q}$, and $\tilde{P}$, into 2×2 blocks (of dimensions $n \times n$, $n \times n_c$, $n_c \times n$, and $n_c \times n_c$):

$$\tilde{V} \triangleq \begin{bmatrix} V_1 & V_{12} B_c^{\mathrm{T}} \\ B_c V_{12}^{\mathrm{T}} & B_c V_2 B_c^{\mathrm{T}} \end{bmatrix}, \quad \tilde{Q} \triangleq \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^{\mathrm{T}} & Q_2 \end{bmatrix}, \quad \tilde{P} \triangleq \begin{bmatrix} P_1 & P_{12} \\ P_{12}^{\mathrm{T}} & P_2 \end{bmatrix} \tag{3.45}$$

Now the Lagrangian may be expanded:

$$\begin{aligned} \mathcal{L}(A_c, B_c, C_c, \tilde{Q}, \tilde{P}) = \mathrm{tr}\{ & Q_1 R_1 + Q_{12} C_c^{\mathrm{T}} R_{12}^{\mathrm{T}} + Q_{12}^{\mathrm{T}} R_{12} C_c + Q_2 C_c^{\mathrm{T}} R_2 C_c \\ & + (AQ_1 + BC_c Q_{12}^{\mathrm{T}}) P_1 + (AQ_{12} + BC_c Q_2) P_{12}^{\mathrm{T}} \\ & + (B_c C Q_1 + A_c Q_{12}^{\mathrm{T}}) P_{12} + (B_c C Q_{12} + A_c Q_2) P_2 \\ & + (Q_1 A^{\mathrm{T}} + Q_{12} C_c^{\mathrm{T}} B^{\mathrm{T}}) P_1 + (Q_1 C^{\mathrm{T}} B_c^{\mathrm{T}} + Q_{12} A_c^{\mathrm{T}}) P_{12}^{\mathrm{T}} \\ & + (Q_{12}^{\mathrm{T}} A^{\mathrm{T}} + Q_2 C_c^{\mathrm{T}} B^{\mathrm{T}}) P_{12} + (Q_{12}^{\mathrm{T}} C^{\mathrm{T}} B_c^{\mathrm{T}} + Q_2 A_c^{\mathrm{T}}) P_2 \\ & + \sum_{i=1}^{p} \gamma_i [(A_i Q_1 A_i^{\mathrm{T}} + B_i C_c Q_{12}^{\mathrm{T}} A_i^{\mathrm{T}} + A_i Q_{12} C_c^{\mathrm{T}} B_i^{\mathrm{T}} \\ & \qquad + B_i C_c Q_2 C_c^{\mathrm{T}} B_i^{\mathrm{T}}) P_1 \\ & \qquad + (A_i Q_1 C_i^{\mathrm{T}} B_c^{\mathrm{T}} + B_i C_c Q_{12}^{\mathrm{T}} C_i^{\mathrm{T}} B_c^{\mathrm{T}}) P_{12}^{\mathrm{T}} \\ & \qquad + (B_c C_i Q_1 A_i^{\mathrm{T}} + B_c C_i Q_{12} C_c^{\mathrm{T}} B_i^{\mathrm{T}}) P_{12} \\ & \qquad + (B_c C_i Q_1 C_i^{\mathrm{T}} B_c^{\mathrm{T}}) P_2] \\ & + V_1 P_1 + V_{12} B_c^{\mathrm{T}} P_{12}^{\mathrm{T}} + B_c V_{12}^{\mathrm{T}} P_{12} + B_c V_2 B_c^{\mathrm{T}} P_2 \} \end{aligned} \tag{3.46}$$

This allows us to take the partial derivatives of the Lagrangian with respect to the compensator parameters.

2 $\dfrac{\partial \mathcal{L}}{\partial A_c} = 0$:

$$P_{12}^{\mathrm{T}} Q_{12} + P_2 Q_2 + P_{12}^{\mathrm{T}} Q_{12} + P_2 Q_2 = 0 \tag{3.47}$$

$$\boxed{P_{12}^{\mathrm{T}} Q_{12} + P_2 Q_2 = 0} \tag{3.48}$$

3 $\frac{\partial \ell}{\partial B_c} = 0$:

$$P_{12}{}^{\mathrm{T}}Q_1C^{\mathrm{T}} + P_2Q_{12}{}^{\mathrm{T}}C^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}Q_1C^{\mathrm{T}} + P_2Q_{12}{}^{\mathrm{T}}C^{\mathrm{T}} \tag{3.49}$$

$$+\sum_{i=1}^{p}\gamma_i(P_{12}{}^{\mathrm{T}}A_iQ_1C_i{}^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}B_iC_cQ_{12}{}^{\mathrm{T}}C_i{}^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}A_iQ_1C_i{}^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}B_iC_cQ_{12}{}^{\mathrm{T}}C_i{}^{\mathrm{T}} + 2P_2B_cC_iQ_1C_i{}^{\mathrm{T}})$$

$$+ P_{12}{}^{\mathrm{T}}V_{12} + P_{12}{}^{\mathrm{T}}V_{12} + 2P_2B_cV_2 = 0$$

$$(P_{12}{}^{\mathrm{T}}Q_1 + P_2Q_{12}{}^{\mathrm{T}})C^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}V_{12} + P_2B_cV_2 \tag{3.50}$$

$$+\sum_{i=1}^{p}\gamma_i(P_{12}{}^{\mathrm{T}}A_iQ_1 + P_{12}{}^{\mathrm{T}}B_iC_cQ_{12}{}^{\mathrm{T}} + P_2B_cC_iQ_1)C_i{}^{\mathrm{T}} = 0$$

If for each i, B_i and C_i are not both nonzero (i.e., the Wiener processes used for control and measurement dependent noise are independent), a closed form expression for B_c can be found. For then we have,

$$\boxed{\begin{aligned}&(P_{12}{}^{\mathrm{T}}Q_1 + P_2Q_{12}{}^{\mathrm{T}})C^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}V_{12} + P_2B_cV_2 \\ &\quad +\sum_{i=1}^{p}\gamma_i(P_{12}{}^{\mathrm{T}}A_iQ_1 + P_2B_cC_iQ_1)C_i{}^{\mathrm{T}} = 0\end{aligned}} \tag{3.51}$$

which leads to,

$$B_c = -P_2{}^{-1}[(P_{12}{}^{\mathrm{T}}Q_1 + P_2Q_{12}{}^{\mathrm{T}})C^{\mathrm{T}} + P_{12}{}^{\mathrm{T}}(V_{12} + \sum_{i=1}^{p}\gamma_iA_iQ_1C_i{}^{\mathrm{T}})]\widehat{V}_2{}^{-1} \tag{3.52}$$

where

$$\widehat{V}_2 \triangleq V_2 + \sum_{i=1}^{p}\gamma_iC_iQ_1C_i{}^{\mathrm{T}} \tag{3.53}$$

4 $\frac{\partial \mathcal{L}}{\partial C_c} = 0$:

$$R_{12}{}^{\mathrm{T}}Q_{12} + R_{12}{}^{\mathrm{T}}Q_{12} + 2R_2C_cQ_2 \tag{3.54}$$

$$+ B^{\mathrm{T}}P_1Q_{12} + B^{\mathrm{T}}P_{12}Q_2 + B^{\mathrm{T}}P_1Q_{12} + B^{\mathrm{T}}P_{12}Q_2$$

$$+ \sum_{i=1}^{p} \gamma_i (B_i{}^{\mathrm{T}}P_1A_iQ_{12} + B_i{}^{\mathrm{T}}P_1A_iQ_{12} + 2B_i{}^{\mathrm{T}}P_1B_iC_cQ_2$$

$$+ B_i{}^{\mathrm{T}}P_{12}B_cC_iQ_{12} + B_i{}^{\mathrm{T}}P_{12}B_cC_iQ_{12}) = 0$$

$$R_{12}{}^{\mathrm{T}}Q_{12} + R_2C_cQ_2 + B^{\mathrm{T}}(P_1Q_{12} + P_{12}Q_2) \tag{3.55}$$

$$+ \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}}(P_1A_iQ_{12} + P_1B_iC_cQ_2 + P_{12}B_cC_iQ_{12}) = 0$$

Assuming as before that for each i, B_i and C_i are not both nonzero,

$$\boxed{\begin{aligned} &R_{12}{}^{\mathrm{T}}Q_{12} + R_2C_cQ_2 + B^{\mathrm{T}}(P_1Q_{12} + P_{12}Q_2) \\ &\quad + \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}}(P_1A_iQ_{12} + P_1B_iC_cQ_2) = 0 \end{aligned}} \tag{3.56}$$

Then we find that

$$C_c = -\hat{R}_2{}^{-1}[B^{\mathrm{T}}(P_1Q_{12} + P_{12}Q_2) + (R_{12}{}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}}P_1A_i)Q_{12}]Q_2{}^{-1} \tag{3.57}$$

where

$$\hat{R}_2 \triangleq R_2 + \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}}P_1B_i \tag{3.58}$$

The equality constraint in (3.35) and the framed equations — (3.44), (3.48), (3.51), and (3.56) — are equivalent to equations (20), (71), (74), (75), and (76), respectively, in (Bernstein and Hyland 1988a). There, the derivation of the first-order

necessary conditions is completed. Essentially, it involves: (1) proving the existence of the inverses in (3.52) and (3.57); (2) expanding both (3.44) and the equality constraint in (3.35) into four blocks; (3) substituting the expressions for B_c and C_c into those expanded equations; and (4) using algebraic manipulations to solve for an expression for A_c and to find equations for the unknowns, $\tilde{Q}$ and $\tilde{R}$, independent of A_c, B_c, and C_c. The necessary conditions in final form are:

$$0 = AQ + QA^{\mathrm{T}} - \mathcal{Q}\widehat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}} + V_1 + \sum_{i=1}^{p} \gamma_i [A_i Q A_i^{\mathrm{T}} + (A_i - B_i \widehat{R}_2^{-1}\mathcal{P})\widehat{Q}(A_i - B_i \widehat{R}_2^{-1}\mathcal{P})^{\mathrm{T}}] \tag{3.59}$$

$$0 = A^{\mathrm{T}}P + PA - \mathcal{P}^{\mathrm{T}}\widehat{R}_2^{-1}\mathcal{P} + R_1 + \sum_{i=1}^{p} \gamma_i [A_i^{\mathrm{T}} P A_i + (A_i - \mathcal{Q}\widehat{V}_2^{-1}C_i)^{\mathrm{T}}\widehat{P}(A_i - \mathcal{Q}\widehat{V}_2^{-1}C_i)] \tag{3.60}$$

$$0 = A_P \widehat{Q} + \widehat{Q}A_P^{\mathrm{T}} + \mathcal{Q}\widehat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}} \tag{3.61}$$

$$0 = A_Q^{\mathrm{T}}\widehat{P} + \widehat{P}A_Q + \mathcal{P}^{\mathrm{T}}\widehat{R}_2^{-1}\mathcal{P} \tag{3.62}$$

where

$$\mathcal{Q} \triangleq QC^{\mathrm{T}} + V_{12} + \sum_{i=1}^{p} \gamma_i A_i (Q + \widehat{Q}) C_i^{\mathrm{T}} \tag{3.63}$$

$$\mathcal{P} \triangleq B^{\mathrm{T}}P + R_{12}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i B_i^{\mathrm{T}} (P + \widehat{P}) A_i \tag{3.64}$$

$$A_Q \triangleq A - \mathcal{Q}\widehat{V}_2^{-1}C, \qquad A_P \triangleq A - B\widehat{R}_2^{-1}\mathcal{P} \tag{3.65}$$

and A_c, B_c, and C_c are rewritten in terms of the unknown variables Q, P, $\widehat{Q}$, and $\widehat{P}$ as:

$$A_c = A - B\hat{R}_2^{-1}\mathcal{P} - \mathcal{Q}\hat{V}_2^{-1}C \tag{3.66}$$

$$B_c = \mathcal{Q}\hat{V}_2^{-1} \tag{3.67}$$

$$C_c = -\hat{R}_2^{-1}\mathcal{P} \tag{3.68}$$

3.5 Modifications for Stratonovich Noise Model

As mentioned in section 3.2, there is a simple transformation which allows us to reinterpret the Itô equation (3.7b) — that is, our state equation — in the sense of Stratonovich. That transformation is given by the following theorem.

Theorem 3.3: Given the vector Stratonovich equation,

$$dx_t = f(x_t,t)dt + G(x_t,t)d\beta_t\,, \quad t \geq t_0 \tag{3.69}$$

where x_t has dimensions $n \times 1$ and β_t is a vector Wiener process with $\mathbb{E}\{d\beta_t d\beta_t^{\mathrm{T}}\} = Q(t)dt$, the equivalent Itô equation for the i^{th} element of dx_t is

$$dx_{it} = \left\{ f_i(x_t,t) + \tfrac{1}{2}\sum_{k=1}^{n}\left[G(x_t,t)V\frac{\partial G(x_t,t)}{\partial x_{it}}\right]_{ki}\right\}dt + [G(x_t,t)d\beta_t]_i\,, \quad t \geq t_0 \tag{3.70}$$

This theorem is given in Jazwinski (1970, p. 131) and is formally proved by Stratonovich (1966). It shows that a correction term must be added to the first term in the state equations. Comparing (3.69) with (3.7b),

$$d\tilde{x}_t = \tilde{A}\tilde{x}_t dt + \left[\tilde{A}_1\tilde{x}_t \;\; \tilde{A}_2\tilde{x}_t \;\; \ldots \;\; \tilde{A}_p\tilde{x}_t \;\; \tilde{G}\right]\begin{bmatrix} dv_t \\ d\beta_t \end{bmatrix} \tag{3.7b}$$

and recalling equations (3.20),

$$f(\tilde{x}_t) = \tilde{A}\tilde{x}_t, \qquad G(\tilde{x}_t) = \left[\tilde{A}_1\tilde{x}_t \;\; \tilde{A}_2\tilde{x}_t \;\; \ldots \;\; \tilde{A}_p\tilde{x}_t \;\; \tilde{G}\right],$$

$$Q = \mathrm{diag}\{\gamma_1, \gamma_2, \ldots, \gamma_p, V\} \tag{3.20}$$

we can substitute the quantities from (3.20) into (3.70) and show that in order to give the multiplicative white noise model a Stratonovich interpretation, we need to make the following substitution,

$$\tilde{A}\tilde{x}_t \leftarrow \tilde{A}\tilde{x}_t + \tfrac{1}{2}\sum_{i=1}^{p}\gamma_i\tilde{A}_i^2\,\tilde{x}_t \tag{3.71}$$

Eliminating the common factor, $\tilde{x}_t$, and expanding $\tilde{A}$ and $\tilde{A}_i$ into 2×2 block form [see (3.9)], we have

$$\begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix} \leftarrow \begin{bmatrix} A & BC_c \\ B_cC & A_c \end{bmatrix} + \tfrac{1}{2}\sum_{i=1}^{p}\gamma_i \begin{bmatrix} A_i^2 + B_iC_cB_cC_i & A_iB_iC_c \\ B_cC_iA_i & B_cC_iB_iC_c \end{bmatrix} \tag{3.72}$$

Then employing once again the assumption that for each i, B_i and C_i are not both nonzero, and examining the (1,1), (1,2), and (2,1) blocks of (3.72), we find that the Stratonovich noise interpretation may be accomplished by means of the three simple substitutions in the plant model,

$$\boxed{\begin{aligned} A &\leftarrow A + \tfrac{1}{2}\sum_{i=1}^{p}\gamma_iA_i^2\,, \\ B &\leftarrow B + \tfrac{1}{2}\sum_{i=1}^{p}\gamma_iA_iB_i\,, \\ C &\leftarrow C + \tfrac{1}{2}\sum_{i=1}^{p}\gamma_iC_iA_i \end{aligned}} \tag{3.73}$$

3.6 Controller Design Summary

The multiplicative white noise model of this chapter is used for the auxiliary inputs of the LQG/PRE error model. This auxiliary input model is combined with the standard cost functional penalties on the auxiliary outputs to produce a new parameter-robust design procedure. The performance/robustness tradeoff is accomplished similarly to LQG/PRE, where the noise intensities are adjusted by means of the scalar parameters γ_i $(i = 1,2,\ldots,p)$, and auxiliary output penalties are adjusted as before, by

modifying the state and control weighting matrices (R_1 and R_2, respectively) as in equation (1.9) and substituting these modified matrices into the controller design equations developed in this chapter.

4. Optimal Reduced-Order Control

One of the great advantages of the parameter robust design techniques discussed in Chapters 2 and 3 (as well as LQG/PRE) over μ-synthesis (see section 1.1B) is the fact that they lend themselves to application of a method of optimal reduced-order controller design, called the *optimal projection equations* (Hyland and Bernstein 1984). As is well-known, the LQG-optimal compensator for an n^{th}-order plant model is also of order n. However, the optimal projection equations allow us to specify a compensator order $n_c < n$ and directly design the optimal compensator of that order, provided a stabilizing $n_c{}^{\text{th}}$-order compensator exists. The optimal projection results will be stated here directly. For a proof, see Hyland and Bernstein (1984) [for the case of no cross-weighting in the cost and no cross-covariance between process and sensor noise], or see Bernstein and Hyland (1988a) for the more general case.

Given the n^{th}-order plant,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + G_1 w(t) \\ y(t) &= Cx(t) + G_2 w(t) \end{aligned} \tag{4.1}$$

with uncorrelated, unit intensity, Gaussian white noise vector, w, define the covariance matrices,

$$V_1 \triangleq G_1 G_1{}^{\mathrm{T}}, \quad V_{12} \triangleq G_1 G_2{}^{\mathrm{T}}, \quad V_2 \triangleq G_2 G_2{}^{\mathrm{T}} \tag{4.2}$$

and the linear-quadratic cost functional,

$$J(A_c, B_c, C_c) = \lim_{T \to \infty} \mathbb{E}\{x_t{}^{\mathrm{T}} R_1 x_t + 2 x_t{}^{\mathrm{T}} R_{12} u_t + u_t{}^{\mathrm{T}} R_2 u_t \, \mathrm{dt}\} \tag{4.3}$$

If a stabilizing $n_c{}^{\text{th}}$-order compensator,

$$\begin{aligned} x_c(t) &= A_c x_c(t) + B_c y(t) \\ u(t) &= C_c x_c(t) \end{aligned} \tag{4.4}$$

exists, the one which minimizes (4.3) is given by the design equations,

$$A_c = \Gamma(A - BR_2^{-1}\mathcal{P} - \mathcal{Q}V_2^{-1}C)G^T \tag{4.5}$$

$$B_c = \Gamma\mathcal{Q}V_2^{-1}$$

$$C_c = -R_2^{-1}\mathcal{P}G^T$$

where

$$\mathcal{Q} \triangleq QC^T + V_{12}\,, \quad \mathcal{P} \triangleq B^TP + R_{12}^T \tag{4.6}$$

Also define

$$A_Q \triangleq A - \mathcal{Q}V_2^{-1}C, \quad A_P \triangleq A - BR_2^{-1}\mathcal{P} \tag{3.65}$$

Then the $n \times n$ matrices Q, P, $\hat{Q}$, $\hat{P}$, and $\tau_\perp$ are the solution to the optimal projection equations,

$$0 = AQ + QA^T - \mathcal{Q}V_2^{-1}\mathcal{Q}^T + V_1 + \tau_\perp \mathcal{Q}V_2^{-1}\mathcal{Q}^T\tau_\perp^T \tag{4.7}$$

$$0 = A^TP + PA - \mathcal{P}^TR_2^{-1}\mathcal{P} + R_1 + \tau_\perp^T\mathcal{P}^TR_2^{-1}\mathcal{P}\tau_\perp \tag{4.8}$$

$$0 = A_P\hat{Q} + \hat{Q}A_P^T + \mathcal{Q}V_2^{-1}\mathcal{Q}^T - \tau_\perp\mathcal{Q}V_2^{-1}\mathcal{Q}^T\tau_\perp^T \tag{4.9}$$

$$0 = A_Q^T\hat{P} + \hat{P}A_Q + \mathcal{P}^TR_2^{-1}\mathcal{P} - \tau_\perp^T\mathcal{P}^TR_2^{-1}\mathcal{P}\tau_\perp \tag{4.10}$$

$$\text{rank}(\hat{Q}) = \text{rank}(\hat{P}) = \text{rank}(\hat{Q}\hat{P}) = n_c \tag{4.11}$$

where $\tau_\perp \triangleq I - \tau$ and the optimal projection matrix, τ, is perhaps most simply described as follows. Define a *balancing transformation* Ψ (Laub 1980) that simultaneously diagonalizes $\hat{Q}$ and $\hat{P}$ by,

$$\Psi^{-1}\hat{Q}\Psi^{-1} = \Psi^T\hat{P}\Psi = \Lambda \tag{4.12}$$

$$\Lambda \triangleq \text{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_{n_c},0,\ldots,0\}$$

Then

$$\tau = \Psi \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} \Psi^{-1} \tag{4.13}$$

Also,

$$G = [I_{n_c} \quad 0]\Psi^{T} \quad \text{and} \quad \Gamma = [I_{n_c} \quad 0]\Psi^{-1} \tag{4.14}$$

Note that the modified Riccati equations (4.7)–(4.8) and the modified Lyapunov equations (4.9)–(4.10) are coupled together by the matrix $\tau_{\perp}$. For the case of a full-order compensator ($n_c = n$), we have $\tau = G = \Gamma = I$ and $\tau_{\perp} = 0$, so the terms involving $\tau_{\perp}$ disappear. In that case (4.7) and (4.8) become the standard observer and regulator Riccati equations of LQG control, and equations (4.5) become the standard state-space solution for the controller. Also, in the full-order case, the Lyapunov equations (4.9) and (4.10) become superfluous. Notably, these equations must be solved for the observability gramian, $\hat{Q}$, and the controllability gramian, $\hat{P}$, if one wishes to balance and truncate the plant to produce a suboptimal reduced-order controller. The coupled structure of the optimal projection equations shows that, in a sense, the balancing and controller design must be carried out simultaneously in order to preserve optimality. In the general case, where possibly $n_c < n$, $\hat{Q}$ and $\hat{P}$ are referred to in the literature as the observability and controllability *pseudogramians*.

Clearly, the optimal reduced-order design equations apply to LQG/PRE, since LQG/PRE differs from standard LQG only in the fact that the values of V_1, V_2, R_1, and R_2 are modified, as in (1.9). Optimal projection may also easily be applied to the frequency-domain method of Chapter 2. The auxiliary input modeling phase of that method merely involves augmentation of the plant dynamics and augmentation of the covariance matrices V_1 and V_2 to account for additional noise sources. The auxiliary output modeling phase only modifies the weighting matrices R_1 and R_2. Therefore, the modified model of the plant, the modified noise model, and the modified cost criterion still constitute a standard LQG problem. That being the case, the optimal projection equations may be applied directly.

The time-domain method of Chapter 3, however, does not conform to the standard LQG framework, because the multiplicative white noise model of the auxiliary

inputs requires a change in the structure of the state and/or output equations of the plant. Fortunately, optimal projection may still be applied. The optimal projection equations, as stated above, may be derived by the Lagrange multiplier method of section 3.4 (Hyland and Bernstein 1984). More specifically, the rank conditions (4.11) are enforced in the course of the algebraic manipulations which follow after taking the partial derivatives of the Lagrangian. In fact, the same rank conditions may be applied to the more general multiplicative white noise problem (Bernstein and Hyland 1988a). These rank conditions did not appear explicitly in Chapter 3, because the algebraic manipulations were referenced to the paper just mentioned. Also, Chapter 3 was concerned only with the case of a full-order controller. Of course, the full-order case is still important, since a restriction, $n_c < n$, on the controller order generally results in an increase in the value of the cost functional, J, and therefore brings about a degradation in performance.

By including multiplicative white noise in the state and output equations of the plant, as in Chapter 3,

$$dx_t = Ax_t dt + \sum_{i=1}^{p} \gamma_i A_i x_t dv_{it} + Bu_t dt + \sum_{i=1}^{p} \gamma_i B_i u_t dv_{it} + G_1 d\beta_t \tag{3.1}$$

$$dy_t = Cx_t dt + \sum_{i=1}^{p} \gamma_i C_i x_t dv_{it} + G_2 d\beta_t \tag{3.2}$$

we obtain a general model which applies to all of the parameter-robust controller design methods discussed in this work. For a method in which multiplicative white noise is not desired, we may set $A_i = B_i = C_i = 0$ in (3.1)–(3.2) to obtain (4.1). The optimal $n_c{}^{\text{th}}$-order compensator (for $n_c \leq n$), if a stabilizing compensator of that order exists, is given by

$$\begin{aligned} A_c &= \Gamma(A - B\hat{R}_2{}^{-1}\mathcal{P} - \mathcal{Q}\hat{V}_2{}^{-1}C)G^{\mathrm{T}} \\ B_c &= \Gamma\mathcal{Q}\hat{V}_2{}^{-1} \\ C_c &= -\hat{R}_2{}^{-1}\mathcal{P}G^{\mathrm{T}} \end{aligned} \tag{4.15}$$

where Q, P, $\hat{Q}$, $\hat{P}$, and $\tau_\perp$ are the solution to the more general optimal projection equations,

$$0 = AQ + QA^{\mathrm{T}} - \mathcal{Q}\hat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}} + V_1 + \tau_\perp \mathcal{Q}\hat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}}\tau_\perp^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i [A_i Q A_i^{\mathrm{T}} + (A_i - B_i \hat{R}_2^{-1}\mathcal{P})\hat{Q}(A_i - B_i \hat{R}_2^{-1}\mathcal{P})^{\mathrm{T}}] \tag{4.16}$$

$$0 = A^{\mathrm{T}}P + PA - \mathcal{P}^{\mathrm{T}}\hat{R}_2^{-1}\mathcal{P} + R_1 + \tau_\perp{}^{\mathrm{T}}\mathcal{P}^{\mathrm{T}}\hat{R}_2^{-1}\mathcal{P}\tau_\perp + \sum_{i=1}^{p} \gamma_i [A_i{}^{\mathrm{T}} P A_i + (A_i - \mathcal{Q}\hat{V}_2^{-1}C_i)^{\mathrm{T}}\hat{P}(A_i - \mathcal{Q}\hat{V}_2^{-1}C_i)] \tag{4.17}$$

$$0 = A_P\hat{Q} + \hat{Q}A_P^{\mathrm{T}} + \mathcal{Q}\hat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}} - \tau_\perp \mathcal{Q}\hat{V}_2^{-1}\mathcal{Q}^{\mathrm{T}}\tau_\perp^{\mathrm{T}} \tag{4.18}$$

$$0 = A_Q{}^{\mathrm{T}}\hat{P} + \hat{P}A_Q + \mathcal{P}^{\mathrm{T}}\hat{R}_2^{-1}\mathcal{P} - \tau_\perp{}^{\mathrm{T}}\mathcal{P}^{\mathrm{T}}\hat{R}_2^{-1}\mathcal{P}\tau_\perp \tag{4.19}$$

$$\mathrm{rank}(\hat{Q}) = \mathrm{rank}(\hat{P}) = \mathrm{rank}(\hat{Q}\hat{P}) = n_c \tag{4.20}$$

The $n_c \times n$ matrices Γ and G, and the $n \times n$ matrix $\tau_\perp$ are determined by $\hat{Q}$ and $\hat{P}$ as described above, and the following definitions from Chapter 3 apply.

$$\hat{V}_2 \triangleq V_2 + \sum_{i=1}^{p} \gamma_i C_i Q_1 C_i^{\mathrm{T}} \tag{3.53}$$

$$\hat{R}_2 \triangleq R_2 + \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}} P_1 B_i \tag{3.58}$$

$$\mathcal{Q} \triangleq QC^{\mathrm{T}} + V_{12} + \sum_{i=1}^{p} \gamma_i A_i (Q + \hat{Q}) C_i{}^{\mathrm{T}} \tag{3.63}$$

$$\mathcal{P} \triangleq B^{\mathrm{T}}P + R_{12}{}^{\mathrm{T}} + \sum_{i=1}^{p} \gamma_i B_i{}^{\mathrm{T}}(P + \hat{P})A_i \tag{3.64}$$

$$A_Q \triangleq A - \mathcal{Q}\hat{V}_2{}^{-1}C, \qquad A_P \triangleq A - B\hat{R}_2{}^{-1}\mathcal{P} \tag{3.65}$$

The numerical solution of the coupled Riccati- and Lyapunov-type equations in this and the previous chapter require an iterative algorithm. The algorithms used to solve these equations will be discussed in Chapter 5.

5. Methods and Algorithms for Controller Design

Chapters 2, 3, and 4 described the basic theory needed for the controller design techniques under study. This chapter is concerned with the practical problems involved in applying the design methods to an actual control system. In the sections that follow, a few of the more important implementation considerations are discussed which arose over the course of this research. Solutions are given to these problems, and algorithms are described in detail.

5.1 Selection of Cost Functional Weighting Matrices and Covariance Matrices

When relatively little is known about the physical meaning of the dynamics of a control system's state-space model, it is common practice to oversimplify the cost and covariance matrices for LQG controller designs. One method is to assign scalar-weighted identity matrices, as follows:

$$R_1 = I, \quad R_2 = \rho_r I, \quad R_{12} = 0 \tag{5.1}$$
$$\mathbb{E}\{vv^{\mathrm{T}}\} = I, \quad \mathbb{E}\{nn^{\mathrm{T}}\} = \rho_v I, \quad \mathbb{E}\{vn^{\mathrm{T}}\} = 0$$

where v and n represent the process and sensor noise vectors, respectively. The idea, then, is to reduce the LQG design process to the selection of two scalar parameters — ρ_r to adjust controller authority, and ρ_v to adjust the tracking speed of the Kalman filter. Although it is generally not possible to know the cost matrices which will give the most desirable response or the covariance matrices which most accurately describe the noise, we may take advantage of what knowledge we have about the simply supported plate under study to take a more meaningful design approach.

A large class of damped flexible structures, including those modeled by finite-element methods, can be expressed in *spatial coordinates* by the following differential equation:

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = f(t) \tag{5.2}$$

where q denotes the n-dimensional spatial displacement vector and f denotes the

n-dimensional applied force vector. M, C, and K are symmetric *mass, damping,* and *stiffness* matrices, respectively. Assuming the structure has *proportional damping* (i.e., $C = k_1 M + k_2 K$, for some scalar constants k_1 and k_2), the matrices M, C, and K can be simultaneously diagonalized by a left and right multiplication. In other words, the system may be decoupled into *modal coordinates* by means of a variable substitution, so that

$$m_i \ddot{x}_i(t) + c_i \dot{x}_i(t) + k_i x_i(t) = f_i(t), \quad i = 1, 2, \ldots, n \tag{5.3}$$

where x_i represents the "displacement" of the i^{th} mode. The proportional damping assumption tends to hold, for all practical purposes, when the damping is very light, as is the case for many flexible structures and in particular for the one studied in this work. Also, the identification procedure used to obtain a model of the simply supported plate assumes a modal system. In terms of modal natural frequencies and damping ratios, ω_i and ζ_i, respectively, an equivalent expression for (5.3) is

$$\ddot{x}_i(t) + 2\zeta_i \omega_i \dot{x}_i(t) + \omega_i^2 x_i(t) = u_i(t), \quad i = 1, 2, \ldots, n \tag{5.4}$$

Then the modal frequencies are expressed in terms of the diagonalized mass and stiffness matrices by

$$\omega_i^2 = \frac{k_i}{m_i} \tag{5.5}$$

The energy contained in each mode is the sum of the modal potential and kinetic energies,

$$E_i = \tfrac{1}{2} k_i x_i^2 + \tfrac{1}{2} m_i \dot{x}_i^2 \tag{5.6}$$

Therefore,

$$E_i \propto \omega_i^2 x_i^2 + \dot{x}_i^2 \tag{5.7}$$

Rather than penalizing all of the states equally, as in (5.1), it seems reasonable to penalize the total energy in the system. Then the controls may be penalized as before,

or individually, to balance the tradeoff between modal energy and control effort. If we order the displacements and velocities into a state vector, $\overline{x} \triangleq [x^{\mathrm{T}} \ \dot{x}^{\mathrm{T}}]^{\mathrm{T}}$, this is accomplished by the cost matrices,

$$R_1 = \begin{bmatrix} \Omega^2 & 0 \\ 0 & I \end{bmatrix}, \quad R_2 = \mathrm{diag}\{\rho_{r_1}, \rho_{r_2}, \ldots \rho_{r_m}\}, \quad R_{12} = 0 \tag{5.8}$$

for a system with m controls, where $\Omega \triangleq \mathrm{diag}\{\omega_1, \omega_2, \ldots, \omega_n\}$. For the system under study there is only one control input, so this technique leaves only one parameter to adjust. In section 5.2 it will be shown how the cross-weighting matrix, R_{12}, may be manipulated to suit other design goals.

The covariance matrices can also be given a more sensible structure, although the modifications involved may not be significant enough to gain much advantage. For controller design purposes the state and output equations for the plant are expressed in modal coordinates:

$$\begin{aligned} \dot{\overline{x}} &= A\overline{x} + Bu + G_1 w \triangleq A\overline{x} + Bu + [g_{11} \ \ g_{12}] \begin{bmatrix} v \\ n \end{bmatrix} \\ y &= C\overline{x} + Du + G_2 w \triangleq C\overline{x} + Du + [g_{21} \ \ g_{22}] \begin{bmatrix} v \\ n \end{bmatrix} \end{aligned} \tag{5.9}$$

It is assumed here that the process and sensor noise are independent of each other. Hence, we have partitioned the exogenous noise vector w into process noise v and sensor noise n. Since sensor noise would only occur in the output equation, $g_{12} = 0$. Also, w is assumed normalized by G_1 and G_2 such that $\mathbb{E}\{ww^{\mathrm{T}}\} = I$.

One might arbitrarily assume that the elements of n each affect one modal output in y and to the same degree, so that $g_{22} = \rho_v{}^{1/2} I$. However, any judgment about the relative intensities and distribution of the sensor noise should be made in *sensor coordinates.* Denote the sensor measurements as a sum of a true signal and sensor noise, as follows:

$$y_{\mathrm{meas}} = y_s + g_s n_s \tag{5.10}$$

The relationship between the sensor outputs and the modal outputs is given by

$$y_{\mathrm{meas}} = \Phi y \tag{5.11}$$

where the *sensor eigenvector* (or *modeshape*) *matrix* Φ has at least as many rows as columns and indicates the relative participation each modal output has on each sensor. The output equation in modal coordinates is approximated by the least squares solution,

$$y = (\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}y_{\mathrm{meas}} = \Phi^{\#}y_s + \Phi^{\#}g_s n_s \tag{5.12}$$

where $\Phi^{\#}$ is the least squares pseudoinverse of Φ, or simply the inverse if Φ is square. Comparison of (5.12) with the output equation in (5.9) reveals that simply setting $g_{22} = {\rho_v}^{1/2} I$ ignores the structure of the modeshape matrix. Instead, for $\mathbb{E}\{n_s n_s\} \triangleq I$ we might reasonably set $g_s = {\rho_v}^{1/2} I$ if all of the sensors are identical and function independently of one another. That gives us

$$g_{22} n = \Phi^{\#} g_s n_s = {\rho_v}^{1/2} \Phi^{\#} n_s \tag{5.13}$$

The original sensor noise vector n should be replaced by the possibly higher dimensional vector n_s, giving

$$g_{22} = {\rho_v}^{1/2} \Phi^{\#} \tag{5.14a}$$

but that does not affect the dimension of $V_2 \triangleq G_2 {G_2}^{\mathrm{T}} = g_{21} {g_{21}}^{\mathrm{T}} + g_{22} {g_{22}}^{\mathrm{T}}$. Therefore,

$$V_2 = g_{21} {g_{21}}^{\mathrm{T}} + \rho_v \Phi^{\#} {\Phi^{\#}}^{\mathrm{T}} = g_{21} {g_{21}}^{\mathrm{T}} + \rho_v (\Phi^{\mathrm{T}}\Phi)^{-1} \tag{5.15}$$

Note that the same result for V_2 may be arrived at by replacing (5.14a) with

$$g_{22} = {\rho_v}^{1/2} (\Phi^{\mathrm{T}}\Phi)^{-1/2} \tag{5.14b}$$

This alteration has no effect on $V_{12} \triangleq G_1 {G_2}^{\mathrm{T}}$ either because, as was already indicated, $g_{12} = 0$. Using (5.14b) allows us to retain the lower dimensional sensor noise vector n.

5.2 Disturbance Cancellation and Inverse Optimal Control

When the process noise disturbance is expected to lie in some limited frequency range, it is a highly inefficient allocation of control effort to attempt to reject that disturbance by merely lowering the control penalty, ρ_r, and sensor noise covariance, ρ_v, until the desired degree of rejection is achieved. That procedure results in greater disturbance rejection at all frequencies and potentially requires an unacceptable amount of control effort (as well as excessive sensitivity to sensor noise) to achieve the specified level of rejection in the anticipated disturbance frequency band. Rather than raise controller authority at all frequencies, it is more desirable to concentrate the control effort where it is needed.

Suppose we have the following plant and colored disturbance model, respectively:

$$\begin{aligned} \dot{x}_p &= Ax_p + Bu + G_w w \\ y &= Cx_p + Du + g_{21}w + g_{22}n \end{aligned} \tag{5.16}$$

$$\begin{aligned} \dot{x}_w &= A_w x_w + B_w v \\ w &= C_w x_w \end{aligned} \tag{5.17}$$

Augmenting the plant with the shaping filter dynamics yields

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} x_p \\ x_w \end{bmatrix} &= \begin{bmatrix} A & G_w C_w \\ 0 & A_w \end{bmatrix}\begin{bmatrix} x_p \\ x_w \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ B_w \end{bmatrix} v \\ y &= \begin{bmatrix} C & g_{21}C_w \end{bmatrix}\begin{bmatrix} x_p \\ x_w \end{bmatrix} + Du + g_{22}n \end{aligned} \tag{5.18}$$

If we assume complete knowledge of the augmented state, the feedback

$$u = -Fx \triangleq -[F_p \quad F_w]\begin{bmatrix} x_p \\ x_w \end{bmatrix} \tag{5.19}$$

produces the closed-loop system and output matrices,

$$A_{cl} = \begin{bmatrix} A - BF_p & G_w C_w - BF_w \\ 0 & A_w \end{bmatrix} \tag{5.20}$$

$$C_{cl} = \begin{bmatrix} C - DF_p & g_{21} C_w - DF_w \end{bmatrix}$$

If g_{21} is nonzero, there are two components of the disturbance — one which enters through G_w and influences the states, and the other which enters by way of g_{21} and feeds through to the outputs. Both components are completely canceled if F_w is chosen such that

$$G_w C_w - BF_w = 0 \tag{5.21a}$$

$$g_{21} C_w - DF_w = 0 \tag{5.21b}$$

Both conditions cannot in general be met simultaneously. However, the type of problem under study has a favorable structure in this regard. For the n_m-mode flexible structure model studied in chapters 6 and 7, the control input and disturbance are forces and affect only the derivatives of the velocity states. Also, the measurements are modal accelerations. Therefore, the matrices B, G_w , D, and g_{21} take the form,

$$B = \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad G_w = \begin{bmatrix} 0 \\ g \end{bmatrix} \tag{5.22}$$

$$D = b, \quad g_{21} = g$$

and both (5.21a) and (5.21b) reduce to:

$$\boxed{bF_w = gC_w} \tag{5.23}$$

A solution, F_w, that satisfies (5.23) cannot be found in general unless the $n_m \times m$ matrix b has rank n_m. This condition requires that the number of actuators be at least as great as the number of modes in the plant model (i.e., $m \geq n_m$). If $m < n_m$ it may not be possible to completely cancel the disturbance, but we can still choose F_w to cancel the effect of the disturbance on m of the modes by selecting the corresponding m rows of b and g to replace the full matrices in (5.23).

If the state feedback (5.19) must pass through a low pass smoothing filter before entering the plant through B, the above technique may still be applied, but it is no longer possible to cancel the disturbance at all frequencies simultaneously. Define the smoothing filter transfer function, $H_l(s) = C_l(sI - A_l)^{-1}B_l$. Then the condition on F_w which produces exact cancellation of the disturbance at frequency ω_d is:

$$bC_l(j\omega_d I - A_l)^{-1}B_l F_w x_w(j\omega_d) = gC_w x_w(j\omega_d) \tag{5.24}$$

The vector $x_w(j\omega_d)$ is found by applying a sinusoid of frequency ω_d to the input of the disturbance shaping filter (5.17) and measuring the response of x_w in the frequency domain. Assuming v is a scalar, as is the case in our single disturbance input model, we set

$$v(t) = \sin(\omega_d t) \tag{5.25}$$

Then, after taking the Fourier transform of (5.17) and solving for x_w we have

$$x(j\omega_d) = (j\omega_d I - A_w)^{-1}B_w v(j\omega_d) \tag{5.26}$$

Substituting (5.26) into (5.24) and eliminating the scalar $v(j\omega_d)$ from both sides of the equation, we arrive at:

$$\boxed{bC_l(j\omega_d I - A_l)^{-1}B_l F_w(j\omega_d I - A_w)^{-1}B_w = gC_w(j\omega_d I - A_w)^{-1}B_w} \tag{5.27}$$

When estimator-based feedback is necessary, infinite disturbance rejection is no longer possible, but the disturbance cancellation technique may still be used to provide good rejection if the estimates of the disturbance states are accurate. Therefore, the effectiveness of this technique depends on the accuracy of our knowledge of the plant (particularly at the disturbance frequency, where the control effort is concentrated) and on the speed of the estimator.

Conceptually, disturbance cancellation depends on the separability of the estimator and regulator designs. The method just discussed assumes that the estimates of the disturbance states are available for feedback so that the corresponding regulator gains can be designed to cancel the disturbance. This restriction presents a problem

when the plant model involves multiplicative white noise (as with the time-domain technique of Chapter 3) or when an optimal reduced order controller (see Chapter 4) is desired. In both cases, the separation principle of LQG does not hold, because the controller solution involves *coupled* Riccati- and Lyapunov-type equations [see (4.16)–(4.19)]. As will be discussed in section 5.5, however, its application is most desirable under precisely these conditions.

Fortunately, disturbance cancellation can be adapted to problems involving a coupled regulator and estimator by applying an idea based on inverse optimal control. The inverse problem of optimal control for linear state feedback problems was first investigated by Kalman (1964). Given a linear state feedback law, $u = -Fx$, the inverse problem is concerned with finding all cost functionals,

$$J = \tfrac{1}{2} \int_0^\infty \left[x(t)^{\mathrm{T}} R_1 x(t) + 2x(t)^{\mathrm{T}} R_{12} u(t) + u(t)^{\mathrm{T}} R_2 u(t) \right] dt \tag{5.28}$$

for which the control law is optimal. That is, given F, find all corresponding matrix triples $\{R_1,R_{12},R_2\}$. Returning to the disturbance cancellation problem, F is considered known, because we can set $F_p = 0$ and solve (5.23) or (5.27) for F_w, assuming for the moment that we have a standard LQG problem. By applying inverse optimal control, we can then solve for cost matrices — R_1, R_{12}, and R_2 — which provide us with a regulator to cancel the anticipated disturbance. These cost matrices may then be used to design a controller with coupled regulator and estimator for problems involving multiplicative white noise or optimal reduced-order control.

Kalman resolved the inverse optimal control problem for single-input systems, with the restriction: $R_{12} = 0$. He showed (Kalman 1964, Theorem 6) that a solution, $\{R_1,R_2\}$, exists if and only if a certain condition on the return difference function — now known as the *Kalman inequality* — holds. The restriction on R_{12} is not a concern for many problems, because any crossweighting may effectively be eliminated by a suitable modification of the system matrix and the addition of artificial state feedback to compensate for that modification. In estimator-based feedback systems where the separation principle does not hold, however, that method of formulating an equivalent problem is no longer valid.

Kreindler and Jameson (1972) showed that if a nonzero crossweighting is allowed, the inverse problem always has a solution. Since they were concerned with the

conventional problem of finding *all* solutions, $\{R_1,R_{12},R_2\}$, their approaches are so general (e.g., allowing Riccati solutions not to be positive definite) that they make it difficult to find a single solution.

When only a single solution is sought, as is the case here, the inverse optimal control problem is much simpler. Let $R_2 = I$ and assume that the state feedback law which provides complete disturbance cancellations given by: $u = -Fx$. Then the integrand of the cost functional (5.28) can be rewritten as follows:

$$\begin{aligned} & x^{\mathrm{T}}R_1x + 2x^{\mathrm{T}}R_{12}u + u^{\mathrm{T}}R_2u \\ & \quad = x^{\mathrm{T}}(R_1 - R_{12}{R_2}^{-1}{R_{12}}^{\mathrm{T}})x + (u + {R_2}^{-1}{R_{12}}^{\mathrm{T}}x)^{\mathrm{T}}R_2(u + {R_2}^{-1}{R_{12}}^{\mathrm{T}}x) \\ & \quad = x^{\mathrm{T}}(R_1 - R_{12}{R_{12}}^{\mathrm{T}})x + x^{\mathrm{T}}({R_{12}}^{\mathrm{T}} - F)^{\mathrm{T}}({R_{12}}^{\mathrm{T}} - F)x \end{aligned} \tag{5.29}$$

Since a negative cost is physically impossible, we must assume that

$$R_1 - R_{12}{R_2}^{-1}{R_{12}}^{\mathrm{T}} = R_1 - R_{12}{R_{12}}^{\mathrm{T}} \geq 0 \tag{5.30}$$

where " ≥ 0" denotes positive semidefiniteness. Therefore, a set of cost matrices for which F minimizes the cost functional (in fact makes it zero) is:

$$\boxed{\{R_1, R_{12}, R_2\} = \{F^{\mathrm{T}}F, F^{\mathrm{T}}, I\}} \tag{5.31}$$

This means that in order to apply the disturbance cancellation technique to multiplicative white noise or optimal reduced-order control problems, we can use the following algorithm:

(1) Solve (5.23) or (5.27) for F_w
(2) Arbitrarily set $F_p = 0$, so that $F = [0 \quad F_w]$
(3) Compute the desired cost matrices according to (5.31)
(4) Design the controller using the cost matrices from step (3)

As in the case of LQG problems, the effectiveness of this algorithm depends, in a sense, on a good estimator, since accurate knowledge of the states was assumed. Even though there is no explicit estimator in these modified LQG controllers, the algorithm stated above has worked well consistently for "large" V_1 (corresponding to a fast estimator).

In the event that some transient suppression is desired in addition to good steady disturbance rejection, the algorithm may be modified by adding a penalty to the plant states. The above algorithm yields

$$R_1 = \begin{bmatrix} 0 & 0 \\ 0 & F_w{}^{\mathrm{T}} F_w \end{bmatrix} \tag{5.32}$$

The desired transient suppression is accomplished by replacing the zero submatrix of the (1,1) block with a sufficiently "large" positive semidefinite matrix. That matrix may be chosen by penalizing modal energy as in section 5.1 (i.e., by choosing some multiple of R_1 from (5.8) to fill the (1,1) block). Experience has shown this modification to the algorithm to be very effective.

5.3 Discrete-Time Controller Design

Continuous-time controllers tend to be more convenient for theoretical work and for frequency-domain analysis, but implementation of a control law on a digital controller requires the design of a discrete-time controller. Ideally, the discrete-time controller is directly designed from sampled-data model of the control system. Since discrete-time design algorithms are readily available for the standard LQG problem, this direct design approach was used to implement controllers on the hardware (see Chapter 7) when neither multiplicative white noise nor optimal projection was involved. The development of the sampled-data version of those modified LQG problems, as well as the algorithms necessary to solve for their respective compensators, is beyond the scope of this work and arguably could be considered a duplication of effort. A good approximation to the directly designed discrete-time controller can be obtained by discretizing the corresponding continuous-time controller. Of course, the approximation is particularly good when the sample rate is high. Therefore, it is fortunate that optimal reduced-order controllers tend to make higher sample rates possible.

The remainder of this section discusses two considerations which should be taken into account in designing a discrete-time controller.

A. Conversion of Cost and Covariance Matrices to Discrete-Time

In sampled-data controller design, the continuous-time model of an actual

continuous-time plant is translated into a discrete-time equivalent, and there is relatively little intuitive understanding for the meaning of the resulting discrete-time variables. Therefore, rather than designing the compensator entirely in discrete-time, it makes more sense to specify the cost functional and noise covariances in continuous-time and then translated them to discrete-time for computing the optimal discrete-time controller. Assume the continuous-time plant model,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + w_1(t) \\ y(t) &= Cx(t) + Du(t) + w_2(t) \end{aligned} \tag{5.33}$$

where the noise covariances are given by:

$$\mathbb{E}\left\{ \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \begin{bmatrix} w_1{}^{\mathrm{T}}(t) & w_2{}^{\mathrm{T}}(t) \end{bmatrix} \right\} = \begin{bmatrix} V_1 & V_{12} \\ V_{12}{}^{\mathrm{T}} & V_2 \end{bmatrix} \delta(t) \triangleq V\delta(t) \tag{5.34}$$

And define the zero-order hold equivalent of the plant:

$$\begin{aligned} x_{k+1} &= \Phi(\Delta t)x_k + \Gamma(\Delta t)u_k + w_{1_k} \\ y_k &= Cx_k + Du_k + w_{2_k} \end{aligned} \tag{5.35}$$

where Δt is the sampling interval, the subscript "k" denotes the sampling instant for $k = 0, 1, 2, \ldots$ (i.e., $x_k \triangleq x(t_k)$, where the k^{th} sampling instant occurs at time t_k), and

$$\begin{aligned} \Phi(\Delta t) &= e^{A \cdot \Delta t} \\ \Gamma(\Delta t) &= \int_0^{\Delta t} \Phi(t)dt \cdot B \end{aligned} \tag{5.36}$$

$$\begin{aligned} w_{1_k} &= \int_{t_k}^{t_{k+1}} \Phi(t - t_k) w_1(t) dt \\ w_{2_k} &= w_2(t_k) \end{aligned} \tag{5.37}$$

Then the discrete-time equivalent to the continuous-time cost functional, (5.28), takes the form,

$$J = \tfrac{1}{2} \sum_{k=0}^{\infty} \left[x_k{}^{\mathrm{T}} \hat{R}_1 x_k + 2x_k{}^{\mathrm{T}} \hat{R}_{12} u_k + u_k{}^{\mathrm{T}} \hat{R}_2 u_k \right] \tag{5.38}$$

where the discrete-time cost matrices are denoted with carats and are given by (Stengel 1986, pp. 276-7):

$$\boxed{\begin{aligned}\hat{R}_1 &= \int_0^{\Delta t} \Phi^{\mathrm{T}}(t) R_1 \Phi(t) dt \\ \hat{R}_{12} &= \int_0^{\Delta t} \Phi^{\mathrm{T}}(t)[R_1 \Gamma(t) + R_{12}] dt \\ \hat{R}_2 &= \int_0^{\Delta t} [\Gamma^{\mathrm{T}}(t) R_1 \Gamma(t) + \Gamma^{\mathrm{T}}(t) R_{12} + {R_{12}}^{\mathrm{T}} \Gamma(t) + R_2] dt\end{aligned}} \tag{5.39}$$

The discrete-time equivalent covariance matrices are computed by their definitions, using (5.34) and (5.37) as follows:

$$\begin{aligned}\boxed{\hat{V}_1} &\triangleq \mathbb{E}\left[{w_1}_k {w_1}_k^{\mathrm{T}}\right] \\ &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \Phi(t - t_k) w_1(t) dt \cdot \int_{t_k}^{t_{k+1}} {w_1}^{\mathrm{T}}(t) \Phi^{\mathrm{T}}(t - t_k) dt\right] \\ &= \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \Phi(t - t_k) \mathbb{E}\left[w_1(t) {w_1}^{\mathrm{T}}(t)\right] \Phi^{\mathrm{T}}(\tau - t_k) d\tau dt \\ &= \int_{t_k}^{t_{k+1}} \Phi(t - t_k) \int_{t_k}^{t_{k+1}} V_1 \delta(t - \tau) \Phi^{\mathrm{T}}(\tau - t_k) d\tau dt \\ &= \int_{t_k}^{t_{k+1}} \Phi(t - t_k) V_1 \Phi^{\mathrm{T}}(t - t_k) dt \\ &= \boxed{\int_0^{\Delta t} \Phi(t) V_1 \Phi^{\mathrm{T}}(t) dt}\end{aligned} \tag{5.40}$$

$$\begin{aligned}\boxed{\hat{V}_{12}} &\triangleq \mathbb{E}\left[{w_1}_k {w_2}_k^{\mathrm{T}}\right] \\ &= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \Phi(t - t_k) w_1(t) dt \cdot {w_2}^{\mathrm{T}}(t_k)\right] \\ &= \int_{t_k}^{t_{k+1}} \Phi(t - t_k) \mathbb{E}\left[w_1(t) {w_2}^{\mathrm{T}}(t_k)\right] dt \\ &= \int_{t_k}^{t_{k+1}} \Phi(t - t_k) V_{12} \delta(t - t_k) dt \\ &= \boxed{V_{12}}\end{aligned} \tag{5.41}$$

$$\hat{V}_2 \triangleq \mathbb{E}\left[w_{2_k} w_{2_k}{}^{\mathrm{T}}\right] \tag{5.42}$$

$$= \mathbb{E}\left[w_2(t_k) w_2{}^{\mathrm{T}}(t_k)\right]$$

$$= V_2 \delta(0) \quad ???$$

From (5.42) we see that the literal translation to discrete-time results in an infinite sensor noise covariance. If (5.42) were used for the controller design computations, the Kalman filter would degenerate to an open circuit. The problem arises from the fact that the δ-correlated, infinite-power white noise model assumed by LQG theory does not reflect reality. Intuitively, it seems desirable to retain the structure of the covariance derived in (5.42), while changing the scalar multiplier to a finite number. Gelb (1984, p.121) argues in favor of using

$$\boxed{\hat{V}_2 = \frac{V_2}{\Delta t}} \tag{5.43}$$

because the resulting discrete-time sensor noise covariance approaches the continuous-time covariance as $\Delta t \to 0$. That was the method used for this study, although the actual scalar multiplier used is not important, since it tends to become a design parameter in practice.

The matrix integrals used to compute the discrete-time cost and covariance matrices are easily evaluated numerically — for example, by means of the forward rectangular rule. Only 20 to 30 intervals were required for good accuracy in the examples of this study, but fewer are necessary for smaller Δt.

B. Computational Delay

Standard LQG compensators (i.e., those with a Kalman filter separate from the regulator) are easily modified to overcome the computational delay inherent in discrete-time controllers. For a one-sampling-interval time delay, the Kalman filter extrapolation and update equations need merely to be reordered to turn the Kalman *filter* into a one-step-ahead *predictor*. The predicted state estimates are then fed back, by means of the regulator, so that the proper state estimates reach the actuators at the proper time.

In problems involving multiplicative white noise and/or optimal projection, the lack of an explicit estimator makes that procedure impossible. If the continuous-time controller is simply discretized, there will be a one-interval time lag between the time of the intended controller command and the instant that command reaches the actuators. For faster sample rates, the problem is smaller and may not be significant. In case the computational delay is significant, it should be modeled in the control system.

A pure one-interval delay is precisely modeled by inserting the irrational transfer function

$$H_D(s) = e^{-s \cdot \Delta t} \tag{5.44}$$

between the commanded control signal leaving the controller and the actual control input entering the plant. However, since $H_D(s)$ is infinite-dimensional, a suitable finite-dimensional approximation must be found — preferably of as low order as possible, because additional dynamics tend to complicate controller design computations and increase controller order. Padé approximations to $e^{-s \cdot \Delta t}$ are useful in this context. A Padé approximation tends to be more accurate than the Taylor series approximation of the same order, and it is easily computed (Bender and Orszag 1978, secs. 8.3 and 8.4). As an example, the second-order Padé approximation to (5.44) is given here:

$$\hat{H}_D(s) = \frac{1 - \frac{1}{2}s \cdot \Delta t + \frac{1}{12}(s \cdot \Delta t)^2}{1 + \frac{1}{2}s \cdot \Delta t + \frac{1}{12}(s \cdot \Delta t)^2} \tag{5.45}$$

$$= \frac{s^2 - \frac{6}{\Delta t}s + \frac{12}{\Delta t^2}}{s^2 + \frac{6}{\Delta t}s + \frac{12}{\Delta t^2}}$$

5.4 Iterative Relaxation Algorithm for Solving Coupled Riccati/Lyapunov Equations

As mentioned in section 4.1, problems involving multiplicative white noise or optimal controller order reduction require an iterative method to find a solution to the coupled extremal equations. A number of different algorithms are described in the literature, but they all belong to one of two categories (or are a hybrid of the two) — *iterative relaxation* methods and *continuation* (or *homotopy*) methods. The idea behind the iterative relaxation approach is to alternately solve the coupled equations for one unknown at a time while treating the other unknowns as constants. The main

advantage of this method of solution is that it requires no prior analytical development, assuming standard Riccati- and Lyapunov-equation solvers are already available, and it is therefore easily and quickly implemented. The main disadvantage is that this type of algorithm is not guaranteed to converge to a solution.

The iterative relaxation algorithm detailed in this section is a variation on one described by Bernstein and Hyland (1988b, pp. 290-91). The algorithm contains two nested loops and is designed to handle the general reduced-order controller problem in the presence of simultaneous state-, control-, and measurement-dependent noise. More specifically, it attempts to solve the general optimal projection equations, (4.16)–(4.20), for Q, P, $\hat{Q}$, $\hat{P}$, and $\tau_\perp$ and subsequently to find the controller state-space matrices — A_c, B_c, and C_c — using (4.15). If a full-order controller is desired, the algorithm is easily modified by removing the outer loop.

Define the *absolute norm* of an $m \times n$ matrix M as follows:

$$\| M \|_A \triangleq \max(M_{ij},\ i = 1,2,\ldots,m;\ j = 1,2,\ldots,n) \tag{5.46}$$

The iterative relaxation algorithm used in this study will now be stated:

(1) If the Stratonovich multiplicative noise interpretation is desired, modify A, B, and C using the substitutions prescribed in (3.73).

(2) Perform initializations:
   (a) Let $\gamma_i = 0$ $(i = 1,2,\ldots,p)$ and $\tau = I_n$ (i.e., $\tau_\perp = 0$).
   (b) Solve Riccati equations (4.16) and (4.17) for initial values of Q and P.
   (c) Solve Lyapunov equations (4.18) and (4.19) for initial $\hat{Q}$ and $\hat{P}$.
   (Note: this initialization corresponds to the standard LQG solution.)

Beginning of outer loop:

(3) Update the optimal projection matrix, τ:

(a) Compute a balancing transformation, Ψ, such that:

$$\Psi^{-1}\hat{Q}\Psi^{-1} = \Psi^{T}\hat{P}\Psi = \Lambda \tag{5.47}$$

$$\Lambda \triangleq \text{diag}\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$$

(Laub 1980). Note: all λ_i $(i = 1,2,\ldots,n)$ are nonnegative real.

(b) Perform a basis rearrangement:

Rearrange the λ_i so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, then rearrange the corresponding n columns of Ψ using that same ordering.

(c) Compute

$$\tau = \Psi \begin{bmatrix} I_{n_c} & 0 \\ 0 & \gamma \end{bmatrix} \Psi^{-1} \tag{5.48}$$

where

$$\gamma \triangleq \text{diag}\left\{\frac{\lambda_{n_c+1}}{\lambda_{n_c}}, \frac{\lambda_{n_c+2}}{\lambda_{n_c}}, \ldots, \frac{\lambda_n}{\lambda_{n_c}}\right\} \tag{5.49}$$

(4) Test for convergence of Q, P, $\hat{Q}$, $\hat{P}$, and τ (condition to leave outer loop):

(a) Compute relative errors, as follows.

$$e_Q \triangleq \|\text{r.h.s. of (4.16)}\|_A / \|V_1\|_A \tag{5.50}$$

$$e_P \triangleq \|\text{r.h.s. of (4.17)}\|_A / \|R_1\|_A$$

$$e_{\hat{Q}} \triangleq \|\text{r.h.s. of (4.18)}\|_A / \|\mathcal{Q}V_2{}^{-1}\mathcal{Q}^{T}\|_A$$

$$e_{\hat{P}} \triangleq \|\text{r.h.s. of (4.19)}\|_A / \|\mathcal{P}^{T}R_2{}^{-1}\mathcal{P}\|_A$$

(b) If $\max\{e_Q,e_P,e_{\hat{Q}},e_{\hat{P}}\} \leq \epsilon_1$ (the outer loop tolerance), go to (14). Otherwise, go to (5).

Beginning of inner loop:

(5) Update $\hat{V}_2$, $\hat{R}_2$, $\mathcal{Q}$, and $\mathcal{P}$ using equations (3.53), (3.58), (3.63), and (3.64).

(6) Update Q:

Solve Riccati equation (4.16) for Q by treating the terms involving γ_i and $\tau_\perp$ as constants, effectively adding them to V_1. The γ_i term in the definition of $\mathcal{Q}$ should be added to V_{12}, effectively defining a modified cross-covariance.

(7) Repeat (5).

(8) Update P:

Solve Riccati equation (4.17) for P by treating the terms involving γ_i and $\tau_\perp$ as constants, effectively adding them to R_1, with one exception. The γ_i term in the definition of $\mathcal{P}$ should be added to R_{12}, effectively defining a modified cross-weighting.

(9) Repeat (5) and update A_P using equation (3.65).

(10) Update $\hat{Q}$:

Solve Lyapunov equation (4.18) for $\hat{Q}$ by treating the term involving $\tau_\perp$ as a constant, effectively adding it to the other constant term.

(11) Repeat (5) and update A_Q using equation (3.65).

(12) Update $\hat{P}$:

Solve Lyapunov equation (4.19) for $\hat{P}$ by treating the term involving $\tau_\perp$ as a constant, effectively adding it to the other constant term.

(13) Test for convergence of Q, P, $\hat{Q}$, and $\hat{P}$ (condition to leave inner loop):

(a) Compute maximum relative change since the last iteration, as follows.

$$\begin{aligned} \Delta Q &\triangleq \|Q^i - Q^{i-1}\|_A / \|Q^i\|_A \\ \Delta P &\triangleq \|P^i - P^{i-1}\|_A / \|P^i\|_A \\ \Delta \hat{Q} &\triangleq \|\hat{Q}^i - \hat{Q}^{i-1}\|_A / \|\hat{Q}^i\|_A \\ \Delta \hat{P} &\triangleq \|\hat{P}^i - \hat{P}^{i-1}\|_A / \|\hat{P}^i\|_A \end{aligned} \tag{5.51}$$

where the superscripts indicate the iteration number of the inner loop — i being the current iteration.

(b) If $\max\{Q,P,\hat{Q},\hat{P}\} \leq \epsilon_2$ (the inner loop tolerance), go to (3). Otherwise, go to (5).

(14) Compute G and Γ using equations (4.14).

(15) Compute A_c, B_c, and C_c using equations (4.15).

We should recall that the optimal projection equations are *necessary* conditions for the solution of optimal reduced-order controllers. According to Hyland and Bernstein (1988b, p. 292), for an n^{th}-order plant with m control inputs and l sensor outputs, there is only one solution if $n_c \geq \min(n,m,l)$, but otherwise there may be as many as

$$\binom{\min(n,m,l)}{n_c} \triangleq \frac{[\min(n,m,l)]!}{[\min(n,m,l)-n_c]! \cdot n_c!}$$

solutions. The choice of initial values for Q, P, $\hat{Q}$, and $\hat{P}$ corresponding to the LQG solution in step (2) is an attempt to begin the iterations as close as possible to the optimal solution sought, therefore minimizing the risk of convergence to a suboptimal solution and reducing computation time. However, for the problems treated in this study, $m = 1$. That allowed the controller order to be specified as low as needed without concern for the existence of multiple solutions to the optimal projection equations.

The most difficult parts of the algorithm to establish are the convergence tests and tolerance specifications. These are important, since a test which is too stringent will never be satisfied, and a test which is too relaxed will not allow the algorithm to converge. In either case, the algorithm becomes an infinite loop. Unfortunately, the tolerances must be adjusted to suit the problem at hand, because the values of the most effective tolerances are somewhat sensitive to the plant model and even to controller authority. Although both tests are in some sense normalized, they are not universally applicable with the same tolerances. When the algorithm is repeatedly applied to the same plant model, acceptable tolerances can be chosen after some experimentation. Otherwise, it is advisable to do one or more of the following: (1) monitor more than one indicator, (2) make the tolerances automatically adaptive, or (3) give the software the

capability to accept manual changes in the tolerances between iterations.

The convergence test used for the inner loop — step (12) — was chosen because it is very cheap computationally to monitor *changes* in values since the last iteration. This type of test has also proven to be quite effective in determining convergence, even when compared to more sophisticated tests. The outer loop convergence test — step (4) — was suggested by Richter and Collins (1989). It seemed to be the logical choice for the outer loop (or for the *only* loop if both are not present), because it monitors the equation errors directly, although it requires much more matrix arithmetic. Still another convergence test for the outer loop was suggested by Hyland and Bernstein (1988b, p. 291):

$$\frac{\text{trace}(\tau) - n_c}{n_c} < \epsilon \tag{5.52}$$

for some tolerance ϵ. This test is based on the property — $\text{trace}(\tau) = n_c$ — when complete convergence has been attained. To see why this property holds, recall equations (4.13) and (4.14), and note that

$$\tau = G^{\mathrm{T}}\Gamma, \quad \Gamma G^{\mathrm{T}} = I_{n_c} \tag{5.53}$$

Then

$$\text{trace}(\tau) = \text{trace}(G^{\mathrm{T}}\Gamma) = \text{trace}(\Gamma G^{\mathrm{T}}) = n_c \tag{5.54}$$

Step (5) — the update of matrix values which appear in the Riccati- and Lyapunov-type equations — is repeated in several places throughout the algorithm in order to give those equations the most recent information. It may be omitted in some places in order to save a little computation time, but the algorithm may diverge if it is not repeated frequently enough.

5.5 Homotopy Algorithm for Solving Optimal Projection Equations

Several homotopy algorithms have been developed to solve the optimal projection equations, but as of this writing they have only been hinted at (e.g., Richter 1987; Richter and Collins 1989) and have not been published in explicit form in archival journals. For that reason, a homotopy algorithm is developed here, although with the

restriction that it applies to optimal projection only (i.e., multiplicative white noise is not included). A similar algorithm for the more general case may be derived based on the principles given in this section, although it would likely be much more complicated. The presence of multiplicative white noise adds several new terms to the Riccati and Lyapunov equations, resulting in the appearance of all five unknowns (Q, P, $\hat{Q}$, $\hat{P}$, and τ) in all four equations when state-, control-, and measurement-dependent noise are present.

Examination of the optimal projection equations (4.7)–(4.11) reveals that the degree of coupling among the Riccati- and Lyapunov-type equations is directly related to the "size" of the matrices V_2 and R_2. When the sensor noise covariances and control penalties are large, the coupling is reduced. However, when large controller authority (i.e., small V_2 and R_2) is desired, the interaction among these equations tends to be great, and the iterative algorithm of section 5.4 is less likely to converge to a solution. This problem frequently occurred when the disturbance cancellation method of section 5.2 was not used and the resulting controller authority needed to meet disturbance rejection specifications was necessarily high. Rather than alternately solving for Q, P, $\hat{Q}$, and $\hat{P}$, repeatedly, the homotopy algorithm replaces this inner loop iteration with numerical integration and finds these solutions (for fixed τ) in a single pass. It does so by continuously deforming the Riccati equation solutions from a known solution (e.g., the LQG solution) into the solution sought.

Homotopy methods of solving complex equations, such as the one described here, are developed in three steps: (1) Find the solution to a simple, but related equation; (2) Express the relationship between the solutions to the simple and complex equations in terms of a differential equation; and (3) (Numerically) integrate the differential equation to obtain the solution to the complex equation. Given a function $F:\mathbb{R}^n \to \mathbb{R}^n$, suppose we seek the solution, u, to $F(u) = 0$. Define a function $H:\mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ such that:

$$\begin{aligned} &\text{(a) } H(u(\alpha),\alpha) = 0 \text{ for } \alpha \in [0,1] \\ &\text{(b) } H(u,1) = F(u) \\ &\text{(c) a solution } u(0) \text{ to } H(u(0),0) = 0 \text{ is known} \end{aligned} \tag{5.55}$$

Then if H is continuous and $(\partial H/\partial u)^{-1}$ exists over the entire interval $\alpha \in [0,1]$, the solution $u(1)$ to $H(u(1),1) = F(u) = 0$ may be found as follows. First, differentiate

$H(u(\alpha),\alpha) = 0$ with respect to α using the chain rule for vectors (Graham 1981, sec. 4.3):

$$\frac{dH(u(\alpha),\alpha)}{\alpha} = \frac{du(\alpha)}{d\alpha} \cdot \frac{\partial H(u(\alpha),\alpha)}{\partial u(\alpha)} + \frac{\partial H(u(\alpha),\alpha)}{\partial \alpha} = 0 \tag{5.56}$$

Then

$$\frac{du(\alpha)}{d\alpha} = -\frac{\partial H(u(\alpha),\alpha)}{\partial \alpha}\left(\frac{\partial H(u(\alpha),\alpha)}{\partial u(\alpha)}\right)^{-1} \tag{5.57}$$

Since $u(0)$ is known, we can treat (5.57) as an initial value problem, integrating $du(\alpha)/d\alpha$ over the interval from 0 to 1 to obtain the solution $u \triangleq u(1)$. For a more complete discussion of homotopy methods, see Richter and DeCarlo (1983).

Before applying the homotopy principle to the modified Riccati equations (4.7)–(4.8), it is advisable to rewrite these equations in vector form in order to avoid the problem of taking derivatives of a matrix with respect to a matrix. Recall equation (4.7):

$$0 = AQ + QA^{\mathrm{T}} - \mathcal{Q}V_2{}^{-1}\mathcal{Q}^{\mathrm{T}} + V_1 + \tau_{\perp}\mathcal{Q}V_2{}^{-1}\mathcal{Q}^{\mathrm{T}}\tau_{\perp}{}^{\mathrm{T}} \tag{4.7}$$

where $\mathcal{Q} \triangleq QC^{\mathrm{T}} + V_{12}$, and assume $\tau_{\perp}$ is fixed. Then the equivalent equation in Q that we wish to solve is:

$$\begin{aligned} F_Q(\mathrm{vec}Q, \mathrm{vec}\mathcal{Q}^{\mathrm{T}}) &= (A \oplus A)\ \mathrm{vec}Q + \mathrm{vec}\ V_1 - \mathrm{vec}(\mathcal{Q}V_2{}^{-1}\mathcal{Q}^{\mathrm{T}}) + \tau^*\mathrm{vec}(\mathcal{Q}V_2{}^{-1}\mathcal{Q}^{\mathrm{T}}) \\ &= 0 \end{aligned} \tag{5.58}$$

where

$$\tau^* \triangleq \tau_{\perp} \otimes \tau_{\perp} \tag{5.59}$$

If the last term in (5.58) [or (4.7)] were not present, we would have a standard Riccati equation, whose solution is easily found. Therefore, a logical choice for the function H is:

$$H_Q(\mathrm{vec}Q, \mathrm{vec}\mathcal{Q}^{\mathrm{T}}, \alpha) = (A \oplus A)\ \mathrm{vec}Q + \mathrm{vec}\ V_1 - \mathrm{vec}(\mathcal{Q}{V_2}^{-1}\mathcal{Q}^{\mathrm{T}}) + \alpha\tau^*\mathrm{vec}(\mathcal{Q}{V_2}^{-1}\mathcal{Q}^{\mathrm{T}})$$
$$= 0 \tag{5.60}$$

For $\alpha = 0$, (5.60) is equivalent to the standard Riccati equation, and for $\alpha = 1$, we recover the modified Riccati equation (4.7). Denote the initial solution for Q (i.e., the LQG solution) by Q^0, corresponding to the constant projection matrix: $\tau^0 = I \Rightarrow \tau^{*0} = 0$. Equation (5.60) is suitable if we wish to solve (4.7) only once, for constant $\tau = \tau^1$, where τ^1 has been computed based on the initial values of $\hat{Q}$ and $\hat{P}$. However, we also wish to solve for τ, so the outer loop of the algorithm will be updating the projection matrix, giving τ^i, $i = 1, 2, \ldots$. Therefore, on the i^{th} iteration we need to continuously deform a known solution Q^{i-1} into Q^i based on our knowledge of τ^i and τ^{i-1}. To that end, consider the function,

$$\begin{aligned} H_Q^i(\mathrm{vec}Q^i, \mathrm{vec}\mathcal{Q}^{i^{\mathrm{T}}}, \alpha) = (A \oplus A)\ \mathrm{vec}Q^i + \mathrm{vec}\ V_1 - \mathrm{vec}(\mathcal{Q}^i{V_2}^{-1}\mathcal{Q}^{i^{\mathrm{T}}}) \\ + \tau^{*i-1}\mathrm{vec}(\mathcal{Q}^{i-1}{V_2}^{-1}\mathcal{Q}^{i-1^{\mathrm{T}}}) \\ + \alpha[\tau^{*i}\mathrm{vec}(\mathcal{Q}^i{V_2}^{-1}\mathcal{Q}^{i^{\mathrm{T}}}) \\ - \tau^{*i-1}\mathrm{vec}(\mathcal{Q}^{i-1}{V_2}^{-1}\mathcal{Q}^{i-1^{\mathrm{T}}})] = 0 \end{aligned} \tag{5.61}$$

and note that it satisfies the conditions in (5.55). For $\alpha = 0$, the solution $(\mathrm{vec}Q^{i-1}, \mathrm{vec}\mathcal{Q}^{i-1^{\mathrm{T}}})$ from the previous iteration satisfies (5.61), and for $\alpha = 1$, (5.61) is equivalent to (5.58). Now we may develop the initial condition problem to solve by taking the derivative of this function with respect to α, as follows:

$$\frac{dH_Q^i}{d\alpha} = \frac{d\mathrm{vec}Q^i}{d\alpha} \cdot \frac{\partial H_Q^i}{\partial \mathrm{vec}Q^i} + \frac{d\mathrm{vec}\mathcal{Q}^{i^{\mathrm{T}}}}{d\alpha} \cdot \frac{\partial H_Q^i}{\partial \mathrm{vec}\mathcal{Q}^{i^{\mathrm{T}}}} + \frac{\partial H_Q^i}{\partial \alpha} = 0 \tag{5.62}$$

By the definition of $\mathcal{Q}$, we have:

$$\frac{d\mathrm{vec}\mathcal{Q}^{i^{\mathrm{T}}}}{d\alpha} = \frac{d}{d\alpha}\ \mathrm{vec}(CQ^i + {V_{12}}^{\mathrm{T}}) = \frac{d}{d\alpha}\ (I \otimes C)\ \mathrm{vec}Q^i = \frac{d\mathrm{vec}Q^i}{d\alpha} \cdot (I \otimes C^{\mathrm{T}}) \tag{5.63}$$

Therefore,

$$\frac{d\mathrm{vec}Q^i}{d\alpha}\left[\frac{\partial H_Q^i}{\partial \mathrm{vec}Q^i} + (I \otimes C^{\mathrm{T}}) \cdot \frac{\partial H_Q^i}{\partial \mathrm{vec}\mathcal{Q}^{i^{\mathrm{T}}}}\right] + \frac{\partial H_Q^i}{\partial \alpha} = 0 \tag{5.64}$$

or:

$$\frac{d\text{vec}Q^i}{d\alpha} = -\frac{\partial H_Q^i}{\partial \alpha}\left[\frac{\partial H_Q^i}{\partial \text{vec}Q^i} + (I \otimes C^{\mathrm{T}}) \cdot \frac{\partial H_Q^i}{\partial \text{vec}\mathcal{Q}^{i^{\mathrm{T}}}}\right]^{-1} \tag{5.65}$$

where

$$\frac{\partial H_Q^i}{\partial \alpha} = \left[\tau^{*i}\text{vec}(\mathcal{Q}^i {V_2}^{-1} \mathcal{Q}^{i^{\mathrm{T}}}) - \tau^{*i-1}\text{vec}(\mathcal{Q}^{i-1} {V_2}^{-1} \mathcal{Q}^{i-1^{\mathrm{T}}})\right]^{\mathrm{T}}$$

$$\frac{\partial H_Q^i}{\partial \text{vec}Q^i} = (A \oplus A)^{\mathrm{T}}$$

$$\frac{\partial H_Q^i}{\partial \text{vec}\mathcal{Q}^{i^{\mathrm{T}}}} = -\left[({V_2}^{-1} \mathcal{Q}^{i^{\mathrm{T}}} \otimes I)_{(n)} + (I \otimes {V_2}^{-1} \mathcal{Q}^{i^{\mathrm{T}}})\right](I - \alpha\tau^{*i})^{\mathrm{T}}$$

The properties used to take these partial derivatives may all be found in Graham (1981). The operator denoted by the subscript "(n)" is also borrowed from Graham (1981, p. 71). It has the effect of reordering the rows of a matrix by taking the first row followed by each subsequent n^{th} row, then the second row followed by each subsequent n^{th} row, etc. Here, n is the order of the (augmented) plant (i.e., $A \in \mathbb{R}^{n \times n}$).

In summary, we may solve the Riccati equation (4.7) for Q^i (given a fixed τ^i) by treating the previous solution, Q^{i-1}, as the initial condition at $\alpha = 0$, then integrating (5.65) over the interval from $\alpha = 0$ to $\alpha = 1$. The starting value, Q^0, is the LQG solution for Q (i.e., the solution for $\tau_\perp = 0$).

The same procedure may be applied to the other Riccati equation, (4.8). From the definition of $\mathcal{P}$ (4.6) we have:

$$\frac{d\text{vec}\mathcal{P}^i}{d\alpha} = \frac{d}{d\alpha}\,\text{vec}(B^{\mathrm{T}}P^i + {R_{12}}^{\mathrm{T}}) = \frac{d}{d\alpha}\,(I \otimes B^{\mathrm{T}})\,\text{vec}P^i = \frac{d\text{vec}P^i}{d\alpha} \cdot (I \otimes B) \tag{5.66}$$

which leads to:

$$\frac{d\mathrm{vec}P^i}{d\alpha} = -\frac{\partial H_P^i}{\partial \alpha}\left[\frac{\partial H_P^i}{\partial \mathrm{vec}P^i} + (I \otimes B)\cdot\frac{\partial H_P^i}{\partial \mathrm{vec}\mathcal{P}^i}\right]^{-1} \tag{5.67}$$

where

$$\frac{\partial H_P^i}{\partial \alpha} = \left[\tau^{*i\mathrm{T}}\mathrm{vec}(\mathcal{P}^{i\mathrm{T}}R_2^{-1}\mathcal{P}^i) - \tau^{*i-1\mathrm{T}}\mathrm{vec}(\mathcal{P}^{i-1\mathrm{T}}R_2^{-1}\mathcal{P}^{i-1})\right]^{\mathrm{T}}$$

$$\frac{\partial H_P^i}{\partial \mathrm{vec}P^i} = A \oplus A$$

$$\frac{\partial H_P^i}{\partial \mathrm{vec}\mathcal{P}^i} = -\left[(R_2^{-1}\mathcal{P}^i \otimes I)_{(n)} + (I \otimes R_2^{-1}\mathcal{P}^i)\right](I - \alpha\tau^{*i})$$

The Lyapunov equations, (4.9)–(4.10), do not require a homotopy solution for fixed τ^i, because: (a) the variables Q and P can be treated as fixed constants, since they have already been solved for on the first pass, and (b) the term involving the projection matrix is not a function of the only remaining variable, $\hat{Q}$ (or $\hat{P}$), and may therefore be treated as part of the constant term. That is, having already solved the modified Riccati equations for Q and P, the modified Lyapunov equations may be treated as standard Lyapunov equations and be solved immediately for $\hat{Q}$ and $\hat{P}$.

The entire homotopy algorithm for the solution of the optimal reduced-order control problem (without multiplicative white noise) is now stated.

(1) Perform initializations:
   (a) Let $\tau^0 = I_n$ (i.e., $\tau_\perp{}^0 = 0$ and $\tau^{*0} = 0$).
   (b) Solve Riccati equations (4.7) and (4.8) for the initial values, Q^0 and P^0.
   (c) Solve Lyapunov equations (4.9) and (4.10) for initial values $\hat{Q}^0$ and $\hat{P}^0$.
   (d) Set the iteration number: $i = 0$.
   (Note: this initialization corresponds to the standard LQG solution.)

Beginning of main loop:

(2) Increment the iteration number: $i \leftarrow i+1$.

(3) Compute the optimal projection matrix, τ^i (based on $\hat{Q}^{i-1}$ and $\hat{P}^{i-1}$), exactly as in step (3) of the algorithm in section 5.4.

(4) Test for convergence of Q^{i-1}, P^{i-1}, $\hat{Q}^{i-1}$, $\hat{P}^{i-1}$, and τ^i (condition to leave main loop):

(a) Compute relative errors, as follows.

$$e_Q \triangleq \|\text{r.h.s. of (4.16)}\|_A / \|V_1\|_A \qquad (5.50)$$

$$e_P \triangleq \|\text{r.h.s. of (4.17)}\|_A / \|R_1\|_A$$

$$e_{\hat{Q}} \triangleq \|\text{r.h.s. of (4.18)}\|_A / \|\mathcal{Q} {V_2}^{-1} \mathcal{Q}^{\mathrm{T}}\|_A$$

$$e_{\hat{P}} \triangleq \|\text{r.h.s. of (4.19)}\|_A / \|\mathcal{P}^{\mathrm{T}} {R_2}^{-1} \mathcal{P}\|_A$$

(b) If $\max\{e_Q, e_P, e_{\hat{Q}}, e_{\hat{P}}\} \leq \epsilon$ (the tolerance), go to (12).
Otherwise, go to (5).

(5) Compute $\mathcal{Q}^{i-1}$ and $\mathcal{P}^{i-1}$ (based on Q^{i-1} and P^{i-1}) using equations (3.63) and (3.64), and compute τ^{*i} using (5.59).

(6) Compute Q^i:

Numerically integrate (5.65) over the interval from $\alpha = 0$ to $\alpha = 1$.

(7) Compute P^i:

Numerically integrate (5.67) over the interval from $\alpha = 0$ to $\alpha = 1$.

(8) Compute $\mathcal{Q}^i$ and $\mathcal{P}^i$ (based on Q^i and P^i) using equations (3.63) and (3.64), and update A_P and A_Q using equations (3.65).

(9) Compute $\hat{Q}$:

Solve Lyapunov equation (4.9) for $\hat{Q}$ by treating the term invloving $\tau_\perp$ as a constant, effectively adding it to the other constant term.

(10) Compute $\hat{P}$:

Solve Lyapunov equation (4.10) for $\hat{P}$ by treating the term invloving $\tau_{\perp}$ as a constant, effectively adding it to the other constant term.

(11) Go to (4).

(12) Compute G and Γ using equations (4.14).

(13) Compute A_c, B_c, and C_c using equations (4.15).

6. Evaluation of Designs

The controller design principles developed in Chapters 2–5 were applied to an FDLTI model of a simply supported plate in order to study their effectiveness. This chapter describes the complete control system model, then evaluates the comparative value of the various parameter-robust and optimal reduced-order controller designs based on this model. Chapter 7 applies the analytical techniques of this chapter to solve the problems of plant uncertainty and high controller order that were encountered in the actual hardware.

6.1 Problem Description

The simply supported plate is made of cold-rolled steel. It is rectangular, of dimensions $.5\text{m} \times .6\text{m} \times 2.9\text{mm}$, and has attached to it twelve accelerometer sensors on one side and two point force actuators on the other side — one for the control input and one for external disturbance generation. The accelerometers are lightweight piezoelectric devices mounted to the plate with wax and are placed in a 3×4 rectangular array with locations devised to assure observability of the first twelve vibrational modes of the plate. The presence of more sensors than observed modes (i.e., *spatial oversampling*) is a redundancy that tends to provide greater accuracy in the sensor measurements. The actuators are electromagnetic shakers with a magnitude frequency response from voltage input command to force output that is essentially constant over the frequency range of interest. They are placed in locations coinciding with two of the sensor positions and chosen such that most of the authority of these shakers is on the first two modes. Also, the shakers are located near node lines of modes three and four, with the intention of limiting the "spillover" of the control signal into the higher frequency modes. For the experiments carried out in Chapter 7, the disturbance signal was produced by a function generator. The zero-order hold control signal was sent through a second-order low pass filter before reaching the control shaker for the purpose of smoothing the signal and preventing aliasing. For more details concerning the hardware configuration of the simply supported plate experiment, see Rubenstein (1991).

The plate with actuators and sensors is described by the standard state-space model,

$$\dot{x}_p(t) = A_p x_p(t) + B_p u(t) + [g_{11} \quad g_{12}] w_p(t) \tag{6.1}$$
$$y(t) = C_p x_p(t) + D_p u(t) + [g_{21} \quad g_{22}] w_p(t)$$

Expanding the matrices in (6.1), we define the modal model,

$$\frac{d}{dt}\begin{bmatrix} x_p(t) \\ x_v(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Omega^2 & -2Z\Omega \end{bmatrix}\begin{bmatrix} x_p(t) \\ x_v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \Phi_u \end{bmatrix} u(t) + \begin{bmatrix} 0 & 0 \\ \Phi_w & 0 \end{bmatrix}\begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \tag{6.2}$$

$$y(t) = \begin{bmatrix} -\Omega^2 & -2Z\Omega \end{bmatrix}\begin{bmatrix} x_p(t) \\ x_v(t) \end{bmatrix} + \Phi_u u + \begin{bmatrix} \Phi_w & \rho_v{}^{1/2}(\Phi^{\mathrm{T}}\Phi)^{-1/2} \end{bmatrix}\begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

The time variables and their dimensions are defined as follows:

$x_p \in \mathbb{R}^{n_m}$ ..... Modal position states
$x_v \in \mathbb{R}^{n_m}$ ..... Modal velocity states
$u \in \mathbb{R}^m$ ..... Control input
$w_1 \in \mathbb{R}^p$ ..... Process noise (disturbance input)
$w_2 \in \mathbb{R}^l$ ..... Sensor noise
$y \in \mathbb{R}^l$ ..... Modal acceleration measurements

For the simply supported plate experiment, $m = p = 1$, because there is only one control shaker and one disturbance shaker. The number of modeled modes, n_m, varies according to the order of the plant for which we wish to design a controller, although it is limited to nine (the total number of modes thus far identified). Since we are modeling n_m modes, the number of modal acceleration measurements is $l = n_m$. This number would be limited to twelve (the number of accelerometers), should there be more than twelve identified modes available. The excess accelerometers provide redundant information about the spatial accelerations from which a more reliable least squares solution to the modal accelerations is computed.

The quantities Ω and Z are diagonal matrices of the modal natural frequencies and damping ratios, respectively. That is,

$$\Omega \triangleq \text{diag}\{\omega_1, \omega_2, \ldots, \omega_{n_m}\} \tag{6.3}$$
$$= 2\pi \cdot \text{diag}\{f_1, f_2, \ldots, f_{n_m}\}$$

$$Z \triangleq \text{diag}\{\zeta_1, \zeta_2, \ldots, \zeta_{n_m}\}$$

where the modal natural frequencies and damping ratios are given in Table 6.1.

Table 6.1: Natural Frequencies and Damping Ratios for the First Nine Modes

i	f_i (Hz)	ζ_i
1	49.447	0.007722826
2	108.96	0.01171460
3	130.25	0.008318498
4	188.53	0.002731109
5	203.25	0.002725023
6	265.62	0.002387554
7	285.78	0.001224449
8	326.08	0.001321583
9	338.30	0.002220888

The column vectors Φ_u and Φ_w are the modeshapes corresponding to the control and disturbance inputs, respectively. They indicate the relative effect each shaker has on the modal amplitudes and are a function of the shaker locations. These modeshapes are related to the sensor modeshape matrix, Φ, introduced in Chapter 5 [see (5.11)]. According to the data provided by the identification procedure, the matrix Φ for a model of the first nine modes is given (to four places past the decimal) by:

$$\Phi = \begin{bmatrix} 0.2068 & -0.5350 & 0.4716 & -0.6245 & -0.4200 & -0.3708 & -0.5257 & -0.6133 & -0.2620 \\ 0.4290 & -0.7181 & 0.1104 & -0.1516 & -0.6814 & 0.4496 & -0.0798 & 0.6420 & -0.4180 \\ 0.3072 & -0.4771 & -0.5241 & 0.6138 & -0.4607 & -0.3951 & 0.4640 & -0.6078 & -0.2086 \\ 0.4001 & -0.2333 & 0.8233 & -0.1619 & 0.3468 & -0.6213 & 0.5346 & -0.2604 & 0.4473 \\ 0.7033 & -0.3062 & 0.1353 & -0.0554 & 0.5246 & 0.6389 & 0.1319 & 0.3292 & 0.6070 \\ 0.4829 & -0.1404 & -0.9120 & 0.4227 & 0.3599 & -0.6038 & -0.6515 & -0.2957 & 0.3680 \\ 0.4468 & 0.2215 & 0.7428 & 0.6210 & 0.3438 & -0.6273 & 0.6039 & 0.3454 & -0.3710 \\ 0.7033 & 0.3664 & 0.1091 & 0.2185 & 0.5573 & 0.6715 & 0.1343 & -0.4100 & -0.5521 \\ 0.4480 & 0.2853 & -0.8733 & -0.4771 & 0.3848 & -0.6201 & -0.6411 & 0.3428 & -0.3926 \\ 0.2803 & 0.4538 & 0.5085 & 1.0755 & -0.5294 & -0.3369 & -0.7840 & 0.6357 & 0.1768 \\ 0.4365 & 0.7018 & 0.0766 & 0.2210 & -0.7149 & 0.4368 & -0.1769 & -0.8063 & 0.3551 \\ 0.1884 & 0.4729 & -0.5331 & -1.0680 & -0.4109 & -0.3970 & 0.7980 & 0.5814 & 0.1820 \end{bmatrix} \tag{6.4}$$

The twelve rows of Φ are the modeshapes for the twelve accelerometer locations. The control and disturbance shakers are each colocated with one of the accelerometers, so Φ_u and Φ_w are given by two of the rows of Φ. Denote the i^{th} row of Φ by ϕ_i. Then

$$\Phi_u = \phi_8{}^{\mathrm{T}}, \quad \Phi_w = \phi_5{}^{\mathrm{T}} \tag{6.5}$$

The expression for g_{22} was developed in section 5.1 [see (5.14b)]. When fewer than nine modes are modeled (i.e., $n_m < 9$), the appropriate columns of Φ are eliminated. For example, in order to model only the first four modes, we take Φ to be equal to the first four columns of Φ in (6.4).

The control input, u, passes through a second-order smoothing filter before reaching the shaker. If we denote the control command entering the filter by u_c, a state-space model of the smoothing filter is:

$$\begin{aligned} \dot{x}_l(t) &= A_l x_l(t) + B_l u_c(t) \\ u(t) &= C_l x_l(t) \end{aligned} \tag{6.6}$$

where

$$A_l = \begin{bmatrix} 0 & 1 \\ -\omega_l^2 & -2\zeta_l\omega_l \end{bmatrix}, \quad B_l = \begin{bmatrix} 0 \\ \omega_l^2 \end{bmatrix}, \quad C_l = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{6.7}$$

$$\omega_l = 2\pi \cdot 120, \quad \zeta_l = .707$$

The disturbance w_1 is modeled as narrowband noise centered about 60 Hz. Conceptually, this was accomplished by passing the fictitious white noise variable v through a second-order noise shaping filter. The state-space model of that shaping filter is as follows:

$$\begin{aligned} \dot{x}_w(t) &= A_w x_w(t) + B_w v(t) \\ u(t) &= C_w x_w(t) \end{aligned} \tag{6.8}$$

where

$$A_w = \begin{bmatrix} 0 & 1 \\ -\omega_w^2 & -2\zeta_w\omega_w \end{bmatrix}, \quad B_w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_w = \begin{bmatrix} 0 & 1 \end{bmatrix} \tag{6.8}$$

$$\omega_w = 2\pi \cdot 60, \quad \zeta_w = .1$$

Controllers are designed from the augmented plant model, comprised of the interconnected plate, smoothing filter, and noise shaping filter dynamics — (6.1), (6.6), and (6.8). The n^{th}-order augmented plant is described by the following state and output equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu_c(t) + G_1 w(t) \\ y(t) &= Cx(t) + G_2 w(t) \end{aligned} \tag{6.9}$$

where

$$x(t) \triangleq \begin{bmatrix} x_p(t) \\ x_l(t) \\ x_w(t) \end{bmatrix}, \quad w(t) \triangleq \begin{bmatrix} v(t) \\ w_2(t) \end{bmatrix} \tag{6.10}$$

$$A = \begin{bmatrix} A_p & B_p C_l & g_{11} C_w \\ 0 & A_l & 0 \\ 0 & 0 & A_w \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ B_l \\ 0 \end{bmatrix}, \quad G_1 = \begin{bmatrix} 0 & g_{12} \\ 0 & 0 \\ B_w & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} C_p & D_p C_l & g_{21} C_w \end{bmatrix}, \quad G_2 = \begin{bmatrix} 0 & g_{22} \end{bmatrix}$$

The noise vector $w(t)$ is assumed to have identity covariance, so the correlations among the individual elements are specified by G_1 and G_2, and the relative intensities of the process and sensor noise are adjusted by means of ρ_v, which scales g_{22} [see (6.2)]. As stated in section 5.1, $g_{12} = 0$, since the sensor noise (w_2) only affects the output equation. The presence of the smoothing filter eliminates any feedthrough from the controller command signal to the outputs. However, for many of the experiments simulated in this chapter, the smoothing filter was ignored in order to thoroughly study the robustness design methods on a simple model before stepping up to a full scale model. Without the filter, the x_l states are eliminated and $u_c = u$.

Figure 6.1 shows the frequency responses of the plate (for $n_m = 4$), smoothing filter, and noise shaping filter just described. The frequency response of the plate is represented by a plot of its maximum singular value. That enables the magnitude response of the 4×1 transfer function matrix to be expressed by a single curve. The response of the noise shaping filter is scaled so as to be visible in the magnitude range of the plot. This curve also represents the frequency content of the assumed disturbance, although the actual scaling depends on the intensity of that disturbance and is considered a design parameter. The 120 Hz cutoff frequency of the smoothing filter was chosen to allow adequate control of the observed modes (modes 1–4) while minimizing excitation of the higher frequency unobserved modes.

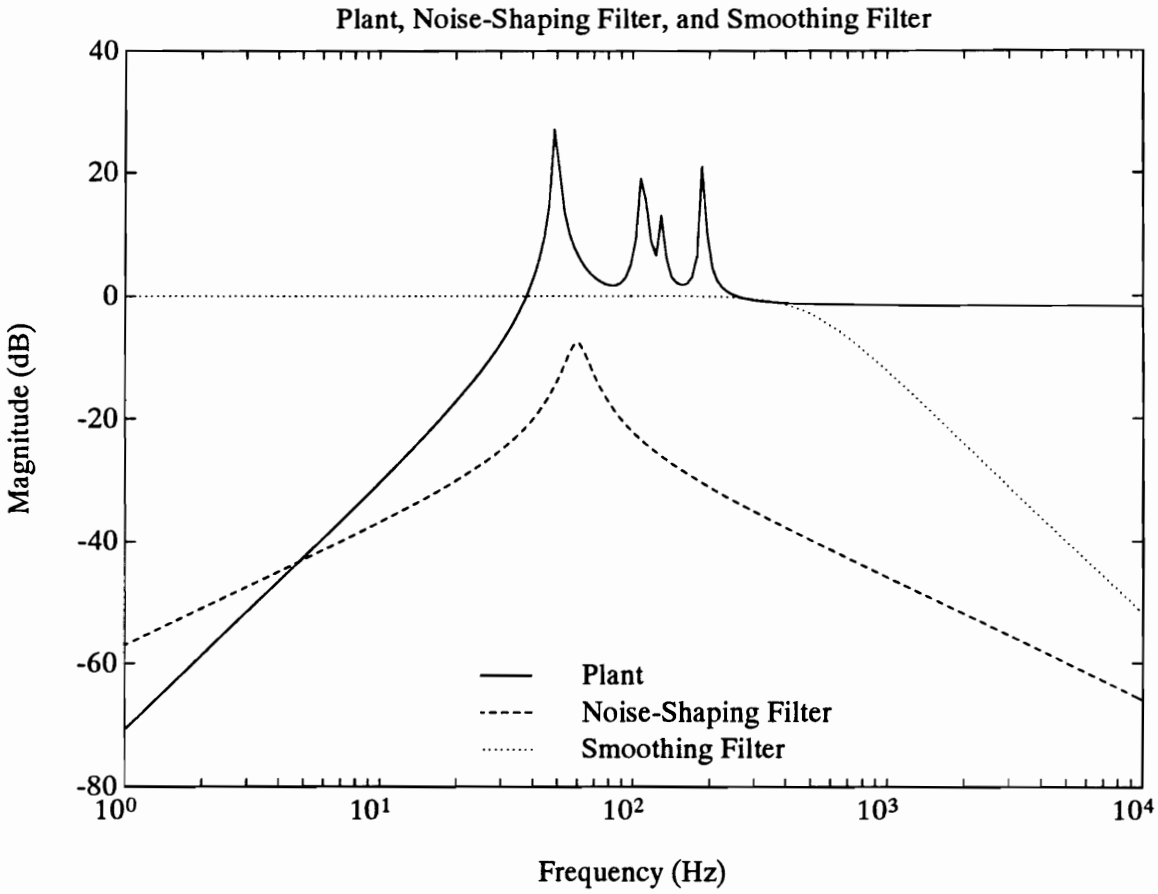

Figure 6.1: Frequency Responses of Augmented Plant Components for 4-Mode Model

6.2 Comparison of Tradeoff for Different Methods

In order to speed up computations and eliminate unnecessary complications, the robust control methods of chapters 2 and 3 were applied to a single mode model of the plate (i.e, $n_m = 1$), and the smoothing filter was omitted from the augmented plant model of (6.9)–(6.10). This continuous-time augmented plant model has order $n = 4$.

An LQG controller was chosen for the baseline design, and after closing the loop it was found that system stability was particularly sensitive to two of the plant parameters — the natural frequency, ω_1, and the scalar control input eigenvector, Φ_u (corresponding to the control shaker location). The cost functional weighting matrices were chosen by combining the modal energy penalty technique of section 5.1 with the disturbance cancellation algorithm of section 5.2 [using (5.23)], yielding:

$$R_1 = \begin{bmatrix} {\omega_1}^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0002 \end{bmatrix}, \quad R_2 = 1, \quad R_{12} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1.0001 \end{bmatrix} \tag{6.11}$$

Since there is only one modal output, there is no need to worry about relative intensities of multiple sensor noise sources. Therefore, g_{22} was simplified to: $g_{22} = {\rho_v}^{1/2}$. Then, as usual, the covariance matrices were given by:

$$V_1 = G_1 {G_1}^{\mathrm{T}}, \quad V_2 = G_2 {G_2}^{\mathrm{T}}, \quad V_{12} = G_1 {G_2}^{\mathrm{T}} \tag{6.12}$$

The sensitivity of system stability with respect to ω_1 and Φ_u varied greatly with the selection of ρ_v. For relatively large ρ_v (i.e., relatively small disturbance intensity), stability is more sensitive to ω_1 and less sensitive to Φ_u, whereas for relatively small ρ_v, the reverse is true.

Both cases — sensitivity to ω_1 and Φ_u — were studied in order to test the effectiveness of the various parameter-robustness methods under different conditions. If one method were clearly and consistently superior to the others, it would stand out during these tests. Since the same auxiliary output weighting technique is common to LQG/PRE, the frequency-domain method of Chapter 2, and the time-domain method of

Chapter 3, the comparative effectiveness of these methods was studied by improving robustness of system stability using only auxiliary input noise. Separately, only auxiliary output weighting was used. Then the effect of combining the two was demonstrated.

A. Uncertainty in Natural Frequency

When the cost matrices in (6.11) are combined with the noise intensity parameter, $\rho_v = 10^{-6}$, a $+3\%$ error in the natural frequency is enough to drive the closed-loop system (with LQG controller) unstable. In order to evaluate the robust controller designs, a specification was made to raise the stability margin to $\pm 10\%$, and the resulting loss in performance was compared for four different methods:

Auxiliary inputs only:

(1) White noise (w.n.)

(2) Frequency-shaped noise (f.s.n.)

(3) Multiplicative white noise (m.w.n.)

Auxiliary outputs only:

(4) Auxiliary output penalty (aux. output)

Note that the term "stability margin", as used here and in the remainder of this thesis, does not denote a gain or phase margin in the sense of classical control theory. Rather, it is an abbreviation for *parameter stability margin* and represents the maximum amount a parameter may deviate from its nominal (i.e., assumed) value without causing system instability.

The natural frequency appears in both the A- and C-matrices, so methods (1), (2), and (4) require that we factor both ΔA and ΔC to obtain [see (1.8)]:

$$M_a = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad M_c = 1, \quad N_a = N_c = \begin{bmatrix} 2.05 & 1 & 0 & 0 \end{bmatrix} \tag{6.13}$$

The natural frequency appears in matrix elements $-\omega_1^2$ and $-2\zeta_1\omega_1$ in both A and C, so it is impossible to find a relative scaling for the corresponding elements of ΔA and ΔC that will hold for any size deviation of ω_1. For a $+5\%$ deviation in ω_1

(or $-2\zeta_1\omega_1$), there is a $+2.05\times 5\%$ deviation in ${\omega_1}^2$ (or $-{\omega_1}^2$). This 2.05 ratio very nearly holds over the entire parameter range of interest, hence the values of N_a and N_c in (6.13).

In applying method (2), the noise shaping filter $H_a(s)$ was computed directly from equation (2.11), without a low-order approximation. The full-order shaping filter did not raise the controller order excessively, because the original plant order was so low. Also, due to the structure of the parameter uncertainties, the same noise shaping filter was used for $H_c(s)$, so no additional dynamics were necessary for that filter. Therefore the controller order for method (2) was $n_c = 8$, as opposed to $n_c = 4$ for the other methods.

The design parameters needed to just meet the $\pm 10\%$ stability margin specification are as follows:

(1) White noise ..... $\mu_a = \mu_c = 400$

(2) Frequency-shaped noise ..... (scale H_a and H_c by a factor of 1100)

(3) Multiplicative white noise..... $\gamma_1 = .64$,

$$A_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2.05 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$C_1 = \begin{bmatrix} 2.05 & 1 & 0 & 0 \end{bmatrix}$$

(4) Auxiliary output penalty ..... $\rho_a = \rho_c = 600$

Figure 6.2 shows the results of the performance/stability robustness tradeoff for all four robust controllers, compared with the baseline LQG controller. Both frequency-shaped noise and multiplicative white noise provided more suitable auxiliary input models than the white noise of LQG/PRE, although the multiplicative white noise design was clearly the best performer of the three. The auxiliary output penalty design gave up the least performance of all in achieving the stability robustness objective. At the nominal value of ω_1, that design, and the multiplicative white noise design, are nearly optimal in performance.

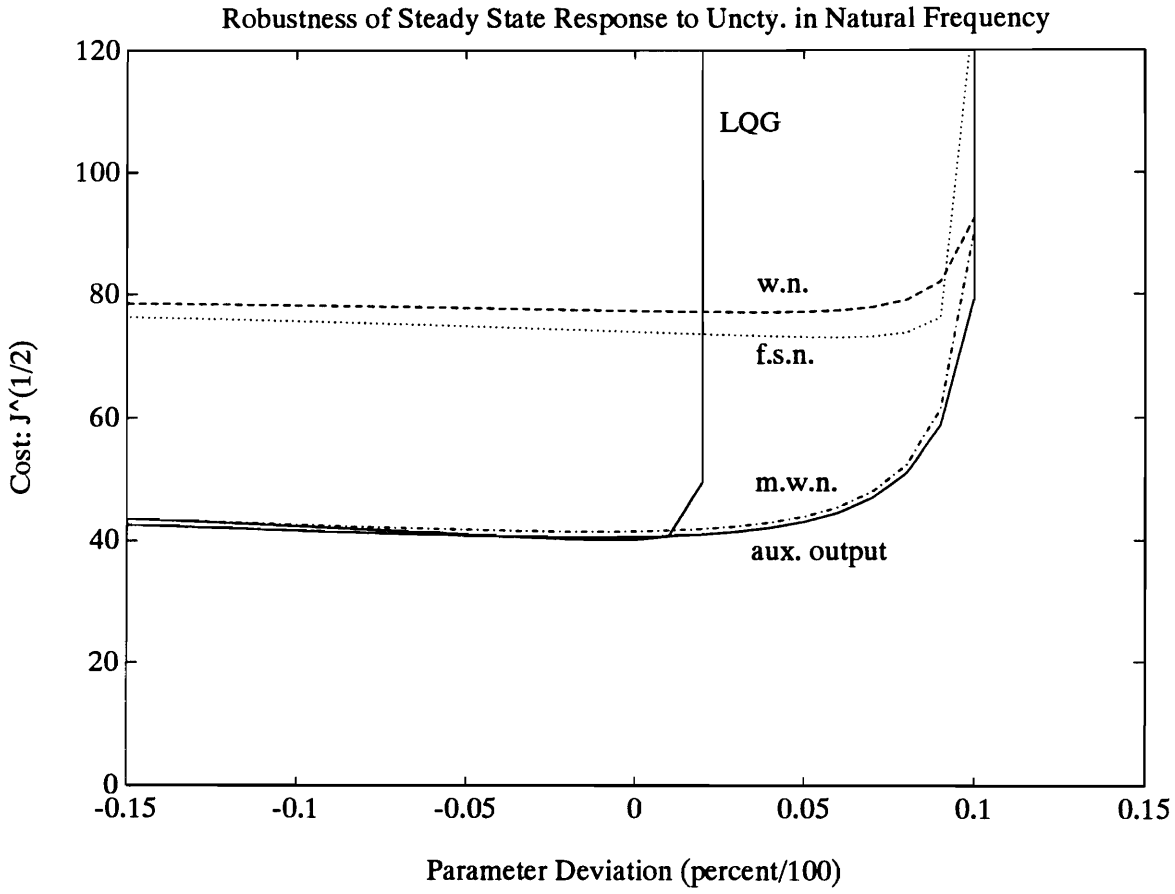

Figure 6.2: Steady State Performance and Stability Tradeoff for Uncertainty in Natural Frequency

The experiment depicted in Figure 6.2 was based on the assumption of steady state process noise — the premise of LQG optimal control. This comparison was repeated with an impulse applied at the process noise port [i.e. the signal v in (6.10)] in order to simulate the effect of transient disturbances on the designs. The results are shown in Figure 6.3. The LQG curve no longer represents optimality at zero parameter deviation, rather it is only an approximation to optimality. In fact, designs (3) and (4) both outperformed the LQG controller in the presence of an impulse disturbance. The multiplicative white noise design performs relatively somewhat better when the disturbance is an impulse, as was expected. The signal levels, and therefore the auxiliary input amplitudes, vary greatly over the course of time, and the multiplicative white noise model takes advantage of this. On the other hand, the finite-energy disturbance input conflicts with the premise of the frequency-shaped noise design, and it performs relatively worse.

B. Uncertainty in Eigenvector

In order to make the closed-loop system sensitive to the control input eigenvector, Φ_u, the sensor noise intensity parameter, ρ_v, was lowered to $\rho_v = 1/(4.9 \times 10^9)$. That resulted in the same baseline stability margin as in the natural frequency uncertainty problem just discussed. A $+3\%$ deviation in Φ_u (in the B_p-matrix only) destabilized the closed-loop system with LQG controller. The matrix $D_p = \Phi_u$ was left unaltered for this study, because hardware experimentation showed uncertainty in B_p to be a particular problem.

The control input eigenvector appears in the second row of the augmented B-column vector, so the obvious choice of a factorization for ΔB was:

$$M_b = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad N_b = 1 \tag{6.14}$$

The same stability margin specification was made as before — to improve that margin to $\pm 10\%$. The following design parameters enabled the four robustness methods to meet that specification:

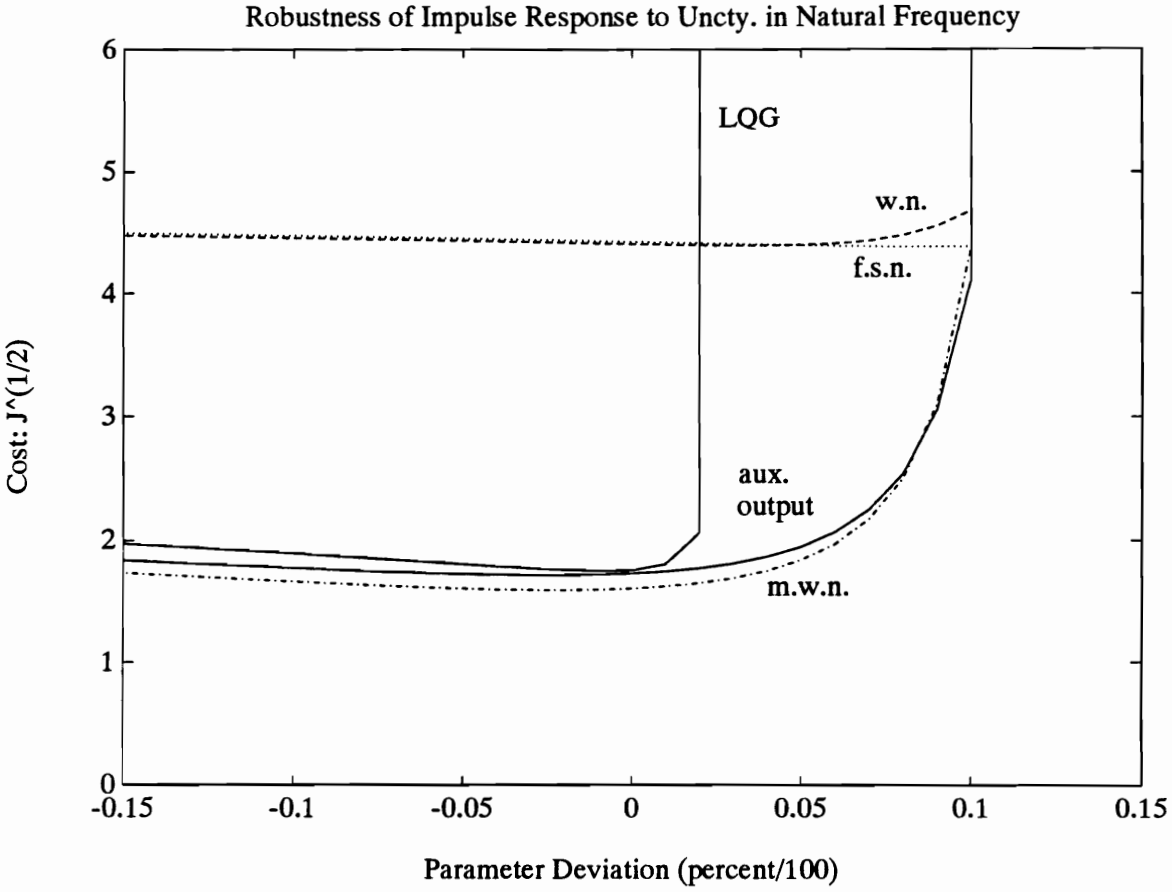

Figure 6.3: Transient Performance and Stability Tradeoff for Uncertainty in Natural Frequency

(1) White noise ..... $\mu_b = 70$

(2) Frequency-shaped noise ..... (scale H_b by a factor of .15)

(3) Multiplicative white noise..... $\gamma_1 = 1.44 \times 10^{-6}$,

$$B_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

(4) Auxiliary output penalty ..... $\rho_b = 3.4$

The steady state disturbance results for this problem are shown in Figure 6.4. The auxiliary output penalty method, rather than being the most effective, was by far the worst performer this time. Among the three auxiliary input models, multiplicative white noise once again gave very good results, but provided no improvement over the simple white noise model. Surprisingly, the frequency-shaped noise model led to a somewhat greater sacrifice in performance than the white noise model.

Plots of the impulse response costs for this example are shown in Figure 6.5. These results are not significantly different from those of the steady state disturbance case.

C. Results for Different Factorizations

Up to now, only the individual components of the parameter robustness methods under study have been compared. A logical procedure for choosing among LQG/PRE, the frequency-domain method of Chapter 2, and the time-domain method of Chapter 3 is to determine the corresponding auxiliary input model that gives the best performance, then to combine that auxiliary input model with the auxiliary output penalty common to all three methods. Once a required stability margin has been specified, the auxiliary input noise intensity (through μ_a, μ_b, and μ_c — assuming we are using LQG/PRE) and the auxiliary output penalty (through ρ_a, ρ_b, and ρ_c) may be applied individually or combined in any number of different proportions to just meet the stability specification. For any single independent parameter uncertainty, adjusting the relative magnitude of the μ- and ρ-scalars is equivalent to choosing different factorizations for ΔA, ΔB, and ΔC, giving the corresponding nonzero elements of either the M- or the N-matrices a relatively greater magnitude. Of course, when more than

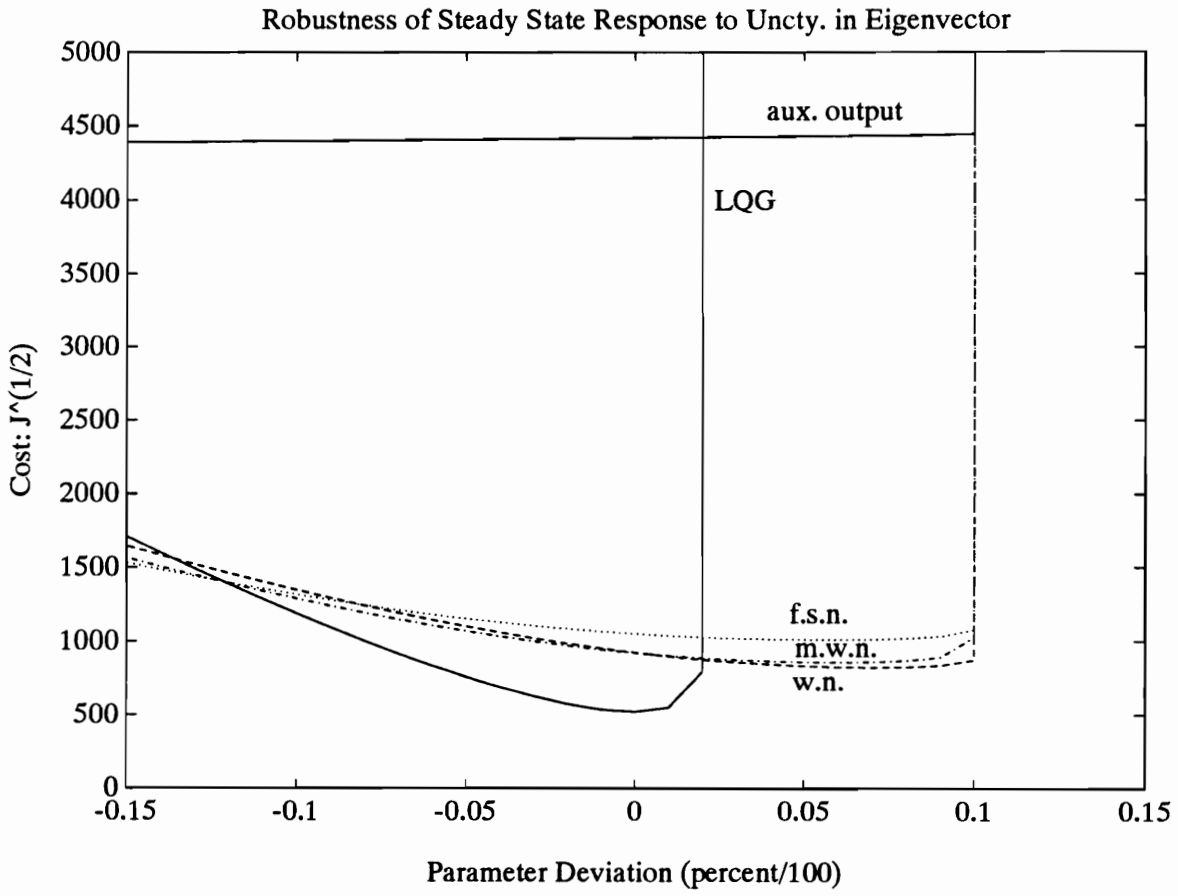

Figure 6.4: Steady State Performance and Stability Tradeoff for Uncertainty in Eigenvector

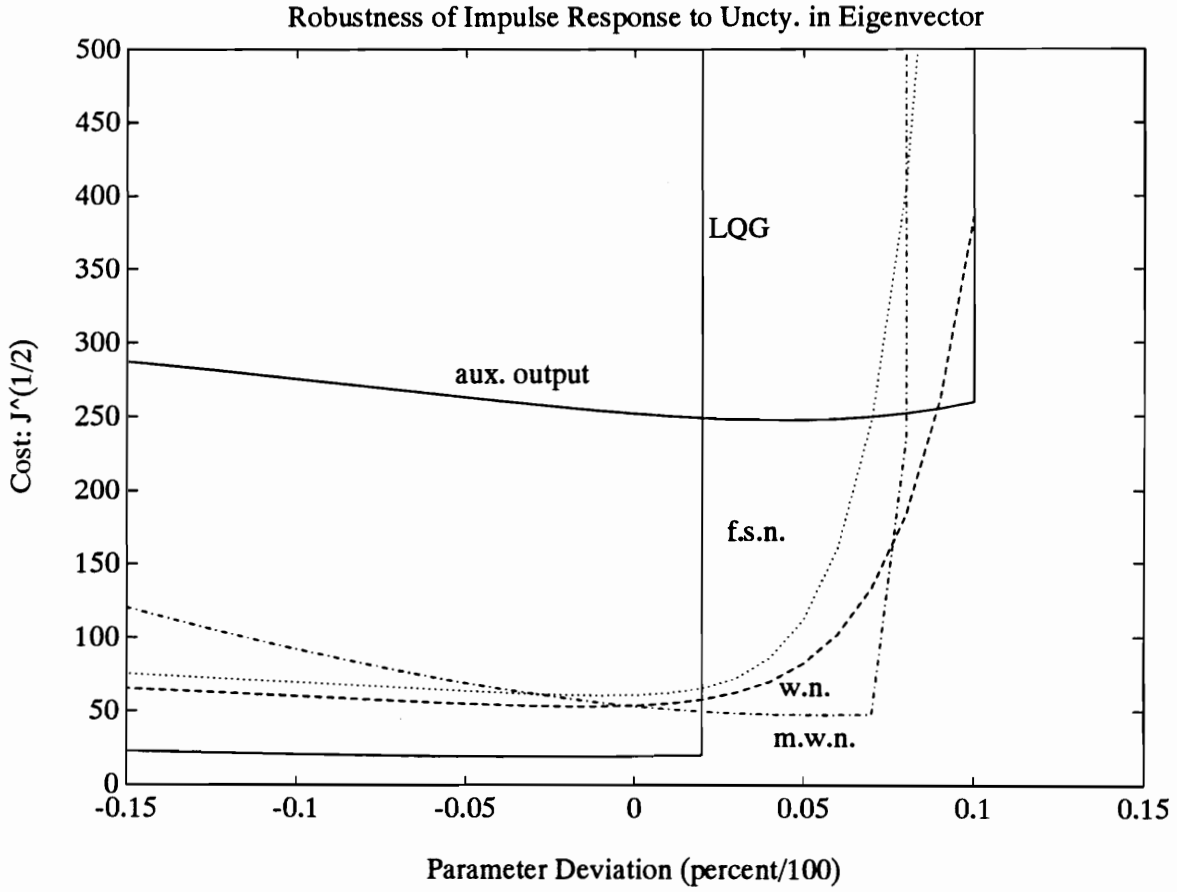

Figure 6.5: Transient Performance and Stability Tradeoff for Uncertainty in Eigenvector

one independent uncertainty is present, the factorizations must be adjusted in order to exercise every degree of freedom available. Analysis of a few different factorizations should reveal one that is very nearly the best possible for any given uncertainty.

This technique was applied to the eigenvector uncertainty problem in section 6.2B. First, the white noise model was chosen for the auxiliary inputs, because of the good performance that results and because of the simplicity of design. Then, holding M_b and N_b constant, the $\pm 10\%$ stability margin was attained using several different combinations of values for μ_b and ρ_b. Those values are detailed in Table 6.2 for eight different controller designs. The performance/stability results of each design are plotted

Table 6.2: Controller Design Parameters Corresponding to Eight Different Factorizations

	Case Number							
	1	2	3	4	5	6	7	8
μ_b	0	10	20	30	40	50	60	70
ρ_b	3.4	0.9	0.5	0.3	0.15	0.1	0.05	0

in Figure 6.6, where the top curve represents Case 1, and the curve for each successive decrease in nominal cost represents the next higher case number. These results are typical of a those from a number of such experiments, including analysis of controllers based on a multiplicative white noise model of the auxiliary inputs. Normally, either the auxiliary input modeling or auxiliary output penalty alone provide virtually the best overall performance possible, although neither one does so consistently. So far, no significant improvement has been achieved by combining auxiliary inputs and outputs.

Blelloch and Mingori (1990) point out that one particular factorization has a certain intuitive value for the problem of multiple natural frequency uncertainties (actually for uncertainties in the $-\omega_i^2$ elements of the A-matrix). Namely, by factoring ΔA such that the nonzero elements of M_a and N_a are equal, the auxiliary outputs lead to a penalty on the elastic strain energy of the flexible structure. However, this approach to factoring ΔA has no bearing on the relative emphasis to be placed on the auxiliary inputs versus the auxiliary outputs. It only decides what relative

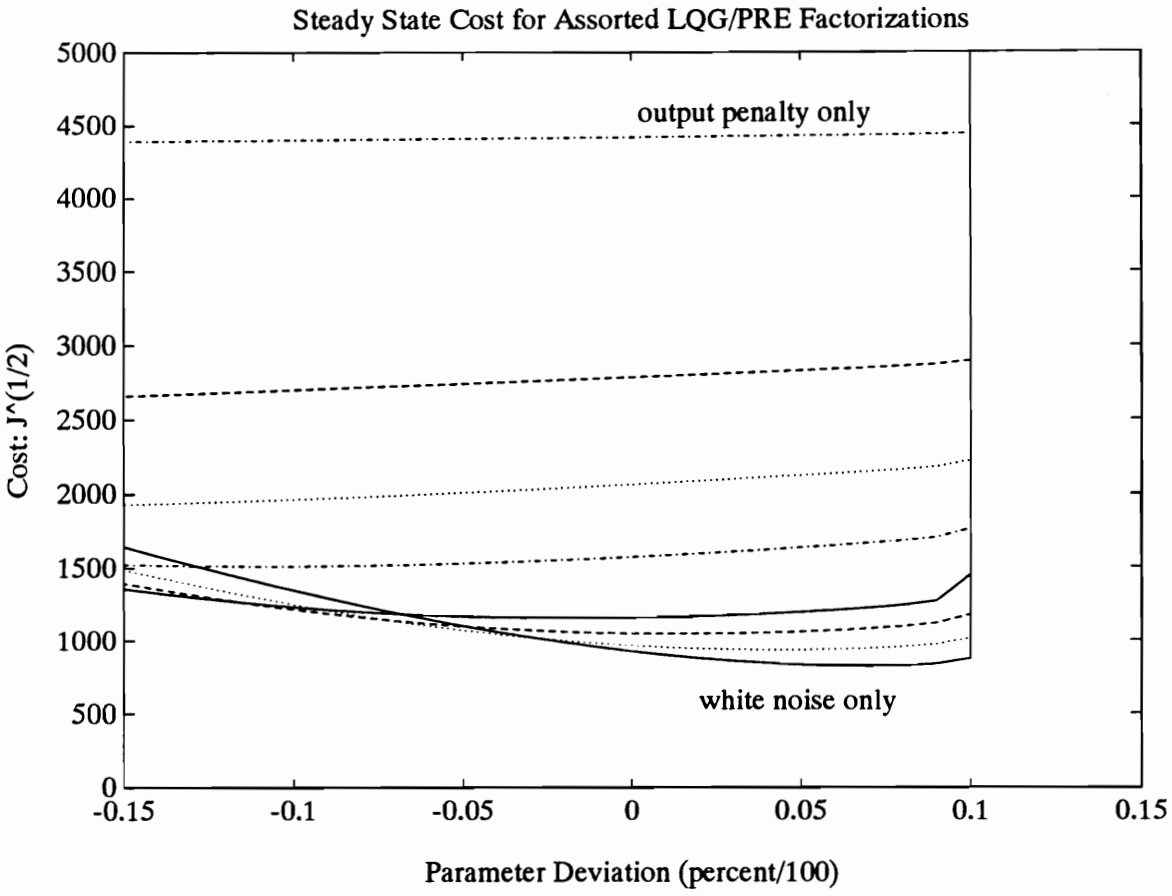

Figure 6.6: Comparative Performance for Several Different LQG/PRE Factorizations

importance should be placed on the various parameter uncertainties. Also, no results were given to draw comparisons with other factorizations.

6.3 Itô vs. Stratonovich Noise

In the examples considered in the previous section, the multiplicative white noise models were interpreted in the sense of Itô. Recall from section 3.5 that in order to interpret multiplicative white noise as Stratonovich noise, we need only to make the following substitutions in the controller design equations:

$$A \leftarrow A + \frac{1}{2}\sum_{i=1}^{p}\gamma_i A_i^2\,, \tag{3.73}$$

$$B \leftarrow B + \frac{1}{2}\sum_{i=1}^{p}\gamma_i A_i B_i\,,$$

$$C \leftarrow C + \frac{1}{2}\sum_{i=1}^{p}\gamma_i C_i A_i$$

The Stratonovich noise interpretation was applied to the natural frequency uncertainty problem of section 6.2A (using auxiliary input modeling only) so as to compare the controller performance with that of the Itô noise design already discussed. For a $\pm 10\%$ stability margin requirement, we have the following matrix substitutions for A and C:

$$A \leftarrow A + \frac{1}{2}\gamma_1 A_1^2 \tag{6.15}$$

$$A \leftarrow A + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.656 & 0.32 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

where on the right-hand side of (6.15):

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -96525 & -4.7987 & 0 & .70334 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -142120 & -75.398 \end{bmatrix}$$

and

$$C \leftarrow C + \tfrac{1}{2}\gamma_1 C_1 A_1 \tag{6.16}$$

$$C \leftarrow C + \begin{bmatrix} .656 & .32 & 0 & 0 \end{bmatrix}$$

where on the right-hand side of (6.16):

$$C = \begin{bmatrix} -96525 & -4.7987 & 0 & .70334 \end{bmatrix}$$

The stability specification was met with auxiliary input noise of such a low intensity that the modifications to A and C were insignificant. The difference in the stability and performance characteristics of the closed-loop systems for the two noise interpretations was therefore not discernible.

6.4 Reduced-Order Controller

A much higher order model was used to investigate the benefits of optimal reduced-order control. This study employed a 4-mode model of the plate, along with the second-order smoothing filter and second-order disturbance shaping filter to create a twelfth-order augmented plant. The design objectives for the reduced-order controllers were threefold: (1) to provide 15 dB of rejection at the assumed disturbance center frequency (60 Hz), (2) to minimize the increase in the linear quadratic cost over that of the full-order design, and (3) to provide some transient suppression along with the steady state rejection.

The disturbance cancellation method of section 5.2 was found to provide over 15 dB of rejection at 60 Hz for a full-order controller. Combining 60 Hz cancellation with a penalty on the modal energies supplied some transient suppression as well. The cost functional matrices which resulted were similar to those used for the robustness studies:

$$R_1 = \mathrm{diag}\{\Omega^2, I_{4\times 4}, 0, 0, 0, 1.0002\}, \quad R_2 = 1, \quad R_{12} = \begin{bmatrix} 0_{11\times 1} \\ 1.0001 \end{bmatrix} \tag{6.17}$$

where the 4×4 natural frequency matrix Ω is defined in (6.3). As before the noise

intensity parameter was chosen to be $\rho_v = 10^{-6}$. The simple model for g_{22} was used: $g_{22} = I$. Then V_1, V_2, and V_{12} were computed from (6.12).

A full-order controller was designed for this model, then its performance, frequency response, and time response characteristics were compared with optimal controllers of orders 4, 3, and 2. Table 6.3 quantifies the comparative performance of all four controllers.

Table 6.3: LQ Cost and Disturbance Rejection vs. Controller Order

Order	12	4	3	2
Rejection	15.9 dB	17.9 dB	[illegible].8 dB	5.3 dB
LQ Cost	1576	1623	1633	2753

Approximately 15 dB of rejection at 60 Hz is retained for a controller order as low as 3, and very little rejection is lost by substituting this 3rd-order controller for a 12th-order controller. However, the 2nd-order controller is markedly worse in performance. Note how the linear quadratic cost rises with each reduction in controller order. On the transition from a 3rd-order to a 2nd-order controller, the increase in cost is very steep.

Figure 6.7 shows the magnitude frequency responses of the first modal acceleration to the disturbance shaker input for the open-loop system and three closed-loop designs. The 3rd-order controller appears to be the design of choice. We see here how closely the closed-loop system behavior with this controller resembles that of the system with full-order controller. When the controller order is restricted to less than three, however, the optimality (in the full-order sense) of the design begins to break down. All three controllers provide some transient suppression by lowering the open-loop peak at 49 Hz due to the lightly damped first mode at that frequency. However, the 2nd-order controller fails to produce a notch at the 60 Hz disturbance center frequency. By subtracting the open-loop response from each of the closed-loop responses, we arrive at the disturbance rejection plots in Figure 6.8. The vertical line is drawn at 60 Hz to indicate where the 15 dB of rejection is desired. The 2nd-order controller is not able to simultaneously dampen the first mode and provide good rejection at 60 Hz. The frequency response plots of the controllers in Figure 6.9

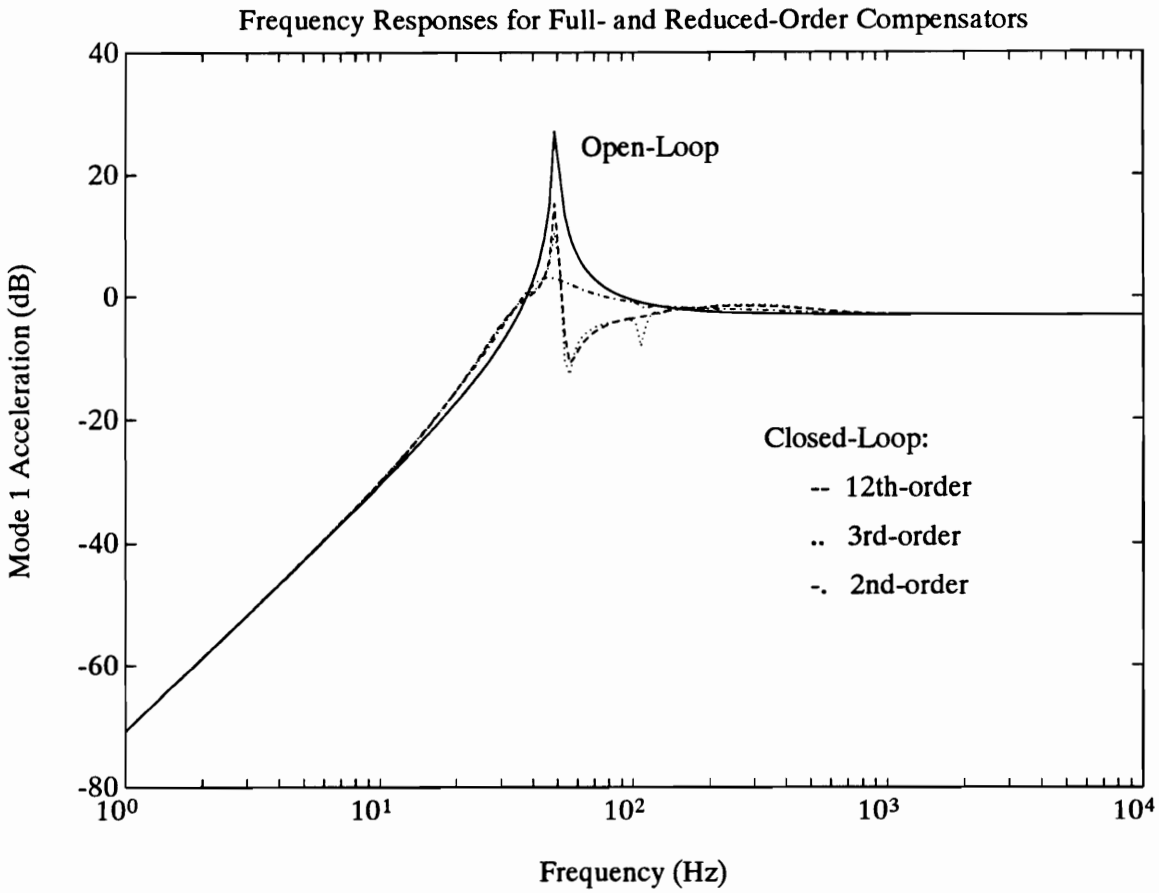

Figure 6.7: Frequency Responses for Full- and Reduced-Order Compensators

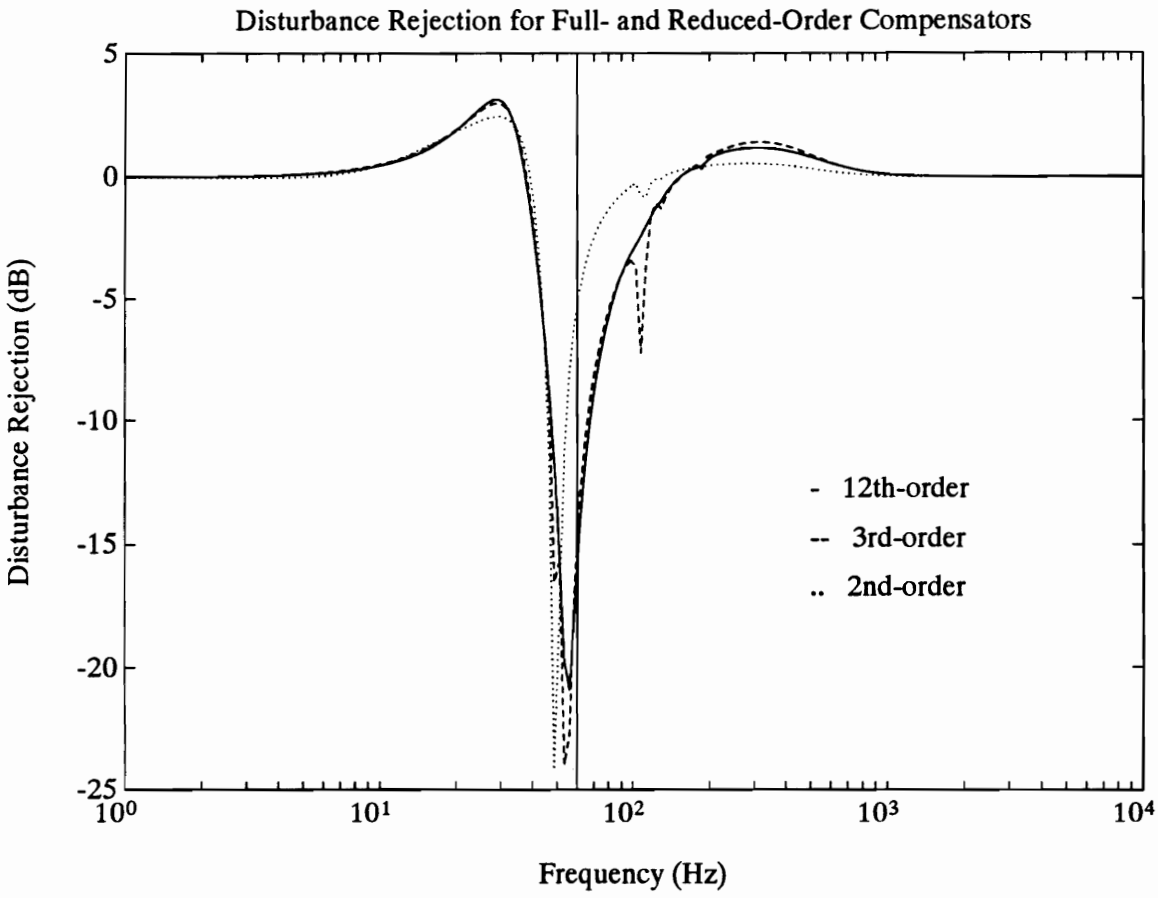

Figure 6.8: Disturbance Rejection for Full- and Reduced-Order Compensators

demonstrate the cause of this limitation. These curves represent the response of the transfer function from the first modal acceleration input to the control output. The response of the 2^{nd}-order controller can approach the full-order response only at the asymptotes. The rapid change of the full-order response in the 49–60 Hz region cannot be matched at the same time by a transfer function with only two poles. In order to compare the time responses of the closed-loop system with 12^{th}- and 3^{rd}-order controllers, a 60 Hz disturbance was applied and the first modal acceleration was measured. These results are shown in Figure 6.10. Again, we see how little performance is sacrificed to achieve this dramatic reduction in controller order.

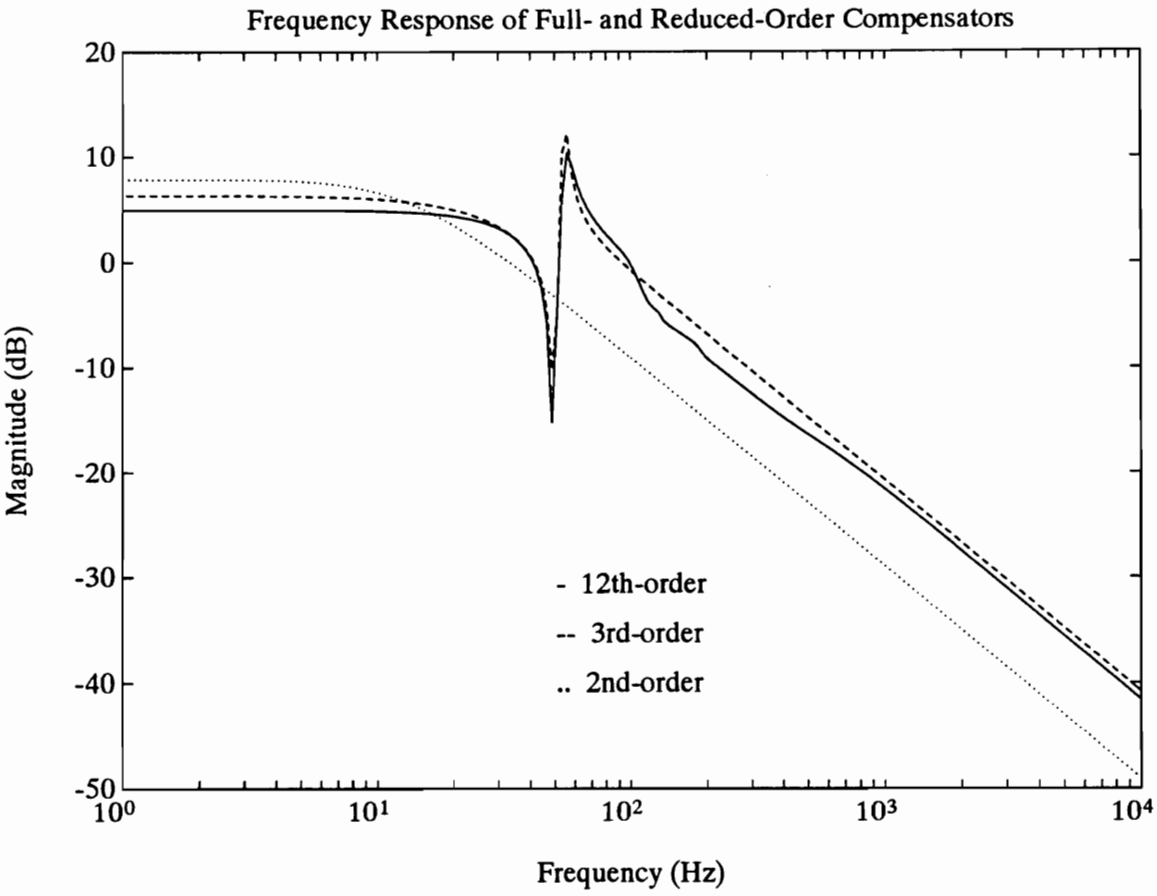

Figure 6.9: Frequency Response of Full- and Reduced-Order Compensators

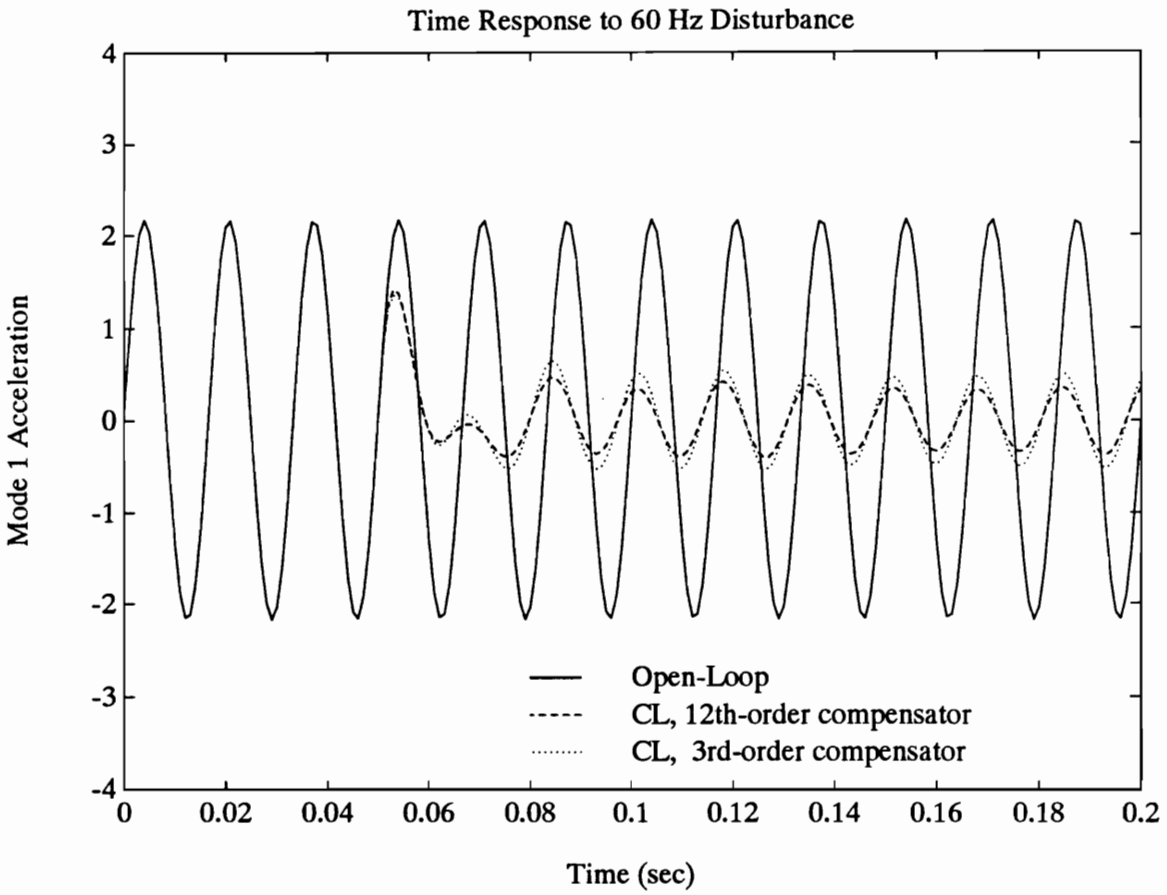

Figure 6.10: Time Response for Full- and Reduced-Order Compensators

7. Application to Simply Supported Plate

7.1 Stability Robustness Problem and its Solution

For experimentation with the simply supported plate hardware, a sinusoidal 60 Hz disturbance was applied at the disturbance shaker. When standard LQG controllers were implemented, very little disturbance rejection could be accomplished. Large controller authorities (in the form of a small penalty, R_2, on the control input) resulted in system instability due modeling errors of the plate. In order to determine possible sources of the modeling errors, the control penalty was gradually lowered and the LQG controller was repeatedly redesigned until instability resulted. Then the sensitivity of the model to deviations in its parameters was analyzed. The augmented plant model from which the controllers were designed was the same 4-mode, 12^{th}-order model described in section 6.4. The covariance matrices were also the same, with the minor exception that g_{22} was computed as in (6.2), yielding a slightly different value for V_2. In order to keep matters simple, direct 60 Hz disturbance cancellation was not attempted. Rather, a penalty was applied to the control input and the modal energy of the first mode only. The plant, cost, and covariance matrices were then translated to zero-order hold equivalent form, and the LQG controllers were designed directly in discrete-time. The following (continuous-time) cost functional matrices resulted in instability of the simply supported plate:

$$R_1 = \text{diag}\{\omega_1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0\}, \quad R_2 = 2 \times 10^5, \quad R_{12} = 0 \tag{7.1}$$

Continuous-time analysis of the closed-loop system model with LQG controller designed from these cost matrices revealed that the system was robust to all of the plate's damping ratios and modal frequencies except one. A -5% or a $+6\%$ deviation in the second natural frequency was found to destabilize the system model. The nominal natural frequency for the second mode (from Table 6.1) is approximately 109 Hz. Therefore, if the plant model were otherwise accurate and the actual second natural frequency of the plate were anywhere outside the range 103.5–115.5 Hz, instability would result. In order to correct this problem, a requirement was placed on the stability margin with respect to ω_2 — to improve the margin to $\pm 10\%$. Three different modified LQG designs were applied to just meet this specification: (1) white noise auxiliary input modeling, (2) multiplicative white noise input modeling, and

(3) auxiliary output penalty. Since the simultaneous application of auxiliary inputs and outputs was found in Chapter 6 to have little success, no combination of the two was attempted. Also, frequency-shaped auxiliary input noise was not considered, because of its relative ineffectiveness during experimentation with the model and because it would result in a controller of order greater than twelve, meaning that the controller would be too slow for implementation without using some kind of order reduction.

The factorization used to compensate for the natural frequency uncertainty is analogous to the one used in section 6.2A. for the 1-mode model:

$$M_a = \begin{bmatrix} 0_{5\times 1} \\ 1 \\ 0_{6\times 1} \end{bmatrix}, \quad M_c = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad N_a = N_c = \begin{bmatrix} 0 & 2.05 & 0_{1\times 3} & 1 & 0_{1\times 6} \end{bmatrix} \qquad (7.2)$$

The design parameters that the three robust designs used to meet the $\pm 10\%$ stability margin requirement are:

(1) White noise ..... $\mu_a = \mu_c = .22$

(2) Multiplicative white noise..... $\gamma_1 = 2.89$,

$$A_1 = \begin{bmatrix} & & & 0_{5\times 12} & & & \\ 0 & 2.05 & 0 & 0 & 0 & 1 & 0_{1\times 6} \\ & & & 0_{6\times 12} & & & \end{bmatrix},$$

$$C_1 = \begin{bmatrix} & & & 0_{1\times 12} & & & \\ 0 & 2.05 & 0 & 0 & 0 & 1 & 0_{1\times 6} \\ & & & 0_{2\times 12} & & & \end{bmatrix}$$

(3) Auxiliary output penalty ..... $\rho_a = \rho_c = 7\times 10^4$

The performance/stability tradeoff results for all three candidate designs is depicted in Figure 7.1, along with the LQG results. Even though the white noise auxiliary input model gave very poor nominal performance in the 1-mode natural frequency uncertainty example of section 6.2A., that same method of robust design sacrificed virtually no nominal performance here and was clearly the best performer overall. The multiplicative white noise design did provide a larger stability margin for positive

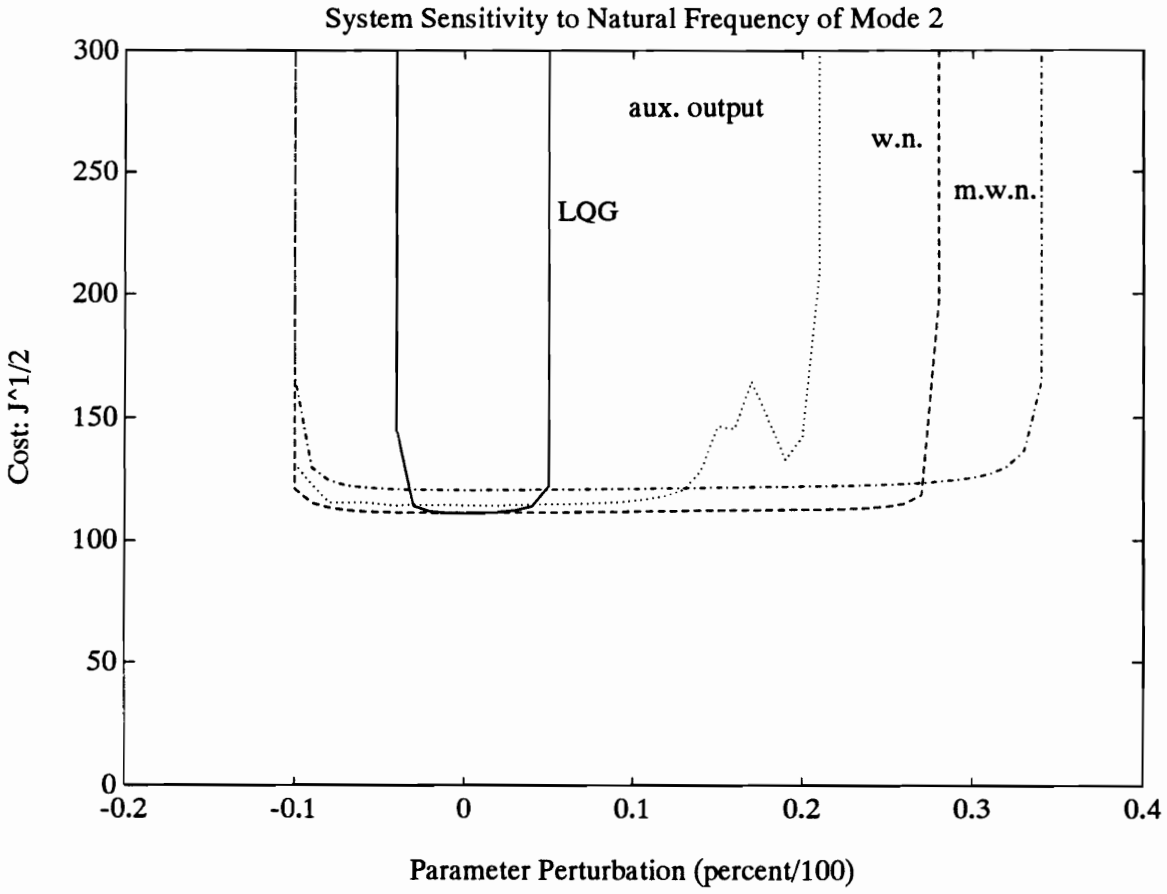

Figure 7.1: Performance/Stability Robustness Tradeoff for Uncertainty in ω_2

deviations in ω_2, but the natural frequency is surely known to within 28%, so the additional stability margin to $+34\%$ is of no value.

The white noise auxiliary input model requires only that we make the following two substitutions for the original covariance matrices:

$$V_1 \leftarrow V_1 + \mu_a M_a M_a^{\mathrm{T}} \tag{7.3}$$
$$V_2 \leftarrow V_2 + \mu_c M_c M_c^{\mathrm{T}}$$

Then we use LQG to arrive at a different Kalman filter (but the same regulator gains). This parameter robust controller design was implemented in the hardware and did in fact stabilize the system. The measured time responses of mode 1 and mode 2 accelerations to the 60 Hz disturbance are shown in Figures 7.2 and 7.3, respectively. After the loop was closed, the LQG controller rejected the response in mode 1, but only temporarily, as mode 2 was driven unstable. The robust controller stabilized mode 2 (and the entire system) while giving up only a small amount of rejection in mode 1.

7.2 Effectiveness of Reduced-Order Controllers

Robust optimal controllers of orders 3 and 4 were designed using the same white noise auxiliary input model and the same cost and covariance matrices used to design the full-order robust controller discussed above. For this level of controller authority, the iterative relaxation method of section 5.4 proved to be sufficient for solving the optimal projection equations. Since a sampled-data version of the optimal projection equations was not available, the reduced-order controllers were designed in continuous-time, then discretized by means of a bilinear transformation. The 3^{rd}- and 4^{th}-order controllers significantly reduced the computational delay caused by the full-order controller, so the one-sample-interval delay was ignored on the first attempt. Then, in each case, a second controller was designed based on a 2^{nd}-order Padé approximation of the delay. This time delay model raised the order of the augmented plant from 12 to 14.

The sample rates for all three controller orders are given in Table 7.1.

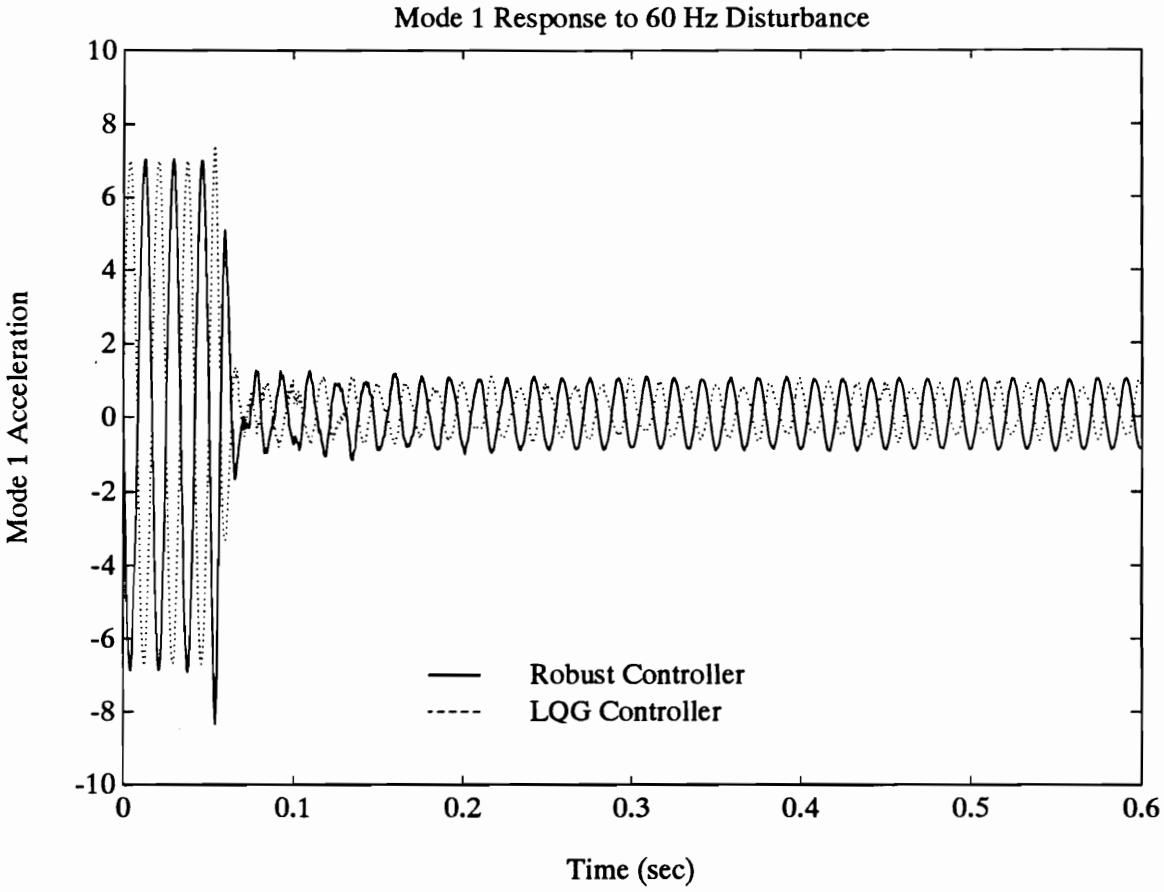

Figure 7.2: Robust Control — Mode 1 Response to 60 Hz Disturbance

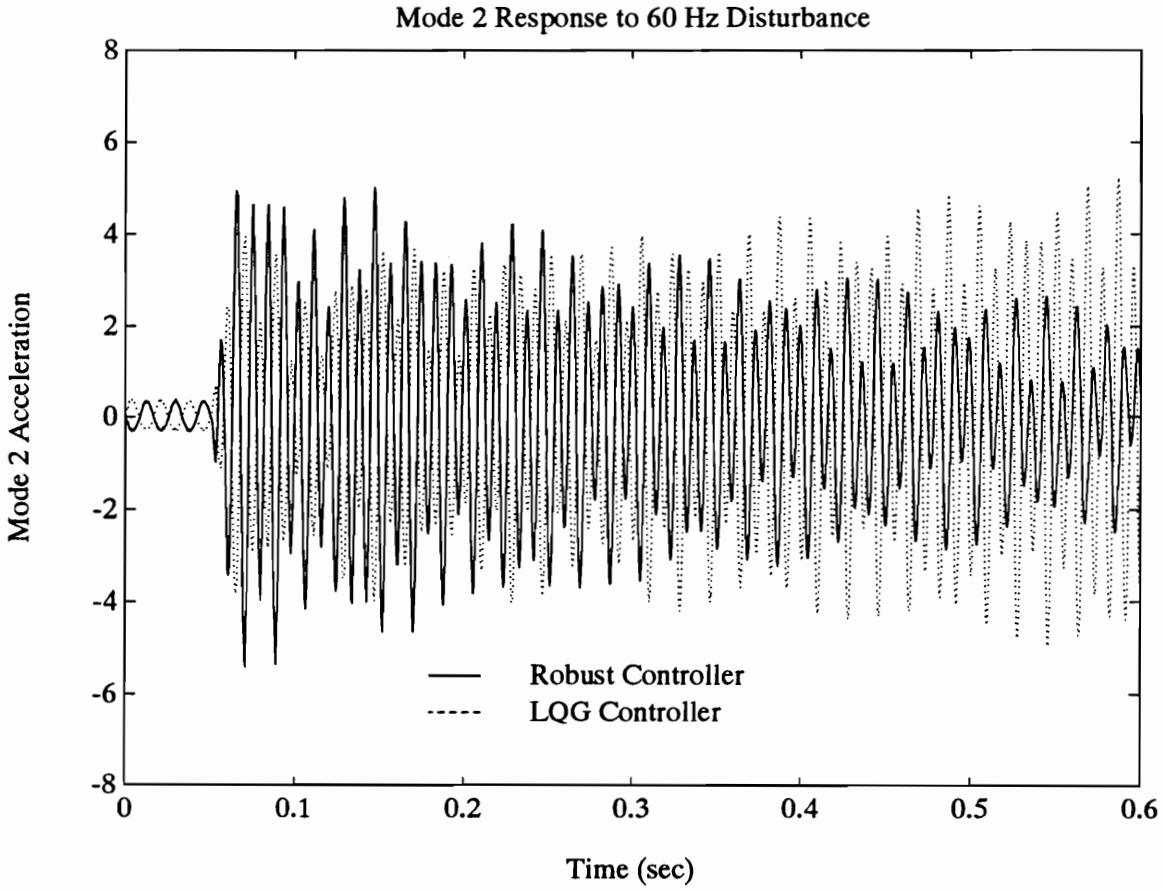

Figure 7.3: Robust Control — Mode 2 Response to 60 Hz Disturbance

Table 7.1: Sample Rates for Full- and Reduced-Order Controllers

Controller Order	12	4	3
Sample Rate (Hz)	1997	3615	4023

The improvement of the reduced-order sample rates over that of the full-order design was not as dramatic as one might expect, because the full-order controller was able to take advantage of the structure of the modal state-space realization of the plant. It exploited the sparseness of the block-diagonal system matrix and the sparseness of the output matrix to speed up the Kalman filter considerably.

In order to examine the effect of computational time delay on control system performance, an unmodeled delay of one sample interval was introduced into the full-order robust controller by preventing the Kalman filter from predicting one step ahead. The delay did not cause instability, but controller performance suffered somewhat. Figure 7.4 shows the comparative responses of the first mode to a 60 Hz disturbance for full-order controllers with and without the time delay.

Even though continuous-time analysis predicted the 3^{rd}-order controller would stabilize the system, it did not, regardless of whether the computational delay was modeled. The 4^{th}-order controller stabilized the system only when the delay was modeled, and in that case gave quite good performance. Again, a 60 Hz disturbance was applied to the plate, and the responses of modes 1 and 2 were measured. In Figures 7.5 and 7.6, these results are compared with those of the full-order controller. The 4^{th}-order controller does give up a small amount of disturbance rejection, but increases the sample rate significantly. The significance of this experiment is its demonstration that parameter robust and reduced-order control can be accomplished simultaneously to stabilize a flexible structure and reject disturbances with minimal computational delay.

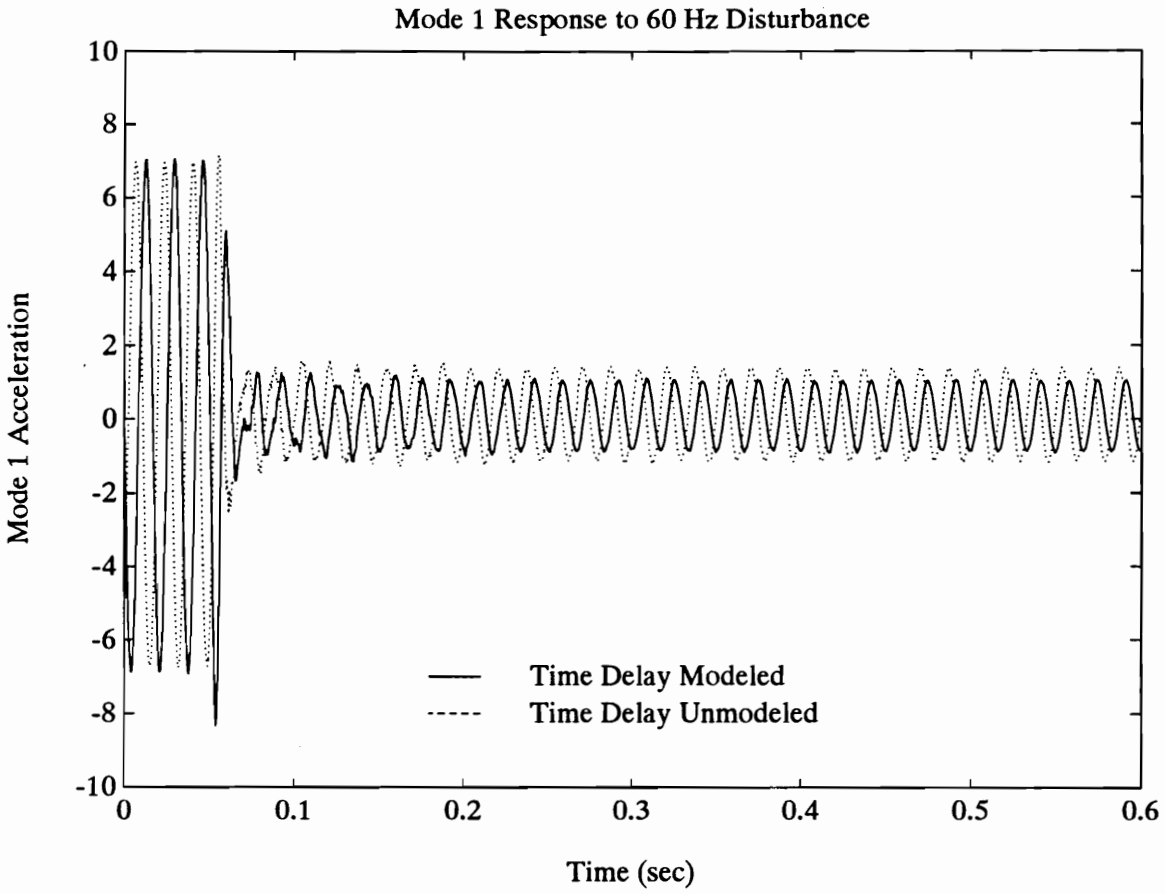

Figure 7.4: Response of Full-Order Robust Controllers With and Without Time Delay

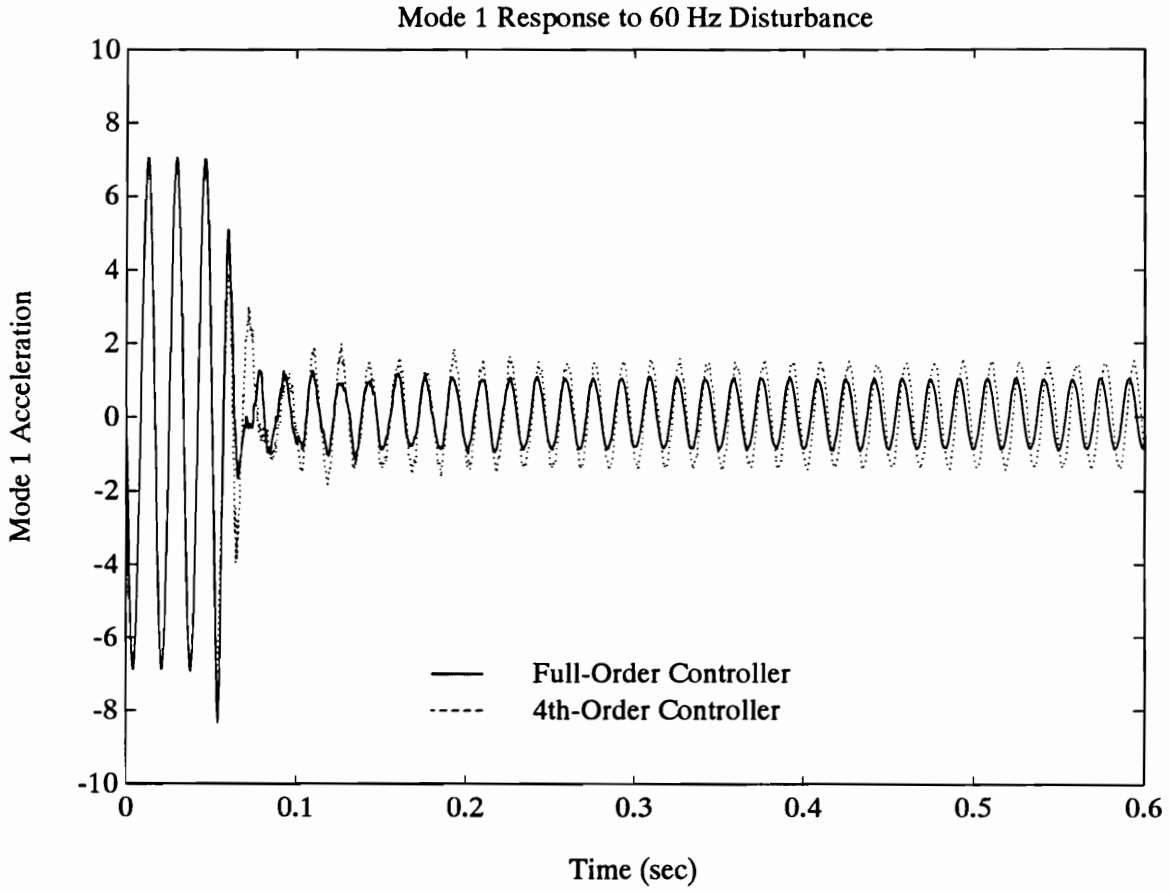

Figure 7.5: Robust Reduced-Order Control — Mode 1 Response to 60 Hz Disturbance

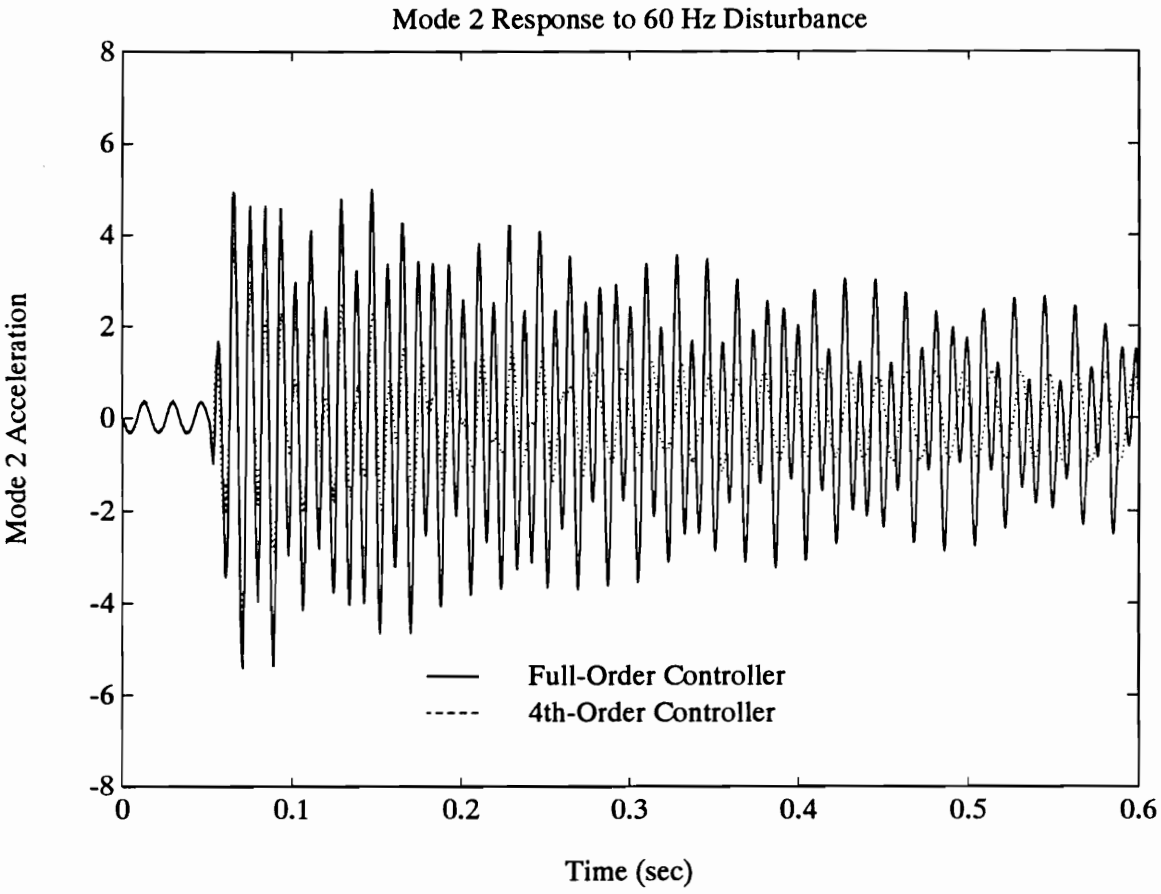

Figure 7.6: Robust Reduced-Order Control — Mode 2 Response to 60 Hz Disturbance

8. Conclusions

As stated in the introduction, the objectives of this work were to compare a number of different parameter robustness techniques based on the LQG/PRE error model, to investigate the limitations on controller order reduction, and to evaluate the combined robust minimal order design.

8.1 Robustness

The few parameter robust control problems considered in Chapters 6 and 7 serve as a counterexample to the proposition that any one of the robust design techniques studied is consistently superior to the others. In fact, the white noise, multiplicative white noise, and auxiliary output designs each provided the best performance for at least one problem. The best performer in one case was found to be the worst performer, by far, in another. Also, the frequency-shaped noise design proved to be a contender at times. Therefore, no strong conclusions can be drawn on the basis of performance alone.

Although the combination of an auxiliary input and auxiliary output model provides greater flexibility in the design, experience seems to indicate that only rarely is any advantage gained from such a combination. Even then, the advantage appears to be very slight. This result is useful, since the elimination of this flexibility from the design allows us to eliminate the complexity of choosing among an infinite number of MLN-factorizations for each parameter uncertainty. Excellent performance and stability characteristics are attained consistently by choosing the better of the two extremes — auxiliary inputs only or auxiliary outputs only. Assuming this simplification, we may discuss the comparative merits of the time-domain method, the frequency-domain method, and LQG/PRE in complete generality by considering the three auxiliary input models and the auxiliary output model each on its own.

The multiplicative white noise model for the auxiliary inputs was the most consistent of the four models in providing good performance for a specified stability margin. Relative to the other models, this one tends to show added improvement when the signal amplitudes of the plant vary greatly over time, such as is the case when the plant is subjected to transient disturbances. However, the performance provided by this model was repeatedly matched or beaten by simpler models.

The frequency-shaped noise model has not proven to be superior for any example studied thus far. Although it has in some cases resulted in a controller with better performance than one designed using a white noise model, the added design complexity would seldom justify the small advantage this model might provide. The intent of frequency shaping is to model the low and high frequency rolloff of the auxiliary input frequency response. However, the frequency content of white noise outside the passband of the system (which poorly models reality) is greatly attenuated and apparently has no significant harmful effect. In the presence of transient disturbances, the frequency-shaped noise model tends to impair performance, since its design relies on the frequency spectrum of an assumed steady state disturbance.

The white noise auxiliary input model and the auxiliary output penalty each yield poor performance in some cases. However, one or the other consistently provides very good performance, and the simplicity of their associated controller design methods has important advantages. Both models require only the modification of cost or covariance matrices in a standard LQG design. Therefore, they do not require the iterative design computations of the multiplicative white noise model, nor do they increase controller order, as does the frequency-shaped noise model. The rapidity of controller design allows more effort to be devoted to analysis, which is crucial when there are multiple independent parameter uncertainties.

8.2 Controller Order

For a 12^{th}-order augmented plant (14^{th}-order with the computational delay model) we were able to design a 3^{rd}-order controller in theory and a 4^{th}-order controller in practice without sacrificing a great deal of performance. Instead of determining the minimal controller order that is practical by successively designing optimal controllers of many different orders, we may examine the frequency response of the full-order controller to find out what order a controller must have to display the significant features of that response. That order serves as a good first iteration to the minimum controller order desirable. The addition of more modes to the model would not tend to raise the order of a controller necessary to provide a good response, because any additional modeled modes would necessarily be less dominant. This effect was demonstrated by the addition of the 2^{nd}-order computational delay model to the plant

in Chapter 7. Not only did the minimum controller order fail to rise, it fell due to the more precise modeling of the plant.

8.3 Parameter Robust Reduced-Order Design

The application of the optimal projection equations to the auxiliary input and output models of LQG/PRE is new and proved to be very successful. Its application to the LQG/PRE error model in general makes these methods of parameter robust controller design much more powerful. The reduction in computational delay makes active vibration control possible when it otherwise may not have been. Also, optimal reduced-order design allows us to model more dynamics to give the controller a more accurate picture of the plant, thus reducing another source of uncertainty — unmodeled dynamics. The additional dynamics may include more disturbance modeling, smoothing filter modeling, and time delay modeling. In the robust reduced-order design of Chapter 7, the sacrifice in performance may have been reduced or eliminated by giving up some of the improvement in sample rate in favor of modeling more high frequency modes. This experiment was not carried out, because it would have required rewiring of the smoothing filter, which was designed for a 4-mode model. However, the 4^{th}-order robust controller as designed did stabilize the system and provide a significant amount of disturbance rejection.

8.4 Directions for Further Study

The most conspicuous shortcoming of the design methods and algorithms presented in this study is the lack of a method of designing optimal reduced-order and multiplicative white noise based controllers directly in discrete-time. This deficiency results in the introduction of two additional sources of error. Firstly, the design of an optimal continuous-time controller followed by the transformation of that controller into a discrete-time equivalent is a suboptimal process. Secondly, for the continuous-time design an approximation is necessary to create a finite-dimensional model of the inevitable computational delay. In a continuous-time augmented plant model, the computational delay is not only imperfectly modeled, it may also require the addition of high order dynamics in order to be modeled adequately, particularly when the plant is multi-input multi-output. The solution of the optimal reduced-order control problem

for sampled-data systems is of particular interest, since it would have widespread application in LQG-based design.

References

Anderson, Brian D. O., and Yi Liu. 1989. Controller reduction: Concepts and approaches. *IEEE Transactions on Automatic Control* 34 (August): 802–812.

Arnold, Ludwig. 1974. *Stochastic Differential Equations: Theory and Applications.* New York: John Wiley & Sons.

Balas, Gary J., and John C. Doyle. 1990. Identification of flexible structures for robust control. *IEEE Control Systems Magazine* 10 (June): 51–58.

Bender, Carl M., and Steven A. Orszag. 1978. *Advanced Mathematical Methods for Scientists and Engineers.* New York: McGraw-Hill.

Bernstein, Dennis S. 1987. Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives. *IEEE Transactions on Automatic Control* 32 (December): 1076–84.

Bernstein, Dennis S., and Scott W. Greeley. 1986. Robust controller synthesis using the maximum entropy design equations. *IEEE Transactions on Automatic Control* 31 (April): 362–4.

Bernstein, Dennis S., and Wassim M. Haddad. 1989. LQG control with an $\mathcal{H}_\infty$ performance bound: A Riccati equation approach. *IEEE Transactions on Automatic Control* 34 (March): 293–305.

Bernstein, Dennis S., and David C. Hyland. 1988a. Optimal projection equations for reduced-order modelling, estimation, and control of linear systems with multiplicative white noise. *SIAM Journal of Optimization Theory and Applications* 58 (September): 387–409.

Bernstein, Dennis S., and David C. Hyland. 1988b. Optimal projection for uncertain systems (OPUS): A unified theory of reduced-order, robust control design. In *Large Space Structures: Dynamics and Control,* edited by S. N. Atluri and A. K. Amos, 263–302. Berlin: Springer-Verlag.

Blelloch, P. A., and D. L. Mingori. 1990. Robust Linear Quadratic Gaussian Control for Flexible Structures. *Journal of Guidance, Control, and Dynamics* 13 (January–February): 66–72.

Dailey, R. Lane. 1990. Lecture notes. *Workshop on $\mathcal{H}_\infty$ and μ Methods for Robust Control,* San Diego, CA.

Doyle, John C. 1984. Lecture notes in advances in multivariable control. *ONR/Honeywell Workshop,* Minneapolis, MN.

Doyle, John C. 1985. Structured uncertainty in control system design. *Proceedings of the 24th Conference on Decision and Control.* Ft. Lauderdale, FL. 260–65.

Doyle, John C., Kathryn Lenz, and Andy Packard. 1987. Design examples using μ-synthesis: Space shuttle lateral axis fcs during reentry. In *Modelling, Robustness and Sensitivity Reduction in Control Systems,* edited by Ruth F. Curtain, 127–54. Berlin: Springer-Verlag.

Doyle, John C., and Gunter Stein. 1981. Multivariable feedback design: Concepts for a classical/modern synthesis. *IEEE Transactions on Automatic Control* 26 (February): 4–16.

Ewing, George M. 1985. *Calculus of Variations with Applications.* New York: Dover.

Gelb, Arthur (editor). 1984. *Applied Optimal Estimation.* Cambridge, MA: M.I.T. Press.

Gikhman, I. I., and A. V. Skorokhod. 1969. *Introduction to the Theory of Random Process.* Philadelphia: W. B. Saunders Company.

Glover, Keith. 1984. All optimal Hankel norm approximations of linear multivariable systems and their L_∞ error bounds. *International Journal of Control.* 39 (June): 1115–93.

Graham, Alexander. 1981. *Kronecker Products and Matrix Calculus: with Applications.* New York: John Wiley & Sons, Inc.

Gupta, Narendra K. 1980. Frequency-shaped cost functionals: Extension of linear-quadratic-Gaussian design methods. *Journal of Guidance, Control, and Dynamics* 3 (November–December): 529–35.

Hyland, David C. 1982. Minimum information stochastic modelling of linear systems with a class of parameter uncertainties. *Proceedings of the American Control Conference.* 620–27.

Hyland, David C., and Dennis S. Bernstein. 1984. The optimal projection equations for fixed-order dynamic compensation. *IEEE Transactions on Automatic Control* 29 (November): 1034–37.

Hyland, David C., and Appasaheb N. Madiwale. 1981. A stochastic design approach for full-order compensation of structural systems with uncertain parameters. *Proceedings of the AIAA Guidance and Control Conference.* Albuquerque, NM. 324–32.

Itô, K. 1944. Stochastic integral. *Proc. Imp. Acad. Tokyo* 20: 519–24.

Jazwinski, Andrew H. 1970. *Stochastic Processes and Filtering Theory.* New York: Academic Press.

Kalman, R. E. 1964. When is a linear control system optimal? *Journal of Basic Engineering* 86 (March): 51–60.

Kreindler, Eliezer, and Antony Jameson. 1972. Optimality of linear control systems. *IEEE Transactions on Automatic Control* 17 (June): 349–51.

Kwakernaak, Huibert, and Raphael Sivan. 1972. *Linear Optimal Control Systems.* New York: John Wiley & Sons.

Laub, Alan J. 1980. Computation of "balancing" transformations. *Proceedings of the 1980 Joint Automatic Control Conference.* Vol. 2, FA8-E.

Lin, Jong-Yin. 1989. *Robust Design of LQG Controllers.* Ph.D. Dissertation. University of California, Los Angeles, CA.

Maciejowski, J. M. 1989. *Multivariable Feedback Design.* Wokingham, England: Addison-Wesley.

Moore, B. C. 1981. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control* 26 (January): 17–32.

Richter, Stephen. 1987. A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: Existence, convergence and global optimality. *Proceedings of the American Control Conference.* 1527–31.

Richter, Stephen, and Emmanuel G. Collins, Jr. 1989. A homotopy algorithm for reduced order compensator design using the optimal projection equations. *Proceedings of the 28th Conference on Decision and Control.* Tampa, FL. 506–11.

Richter, Stephen L., and Raymond A. DeCarlo. 1983. Continuation methods: Theory and applications. *IEEE Transactions on Automatic Control* 28 (June): 660–65.

Rubenstein, Stephen P. 1991. *An Experiment in State-Space Vibration Control of Steady Disturbances on a Simply Supported Plate.* Ph.D. Dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

Sievers, Lisa A., and Andreas H. von Flotow. 1989. Comparison of two LQG-based methods for disturbance rejection. *Proceedings of the 28th Conference on Decision and Control.* Tampa, FL. 483–85.

Stengel, Robert F. 1986. *Stochastic Optimal Control: Theory and Application.* New York: John Wiley & Sons.

Stratonovich, R. L. 1966. A new representation for stochastic integrals and equations. *J. SIAM Control* 4: 362–71.

Tahk, Minjea, and Jason L. Speyer. 1987. Modeling of parameter variations and asymptotic LQG synthesis. *IEEE Transactions on Automatic Control* 32 (September): 793–801.

Tahk, Minjea, and Jason L. Speyer. 1989. Parameter robust linear-quadratic-Gaussian design synthesis with flexible structure control applications. *Journal of Guidance, Control, and Dynamics* 12 (July–August): 460–68.

Wong, Eugene, and Moshe Zakai. 1965. On the relation between ordinary and stochastic differential equations. *International Journal of Engineering Science* 3: 213–29.

Wonham, W. Murray. 1967. Optimal stationary control of a linear system with state-dependent noise. *SIAM Journal on Control.* 5: 486–500.

Wonham, W. Murray. 1968. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control.* 6: 681–697.

Wonham, W. Murray. 1970. Random differential equations in control theory. In *Probabilistic Methods in Applied Mathematics,* Vol. 2, edited by A. T. Bharucha-Reid, 131–212. New York: Academic Press.

Vita

Stephen H. Jones was born August 12, 1960, in Greensboro, North Carolina. After undergraduate study at the University of North Carolina and the Georg-August-Universität in Göttingen, West Germany, he completed the B.S.E.E. degree in 1983 at the University of Tennessee. Beginning in 1983 he worked as an aircraft fire control engineer at General Dynamics Corporation in Fort Worth, Texas, and later as a missile guidance and control engineer at Coleman Research Corporation in Orlando, Florida. He earned the M.S. and Ph.D. degrees in electrical engineering from Virginia Polytechnic Institute and State University in 1987 and 1991, respectively, where he served as a teaching and research assistant. Mr. Jones is a member of the AIAA, the IEEE and the IEEE Control Systems Society. His current interests include flight controls, guidance, and navigation, as well as orbital and flight dynamics.

Stephen H. Jones