Robust Controller Design: Minimizing Peak-to-Peak Gain

Munther A. Dahleh
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

Foreward

This research was conducted by Professor Munther A. Dahleh and his group at the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, MIT. This research was supported in part by Wright Patterson Air Force Base, under grant No. F33615-90-c-3608, for the period 8/1990-12/1991.

Several people were involved in the research documented in this report. These are Prof. Khammash, Dr. Diaz-Bobillo, Prof. Shamma, Prof. Voulgaris and Prof. Valavani. Contacts with Wright Patterson were done with Dr. Siva Banda.

Contents

Chapter 1

Introduction

In this report, we address the general problem of designing controllers that minimize the maximum peak-to-peak gain, otherwise known as the ℓ_1 optimal control problem, in the presence of structured uncertainty. Four different problems are discussed:

1. Controller design in the presence of structured uncertainty with a general discussion on the synthesis of ℓ_1 optimal controllers.

2. The ℓ_1 State-feedback problem.

3. The advantages of nonlinear controllers in minimizing the ℓ_∞ induced norm.

4. Peak-to-Peak performance for slowly varying systems.

These problems are addressed in the following four chapters. Chapter 1 is taken from the paper written by Dahleh and Khammsh [8], Chapter 2 is taken from the paper written by Diaz-Bobillo and Dahleh [23], Chapter 3 is taken from the paper written by Dahleh and Shamma [14] and finally Chapter 4 is taken from the paper written by Voulgaris, Dahleh and Valavani [59].

1

Chapter 2

Robust Controller Design in the presence of Structured Uncertainty

This chapter addresses the problem of designing feedback controllers to achieve good performance in the presence of structured plant uncertainty and bounded but unknown disturbances. A general formulation for the performance robustness problem is presented and exact computable conditions are furnished. These conditions are then utilized for synthesizing robust controllers which involves solving ℓ_1 optimization problems. These solutions are computed using the duality theory of Lagrange multipliers. Approximations and computational issues are discussed.

2.1   Introduction

The objective of Robust Control is to provide in a quantitative way the fundamental limitations and capabilities of controller design in order to achieve good performance requirements in the presence of uncertainty. Even though a real system is not uncertain, it is desirable to think of it as such to reflect our imprecise or partial knowledge of its dynamics. On the other hand, uncertainty in the noise and disturbances can be cast under "real uncertainties," as it is practically impossible to provide exact models for such inputs.

Many of the design specifications tend to be concerned with amplitudes of signals. For instance, tracking, disturbance rejection, actuator authority, all result in specifications concerning the maximum amplitudes of signals. On the other hand, disturbances and

noise are usually persistent, bounded, otherwise unknown. This environment motivates a Peak-to-Peak kind of specifications, which is the theme of the ℓ_1 theory.

In this chapter, a general framework for designing controllers that achieve robust Peak-to-Peak performance in the presence of plant perturbations is presented. First, computable necessary and sufficient conditions for performance robustness are presented. The connections between these conditions and spectral properties of positive matrices are highlighted and utilized to simplify the computations. These conditions are in turn used for the synthesis problem which will involve iterative solutions of ℓ_1 minimization problems, the solution of which is obtained by using the duality theory of Lagrange multipliers.

The ℓ_1 problem, formulated in [58], was solved in [10, 11]. The theory was further developed in [17, 23, 24, 45, 54, 55]. The robust stabilization problem in the presence of ℓ_∞-stable perturbations was first analyzed in [9] in the case of unstructured perturbations. In [32], a performance objective was added to the robust stability requirement in the unstructured perturbations case and conditions were provided for robust performance and stability. This led the way to the development of exact necessary and sufficient conditions for robust performance in the presence of structured perturbations [33, 34, 35]. Most of the above results have continuous-time analogs.

There are a number of contributions in this chapter. On one hand, it presents a unified framework for designing robust controllers in the presence of structured uncertainty. Non-conservative conditions to guarantee robust performance are developed directly in terms of the spectral radius of certain matrices capturing the structure of the perturbations. Exact relations between these conditions and linear matrix inequality conditions are then established. On the other hand, the use of linear programming in synthesizing robust controllers is highlighted through the application of the theory of Lagrange multipliers. Through this simple formulation, problems that admit a finite-dimensional equivalence become quite transparent. For the rest of the problems, the theory proves to be quite instrumental in providing upper and lower approximations of the exact problem.

This chapter puts together all of the above development in a way that makes the theory readily usable for design. In general, details that appeared elsewhere will not be presented, however, simple and intuitive proofs of the main ideas will be. Similarities and contrasts between this theory and the μ formalism will also be highlighted.

3

2.2 Preliminaries

First, some notation regarding standard concepts for input/output systems. For more details, consult [18, 60] and references therein.

- ℓ_∞ denotes the set of all sequences $f = \{f_0, f_1, f_2, \ldots\}$, $f_k \in \mathbb{R}^N$, so that

$$\|f\|_{\ell_\infty} = \sup_k |f(k)|_\infty < \infty,$$

  where $|f(k)|_\infty$ is the standard ℓ_∞ norm on vectors. Also, $\ell_{\infty,e}$ denotes the extended space of all sequences in $\mathbb{R}^N$ and $\ell_{\infty,e} \backslash \ell_\infty$ denotes the set difference.

- $\ell_p, p \in [1, \infty)$, denotes the set of all sequences so that

$$\|f\|_{\ell_p} = \left(\sum_k |f(k)|_p^p \right)^{1/p} < \infty.$$

- c_0 denotes the subspace of ℓ_∞ of sequences converging to zero.

- S denotes the backward shift operator (unit time delay).

- P_k denotes the kth-truncation operator on $\ell_{\infty,e}$:

$$P_k \colon \{f_0, f_1, f_2, \ldots\} \longrightarrow \{f_0, \ldots, f_k, 0, \ldots\}$$

- A nonlinear operator $H : \ell_{\infty,e} \longrightarrow \ell_{\infty,e}$ is *causal* if

$$P_k H = P_k H P_k, \quad \forall k = 0, 1, 2, \ldots,$$

  strictly causal if

$$P_k H = P_k H P_{k-1}, \quad \forall k = 0, 1, 2, \ldots,$$

  time-invariant if it commutes with the shift operator ($HS = SH$), and ℓ_p *stable* if

$$\|H\| = \sup_k \sup_{\substack{f \in \ell_{p,e} \\ P_k f \neq 0}} \frac{\|P_k H f\|_{\ell_p}}{\|P_k f\|_{\ell_p}} < \infty.$$

  The quantity $\|H\|$ is called the *induced operator norm over* ℓ_p.

- $\mathcal{L}_{TV}$ denotes the set of all *linear* causal ℓ_∞-stable operators. This space is characterized by infinite block lower triangular matrices of the form

$$\begin{pmatrix} H_{00} & & 0 & \\ H_{10} & H_{11} & & \\ \vdots & \vdots & \ddots & \end{pmatrix}$$

where H_{ij} is a $p \times q$ matrix. This infinite matrix representation of H acts on elements of ℓ_∞^q by multiplication, i.e., if $u \in \ell_\infty^q$, then $y := Hu \in \ell_\infty^p$ where $y(k) = \sum_{j=0}^{k} H_{kj} u(j) \in \mathbb{R}^p$. The induced norm of such an operator is given by:

$$\|H\|_{\mathcal{L}_{TV}} = \sup_i |(H_{i1} \ldots H_{ii})|_1$$

where $|A|_1 = \max_i \sum_j |a_{ij}|$.

- $\mathcal{L}_{TI}$ denotes the set of all $H \in \mathcal{L}_{TV}$ which are *time-invariant*. It is well known that $\mathcal{L}_{TI}$ is isomorphic to ℓ_1 and the matrix representation of the operator has a Toeplitz structure. Every element in $\mathcal{L}_{TI}$ is associated with a λ-transform defined as

$$\hat{H}(\lambda) = \sum_{k=0}^{\infty} H(k) \lambda^k.$$

The collection of all such transforms is usually denoted by $\mathbf{A}$, which will be equipped with the same norm as the ℓ_1 norm.

Throughout this chapter, systems are thought of as operators. So the composition of two operators G, H is denoted as GH. If both are time-invariant then $GH \in \ell_1$ (or $\mathcal{L}_{TI}$), and the induced norm is denoted by $\|GH\|_1$. When the λ-transform is referred to specifically, we use the notation $\hat{H}$ for the transform of H. Also, all operator spaces are matrix-valued functions whose dimensions will be suppressed in general whenever understood from the context.

Let X be a normed linear space. The space of all bounded linear functionals on X is denoted X^*, equipped with the natural induced norm; X^* is always complete. It is convenient to put on X^* a weaker topology which makes $X^{**} = X$. This is the *weak*-topology*.

Dual of $\ell_p, 1 \le p < \infty$: The dual of ℓ_p is ℓ_q, where $\frac{1}{p} + \frac{1}{q} = 1$. The characterization is given by the following theorem.

5

Theorem 2.2.1 *Every bounded linear functional f on $\ell_p, 1 \leq p < \infty$, is representable uniquely in the form*

$$f(x) = \sum_{i=0}^{\infty} x_i y_i$$

where $y = (y_i)$ is an element in ℓ_q. Furthermore, every element of ℓ_q defines a member of ℓ_p^* in this way and

$$||f|| = ||y||_q$$

The above definitions are extended for vector-valued sequences and matrix-valued sequences in the obvious way.

In this chapter, we will give a solution to the ℓ_1 synthesis problem by using the theory of Lagrange multipliers. Many people are quite familiar with this theory for finite-dimensional optimization problems, and in the sequel, we will review the basic duality theorem for infinite-dimensional problems. For a more thorough treatment, see [43].

Let X be a vector space. A *convex cone* P is a convex set such that if $x \in P$ then $\alpha x \in P$ for all real $\alpha \geq 0$. Given such P, it is possible to define an ordering relation on X as follows: $x \geq y$ if and only if $x - y \in P$. Then it is natural to define a dual cone P^* (with an abuse of notation) inside X^* in the following way:

$$P^* = \{x^* \in X^* | < x, x^* > \geq 0 \ \forall x \in P\}.$$

This in turn defines an ordering relation on X^*.

Let f be a convex function from X to $\mathbb{R}$ and G a convex map from X to another normed space Z. Also, let Ω be a convex subset of X. Assume that there exists $x_1 \in X$ such that $G(x_1) < 0$ (the inequality with respect to some cone in Z). This is generally known as the *regularity assumption*. Define the minimization problem:

$$\mu_0 = \inf f(x) \quad \text{subject to} \quad x \in \Omega, \ G(x) \leq 0.$$

The Lagrange multiplier theory basically says that this constraint optimization problem can be transformed to an unconstrained problem over $x \in \Omega$. Precisely, there exists an element $z_0^* \geq 0$ in Z^* (with respect to the dual cone), so that

$$\mu_0 = \inf_{x \in \Omega} \{f(x) + < G(x), z_0^* >\}$$

The element z_0^* is precisely the Lagrange multiplier. Equivalently,

$$\mu_0 = \sup_{z^* \geq 0} \inf_{x \in \Omega} \{ f(x) + < G(x), z^* > \}.$$

In the case where the infimization problem contains equality constraints, we will replace them by two inequality constraints. Care should be taken in this case since the assumption that the constraint set has an interior point will be violated; however under mild assumptions, if the equality constraints are given in terms of linear operators, the result will still hold without the regularity conditions.

2.3 Why the ℓ_∞ signal norm?

In many real-world applications, output disturbance and/or noise is persistent, i.e., continues acting on the system as long as the system is in operation. This implies that such inputs have infinite energy, and thus one cannot model them as "bounded-energy signals." Nevertheless, one can get a good estimate on the maximum amplitude of such inputs. Examples where bounded disturbances arise in practical situations are abundant. Wind gusts facing an aircraft in flight can be viewed as bounded disturbances. Without a correcting control action, such disturbances will cause the aircraft to deviate from its set path. An automobile driven over an unpaved road experiences disturbances due to the irregularity of the course. Such disturbances, although persistent, are clearly bounded in magnitude. In process control, level measurements of a boiling liquid are corrupted by bounded disturbances due to the constant level fluctuations of the liquid. Because such disturbances are so frequent, a mathematical model describing them is essential. The ℓ_∞ norm is clearly the most natural choice for measuring the size of such disturbances. In general, we will assume that the disturbance is the output of a linear time-invariant (LTI) filter subjected to signals of magnitude less than or equal to one, i.e.,

$$d = Ww, \quad ||w||_\infty \leq 1.$$

Not only is the ℓ_∞ norm useful for measuring input signal size, but it can also be very useful as a measure for the size of output signals. For example, in many applications it is crucial that the tracking error never exceeds a certain level at *any* time. While this requirement cannot be captured by using the ℓ_2 norm, it can be stated explicitly as a condition on the ℓ_∞ norm of the error signal. Another situation when the ℓ_∞ norm is

useful is when the plant, or any other device in the control loop, has a maximum input rating which should not be exceeded. This translates directly to a requirement on the ℓ_∞ norm of that input. An example of such a requirement appears in the next section. In addition, the ℓ_∞ norm plays an important role in designing controllers for nonlinear systems. Since most of the nonlinear controller designs are based on linearization, the linear model gives a faithful representation of the system only if the states remain close to the equilibrium point, a requirement captured directly in terms of the ℓ_∞ norm.

2.4 The ℓ_1 Norm

While the ℓ_∞ norm is used as a measure of signal size, the ℓ_1 norm is used to measure a system's amplification of ℓ_∞ input signals. Let T be an LTI system given by

$$z(t) = (Tw)(t) = \sum_{k=0}^{t} T(k)w(t-k).$$

The inputs and outputs of the system are measured by their maximum amplitude over all time, otherwise known as the ℓ_∞ norm, i.e.,

$$\|w\|_\infty = \max_j \sup_k |w_j(k)|.$$

The ℓ_1 norm of the system T is precisely equal to the maximum amplification the system exerts on bounded inputs. This measure defined on the system T is known as the *induced operator norm* and is mathematically defined as

$$\|T\| = \sup_{\|w\|_\infty \leq 1} \|Tw\|_\infty = \|T\|_1,$$

where $\|T\|_1$ is the ℓ_1-norm of the pulse response and is given by

$$\|T\|_1 = \max_i \sum_j \sum_k |t_{ij}(k)|.$$

A system is said to be ℓ_∞- stable if it has a bounded ℓ_1 norm, and the space of all such systems will be denoted by ℓ_1. From this definition, it is clear that the system attenuates inputs if its ℓ_1 norm is strictly less than unity.

In the case where the inputs and outputs of the linear system are measured by the ℓ_2 norm, then the gain of the system is given by the H_∞ norm and is given by [20, 28, 57, 62]

$$\|\hat{T}\|_\infty = \sup_{0 \leq \theta \leq 2\pi} \sigma_{max}(\hat{T}(e^{i\theta})).$$

8

The two induced norms are related by [4]

$$\|\hat{T}\|_\infty \leq C_1 \|T\|_1 \leq C_2(N)\|\hat{T}\|_\infty,$$

where C_1 is a constant depending only on the dimension of the matrix T, and C_2 is a linear function of the McMillan degree N of $\hat{T}$. In other words, every system inside ℓ_1 is also inside H_∞, but the converse is not true. This means that there exist ℓ_2 stable LTI systems that are not ℓ_∞ stable; an example is the function with the λ-transform given by [4];

$$\hat{T}(\lambda) = e^{\frac{1}{1-\lambda}}.$$

Thus, for LTI systems, minimizing the ℓ_1 norm guarantees that the H_∞ norm is bounded. This means that this system will have good ℓ_2-disturbance rejection properties as well as ℓ_∞-disturbance rejection properties. Also, the ℓ_1 norm is more closely allied with BIBO stability notions and hence quite desirable to work with. The disadvantage in working with the ℓ_1 norm is the fact that it is a Banach space of operators operating on a Banach space, not a Hilbert space itself. Many of the standard tools are not usable; however, this chapter will present new techniques for handling problems of this kind.

2.5 Prototype Problems

In this section we demonstrate the advantages of using the ℓ_∞ signal norm by presenting a few prototype problems. For each problem, certain control objectives related to the ℓ_∞ norm are to be met. These problems demonstrate the advantages of using the ℓ_∞ signal norm as a means of capturing time-domain specifications in an uncertain environment. Later on, it will be shown how all such problems can be treated in a unified manner under a single framework. We shall then develop mathematical techniques for obtaining solutions for all problems which fit within that framework.

2.5.1 Disturbance Rejection Problem

Consider the system in Fig. 2.1. Here P_o is a plant and K a controller, both LTI.

The system is subjected to bounded disturbances which are reflected at the plant output. As mentioned earlier, these disturbances are assumed to be the output of a time-invariant filter W_1 reflecting the frequency content of such disturbances. The control objective in this case will be to find a controller K which satisfies the following:

9
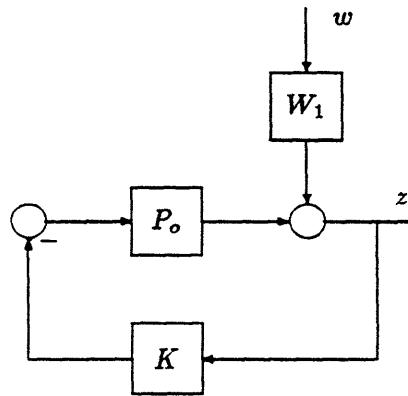
Figure 2.1: Disturbance Rejection Problem

1. K internally stabilizes the feedback system.

2. The effect of the disturbances at the plant output is minimized, i.e., K minimizes $\sup_{\|w\|_\infty \le 1} \|z\|_\infty$.

2.5.2 Command Following in the Presence of Input Saturation

The command following problem is equivalent to the disturbance rejection problem. Consider the system in Fig. 2.2. The plant, P, suffers from saturation nonlinearities at its
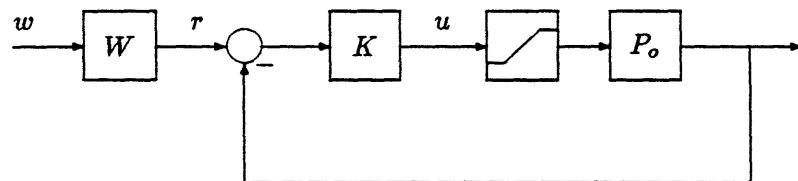
Figure 2.2: Command Following with Input Saturation

input. Therefore, it can be viewed as having two components: a saturation component, $Sat(.)$, and an LTI component, P_o. The saturation component is defined as follows:

$$Sat(u) = \begin{cases} u & |u| < U_{\max} \\ U_{\max} & |u| \geq U_{\max}. \end{cases}$$

As a result the plant is described as $P = P_o\, Sat(\cdot)$. Because of the presence of the saturation, the plant input, u, must not be allowed to exceed U_{max}. This requirement can be captured in a natural way using the ℓ_∞ norm of u. In other words, u must satisfy $\|u\|_\infty \leq U_{max}$.

The command, r, is to be followed at the plant output. It is not fixed but rather can be any command in the set

$$\{r = Ww \; : \; \|w\|_\infty \leq 1\},$$

where W reflects the frequency content of the desired commands and is typically a low pass filter.

The control objectives can now be stated more precisely. It is desired to find a controller K so that:

1. K internally stabilizes the system.

2. $\|u\|_\infty \leq U_{max}$.

3. y follows r uniformly in time to within a maximum error level of $\gamma > 0$, i.e.,
   $\|y - r\|_\infty \leq \gamma$.

2.5.3 Robust Disturbance Rejection

In the previous two problems, the plant was assumed to be known exactly. This is rarely the case due to unmodelled dynamics, parameter variations, etc. When the controller designed for a nominal plant model is implemented on the real system, there are no guarantees on the resulting performance of the system. Even requirements as basic as stability may not be met. The deviation from the expected behavior of the system clearly depends on the accuracy of the model. Since modelling uncertainty is inevitable, it is imperative to include stability and performance robustness to model uncertainty as a design objective. We now take a second look at the disturbance rejection problem
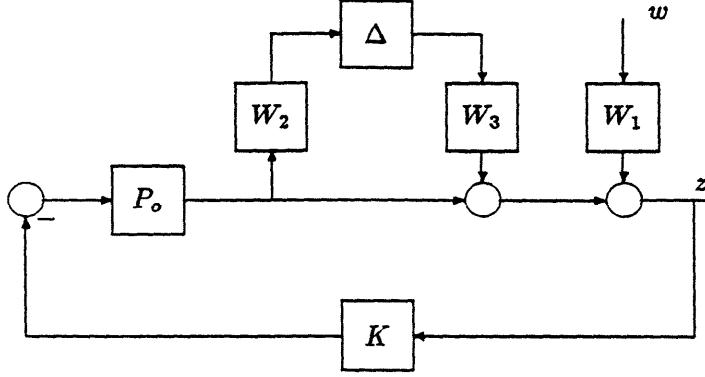
Figure 2.3: Robust Disturbance Rejection Problem

discussed earlier. Instead of considering a single nominal time-invariant plant, P_o, we shall instead consider a collection of plants. The class of plants considered is taken to be

$$\Pi := \left\{ P = P_o + W_3 \Delta W_2 : \Delta \text{ is causal and } \|\Delta\| := \frac{\|\Delta u\|_\infty}{\|u\|_\infty} \leq 1 \right\},$$

where W_1 and W_2 are time-invariant weighting functions. In this definition, the plant perturbation, Δ, may be time-varying and/or nonlinear. Any plant belonging to this plant class is said to be admissible. Note that when $\Delta = 0$, we recover the nominal LTI plant. Consequently, the collection of admissible plants, Π, may be viewed as a ball of plants centered around the nominal time-invariant plant model. If a system property, such as stability, holds for all admissible plants it is said to be *robust*. We now add to our original disturbance rejection problem a new objective: robustness. In other words, the controller K is now required to perform the following tasks:

1. K internally stabilizes all admissible plants, i.e., all plants in the class Π.

2. K minimizes the effect of the disturbance w on the magnitude of the output for the worst possible admissible plant, i.e., K minimizes $\sup\limits_{P \in \Pi} \sup\limits_{\|w\|_\infty \leq 1} \|y\|_\infty$.

2.5.4 Robustness in the Presence of Coprime Factor Perturbations

Another approach to the representation of plant uncertainty is through coprime factor perturbations [5, 26]. Let $P_o = N M^{-1}$ be a coprime factorization of the nominal plant.
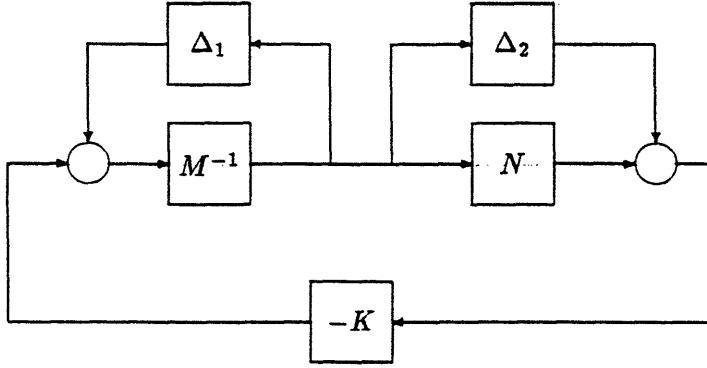
Figure 2.4: Coprime Factor Perturbations

The graph of the plant P_o over the space ℓ_∞ is define as the image of the space ℓ_∞ under the map G_{P_0} where

$$G_{P_o} : \ell_\infty \mapsto \ell_\infty \times \ell_\infty$$
$$G_{P_o} u = \begin{bmatrix} Mu \\ Nu \end{bmatrix}.$$

The class of admissible plants can be defined as those plants whose graph is perturbed in following way:

$$\Pi = \left\{ P : G_P = \begin{bmatrix} M + \Delta_1 \\ N + \Delta_2 \end{bmatrix}, \|\Delta_1\| \leq 1, \|\Delta_2\| \leq 1 \right\}.$$

This plant class can be viewed as that obtained by perturbing the plant numerator and the plant denominator independently as shown in Fig. 2.4. The main objective in this case is to find a controller K which stabilizes all plants in the class Π.

2.5.5 A Multiobjective Control Problem

In almost all practical control problems, more than one objective must be met simultaneously. Perhaps one of the most attractive features of the present approach is its ability to handle multiple objectives in a natural way. As an example of a multiple objective problem consider the system in Fig. 2.5. In the figure the plant is subjected to multiplicative output perturbations. In addition it has a saturation nonlinearity at its input of

13
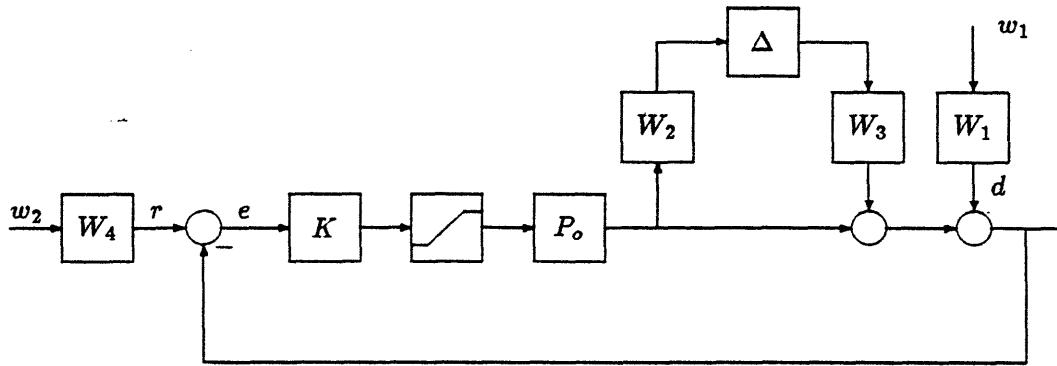
Figure 2.5: Multiobjective Problem

the type discussed earlier. A command input, r, is applied while a bounded disturbance, d, is acting at the plant output. The objectives in this problem are a combination of those objectives in the first three problems discussed earlier. Aside from stabilizing all admissible plants, the controller must also ensure that the plant input, u, never exceeds its maximum, U_{max}, despite the presence of the output disturbance, the command input, and the plant uncertainty. Furthermore, the tracking error in this unfriendly environment must be maintained at a minimum level for all time. These requirements on the controller are summarized as follows:

1. K stabilizes all plants in Π.

2. K is chosen so that $\displaystyle\sup_{\|w_i\|_\infty \leq 1} \sup_{P \in \Pi} \|u\|_\infty \leq U_{max}$.

3. K is chosen so that $\displaystyle\sup_{\|w_i\|_\infty \leq 1} \sup_{P \in \Pi} \|e\|_\infty$ is minimized.

Comment: It is possible in this formulation to include time-varying weights with which one can emphasize certain periods of the time response. The general framework and solutions presented in the sequel generalize in the presence of such weights; however, we will restrict our discussion to the time-invariant case.

14
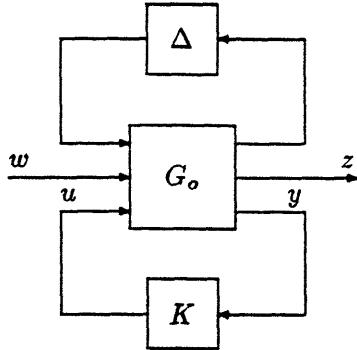
Figure 2.6: The Robust Performance Problem

2.6 A General Formulation: The Robust Performance Problem

In the previous section, we have formulated sample control problems which reflect various practical control requirements. Two assumptions were embedded in the problem statement. The first of these is that the command signals and the disturbance signals do not necessarily decay in time but can instead persist over all time so long as they are bounded. This is a fairly realistic assumption and leads to the adoption of the ℓ_∞ norm to measure the signal size. The second consists of requiring the regulated signals of interest to have small *maximum* amplitudes. Thus, once more, the ℓ_∞ signal norm is used as a measure for signal size, but this time it is the regulated output signals which are being measured. When considering that quite often the output of interest is a tracking error, plant input and/or plant output, it becomes clear that limiting the maximum value these signals can achieve is desirable if not necessary. As a means for obtaining a unifying framework for formulating and solving a wide variety of problems with ℓ_∞ signal norms and ℓ_∞ induced-norm bounded perturbations, we set up the Robust Performance Problem. All the prototype problems discussed in the previous section are special cases of this general problem. So consider the system in Fig. 2.6: Δ models the system uncertainty, K is the controller, and G_o contains the remaining part of the system. It is assumed that Δ belongs to the following class:

$$\mathcal{D}(n) := \{\Delta = diag(\Delta_1, \ldots, \Delta_n) \ : \ \Delta_i \text{ is causal and } \|\Delta_i\| \leq 1\}.$$

15

Here $\|\Delta_i\|$ is the induced ℓ_∞ norm, i.e., $\|\Delta_i\| = \sup_{u \neq 0} \frac{\|\Delta_i u\|_\infty}{\|u\|_\infty}$. In the sequel, the Δ_i's are assumed to be SISO for simplicity. There is no time-invariance restriction on the perturbations, and hence time-varying and/or nonlinear perturbations are allowed. The diagonal structure of the perturbations is essential for incorporating information about the *location* of the system uncertainty. For example actuator unmodelled dynamics are not related to sensor unmodelled dynamics or to the plant's unmodelled high-frequency dynamics, and should not be modeled by the same perturbation block. By isolating the independent sources of uncertainty, a more realistic and less conservative system model is obtained. This is the main reason for considering structured perturbations. While Δ models the uncertain part of the system, G_o is the known part of the system with the exception of the controller, and it is a 3×3 block matrix. The actual system is an element in the upper linear fractional connection of G_o and the admissible Δ's. So included in G_o is the nominal plant/plants, any input and output weighting functions, and any weighting functions on the perturbations. We shall restrict the weights and the nominal plant to be LTI. As a result, G_o is LTI. The signal w denotes all exogenous inputs, including the command inputs and the disturbance inputs which are assumed to be in ℓ_∞, while z denotes the regulated outputs. Both w and z are allowed to be vector signals. From now on, we shall refer to the map taking w to z as $\mathcal{T}_{zw}$. The induced ℓ_∞ norm of $\mathcal{T}_{zw}$ is defined as follows:

$$\|\mathcal{T}_{zw}\| := \sup_{w \neq 0} \frac{\|\mathcal{T}_{zw} w\|_\infty}{\|w\|_\infty} = \sup_{w \neq 0} \frac{\|z\|_\infty}{\|w\|_\infty}.$$

Finally, the controller K is assumed to be LTI. We are now ready to state the Robust Performance Problem.

Robust Performance Problem: *Find a controller K so that*

1. *The system achieves robust stability, i.e., K internally stabilizes the system for all admissible perturbations, i.e., for all Δ in $\mathcal{D}(n)$.*

2. *The system achieves robust performance, i.e., K is chosen so that*

$$\sup_{\Delta \in \mathcal{D}(n)} \|\mathcal{T}_{zw}\| < 1.$$

As mentioned earlier, the prototype problems discussed can all fit in this framework. As an example, for the Disturbance Rejection Problem since the number of perturbation

16

blocks, n, is zero, G_o has only two inputs w and u, and two outputs z and y. As a result G_o has the form:

$$G_o = \begin{pmatrix} W_1 & -P_o \\ W_1 & -P_o \end{pmatrix}$$

This is referred to as the nominal performance problem.

On the other hand, for the robust disturbance rejection problem n will be 1. Thus, G_o has an additional input fed from the perturbation output, and an additional output feeding the perturbation input. It follows that G_o has the following structure:

$$G_o = \begin{pmatrix} 0 & 0 & -W_2 P_o \\ W_3 & W_1 & -P_o \\ W_3 & W_1 & -P_o \end{pmatrix}.$$

And so on.

2.7 Robustness Conditions

Having stated the Robust Performance Problem, we can now focus our attention on its solution. In particular, we shall develop necessary and sufficient conditions for achieving both performance robustness and stability robustness. These conditions will be used for the robustness *analysis* of the system at hand. In this case, the controller is assumed given and fixed and its effect on the robustness of the system is investigated. The same conditions developed for robustness analysis are used to develop techniques for the synthesis of robust controllers.

We begin by discussing the robustness analysis issue. Suppose we are given a nominal system G_o, a perturbation class $\mathcal{D}(n)$, and a controller K connected as shown in Fig. 2.6. We can incorporate G_o and K together and view them as one system, M, as shown in Fig. 2.7. Thus M will have two inputs and two outputs. We will assume that the controller K stabilizes the nominal system G_o; otherwise robust stability and hence performance clearly will not be achieved. Consequently, M will be LTI and stable. We will say the system in Fig. 2.7 achieves *robust stability* if it is stable for all admissible perturbations. We will say that it achieves *robust performance* if, in addition, $\|T_{zw}\| < 1$ for all admissible perturbations. We can now state the following problem whose solution is provided in the next two sections:

Robustness Analysis Problem   *Under what conditions on M will the system in Fig. 7 achieve robust performance?*

17
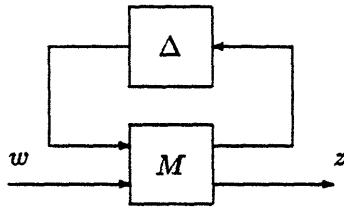
Figure 2.7: Robust Performance Problem in M-Δ form

2.7.1 Stability Robustness vs. Performance Robustness

It is an interesting fact that a robust performance problem can be transformed to a robust stability problem. This has been shown in [22, 21, 47] when the perturbations are LTI with an L_2 induced-norm. This remains true in our case as well, although the method of proof is quite different. To elaborate further on this relationship between stability robustness and performance robustness consider the two systems shown in Fig. 2.8. System I corresponds to a performance robustness problem, while System II is formed
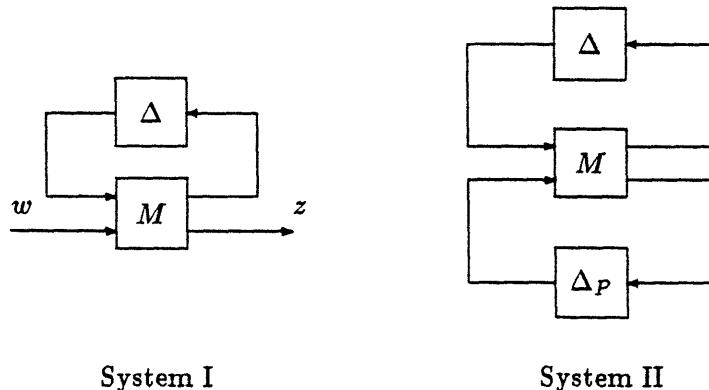
System I   System II

Figure 2.8: Stability Robustness vs. Performance Robustness

18

from System I by feeding z back to w through a fictitious perturbation, Δ_P, satisfying $\|\Delta_P\| \leq 1$. As a result, System II has $\mathcal{D}(n+1)$ as its perturbation class. We can now ask the following question: How does the performance robustness of System I relate to the stability robustness of System II ? One aspect of the relationship between the two notions of stability is fairly obvious: performance robustness of System I implies stability robustness of System II. This is quite easy to see. Since robust performance is equivalent to the norm of the map between w and z being less than one, the Small Gain Theorem can be used to establish the stability of System II for all $\|\Delta_P\| \leq 1$, or equivalently to establish the robust stability of System II. Equally important, the relation between stability robustness and performance robustness holds the other way as well. In other words, stability robustness of System II implies performance robustness of System I. This direction is not as obvious as the first one. The proof follows from certain results on the robustness of time-varying systems.

2.7.2 Stability Robustness Conditions

Because performance robustness is equivalent to stability robustness in the sense discussed earlier, we need only discuss stability robustness. Specifically, we can consider the interconnection of a stable LTI system, M, with a structured perturbation $\Delta \in \mathcal{D}(n)$ in Fig. 2.9, and seek necessary and sufficient conditions for the stability robustness of the system. Since M and Δ are both stable, the internal stability of the system is equivalent
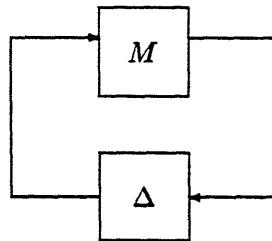
Figure 2.9: Stability robustness problem

to the map $I - M\Delta$ having a stable inverse, one which maps ℓ_∞ to itself with a finite gain. When the signal norm is the ℓ_2 norm and the perturbations are time-invariant, the conditions are provided by the Structured Singular Value, the function μ [22]. In

19

particular, robustness is achieved iff $\sup\limits_{0\leq\theta\leq 2\pi}\ \mu(M(e^{j\theta})) < 1$. In our formulation, it turns out that the conditions are much simpler and easier to compute than μ. Before we can present these conditions we need to define a certain nonnegative matrix, $\widehat{M}$, which depends solely on M. Recalling M has n inputs and n outputs, it can be partitioned as follows:

$$M = \begin{bmatrix} M_{11} & \cdots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \cdots & M_{nn} \end{bmatrix}.$$

Each M_{ij} is LTI and stable, and thus $M_{ij} \in \ell_1$. Clearly $\|M_{ij}\|_1$ can be computed with arbitrary accuracy by considering finite truncations of M_{ij} as approximations. We can now define $\widehat{M}$ as follows:

$$\widehat{M} = \begin{bmatrix} \|M_{11}\|_1 & \cdots & \|M_{1n}\|_1 \\ \vdots & & \vdots \\ \|M_{n1}\|_1 & \cdots & \|M_{nn}\|_1 \end{bmatrix}.$$

One of the most interesting aspects of the robustness problem formulated here is the role which $\widehat{M}$ plays in the system robustness. This is presented in the next theorem due to Khammash and Pearson [33, 34, 35].

Theorem 2.7.1 *The system in Fig. 2.9 possesses robust stability iff any one of the following holds:*

1. $\rho(\widehat{M}) < 1$, *where $\rho(.)$ denotes the spectral radius.*

2. $x \leq \widehat{M}x$ *and* $x \geq 0$ *imply that* $x = 0$, *where the vector inequalities are to be interpreted componentwise.*

3. $\inf\limits_{R\in\mathcal{R}} \|R^{-1}MR\|_1 < 1$, *where* $\mathcal{R} := \{diag(r_1,\ldots,r_n) : r_i > 0\}$.

One of the main contributions of this theorem is that it provides simple and exact conditions for testing the system's stability robustness regardless of the number of perturbation blocks, n. While the three conditions in the theorem are equivalent, each provides a different perspective and has certain advantages over the others. For example, the spectral radius condition is in general the easiest to compute. It is particularly useful when n is large since it can be computed efficiently using power methods. Specifically, given an $\widehat{M}$ which is assumed primitive (i.e., $\widehat{M}^k > 0$ for some integer k), then it satisfies the following:

$$\min_i \frac{(\widehat{M}^{k+1}x)_i}{(\widehat{M}^k x)_i} \leq \rho(\widehat{M}) \leq \max_i \frac{(\widehat{M}^{k+1}x)_i}{(\widehat{M}^k x)_i}$$

for any vector $x > 0$. Furthermore, the upper and lower bounds both converge to $\rho(\widehat{M})$ as k goes to infinity. If $\widehat{M}$ were not primitive, it can be perturbed slightly to become primitive.

Whereas the spectral radius test provides a yes or no answer concerning system robustness, the second test involving the Linear Matrix Inequality (LMI) is most useful for providing information about the effect of the individual entries of $\widehat{M}$ on the overall robustness of the system. This is achieved by translating the LMI condition into n algebraic conditions stated explicitly in terms of the entries of $\widehat{M}$. This is best demonstrated by an example. Suppose $\widehat{M}$ is a 2×2 matrix corresponding to a certain robustness problem with $n = 2$. The LMI condition states that robust stability iff the system

$$x_1 \leq \|M_{11}\|_1 x_1 + \|M_{12}\|_1 x_2$$
$$x_2 \leq \|M_{21}\|_1 x_1 + \|M_{22}\|_1 x_2$$

has no solution $x = (x_1, x_2) \in [0, \infty) \times [0, \infty) \backslash \{0\}$. Among other things, this implies that $\|M_{11}\|_1 < 1$; otherwise $x = (1, 0)$ would be a solution for the two inequalities. The first inequality can be rewritten as

$$x_1 \leq \frac{\|M_{12}\|_1}{1 - \|M_{11}\|_1} x_2.$$

When combined with the second inequality, we have that

$$x_2 \leq \left(\|M_{21}\|_1 \frac{\|M_{12}\|_1}{1 - \|M_{11}\|_1} + \|M_{22}\|_1 \right) x_2$$

has no solution in $(0, \infty)$, which is equivalent to

$$\|M_{21}\|_1 \frac{\|M_{12}\|_1}{1 - \|M_{11}\|_1} + \|M_{22}\|_1 < 1.$$

This last condition, together with the condition that $\|M_{11}\|_1 < 1$, is therefore necessary for the inequality robustness conditions to hold. Be retracing our steps backwards, it becomes clear that they are also sufficient. This procedure of constructing explicit norm conditions from the second robustness condition can be repeated in the same way for any n.

Finally, the third robustness condition is useful for robust controller synthesis. This will be discussed in more detail later on.

Equivalence of the Robustness Conditions in Theorem 2

Before we shed more light on the proof of Theorem 2, we will show that the three, apparently unrelated, conditions in the statement of the theorem are indeed equivalent. We will show that $1 \Leftrightarrow 2$ and that $1 \Leftrightarrow 3$. For simplicity, we will do this for an irreducible $\widehat{M}$. So suppose that $\rho(\widehat{M}) < 1$. It follows that $(I - \widehat{M})^{-1}$ exists. Since $(I - \widehat{M})^{-1} = I + \widehat{M} + \widehat{M}^2 + \ldots$, all of its entries will be positive. Now if $x \geq 0$ is such that $x \leq \widehat{M}x$, or equivalently, $(I - \widehat{M})x \leq 0$, then multiplying both sides by $(I - \widehat{M})^{-1}$ implies that $x \leq 0$. Thus x must be zero. This is what 2 states. To show that 2 implies 1, suppose 1 does not hold, i.e., that $\rho(\widehat{M}) \geq 1$. The Perron-Frobenius theory for nonnegative matrices states that $\rho(\widehat{M})$ is itself an eigenvalue of $\widehat{M}$. Moreover, associated with $\rho(\widehat{M})$ we can find an eigenvector $x' > 0$. This implies that $\rho(\widehat{M})x' = \widehat{M}x'$, which in turns implies that 2 does not hold. Thus, we have demonstrated that $1 \Leftrightarrow 2$.

We now show $1 \Leftrightarrow 3$ by showing that $\rho(\widehat{M}) = \inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1$. By definition,

$$\|R^{-1}MR\|_1 = \max_i \sum_{j=1}^{n} \|(R^{-1}MR)_{ij}\|_1 = \max_i \sum_{j=1}^{n} \frac{r_j}{r_i} \|M_{ij}\|_1.$$

The expression on the right is also equal to the induced norm of the matrix $R^{-1}\widehat{M}R$ as a map from $(\mathbb{R}^n, \|.\|_\infty)$ to itself. Referring to this norm by $|.|_1$, we therefore have $\|R^{-1}MR\|_1 = |R^{-1}\widehat{M}R|_1$. Since any matrix norm bounds from above the spectral radius of that matrix we have:

$$\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 = \inf_{R \in \mathcal{R}} |R^{-1}\widehat{M}R|_1 \geq \rho(R^{-1}\widehat{M}R) = \rho(\widehat{M}).$$

But if we choose $R = diag(r'_1, \ldots, r'_n)$, where $(r'_1, \ldots, r'_n)^t$ is the positive eigenvector corresponding to the eigenvalue $\rho(\widehat{M})$, the inequality becomes an equality and the equivalence between 1 and 3 is established. It is interesting to note that for the optimum scalings $R = diag(r'_1, \ldots, r'_n)$, all the rows of $R^{-1}MR$ have the same norm. As will be demonstrated shortly, this fact is used to show why condition 3 in the above theorem is necessary for system robustness.

Proof of Necessity and Sufficiency

When $n = 1$, the spectral radius condition in the theorem above reduces to the condition $\|M\|_1 < 1$. A simple application of the Small Gain Theorem shows that this condition is sufficient for stability. Necessity has been shown by Dahleh and Ohta [9]. For n larger
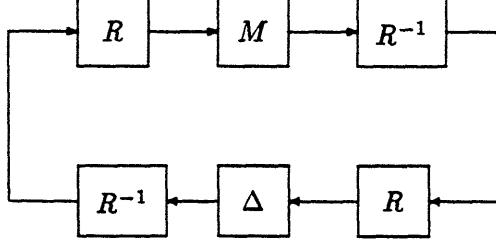
Figure 2.10: Scaled System

than one, we now show that $\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 < 1$ is sufficient for robust stability. We do this with the aid of Fig. 2.10 obtained via the addition of scalings R and R^{-1}, where $R \in \mathcal{R}$. Clearly, the robustness of this system and that in Fig. 9 are equivalent in the sense that if one is robustly stable, then so is the other one. Moreover, for the system in Fig. 2.10, $R\Delta R^{-1}$ belongs to $\mathcal{D}(n)$ whenever Δ belongs to $\mathcal{D}(n)$, and thus $\|R\Delta R^{-1}\| < 1$. That being the case, the Small Gain Theorem can be invoked to conclude that $\|R^{-1}MR\|_1 < 1$ is sufficient for robust stability. This holds for any $R \in \mathcal{R}$. The least conservative sufficient condition obtainable in that manner is

$$\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 < 1.$$

We now demonstrate that $\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 < 1$ is necessary for robust stability. For simplicity, we do this for the case $n = 2$. The approach will be to show how one can construct a destabilizing perturbation $\Delta \in \mathcal{D}(2)$ whenever $\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 \geq 1$. So suppose that $\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_1 \geq 1$. We have previously shown that this infimum is in fact a minimum, and it is achieved by an optimum scaling, R, obtained from the eigenvector corresponding to $\rho(\widehat{M})$. It was also shown that the two rows of $R^{-1}MR$ will have equal norms. This can be expressed as follows:

$$\|(R^{-1}MR)_1\|_1 = \|(R^{-1}MR)_2\|_1 = \|R^{-1}MR\|_1 \geq 1.$$

where $(R^{-1}MR)_i$ denotes the ith row of $R^{-1}MR$. The system $R^{-1}MR$ appears in Fig. 2.11 and has as its input $\xi = (\xi_1, \xi_2)$ and output $z = (z_1, z_2)$. In the figure, $y = (y_1, y_2)$ consists of the output $z = (z_1, z_2)$ after a bounded signal, the output of a sign function, has been added to it. This bounded signal will be interpreted as an
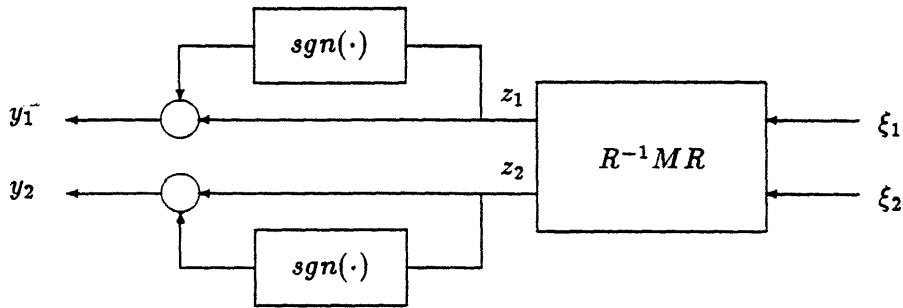
Figure 2.11: Scaled System with Constructed Input ξ

external signal injected for stability analysis. The strategy taken will be to construct ξ satisfying the two requirements:

1. ξ is unbounded.

2. ξ results in a signal y which satisfies $\|P_k\xi_i\|_1 \leq \|P_k y_i\|_1$ for $i = 1, 2$, where P_k is the truncation operator which acts on sequences by preserving the first $k+1$ terms and setting the rest to zero.

The first requirement on ξ guarantees that if an admissible perturbation Δ were to map y to ξ, it would be a destabilizing one because the bounded external signal would have produced an unbounded internal signal ξ. The second requirement, guarantees that such an admissible perturbation exists. In other words, if ξ and y satisfy the second condition, then it is possible to find Δ_i, for $i = 1, 2$, so that Δ_i is causal, has induced norm less than or equal to one, and satisfies $\Delta_i y_i = \xi_i$. If the first requirement is also met, this Δ will be a destabilizing perturbation.

For simplicity we shall assume that all M_{ij}'s have finite impulse response of length, say N. The construction of ξ proceeds as follows. While maintaining $|\xi_i(k)| \leq 1$ for $k = 0, \ldots, N-1$, the first N components of ξ can be constructed so as to achieve $\|(R^{-1}MR)_1\|_1$. Since $\|(R^{-1}MR)_1\|_1 \geq 1$, this implies that $\|P_{N-1}z\|_\infty \geq 1$, which in turn implies that $\|P_{N-1}y\|_\infty \geq 2$. Next, while still maintaining $|\xi_i(k)| \leq 1$, we pick the next N components of ξ so as to achieve the second row norm, $\|(R^{-1}MR)_2\|_1$. As a result we have $\|P_{2N-1}z_i\|_\infty \geq 1$ which implies that $\|P_{2N-1}y_i\|_\infty \geq 2$. Note that the second requirement on ξ has been met for $k = 0, \ldots, 2N-1$. In addition, because of the

way the first $2N$ terms of ξ have been constructed, we have

$$\|P_{2N-1}y_i\|_\infty \geq \|P_{2N-1}\xi_i\|_\infty + 1 \qquad i = 1, 2.$$

This allows us to relax the restriction on $|\xi_i(k)|$ for $k > 2N - 1$ without violating the second requirement on ξ. Specifically, we now allow $|\xi_i(k)|$ to be as large as 2 for $k = 2N, \ldots, 4N - 1$. In the same way as before we can pick $\xi(k)$ for this range of k so that we satisfy

$$\|P_{4N-1}y_i\|_\infty \geq \|P_{4N-1}\xi_i\|_\infty + 1 \qquad i = 1, 2.$$

which allows us to increase $|\xi_i(k)|$ by 1 for the next $2N$ components of ξ, and repeat the whole procedure again. From this construction, it is clear that when ξ is completely specified it will be unbounded and hence meets the first requirement. The second requirement is also met since all along $\xi_i(k)$ was chosen carefully so as not to become too large too soon.

It should be mentioned that the destabilizing perturbation can be taken to be linear time-varying (LTV), or it can instead be nonlinear time-invariant. So the spectral radius condition for robustness is also necessary and sufficient whenever the class of perturbation is restricted to include norm-bounded nonlinear time-invariant perturbations.

Construction of the Destabilizing Perturbation

In the previous section, we have claimed that given $\xi = \{\xi(i)\}_{i=0}^\infty \in \ell_{\infty,e}$ and $y = \{y(i)\}_{i=0}^\infty \in \ell_{\infty,e}$ so that $\|P_k\xi_i\|_\infty \leq \|P_{k-1}y_i\|_\infty \forall k$, and for $i = 1, 2$, then there exists $\Delta = diag(\Delta_1, \Delta_2)$ such that $\Delta y = \xi$ and $\|\Delta_i\| \leq 1$. Such a Δ was shown to be a destabilizing perturbation. In this section, we prove this claim by explicitly constructing the perturbation Δ. It turns out that Δ_i can be either LTV or nonlinear and time-invariant. We shall construct Δ_1 to be of the former type, while Δ_2 will be of the latter type.

So suppose we are given $\xi_1 = \{\xi_1(i)\}_{i=0}^\infty \in \ell_{\infty,e}$ and $y_1 = \{y_1(i)\}_{i=0}^\infty \in \ell_{\infty,e}$ such that $\|P_k\xi_1\|_\infty \leq \|P_k y_1\|_\infty \forall k$. The construction of Δ_1 is trivial if $y_1 = 0$: just pick Δ_1 itself to be zero. So assume $y_1 \neq 0$. We start the construction of Δ_1 by identifying a subsequence of y_1, say $(y_1(i_1), y_1(i_2), \ldots)$ which, depending on y_1, may or may not be finite. This subsequence may be defined recursively in the following manner: Let i_1 be the smallest integer such that $y_1(i_1) \neq 0$. Given $y_1(i_n)$, let i_{n+1} be the smallest integer

greater that i_n such that $|y_1(i_{n+1})| \geq |y_1(i_n)|$. Using the $\xi_1(i)$'s and $y_1(i_j)$'s we are now ready to construct Δ_1 through specifying its matrix kernel representation as follows:

$$
\Delta_1 = \begin{pmatrix}
\ddots & & & & & & & & \\
 & \frac{\xi_1(i_1)}{y_1(i_1)} & & & & & & & \\
 & \vdots & 0 & & & & & & \\
 & \vdots & \vdots & \ddots & & & & & \\
 & \frac{\xi_1(i_2-1)}{y_1(i_1)} & 0 & \cdots & 0 & & & & \\
 & & & & & \frac{\xi_1(i_2)}{y_1(i_2)} & & & \\
 & & & & & \vdots & 0 & & \\
 & & & & & \vdots & \vdots & \ddots & \\
 & 0 & & & & \frac{\xi_1(i_3-1)}{y_1(i_2)} & 0 & \cdots & 0 \\
 & & & & & & & & \frac{\xi_1(i_3)}{y_1(i_3)} \\
 & & & & & & & & \vdots & \ddots
\end{pmatrix}
$$

Notice that each row of the above matrix has at most one nonzero element, which, by the choice of the $y_1(i_j)$'s, will have its absolute value less than or equal to one. This implies $\|\Delta_1\| \leq 1$. Moreover, Δ_1 is clearly causal and it can easily be checked that $\Delta_1 y_1 = \xi_1$, which is what we wanted to show.

We now construct a nonlinear, time-invariant, and causal perturbation Δ_2. As before, Δ_2 must be so that $\|\Delta_2\| \leq 1$ and $\Delta_2 y_2 = \xi_2$. Let Δ_2 be defined as follows:

$$
(\Delta_2 f)(k) = \begin{cases} y(k-i) & \text{if for some integer } i \geq 0, \ P_k f = P_k S_i \xi_2 \\ 0 & \text{otherwise.} \end{cases}
$$

Note that Δ_2 maps y_2 to ξ_2 and $\|\Delta_2\| \leq 1$.

2.7.3 Comparisons

It is worthwhile comparing the class of perturbations that have gain less than unity over ℓ_2 (which arise in the standard μ) with the class of perturbations that have gain less than unity over ℓ_∞. If the perturbations are restricted to time-invariant ones, the ℓ_∞-stable perturbations with gain less than unity lie inside the unit ball of ℓ_2-stable perturbations (for the multivariable case, the unit ball will be scaled by a constant). This follows directly from the norm inequality between ℓ_1 and H_∞. If the perturbations are allowed to be time-varying, then the two sets are not comparable. Earlier, an example was presented that shows that the H_∞ ball is larger than the ℓ_1 ball. On the other hand, the

Perturbation class	$\mu(M) < 1$	$\displaystyle\inf_{R \in \mathcal{R}} \|R^{-1}MR\|_{H\infty} < 1$	$\rho(\widehat{M}) < 1$
NLTV, bounded ℓ_2-gain	nec	suff	suff
NLTV, bounded ℓ_∞-gain	nec	nec	nec and suff
NLTI, bounded ℓ_2-gain	nec	suff	suff
NLTI, bounded ℓ_∞-gain	nec	nec	nec and suff
LTV, bounded ℓ_2-induced norm	nec	suff	suff
LTV, bounded ℓ_∞-induced norm	nec	nec	nec and suff
LTI, bounded ℓ_2-induced norm	nec and suff	suff	suff
LTI, bounded ℓ_∞ induced norm	nec and suff	suff	suff

Table 2.1: Comparisons between different robustness criteria

operator Δ defined by

$$(\Delta f)(k) = f(0)$$

is ℓ_∞ stable but not ℓ_2 stable.

A question which might arise is, how do the derived robustness conditions differ from the Structured Singular Value? The answer lies in the class of perturbations assumed. While the perturbations here may be nonlinear time-varying (NLTV), nonlinear time-invariant (NLTI), or LTV for the conditions to be necessary and sufficient, μ theory gives necessary and sufficient conditions only for LTI perturbations. In terms of computation, the robustness test proposed here is much easier to compute and gives exact answers for any number of perturbation blocks, n. On the other hand, μ is much harder to compute especially since for $n > 3$ only an upper bound can be computed. One can use the small gain theorem to get sufficient conditions for robust stability in the presence of NLTV ℓ^2 induced norm-bounded perturbations in the same way it was done for the A norm. In this case, a sufficient condition would be $\inf_{R \in \mathcal{R}} \|R^{-1}\hat{M}R\|_{H_\infty} < 1$. It is not known whether this condition is also necessary. However, it is not sufficient to guarantee robustness when perturbations of the type considered in this chapter are present, i.e., for ℓ_∞ induced norm-bounded perturbations. In contrast, robustness in the presence of ℓ_∞ induced-norm bounded perturbations does imply robustness to ℓ_2 induced-norm bounded perturbations. The relationship between the various robustness conditions is summarized in Table 1 (In the table: nec, suff respectively mean necessary and sufficient).

In terms of robust controller synthesis, the controller must be chosen so that $\rho(\widehat{M})$ is minimized. The dependence of M on the controller is reflected through the Youla

parameter, Q, since M can be expressed as [28, 57, 61]

$$M = M(Q) = T_1 - T_2 Q T_3,$$

where the T_i's depend only on G_o. Because $\rho(\widehat{M}) = \inf_{R \in \mathcal{R}} \|R^{-1} M R\|_1$, the robustness synthesis problem becomes one of finding

$$\inf_{Q \text{ stable}} \inf_{R \in \mathcal{R}} \|R^{-1} M(Q) R\|_1.$$

With Q stable and fixed, we have seen that picking the eigenvector associated with $\rho(\widehat{M(Q)})$ will yield the minimum value over all scalings in $\mathcal{R}$. When R is fixed, we have an ℓ_1-norm minimization problem. This problem and its solution will be discussed in the remaining part of the chapter. So the approach which will be taken to solving the robustness synthesis problem is to start with an initial $R \in \mathcal{R}$. For that R we find the optimal Q resulting from the norm minimization problem. We then fix that Q and solve for the optimal R associated with this new Q and so on. Since at each step the objective function gets smaller and smaller, and since it is bounded from below by zero it is guaranteed to converge to some value. Unfortunately, this value may not be the global minimum. If at that point, a satisfactory level of performance robustness has been reached, we can stop and use the final Q to construct the controller. Otherwise, the iteration process should be restarted with a different initial scaling matrix in $\mathcal{R}$. This scheme is similar to the so-called $D - K$ iteration used in the μ-synthesis technique [22, 21]. The main difference is that while the scales used for μ-synthesis are frequency dependent and a convex optimization problem must be solved at each frequency, the scalings here are not frequency dependent and can be readily found by computing the eigenvector associated with $\rho(\widehat{M})$. Such a computation can be done very effectively using power methods, and no optimization problem need be solved to find the optimal scalings.

2.8 Synthesis of the ℓ_1 controller

As stated earlier, the ℓ_1 minimization problem is given by:

$$\mu_0 = \inf_{Q \text{ stable}} \|T_1 - T_2 Q T_3\|_1 \qquad (OPT)$$

In this section, we will show that this problem is equivalent to a linear programming problem in an infinite-dimensional space. By utilizing the duality theory of Lagrange multipliers, it is shown that in some cases the linear programs are in fact finite-dimensional

and thus exact solutions for (OPT) can be obtained. For the rest of the cases, the duality theory provides upper and lower approximations of the optimal solution. The use of the Lagrange multiplier theory highlights the strong resemblance between the ℓ_1 problem and standard linear programming problems.

The admissible subspace $\mathcal{S}$ is defined as:

$$\mathcal{S} = \{R \in \ell_1^{m \times n} | R = T_2 Q T_3, Q \text{ is stable}\}$$

The ℓ_1 problem can be interpreted as a *distance problem*: Find an element in the subspace $\mathcal{S}$ which is closest to the fixed element T_1, where distance is measured in the ℓ_1-norm. Previous work [10, 11] used the duality theory for distance problems to arrive at a solution for (OPT). Here we take an alternate approach using Lagrange multiplier theory, which is in fact more intuitive and transparent, to arrive at similar conclusions.

2.8.1 Characterization of the Subspace S

In the discussion below, it is assumed that $\hat{T}_2$ has full column rank $=n_2$, and $\hat{T}_3$ has full row rank $=n_3$. It is evident that this captures the most general situation since if either of these conditions does not hold, we can perform inner-outer factorizations on $\hat{T}_2$ and $\hat{T}_3$ and absorb the extra degree of freedom in $\hat{Q}$. Also, it is assumed that there exist n_2 rows of $\hat{T}_2$ and n_3 columns of $\hat{T}_3$ which are linearly independent for all λ on the unit circle. This assumption simplifies the exposition although it is not necessary. In general, it is enough to assume the above for 1 point on the unit circle [55]. Under this assumption, $\hat{T}_2$ and $\hat{T}_3$ can be written in the following form without loss of generality (possibly requiring the interchange of inputs and/or outputs):

$$\hat{T}_2 = \begin{pmatrix} \hat{T}_{21} \\ \hat{T}_{22} \end{pmatrix}$$

$$\hat{T}_3 = \begin{pmatrix} \hat{T}_{31} & \hat{T}_{32} \end{pmatrix}$$

where $\hat{T}_{21}$ has dimensions $n_2 \times n_2$ and is invertible and $\hat{T}_{31}$ has dimensions $n_3 \times n_3$ and is invertible. Moreover, $\hat{T}_{21}$ and $\hat{T}_{31}$ have no transmission zeros on the unit circle. Thus $\hat{R} = \hat{T}_2 \hat{Q} \hat{T}_3$ can be written:

$$\hat{R} = \begin{pmatrix} \hat{T}_{21} \\ \hat{T}_{22} \end{pmatrix} \hat{Q} \begin{pmatrix} \hat{T}_{31} & \hat{T}_{32} \end{pmatrix} = \begin{pmatrix} \hat{R}_{11} & \hat{R}_{12} \\ \hat{R}_{21} & \hat{R}_{22} \end{pmatrix}$$

The objective is to obtain a characterization of the feasible set $\mathcal{S}$. Notice that $\hat{Q}$ can be uniquely determined from the equality $\hat{R}_{11} = \hat{T}_{21} \hat{Q} \hat{T}_{31}$. As was shown in [10, 44], the

choice of $\hat{R}_{11}$ is constrained by the zeros of $\hat{T}_{21}$, $\hat{T}_{31}$ that are inside the unit disc. There is only a finite number of such zeros, and each zero is interpreted as a bounded linear functional on R_{11}. In the sequel, we use the following terminology:

Definition 2.8.1 *A transfer function $\hat{G}$ interpolates $\hat{T}_{21}$, $\hat{T}_{31}$ if $\hat{T}_{21}^{-1}\hat{G}\hat{T}_{31}^{-1}$ is stable.*

The motivation for this terminology stems from the fact that for $\hat{T}_{21}^{-1}\hat{G}\hat{T}_{31}^{-1}$ to be stable, $\hat{G}$ must have zeros at the same locations and directions as the zeros of $\hat{T}_{21}$ and $\hat{T}_{31}$. Each zero is in fact a bounded linear functional that annihilates the element G, and thus has a representation inside the dual space of ℓ_1, with the appropriate dimension. If these functionals are inside c_0, then we can view G as the annihilator in the dual of c_0. For example, let $\hat{G}(a) = 0$, where G is SISO, and $|a| < 1$. By definition of $\hat{G}$, we have $\hat{G}(a) = \sum_{k \geq 0} g(k)a^k = 0$. Define $z_a = (1, a^2, a^3, ...) \in c_0$, then the interpolation condition can be expressed as $< z_a, G > = 0$. If a is a complex number, then two functionals are defined, the real of z_a and the imaginary of z_a. The multivariable case carries more details, but the basic idea is the same (see [10, 44]).

The choice of $\hat{R}_{11}$ is constrained further so that the rest of the equations are still consistent, which in turn dictates a set of constraints on the rest of the elements of R. Define the following coprime polynomial factorizations:

$$
\begin{aligned}
\hat{T}_{22}\hat{T}_{21}^{-1} &= \hat{\tilde{D}}_2^{-1}\hat{\tilde{N}}_2 \\
\hat{T}_{31}^{-1}\hat{T}_{32} &= \hat{N}_3\hat{D}_3^{-1}.
\end{aligned}
\tag{3.3}
$$

Using these definitions, we state the following result characterizing the feasible set $\mathcal{S}$ for this case [44].

Theorem 2.8.1 *Given $\hat{T}_2, \hat{T}_3$ with the assumptions as above, and $\hat{R} \in \mathbf{A}$, there exists $\hat{Q} \in \mathbf{A}$ satisfying $\hat{R} = \hat{T}_2\hat{Q}\hat{T}_3$ if and only if:*

$$
i). \quad \begin{pmatrix} -\hat{\tilde{N}}_2 & \hat{\tilde{D}}_2 \end{pmatrix} \begin{pmatrix} \hat{R}_{11} & \hat{R}_{12} \\ \hat{R}_{21} & \hat{R}_{22} \end{pmatrix} = 0
$$

$$
ii). \quad \begin{pmatrix} \hat{R}_{11} & \hat{R}_{12} \end{pmatrix} \begin{pmatrix} -\hat{N}_3 \\ \hat{D}_3 \end{pmatrix} = 0
$$

$iii).$ $\hat{R}_{11}$ interpolates $\hat{T}_{21}$ and $\hat{T}_{31}$

The conditions shown in parts i, ii are convolution constraints on the ℓ_1 sequence R. The interpolation condition in the last part can be tightened, since only the common zeros of $\hat{T}_{21}$ and $\hat{T}_{22}$ need to be interpolated.

The discussion above shows that the characterization of $\mathcal{S}$ can be summarized by defining two operators,

$$\mathcal{V} : \ell_1^{m \times n} \to \mathbb{R}^s$$

and

$$\mathcal{C} : \ell_1^{m \times n} \to \ell_1^r$$

where s, r are some integers. The first operator captures the interpolation constraints, and thus has a finite dimensional range, and the second captures the convolution constraints. These two operators can be constructed in a straightforward fashion, bookkeeping being the only difficulty. To overcome this problem, it is helpful to think of R as a vector rather than a matrix. To illustrate this, let the operator $\mathcal{W}$ be a map from $\ell_1^{m \times n}$ to ℓ_1^{mn} defined as follows:

$$(\mathcal{W}R)(k) = \begin{pmatrix} r_{11}(k) \\ \vdots \\ r_{m1}(k) \\ r_{21}(k) \\ \vdots \\ r_{mn}(k) \end{pmatrix}.$$

The operator $\mathcal{W}$ is a one-to-one and onto operator, whose inverse is equal to its adjoint (a fact used later). It simply re-arranges the variables in R. The conditions on R presented in the above theorem can be written explicitly in terms of each component of R.

To construct the first operator $\mathcal{V}$, recall that each interpolation condition is interpreted as a bounded linear functional on R. By stacking up these functionals, the opertor $\mathcal{V}$ is constructed. For example, suppose $\hat{T}_{21}$ and $\hat{T}_{31}$ are SISO and both have N zeros a_i in the open unit disc. Then the matrix $\mathcal{V}$ is given by $\mathcal{V} = V_\infty \mathcal{W}$ where

$$V_\infty = \begin{pmatrix} Re(a_1^0) & 0 & 0 & 0 & Re(a_1^1) & 0 & \dots & Re(a_1^j) & 0 & \dots \\ Im(a_1^0) & 0 & 0 & 0 & Im(a_1^1) & 0 & \dots & Im(a_1^j) & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Re(a_N^0) & 0 & 0 & 0 & Re(a_N^1) & 0 & \dots & Re(a_N^j) & 0 & \dots \\ Im(a_N^0) & 0 & 0 & 0 & Im(a_N^1) & 0 & \dots & Im(a_N^j) & 0 & \dots \end{pmatrix} \quad j = 0,1,2,\dots$$

For the second operator, $\mathcal{C}$, recall that convolution can be interpreted as multiplication by a block Toeplitz matrix, in this case with finite memory since $\tilde{N}_2, \tilde{D}_2, N_3$ and D_3 all have finite length (the corresponding λ-transform is a polynomial). By simple rearrangement,

the operator is constructed with its image inside ℓ_1^r. Hence $\mathcal{C}$ is given by $\mathcal{C} = \mathcal{TW}$, where $\mathcal{T}$ is a block lower triangular matrix. For a detailed example, see [11, 44].

To illustrate the construction of the operator $\mathcal{T}$, consider as an example the coprime-factor perturbation problem considered earlier for a SISO. The condition for stability robustness is given by [5]

$$\left\| \begin{bmatrix} \tilde{V} - Q\tilde{N} & -\tilde{U} + Q\tilde{M} \end{bmatrix} \right\|_1 \leq 1.$$

In this case, $T_2 = 1$ and $T_3 = (\tilde{N} \quad -\tilde{M})$. Since $\tilde{M}^{-1}\tilde{N} = NM^{-1}$ with N, M coprime, the conditions in the above theorem translate to

$$(R_{11} \quad R_{12}) \begin{pmatrix} M \\ N \end{pmatrix} = 0$$

The matrix $\mathcal{T}$ is then given by:

$$\mathcal{T} = \begin{pmatrix} (m(0) & n(0)) & 0 & 0 & 0 & 0 & \dots \\ (m(1) & n(1)) & (m(0) & n(0)) & 0 & 0 & 0 & \dots \\ (m(2) & n(2)) & (m(1) & n(1)) & (m(0) & n(0)) & 0 & 0 & \dots \\ (m(3) & n(3)) & (m(2) & n(2)) & (m(1) & n(1)) & (m(0) & n(0)) & 0 & \dots \\ & \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots \end{pmatrix}$$

It is interesting to note that in this example the operator $\mathcal{C}$ captures all the conditions and no interpolation conditions are needed. The conditions presented in the theorem can be redundant, and can be significantly reduced [55].

The subspace $\mathcal{S}$ is then the set of all elements $R \in \ell_1^{m \times n}$ so that $\mathcal{V}R = 0$ and $\mathcal{C}R = 0$. Let $b_1 = \mathcal{V}T_1$, $b_2 = \mathcal{C}T_1$, and $\Phi = T_1 - R$. The ℓ_1 optimization problem can be restated as:

$$\inf_{\Phi \in \ell_1^{m \times n}} \|\Phi\|_1 \qquad \text{subject to} \quad \mathcal{V}\Phi = b_1, \quad \mathcal{C}\Phi = b_2 \qquad (OPT).$$

2.8.2 Relations to Linear Programming

It is well-known that in finite-dimensional spaces ℓ_1-norm minimization is equivalent to linear programming. This turns out to be true in general, and can be justified as follows: Let $\Phi = \Phi^1 - \Phi^2$, with $\phi_{ij}^1(k), \phi_{ij}^2(k) \geq 0$. The norm is then replaced by the function $\max_i \sum_{j,k} \phi_{ij}^1(k) + \phi_{ij}^2(k)$. Define the operator $\mathcal{N} : \ell_1^{m \times n} \to \mathbb{R}^m$ by $(\mathcal{N}\Phi)_i = \sum_{j,k} \phi_{ij}(k)$.

32

The following problem is easily seen to be equivalent to (OPT):

$$\inf \mu$$

subject to

$$\mathcal{N}(\Phi^1 + \Phi^2) - \mu e \leq 0$$

$$\mathcal{V}(\Phi^1 - \Phi^2) = b_1$$

$$\mathcal{C}(\Phi^1 - \Phi^2) = b_2$$

$$\phi_{ij}^1(k), \phi_{ij}^2(k) \geq 0$$

where $e \in \mathbb{R}^m$ and $e^T = (1, 1, \ldots, 1)$. It is interesting to notice that if Φ^1, Φ^2 were restricted to finite impulse response sequences, the above problem is readily a linear programming problem. This will turn out to be a crucial observation in obtaining approximate solutions, as will be described later on.

2.8.3 Lagrange Multiplier Formulation

Let $X = \ell_1^{m \times n} \times \ell_1^{m \times n} \times \mathbb{R}$ and $Z = \mathbb{R}^m \times \mathbb{R}^s \times \mathbb{R}^s \times \ell_1^r \times \ell_1^r$. Let P_X, P_Z denote the positive cones inside X, Z consisting of elements with nonnegative pointwise components. Define the operator $\mathcal{A} : X \to Z$, decomposed conformally with X and Z, and the vector $b \in Z$ as follows:

$$\mathcal{A} = \begin{pmatrix} \mathcal{N} & \mathcal{N} & -e \\ \mathcal{V} & -\mathcal{V} & 0 \\ -\mathcal{V} & \mathcal{V} & 0 \\ \mathcal{C} & -\mathcal{C} & 0 \\ -\mathcal{C} & \mathcal{C} & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ b_1 \\ -b_1 \\ b_2 \\ -b_2 \end{pmatrix}.$$

Define the linear functional $c^* = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ on X. With these definitions, (OPT) becomes:

$$\inf <x, c^*>$$

subject to

$$\mathcal{A}x \leq b$$

$$x \in X, x \geq 0,$$

where $x \in X$ has the form

$$x = \begin{pmatrix} \Phi^1 \\ \Phi^2 \\ \mu \end{pmatrix}.$$

All the inequalities should be interpreted with respect to the positive cones. It is interesting that with the above definitions, (OPT) looks very much like a standard linear programming problem, with the exception that the number of variables and constraints is infinite.

The Lagrange multiplier is an element inside Z^*, the dual space of Z which can be identified as: $Z^* = \mathbb{R}^m \times \mathbb{R}^s \times \mathbb{R}^s \times c_0^r \times c_0^r$. (Here we have assumed that Z is equipped with the $weak^*$ topology, not the norm topology.) The dual cone P_Z^* again consists of the nonnegative elements in Z^*. The Lagrangian can be defined as

$$L(x, z^*) = \{<x, c^*> + <\mathcal{A}x - b, z^*>\}$$
$$= \{<x, c^* + \mathcal{A}^* z^*> - <b, z^*>\}$$

where $\mathcal{A}^* : Z^* \to X^*$ is the adjoint operator of $\mathcal{A}$. From the theory of Lagrange multipliers [43], the minimum solution can be obtained by performing an unconstrained minimization of L, i.e.,

$$\mu_0 = \sup_{z^* \geq 0} \inf_{x \geq 0} \{<x, c^* + \mathcal{A}^* z^*> - <b, z^*>\}$$

Clearly for μ_0 to be finite, i.e $\mu_0 > -\infty$, $c^* + \mathcal{A}^* z^* \geq 0$ and hence the above infimization is achieved for $x = 0$. This gives a dual formulation of (OPT) summarized as:

$$\mu_0 = \sup_{z^* \geq 0} <b, -z^*> \qquad \text{subject to} \quad c^* + \mathcal{A}^* z^* \geq 0 \qquad (DOPT).$$

To evaluate this explicitly, let $\mathcal{A}^*$, z^* be given by:

$$\mathcal{A}^* = \begin{pmatrix} \mathcal{N}^* & \mathcal{V}^* & -\mathcal{V}^* & \mathcal{C}^* & -\mathcal{C}^* \\ \mathcal{N}^* & -\mathcal{V}^* & \mathcal{V}^* & -\mathcal{C}^* & \mathcal{C}^* \\ -e^T & 0 & 0 & 0 & 0 \end{pmatrix} \qquad z^* = \begin{pmatrix} \eta \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

By direct substitution, (DOPT) is converted to

$$\mu_0 = \sup <b_1, \alpha_1 - \alpha_2> + <b_2, \beta_1 - \beta_2>$$

subject to

$$\mathcal{N}^* \eta + \mathcal{V}^*(\alpha_1 - \alpha_2) + \mathcal{C}^*(\beta_1 - \beta_2) \geq 0$$

$$\mathcal{N}^* \eta - \mathcal{V}^*(\alpha_1 - \alpha_2) - \mathcal{C}^*(\beta_1 - \beta_2) \geq 0$$

$$\sum_{i=1}^{m} \eta_i \leq 1$$

$$\alpha_1, \alpha_2, \beta_1, \beta_2, \eta \geq 0.$$

Finally, substituting $\alpha = \alpha_1 - \alpha_2$ and $\beta = \beta_1 - \beta_2$ we get

$$\mu_0 = \sup\ <b_1, \alpha> + <b_2, \beta>$$

subject to

$$-\mathcal{N}^*\eta \leq \mathcal{V}^*\alpha + \mathcal{C}^*\beta \leq \mathcal{N}^*\eta \qquad\qquad (DOPT)$$

$$\sum_{i=1}^{m} \eta_i \leq 1, \eta \geq 0$$

$$\alpha \in \mathbb{R}^s, \quad \beta \in c_0^r.$$

This dual formulation sheds a new light on the optimization problem. In our context, it will provide two important results: the existence of finite-dimensional duals for specific classes of problems, and the ability to construct suboptimal solutions that are within a prescribed ϵ from the actual minimum.

Comment: The computation of the adjoint operators is quite simple once the operators are already constructed. Recall that $\mathcal{V} = V_\infty \mathcal{W}$; hence the adjoint operator $\mathcal{V}^* = \mathcal{W}^{-1}V_\infty^T$. Similarly, $\mathcal{C}^* = \mathcal{W}^{-1}\mathcal{T}^T$. Matrix representations of the operator $\mathcal{N}$ and its adjoint are obtained in a similar fashion.

2.8.4 Exact Solutions for a Class of Problems

Let the space $\mathcal{S}$ be characterized solely by interpolation conditions. This is the situation when both T_2 and T_3 have full row rank and column rank respectively. In this case $\mathcal{C} = 0$ and $b_2 = 0$. The dual problem (DOPT) involves only a finite number of variables and thus it is a finite-dimensional problem. The constraints however are infinite. Since the elements of $\mathcal{V}^*$ were constructed from zeros inside the unit disc, the entries will eventually decay and only a finite number of the constraints are active. A bound on the number of such constraints can be derived [10]. The problem is now a standard finite-dimensional linear program, which can be solved exactly. The solution to the primal problem (OPT) can be constructed either by the alignment conditions, or by observing that the dual of (DOPT) is exactly the primal problem.

2.8.5 Approximation

In the sequel, we will assume that $\mathcal{C}T_1 = b_2$ is a finite impulse response sequence. This condition is equivalent to saying that there exists a FIR feasible solution for (OPT). If this condition is not satisfied, then the problem can be modified so that the condition

will hold [11, 44]. Upper approximations of μ_0 can be readily obtained from the primal problem. Define $\bar{\mu}_N$ as follows:

$$\bar{\mu}_N = \min \mu$$

subject to

$$\mathcal{N}(\Phi^1 + \Phi^2) - \mu e \leq 0$$

$$\mathcal{V}(\Phi^1 - \Phi^2) = b_1$$

$$\mathcal{C}(\Phi^1 - \Phi^2) = b_2$$

$$\phi^1_{ij}(k), \phi^2_{ij}(k) \geq 0$$

$$\phi^1_{ij}(k) = 0, \phi^2_{ij}(k) = 0 \,\forall k > N.$$

Since $\mathcal{C}$ is constructed from FIR sequences, this optimization will involve a finite number of variables and a finite number of constraints. It is evident that $\bar{\mu}_N$ is a non-increasing sequence satisfying $\mu_0 \leq \bar{\mu}_N$ for all N. Also, since a feasible FIR solution exists, then $\bar{\mu}_N$ is finite for N large enough. Since FIR solutions are dense, it follows that $\bar{\mu}_N \to \mu_0$ as $N \to \infty$. For each $\bar{\mu}_N$ a solution for the primal problem can be constructed. The difficulty with this procedure is that it is not clear how far the solution is from optimal at any given N. This will be overcome by presenting lower approximations of the problem.

It is interesting to notice that the dual of this problem is obtained through truncating the constraints of the dual problem (DOPT). Another approximation obtained from the dual problem can be obtained by truncating the variables $\beta \in c_0^r$ [5, 55]. Define $\underline{\mu}_N$ as follows:

$$\underline{\mu}_N = \max < b_1, \alpha > + < b_2, \beta >$$

subject to

$$-\mathcal{N}^* \eta \leq \mathcal{V}^* \alpha + \mathcal{C}^* \beta \leq \mathcal{N}^* \eta$$

$$\sum_{i=1}^{m} \eta_i \leq 1, \eta \geq 0$$

$$\alpha \in \mathbb{R}^s, \ \beta \in c_0^p, \ \beta(k) = 0 \,\forall k > N.$$

It is evident that $\underline{\mu}_N \leq \mu_0$ and that $\underline{\mu}_N \to \mu_0$ as $N \to \infty$. The former assertion is due to the fact that the new problem has fewer degrees of freedom, and the later is due to the fact that finite sequences are dense in c_0. The above problem is not immediately a finite-dimensional problem—the constraints due to the operator $\mathcal{V}^*$ are still infinite; however, only a finite subset of these are active as it was in the case where $\mathcal{C}$ was equal

to 0. A complete discussion of the computation of this problem is given in [55]. Clearly, there is no feasible solution for the primal problem for any of the $\underline{\mu}_N$'s.

2.8.6 Computations

In the case where $C = 0$, the ℓ_1 minimization problem is solved exactly. In all other cases, only approximate solutions are obtained through obtaining upper and lower approximations of μ_0. The major computational burden is in fact obtaining the operator $\mathcal{V}$, since it requires the computation of the zeros of $\hat{T}_{21}$, $\hat{T}_{31}$ and their multiplicities. Work on the computational aspect of this problem is in progress [24].

To obtain fast solutions that do not necessarily capture the structure of this problem, one can follow the approach in [3] in which one seeks direct FIR solutions for Q. This problem can be posed as a linear programming problem which can approximate the actual solution arbitrarily closely. However, unless one invokes duality, the difference between the approximate and actual value of μ_0 remains unknown.

It is interesting to note that exact solutions for special problems with $C \neq 0$ have been constructed in [54]. Although existence of ℓ_1-optimal solutions is guaranteed (under mild conditions, namely no interpolations on the unit circle), it is not known whether these solutions are rational or not. If $C = 0$ optimal solutions are FIR, and hence rational. The general case is still an active area of research.

2.9 Conclusions

This chapter gives an overview of the problem of synthesizing optimal controllers to deliver performance specifications in the time domain, in the presence of bounded but unknown exogenous inputs. A general framework for the robust performance problem is presented from which necessary and sufficient conditions are derived. These conditions were related to the spectral radius of a matrix constructed from the configuration of the closed-loop system. Alternate equivalent conditions are also discussed in terms of linear matrix inequalities. These conditions are in turn used in the synthesis problem, which requires the solution of an ℓ_1 optimal control problem. A solution of this problem using the duality theory of Lagrange multipliers is used. This approach highlights in a non-trivial way the relations between ℓ_1 optimization problems for infinite-dimensional systems and infinite linear programming problems. In fact, the solutions presented ex-

ploit the problem structure and do not rely on some general theory for solving infinite linear programming problems, since such a theory does not exist.

This chapter discusses only discrete-time problems. The interest in discrete-time systems stems from the fact that most controllers these days are digital controllers and are interfaced with the continuous-time plant through A/D and D/A converters. A better formulation should have a hybrid system consisting of both continuous- and discrete-time dynamics. Such systems have recently received considerable attention from the control community and are known as sampled-data systems. A formulation of the ℓ_1 sampled-data problem can be found in [1, 25, 36, 53] in which it is shown that synthesizing a digital controller for a continuous-time plant can be done by solving a purely discrete-time problem. This motivates the earlier discussion.

There are other related problems that are not discussed in this chapter. The problem of designing controllers for tracking a specific trajectory is an important problem and was solved in [12]. The ℓ_1 synthesis approach has also been extended for periodic and multi-rate sampled plants [15]. Also, this theory was successfully incorporated as part of an adaptive control scheme, in which the stability of the closed loop system was guaranteed for a larger set of plant uncertainty [16, 59]. Finally, a case-study for the applicability of this theory was reported in [13] in which a ℓ_1 controller was designed for a model of the $X - 29$ aircraft.

A pressing research problem is the understanding of the structure of the optimal ℓ_1 controllers. Such an understanding will not only add insight into the problem, but will also offer simpler ways of computing the optimal solution. This has been the case for the H_∞ and H_2 problems. Some interesting results in that direction are reported in [54] in which exact solutions for the infinite-dimensional linear programs arising in some special non-square problems have been computed. Also, it was shown in [23] that optimal solutions may require a dynamic controller even though all the states are available. The existence of some separation structure of the ℓ_1 problem (similar to that of the H_∞ problem [20]) is still under investigation.

Another important research direction is the synthesis problem by exactly minimizing the spectral radius function, rather than the iterative scheme suggested. The iterative scheme is guaranteed to converge only to local minima and hence there is a need for looking for another approach for minimizing this function.

In this chapter, a comparison between the spectral radius function and μ is sketched.

At this point, it is not known whether there exist examples in which the two methods exhibit extreme behavior. Research in that direction is currently in progress.

Chapter 3

State Feedback ℓ_1-Optimal Controllers can be Dynamic

This chapter considers ℓ_1-optimal control problems given by discrete-time systems with full state feedback, scalar control and scalar disturbance. Motivation stems from the central role that this problem structure played in the development of the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theories. First, systems with a scalar regulated output are studied (singular problems). Sufficient conditions are given, based on the non-minimum phase zeros of the transfer function from the control to the regulated output, for the existence of a static ℓ_1-optimal controller. A simple way to compute the static gain is provided, using pole placement ideas. It is shown, however, that having full state information does not prevent the ℓ_1-optimal controller from being dynamic in general, and that examples with arbitrarily high order optimal controllers can be easily constructed.

Second, problems with two regulated outputs, one of them being the scalar control, are considered (non-singular problems). It is shown, by means of a class of fairly general examples for which exact ℓ_1-optimal solutions are constructed, that such problems may not have static controllers that are ℓ_1-optimal. Thus concluding that a "separation structure" does not occur in these problems in general.

3.1 Introduction

Since Dahleh and Pearson ([10],[11]) presented the solution to the ℓ_1 optimal control problem, there has been increasing interest in understanding the basic properties of such problems ([5],[44], [45] and [54]). Considering that in the case of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimization

([20],[65]), state feedback optimal controllers have a very special structure (i.e. static), it seems only natural to ask how full state information affects the ℓ_1 optimal solution. In particular, under what conditions (if any) there exist a static linear controller that achieves optimality. This chapter presents results regarding this question, for systems with scalar control and scalar disturbance. In particular, two different types of problems within this class of systems are considered: a) those with a scalar regulated output, denoted as *singular problems*, and b) those with two regulated outputs, denoted as *non-singular problems*, where one of the outputs is the scalar control signal. For systems in a), it is shown that there exists a static controller which is ℓ_1-optimal if the non-minimum phase zeros of the transfer function from the control input to the regulated output satisfy a simple algebraic condition. Violating such condition, however, may result in a dynamic ℓ_1-optimal controller of possibly high order (generally when the non-minimum phase zeros are "close" to the unit circle). For systems in b), it is shown by means of an example that optimal controllers are dynamic in a broad class of cases which are common in control design. The difficulty in analyzing the non-singular problem is that it is not straightforward to compute the optimal solution, as it is the case with a singular (i.e. square) one. For the given non-singular example, the optimal solution is constructed and shown to require a dynamic compensator.

The chapter is organized as follows. Section 2 formulates the singular problem along with some basic notation. Sections 3 and 4 present results corresponding to singular problems involving minimum and non-minimum phase plants respectively. Section 5 examines a non-singular problem by means of a general example, followed by the conclusions in Section 6.

3.2   Problem Formulation

Consider the following state-space minimal realization of a full state feedback system with scalar input disturbance, scalar control, scalar regulated output:

$$\left(\begin{array}{c|cc} A & b_1 & b_2 \\ \hline c_1 & 0 & d_{12} \\ I & 0 & 0 \end{array} \right)$$

where $A \in \mathbb{R}^{n \times n}$, b_1 and $b_2 \in \mathbb{R}^{n \times 1}$, $c_1 \in \mathbb{R}^{1 \times n}$, and $d_{12} \in \mathbb{R}$. For any internally stabilizing controller k, let $\phi = \{\phi(0), \phi(1), \phi(2), \ldots\}$ denote the closed-loop pulse response

41

sequence from the disturbance to the regulated output. Then, the problem can be stated as follows:

$$\inf_{k-stab.} \|\phi\|_1 \tag{3.1}$$

where $\|\phi\|_1 \overset{\text{def}}{=} \sum_i |\phi(i)|$. Using standard results in the parameterization of all stabilizing controllers (see [28]), problem (3.1) can be rewritten as follows:

$$\inf_{q \in \ell_1^{1 \times n}} \|h - u * q * v\|_1 \tag{3.2}$$

where $\ell_1^{m \times n}$ indicates the space of all $m \times n$ matrices with entries in ℓ_1 and $*$ denotes convolution. Thus, h and $u \in \ell_1$, and $v \in \ell_1^{n \times 1}$. Let the λ-transform of a right-sided real sequence $x = \{x(0), x(1), x(2), \ldots\}$ be defined as

$$\hat{x}(\lambda) = \sum_{k=0}^{\infty} x(k) \lambda^k$$

where λ represents the unit delay. Then, a state-space realizations for h, u and v can be found by using the state-space formulas in [28] with the observer gain matrix, H, equal to $-A$. For this specific choice, the realizations are:

$$\hat{h}(\lambda) = \lambda [A_f, A b_1, c_1 + d_{12} f, c_1 b_1] \tag{3.3}$$

$$\hat{u}(\lambda) = [A_f, b_2, c_1 + d_{12} f, d_{12}] \tag{3.4}$$

$$\hat{v}(\lambda) = [0, b_1, I, 0] = \lambda b_1 \tag{3.5}$$

where $\hat{h}(\lambda)$, $\hat{u}(\lambda)$ and $\hat{v}(\lambda)$ denote the λ-transform of h, u and v; $A_f \overset{\text{def}}{=} A + b_2 f$,

$$[A, B, C, D] \overset{\text{def}}{=} \lambda C (I - \lambda A)^{-1} B + D$$

and f is chosen so that all the eigenvalues of A_f are inside the unit disk.

The following result, which will be needed in the next section, is proved in [10]

Theorem 3.2.1 *Assuming $\hat{u}(\cdot)$ and $\hat{v}(\cdot)$ have no left and right zeros respectively on the unit circle, there exists $q_{opt} \in \ell_1^{1 \times n}$ that achieves the optimal norm in problem (3.2). Moreover, the closed-loop optimal pulse response, $\phi_{opt} = h - u * q_{opt} * v$, has finite support.*

3.3   Singular Problems with Minimum-Phase Plants

This section considers the case where the transfer function from the control input to the regulated output is minimum-phase except for an integer number of unit delays

(i.e. zeros at the origin in the λ-plane). It will be assumed throughout that (A, b_2) is reachable.

Theorem 3.3.1 *For such a system, the static feedback gain, f^*, that places the eigenvalues of $(A + b_2 f^*)$ at the exact location of the minimum-phase zeros of $[A, b_2, c_1, d_{12}]$ and the rest at the origin is ℓ_1-optimal.*

Proof. Consider using f^* as the state feedback gain in the parameterization described above. Then, after carrying out all stable pole-zero cancellations,

$$\hat{u}(\lambda) = \gamma_r \, \lambda^r$$

where r is the number of unit delays in $[A, b_2, c_1, d_{12}]$ and γ_r is a scalar depending on r. In what follows, the cases where $r = 0$ and $r > 0$ will be treated separately.

i) If $r = 0$, then $d_{12} \neq 0$, $c_1 + d_{12} f^* = 0$, and $\hat{u}(\lambda) = d_{12}$. Also, from equation (3),

$$\hat{h}(\lambda) = c_1 b_1 \lambda = (c_1 b_1 + d_{12} f^* b_1 - d_{12} f^* b_1)\lambda = -d_{12} f^* b_1$$

$$\implies \hat{\phi}(\lambda) = -d_{12} f^* b_1 \lambda - d_{12} \hat{q}(\lambda) b_1 \lambda$$

Thus, the ℓ_1-optimal solution is given by $\hat{q}_{opt}(\lambda) = -f^*$, and $\hat{\phi}_{opt}(\lambda) = 0$. Furthermore, using the state-space equations in [28] for computing the optimal controller, it can be shown after a little algebra that $\hat{k}_{opt}(\lambda) = f^*$.

ii) If $r > 0$, then $d_{12} = 0$, $c_1 A_{f^*}^r = 0$ by construction since (A, b_2) is reachable. Also $\hat{u}(\lambda) = c_1 A_{f^*}^{r-1} b_2 \lambda^r$. Again, from equation 3,

$$\hat{h}(\lambda) = c_1 b_1 \lambda + c_1 A b_1 \lambda^2 + c_1 A_{f^*} A b_1 \lambda^3 + \cdots + c_1 A_{f^*}^{r-1} A b_1 \lambda^{r+1}$$

Therefore, the closed-loop pulse response is given by

$$\begin{aligned}
\hat{\phi}(\lambda) &= c_1 b_1 \lambda + c_1 A b_1 \lambda^2 + c_1 A_{f^*} A b_1 \lambda^3 + \cdots \\
&\quad + c_1 A_{f^*}^{r-2} b_1 \lambda^r + c_1 A_{f^*}^{r-1}(A - b_2 \hat{q}(\lambda)) b_1 \lambda^{r+1}
\end{aligned}$$

Clearly, q does not affect the first $r + 1$ elements of ϕ (i.e. $\phi(i), i = 0, 1, \ldots, r$). Then, the best possible choice of q, in the sense of minimizing the ℓ_1-norm of ϕ, is the one that makes $\phi(i) = 0$ for $i = r + 1, r + 2, \ldots$, and is achieved by letting $\hat{q}_{opt}(\lambda) = -f^*$, since $\phi(r + 1) = c_1 A_{f^*}^r b_1 = 0$. Again, the corresponding ℓ_1-optimal controller is f^*. ∎

43

Corollary 3.3.1 *The ℓ_1-optimal closed-loop transfer function of the system considered in Theorem 3.3.1 (with $r > 0$) is given by:*

$$\hat{\phi}_{opt}(\lambda) = c_1 \sum_{i=1}^{r} A^{i-1} \lambda^i b_1$$

Proof. It follows from the fact that $c_1 A^i_{f\bullet} b_2 = 0$ for $i = 0, 1, \ldots, r - 2$. The details are left to the reader. ∎

Put in words, Theorem 3.3.1 says that there is nothing the controller can do to invert the delays in the system. It can, however, cancel the rest of the dynamics of the system due to the absence of non-minimum phase zeros in the transfer function from the control input to the regulated output. This results in an optimal closed-loop pulse response that is equal to the open loop pulse response in its first $(r + 1)$ elements and zero thereafter. It is also worth noting that Theorem 3.3.1 is directly applicable to the discrete-time LQR problem, where $\sum_i \phi_i^2$ is minimized. More precisely, the asymptotic LQR solution (see [40]) where the weight on the control tends to zero (i.e. cheap control problem) is identical to that of Theorem 3.3.1.

3.4 Singular Problems with Non-minimum Phase Plants

This section considers those cases where $[A, b_2, c_1, d_{12}]$ has r non-minimum phase zeros not necessarily at the origin (i.e. $\lambda = 0$).

Again, we use the same parameterization as in the previous section. That is, we choose f^* to place $(n - r)$ eigenvalues of A_f^* at the exact location of the minimum phase zeros of $[A, b_2, c_1, d_{12}]$ and the rest (r) at the origin. Then, from the discussion in section 3, $\hat{h}(\lambda)$ is polynomial in λ and of order $(r + 1)$, $\hat{u}(\lambda)$ is polynomial too, but of order r, and $\hat{v}(\lambda)$ is simply λb_1. Therefore, the closed-loop transfer function can be written as follows:

$$\hat{\phi}(\lambda) = \left(g_1 \prod_{i=1}^{r} (\lambda - \alpha_i) - g_2 \prod_{j=1}^{r} (\lambda - \beta_j) \tilde{q}(\lambda) \right) \lambda \overset{\text{def}}{=} \tilde{\phi}(\lambda) \lambda \tag{3.6}$$

where $g_1, g_2 \in \mathbb{R}$, α_i's are the zeros of $\hat{h}$, β_j's are the (non-minimum phase) zeros of $\hat{u}$ and $[A, b_2, c_1, d_{12}]$, and $\tilde{q}(\lambda) \overset{\text{def}}{=} \hat{q}(\lambda) b_1 \in \ell_1$. Note that $\|\phi\|_1 \equiv \|\tilde{\phi}\|_1$. Also, by Theorem 3.2.1, $\tilde{\phi}_{opt}(\lambda)$ is polynomial in λ, which implies that $\tilde{q}_{opt}(\lambda)$ is polynomial in

44

λ. Thus, the optimization problem is equivalent to the following linear programming (primal) problem: for a sufficiently large but finite s,

$$\min_{\tilde{\phi}} \sum_{i=0}^{s} \left| \tilde{\phi}(i) \right| \tag{3.7}$$

$$s.t. \quad \sum_{i=0}^{s} \tilde{\phi}(i) \, \beta_j^i = g_1 \prod_{i=1}^{r} (\beta_j - \alpha_i) \, , \; j = 1, 2, \ldots, r$$

In the above we have assumed that the β_j's are simple zeros to simplify the formulation of the interpolation conditions. The following results, however, carry over to the more general case.

The following theorem by Deodhare and Vidyasagar [17] will prove useful. It is stated with no proof.

Theorem 3.4.1 *The support of $\tilde{\phi}$ in (3.7), denoted as $(s + 1)$, equals the number of constraints r, if*

$$\sum_{i=0}^{r-1} |a_i| < 1 \tag{3.8}$$

where $\prod_{j=1}^{r} (\lambda - \beta_j) = \lambda^r + a_{r-1}\lambda^{r-1} + \cdots + a_1\lambda + a_0$.

Now we are ready to present the next result.

Theorem 3.4.2 *Let $[A, b_2, c_1, d_{12}]$ have r non-minimum phase zeros, then if (3.8) is satisfied, f^* is ℓ_1-optimal.*

Proof. By Theorem 3.4.1, $\tilde{\phi}_{opt}(\lambda)$ is of order $(r - 1)$. Then, considering the order of each term in (3.6), it is clear that $\hat{q}_{opt}(\lambda)$ has to be constant and such that $\tilde{\phi}(r) = 0$. Using the state-space formulas (3), (4) and (5),

$$\begin{aligned}
0 = \tilde{\phi}(r) &= (c_1 + d_{12}\, f^*)A_{f^*}^{r-1}(Ab_1 - b_2\, \tilde{q}_{opt}(0)) \\
&= (c_1 + d_{12}\, f^*)A_{f^*}^{r-1}(A - b_2\, q_{opt}(0))b_1
\end{aligned}$$

But, by construction, $(c_1 + d_{12}f^*)A_{f^*}^{r} = 0$ due to the stable pole-zero cancellations and the fact that the rest of the poles are placed at the origin. Therefore, $\hat{q}_{opt} = -f^*$ is the required value, and $\hat{k}_{opt} = f^*$. ∎

Observation: It remains to consider those cases where the non-minimum phase zeros of $[A, b_2, c_1, d_{12}]$ are such that they violate condition (3.8). Theorem 3.4.1 established a

sufficient condition to determine the order of the optimal response. If (3.8) is violated, the optimal closed-loop response may be of higher order, possibly greater than n, but still polynomial. If that is the case, then the ℓ_1-optimal controller is necessarily dynamic, since the highest order polynomial response that a static controller can generate is n by placing all closed-loop poles of the plant at the origin. Any polynomial response of order greater than n, say N, requires a dynamic compensator of at least order $N - n$. Thus, $\sum_{i=0}^{r-1} |a_i| \geq 1$ can be viewed as a necessary condition for the optimal controller to be dynamic.

The following example shows that a large class of state feedback singular problems have this property.

EXAMPLE 1: Consider the following parameterized family of plants (with parameter κ),

$$P_\kappa(\lambda) = \frac{\lambda(\kappa\lambda^2 - 2.5\lambda + 1)}{(1 - 0.2\lambda)(23\lambda^2 - 2.5\lambda + 1)}$$

Assume that the controller has access to the state vector and that the disturbance acts at the plant input. The non-minimum phase zeros relevant to this theory are given by the roots of $\kappa\lambda^2 - 2.5\lambda + 1$, as a function of κ. It is easy to see that for $\kappa > 3.5$ condition (3.8) is satisfied and the optimal controller is f^*. By applying the methods of [10], it can be shown that for $\kappa = 3.5$ the optimal solution is no longer unique. Actually two possible solutions with $\|\phi_{opt}\|_1 = 7$ are:

$$\hat{\phi}_{opt_{\kappa=3.5}} = \begin{cases} \lambda - 2.5\lambda^2 + 3.5\lambda^3 \\ \lambda - 1.1\lambda^2 + 4.9\lambda^4 \end{cases}$$

The first is achieved with f^* while the second requires a first order controller. (The non-uniqueness is related to the occurrence of weakly redundant constraints in the linear program.) Note that for this value of κ, the left hand side of (3.8) is equal to one.

For $1.5 < \kappa < 3.5$ condition (3.8) is violated and the optimal solution has the following general form:

$$\hat{\phi}_{opt_{1.5<\kappa<3.5}} = \lambda + \phi_\kappa(2)\lambda^2 + \phi_\kappa(N_\kappa)\lambda^{N_\kappa}$$

As $\kappa \searrow 1.5$, one of the non-minimum phase zeros approaches the boundary of the unit disk while $\phi_\kappa(2) \longrightarrow -1.5$, $\phi_\kappa(N_\kappa) \longrightarrow 0.5$, and, most remarkably, $N_\kappa \nearrow \infty$. This implies that the optimal controller can have arbitrarily large order. For instance, if $\kappa = 1.51$, then

$$\hat{\phi}_{opt_{\kappa=1.51}} \simeq \lambda - 1.4907\lambda^2 + 0.5776\lambda^{12}$$

46

and the optimal compensator is of order 9. It is also interesting to point out that for $\kappa < 1.5$ one of the non-minimum phase zeros leaves the unit disk and condition (3.8) is again satisfied. In this case, $\hat{\phi}_{opt_{\kappa<1.5}} = \lambda - 1.5\lambda^2$ and $\hat{k}_{opt} = f^*$. With regard to the optimal norm, it drops from a value arbitrarily close but greater than 3 to a value of 2.5 in the transition.

Similar behavior has been reported in [46], for the case of sensitivity minimization through output feedback. The above example shows that the nature of such solutions have comparable characteristics even under full state feedback. There is one difference, however, which reflects the structure added to the problem. In [46] a parameterized family of first order systems was constructed with arbitrarily high order optimal controllers, while this setup requires at least a second order plant with two non-minimum phase zeros away from the origin. Note that condition (3.8) is automatically satisfied otherwise.

3.5   A Non-Singular Problem

So far we have considered problems with a scalar regulated output. One could argue that sensitivity minimization problems, such as the one in the above example, where a measure of the control effort is not included in the cost functional (i.e. singular problems), may have peculiar solutions that could hide the structure of the more general non-singular case. To clarify this point, we will consider a variation of the above example by including the control effort in the cost functional. That is,

$$\mu = \inf_{k-stab.} \left\| \begin{array}{c} \phi_1 \\ \gamma\,\phi_2 \end{array} \right\|_1 \stackrel{\text{def}}{=} \inf_{k-stab.} max(\|\phi_1\|_1, \gamma\|\phi_2\|_1) \qquad (3.9)$$

where ϕ_1 represents the closed-loop map from the disturbance to the output of the plant, ϕ_2 represents the closed-loop map from the disturbance to the control input, and γ is a positive scalar weight. The fact that there are two regulated outputs and only a scalar control makes this problem of the *bad rank* class (i.e. two-block column problem, see [11] and [44]). This implies that a linear programming formulation of the solution will have, in general, an infinite number of non-zero variables and active constraints (Theorem 3.2.1 no longer holds) making the construction of exact solutions a non trivial task. For the following example, however, it is shown that the optimal response has finite support, and that an exact solution can be computed by the methods in [54] and [24].

EXAMPLE 2: Consider problem (3.9) for the parameterized family of plants of Example 1. By expanding each term, Equation 3.9 can be rewritten as:

$$\bar{\mu} = \inf_{q \in \ell_1^{1 \times n}} \left\| \begin{pmatrix} h_1 \\ \gamma h_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ \gamma u_2 \end{pmatrix} * q * v \right\|_1 \tag{3.10}$$

where, according to the previous parameterization (and using the same notation),

$$
\begin{aligned}
\hat{h}_1(\lambda) &= \lambda[A_{f^*}, Ab_1, c_1 + d_{12}f^*, c_1 b_1] \\
\hat{h}_2(\lambda) &= \lambda[A_{f^*}, Ab_1, f^*, 0] \\
\hat{u}_1(\lambda) &= [A_{f^*}, b_2, c_1 + d_{12}f^*, d_{12}] \\
\hat{u}_2(\lambda) &= [A_{f^*}, b_2, f^*, 1] \\
\hat{v}(\lambda) &= \lambda b_1
\end{aligned}
\tag{3.11}
$$

With the particular problem data and $\kappa = 2$, $\hat{v}(\lambda)$ has a right zero at the origin and $(\hat{u}_1(\lambda) \quad \gamma \hat{u}_2(\lambda))^T$ has no left zeros. Then, the optimization problem can be posed in the primal space, ℓ_1, as follows ([11]):

$$\mu = \inf_{\phi_1, \phi_2} \left\| \begin{matrix} \phi_1 \\ \gamma \phi_2 \end{matrix} \right\|_1 \tag{3.12}$$

subject to:

$$
\begin{aligned}
\phi_1(0) &= 0 \\
\phi_2(0) &= 0 \\
(u_2 * \phi_1 - u_1 * \phi_2)(k) &= (u_2 * h_1 - u_1 * h_2)(k), \quad k = 0, 1, 2, \ldots
\end{aligned}
\tag{3.13}
$$

Or in the dual space, ℓ_∞, as

$$\mu = \sup_{\alpha, \beta_1, \beta_2, \tau_1, \tau_2} \sum_{i=0}^{\infty} \alpha(i)(u_2 * h_1 - u_1 * h_2)(i) \tag{3.14}$$

subject to:

$$
\begin{aligned}
\left| \left(\sum_{i=0}^{\infty} \alpha(i+k) u_2(i) \right) + (0^k)\beta_1 \right| &\leq \tau_1 \\
\left| \left(\sum_{i=0}^{\infty} \alpha(i+k) u_1(i) \right) - (0^k)\beta_2 \right| &\leq \tau_2 \\
\tau_1 + \tau_2 &\leq 1
\end{aligned}
\tag{3.15}
$$

for $k = 0, 1, 2, \ldots$, where $\alpha \in \ell_\infty$ and $\beta_1, \beta_2, \tau_1, \tau_2 \in \mathbb{R}$.

Let $\bar{\mu}_N$ denote the value of (3.12) when the constraints $\phi_i(k) = 0$ for all $k > N, i = 1, 2$ are appended to (3.13), and let $\underline{\mu}_M$ denote the value of (3.14) when the constraints $\alpha(k) = 0$ for all $k > M$ are appended to (3.15). Then, clearly

$$\underline{\mu}_M \leq \mu \leq \bar{\mu}_N \tag{3.16}$$

48

for all positive integers M and N. We will refer to these problems as the *truncated primal* and the *truncated dual* problem respectively. Next, let $\gamma = 0.1$, $N = 5$ and $M = 13$, then the following are exact solutions to the truncated primal and truncated dual problems (within 15 digits accuracy):

$$\begin{aligned}
\hat{\phi}_1(\lambda) &= \lambda - \tfrac{887}{558}\lambda^2 + \tfrac{631}{558}\lambda^4 + \tfrac{308}{558}\lambda^5 \\
\hat{\phi}_2(\lambda) &= -\tfrac{998.6}{558}\lambda + \tfrac{11895.4}{558}\lambda^2 + \tfrac{8955.4}{558}\lambda^3 + \tfrac{1282.2}{558}\lambda^4 - \tfrac{708.4}{558}\lambda^5
\end{aligned} \qquad (3.17)$$

$$\begin{pmatrix} \tau_1 \\ \tau_2 \\ \beta_1 \\ \beta_2 \\ \alpha(0) \\ \vdots \\ \alpha(13) \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0.970131421744327 \\ 0.0298685782556741 \\ 18.1617920759050 \\ 42.75388829151732 \\ 0 \\ 37.7844820452347 \\ -3.29789187316069 \\ -1.65259869560944 \\ -0.401868143803625 \\ 0.309029878922354 \\ 0.572287131426917 \\ 0.545909685694633 \\ 0.411177830532670 \\ 0.255951734446358 \\ 0.129285041919450 \\ 0.0273249362257737 \\ -0.0155520615496705 \\ -0.0181682559221378 \\ 0 \\ \vdots \end{pmatrix} \qquad (3.18)$$

It is easy to verify using the values in (3.17) and (3.18) that $\underline{\mu}_{13} = \overline{\mu}_5 = 1192/279 \simeq 4.2724$, thus, from (3.16), $\mu = 1192/279$ and (3.17) is the exact solution to the full primal problem (3.12). Therefore, since such solution has finite support and is of fifth order, the optimal controller is necessarilly dynamic and of second order. Also note that the optimal closed loop response is such that $\|\phi_1\|_1 = \gamma\|\phi_2\| = \mu$.

It is also interesting to consider the singular problem corresponding to this example (i.e. $\kappa = 2$ and $\gamma = 0$). The optimal solution (which is obtained by eliminating the second row and solving the resulting *good rank* problem) is given by:

$$\begin{aligned}
\hat{\phi}_1(\lambda) &= \lambda - \tfrac{90}{68}\lambda^2 + \tfrac{128}{68}\lambda^5 \\
\hat{\phi}_2(\lambda) &= -\tfrac{103.6}{68}\lambda + \tfrac{1446}{68}\lambda^2 + \tfrac{1394.4}{68}\lambda^3 + \tfrac{1136}{68}\lambda^4 - \tfrac{294.4}{68}\lambda^5
\end{aligned} \qquad (3.19)$$

49

where $\|\phi_1\|_1 = 286/68 \simeq 4.2059$ while $\|\phi_2\|_1 = 4374.4/68 \simeq 64.3294$ is clearly larger since it was left out of the optimization. In fact, the above solution is valid for $\gamma \in [0, 286/4374.4]$ since for any γ in such interval $\|\phi_1\|_1 \geq \gamma\|\phi_2\|_1$. Moreover, for any such γ, the ℓ_1-optimal controller is dynamic and of second order since the optimal ϕ_1 is polynomial and of fifth order. This alone constitutes a family of problems, parameterized by $\gamma \in [0, 286/4374.4]$, requiring dynamic optimal controllers.

All this indicates that given a non-singular (two-block) problem, the optimal controller may very well be dynamic, whether or not the two regulated outputs impose conflicting goals (i.e. active constraints). Further, it can be shown that even when the corresponding singular problem has a static optimal controller, the non-singular problem may require a dynamic one. This will happen only if γ is large enough to make the second row of the cost functional active in the optimization.

A last question remains to be answer: given a full-state feedback problem with a dynamic ℓ_1-optimal controller, is it possible to find a static controller that achieves an ℓ_1-norm arbitrarily close to the optimal? Again, it is easy to show via a counter example (numerical) that this is not the case. In fact, a simple second order problem can show that the gap between the norms achieved by the optimal and the static-optimal controller can be significant.

3.6 Concluding Remarks

This chapter presented a study of the ℓ_1 optimization problem for systems with full state feedback, scalar disturbance and scalar control. Two classes of problems were considered: a) singular problems with a scalar regulated output, and b) non-singular two-block problems with two regulated outputs, one of them being the control sequence. The main purpose of the study was to determine whether or not there is always a static controller which is ℓ_1-optimal. In the case of singular problems, a sufficient condition was given, based on the non-minimum phase zeros of the transfer function from the control to the regulated output, for the existence of a static ℓ_1-optimal controller. The optimal gain is such that it places a subset of the closed-loop poles at the exact location of the minimum phase zeros of the transfer function from the disturbance to the regulated output and the rest at the origin. Then, it was shown by means of general examples, that both singular as well as non-singular problems may require dynamic optimal controllers

of arbitrarily high order, in spite of the perfect state information. This adds to the observations made in [46] where singular problems with output feedback were considered. In fact, it can be shown using similar arguments that full information problems (where the disturbance is measured exactly) also have these characteristics.

Although the systems in question were simple, it is safe to conclude that more complex MIMO state feedback ℓ_1 optimization problems will also have these characteristics in general. Therefore, it is doubtful that the study of the full state feedback problem will render a "separation structure" similar to the ones found in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimization theory ([20], [65]).

Chapter 4

Rejection of Persistent Bounded Disturbances: Nonlinear Controllers

This chapter considers nonlinear time-varying (NLTV) compensation for linear time-invariant (LTI) plants subject to persistent bounded disturbances. It is shown using two different approaches that using NLTV compensation instead of LTI compensation does not improve the optimal rejection of persistent bounded disturbances. The first approach is to derive a bound on the achievable performance over all stabilizing NLTV controllers. Using results from ℓ_1-optimal control, it follows that in some special cases this bound can be achieved by LTI compensation. This approach involves the introduction of an operator analogous to the Hankel operator in $\mathcal{H}^\infty$-optimal control and is of independent interest. The second approach is to assume the NLTV controller is sufficiently smooth to admit a time-varying linearization. This time-varying linearization is then used to construct an LTI controller which achieves the same performance as the original NLTV controller. These results extend previous work by the authors regarding linear time-varying compensation.

Notation

LTI := linear time-invariant

LTV := linear time-varying

NLTV := nonlinear time-varying

$$\ell_\infty := \left\{ f = (\ldots, f(-1), f(0), f(1), f(2), \ldots) : \|f\|_\infty \stackrel{\text{def}}{=} \sup_n |f(n)| < \infty \right\}$$

$$\ell_\infty(Z_+) := \{ f \in \ell_\infty : f(n) = 0, \forall n < 0 \}$$

$$\ell_1 := \left\{ f \in \ell_\infty : \|f\|_1 \stackrel{\text{def}}{=} \sum_n |f(n)| < \infty \right\}$$

$$c^0 = \left\{ f \in \ell^\infty \big| \lim_{k \to \infty} x(k) = 0 \right\}$$

$$\|T\| := \sup_{\substack{f \in \ell_\infty \\ f \neq 0}} \frac{\|Tf\|_\infty}{\|f\|_\infty}$$

$$\Pi_- f(n) = \begin{cases} f(n), & n \leq 0; \\ 0, & n > 0. \end{cases}$$

4.1 Problem Statement

In this chapter, we consider the use of NLTV compensation to achieve optimal disturbance rejection with LTI plants. This problem has been considered in [2, 27, 37, 39, 51]. In [27, 37, 39], it was shown that NLTV compensation does not improve the optimal rejection of finite-energy (i.e., ℓ^2) disturbances. In [51], it was shown that LTV compensation does not improve the optimal rejection of persistent bounded (i.e., ℓ_∞) disturbances. This was extended to continuous-time systems in [2]. Possible advantages of NLTV control are discussed in [38] and references contained therein.

The results of [2, 51] hold for LTV compensation only. In this chapter, we consider *nonlinear* compenstation with persistent bounded disturbances. It is shown using two approaches that NLTV compensation again does not improve the optimal disturbance rejection of ℓ_∞ disturbances. The first approach involves the introduction of an operator analogous to the Hankel operator in $\mathcal{H}^\infty$-optimal control which is of independent interest. This operator leads to a bound on the achievable performance which can be achieved by LTI compensation. The second approach uses a linearization of the nonlinear controller

to construct an LTI controller which achieves the same performance as the original NLTV controller.

In the discussion that follows, familiarity with the disturbance rejection problem framework and related notions of stabilization, causality, and well-posedness is assumed (cf., [28, 60]). In particular, unless otherwise specified, all operators are norm-bounded causal mappings over signals with support $(-\infty, \infty)$.

To set up the problem, let $T_{zw}(K)$ denote the closed-loop mapping from the exogenous disturbances, w, to the regulated variables, z, as a function of the controller K. Let $\mathcal{N}_{TV}$ denote all norm-bounded causal NLTV operators on ℓ_∞. Let $\mathcal{L}_{TV}$ denote the subset of $\mathcal{N}_{TV}$ which are linear. Similarly, let $\mathcal{L}_{TI}$ denote the subset of $\mathcal{L}_{TV}$ which are time-invariant. The λ transform of an element $H \in \mathcal{L}_{TI}$ will be denoted by $\hat{H}(\lambda)$. It can be shown (cf., [28, 56]) that the problem reduces to comparing the following quantities:

$$\mu_{NL} \overset{\text{def}}{=} \inf \{\|T_{zw}(K)\| : K \text{ is any stabilizing NLTV controller}\}$$
$$= \inf_{Q \in \mathcal{N}_{TV}} \|T_1 - T_2 Q T_3\|.$$

$$\mu_{TV} \overset{\text{def}}{=} \inf \{\|T_{zw}(K)\| : K \text{ is any stabilizing LTV controller}\}$$
$$= \inf_{Q \in \mathcal{L}_{TV}} \|T_1 - T_2 Q T_3\|.$$

$$\mu_{TI} \overset{\text{def}}{=} \inf \{\|T_{zw}(K)\| : K \text{ is any stabilizing LTI controller}\}$$
$$= \inf_{Q \in \mathcal{L}_{TI}} \|T_1 - T_2 Q T_3\|.$$

Here, $T_{1,2,3} \in \mathcal{L}_{TI}$ are discrete-time multiple-input/multiple-output systems determined by the discrete-time LTI plant and disturbance rejection problem under consideration. In the remainder of this chapter, any extra assumptions on $T_{1,2,3}$ will be introduced as needed.

The following theorem concerns LTV compensation:

Theorem 4.1.1 ([51]) $\mu_{TV} = \mu_{TI}$.

In this chapter, we show that under certain conditions $\mu_{NL} = \mu_{TI}$.

54

4.2 Main Results

4.2.1 A Hankel-like Operator

In this section, a Hankel-like operator (cf., [28]) is defined for general operators on ℓ_∞. This Hankel-like operator leads to a lower bound on the achievable performance over NLTV compensators. In the special case where $T_2 = I$, it is shown that this bound may be achieved by an LTI compensator. In this section, $T_{1,2,3}$ are assumed to be single-input/single-output with the exception of Example 2.2.

Let $a_1, \ldots, a_n, b_1, \ldots, b_m$ be the zeros of the transfer functions of T_2 and T_3 respectively inside in the open unit disc. For simplicity, assume they are real and distinct. The forthcoming analysis still goes through in the general case. Define the functions

$$u_{a_j}(k) = a_j^{-k} \qquad \forall k \leq 0, j = 1, \ldots, n$$

$$v_{b_j}(k) = b_j^{-k} \qquad \forall k \leq 0, j = 1, \ldots, m$$

Let

$$\mathcal{U} = span\{u_{a_i}\}$$

$$\mathcal{V} = span\{v_{b_j}\}$$

$\mathcal{U}$ and $\mathcal{V}$ are subspaces inside ℓ_∞ supported on the nonpositive integers. Given any operator H on ℓ_∞, define a Hankel-like operator as follows

$$\Gamma_H : \mathcal{U} + \mathcal{V} \to \ell_\infty$$

$$u + v \to \Pi_-(u * He + e * Hv)$$

where Π_- denotes the projection on the nonpositive integers, $*$ denotes convolution, and e denotes the unit pulse at the origin,

$$e(k) = \begin{cases} 1, & k = 0; \\ 0, & k \neq 0. \end{cases}$$

In the case of an LTI operator H, the operator Γ_H has a simple representation. Let $u + v \in \mathcal{U} + \mathcal{V}$, then

$$u + v = \sum_{i=1}^{n} \alpha_i u_{a_i} + \sum_{i=1}^{m} \beta_i v_{b_i}$$

55

and

$$\Pi_-(H(u+v)) = \sum_{i=1}^{n} \alpha_i \hat{H}(a_i) u_{a_i} + \sum_{i=1}^{m} \beta_i \hat{H}(b_i) u_{b_i}$$

The norm of $\Gamma_{\vec{H}}$ is defined as

$$\|\Gamma_H\| = \sup_{f \in \mathcal{U}+\mathcal{V}, f \neq 0} \frac{\|\Gamma_H f\|_\infty}{\|f\|_\infty}$$

In the case where H is LTI, this norm can be computed exactly via solving a linear programming problem. This is captured in the following proposition.

Proposition 4.2.1 *Let H be an LTI operator with λ-transform $\hat{H}$. Then*

$$\|\Gamma_H\| = \max_{\alpha,\beta} \sum_{i=1}^{n} \alpha_i \hat{H}(a_i) + \sum_{i=1}^{m} \beta_i \hat{H}(b_i)$$

subject to

$$\left| \sum_{i=1}^{n} \alpha_i a_i^k + \sum_{i=1}^{m} \beta_i a_i^k \right| \leq 1, \quad \forall k \geq 0.$$

Proof. By direct computation,

$$\|\Gamma_H\| = \max_{\alpha,\beta} \max_{k \leq 0} \left| \sum_{i=1}^{n} \alpha_i \hat{H}(a_i) a_i^{-k} + \sum_{i=1}^{m} \beta_i \hat{H}(b_i) b_i^{-k} \right|$$

subject to

$$\left| \sum_{i=1}^{n} \alpha_i a_i^{-k} + \sum_{i-1}^{m} \beta_i b_i^{-k} \right| \leq 1, \forall k \leq 0.$$

It remains to be shown that the function to be maximized achieves the maximum at $k = 0$. To prove this, assume it achieves the maximum at $k = k^*$. Let $\tilde{\alpha}_i = \alpha_i a_i^{-k^*}$, $\tilde{\beta}_i = \beta_i b_i^{-k^*}$. Then $\tilde{\alpha}_i, \tilde{\beta}_i$ are feasible solutions which gives the same value at $k = 0$. ∎

The following theorem establishes the connection between ℓ^1-optimal control and the above Hankel-like operator.

Theorem 4.2.1 ([10])

$$\mu_{TI} = \|\Gamma_{T_1}\|.$$

It is interesting to notice that when Q is LTI, one can easily show that $||\Gamma_{T_1}||$ is a lower bound for μ_{TI} as follows. Let $\Phi = T_1 - T_2 Q T_3$ with impulse response $\phi \in \ell^1$. Then

$$
\begin{aligned}
||\Phi|| \;=\; ||T_1 - T_2 Q T_3|| &\geq \sup_{u \in \mathcal{U}, v \in \mathcal{V}, ||u + v||_\infty \leq 1} ||\Pi_-(u * \phi * e) + \Pi_-(e * \phi * v)||_\infty \\
&= ||\Gamma_{T_1}||_\infty
\end{aligned}
$$

This lower bound is not valid in general if $Q \in \mathcal{N}_{TV}$. This poses a serious problem in proving the general result we desire. In the special case where $T_2 = I$ (and hence $\mathcal{U} = 0$), the lower bound is valid and the desired result can be proved. Of course this includes the case where T_2^{-1} is stable.

Theorem 4.2.2 If $T_2 = I$ then $\mu_{NL} = \mu_{TI}$.

Proof. For all $v \in \mathcal{V}$ with $||v||_\infty \leq 1$,

$$
\begin{aligned}
||T_1 - Q T_3|| &\geq ||\Pi_-(T_1 - Q T_3)v||_\infty \\
&= ||\Pi_- T_1 v - \Pi_- Q \Pi_- T_3 v||_\infty = ||\Pi_- T_1 v||_\infty
\end{aligned}
$$

The above is true since $\hat{T}_3(b_i) = 0$. Hence,

$$
||T_1 - Q T_3|| \geq ||\Gamma_{T_1}|| = \mu_{TI}
$$

However, the lower bound is achieved by an LTI Q. ∎

While the conditions of Theorem 2.2 are not the most general, there are in fact some interesting problems in which T_2 has a stable inverse. Below are a few examples.

Example 4.2.1 Weighted input-sensitivity minimization for a stable plant. The map from the reference input to the input of the plant, with a controller in the feedback loop, is given by $S_i = W_1(I + KP)^{-1} W_2$. Incorporating the parameterization of all stabilizing controllers, S_i is given by

$$
S_i = W_1(I - QP)W_2
$$

Both W_1 and W_2 are assumed to have a stable inverse. The result above implies that nonlinear controllers will not offer any advantage in ℓ^1 tracking problems with stable

plants. The parallel result for output sensitivity is still open. ■

Example 4.2.2 Robust stability with coprime factor perturbations. There is
an important reason for considering this example. Even though the result we presented
earlier is for the square case, it is still valid for the non-square case, i.e for the case where
T_3 is a row vector. In this example we will sketch the proof of this result in this special
problem. The general non-square case follows in the same way.

Define the following class of plants:

$$\Omega = \{P | P = (N + \Delta_2)(M + \Delta_1)^{-1} \ and \ \|\Delta_i\| < 1\}$$

with Δ_i being ℓ^∞ bounded LTV operators and $P_0 = NM^{-1}$ satisfying the Bezout identity

$$\begin{pmatrix} \tilde{V} & \tilde{U} \\ \tilde{N} & \tilde{M} \end{pmatrix} \begin{pmatrix} M & -U \\ -N & V \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

A sufficient condition for robustly stabilizing the above family with any NLTV controller
is given by

$$\inf_{Q \in \mathcal{N}_{TV}} \|[\tilde{V} \quad \tilde{U}] + Q[\tilde{N} \quad \tilde{M}]\| \le 1$$

This condition is also necessary if the controllers are restricted to be linear, possibly
time-varying [5, 51]. The necessity of this condition for NLTV controllers is as follows.
First, the underlying notion of stability is finite-gain stability over c^0 rather than ℓ^∞.
Second, the operator Q is restricted to be continuous and have *pointwise fading-memory*
[52].

We note that the construction in [5, 51, 52] leads to a construction of admissible LTV
Δ_i such that either of the following conditions occurs. The first condition is that the
plant $(N + \Delta_2)(M + \Delta_1)^{-1}$ has an internal cancellation. That is, the operators $M + \Delta_1$
and $N + \Delta_2$ are no longer coprime. This corresponds to an admissible plant which is not
stabilizable. The second is that the admissible plant $(N + \Delta_2)(M + \Delta_1)^{-1}$ is stabilizable,
but not using the particular Q with the property $\|[\tilde{V} \quad \tilde{U}] + Q[\tilde{N} \quad \tilde{M}]\| > 1$.

It turns out that above infimization is achieved via a linear time-invariant Q. Define
the subspace $\mathcal{V}$ (inside $\ell^\infty \times \ell^\infty$) as follows:

$$\mathcal{V} = \{v = \begin{pmatrix} M \\ -N \end{pmatrix} x, \ x \in c^0, \ x(k) = 0 \quad \forall k \ge 1\}$$

Then for any $Q \in \mathcal{N}_{TV}$ and $v \in \mathcal{V}$, it is true that

$$\|[\tilde{V} \quad \tilde{U}] + Q[\tilde{N} \quad \tilde{M}]v\|_\infty = \|x\|_\infty \geq |x(0)|$$

Equivalently,

$$\|[\tilde{V} \quad \tilde{U}] + Q[\tilde{N} \quad \tilde{M}]\| \geq \sup_{v \in \mathcal{V}} \frac{|x(0)|}{\|v\|_\infty}$$

which was shown in [5] to be achieved by an LTI Q. The generalization to arbitrary T_3 follows in a similar fashion. ∎

So far, there does not exist a general result that proves or disproves the general case where T_2 does not have a stable inverse. In the sequel, a smaller lower bound on $\|T_1 - T_2 Q T_3\|$ is furnished. However, it is not evident that there exists a causal Q that achieves the bound.

Theorem 4.2.3 *Let $T_3 = I$. Then*

$$\mu_{NL} \geq \sup_{u \in \mathcal{U},\, u \neq 0} \frac{\|\Gamma_{T_1} u\|_1}{\|u\|_1}$$

Proof. By direct computation, with $u \in \mathcal{U}$,

$$
\begin{aligned}
\|T_1 - T_2 Q\| &\geq \|(T_1 - T_2 Q)f\|_\infty && \forall f \in \ell_\infty(Z_+),\ \|f\|_\infty \leq 1 \\
&\geq \|u * (T_1 - T_2 Q)f\|_\infty && \|u\|_1 \leq 1 \\
&\geq \|\Pi_-(u * (T_1 - T_2 Q)f)\|_\infty \\
&= \|\Pi_-(u * T_1 f)\|_\infty,
\end{aligned}
$$

This leads to

$$
\begin{aligned}
\|T_1 - T_2 Q\| &\geq \sup_{\|u\|_1 \leq 1, u \in \mathcal{U}} \sup_{\|f\|_\infty \leq 1} \|\Pi_-(u * T_1 f)\| \\
&= \sup_{\|u\|_1 \leq 1, u \in \mathcal{U}} \|\Gamma_{T_1} u\|_1
\end{aligned}
$$

∎

The interpretation of this bound is as follows. Fix any $f \in \ell_\infty(Z_+)$, then there exist a Q_f such that

$$\|(T_1 - T_2 Q_f)f\|_\infty = \sup_{\|u\|_1 \leq 1, u \in \mathcal{U}} \|\Pi_- u * T_1 f\|_\infty$$

This Q_f however may very well be a non-causal function of f, and hence does not qualify as a candidate solution for the original problem. A consequence of this theorem is that in the case of a fixed input minimization [12], nonlinear time varying compensation does not improve the performance. This is clear from the fact that the above lower bound is valid for each f regardless of Q and can be achieved with Q time-invariant [12].

4.2.2 Linearization

In this section, we show that the use *smooth* NLTV compensation instead of LTI compensation does not improve the achievable rejection of persistent bounded disturbances. The systems $T_{1,2,3}$ are now assumed to be multi-input/multi-output.

The smoothness condition in this context is in terms of the compensation being *linearizable*. The following definition is adapted from [60, Chapter 7].

Definition 4.2.1 *An operator $H \in \mathcal{N}_{TV}$ is **linearizable** if there exists an linear operator $H_L \in \mathcal{L}_{TV}$ so that*

$$\lim_{\alpha \to 0} \sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|Hf - H_L f\|_\infty}{\|f\|_\infty} = 0.$$

*In this case, H_L is called the **linearization** of H.*

The main result of this section is as follows:

Theorem 4.2.4 *Let μ_{NL} be defined as in (4.1) with the infimization being over all $Q \in \mathcal{N}_{TV}$ which are **linearizable**. Then $\mu_{TI} = \mu_{NL}$.*

Proof. Let $Q \in \mathcal{N}_{TV}$ be linearizable, and let

$$\|T_1 - T_2 Q T_3\| = \mu.$$

We will show that there exists a $\tilde{Q} \in \mathcal{L}_{TI}$ so that

$$\|T_1 - T_2 \tilde{Q} T_3\| \leq \mu.$$

60

Towards this end, let Q_L denote the linearization of Q. Then from Definition 4.2.1, given any $\varepsilon > 0$, there exists an $\alpha > 0$ such that

$$\sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|T_2 Q T_3 f - T_2 Q_L T_3 f\|}{\|f\|_\infty} \leq \varepsilon.$$

Then

$$
\begin{aligned}
\mu &\geq \sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|(T_1 - T_2 Q T_3)f\|}{\|f\|_\infty} \\
&= \sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|(T_1 - T_2 Q_L T_3)f - (T_2 Q T_3 - T_2 Q_L T_3)f\|}{\|f\|_\infty} \\
&\geq \sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|(T_1 - T_2 Q_L T_3)f\|}{\|f\|_\infty} - \sup_{\substack{\|f\|_\infty \leq \alpha \\ f \neq 0}} \frac{\|(T_2 Q T_3 - T_2 Q_L T_3)f\|}{\|f\|_\infty} \\
&\geq \|T_1 - T_2 Q_L T_3\| - \varepsilon.
\end{aligned}
$$

Since ε is arbitrary, it follows that

$$\|T_1 - T_2 Q_L T_3\| \leq \mu.$$

Upon applying Theorem 4.1.1, there exists a $\tilde{Q} \in \mathcal{L}_{TI}$ so that

$$\|T_1 - T_2 \tilde{Q} T_3\| \leq \mu.$$

■

The idea in the proof of Theorem 4.2.4 is first to show that LTV compensation gives the same performance as linearizable NLTV compensation. We then use the results from [51] to show that LTI compensation gives the same performance as linearizable NLTV compensation.

4.3   Concluding Remarks

Even for the problem of disturbance rejection, nonlinear controllers can offer some advantage as seen in the following example. Let z denote the unit delay operator. Let

$T_3 = I$, and let

$$T_1 - T_2Q = \begin{pmatrix} 1 \\ z \\ z^2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} Q.$$

Then for any $w \in \ell_\infty$,

$$((T_1 - T_2Q)w)(n) = \begin{pmatrix} w(n) \\ w(n-1) \\ w(n-2) \end{pmatrix} - \begin{pmatrix} (Qw)(n) \\ (Qw)(n) \\ (Qw)(n) \end{pmatrix}.$$

Given this structure, an optimal Q may be constructed as follows. Define

$$\overline{w}(n) \overset{\text{def}}{=} \max\left(w(n), w(n-1), w(n-2)\right),$$

$$\underline{w}(n) \overset{\text{def}}{=} \min\left(w(n), w(n-1), w(n-2)\right).$$

Then set

$$(Q_{NL}w)(n) = \frac{\overline{w}(n) + \underline{w}(n)}{2}.$$

It can be shown that this selection of Q_{NL} leads to $\|T_1 - T_2Q_{NL}\| = 1$. This choice of Q is nonlinear. However, the same norm can be achieved by using the *linear* $Q = 0$. Nevertheless, the compensator Q_{NL} achieves *better* performance in the sense that *signal-by-signal*, the response using Q_{NL} is smaller than using $Q = 0$. That the two choices lead to the same norm means there exists a signal so that the responses are the same size. Note that the choice of Q_{NL} is *not* differentiable. Thus, the performance is not characterized by the small signal behavior.

A comment is in order regarding the use of induced norms to assess the performance of nonlinear feedback systems. For linearizable systems, the overall performance is at best the "small-signal" performance. Thus, it seem natural that linear controllers would perform as well as linearizable nonlinear controllers.

It turns out that the use of induced norms to assess performance may be too restrictive in the presence of nonlinear compensation. The reason is that this definition requires the *ratio* of the error-norm to the disturbance-norm to be small without regard to the size of the magnitude of the disturbances. More precisely, it may be that the regulated variable is small while the ratio of error-norm to disturbance-norm is large. This leads to questioning the utility of induced norms to quantify performance in nonlinear systems.

One alternative is to consider the worst case performance over a given class of disturbances. For example, let $\mathcal{W}$ denote some bounded set of disturbances. Then define the performance measure

$$\mu \overset{\text{def}}{=} \sup_{w \in \mathcal{W}} \|T_{zw} w\|.$$

Such a performance objective is particularly well-suited to nonlinear systems. It avoids using induced norms and addresses directly the desired goal of keeping regulated variables small. Furthermore, it allows the class of disturbances to be defined as desired. For example, one may define $\mathcal{W}$ as

$$\mathcal{W} = \left\{ w \in \ell_\infty : \|w\| \leq c_1 \text{ and } \sum_n |w(n)|^2 \leq c_2 \right\}.$$

This definition allows both a magnitude and energy bound on the disturbances of interest. Such notions of performance have been considered in [48, 49].

Acknowledgments The authors thank Paul Middleton for suggesting the preceding Q_{NL} example.

Chapter 5

On Slowly Varying Systems: ℓ_∞ to ℓ_∞ Performance and Implications to Robust Adaptive Control

In this chapter we present a result on the ℓ_∞ to ℓ_∞ performance of slowly time varying systems. In particular we show that the performance of such systems cannot be much worse than that of the frozen-time systems which are time invariant. This result is used to characterize a class of indirect adaptive controllers that can stabilize a time invariant system subjected to both parametric and ℓ_∞ to ℓ_∞ bounded unstructured uncertainty. Pertaining to this class of controllers, a particular r indirect adaptive scheme is proposed that provides the greatest upper bound on the size of the unstructured uncertainty for which stability is ensured.

5.1 Introduction

The problem of controlling slowly time-varying systems arises in many applications. The main paradigm is in gain-scheduling where the plant is time-varying and at successive points in time a controller is designed to satisfy certain stability and performance specifications based on the "frozen-time" system which is time invariant (LTI). Therefore, the resulting controller is itself time-varying. However, it is expected that if the rate of time variation is small enough then the frozen-time properties carry on to the overall time-

varying system. In other words, it is expected that the stability of "frozen-time" designs will guarantee stability of the global time-varying system and also that the performance of the global system cannot be considerably worse than that of the frozen designs. As a matter of fact, these expectations have not only been confirmed in practice but also in theory by the work of several researchers in this area for example [6, 18, 27, 37, 50, 63, 64].

In this chapter we continue the work of [6] that was centered at the stability issue and extend it to capture the performance part in a bounded input to bounded output (i.e. ℓ_∞ to ℓ_∞) sense. We use the input-output framework of [6] that allows infinite-dimensional plants and controllers. Hence, the need of a fixed degree is not apparent. The main result of this chapter is given for single-input single-output (SISO) discrete slowly varying systems. It states that the ℓ_∞ to ℓ_∞ performance of the global time varying system cannot be much worse than the worst frozen-time ℓ_∞ to ℓ_∞ performance given that the rates of variation of the plant and the controller are sufficiently small. Moreover, given the continuity properties of the optimal ℓ_1 design established in [7] it follows that optimal ℓ_1 [10, 11] frozen-time design yields an upper bound on the ℓ_∞ to ℓ_∞ p performance of the global system. Our main result is in parallel with these in [63, 64] however our derivation is more direct and suited to the ℓ_∞ to ℓ_∞ disturbance rejection.

An important application of our main result is in robust adaptive control. In particular we characterize a class of indirect adaptive controllers that can stabilize systems that contain both parametric and unstructured uncertainty. The unstructured uncertainty enters the system in the form of bounded-input, bounded-output operators perturbing the coprime factors of the plant. This class of stabilizing controllers is obtained by frozen-time controllers that stabilize the estimated models at each time of the plant. The estimated model is obtained via a parameter estimation algorithm which produces slowly varying estimates. The conclusion is that if the frozen time designs stabilize the estimated model together with the unstructured uncertainty (possibly after some initial transient period) then stability of the adaptive scheme is guaranteed. This result is similar to the one reported in [41] where the auhtors use a continuous time framework and a different characterization of the size of the uncertainty. Finally, among this class of controllers we present a particular adaptive scheme that requires for stability the least

65

conservative bound on the ℓ_∞ to ℓ_∞ "gain" of the unstructured perturbations. This scheme is a modification of the ℓ_1 adaptive scheme found in [7].

5.2 Preliminary Definitions

In this chapter the following notation is used:

$|x|_2$: The Euclidian norm of the finite dimensional real vector x.

$\overline{\sigma}[A]$: The maximum singular value of the matrix A.

$\ell_{1m\times n}$: The normed linear space of all $m \times n$ matrices H each of whose entries is a right sided, absolutely summable real sequence $H_{ij} = (H_{ij}(k))_{k=0}^\infty$. The norm is defined as:

$$\|H\|_{\ell_{1m\times n}} = \max_i \sum_{j=1}^n \sum_{k=0}^\infty |H_{ij}(k)|$$

$\mathcal{A}^*_{m\times n}$: The normed linear space of all $m \times n$ matrices H each of whose entries is a right sided, magnitude bounded real sequence $H_{ij} = (H_{ij}(k))_{k=0}^\infty$. The norm is defined as:

$$\|H\|_{\mathcal{A}^*_{m\times n}} = \sum_{i=1}^m \max_j (\sup_k |H_{ij}(k)|)$$

$c^0_{m\times n}$: The subspace of $\mathcal{A}^*_{m\times n}$ consisting of all elements which converge to zero.

ℓ_m^∞ : The space of real $m \times 1$ vectors u each of whose components is a magnitude bounded real sequence $(u_i(k))_{k=0}^\infty$. The norm is defined as:

$$\|u\|_{\ell_m^\infty} = \max_i (\sup_k |u_i(k)|)$$

$\ell_m^{\infty,e}$: The space of real $m \times 1$ vector valued sequences.

$\hat{H}(\lambda)$: The λ-transform of a right sided $m \times n$ real sequence $H = (H(k))_{k=0}^\infty$ defined as:

$$\hat{H}(\lambda) = \sum_{k=0}^\infty H(k)\lambda^k$$

$\mathcal{A}_{m\times n}$: The real normed linear space of all $m \times n$ matrices $\hat{H}(\lambda)$ so that $\hat{H}(\lambda)$ is the

λ-transform of an $\ell_{1m \times n}$ sequence H. This space is isometrically isomorphic to $\ell_{1m \times n}$.

$\mathcal{L}_{TV}^{m \times n}$: The space of all linear bounded and causal maps from $\ell_{\infty n}$ to $\ell_{\infty m}$. We refer to these operators as stable.

$\mathcal{L}_{TI}^{m \times n}$: The subspace of $\mathcal{L}_{TV}^{m \times n}$ consisting of the maps that commute with the shift operator (i.e. the time invariant maps). This space is isometrically isomorphic to $\mathcal{A}_{m \times n}$.

Π_m^k : The kth-truncation operator on $\ell_m^{\infty,e}$ defined as:

$$\Pi_m^k : \{u(0), u(1), \ldots\} \longrightarrow \{u(0), \ldots, u(k), 0, 0, \ldots\}$$

Λ_m : The right shift operator on $\ell_m^{\infty,e}$ i.e.

$$\Lambda_m : \{a(0), a(1), \ldots\} \longrightarrow \{0, a(0), a(1), \ldots\}$$

Note: We will often drop the m and n in the above notation when the dimension is not important or when it is clear from the context. Also, subscripts on the norms are dropped when there is no ambiguity.

Let T be an operator in $\mathcal{L}_{TI}$ with transform representation

$$\hat{T}(\lambda) = \sum_{i=0}^{\infty} T(i)\lambda^i.$$

Definition 5.2.1 *The Integral Time Absolute Error ITAE associated with T is defined as*

$$ITAE(T) = \sum_{k=0}^{\infty} k|T(k)|.$$

If T' is the LTI operator associated with the derivative $\hat{T}'(\lambda) = \frac{d\hat{T}(\lambda)}{d\lambda}$ then it follows that

$$ITAE(T') = \|T'\|.$$

Given a sequence of LTI operators $\{A_t\}_{t=0}^{\infty}$ where each A_t is a map from $\ell^{\infty,e}$ to $\ell^{\infty,e}$ we can generate a time varying operator A as $(Ay)(t) = (A_t y)(t)$, $t = 0, 1, \ldots$.

Definition 5.2.2 *The operator A is called slowly time-varying if there is a constant $\gamma > 0$ so that*

$$\|A_t - A_\tau\| \leq \gamma |t - \tau| \quad \forall t, \tau.$$

This is denoted by $A_t \in STV(\gamma)$.

If $A_t \in \mathcal{L}_{TI}$ for all t and also the $\mathcal{L}_{TI}$ norm is bounded uniformly in t then $A \in \mathcal{L}_{TV}$ and $norm A = \sup_t \|Pi^t A_t\|$.

5.3 Problem Definition

The problem we want to analyze is the stability and performance of the feedback system in Figure 5.1 where P is a slowly varying plant and C is a controller obtained by "frozen-time" control. Specifically, the plant P is defined as $P = A^{-1}B$ where A, B are slowly varying operators associated with the sequences $\{A_t\}, \{B_t\}$, $t = 0, 1, 2, \ldots$ of LTI stable operators respectively and with A^{-1} being well defined. Hence, the plant model is

$$y(t) = (Pu)(t) = (A^{-1}Bu)(t), \quad t = 0, 1, 2, \ldots$$

or, equivalently,

$$(A_t y)(t) = (B_t u)(t), \quad t = 0, 1, 2, \ldots .$$

We refer to the LTI system $P_t = A_t^{-1}B_t$ as the "frozen-time" plant. The controller is given as $C = L^{-1}M$ where L, M are associated with the sequences $\{L_t\}, \{M_t\}$, $t = 0, 1, 2, \ldots$ of LTI stable operators i.e., $(Ly)(t) = (L_t y)(t)$ and $(My)(t) = (M_t y)(t)$. Moreover, L_t, M_t are so that the LTI controller defined as $C_t = L_t^{-1}M_t$ stabilizes the frozen time plant P_t. The controller operates as

$$(Cy)(t) = (L^{-1}My)(t), \quad t = 0, 1, 2, \ldots$$

or, equivalently,

$$(L_t u)(t) = (M_t y)(t), \quad t = 0, 1, 2, \ldots .$$

The question we want to answer is under what conditions the feedback loop is stable and, if so, what is the relation between the performance of the frozen-time pair (P_t, C_t) and the actual time varying feedback pair (P, C). This is done in the following section.
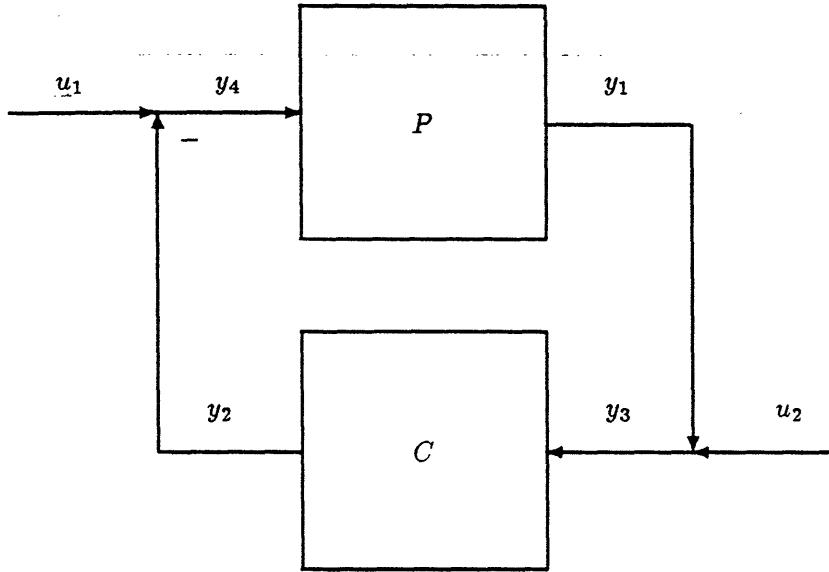
Figure 5.1: Feedback Loop of the Pair (P, C)

5.4 Main Result

In [6] an input-output point of view was taken to prove that, under the assumption of sufficiently small rate of variation, stability of the frozen-time feedback pair (P_t, C_t) implies stability of the pair (P, C). Yet, the performance part of the problem was not investigated. In the sequel we take the same point of view as in [6] and extend the results in [6] to capture the performance issue. In particular, for the system in Figure 5.1 define the stable LTI operator for each $t = 0, 1, 2, \ldots$

$$G_t = L_t A_t + M_t B_t.$$

Since C_t stabilizes P_t then $H_t = G_t^{-1} \in \mathcal{L}_{TI}$. Now let S^{ij} represent the map from u_j to y_i in the system of Figure 5.1 and S_t^{ij} the (LTI) map from u_j to y_i for the frozen system (P_t, C_t). The following theorem which is an extension of Theorem 1 in [6] supplies the answer to our problem.

Theorem 5.4.1 *Assume the following:*

69

1. The operators defining the plant P are slowly time-varying with rates γ_A, γ_B i.e., $A_t \in STV(\gamma_A), B_t \in STV(\gamma_B)$.

2. The operators defining the controller C are slowly time-varying with rates γ_L, γ_M i.e., $L_t \in STV(\gamma_L), M_t \in STV(\gamma_M)$.

3. The $\mathcal{L}_{TI}$ norms and the ITAE of the operators A_t, B_t, L_t, M_t are uniformly bounded in t; this of course means that $A, B, L, M, \in \mathcal{L}_{TV}$

4. The $\mathcal{L}_{TI}$ norms and the ITAE of the operator $H_t = G_t^{-1}$ are uniformly bounded in t.

Then, for a given $\epsilon > 0$, there exists a nonzero constant γ so that, if $\gamma_A, \gamma_B, \gamma_L, \gamma_M \leq \gamma$, the closed loop system is internally stable and

$$(1 - \epsilon)\|S^{ij}\| \leq \sup_t \|S_t^{ij}\| + \epsilon.$$

Proof. The proof of the stability part is given in [6]. Here we repeat in brief the main steps because we will use them to prove the claim for the performance. The closed loop equations for the system in Figure 5.1 are as follows:

$$(A_t y_1)(t) = (B_t(u_1 - y_2))(t)$$

$$(L_t y_2)(t) = (M_t(u_2 + y_1))(t)$$

$$A_t L_t + M_t B_t = G_t$$

By adding subtracting and grouping terms we finally arrive [6] at

$$\begin{pmatrix} G + X & Y \\ -Z & G + W \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} LB & -BM \\ MB & AM \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{CL}$$

where G is the operator in $\mathcal{L}_{TV}$ associated with the family $\{G_t\}$ and X, Y, Z, W are "perturbation" operators which are due to the time variation of the system P. As indicated in lemmas 1 and 2 in [6] these operators have $\mathcal{L}_{TV}$ norm bounded by the term $\gamma \times$ constant where $\gamma = \max(\gamma_A, \gamma_B, \gamma_L, \gamma_M)$ and the constant depends on the uniform

bounds of assumption 3 of the Theorem 5.4.1; i.e., there are constants $c_X, c_Y, c_Z, c_W > 0$ so that

$$\|X\| \le \gamma c_X, \quad \|Y\| \le \gamma c_Y, \quad \|Z\| \le \gamma c_Z, \quad \|W\| \le \gamma c_W$$

Now, from the first equation in (CL) we have

$$G y_1 + X y_1 + Y y_2 = v$$

where $v = LBu_1 - BMu_2$. If we fix some t, G_t is a LTI operator; adding and subtracting this operator in the above operator equation we obtain

$$G_t y_1 + (G - G_t) y_1 + X y_1 + Y y_2 = v$$

or since $H_t = G_t^{-1}$ we obtain

$$y_1 + H_t (G - G_t) y_1 + H_t X y_1 + H_t Y y_2 = H_t v.$$

Evaluating this operator equation at time t we obtain

$$y_1(t) + (H_t(G - G_t) y_1)(t) + (H_t X y_1)(t) + (H_t Y y_2)(t) = (H_t v)(t).$$

Define the operator H as $(Hz)(\tau) = (H_\tau z)(\tau)$, $\tau = 0, 1, 2, \ldots$. Also define the operator R as $(R y_1)(\tau) = (H_\tau (G - G_\tau) y_1)(\tau)$, $\tau = 0, 1, 2, \ldots$. Rewritting the above equation in operator form we have

$$y_1 + R y_1 + H X y_1 + H Y y_2 = H v.$$

Similarly working with the second equation, letting $w = MBu_1 + AMu_2$ and putting both equations together in operator form we get

$$(I + F) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} Hv \\ Hw \end{pmatrix}$$

where

$$F = \begin{pmatrix} R + HX & HY \\ -HZ & R + HW \end{pmatrix}.$$

Note that from the uniform bound assumption on H_τ it follows that $H \in \mathcal{L}_{TV}$ and therefore the norms of the operators HX, HZ, HY, HW can be bounded by $\gamma \times$ constant.

Also, utilizing the fact that the ITAE of H_τ is uniformly bounded it is shown in [6] that the norm of R is bounded in the same way i.e., $\|R\| \le \gamma \times$ constant. The stability of the loop then, follows from the small gain theorem for sufficiently small γ.

We now come to the performance part. We will prove our claim for the maps S^{12}, S^{22}; the proof for any other map is completely analogous. Let $u_1 = 0$ and let $\|u_2\| \le 1$. Then from the system equations we get

$$y_1(t) = -(H_t BM u_2)(t) - (H_t X y_1)(t) - (H_t Y y_2)(t) - (H_t(G - G_t)y_1)(t)$$

Consider now the frozen LTI feedback system at time t i.e., (P_t, C_t) subjected to the same input u_2 and let y_{1t} denote the output that corresponds to y_1 in the time varying loop (P, C). Then evaluating y_{1t} at t we have

$$y_{1t}(t) = -(H_t B_t M_t u_2)(t).$$

Subtracting the above two equations we obtain

$$y_{1t}(t) - y_1(t) = (H_t(BM - B_t M_t)u_2)(t) + (H_t X y_1)(t) + (H_t Y y_2)(t) + (H_t(G - G_t)y_1)(t).$$

The idea here is to bound $|(H_t(BM - B_t M_t)u_2)(t)|$ by $\gamma \times$ constant. For this purpose define the operator $K \in \mathcal{L}_{TV}$ as

$$(Kz)(\tau) = (B_\tau M_\tau z)(\tau) \quad \tau = 0, 1, 2, \dots$$

then

$$(H_t(BM - B_t M_t)u_2)(t) = (H_t(BM - K)u_2)(t) + (H_t(K - B_t M_t)u_2)(t).$$

By lemma 1 in [6] and the fact that H_t has norm uniformly bounded it follows that

$$|(H_t(BM - K)u_2)(t)| \le \gamma c_1$$

with c_1 a positive constant. For the term $(H_t(K - B_t M_t)u_2)(t)$ we have the following:

$$\begin{aligned}
\|B_\tau M_\tau - B_t M_t\| &\le \|B_\tau\|\|M_\tau - M_t\| + \|M_t\|\|B_t - B_\tau\| \\
&\le \|B_\tau\|\gamma_M|t - \tau| + \|M_t\|\gamma_B|t - \tau|
\end{aligned}$$

72

$$\leq \quad \gamma c_2 |t - \tau|.$$

Hence, if $z(\tau) = ((K - B_t M_t) u_2)(\tau)$, then $|z(\tau)| \leq \gamma c_2 |t - \tau|, \quad \tau = 0, 1, 2, \ldots$ with $c_2 > 0$. But then from the fact that H_t has bounded (uniformly in t) ITAE it follows as in theorem 1 of [6] that

$$
\begin{aligned}
|(H_t (K - B_t M_t) u_2)(t)| &= |\sum_{\tau=0}^{\tau=t} h_t(t - \tau) z(\tau)| \\
&\leq \gamma c_2 \sum_{\tau=0}^{\tau=t} |h_t(\tau)| \tau \\
&\leq \gamma c_3, \quad c_3 > 0.
\end{aligned}
$$

Now, looking at the rest of the terms and since $\|u_2\| \leq 1$ we have $|(H_t X y_1)(t)| \leq \gamma c_4 \|S^{12}\|$, $|(H_t Y y_2)(t)| \leq \gamma c_5 \|S^{22}\|$ and $|(H_t (G - G_t) y_1)(t)| \leq \gamma c_6 \|S^{12}\|$ so putting everything together it follows that there are $c, c_{12}, c_{22} > 0$ so that

$$|y_1(t) - y_{1t}(t)| \leq \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\|$$

or since $\|u_2\| \leq 1$ then $|y_{1t}(t)| \leq \|S_t^{12}\|$ and therefore

$$\sup_t |y_1(t)| \leq \sup_t \|S_t^{12}\| + \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\|$$

and since u_2 is arbitrary

$$\|S^{12}\| \leq \sup_t \|S_t^{12}\| + \gamma c + \gamma c_{12} \|S^{12}\| + \gamma c_{22} \|S^{22}\|.$$

Similarly working for $\|S^{22}\|$ we get

$$\|S^{22}\| \leq \sup_t \|S_t^{22}\| + \gamma k + \gamma k_{22} \|S^{22}\| + \gamma k_{12} \|S^{12}\|.$$

Now noting that $\|H_t\|$ is uniformly bounded then $\sup_t \|S_t^{12}\|, \sup_t \|S_t^{12}\| < \infty$ and hence by assumming γ sufficiently small the proof of the theorem is complete. ∎

The above theorem, roughly speaking, indicates that if the rates of variation of the plant and the controller are sufficiently small then frozen time control would not only provide stability but also the resulting performance cannot be much worse than the

worst frozen time design. In [7] it was shown under certain assumptions of existence and uniqueness that the ℓ_1 design methodology produces optimal frozen-time LTI controllers for the frozen-time plant that possess the slow variation property given that the plant is slowly varying. Hence an upper bound on the achievable $\|S^{ij}\|$ can be obtained by considering ℓ_1-optimal [10, 11] frozen time control i.e., by considering $\sup_t \|S^{ij}_t\|$ obtained by ℓ_1 optimal designs.

Remark

A natural question that arises in the case where the plant P is slowly time varying is whether optimal frozen-time design at each time t will result in an optimal or near-optimal design (depending on the rate of variation) for the time-varying system. Although it is tempting to conjecture that, if the rate of variation is sufficiently small then the optimal performance cannot be far from the performance provided by optimal frozen-time control at each time t, the following example shows that this might not be true: Consider the plant $P \in \mathcal{L}_{TV}$ defined by the sequences $\{A_t\}, \{B_t\}$ where $\hat{A}_t(\lambda) = 1\, \forall t$, $\hat{B}_t(\lambda) = 2\lambda + 1$ for $t = 0, 1$ and $\hat{B}_t(\lambda) = 2\lambda + (1 + \gamma t)$ for $2 \leq t \leq T = [1/\gamma + 1]$, $\hat{B}_t(\lambda) = 2\lambda + (1 + \gamma T)$ for $t > T$, with $\gamma > 0$. The resulting Toeplitz representation of P is

$$
P = \begin{pmatrix}
1 & & & & & & \\
2 & 1 & & & & & \\
 & 2 & 1+\gamma & & & & \\
 & & 2 & 1+2\gamma & & & \\
 & & & \ddots & & \ddots & \\
 & & & & 2 & 1+T\gamma & \\
 & & & & & 2 & 1+T\gamma \\
 & & & & & & \ddots & \ddots
\end{pmatrix}
$$

Clearly, for this P we have $\|B_t - B_\tau\| \leq \gamma |t - \tau| \; \forall t, \tau$. Suppose we are interested in minimizing the $\mathcal{L}_{TV}$ norm of the sensitivity map $S = (1 + PC)^{-1}$. Then as it is well known [19, 57, 61] the optimization problem transforms to

$$
\inf_{Q \in \mathcal{L}_{TV}} \|1 - PQ\|.
$$

74

For this particular P we have that $P^{-1} \in \mathcal{L}_{TV}$ since $P_t = B_t$ is eventually $(t > T)$ stably invertible. To view this, let P^{-1} be represented by the lower diagonal structure

$$P^{-1} = \begin{pmatrix} q(0,0) & 0 & \cdots \\ q(1,0) & q(1,1) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Then the following recursion holds:

$$q(i,j) = -2q(i,i)q(i-1,j), \; i > 0, \; j = 0,1,\ldots,i-1$$

with $q(0,0) = 1$, $q(i,i) = 1/(1 + (i-1)\gamma)$ for $i = 1,\ldots,T$ and $q(i,i) = 1/(1 + T\gamma)$ for $i \geq T$. Note that for $i \geq T$ we have that $|q(i,i)| = |q(T,T)| < 1$. Therefore, for any $k = 1,2,\ldots$ we have

$$\sum_{j=0}^{T+k} |q(T+k,j)| \; \leq (\max_{0 \leq j \leq T-1} |q(T-1,j)|)|q(T,T)|^k + \sum_{j=1}^{k} |q(T,T)|^j$$

$$\leq c_1 |q(T,T)| + \frac{1}{1 - |q(T,T)|}.$$

This evidently shows that $P^{-1} \in \mathcal{L}_{TV}$. Hence, by choosing $Q_o = P^{-1}$ we can make $\|1 - PQ_o\| = 0$ for any γ. On the other hand, using ℓ_1 optimal frozen time design yields [10, 11] $S_{t=0} = 1, S_{t=1} = 1, \ldots, S_{t=T} = 0, S_{t=T+1} = 0, \ldots$. The reason for $S_{t=0}, \ldots S_{t=T-1} \neq 0$ is of course the unstable zero of $P_{t=0}, \ldots, P_{t=T}$ at $\lambda = (1 + t\gamma)/2$ for all γ. Moreover, the resulting frozen time based controller will yield a performance $\|S\| \geq 1$ for any $\gamma > 0$ no matter how small, since the system will behave exactly as the frozen LTI one for $t = 0, 1$.

5.5 Application to Robust Adaptive Control

In this section we utilize the main result of the previous section in order to design a controller for a LTI system which contains both parametric (structured) and unstructured uncertainty.

5.5.1 Problem statement

The problem we want to resolve is as follows:

We are given the single-input, single-output discrete system

$$((A_o + \Delta_A)y)(t) = ((B_o + \Delta_B)u)(t) + d(t)$$

75

where A_o, B_o are operators in $\mathcal{L}_{TI}$ with a polynomial λ-transform representation

$$\hat{A}_o(\lambda) = 1 + a_o(1)\lambda + a_o(2)\lambda^2 + \ldots + a_o(m_1)\lambda^{m_1},$$

$$\hat{B}_o(\lambda) = b_o(1)\lambda + b_o(2)\lambda^2 + \ldots + b_o(m_2)\lambda^{m_2},$$

with the coefficients of $\hat{A}_o(\lambda)$ and $\hat{B}_o(\lambda)$ not known a priori; Δ_A and Δ_B are unknown, possibly time-varying, operators in $\mathcal{L}_{TV}$ i.e., $\|\Delta_A\|, \|\Delta_B\| < \infty$; finally, d is a bounded disturbance i.e., $|d(t)| < D, \forall t = 0, 1, 2, \ldots$ for some $D > 0$. We assume the following a priori knowledge

Assumption 5.5.1 *The integer $n = \max(m_1, m_2)$ is known.*

Assumption 5.5.2 *The coefficients of $\hat{A}_o(\lambda), \hat{B}_o(\lambda)$ lie in a compact convex set Θ which is known. Moreover, the above polynomials are coprime (no common zeros) for all possible values of their coefficients.*

Assumption 5.5.3 *The bound D and some bound D_Δ so that $\|(\Delta_B \; \Delta_A)\| \leq D_\Delta$ are known.*

Our task is to find a controller C that stabilizes the system in the presence of the bounded disturbance d. The situation is depicted in Figure 5.2 where $P_o = A_o^{-1} B_o$.

5.5.2 An Indirect Control Scheme

The system equations can be rewritten as

$$y(t) = \phi(t-1)^T \theta_o + d(t) + \left((\Delta_B \; \Delta_A) \begin{pmatrix} u \\ y \end{pmatrix}\right)(t)$$

where

$$\theta_o = (-a_o(1) \ldots - a_o(n) \; b_o(1) \ldots b_o(n))^T$$

$$\phi(t-1)^T = (y(t-1) \ldots y(t-n) \; u(t-1) \ldots u(t-n)).$$

The approach we will use to design the controller is an indirect adaptive scheme [30] which is a generalization of the one in [7] to include unstructured uncertainty. In particular, we will use a parameter estimation scheme to supply at each time t estimates

76
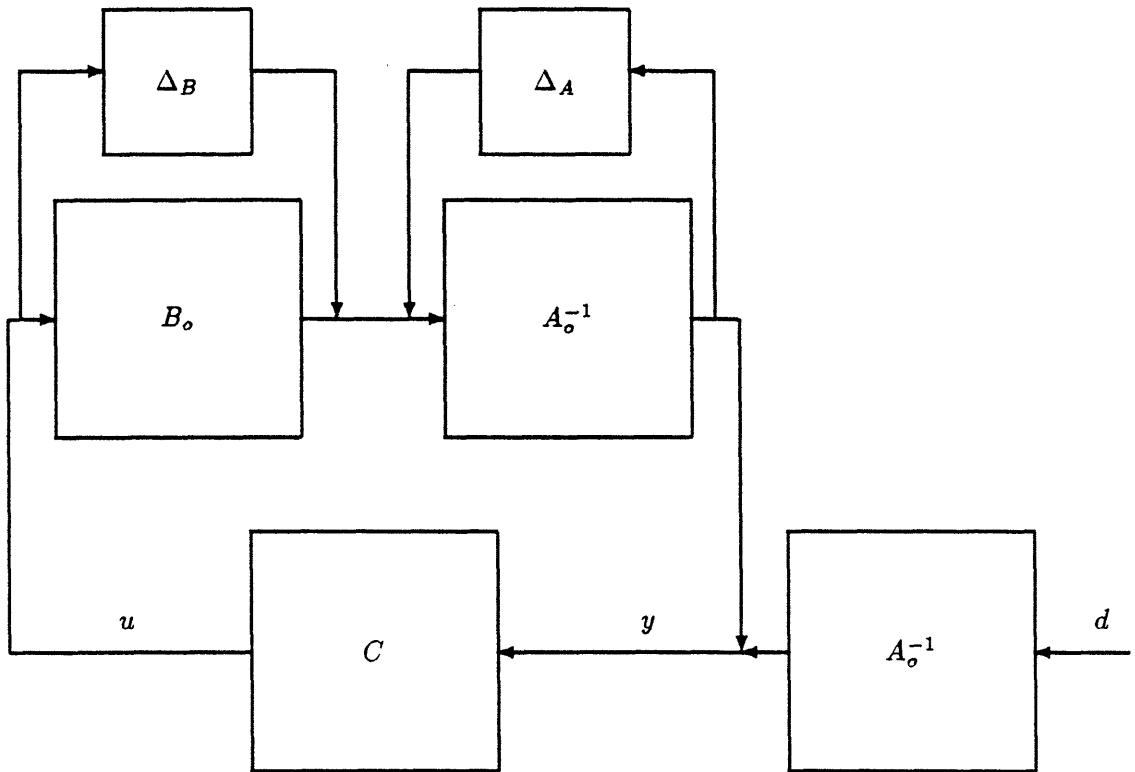
Figure 5.2: The Feedback Loop of (P_o, C) with Unstructured Uncertainty

θ_t for θ_o. The controller C will be designed based on frozen designs C_t so that C_t stabilizes the estimated system at time t; the properties of slow variation of the estimates produced by the parameter estimation scheme together with the main result of the previous section will guarantee stability.

More specifically, let $\delta(t) = ((\Delta_B\ \Delta_A) \begin{pmatrix} u \\ y \end{pmatrix})(t)$ then the equation for the model is

$$y(t) = \phi(t-1)^T\theta_o + d(t) + \delta(t)$$

with $|\delta(t)| \le D^1(t)$ and $|d(t)| < D$, $\forall t = 0, 1, 2, \ldots$ where

$$D^1(t) = D_\Delta \max_{0 \le \tau \le t}(|u(\tau)|, |y(\tau)|).$$

The parameter estimation scheme to be used is a robustified least squares algorithm with dead-zone found in [42] which is a modification of the one in [29]. Define for each estimate θ_t the error signal

$$e(t) = y(t) - \phi(t-1)^T\theta_{t-1}$$

then the algorithm is as follows

$$\theta_t = \theta_{t-1} + \frac{v(t)P(t-2)\phi(t-1)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)}e(t)$$

with

$$P(t-1) = P(t-2) - \frac{v(t)P(t-2)\phi(t-1)\phi(t-1)^T P(t-2)}{1 + \phi(t-1)^T P(t-2)\phi(t-1)}$$

where θ_0 and $P(-1)$ are initial guesses with $P(-1) = P(-1)^T > 0$, and where $v(t) = \alpha s(t)$ with

$$s(t) = f(\beta(D^1(t) + D), e(t))/e(t),$$

where we choose $\alpha \in (0, 1)$, β is defined by $\beta = \sqrt{1/(1-\alpha)}$ and $f(\cdot, \cdot)$ is the dead-zone function

$$f(x, y) = \begin{cases} |y| - |x|, & \text{if } |x| < |y| \\ 0, & \text{otherwise} \end{cases}.$$

The full set of details of the algorithm can be found in [42]. The properties of the algorithm that will be used for stability of the adaptive scheme are

1.

$$\lim_{t \to \infty} \frac{f^2(\beta(D^1(t) + D), |e(t)|)}{1 + \phi^T(t-1)P(t-2)\phi(t-1)} = 0$$

2.

$$\lim_{t \to \infty} |\theta_t - \theta_{t-1}|_2 = 0$$

3.

$$P(t) > 0, \ \overline{\sigma}[P(t)] \leq \overline{\sigma}[P(-1)] < \infty, \ \forall t$$

We should note that constraining the estimates θ_t to lie in Θ as in [30] does not change properties 1,2,3 of the algorithm. The parameter β in the estimation scheme will be taken close to 1 i.e., $1 < \beta < 1 + \eta$, where η is sufficiently small (to be established in the sequel) and positive. The following generalized "key technical lemma" [7, 30] gives conditions for stability of the adaptive scheme.

Lemma 5.5.1 *Assume the following*

1. *there are constants $c_1 \geq 0, c_2 > 0$ and some time instant T_1 so that for all $t \geq T_1$*

$$|\phi(t)|_2 \leq c_1 + c_2 \max_{T_1 \leq \tau \leq t} |e(t)|,$$

2.

$$\lim_{t \to \infty} \frac{f^2(\beta(D^1(t) + D), |e(t)|)}{1 + \phi^T(t-1)P(t-2)\phi(t-1)} = 0,$$

3. *there are constants $k_1 \geq 0, k_2 > 0$ with $k_2 < \frac{1}{\beta}$ and some time T_2 so that for all $t \geq T_2$*

$$D^1(t) \leq k_1 + k_2 \max_{T_2 \leq \tau \leq t} |e(\tau)|.$$

Then the sequence $\{e(t)\}$ is bounded and, therefore, $\{y(t)\}, \{u(t)\}$ are bounded.

Proof. Assume $\{e(t)\}$ is unbounded and let the subsequence $\{e(t_n)\}$ be so that $\lim_{n \to \infty} |e(t_n)| = \infty$ with $|e(t_0)| < |e(t_1)| < \dots$. Then there is some n_0 so that $\forall n > n_0$

$$D^1(t_n) \leq k_1 + k_2 |e(t_n)|$$

$$|\phi(t_n)|_2 \le c_1 + c_2|e(t_n)|$$

$$f(\beta(D^1(t_n) + D), e(t_n)) = |e(t_n)| - \beta(D^1(t_n) + D) > 0.$$

But then if

$$a_n = \frac{f^2(\beta(D^1(t_n) + D), |e(t_n)|)}{1 + \phi^T(t_n - 1)P(t_n - 2)\phi(t_n - 1)}$$

we have

$$a_n \ge \frac{(|e(t_n)| - \beta(D^1(t_n) + D))^2}{1 + |\phi(t_n - 1)|_2^2 \overline{\sigma}[P(-1)]}$$

therefore

$$\limsup_{n \to \infty} a_n \ge \frac{(1 - \beta k_2)^2}{c_2^2 \overline{\sigma}[P(-1)]} > 0$$

which contradicts assumption 2. ■

The above lemma guarantees boundedness of the signals provided that the three assumptions hold. Hence, if the control signal u is such that the assumptions of Lemma 5.5.1 are satisfied, then the above weak form of stability [7] of the system is obtained. In the sequel we show that under certain conditions, using frozen time control for the estimated system at each time t generates a control sequence u that satisfies the assumptions of Lemma 5.5.1 and hence weak-stability is guaranteed. This is done as follows: Rewriting the equation for the error e we obtain

$$(A_{t-1}y)(t) = (B_{t-1}u)(t) + e(t)$$

where $A_{t-1}, B_{t-1} \in \mathcal{L}_{TI}$ are defined by the estimate $\theta_{t-1} = (-a_{t-1}(1) \ldots -a_{t-1}(n) \, b_{t-1}(1) \ldots b_{t-1}(n))^T$ as

$$\hat{A}_{t-1}(\lambda) = 1 + a_{t-1}(1)\lambda + a_{t-1}(2)\lambda^2 + \ldots + a_{t-1}(n)\lambda^n,$$

$$\hat{B}_{t-1}(\lambda) = b_{t-1}(1)\lambda + b_{t-1}(2)\lambda^2 + \ldots + b_{t-1}(n)\lambda^n.$$

Therefore, u and y can be considered as the input and output of a "fictitious" time-varying system defined above (Figure 5.3) subjected to the disturbance e. Suppose now that the controller C provides stability for the fictitious system and also has the property $\|S^{ue}\|, \|S^{ye}\| \le 1/(D_\Delta(1 + \eta))$ where S^{ue} and S^{ye} are the maps from e to u and y respectively and $\eta > 0$ with $1 < \beta < 1 + \eta$. Then, the assumptions of Lemma 5.5.1 are
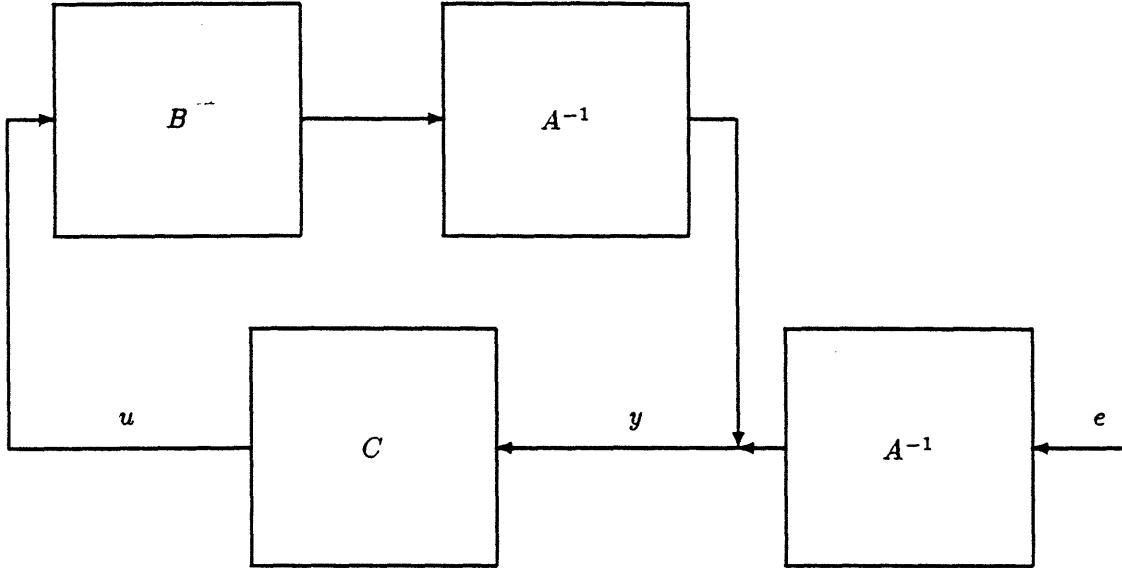
Figure 5.3: The "Fictitious" Feedback Loop

satisfied: the validity of assumption 1 follows from the stability of the fictitious system, assumption 2 is satisfied from the properties of the parameter estimation algorithm and, finally, assumption 3 is fulfilled since

$$D^1(t) = D_\Delta \max_{0 \leq \tau \leq t} (|u(\tau)|, |y(\tau)|) \leq D_\Delta \max(\|S^{ue}\|, \|S^{ye}\|) \max_{0 \leq \tau \leq t} |e(\tau)|$$

or

$$D^1(t) \leq k_2 \max_{0 \leq \tau \leq t} |e(\tau)|, \quad k_2 = 1/(1 + \eta) < 1/\beta.$$

In fact, we can relax the norm requirements $\|S^{ue}\|, \|S^{ye}\| \leq 1/(D_\Delta(1 + \eta))$ on the maps S^{ye}, S^{ue} by imposing the same condition for a delayed version of the fictitious system. To view this, suppose the controller is defined by LTI stable operators as

$$(L_t u)(t) = (M_t y)(t) + r(t)$$

where $\hat{L}_t(\lambda), \hat{M}_t(\lambda)$ are coprime polynomials of degree at most N for all t. Without loss of generality we can take $N \geq n$. Let Λ be the right shift operator and assume that there is a time index T so that the delayed maps $\Lambda^{-T} S^{ue} \Lambda^T$, $\Lambda^{-T} S^{ye} \Lambda^T$, $\Lambda^{-T} S^{ur} \Lambda^T$,

$\Lambda^{-T} S^{yr} \Lambda^T$ are in $\mathcal{L}_{TV}$ i.e., if the inputs $e, r \in \ell_\infty$ are delayed by T then the resulting y, u are in ℓ_∞. Moreover, let

$$\|\Lambda^{-T} S^{ue} \Lambda^T\|, \|\Lambda^{-T} S^{ye} \Lambda^T\| \leq 1/(D_\Delta(1 + \eta)).$$

Defining the state vector

$$x(t) = (y(t - N) \; u(t - N) \; y(t - N + 1) \; u(t - N + 1) \; \ldots \; y(t - 1) \; u(t - 1))^T$$

we obtain the state space description

$$x(t+1) = \begin{pmatrix} 0 & I & 0 & \ldots & 0 \\ 0 & 0 & I & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & I \\ \Phi_1(t) & \Phi_2(t) & \ldots & \Phi_{N-1}(t) & \Phi_N(t) \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ I \end{pmatrix} w(t)$$

where $\Phi_i(t)$ are 2×2 matrices obtained from the coefficients of $\hat{A}_t(\lambda), \hat{B}_t(\lambda), \hat{L}_t(\lambda), \hat{M}_t(\lambda)$ and $w(t) = (e(t) \; r(t))^T$. Clearly, this is a completely reachable state space representation of the closed loop. Therefore, we can ensure that the initial input $\{e(t)\}_{t=0}^T$ does not produce unbounded u and/or y for, otherwise, it contradicts the stability of the delayed system: any state at $T + N$ is reachable by some $w \in \ell_\infty$ of the form $w = \{0, \ldots, 0, w(T), w(T + 1), \ldots, w(T + N), 0, \ldots\}$ and a zero state at T; hence the initial input $\{e(t)\}_{t=0}^T$ cannot drive the system to a state at $T + N$ that results to unbounded $x(t)$, $t \geq T$ because then a bounded w as above applied to the delayed system would yield an unbounded y and u which is a contradiction. But then we can pick some nonzero c_1, k_1 to account for the initial input $\{e(t)\}_{t=0}^T$ and have the conditions of Lemma 5.5.1 satisfied for $T_1 = T_2 = T$. This in turn will guarantee weak stability of the adaptive scheme.

5.5.3   A Class of Stabilizing Controllers

Property 2 of the estimation algorithm shows that the parameter estimates will eventually vary arbitrarily slowly; hence since $\|A_t - A_{t-1}\| \leq |\theta_t - \theta_{t-1}|_2$ and $\|B_t - B_{t-1}\| \leq |\theta_t - \theta_{t-1}|_2$ it follows that eventually $A_t, B_t \in STV(\gamma)$ for some $\gamma > 0$ arbitrarily small. Utilizing now the results of Theorem 5.4.1 and Lemma 5.5.1 we are

able to characterize a class of stabilizing controllers for the original system of Figure 5.2. This is done in the following theorem.

Theorem 5.5.1 *Assume that for each t the frozen time controller $C_t = L_{t-1}^{-1} M_{t-1}$ stabilizes the frozen time LTI system given by $A_{t-1} y = B_{t-1} u + e$. Also let the following be true*

1. *There are constants c_M, $c_L > 0$ so that*

$$\|M_t - M_{t-1}\| \le c_M |\theta_t - \theta_{t-1}|_2, \quad \|L_t - L_{t-1}\| \le c_L |\theta_t - \theta_{t-1}|_2.$$

2. *The degrees of $\hat{L}_t(\lambda)$, $\hat{M}_t(\lambda)$ as well as the ITAE of M_t, L_t are uniformly bounded in t.*

3. *The $\mathcal{L}_{TI}$ norm and the ITAE of $H_t = (L_t A_t + M_t B_t)^{-1}$ are uniformly bounded in t.*

4. *There is a $\epsilon > 0$ and a time index $T_0 > 0$ so that*

$$\sup_{t \ge T_0} \| \begin{pmatrix} S_t^{ue} \\ S_t^{ye} \end{pmatrix} \| \le \frac{1-\epsilon}{D_\Delta(1+\eta)} - \epsilon$$

where $S_t^{ue} = M_{t-1} H_{t-1}$, $S_t^{ye} = L_{t-1} H_{t-1}$.

Then the control law $u(t) = (Cy)(t)$ where $(L_{t-1} u)(t) = (M_{t-1} y)(t)$, $t = 0, 1, \ldots$ yields a weakly stable adaptive system.

Proof. The proof of the theorem follows the same steps as Theorem 5.4.1. First, since the estimates should lie in the compact set Θ and $\hat{A}_t(\lambda)$, $\hat{B}_t(\lambda)$ have degree n then A_t, B_t have uniformly bounded norms and ITAE [7]. Also, from assumption 1 the $\mathcal{L}_{TI}$ norms of M_t, L_t will be uniformly bounded. Note that $S_t^{ue} = M_{t-1} H_{t-1}$, $S_t^{ye} = L_{t-1} H_{t-1}$ are precisely the maps from e to u and from e to y respectively in the frozen LTI system. Since the rate of variation of the estimates converges to zero then there exists some time index $T \ge T_0$ after which the rate of variation is sufficiently small to guarantee stability of the delayed fictitious system and, moreover, the delayed performance conditions

$$(1-\epsilon) \|\Lambda^{-T} S^{ue} \Lambda^T\| \le \sup_{t \ge T} \|S_t^{ue}\| + \epsilon$$

83

and

$$(1 - \epsilon)\|\Lambda^{-T} S^{ye} \Lambda^T\| \leq \sup_{t \geq T} \|S^{ye}_t\| + \epsilon.$$

The above assertion can be proved exactly as in Theorem 5.4.1. But then

$$\|\Lambda^{-T} S^{ye} \Lambda^T\| \leq 1/(D_\Delta(1 + \eta)), \ \|\Lambda^{-T} S^{ue} \Lambda^T\| \leq 1/(D_\Delta(1 + \eta))$$

and hence by the key technical Lemma 5.5.1 the proof is complete. ∎

Note that the requirements on the frozen-time LTI maps $S^{ue}_t = M_{t-1} H_{t-1}$, $S^{ye}_t = L_{t-1} H_{t-1}$ are exactly the stability-robustness requirements in the presence of the co-prime factor perturbations Δ_A, Δ_B of "almost" the same magnitude (for small enough η, ϵ) in the estimated LTI system. Hence the meaning of the above theorem is the following: Given a design methodology that produces controllers C_t which are Lipschitz-continuous with respect to the parameter estimate θ_t, then stability of the frozen-time feedback loop of $(A_{t-1} + \Delta_A)^{-1}(B_{t-1} + \Delta_B)$ and C_t will result to weak stability of the adaptive scheme. Next, we present an indirect adaptive scheme which produces frozen-time controllers that possess the required continuity properties.

5.5.4 The ℓ_1 Adaptive Algorithm

Here we present briefly a generalization of the ℓ_1 adaptive scheme of [7] and give sufficient conditions for stability. The scheme utilizes frozen-time controllers $C_{t+1} = L_t^{-1} M_t$ at each $t + 1$ that stabilize $P_t = A_t^{-1} B_t$ and minimize the following criterion

$$\inf_{C_{t+1} \text{ stabilizing}} \left\| \begin{pmatrix} S^{ue}_{t+1} \\ S^{ye}_{t+1} \end{pmatrix} \right\| = \inf_{C_{t+1} \text{ stabilizing}} \left\| \begin{pmatrix} C_{t+1} A_t^{-1} (1 + P_t C_{t+1})^{-1} \\ A_t^{-1} (1 + P_t C_{t+1})^{-1} \end{pmatrix} \right\| \overset{\text{def}}{=} \mu(\theta_t).$$

Employing the parameterization of all stabilizing controllers [28] for P_t we transform the problem to

$$\mu(\theta_t) = \inf_{Q_t} \left\| \begin{pmatrix} Y_t \\ X_t \end{pmatrix} + \begin{pmatrix} A_t \\ B_t \end{pmatrix} Q_t \right\|$$

where X_t, Y_t are polynomials in $\mathcal{L}_{TI}$ satisfying the Bezout identity

$$X_t A_t - Y_t B_t = 1.$$

All stabilizing controllers are obtained as

$$C_{t+1} = \frac{Y_t + A_t Q_t}{X_t + B_t Q_t}.$$

Since A_t, B_t are coprime then the only restriction on $K_t = \begin{pmatrix} A_t \\ B_t \end{pmatrix} Q_t$ is

$$(-B_t \quad A_t) K_t = 0$$

which implies that the only interpolation [5] on the closed loop $\Psi_t = \begin{pmatrix} S_{t+1}^{ue} \\ S_{t+1}^{ye} \end{pmatrix}$ is

$$(-B_t \quad A_t) \Psi_t = 1.$$

As indicated in [5] the problem can be transformed to a semiinfinite linear programming problem using duality. The solution to the latter can then be computed with arbitrary accuracy by truncating the constraints or the variables. In particular, the resulting problem is as follows [5]

$$\mu(\theta_t) = \sup_{x \in c^0} x(0)$$

subject to

$$\left\| \begin{pmatrix} -b_t(0) & -b_t(1) & \ldots & -b_t(n) & 0 & 0 & \ldots \\ 0 & -b_t(0) & -b_t(1) & \ldots & -b_t(n) & 0 & \ldots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ a_t(0) & a_t(1) & \ldots & a_t(n) & 0 & 0 & \ldots \\ 0 & a_t(0) & a_t(1) & \ldots & a_t(n) & 0 & \ldots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \\ \vdots \end{pmatrix} \right\|_{c^0} \leq 1.$$

¿From the above formulation continuity of the cost $\mu(\theta_t)$ with respect to parameters changes i.e., θ_t is easy to be established. This does not automatically imply that the assumptions 1, 2, 3 of Theorem 5.5.1 are satisfied. What we need is Ψ_t to be continuous with respect to θ_t and also to have a uniform degree bound. These requirements might not be satisfied in this complete generality; for example when the solution is not unique then continuity is immediately destroyed. Hence, additional assumptions might be needed. In the case however, where $\Delta_B = 0$ this is not needed. In this case, the problem becomes

$$\inf_{C_{t+1} \text{ stabilizing}} \| A_t^{-1}(1 + P_t C_{t+1})^{-1} \| \stackrel{\text{def}}{=} \mu_A(\theta_t).$$

85

As shown in [7] the finite dimensionality of the LTI system P_t, the compactness of Θ together with the properties of the optimal ℓ_1 solution serve to satisfy assumptions 1,2,3 of Theorem 5.5.1. For assumption 4 to hold the following condition suffices:

$$\exists\, \epsilon > 0 : \mu_A \stackrel{\text{def}}{=} \sup_{\theta \in \Theta} \mu_A(\theta) \leq \frac{1-\epsilon}{D_\Delta(1+\eta)} - \epsilon \tag{C}.$$

Note that $\mu_A < \infty$ since P_θ is finite dimensional and Θ is compact. Conversely, from the above condition we can evaluate the bound D_Δ of $\|\Delta_A\|$ for which the ℓ_1 indirect adaptive scheme guarantees stability. Namely,

$$D_\Delta < \frac{1}{\mu_A} \quad .$$

This is so because then there are $\epsilon, \eta > 0$ so that condition (C) holds. We should emphasize that, pertaining to this particular class of indirect adaptive controllers, the ℓ_1 adaptive scheme provides the greatest upper bound on the size of $\|\Delta_A\|$ namely $1/\mu_A$ for which stability is guaranteed. We do not however claim that this adaptive scheme is the optimal one. Also note that even if $\Delta_B = 0$ the plant model captures a wide class of uncertain systems. Finally, we should stress that in the case where $\Delta_B \neq 0$ if the continuity assumptions are satisfied then a bound on D_Δ for which stability is guaranteed is

$$D_\Delta < \frac{1}{\sup_{\theta \in \Theta} \mu(\theta)} \quad .$$

5.6 Conclusions

In this chapter we presented a ℓ_∞ to ℓ_∞ performance result in the case of slowly time varying systems. We showed that the performance of a slowly varying system cannot be much worse than that of the frozen time systems. Our approach was an input-output approach established in [6]. We utilized this result to characterize a class of adaptive indirect controllers t hat stabilize a time invariant system which is subjected to both parametric and unstructured uncertainty. Also, among a class of indirect adaptive controllers, we proposed an indirect adaptive scheme that provides the greatest upper bound on the size of unstructured uncertainty for which stability is guaranteed.

Bibliography

[1] B.A. Bamieh, M.A. Dahleh and J.B. Pearson, "Minimization of the L^∞-Induced Norm for Sampled-Data Systems," *submitted.*

[2] H. Chapellat, M. Dahleh, and S.P. Bhattacharyya, Optimal Disturbance Rejection for Periodic Systems, Technical Report #89–019, TAMU, College Station, TX, 1989.

[3] S.P. Boyd and C.H. Barratt, *Linear Controller Design: Limits of Performance*, Prentice Hall, N.J., 1991.

[4] S.P. Boyd and J.C. Doyle, "Comparison of Peak and RMS Gains for Discrete-Time Systems," *Systems & Control Letters,* 9, 1-6, 1987.

[5] M.A. Dahleh, "BIBO Stability Robustness for Coprime Factor Perturbations," *IEEE Trans. Automat. Contr.,* Vol 37, pp. 352-355, 1992.

[6] M. Dahleh and M.A. Dahleh. "On slowly time varying systems," *Automatica,* Vol 27, No 1, January 1991.

[7] M. Dahleh and M.A. Dahleh. "Optimal rejection of persistent and bounded disturbances: Continuity properties and adaptation," *IEEE Trans. A-C,* Vol AC-35, June 1990.

[8] M.A. Dahleh and M.H. Khammash. "Controller Design for Plants with Structured Uncertainty," to appear in *Automatica Special issue on robust control.*

[9] M.A. Dahleh and Y. Ohta, "A Necessary and Sufficient Condition for Robust BIBO Stability," *Systems & Control Letters,* 11, 271-275, 1988.

[10] M.A. Dahleh and J.B. Pearson, "l^1 Optimal Feedback Controllers for MIMO Discrete-Time Systems," *IEEE Trans. A-C* **32**, 314-322, 1987.

[11] M.A. Dahleh and J.B. Pearson, "Optimal Rejection of Persistent Disturbances, Robust Stability and Mixed Sensitivity Minimization," *IEEE Trans. Automat. Contr.* **33**, 722-731, 1988.

[12] M.A. Dahleh and J.B. Pearson. "Minimization of a Regulated Response to a Fixed Input," *IEEE Trans. Automat. Contr.* **33**, 924-930, 1988.

[13] M.A. Dahleh and D. Richards, "Application of Modern Control Theory on a Model of the X-29 Aircraft," LIDS, MIT, report No. LIDS-P-1932, 1989.

[14] M.A.Dahleh and J.S. Shamma. "Rejection of Persistent Bounded Disturbances: Nonlinear Controllers," *Sys. and Contr. Letters*, Vol 18, pp. 245-252, 1992.

[15] M.A. Dahleh, P. Voulgaris and L. Valavani, "Optimal and Robust Controllers for Periodic and Multi-Rate Systems," *to appear in IEEE Trans. Automat. Contr.*, January 1992.

[16] M. Dahleh and M.A. Dahleh, "Optimal Rejection of Persistent and Bounded Disturbances: Continuity Properties and Adaptation," *IEEE Trans. A-C* **35**, 687-696, 1990.

[17] G. Deodhare and M. Vidyasasgar. "Some Results on ℓ_1-Optimality of Feedback Control Systems: The SISO Discrete-Time Case", *IEEE Trans. A-C*, **35**, 1082-1085, 1990.

[18] C.A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, Inc, N.Y., 1975.

[19] C.A. Desoer, R-W. Liu, J. Murray and R. Saeks. "Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis," *IEEE Trans. on Automatic Control*, vol. AC-25, no. 3, pp. 399-412, June 1980.

[20] J.C. Doyle, K. Glover, P.P. Khargonekar and B.A. Francis, "State Space Solutions to Standard H_2 and H_∞ Control Problems," *IEEE Trans. A-C* **34**, 831-847, 1989.

[21] J.C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE-Trans. A-C* **26**, 4-16, 1981.

[22] J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainty," *IEEE Proceedings*, **129**, PtD, 242-250, 1982.

[23] I. Diaz-Bobillo and M.A. Dahleh, "State Feedback ℓ_1 Optimal Controllers Can be Dynamic," *Sys. and Contr. Letters*, Vol 19, pp. 87-93, 1992.

[24] I. Diaz-Bobillo and M.A. Dahleh, "Minimization of the Maximum Peak-to-Peak Gain: The General Multiblock Problem," *Submitted.*

[25] G. Dullerud and B.A. Francis, "$\mathcal{L}^1$ Analysis and Design of Sampled-Data Systems," *IEEE Trans. Auto. Control*, to appear April, 1992.

[26] K. Glover and D. McFarlane, "Robust stabilization of normalized coprime factor plant description with H_∞ bounded uncertainty," *IEEE Trans. Auto. Control*, **34**, 831-830, 1989.

[27] A. Feintuch and B.A. Francis, Uniformly Optimal Control of Linear Time-Varying Systems, *Systems & Control Letters* **5**, 1985, pp. 67–71.

[28] B.A. Francis, *A Course in H_∞ Control Theory,* Springer-Verlag, 1987.

[29] G.C. Goodwin, D.J. Hill, D.Q. Mayne and R.H. Middleton. "Adaptive Robust Control," Tech. Rep. EE8544, Dec. 1985, University of Newcastle, NSW 2308 Australia.

[30] G.C. Goodwin and K.S. Sin. *"Adaptive Filtering Prediction and Control,"* Prentice-Hall, Engelwood Cliffs, NJ.

[31] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[32] M. Khammash and J.B. Pearson, "Robust Disturbance Rejection in ℓ^1-Optimal Control Systems," *Systems & Control Letters*, **14**, 93-101, 1990.

[33] M. Khammash and J. B. Pearson, "Performance Robustness of Discrete-Time Systems with Structured Uncertainty," *IEEE Trans. A-C* **36**, 398-412, 1991.

[34] M. Khammash and J. B. Pearson, "Robustness Synthesis for Discrete-Time Systems with Structured Uncertainty," *Proc. ACC*, 2720-2724, 1991.

[35] M. Khammash and J. B. Pearson, "Robust Controller Design in the Presence of Structured Uncertainty," preprint.

[36] M. Khammash, "Necessary and Sufficient Conditions for the Robustness of Time-Varying Systems with Applications to Sampled-Data Systems," *submitted.*

[37] P.P. Khargonekar and K.R. Poolla, Uniformly Optimal Control of Linear Time-Varying Plants: Nonlinear Time-Varying Controllers, *Systems & Control Letters* **5**, 1986, pp. 303–308.

[38] P.P. Khargonekar and K.R. Poolla, Robust Control of Linear Time-Invariant Plants Using Switching and Nonlinear Feedback, *Proceedings of the 28th IEEE Conference on Decision and Control*, 1989, pp. 2205–2207.

[39] P.P. Khargonekar, K.R. Poolla, and A. Tannenbaum, Robust Control of Linear Time-Invariant Plants by Periodic Compensation, *IEEE Transactions on Automatic Control*, Vol. AC-30, 1985, pp. 1088–1096.

[40] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, (Wiley-Interscience, New York, 1972).

[41] J.M. Krause, P.P. Khargonekar, G. Stein. "Robust adaptive control: Stability and asymptotic performance," *Proceedings of the 28th CDC*, Tampa, Florida, December 1989.

[42] R. Lamaire. "Robust time and frequency domain estimation methods in adaptive control," PhD Thesis LIDS-TH- 1674, MIT, June 1987.

[43] D.G. Luenberger, *Optimization by Vector Space Methods*, John Wiley and Sons, N.Y., 1969.

[44] J.S. McDonald and J.B. Pearson, "ℓ_1-Optimal Control of Multivariable Systems with Output Norm Constraints," *Automatica* **27**, 317-329, 1991.

[45] M.A. Mendlovitz, "A Simple Solution to the ℓ_1 Optimization Problem," *Systems & Control Letters*, **12**, 461-463, 1989.

[46] D.G. Meyer, Two Properties of ℓ_1 Optimal Controllers, *IEEE Trans. Automat. Control* **33** (1988) 876-878.

[47] A. Packard and J.C. Doyle. "The Complex Structured Singular Value," *to appear in Automatica special issue on robust control.*

[48] K. Poolla and J.S. Shamma, Asymptotic Performance through Adaptive Robust Control, *Proceedings of the 29th IEEE Conference on Decision and Control*, Honolulu, HA, December 1990.

[49] J.S. Shamma, Performance Limitations in Sensitivity Reduction for Nonlinear Plants, *Systems & Control Letters* **17**, 1991, pp. 43–47.

[50] J.S. Shamma and M. Athans. "Analysis of gain scheduled control for nonlinear plants," *IEEE Trans. A-C*, Vol 25, No 8, August 1990.

[51] J.S. Shamma and M.A. Dahleh, Time-Varying vs. Time-Invariant Compensation for Rejection of Persistent Bounded Disturbances and Robust Stabilization, *IEEE Transactions on Automatic Control*, Vol. 36, No. 7, July 1991, pp. 838–847.

[52] J.S. Shamma and R. Zhao, "Fading Memory and Necessity of the Small Gain Theorem," *30th IEEE Conference on Decision and Control*, December 1991.

[53] N. Sivashankar and P. Khargonekar, "$\mathcal{L}_\infty$-induced norm of sampled-data systems," *Proc. of the CDC*, 1991.

[54] O.J. Staffans, "Mixed Sensitivity Minimization Problems with Rational ℓ_1-optimal solutions," *Journal of Optimization Theory and Applications*, **70**, 173-189, 1991.

[55] O.J. Staffans, "On the Four-block Model Matching Problem in ℓ_1," Helsinki University of Technology, Espoo, Report A289, 1990.

[56] M.S. Verma, Coprime Factorizational Representations and Stability of Nonlinear Feedback Systems, *International Journal of Control*, Vol. 48, 1988, pp. 897–918.

[57] M. Vidyasagar, *Control Systems Synthesis: A Factorization Approach*, MIT press, 1985.

[58] M. Vidyasagar "Optimal Rejection of Persistent Bounded Disturbances," *IEEE-Trans. A-C* **31**, 527-534, 1986.

[59] P. Voulgaris M.A. Dahleh and L. Valavani," Slowly-Varying Systems: ℓ_∞ to ℓ_∞ Performance and Implications to Adaptive Control," *submitted*.

[60] J. C. Willems, *The Analysis of Feedback Systems*, MIT Press, Cambridge, 1971.

[61] D.C. Youla, H.A. Jabr, and J.J. Bongiorno, "Modern Wiener-Hopf Design of Optimal Controllers–part 2: The Multivariable Case," *IEEE-Trans. A-C* **21**, 319-338, 1976.

[62] G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses," *IEEE-Trans. A-C* **26**, 301-320, 1981.

[63] G. Zames and L.Y. Wang. "Local-Global Double Algebras for Slow H^∞ Adaptation: Part I-Inversion and Stability," *IEEE Trans. A-C*, Vol AC-36, February 1991.

[64] L.Y. Wang and G. Zames. "Local-Global Double Algebras for Slow H^∞ Adaptation: Part II-Optimization of Stable Plants," *IEEE Trans. A-C*, Vol AC-36, February 1991.

[65] K. Zhou and P. Khargonekar, An algebraic Riccati equation approach to $\mathcal{H}_\infty$ optimization, *Systems & Control Letters* **11** (1988) 85-91.