CDS

TECHNICAL MEMORANDUM NO. CIT-CDS 96-010
June, 1996

"Sets and Constraints in the Analysis Of Uncertain Systems"

Fernand Paganini

Control and Dynamical Systems
California Institute of Technology
Pasadena, CA 91125

Sets and Constraints in the Analysis of Uncertain Systems

Thesis by

Fernando Paganini

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1996
Submitted December 1, 1995

To Malena

Acknowledgements

My first thanks go to my advisor, John Doyle. It has become customary for anyone who has had close collaboration with John to credit him with the "big picture", and I am no exception. His main contribution to my education has been, however, that instead of providing this picture as a given and rewarding me for purely technical contributions, he encouraged me to seek a role in defining this picture; I have still a long way to go in the pursuit of this goal, but this outlook on academic research has made me grow in the direction which added the most to the intellectual resources I brought to Caltech.

It has been a privilege to share much of this road with many outstanding people in the Controls group at Caltech. I particularly wish to thank Raff D'Andrea, who provided his quick "parallel" thinking and his enthusiasm to "figure it out for ACC" which pushed me forward in many key opportunities. We also shared countless *mates* at the lab, soccer games, and philosophical conversations in the office, which helped make my days (and nights) more enjoyable. With Geir Dullerud we cultivated our common taste for mathematics; I could always turn to him for guidance and some humor on the intricate ways of the academic world, for which I am very grateful. I also wish to thank Matt Newlin, who shared with me some of his vast knowledge, Richard Murray, who often found time for me in his impossible schedule, and Stefano Soatto, who pointed me to the Bartlett test, which turned out to give the right answer after all. Some visitors such as Munzer Dahleh and Sasha Megretski provided useful technical input at some strategic times.

This thesis would not have been possible without the solid education I brought from the Universidad de la Republica in Montevideo, Uruguay. The contributors are too many to give names, but I wish to thank them all for their proud and obstinate effort to provide, against many odds, a quality of education which, I have found, should envy none in the whole world.

At a more fundamental level, this thesis has been built on the foundations I received from my family. My parents, Omar and Celia, and my brothers Omi and Juanjo, helped me grow up in an environment of caring, support and intellectual stimulation. From my grandfather, Fernando Herrera Ramos, I received a little of his relentless drive to go forward; his enduring memory is present in these pages.

My wife, Malena, gave me the love, unconditional support and enthusiasm for life which carried us both through this, which is truly our joint achievement. From our greater achievement, Rafael, I received the fresh breath of Life which is infinitely richer than our best theories, and whose gratuity surpasses our most consummate efforts.

Abstract

This thesis is concerned with the analysis of dynamical systems in the presence of model uncertainty. The approach of robust control theory has been to describe uncertainty in terms of a structured *set* of models, and has proven successful for questions, like stability, which call for a worst-case evaluation over this set. In this respect, a first contribution of this thesis is to provide robust stability tests for the situation of combined time varying, time invariant and parametric uncertainties.

The worst-case setting has not been so attractive for questions of disturbance rejection, since the resulting performance criteria (e.g., $\mathcal{H}_\infty$) treat the disturbance as an adversary and ignore important spectral structure, usually better characterized by the theory of stochastic processes. The main contribution of this thesis is to show that the set-based methodology can indeed be extended to the modeling of white noise, by employing standard statistical tests in order to identify a typical set, and performing subsequent analysis in a worst-case setting. Particularly attractive sets are those described by quadratic signal constraints, which have proven to be very powerful for the characterization of unmodeled dynamics. The combination of white noise and unmodeled dynamics constitutes the Robust $\mathcal{H}_2$ performance problem, which is rooted in the origins of robust control theory. By extending the scope of the quadratic constraint methodology we obtain a solution to this problem in terms of a convex condition for robustness analysis, which for the first time places it on an equal footing with the $\mathcal{H}_\infty$ performance measure.

A separate contribution of this thesis is the development of a framework for analysis of uncertain systems in implicit form, in terms of equations rather than input-output maps. This formulation is motivated from first principles modeling, and provides an extension of the standard input-output robustness theory. In particular, we obtain in this way a standard form for robustness analysis problems with constraints, which also provides a common setting for robustness analysis and questions of model validation and system identification.

Table of Contents

List of Figures

Chapter 1

Introduction

The field of control has evolved over this century from its origins in the study of feedback amplifiers, into a broad discipline concerning issues of modeling, dynamics, optimization and feedback. On one side lie complex engineering problems, such as regulation in chemical processes, trajectory tracking for robot manipulators, stabilization of high performance aircraft, or dynamics of queueing systems. On the other side lie tools from virtually every mathematical discipline, from dynamical systems and differential geometry to stochastic processes and operator theory. In the middle of this, the task of the control theorist is to abstract a problem of significance in engineering, cast it in an appropriate mathematical setting, and derive a solution, by which is meant a practically computable method of evaluation of the problem at hand. This eclectic mix of disciplines has made control theory the home of people, like the author, who have found it difficult to choose between the fascinating worlds of engineering and mathematics.

While engineers are mainly concerned with real-world problems, and mathematicians with the logical consistency of their abstractions, it is the job of those who attempt to apply mathematics to the real world to deal with the fundamental gaps between theory and practice, which reflect themselves in *uncertainty* about the behavior of a real system when one is given a mathematical prediction. This is particularly the case for control theory, which treats the question of feedback, a technique used by both natural and artificial systems to obtain reliability in spite of faulty predictions. A properly design feedback compensator will effectively reduce the sensitivity of the system to certain sources of uncertainty, but at the expense of increased sensitivity to other unmodeled effects, e.g. in another frequency band. Consequently, a theory of feedback must provide means to quantify these tradeoffs, which can only be achieved if in addition to a mathematical model, one utilizes some form of quantification of model uncertainty.

This thesis is concerned with the analysis of dynamical systems from the point of view of the effect of uncertainty. The fundamental challenge in this area has been to refine as

much as possible the uncertainty description in a model of a complex system, compatible with the possibility of a tractable evaluation of its effect. Two competing strategies have appeared for the characterization of uncertainty: to describe a set of models and perform worst-case analysis, or to assign a probability distribution to the errors and pursue analysis via expected values. The former, the method of choice in *robust* control, has been successful in quantifying the effects of unmodeled dynamics and its impact on the question of stability margins. The latter, pursued by *stochastic* control, has had success in the non-conservative evaluation of the effects of noise, and has also been the method of preference for system identification based on data. The historical difficulties in the combination of the two, which is required in practical problems, highlights one of the shortcomings of a research field based on such a diverse set of mathematical tools, and constituted by communities which speak different languages.

One of the main contributions of this thesis is to provide a set of tools for these kinds of problems which call for a combination of worst-case evaluation over sets of systems, simultaneously with evaluation of the effects of white noise. The most relevant example in control theory is the so-called "Robust $\mathcal{H}_2$" problem, which originates in the efforts in the 1970s to provide stability margins for the so-called LQG regulators, designed from the point of view of white noise rejection. The lack of a tractable solution fueled the development of robust control theory with regard to other, more conservative methods of performance evaluation, like $\mathcal{H}_\infty$ control, which were deemed more compatible with evaluation of margins; for a historical account see Chapter 6. In this thesis we provide a solution to the Robust $\mathcal{H}_2$ problem, based on a combination of statistical tests for white noise, and characterizations of uncertainty in terms of quadratic signal constraints. We reduce the problem to computationally tractable conditions which are of the same nature as those in mainstream robust control, thus closing a cycle of research originated in the 1970s.

From a broader perspective, the mathematical effort involved in the solution of this problem may appear excessive if the only application were to a specific feedback control problem; for this reason we investigate in this thesis which more general problems can be addressed with this type of tools. In particular, the uncertainty modeling techniques can be applied beyond the confines of the traditional feedback control configuration, and into questions of component modeling and system design. For these problems we introduce robustness analysis in *implicit* form, where systems are described in terms of uncertain equations. This approach yields a very general analysis theory for linear, uncertain system involving constraints, and in particular contributes to a more unified approach to robust control and system identification.

1.1 Organization of the Thesis

In addition to the present introductory Chapter 1, and Chapter 9 which contains the concluding remarks, the main body of the thesis is developed in seven Chapters 2 through 8, and two Appendices, A and B. All the material is strongly interrelated in that similar problems are treated with similar techniques. Some separate units of interest can be isolated, however, as represented in the diagram of Figure 1.1.

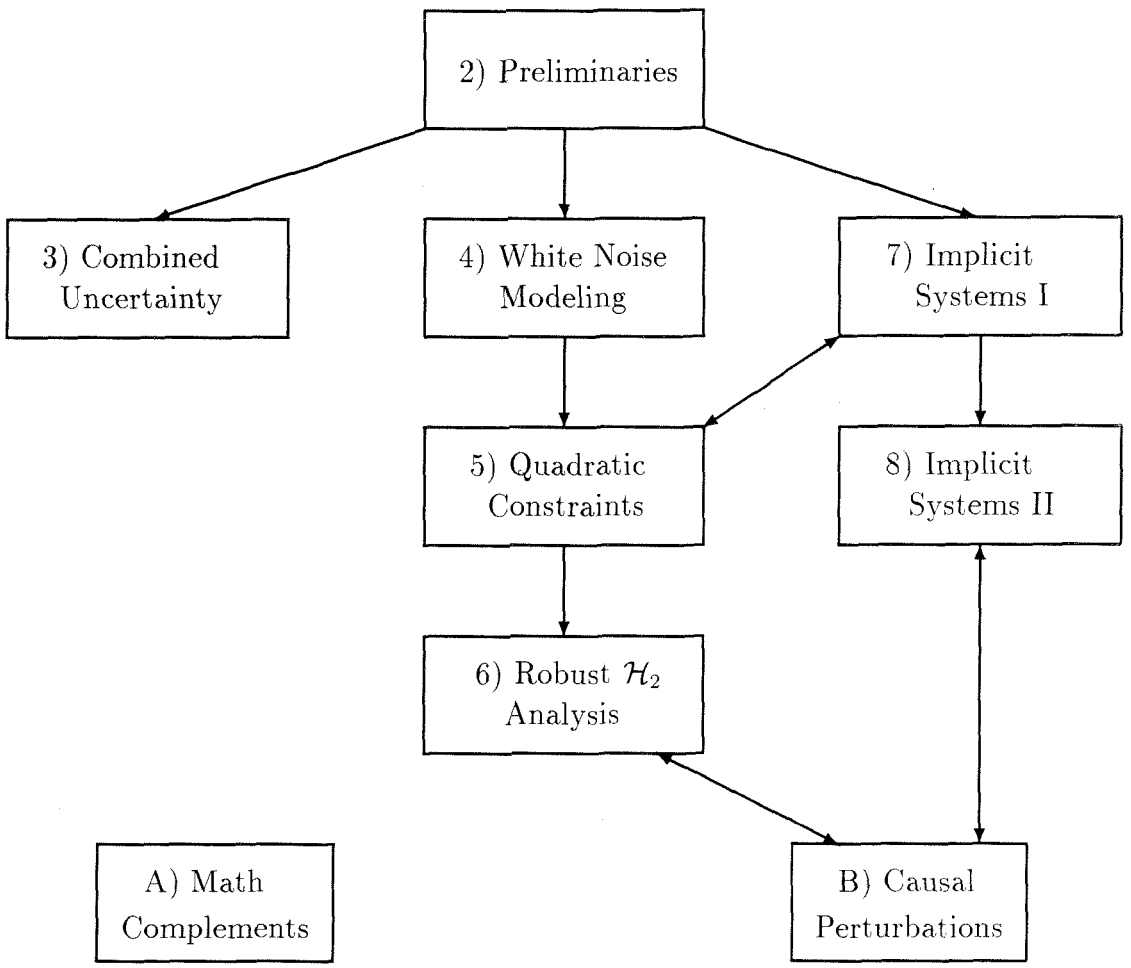


Figure 1.1: Organization of the thesis

Chapter 2 contains general background material on robust control theory, with a review of previous work related to this thesis. The remaining chapters have been divided in three columns in Figure 1.1, corresponding to areas of common interest which are largely independent. The arrows represent the strongest interrelations between the chapters.

Chapter 3 covers robust stability analysis under a combination of time invariant, time varying and parametric uncertainty classes, and is mostly separate from the subsequent material. It is included at this point since the methods employed are the most closely related to those reviewed in Chapter 2.

The sequence of Chapters 4, 5 and 6 contains the main results of this thesis. In Chapter 4 a methodology for modeling of white noise signals based on sets is introduced, and analyzed in relation to the more standard stochastic approach. These descriptions, as well as characterizations of unmodeled dynamics, rely on the Integral Quadratic Constraint formulation which is studied in Chapter 5, extending the methods introduced in [94, 50] to more general constraints. The approach leads to the solution of the Robust $\mathcal{H}_2$ performance analysis problem which is obtained in Chapter 6.

Chapters 7 and 8 develop a framework for analysis of uncertain systems described in implicit form. Chapter 7 discusses the motivation for this approach; in this respect an important reason for the generality of this setting is the close relationship between implicit uncertain systems and quadratic constraints as in Chapter 5. In Chapter 8, robustness analysis methods are obtained for these implicit systems.

Appendix A summarizes some results from mathematics (mostly convex analysis) which are used in different points of the thesis. In Appendix B we have collected some of the more technical proofs, in particular those referring to the construction of causal perturbations.

Chapter 2

Preliminaries

This chapter contains background material for the entire thesis. The aim is to give engineering motivation, to establish general notation, and to provide a summarized review of previous work on which the presentation will rely.

2.1 The Robust Control Framework

2.1.1 Modeling and Uncertainty

Any discipline which deals with mathematical models for physical reality is sensitive to the issue of the potential mismatch between the model and the real system, usually referred to as model uncertainty.

In the physical sciences, very accurate models are the objective in themselves. To obtain these "physical laws" one often distills the phenomenon to its simplest form. In this context, uncertainty is interpreted in a narrow sense as referring to the limits in the predictive power of the best available models. For example, the uncertainty associated with prediction in a chaotic systems, or the Uncertainty Principle in quantum mechanics refer to fundamental limitations in predictability.

Models play a different role in engineering science; they are tools employed in analysis, simulation and design of complex, artificial systems. Consequently, model fidelity must be traded off with the complexity of the modeling process and the tractability of the resulting mathematical and computational problems. From this point of view the best model is the simplest summary of the main aspects of the physical system which are relevant to the engineering question at hand. Correspondingly, the term "uncertainty" is used here in a broader sense: it not only describes what one is fundamentally unable to predict, but also, and often predominantly, many aspects of the system which one has chosen to neglect or simplify. For uncertainty in this broad sense, there is by definition no detailed model, but often the

modeling process yields a crude description which allows one to assess its implications on the overall system. These descriptions of uncertainty appear commonly and in various forms in engineering models, whether they result from "black box" system identification techniques, from "first principles" models obtained by application and simplification of physical laws, or a combination thereof.

As remarked in Chapter 1, the issue of uncertainty is at the heart of control engineering, since a feedback configuration can significantly affect the sensitivity of the system behavior to uncertainty at the component level. This is the main motivation for the construction of feedback systems, but also the main potential danger as unmodeled effects can, for example, lead to instability. Consequently, to perform good designs, the control engineer must be furnished with rich descriptions of uncertainty and tools to assess their impact in a complex system. It should be clear from the nature of these descriptions that no hard "guarantees" can result from this assessment; ultimately, the control engineer must be the final mediator between the mathematics and the real system.

2.1.2 A Survey of Uncertainty Descriptions

Traditional methods for uncertainty characterization in dynamical systems include parametric uncertainty, disturbance signals, and system perturbations to account for unmodeled dynamics. We now describe how these typically arise in modeling. For more motivation we refer to [21, 98, 15].

Parametric Uncertainty

Parameters are present in most engineering models. representing a physical quantity which can be assumed to be a real constant within the range of validity of the model. The following are some reasons for uncertainty in the value of a parameter:

- It could be obtained directly or indirectly from experimental data. which leads to statistical deviations.
- It could represent a standardized component (e.g. electrical resistor) subject to manufacturing tolerances.
- It could represent an operating condition which varies in an unforeseen way. A constant parameter is a reasonable model when this variation is very slow (e.g. ambient temperature). In other cases the rate of variation of the operating condition is comparable to the modeled dynamics (e.g. aerodynamic coefficient of an airplane executing a sharp maneuver). In this case a *time-varying* parameter may be preferred.

The most straightforward representation of parametric uncertainty is in terms of an interval of the real line, such as $p = p_0 + k_p\delta$, $\delta \in [-1, 1]$. In models of linear dynamical systems, it is common to encounter rational dependence of a transfer function on an uncertain parameter. This dependence can often be represented in terms of a system block diagram. For example, the configuration of Figure 2.1 represents the transfer function $\frac{1}{s+p}$, where p varies as above. In Section 2.1.3 we introduce more general feedback configurations which allow for higher order rational dependence.

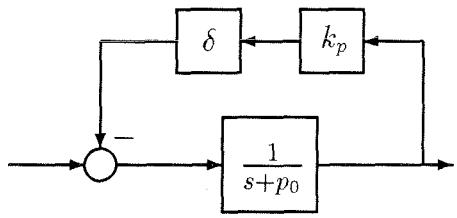


Figure 2.1: Block diagram representation for parametric uncertainty.

Disturbance Signals

Another commonly used method to account for model uncertainty is the injection of disturbances, which are thought of as generated by an external process. Some ways in which they arise are

- To account for microscopic fluctuations which are not included in a large scale model (e.g., wind turbulence, thermal noise in a circuit).
- To describe more systematic effects which are neglected in a simplified model (e.g. ripple in a voltage source, quantization error in an A/D converter).
- In identified models, frequently used as an error signal needed to account for the data.

The two standard choices for characterization of disturbances are in terms of a stochastic process, or in terms of a set of signals. These two competing methodologies will be further discussed in Chapter 4.

Unmodeled Linear Dynamics

The most commonly used dynamical system model for purposes of control is a linear, finite dimensional time invariant system, which is equivalent to a set of linear ODEs, preferably of

low order. Assuming for now that nonlinear effects are negligible, a low order approximation amounts to neglecting linear dynamics, in particular distributed effects.

One example where this may arise is in system identification in the frequency domain (see, e.g., [45]): Figure 2.2 shows points in a Nyquist plot obtained empirically, and a corresponding low order transfer function $G(j\omega)$ which approximates the data. A convenient way to express the remaining uncertainty in this model is to cover the points by a set of transfer functions of the form

$$G(j\omega)(1 + W(j\omega)\Delta(j\omega)), \tag{2.1}$$

where $\Delta(j\omega)$ can be considered to be *unmodeled dynamics* , i.e. an unspecified transfer function normalized such that $|\Delta(j\omega)| \leq 1$, and $W(j\omega)$ is a frequency weighting function.

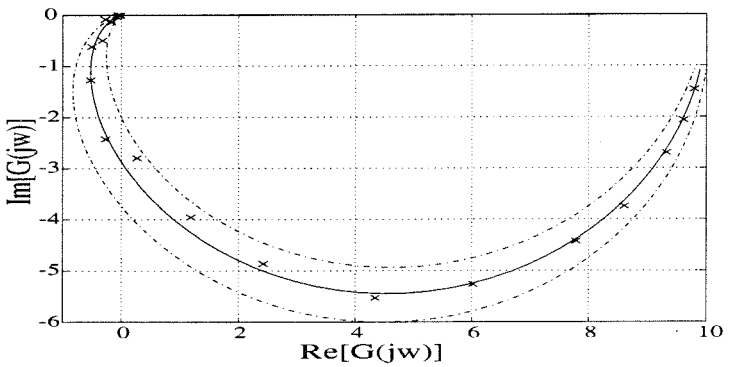


Figure 2.2: Characterization of unmodeled dynamics.

This kind of uncertainty description can also appear in models derived from first principles which are subsequently reduced to lower order, with a bound on the model error (see [38]).

Unmodeled Nonlinear Dynamics

If nonlinearities are very significant in the range of operation, the model itself must be chosen to be nonlinear. Uncertainty descriptions for nonlinear models are not very well developed, and are one of the main open challenges for a satisfactory theory of robust nonlinear control. There is, however, an intermediate situation, in which a linear model provides a reasonable approximation but one wishes to have bounds on the linearization error. This procedure is depicted in Figure 2.3 below, where a nonlinear static relationship between the variables X and Y is linearized around the operating point (X_0, Y_0). To bound the linearization error one can employ a "conic sector" description of the nonlinearity (see [96, 20, 89]) by the following bound on the incremental variables $y = Y - Y_0$, $x = X - X_0$:

$$|y - mx| \leq k_n|x|. \tag{2.2}$$

This imposes that the characteristic falls in the cone of Figure 2.3, which is valid for some range of x, y.

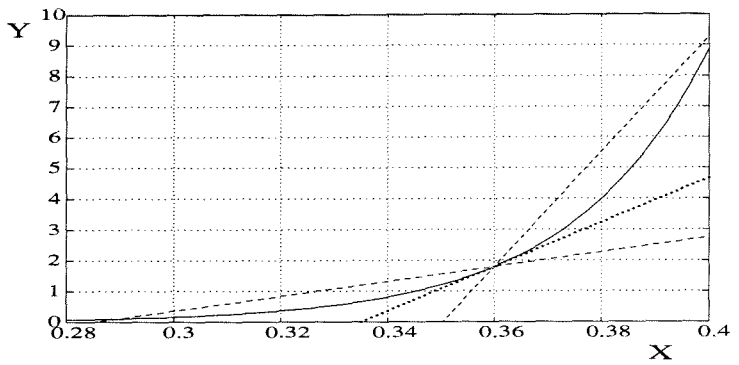


Figure 2.3: Linearization and conic sector bounds

The previous bound is a static constraint, which holds for every time instant, and is usually difficult to analyze. It is common to use instead the constraint

$$\|\mathbf{P^T} e\| \leq \|\mathbf{P^T} x\|, \tag{2.3}$$

where $\mathbf{P^T}$ denotes truncation in time (see (2.11)), and $e = \frac{y - mx}{k_n}$. (2.3) is equivalent to

$$e = \delta_n x, \tag{2.4}$$

with δ_n a contractive, causal nonlinear operator with memory, rather than static. This weaker assumption can be given a conic interpretation in signal space [96], and may be useful to account for dynamic effects which are not described in the static equations.

Remark:

In many important robust control problems, the analysis of such a description can be pursued with linear techniques, by replacing δ_n by a linear time-varying operator. Suppose we wish to answer a question of the type

> $\mathcal{Q}$: Does there exist a operator δ, $\|\delta\| \leq 1$, and nonzero signals x, e satisfying $e = \delta x$ in addition to some other system equations?

If one shows the answer to $\mathcal{Q}$ is negative over arbitrary causal linear operators, the same must hold for a nonlinear operator δ_n as considered in (2.4). The reason is that for each pair of signals e, x such that condition (2.3) holds, one can find a linear time-varying, causal and contractive operator which maps x to e (see [65], and also the related result in Lemma 5.1).

Many robustness analysis questions can be cast in the form $\mathcal{Q}$ (see Chapter 7), therefore this reasoning applies.

2.1.3 A Unified Framework for Uncertainty and Control

When modeling a complex system built from various components, the most natural approach is to construct smaller scale, subsystem models, and subsequently interconnect them to obtain the full system model. The various descriptions of uncertainty outlined above arise naturally at the component modeling stage; when these are combined together, they yield a highly structured description of the overall model uncertainty.

As an example, consider the block diagram of Figure 2.4 which is a model of a control system. The plant is given by a nominal model $\mathbf{H}$, an uncertain parameter δ_p represented as in Figure 2.1, and an unmodeled dynamics perturbation, represented by $\mathbf{\Delta}_2$. The output is measured by a sensor $\mathbf{F}$, which adds some sensor noise n. r is the reference signal, and the tracking error is fed into the controller $\mathbf{K}$, which acts on the plant. The block $\mathbf{\Delta}_1$ describes uncertainty (e.g. nonlinearities) in the actuator.

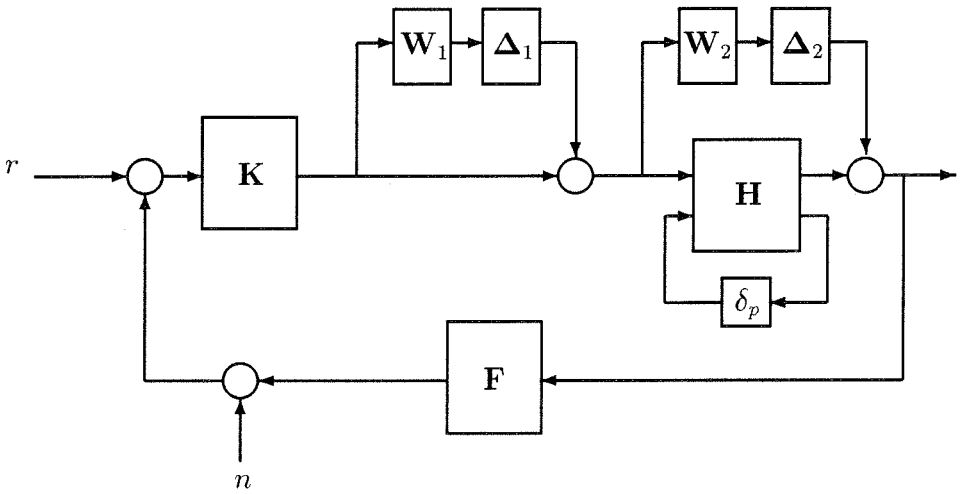


Figure 2.4: Control system with uncertainty

Given such a configuration, one wishes to address the question of robust control, i.e. controller design with satisfactory performance in the presence of the various uncertainty sources. At this level of description, it is not clear how to proceed with this task, and it appears that a solution would be specific to this configuration. In order to obtain more systematic and general tools, it is necessary to obtain a standard form, unifying the control problem and the various uncertainty sources in a common mathematical language.

We have already made the fundamental step in that direction, by representing both uncertainty and control in block diagram form; the final step is to isolate the uncertainty

blocks and the controller and redraw the block diagram in the standard configuration [54, 98] of Figure 2.5.

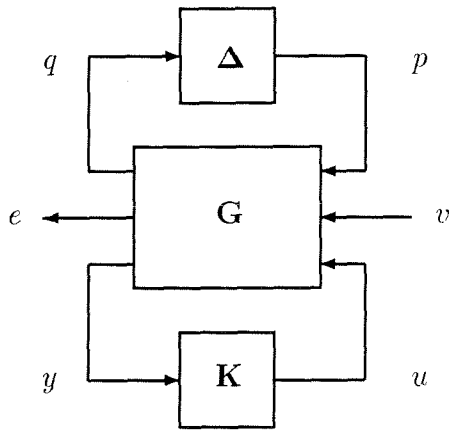


Figure 2.5: Standard robust control configuration

In this diagram, the generalized plant $\mathbf{G}$ contains the plant $\mathbf{H}$, the sensors $\mathbf{F}$ and the description of the overall system interconnection, including the weighting functions $\mathbf{W}_1$, $\mathbf{W}_2$. $\mathbf{K}$ is the controller as before, and the external signals v include in this case the reference r and the noise signal n. $\mathbf{\Delta}$ is a *structured* uncertainty operator of the form

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{\Delta}_1 & 0 & 0 \\ 0 & \mathbf{\Delta}_2 & 0 \\ 0 & 0 & \boldsymbol{\delta}_p \end{bmatrix}. \tag{2.5}$$

This spatial structure (which originates in the work of Safonov [74] and Doyle [28]) represents the different locations in which the blocks arise in Figure 2.4, by assigning them different signal channels. Also the blocks are of different nature: in the example, δ_p is a real parameter, and $\mathbf{\Delta}_1$, $\mathbf{\Delta}_2$ are dynamic perturbations which could be linear time invariant, or nonlinear.

The configuration of Figure 2.5 provides a general framework for linear robust control, flexible enough to encompass most theoretical results in this area.

The controller $\mathbf{K}$ is chosen to provide internal stability and performance, for all values of the perturbation $\mathbf{\Delta}$. The notion of performance typically considered in this setting is some measure of rejection of the signals v: the outputs ϵ are chosen as error variables, and performance means that the gain from v to ϵ is small in some sense, to be made precise in Section 2.2. Robustness *analysis* is the question of whether the desired properties are satisfied for a given controller. Robust *synthesis* is the task of obtaining the controller in

regard to these objectives. After providing a formal mathematical setting for these problems, we will review the most significant results on robustness analysis and synthesis.

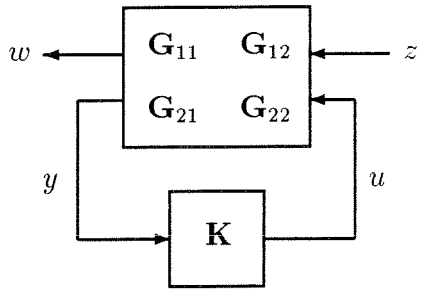


Figure 2.6: The controller feedback loop

It will be useful to extract two reduced configurations from Figure 2.5. The first one is obtained by removing the uncertainty and leads to the system of Figure 2.6, where w, z represent the joint signals (q, e) and (p, v) respectively. The closed loop map from z to w is denoted by $\mathbf{M}$; its dependence on $\mathbf{K}$ is given by a lower Linear Fractional Transformation (LFT, see [69, 98]), defined as

$$\mathbf{M} = \mathbf{G} \star \mathbf{K} := \mathbf{G}_{11} + \mathbf{G}_{12}\mathbf{K}(\mathbf{I} - \mathbf{G}_{22}\mathbf{K})^{-1}\mathbf{G}_{21}. \tag{2.6}$$

Replacing $\mathbf{M}$ back into the uncertain system of Figure 2.5, we obtain the robustness analysis configuration of Figure 2.7. The closed loop map from v to e is given by the upper LFT

$$\mathbf{\Delta} \star \mathbf{M} := \mathbf{M}_{22} + \mathbf{M}_{21}\mathbf{\Delta}(\mathbf{I} - \mathbf{M}_{11}\mathbf{\Delta})^{-1}\mathbf{M}_{12}. \tag{2.7}$$

Conditions for these operations to be well defined will be given in Section 2.2.

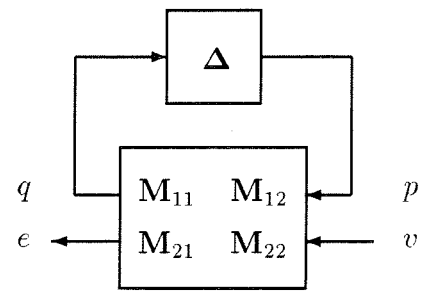


Figure 2.7: Robustness analysis configuration

2.2 Mathematical Formulation and Notation

This thesis adopts the discrete time setting for dynamical systems. A significant portion of the material can, however, be extended to continuous-time systems, as will be remarked throughout the presentation.

2.2.1 Signals and Operators

In this thesis, a signal is a vector-valued sequence, and for simplicity considered in general to be complex-valued (in $\mathbb{C}^n$); the restriction to real sequences will be indicated when necessary. The time index varies in the set of integers $\mathbb{Z}$, or sometimes in the non-negative subset $\mathbb{Z}^+$. We will use l_2 signal norms: l_2^n denotes the Hilbert space of square summable sequences over $\mathbb{Z}$ or $\mathbb{Z}^+$, with the inner product

$$\langle u, v \rangle := \sum_{t \in \mathbb{Z}} u^*(t) v(t), \tag{2.8}$$

where vector signals are represented in column form, and $*$ denotes conjugate transpose. The Fourier transform

$$v(e^{j\omega}) = \sum_{t \in \mathbb{Z}} v(t) e^{j\omega t} \tag{2.9}$$

provides an isomorphism between the complex space $l_2^n(\mathbb{Z})$ and the space of square integrable functions on the unit circle $\mathbb{T}$, with respect to the normalized Lebesgue measure $\frac{d\omega}{2\pi}$. This permits an identification between a signal $v(t)$ and its Fourier transform $v(e^{j\omega})$, with

$$\|v\|^2 = \sum_{t \in \mathbb{Z}} |v(t)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |v(e^{j\omega})|^2 dw. \tag{2.10}$$

In many situations one does not want to impose a priori that signals are norm bounded; one way to allow signals which "blow up at $+\infty$" is to introduce as in [96, 20, 89] the space $l_{2e} = \{v(t) : \mathbf{P^T} v \in l_2^n \ \ \forall T\}$, where $\mathbf{P^T}$ denotes the truncation operator

$$\mathbf{P^T} : v(t) \rightarrow \begin{cases} v(t) & \text{if} \quad t \leq T \\ 0 & \text{if} \quad t > T \end{cases} . \tag{2.11}$$

We now introduce operators between the signal spaces. Three important properties of these operators are causality, boundedness and linearity. An operator $\mathbf{H}$ on l_2 or l_{2e} is causal if $\mathbf{P^T H P^T} = \mathbf{P^T H}$ for all T.

An operator $\mathbf{H} : l_2 \rightarrow l_2$ is bounded if there exists a finite γ such that $\|\mathbf{H}v\|_2 \leq \gamma \|v\|_2$ for all $v \in l_2$. A bounded operator on l_2 which is causal can always be extended to l_{2e}, satisfying the boundedness property $\|\mathbf{P^T H} v\| \leq \gamma \|\mathbf{P^T} v\|$ for all $v \in l_{2e}$, and all T.

Linearity is defined in the usual vector space sense. From now on we will only consider linear operators (see the Remark on Section 2.1.2). These can be given an infinite matrix representation

$$\mathbf{H} = \begin{bmatrix} h_{00} & h_{01} & h_{02} & \cdots \\ h_{10} & h_{11} & h_{12} & \cdots \\ h_{20} & h_{21} & h_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{2.12}$$

where each $h_{ij} \in \mathbb{C}^{m \times n}$ ((2.12) is for a time index in $\mathbb{Z}^+$, for $\mathbb{Z}$ a doubly infinite matrix is used). Causality corresponds to a matrix of lower triangular form.

The class of linear, bounded operators $\mathbf{H} : l_2^n \to l_2^m$ is denoted $\mathcal{L}(l_2^n, l_2^m)$ or simply $\mathcal{L}(l_2)$. The subset of causal operators in $\mathcal{L}(l_2)$ is denoted $\mathcal{L}_c(l_2)$. Every operator in $\mathcal{L}_c(l_2)$ can be extended to l_{2e}; we will denote $\mathcal{L}_c(l_{2e})$ when the extended domain must be made explicit. The l_2-induced norm of an operator in $\mathcal{L}(l_2^n, l_2^m)$ is given by

$$\|\mathbf{H}\|_{l_2 \to l_2} := \sup_{\|v\|_2 = 1} \|\mathbf{H}v\|_2 = \sup_{T \in \mathbb{Z}^+} \bar{\sigma}\left(H_T\right), \tag{2.13}$$

where $\bar{\sigma}$ denotes maximum singular value, and H_T is the truncation of the infinite matrix to the interval $[-T, T]$ for a time index in $\mathbb{Z}$, or to $[0, T]$ for $\mathbb{Z}^+$. For example, in the latter case

$$H_T = \begin{bmatrix} h_{00} & \cdot & h_{0T} \\ \vdots & \ddots & \cdots \\ h_{T0} & \cdots & h_{TT} \end{bmatrix}. \tag{2.14}$$

The adjoint $\mathbf{H}^*$ of an operator $\mathbf{H} \in \mathcal{L}(l_2)$ is defined to satisfy the identity

$$\langle u, \mathbf{H}v \rangle = \langle \mathbf{H}^* u, v \rangle. \tag{2.15}$$

An operator $\mathbf{H}$ is linear, time invariant (LTI) if $\mathbf{H}\boldsymbol{\lambda} = \boldsymbol{\lambda}\mathbf{H}$, where $\boldsymbol{\lambda}$ is the unit delay operator. For an LTI operator, the infinite matrix representation is constant along diagonals, with $h_{t\tau} = h(t - \tau)$. Equivalently, the operator is described by a convolution kernel $h(t)$:

$$(\mathbf{H}v)(t) = \sum_{\tau = -\infty}^{\infty} h(t - \tau) v(\tau). \tag{2.16}$$

In the frequency domain, a bounded LTI operator $\mathbf{H} \in \mathcal{L}(l_2^n, l_2^m)$ can be identified with a multiplication operator $(\mathbf{H}v)(e^{j\omega}) = H(e^{j\omega})v(e^{j\omega})$ by a matrix transfer function $H(e^{j\omega})$ in the space

$$L_\infty^{m \times n}(\mathbb{T}) = \{F : \mathbb{T} \to \mathbb{C}^{m \times n}, \text{ such that } \|F\|_\infty = \operatorname*{ess\,sup}_{e^{j\omega} \in \mathbb{T}} \bar{\sigma}\left(F(e^{j\omega})\right) < \infty\}. \tag{2.17}$$

Also, $\|\mathbf{H}\|_{l_2 \to l_2} = \|H\|_\infty$. The causal subspace $\mathcal{L}_c(l_2^n, l_2^m)$ can be identified with the Hardy space $\mathcal{H}_\infty^{m \times n} \subset L_\infty^{m \times n}(\mathbb{T})$ of matrix-valued functions on $\mathbb{T}$ which admit an analytic continuation to the unit disk $\mathbb{D}$. Background on Hardy spaces can be found in [34] and references therein. If $H(\lambda)$, $\lambda \in \mathbb{D}$ is such an analytic function, the Taylor expansion is given by

$$H(\lambda) = \sum_{t=0}^{\infty} h(t)\lambda^t, \tag{2.18}$$

where the function $h(t)$ is the convolution kernel (impulse response matrix) in (2.16) corresponding to $\mathbf{H}$. The delay operator $\boldsymbol{\lambda}$ in l_2^1 has transfer function λ. We denote by $\mathcal{RL}_\infty$ ($\mathcal{RH}_\infty$) the subspace of $\mathcal{L}_\infty$ ($\mathcal{H}_\infty$) consisting of rational functions; the corresponding operators are called finite-dimensional LTI.

In some robustness analysis questions to be considered, we will find it useful to introduce classes of operators that lie in between $\mathcal{L}(l_2)$ and the LTI subset. Based on [64], we define the normalized set

$$\boldsymbol{B}^\nu := \{\boldsymbol{\Delta} \in \mathcal{L}(l_2) : \|\boldsymbol{\Delta}\| \leq 1, \|\boldsymbol{\lambda}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\lambda}\| \leq \nu\}. \tag{2.19}$$

The parameter ν provides a natural way of measuring the time variation of an operator. For $\nu = 0$, $\boldsymbol{B}^\nu$ is the unit ball of LTI operators; for $\nu = 2$ it includes arbitrary contractions. For intermediate values we will say that the operators in $\boldsymbol{B}^\nu$ have time variation slower than ν.

2.2.2 Systems, Feedback, Stability and Performance

This thesis is concerned with dynamical systems, which can be abstracted mathematically as a relationship between signals, i.e., variables evolving in time. The prevailing viewpoint in system and control theory is to consider this relationship to be of the cause-effect kind: some input variables, which are free to be manipulated, determine the evolution of the remaining output variables. In other words an (Input-Output) system is an operator as considered in Section 2.2.1. In the final chapters of this thesis we will consider an alternative viewpoint for systems, which does not take the input-output partition as a starting point. For the time being, however, we adhere to the standard formulation.

For the special case of a finite dimensional LTI system, one can also write an internal state-space representation, given by the difference equations

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t). \end{aligned} \tag{2.20}$$

Its transfer function is denoted by

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := D + C\lambda(I - \lambda A)^{-1}B = \lambda I \star S. \tag{2.21}$$

The rightmost expression in (2.21) is another example of an upper LFT, of the same form as (2.7), with $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$. The state space system is called internally stable if the spectral radius $\rho(A) < 1$. In this case the transfer function is in $\mathcal{RH}_\infty$.

Figure 2.8 depicts a standard feedback configuration between two input-output systems. The interconnection, with injected disturbance signals d_1, d_2, is equivalent to the equations

$$\begin{bmatrix} \mathbf{I} & -\mathbf{H} \\ -\mathbf{K} & \mathbf{I} \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}. \tag{2.22}$$

We now introduce the concepts of well posedness and stability of the interconnection; these are simplified versions of those given in [89] for more general nonlinear systems.

Definition 1 *The feedback interconnection of Figure 2.8 is well posed if for any d_1, d_2 in l_{2e}, there exist unique solutions u, y to the equations (2.22), which depend causally on d_1, d_2. Equivalently, the operator $\left[\begin{smallmatrix} \mathbf{I} & -\mathbf{H} \\ -\mathbf{K} & \mathbf{I} \end{smallmatrix}\right]$ has a causal inverse on l_{2e}. If in addition this inverse is bounded, i.e., in $\mathcal{L}_c(l_{2e})$, the feedback interconnection is said to be stable.*

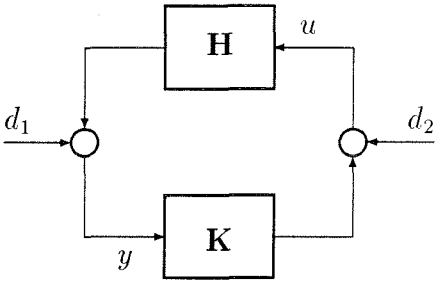


Figure 2.8: Feedback interconnection

The concept of stability applies analogously to other feedback configurations. For the one in Figure 2.6, we inject disturbance signals d_1, d_2 at the interconnection points between $\mathbf{G}$ and $\mathbf{K}$, and require that the map between d_1, d_2, z and all the remaining signals be in $\mathcal{L}_c(l_{2e})$. In particular, the LFT $\mathbf{M} = \mathbf{G} \star \mathbf{K}$ of (2.6) is well defined and in $\mathcal{L}_c(l_{2e})$.

In the sequel, assume that $\mathbf{G}$ and $\mathbf{K}$ are LTI and the interconnection is stable; then $M(\lambda)$ is in $\mathcal{H}_\infty$. We now give a more precise description of the performance objectives for the design of $\mathbf{K}$. As remarked before, it is standard to specify the performance as a measure of disturbance rejection.

One such measure is the $\mathcal{H}_2$ performance criterion. The $\mathcal{H}_2$ norm of an LTI system $\mathbf{H}$ is defined as

$$\|\mathbf{H}\|_2 := \left(\frac{1}{2\pi} \int_0^{2\pi} \text{trace}(H(e^{j\omega})^* H(e^{j\omega}))dw \right)^{\frac{1}{2}}. \tag{2.23}$$

This norm captures the response of the system to signals of flat spectrum (white signals), which arise from stationary white noise or from impulsive signals. The $\mathcal{H}_2$ control problem (also known as LQG control) consists of designing a stabilizing $\mathbf{K}$ for the configuration of Figure 2.6, which minimizes the $\mathcal{H}_2$ norm of the closed loop $\mathbf{M} = \mathbf{G} \star \mathbf{K}$.

The main alternative for disturbance rejection is to consider an induced operator norm, i.e. the worst case gain of a system in some signal norm. In our case of l_2 signal norms, the requirement is that the l_2-induced norm of $\mathbf{M}$ be small; in the LTI case this corresponds to the $\mathcal{H}_\infty$ norm defined in (2.17). The $\mathcal{H}_\infty$ control problem, motivated by Zames [97], is to find a stabilizing $\mathbf{K}$, minimizing the $\mathcal{H}_\infty$ norm of $\mathbf{M} = \mathbf{G} \star \mathbf{K}$. The other standard choice is $\mathcal{L}_1$ control (see [15]) which arises from the l_∞ norm for signals. These induced norm alternatives have been more popular in the robust control literature, since they combine more transparently with system uncertainty. A general discussion of the relative merits of these performance measures will be given in Chapters 4 and 6.

In the absence of uncertainty, these optimal control problems have been completely solved in the literature (see [3, 34, 26, 15]); this means that given a generalized plant $\mathbf{G}$, there is a computationally tractable method for constructing the controller $\mathbf{K}$ in regard to these performance measures applied to $\mathbf{G} \star \mathbf{K}$.

2.2.3 Uncertainty and Robustness

Returning to the system of Figure 2.5, we now characterize the perturbation $\mathbf{\Delta}$ which encompasses the system uncertainty. The motivation of Section 2.1 leads us to consider a structured $\mathbf{\Delta}$ of block diagonal form,

$$\mathbf{\Delta} = \text{diag}\left[\delta_1 \mathbf{I}_{n_1}, \ldots, \delta_L \mathbf{I}_{n_L}, \mathbf{\Delta}_{L+1}, \ldots, \mathbf{\Delta}_{L+F}\right]. \tag{2.24}$$

The notation of (2.24) represents the fact that the blocks act independently on the spatial components of a vector-valued signal. More precisely, if $q = \text{col}(q_1, \ldots, q_{L+F})$ denotes the partition of the column vector signal q in dimensions compatible with the blocks in $\mathbf{\Delta}$, then (2.24) means that $p = \mathbf{\Delta} q = \text{col}(\delta_1 \mathbf{I}_{n_1} q_1, \ldots, \mathbf{\Delta}_{L+F} q_{L+F})$.

The blocks in $\mathbf{\Delta}$ can in general represent real parameters or dynamic (LTI, linear time-varying, nonlinear) perturbations. Two types of uncertainty blocks are included in (2.24): full blocks $\mathbf{\Delta}_{L+j}$ which are general multivariable operators, and repeated scalar blocks $\delta_i \mathbf{I}_{n_i}$ where the same scalar operator acts on each of the channels. These repeated perturbations

are useful for representing rational dependence on an uncertain parameter, or for the case when the frequency variable λ is represented in LFT form as in (2.21).

The results in this thesis are based on linear techniques and apply in principle to linear perturbation structures alone. Many of the results admit extensions to the case of nonlinear operators, following the general Remark made in Section 2.1.2. This issue will not be pursued further in this thesis; from now on, all perturbation structures will be a subset of the class of structured linear operators, given by

$$\boldsymbol{\Delta}_{\boldsymbol{NC}}^{\boldsymbol{LTV}} = \{\boldsymbol{\Delta} \in \mathcal{L}(l_2) : \boldsymbol{\Delta} = \mathrm{diag}\,[\delta_1 \mathbf{I}_{n_1}, \ldots, \delta_L \mathbf{I}_{n_L}, \boldsymbol{\Delta}_{L+1}, \ldots, \boldsymbol{\Delta}_{L+F}]\}. \tag{2.25}$$

The subscript $\boldsymbol{NC}$ means that the class contains non-causal perturbations. In robustness problems, we will usually restrict perturbations to be causal (an exception is in Chapter 7, see also the discussion in Appendix B), writing

$$\boldsymbol{\Delta}^{\boldsymbol{LTV}} = \{\boldsymbol{\Delta} \in \mathcal{L}_c(l_2) : \boldsymbol{\Delta} = \mathrm{diag}\,[\delta_1 \mathbf{I}_{n_1}, \ldots, \delta_L \mathbf{I}_{n_L}, \boldsymbol{\Delta}_{L+1}, \ldots, \boldsymbol{\Delta}_{L+F}]\}. \tag{2.26}$$

The superscript LTV (linear, time varying) is included to highlight the distinction with the subset of LTI structured operators,

$$\boldsymbol{\Delta}^{\boldsymbol{LTI}} = \{\boldsymbol{\Delta} \in \boldsymbol{\Delta}^{\boldsymbol{LTV}} : \boldsymbol{\lambda\Delta} = \boldsymbol{\Delta\lambda}\}. \tag{2.27}$$

In correspondence with (2.19), we introduce the intermediate class

$$\boldsymbol{\Delta}^{\boldsymbol{\nu}} = \{\boldsymbol{\Delta} \in \boldsymbol{\Delta}^{\boldsymbol{LTV}} : \|\boldsymbol{\lambda\Delta} - \boldsymbol{\Delta\lambda}\| \leq \nu\}. \tag{2.28}$$

Other uncertainty classes will also be considered, with different combinations of these operator classes, or some blocks restricted to be real parameters, always with the spatial structure (2.24). The generic notation $\boldsymbol{\Delta}$ will be used for the class under consideration, which will be made clear in each case. The normalized ball of uncertainty, in the sense of the l_2-induced norm, is denoted by $\boldsymbol{B}_{\boldsymbol{\Delta}} = \{\boldsymbol{\Delta} \in \boldsymbol{\Delta} : \|\boldsymbol{\Delta}\|_{l_2 \to l_2} \leq 1\}$.

Assume now that $\mathbf{K}$ has been chosen, and stabilizes the system in the absence of uncertainty, so that the closed map $\mathbf{M} = \mathbf{G} \star \mathbf{K}$, between the inputs $z = \mathrm{col}(p, v)$ and the outputs $w = (q, e)$ is in $\mathcal{L}_c(l_2)$. The resulting robustness analysis configuration is given in Figure 2.7.

The first question to address is whether stability is preserved in the presence of uncertainty; this leads to the following definition:

Definition 2 *Suppose that* $\mathbf{M} \in \mathcal{L}_c(l_2)$. *The system of Figure 2.7 is robustly stable if* $\mathbf{I} - \boldsymbol{\Delta}\mathbf{M}_{11} : l_2 \to l_2$ *has a causal, bounded inverse for all* $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}}$. *Robust stability is uniform if*

$$\sup_{\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}}} \|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M}_{11})^{-1}\| < \infty. \tag{2.29}$$

The previous definition of robust stability is equivalent to requiring stability in the sense of Definition 1 for all perturbations in the normalized ball of uncertainty; this follows from the fact that $\mathbf{M}$ and $\mathbf{\Delta}$ are themselves bounded operators. Some technical results require the mild strengthening of uniformity given in (2.29); in most cases, however, this condition is no stronger than robust stability (see Appendix B).

If robust stability holds, the closed loop map from v to e is well defined for all $\mathbf{\Delta}$ and given by the upper LFT $\mathbf{\Delta} \star \mathbf{M}$ as in (2.7). A performance specification can then be imposed on the map $\mathbf{\Delta} \star \mathbf{M}$. The system will have robust performance if the specification holds for all $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}$. For the case of $\mathcal{H}_2$ performance, the robustness problem is the subject of Chapter 6. The following definition applies to the more established $\mathcal{H}_\infty$ case.

Definition 3 *The uncertain system is said to have robust $\mathcal{H}_\infty$ performance if it is robustly stable, and*

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}} \|\mathbf{\Delta} \star \mathbf{M}\|_{l_2 \to l_2} < 1. \tag{2.30}$$

Remark: Strictly speaking, the l_2 induced norm can be called $\mathcal{H}_\infty$ only in the case where $\mathbf{\Delta} \star \mathbf{M}$ is LTI. Definition 3 involves therefore a language abuse for non-LTI uncertainty.

2.3 Robustness Analysis

Robust stability and performance analysis under structured uncertainty has been the focus of extensive research; this section contains a (by no means exhaustive) summary of previous work related to this thesis.

2.3.1 The Small Gain Theorem

The small-gain theorem was obtained by Zames [96] and Sandberg [75] in the 1960s as a criterion for assessing stability of nonlinear feedback systems based on input-output information. This point of view has proven to be particularly adequate for analysis of uncertain systems, as was already suggested in [96]. In fact variations of the small-gain argument lie at the heart of most methods for robustness analysis.

Consider the feedback configuration of Figure 2.9, where we assume that $\mathbf{M}$ and $\mathbf{\Delta}$ are in $\mathcal{L}_c(l_2)$[1]. The small-gain theorem states that if $\|\mathbf{M}\|\|\mathbf{\Delta}\| < 1$, then $\mathbf{I} - \mathbf{\Delta}\mathbf{M}$ is invertible in $\mathcal{L}_c(l_2)$ and therefore the feedback interconnection is stable. The result holds for any induced system norm, in our case for the l_2-induced norm.

[1] Nonlinear systems could also be considered, see [96].

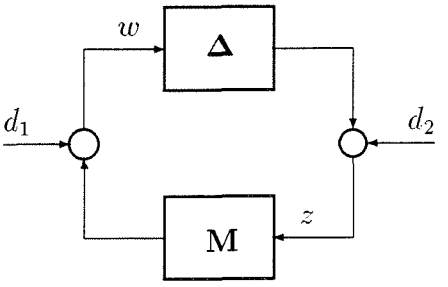


Figure 2.9: Setup for robust stability analysis

In the case where $\mathbf{\Delta}$ is representing uncertainty, normalized to $\|\mathbf{\Delta}\| \leq 1$, then a consequence of the small gain theorem is that $\|\mathbf{M}\| < 1$ is a sufficient condition for robust stability, regardless of the structure of the uncertainty. The condition corresponds in fact to the case of unstructured uncertainty, but may be conservative if there is structure in $\mathbf{\Delta}$.

One standard consequence of the small gain theorem is that robust $\mathcal{H}_\infty$ performance can be converted to robust stability with an additional uncertainty block. If the system in Figure 2.10(a) has robust $\mathcal{H}_\infty$ performance, then for every $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}$, $\|\mathbf{\Delta} \star \mathbf{M}\|_{l_2 \to l_2} < 1$; from small-gain, the configuration of Figure 2.10(b) is robustly stable, where $\mathbf{\Delta}_P$ is an arbitrary contractive operator in $\mathcal{L}_c(l_2)$.

If $\mathbf{M}$ is LTI, the converse implication also holds under fairly general assumptions on $\mathbf{\Delta}$. If robust performance fails on Figure 2.10(a), $\mathbf{\Delta}_P$ can be chosen to violate robust stability in Figure 2.10(b). In the case of $\mathbf{\Delta}$ LTI, this follows from a frequency domain argument (see [98]); for the LTV case see [43, 76] or the results in Appendix B.

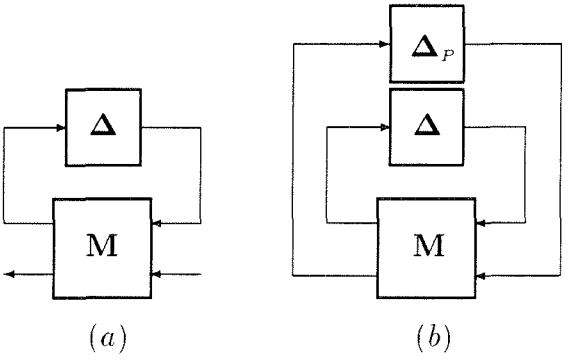


Figure 2.10: Robust performance restated as robust stability

As a consequence, a robust stability theory under structured uncertainty, directly encompasses the corresponding robust $\mathcal{H}_\infty$ performance theory.

2.3.2 Constant Matrix Analysis and μ

As we have seen, robust stability analysis amounts to testing the invertibility in $\mathcal{L}_c(l_{2e})$ of an operator $\mathbf{I} - \mathbf{\Delta M}$, where $\mathbf{M}$ is LTI and stable, and $\mathbf{\Delta}$ is a structured operator as in (2.24). Also, robust $\mathcal{H}_\infty$ performance reduces to a similar condition.

The objective is to provide non conservative tests for robust stability which take in to account the structure (2.24). The natural starting point is to study the constant matrix case $M, \Delta \in \mathbb{C}^{n\times n}$; as a convention, the boldface font is reserved for operators, and normal font is used for matrices. For example, $\mathit{\Delta} = \{\mathrm{diag}\,[\delta_1 I_{n_1}, \ldots, \delta_L I_{n_L}, \Delta_{L+1}, \ldots, \Delta_{L+F}]\} \subset \mathbb{C}^{n\times n}$ denotes the set of structured matrices, and we investigate the invertibility of $I - M\Delta$ with $\Delta \in \mathit{\Delta}$. This leads to the definition of the structured singular value μ, introduced by Doyle in [28] (see also [55, 31, 95]):

Definition 4 *The structured singular value $\mu_{\mathit{\Delta}}(M)$ of a matrix M with respect to a structure $\mathit{\Delta}$ is defined as $\mu_{\mathit{\Delta}}(M) := 0$ if no $\Delta \in \mathit{\Delta}$ makes $I - \Delta M$ singular, otherwise*

$$\mu_{\mathit{\Delta}}(M) := \frac{1}{\min\{\bar{\sigma}(\Delta) : \Delta \in \mathit{\Delta}, \det(I - \Delta M) = 0\}}. \tag{2.31}$$

Thus $I - \Delta M$ is invertible for all $\Delta \in B_{\mathit{\Delta}}$ if and only if $\mu_{\mathit{\Delta}}(M) < 1$. Two different cases are of interest: complex μ [55] for complex Δ, and real or mixed μ [31, 95], where some of the blocks in Δ are restricted to be real (see (2.38) below). We shall use a common generic notation μ; the distinction will be made explicit whenever is necessary. Complex μ is always a continuous function of M; mixed μ is not always continuous, but continuity can be shown (see [57]) if there is complex uncertainty entering the problem in a non-trivial way. The following basic property relating μ to LFTs is known as the *main loop theorem* (see [55]):

Lemma 2.1 *Given a block structure $\mathit{\Delta} = diag[\mathit{\Delta}_1, \mathit{\Delta}_2]$ and a matrix M, suitably partitioned as $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$, then the following are equivalent:*

$$\mu_{\mathit{\Delta}}(M) < 1; \tag{2.32}$$

$$\mu_{\mathit{\Delta}_1}(M_{11}) < 1, \max_{\bar{\sigma}(\Delta_1)\le 1} \mu_{\mathit{\Delta}_2}(\Delta_1 \star M) < 1; \tag{2.33}$$

$$\mu_{\mathit{\Delta}_2}(M_{22}) < 1, \max_{\bar{\sigma}(\Delta_2)\le 1} \mu_{\mathit{\Delta}_1}(M \star \Delta_2) < 1, \tag{2.34}$$

where $\Delta_1 \star M$ and $M \star \Delta_2$ are defined as in (2.7) and (2.6), respectively.

Another useful property (see [55]) is that in the case of complex perturbations, the maxima in (2.33) and (2.34) are achieved for unitary Δ_i; this resembles the maximum modulus theorem for analytic functions.

We now introduce an upper bound for the structured singular value, based on a small-gain property refined by scaling matrices which commute with the structure Δ. The set of matrices with this property is denoted by

$$\mathbb{X} = \{X \in \mathbb{C}^{n\times n} : X = \text{diag}\,[X_1, \ldots, X_L,\, x_{L+1} I_{m_1}, \ldots, x_{L+F} I_{m_F}]\}. \tag{2.35}$$

The subsets of invertible, hermitian and positive definite matrices in $\mathbb{X}$ will be denoted respectively by $\mathbb{X}^I$, $\mathbb{X}^H$, and $\mathbb{X}^P$. Define

$$\hat{\mu}_\Delta(M) := \inf_{X\in\mathbb{X}^I} \bar{\sigma}\left(XMX^{-1}\right) = \inf_{X\in\mathbb{X}^P} \bar{\sigma}\left(XMX^{-1}\right). \tag{2.36}$$

Then $\mu_\Delta(M) \le \hat{\mu}_\Delta(M)$. Equivalently, the Linear Matrix Inequality (LMI) condition

$$\exists X \in \mathbb{X}^P : \quad M^*XM - X < 0 \tag{2.37}$$

is sufficient for $\mu_\Delta(M) < 1$. This convex feasibility condition is attractive for computation, as will be discussed below. It can also be refined for the case of mixed uncertainty structures. If Δ is of the form

$$\Delta = \text{diag}\,[\delta_1 I, \ldots, \delta_{L_R} I, \delta_{L_R+1} I \ldots, \delta_{L_R+L_C} I, \Delta_{L+1}, \ldots, \Delta_{L+F}], \tag{2.38}$$

where $\delta_1 \ldots \delta_{L_R} \in \mathbb{R}$ (i.e. the first L_R scalar times identity blocks are real), then condition (2.37) can be tightened (see [31, 95]) to

$$\exists X \in \mathbb{X}^P,\ G \in \mathcal{G} : \quad M^*XM - X + j(M^*G - G^*M) < 0, \tag{2.39}$$

where the matrices G are of the form $G = \text{diag}[G_1, \ldots, G_{L_R}, 0, \ldots, 0]$ with $G_i = G_i^*$.

2.3.3 Robust Stability Tests

With the notation developed from the constant matrix case, we now summarize known conditions for analysis of the robust stability question in Figure 2.9, where $\mathbf{M}$ is assumed to be always a finite dimensional LTI, stable system.

μ Conditions

The first result (see [55, 98]) is that robust stability over the set $\boldsymbol{B_{\Delta LTI}}$ is equivalent to the complex μ-test across frequency

$$\max_\omega \mu_\Delta(M(e^{j\omega})) < 1. \tag{2.40}$$

An analogous condition holds if the uncertainty structure consists of a combination of real parametric and LTI perturbations, with complex μ replaced by mixed μ.

Condition (2.40) involves a test across frequency; given a state space realization of $M(\lambda) = \lambda I \star S$ as in (2.21), an equivalent μ-test on a single matrix given in [55] is

$$\mu_{\Delta_S}(S) < 1, \tag{2.41}$$

where the matrices in Δ_S are of the form $\Delta_S = \text{diag}[\lambda I, \Delta]$. This condition follows from applying Lemma 2.1, and the maximum modulus-type property of μ, to the test (2.40).

Convex Conditions and their Interpretation

We now apply the μ-upper bound to conditions (2.40) and (2.41). In the case of (2.40), we have the option of using constant or frequency dependent scaling matrices. This gives the following two tests from [55]:

$$\inf_{X \in \mathbb{X}^I} \max_{\omega} \bar{\sigma}\left(X M(e^{j\omega}) X^{-1}\right) < 1, \tag{2.42}$$

$$\max_{\omega} \hat{\mu}_{\Delta}(M(e^{j\omega})) = \max_{\omega} \inf_{X(e^{j\omega}) \in \mathbb{X}^I} \bar{\sigma}\left(X(e^{j\omega}) M(e^{j\omega}) X(e^{j\omega})^{-1}\right) < 1. \tag{2.43}$$

Alternatively, (2.43) can be written as

$$\inf_{X \in \boldsymbol{X}^I} \|X M X^{-1}\|_{\infty} < 1, \tag{2.44}$$

where we define the set of scaling transfer functions

$$\boldsymbol{X} := \{X : \mathbb{T} \rightarrow C^{n \times n} : X(e^{j\omega}) \in \mathbb{X} \text{ for each } \omega\}. \tag{2.45}$$

It can be shown (see [64]) that without loss of generality, the functions in $\boldsymbol{X}$ can be restricted to be rational functions $X(\lambda)$ such that X and X^{-1} are in $\mathcal{RH}_{\infty}$. The test (2.42) is a special case of (2.44) when we additionally impose that X is constant.

The state-space test (2.41) gives the upper bound

$$\inf_{X_S \in \mathbb{X}_S^I} \bar{\sigma}\left(X_S S X_S^{-1}\right) < 1, \tag{2.46}$$

where the matrices in $\mathbb{X}_S$ are of the form $X_S = \text{diag}[X_0, X]$, with X_0 a full matrix of the dimension of the state space, and $X \in \mathbb{X}$.

It is shown in [55] that (2.46) is equivalent to (2.42); also, (2.42) implies (2.43) which in turn implies (2.40), so conditions (2.42, 2.43, 2.46) are all sufficient for robust stability under LTI uncertainty. Also, for mixed real parametric/LTI perturbations, these conditions can be tightened by use of G-scales as in (2.39).

Recent results have shown that the constant scales test (2.42), or equivalently (2.46) has a very natural interpretation, as a necessary and sufficient condition for robust stability under the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$. The necessity was shown independently by Shamma [76] and Megretski [50] for the case of full blocks; for the case of δI blocks see Chapters 5 and 8.

Turning to the frequency dependent test (2.43) Poolla and Tikku [64] have recently shown that (2.43) is necessary and sufficient for robust stability against the class of *arbitrarily slowly varying* structured perturbations, in the sense introduced in (2.19). More precisely, (2.43) holds if and only if there exists $\nu > 0$ such that the system is robustly stable with uncertainty in $\boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}}$. This result is striking because it is known that for a sufficient number of uncertainty blocks, conditions (2.40) and (2.43) are different: robust stability conditions are in some sense discontinuous at $\nu = 0$.

Remark: All the frequency domain tests presented above have direct counterparts for the case of continuous time systems. As for the state space tests, (2.41) and (2.46) apply only to discrete time since they rely on the fact that λ varies in the unit disk, but state-space tests in continuous time can be obtained by a bilinear frequency transformation.

Computational Issues

The practical value of a mathematical condition such as those presented above is largely determined by the tractability of the resulting computation. In regard to robustness analysis, it can be generally said that the more specialized the uncertainty structure, the harder the computation.

In particular, conditions in terms of the structured singular value μ have been shown to be NP hard, in the sense of the theory of computational complexity [36]. This is well established (see [11]) for real and mixed μ, and recent results [84] suggest that this also applies to purely complex μ. It is generally accepted (see [36]) that NP hardness means that the condition cannot be computed exactly in the worst case without entirely unacceptable growth in computation cost with problem size. Hence for moderately large problems, one is forced to relax the requirement for exact computation and rely on upper and lower bounds.

Efficient lower bound algorithms for both complex and mixed μ have been developed by many researchers (see [55, 95, 53, 5]); although they have no guaranteed performance, they appear to behave well on empirical tests.

The upper bounds for μ of a constant matrix can be stated in terms of the Linear Matrix Inequality (LMI) conditions (2.37) or (2.39). The evaluation of such conditions is a convex feasibility problem which has been extensively studied in recent years [10, 35]; efficient interior point methods are available, which have moderate growth with problem size. This

means that a robustness test which is equivalent to a single LMI, such as (2.46) for robust stability over $\boldsymbol{B}_{\boldsymbol{\Delta LTV}}$, can be evaluated efficiently for moderate problem sizes.

In comparison, the frequency dependent test (2.43) is equivalent to an infinite dimensional convex feasibility problem, given by the LMI across frequency

$$\forall \omega,\ \exists X(e^{j\omega}) \in \mathbb{X}^P :\quad M(e^{j\omega})^* X(e^{j\omega}) M(e^{j\omega}) - X(e^{j\omega}) < 0. \tag{2.47}$$

There are two standard approaches for handling the infinite dimensionality of this type of conditions, and turn them into finite dimensional LMIs. One way is to impose (2.47) at a grid of frequencies; this yields a series of decoupled LMI problems. Although this approximation offers no hard guarantees, since it is based on the frequency domain it allows for engineering judgement to be used in choosing the number and location of the grid-points.

The alternative is to select a finite set of rational basis functions for $X(e^{j\omega})$ and restrict the search to the span of these functions. Condition (2.47) will then depend linearly on a finite number of unknowns, and fixed rational functions of frequency. An application of the Positive Real Lemma (see, e.g. [10]) converts this condition to a single LMI in state space; this procedure offers a guaranteed sufficient condition for (2.47), but is computationally intensive since the problem is coupled. For more details see [4].

Given these computational properties, and the interpretation for the test (2.47) in terms of arbitrarily slowly varying uncertainty, it may be questioned whether μ-analysis conditions such as (2.40) should be used at all. There are still good reasons, however, to formulate a problem in terms of μ. In the first place, although the upper bounds have guaranteed polynomial-time computation, the size of the problems can be very large and render the computation impractical. In these cases one often relies heavily on lower bound algorithms to compute the analysis. More importantly, if there is parametric uncertainty in the problem, the upper bounds may be substantially conservative (there is no corresponding slowly-varying interpretation). Lower bound algorithms provide a fast method to obtain "bad" parameter values, and can be further employed to assess this conservatism and, if desired, pursue a more refined analysis by branch and bound techniques (see [53]).

2.4 Robust Synthesis

In this section we briefly discuss the synthesis problem of designing the controller $\mathbf{K}$ to achieve robust performance in Figure 2.5. The standard theory refers to robust $\mathcal{H}_\infty$ performance, which, given the results of Section 2.3.1, reduces to robust stability. Therefore we will only discuss robust stabilization: finding $\mathbf{K}$ such that the system of Figure 2.9 is robustly stable, where $\mathbf{M} = \mathbf{G} \star \mathbf{K}$.

A first condition for synthesis is provided by the small-gain theorem: if $\mathbf{K}$ is designed to give $\|G \star K\|_\infty < 1$, then it is robustly stabilizing. This implies that the $\mathcal{H}_\infty$ control problem can be used for robust synthesis.

If one wishes to exploit the structure of the uncertainty, however, then the synthesis problem becomes harder. For example, (2.44) leads to the synthesis problem

$$\inf_{X,K} \|X(G \star K)X^{-1}\|_\infty \ . \tag{2.48}$$

No efficient algorithm is available for the global minimization in (2.48), which is not a convex optimization problem, except in very special cases (see e.g. [68, 18]).

In the general case, there is strong indication that this problem is hard since it is closely related to bilinear matrix inequalities (BMIs) which are known to be NP hard. The main general procedure for attacking this problem is by iteration (see, e.g., [80]) between the following two tractable problems:

1. For fixed X, optimize in K which is an $\mathcal{H}_\infty$ control problem.

2. For fixed K, obtain new scales by solving the analysis problem. In the constant scales case (2.42), this amounts to solving an LMI. In the case of (2.43), one must find a rational scaling function: this is done by fitting points obtained from a grid, or in terms of basis functions.

This procedure is guaranteed to reduce the cost at every step, provided there are no restrictions in order of $X(\lambda)$ for the frequency-dependent case, but not to converge to the global minimum. Therefore it is best seen as a tool to improve initial designs, rather than a procedure to obtain blindly a robust controller.

Given the difficulties in the structured synthesis problem, it is questionable whether such a specialized design methodology is warranted in a practical system. At this point in time, it appears that the methods of robust control theory under highly structured perturbations are best employed as analysis tools, to validate designs based on simpler uncertainty descriptions or more heuristic methods.

Chapter 3

Robustness Analysis of Combined Dynamic Uncertainty Classes

Different classes of uncertainty description were introduced in Chapter 2 and motivated at the component modeling stage. When a full system model is constructed from subsystems as exemplified in Section 2.1.3, one obtains a decentralized perturbation as considered in Section 2.2.3. In addition to the spatial structure, it is natural to expect a combination of the various types of uncertainty blocks: real parameters, linear time invariant (LTI) dynamic perturbations, and linear time-varying (LTV) parameters or operators.

Most of the robust stability tests reviewed in Section 2.3 refer to uncertainty structures ($\boldsymbol{\Delta}^{LTI}$, $\boldsymbol{\Delta}^{LTV}$, $\boldsymbol{\Delta}^{\nu}$) where all the blocks are of the same dynamic type. The only exception to this considered in Section 2.3.3 is the case of combined real parameters and LTI perturbations, which is captured by a mixed-μ test [31, 95].

This chapter will consider the general case of mixed linear uncertainty classes, in particular where LTV blocks are also present in the combination. The inclusion of time-varying uncertainty precludes a straightforward frequency domain μ-analysis as in mixed LTI/parametric problems. For this reason, we propose an augmentation procedure across a number of frequencies, which is related to a lifting technique for μ analysis by Bercovici et al. [7], and to a power distribution lemma from Poolla and Tikku [64].

The main result in this chapter, presented in Section 3.1, is to show that this augmented μ test is a necessary and sufficient condition for robust stability under combined LTV/LTI uncertainty. The augmentation also provides an alternative for the formulation of convex upper bounds, discussed in Section 3.2. Real parametric uncertainty can also be incorporated in this procedure, as shown in Section 3.3. Some examples which demonstrate the results are given in Section 3.4.

As before, we will concentrate on the question of robust stability. The same conditions apply to robust $\mathcal{H}_\infty$ performance problems, for the reasons explained in Section 2.3.1.

3.1 A μ Test for Mixed LTV/LTI Analysis

3.1.1 A Power Distribution Lemma

The following lemma from Poolla and Tikku [64] provides a useful characterization of time varying perturbations, which will be used in this chapter.

Lemma 3.1 *Let $0 \leq \omega_1 < \ldots < \omega_r \leq \pi$ be distinct frequencies. If the vector signals*

$$q(t) = \sum_{k=1}^{r} q_k e^{j\omega_k t}, \qquad p(t) = \sum_{k=1}^{r} p_k e^{j\omega_k t} \tag{3.1}$$

satisfy the power inequality $\sum_{k=1}^{r} |q_k|^2 \geq \sum_{k=1}^{r} |p_k|^2$, then there exists a linear time-varying, causal operator $\boldsymbol{\Delta} \in \mathcal{L}_c(l_{2e})$ such that

$$\begin{aligned} &(i) \quad \|\boldsymbol{\Delta}\| \leq 1, \\ &(ii) \quad \|\boldsymbol{\lambda}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\lambda}\| \leq \nu = 2\sin(\frac{\omega_r - \omega_1}{2}), \\ &(iii) \quad \boldsymbol{\Delta} q = p + p^{tr},\ p^{tr} \in l_2. \end{aligned}$$

Heuristically, this lemma says that provided that the total power of q is greater than that of p, a contractive LTV operator can rearrange the power between frequencies, mapping q to p in steady state. In contrast, for a contractive LTI operator to verify (iii) would require an inequality in power at every frequency (i.e. $|q_k| \geq |p_k|$ for every k). Also, the lemma provides a bound on the required time variation ν in terms of the amount of frequency shifting performed, an issue which will be reconsidered later in Chapter 5.

3.1.2 Problem Formulation and Augmented Structures

Sections 3.1 and 3.2 refer to the system of Figure 2.9, where $\mathbf{M}$ is, as usual, a finite dimensional LTI, stable discrete time system. The uncertain perturbation $\boldsymbol{\Delta}$ is now assumed to contain a combination of LTV and LTI blocks. We will write

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Phi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\theta} \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{M}_\Phi & \mathbf{M}_{\Phi\theta} \\ \mathbf{M}_{\theta\Phi} & \mathbf{M}_\theta \end{bmatrix}, \tag{3.2}$$

where $\boldsymbol{\Phi}$ is a causal LTV operator in a structured class $\boldsymbol{B}_{\boldsymbol{\Phi}^{LTV}} \subset \mathcal{L}_c(l_2^m)$, and $\boldsymbol{\theta}$ is a causal LTI operator in $\boldsymbol{B}_{\boldsymbol{\theta}^{LTI}} \subset \mathcal{L}_c(l_2^r)$. As in Chapter 2, the corresponding constant matrices and their structures will be represented by the same symbol in normal font, rather than boldface. $\mathbb{X}^\Phi$, $\mathbb{X}^\theta$ will denote the sets of scaling matrices corresponding to each structure.

An important integer parameter determined by the structure is the dimension d of the real vector space $\mathbb{X}^{\Phi_H}$ of hermitian scaling matrices in $\mathbb{X}^{\Phi}$. In the most common case where Φ consists only of full blocks, d is the number of blocks.

The objective is to obtain an extension of the exact μ condition (2.40) for LTI analysis, to the case of mixed LTV/LTI perturbations. Since LTV structures are usually characterized by tests (2.42), (2.46) in terms of $\hat{\mu}$, rather than μ, it is not obvious that a μ test would exist for the mixed LTV/LTI case (and in particular, for the LTV case). The main idea to obtain this μ condition is inspired in work by Bercovici et al. [7], where an augmentation or *lifting* in the structure converts the upper bound $\hat{\mu}$ to μ of a larger matrix. The results in [7] apply to constant matrices and are based on operator-theoretic methods, but can also be obtained as a corollary of the more general dynamic results to be presented in this chapter, which will be proved by convex analysis techniques.

We first consider the augmented matrix

$$\tilde{M}(\omega_1, \ldots, \omega_d) = \mathrm{diag}[M(e^{j\omega_1}), \ldots, M(e^{j\omega_d})] \tag{3.3}$$

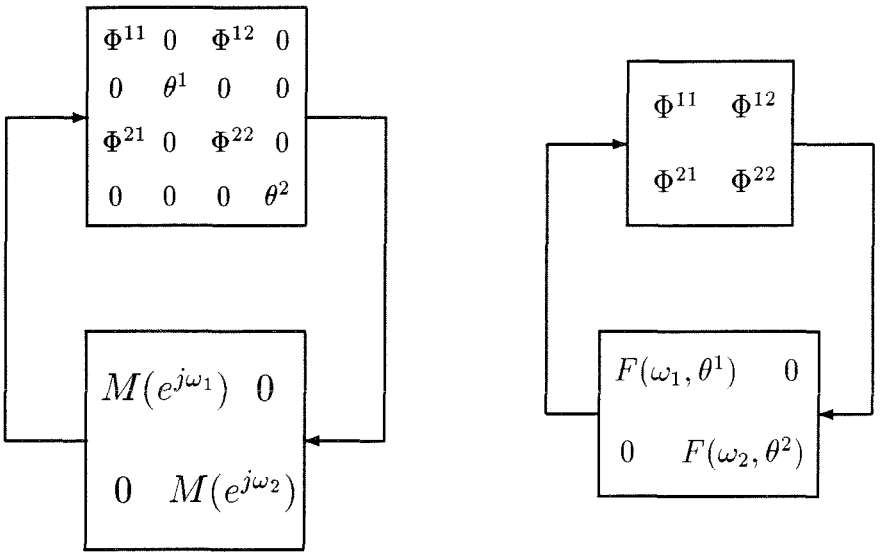
which represents the system $\mathbf{M}$ at a fixed number d of frequencies. Next we introduce the following augmented structures in the space of complex matrices:

$$\tilde{\Phi} = \begin{bmatrix} \Phi^{11} & \ldots & \Phi^{1d} \\ \vdots & \ldots & \vdots \\ \Phi^{d1} & \ldots & \Phi^{dd} \end{bmatrix}, \quad \Phi^{ij} \in \Phi, \quad i,j = 1 \ldots d, \tag{3.4}$$

$$\tilde{\Delta} = \begin{bmatrix} \Delta^{11} & \ldots & \Delta^{1d} \\ \vdots & \ldots & \vdots \\ \Delta^{d1} & \ldots & \Delta^{dd} \end{bmatrix}, \quad \begin{cases} \Delta^{ii} \in \Delta, & i = 1 \ldots d \\ \Delta^{ij} = \mathrm{diag}[\Phi^{ij},\, 0], & i \neq j \end{cases} . \tag{3.5}$$

The augmented structures $\tilde{\Phi}$, $\tilde{\Delta}$, are d times larger than the corresponding Φ, Δ. For the case of $\tilde{\Phi}$, it is obtained simply by considering d^2 copies of Φ, in matrix form. $\tilde{\Delta}$, which contains $\tilde{\Phi}$ as a submatrix, is obtained in a similar fashion, the only difference being that the time invariant blocks θ are only copied along the diagonal, and the rest of the entries are set to zero. As an illustration, Figure 3.1 (a) contains the augmented configuration for the case $d = 2$.

These structures have a "frequency shifting" interpretation which relates to the remarks made in regard to Lemma 3.1: the augmentation corresponds to considering a system at a number d of frequencies, and the different treatment of LTI and LTV blocks is due to the fact that only the time-varying perturbations are allowed to transfer power between frequencies; this is represented in $\tilde{\Phi}$ by the off-diagonal terms.



(a) (b)

Figure 3.1: Augmented representations

It is also convenient to consider the configuration of Figure 3.1 (b), where the LTI portion is included along with M. Define the matrix function

$$F(\omega, \theta) := M(e^{j\omega}) * \theta = M_\Phi + M_{\Phi\theta}\theta(I - M_\theta\theta)^{-1}M_{\theta\Phi}. \tag{3.6}$$

For this LFT to be well defined for $\theta \in B_\Theta$, the following condition must hold:

$$\max_\omega \mu_\Theta(M_\theta(e^{j\omega})) < 1. \tag{3.7}$$

Under this condition, $F(\omega, \theta)$ is continuous for $\omega \in [-\pi, \pi]$, $\theta \in B_\Theta$. For given (ω_k, θ_k), $k = 1 \ldots d$, the matrix

$$\tilde{F}(\omega_1, \ldots, \omega_d, \theta^1, \ldots, \theta^d) := \mathrm{diag}[F(\omega_1, \theta^1) \ldots F(\omega_d, \theta^d)] \tag{3.8}$$

is obtained by LFT between the θ portion of $\tilde{\Delta}$ and the matrix $\tilde{M}(\omega_1, \ldots, \omega_d)$. Figure 3.1 ($b$) represents $\tilde{F}$ and the remaining matrix $\tilde{\Phi}$, in the case $d = 2$.

3.1.3 Main Result

The main result in this chapter is to show that μ conditions on the augmented matrix structures provide an exact characterization for robust stability of the system (3.2). The proof will rely on Lemma 3.1, techniques from μ-analysis, and some convex analysis properties which can be found in Appendix A.

Theorem 3.2 *In reference to the system* (3.2)*, the following are equivalent:*

(a) The system is uniformly robustly stable.

(b) Condition (3.7) *holds, and with* ω_k *varying in* $[-\pi, \pi]$*, and* $\theta^k \in B_\Theta$*,*

$$\max_{\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d} \mu_{\tilde{\Phi}}(\tilde{F}(\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d)) < 1. \tag{3.9}$$

(c) With ω_k *varying in* $[-\pi, \pi]$*,*

$$\max_{\omega_1,\ldots,\omega_d} \mu_{\tilde{\Delta}}(\tilde{M}(\omega_1,\ldots,\omega_d)) < 1. \tag{3.10}$$

Proof: The equivalence of (b) and (c) is a simple consequence of the main loop theorem (Lemma 2.1). In fact, for each $\omega_1,\ldots,\omega_d$, Lemma 2.1 implies that

$$\mu_{\tilde{\Delta}}(\tilde{M}(\omega_1,\ldots,\omega_d)) < 1 \iff \begin{cases} \mu_\Theta(M_\theta(e^{j\omega_k})) < 1, & k = 1\ldots d \\ \max_{\theta^1,\ldots,\theta^d} \mu_{\tilde{\Phi}}(\tilde{F}(\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d)) < 1 \end{cases}. \tag{3.11}$$

Maximization over $\omega_1,\ldots,\omega_d$ gives (3.10) on the left hand side, and (3.7) and (3.9) on the right hand side; the maxima are achieved due to the continuity of complex μ.

The proof of the equivalence of (a) and (b) will be written for simplicity in the case of LTV uncertainty with only full blocks, of the form $\Phi = \mathrm{diag}[\Phi_1,\ldots\Phi_d]$. The extension to the case of $\boldsymbol{\delta I}$ blocks in $\boldsymbol{\Phi}$ mimics other proofs in this thesis (see Lemma 5.2, the remarks after Theorem 6.6, and Theorem 8.4).

The following lemma gives an interpretation of the uncertainty structure $\tilde{\Phi}$.

Lemma 3.3 *Assume* (3.7) *holds. Given* $\omega_1,\ldots,\omega_d$*,* $\theta^1,\ldots,\theta^d$ *the following are equivalent:*

(i) $\mu_{\tilde{\Phi}}(\tilde{F}(\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d)) \geq 1$*.*

(ii) $\exists\, p^1,\ldots,p^d \in \mathbb{C}^m$*, not all zero such that*

$$\sum_{k=1}^{d} |(F(\omega_k,\theta^k)p^k)_l|^2 \geq \sum_{k=1}^{d} |(p^k)_l|^2, \quad l = 1\ldots d. \tag{3.12}$$

Proof (Lemma): (i) is equivalent to the existence of $\tilde{\Phi} \in \tilde{\Phi}$, $\bar{\sigma}(\tilde{\Phi}) \leq 1$ and a column vector $p = \mathrm{col}(p^1,\ldots,p^d) \neq 0$ such that $(I - \tilde{\Phi}\tilde{F}(\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d))p = 0$, or equivalently

$$\begin{bmatrix} p^1 \\ \vdots \\ p^d \end{bmatrix} = \begin{bmatrix} \Phi^{11} & \ldots & \Phi^{1d} \\ \vdots & \ldots & \vdots \\ \Phi^{d1} & \ldots & \Phi^{dd} \end{bmatrix} \begin{bmatrix} F(\omega_1,\theta^1)p^1 \\ \vdots \\ F(\omega_d,\theta^d)p^d \end{bmatrix}. \tag{3.13}$$

Each Φ^{ij} in (3.13) is in turn block diagonal with d subblocks, which impose a partition on the column vectors p^k, $F(\omega_k,\theta^k)p^k$. The elements in this partition are henceforth denoted by the subindex l. The equations and variables in (3.13) can be reordered, putting together $p_l^1,\ldots,p_l^d$ and $(F(\omega_1,\theta^1)p^1)_l,\ldots,(F(\omega_d,\theta^d)p^d)_l$ for each l, which reduces the structure of $\tilde{\Phi}$ to d full blocks in this new order. Now $\bar{\sigma}(\tilde{\Phi}) \leq 1$ is equivalent to the norm inequalities (3.12). ■

$(a)\Longrightarrow(b)$

Condition (3.7) is clearly necessary for robust stability; if it did not hold, the standard result (2.40) implies that the system could be destabilized by LTI perturbations alone. Therefore $F(\omega,\theta)$ is well defined and continuous.

Assume by contradiction, that there exist $\omega_1,\ldots,\omega_d \in [-\pi,\pi]$, $\theta^1,\ldots,\theta^d \in B_\Theta$ with $\mu_{\tilde{\Phi}}(\tilde{F}(\omega_1,\ldots,\omega_d,\theta^1,\ldots,\theta^d)) \geq 1$. Lemma 3.3 yields $p^1,\ldots,p^d$ satisfying (3.12).

Fix $\epsilon > 0$; by continuity of F we can perturb the $\theta^1,\ldots,\theta^d$ to have $\bar{\sigma}(\theta^k) < 1$ (strict inequality), and the $\omega_1,\ldots,\omega_d$ to make them distinct, satisfying

$$\sum_{k=1}^{d} |(F(\omega_k,\theta^k)p^k)_l|^2 \geq (1-\epsilon)^2 \sum_{k=1}^{d} |(p^k)_l|^2, \quad l = 1\ldots d. \tag{3.14}$$

Since $\bar{\sigma}(\theta^k) < 1$, we are in the conditions of an interpolation result given in [29], which states that there exists a causal, stable, rational LTI perturbation $\boldsymbol{\theta} \in \boldsymbol{B_\Theta}$ [1], satisfying $\theta(e^{j\omega_k}) = \theta^k$. Introduce

$$p(t) = \sum_{k=1}^{d} p^k e^{-j\omega_k t}, \quad q(t) = \sum_{k=1}^{d} (F(\omega_k,\theta^k)p^k)e^{-j\omega_k t}, \tag{3.15}$$

and define $\mathbf{F} = \mathbf{M} \star \boldsymbol{\theta}$, which has transfer function $F(w,\theta(e^{j\omega}))$ from (3.6). Then $\mathbf{F}p = q + e$, where $e(t) \in l_2$ is a transient term. For any l, (3.14) implies that the power of the l-th component of $q(t)$, is greater than the power of the l-th component of $p(t)$ times $(1-\epsilon)$. We can invoke Lemma 3.1 to show the existence of a causal time-varying operator $\boldsymbol{\Phi}_l$ which maps $q(t)_l$ to $(1-\epsilon)p(t)_l$, up to a transient term. Constructing $\boldsymbol{\Phi} = \mathrm{diag}[\boldsymbol{\Phi}_1 \ldots \boldsymbol{\Phi}_d]$, we obtain

$$(\mathbf{I} - \boldsymbol{\Phi}\mathbf{F})p = \epsilon p + \bar{e}, \tag{3.16}$$

where $\bar{e}(t) \in l_2$ is transient. This implies that

$$\|(\mathbf{I} - \boldsymbol{\Phi}(\mathbf{M} \star \boldsymbol{\theta}))^{-1}\| = \|(\mathbf{I} - \boldsymbol{\Phi}\mathbf{F})^{-1}\| \geq \frac{1}{\epsilon}. \tag{3.17}$$

[1] [29] constructs a perturbation in the disk algebra $\mathcal{A}(\mathbb{T})$ interpolating a countable number of frequencies; in the case of a finite number of frequencies it can be chosen to be rational.

Now, it follows routinely from (3.2) that

$$(\mathbf{I}-\boldsymbol{\Delta}\mathbf{M})^{-1}=\begin{bmatrix}\mathbf{I}-\boldsymbol{\Phi}(\mathbf{M}\star\boldsymbol{\theta}) & \mathbf{0}\\ -\boldsymbol{\theta}\mathbf{M}_{\theta\Phi} & \mathbf{I}-\boldsymbol{\theta}\mathbf{M}_{\theta}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{I} & \boldsymbol{\Phi}\mathbf{M}_{\Phi\theta}(\mathbf{I}-\boldsymbol{\theta}\mathbf{M}_{\theta})^{-1}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}. \tag{3.18}$$

Also, condition (3.7) implies that

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{B_\Theta}}\|(\mathbf{I}-\boldsymbol{\theta}\mathbf{M}_{\theta})^{-1}\|<\infty, \tag{3.19}$$

which implies that the second term on the right hand side of (3.18), and its inverse, are uniformly bounded in norm.

Letting $\epsilon\rightarrow 0$, the norm in (3.17) goes to infinity, which implies $\sup_{\Delta}\|(\mathbf{I}-\boldsymbol{\Delta}\mathbf{M})^{-1}\|=\infty$, violating uniform robust stability.

$(b)\Longrightarrow(a)$:

We introduce some more notation. For $\omega\in[-\pi,\pi]$, $\theta\in B_\Theta\subset\mathbb{C}^{p\times p}$, $p\in\mathbb{C}^m$, define

$$\sigma_l(\omega,\theta,p)=|(F(\omega,\theta)p)_l|^2-|p_l|^2,\qquad l=1\ldots d, \tag{3.20}$$

$$\Lambda(\omega,\theta,p)=(\sigma_1(\omega,\theta,p),\ldots,\sigma_d(\omega,\theta,p))\in\mathbb{R}^d, \tag{3.21}$$

$$\nabla=\{\Lambda(\omega,\theta,p):\ \omega\in[-\pi,\pi],\ \theta\in B_\Theta,\ p\in\mathbb{C}^m,\ |p|=1\}. \tag{3.22}$$

From (3.7), $\Lambda(\omega,\theta,p)$ is continuous in its 3 variables, therefore ∇ is compact in $\mathbb{R}^d$, and so is its convex hull $\mathrm{co}(\nabla)$.

Claim: If $\mathbb{R}_0^+$ is the set of nonnegative real numbers,

$$\mathrm{co}(\nabla)\cap(\mathbb{R}_0^+)^d=\emptyset. \tag{3.23}$$

In fact, if (3.23) does not hold, then we can find a point Λ in the boundary of $\mathrm{co}(\nabla)$, which falls inside $(\mathbb{R}_0^+)^d$. Invoking Lemma A.2, Λ is a convex combination of d points in ∇. Therefore there exist $\omega_1,\ldots,\omega_d$, $\theta^1,\ldots,\theta^d$, $p^1,\ldots,p^d$ such that

$$\Lambda=\sum_{k=1}^{d}\alpha_k\Lambda(\omega_k,\theta^k,p^k)\in(\mathbb{R}_0^+)^d,\quad \alpha_k\geq 0,\quad \sum_{k=1}^{d}\alpha_k=1. \tag{3.24}$$

Substitution in (3.24) with the definitions (3.20), (3.21) leads to

$$\sum_{k=1}^{d}\alpha_k|(F(\omega_k,\theta^k)p^k)_l|^2\geq\sum_{k=1}^{d}\alpha_k|(p^k)_l|^2,\qquad l=1\ldots d, \tag{3.25}$$

which is the same as (3.12) with p^k substituted by $\sqrt{\alpha_k}p^k$. This violates (b) by Lemma 3.3, so the claim is proved.

$\mathrm{co}(\nabla)$ and $(\mathbb{R}_0^+)^d$ are disjoint closed convex sets in $\mathbb{R}^d$, and $\mathrm{co}(\nabla)$ is compact. Therefore, by Lemma A.3 there exists $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\alpha, \beta \in \mathbb{R}$ such that

$$\langle x, \Lambda \rangle \leq \alpha < \beta \leq \langle x, y \rangle \quad \forall \Lambda \in \nabla, y \in (\mathbb{R}_0^+)^d. \tag{3.26}$$

Given the special structure of the cone $(\mathbb{R}_0^+)^d$, we can choose $\beta = 0$ and $x_1 > 0, \ldots, x_d > 0$. Using (3.21), we have

$$x_1\sigma_1(\omega, \theta, p) + \ldots + x_d\sigma_d(\omega, \theta, p) \leq \alpha < 0 \quad \forall \omega \in [-\pi, \pi],\ \theta \in B_\Theta,\ p \in \mathbb{C}^m,\ |p| = 1. \tag{3.27}$$

Define $X = \mathrm{diag}[x_1 I_{m_1} \ldots x_d I_{m_d}]$ which belongs to X^Φ, then (3.20) and (3.27) give

$$|X^{\frac{1}{2}} F(\omega, \theta) p|^2 - |X^{\frac{1}{2}} p|^2 \leq \alpha < 0 \quad \forall \omega,\ \theta,\ p. \tag{3.28}$$

It follows that $\gamma := \max_{\omega,\theta} \bar{\sigma}\left(X^{\frac{1}{2}} F(\omega, \theta) X^{-\frac{1}{2}}\right) < 1$. For any LTI $\boldsymbol{\theta} \in \boldsymbol{B_\Theta}$, and any LTV $\boldsymbol{\Phi} \in \boldsymbol{B_\Phi}$, we obtain

$$\|X^{\frac{1}{2}} \boldsymbol{\Phi}(\mathbf{M} \star \boldsymbol{\theta}) X^{-\frac{1}{2}}\| \leq \|\boldsymbol{\Phi}\| \|X^{\frac{1}{2}} (\mathbf{M} \star \boldsymbol{\theta}) X^{-\frac{1}{2}}\| \leq \|X^{\frac{1}{2}} F(\omega, \theta(e^{j\omega})) X^{-\frac{1}{2}}\|_\infty \leq \gamma. \tag{3.29}$$

This implies

$$\| \left(\mathbf{I} - X^{\frac{1}{2}} \boldsymbol{\Phi}(\mathbf{M} \star \boldsymbol{\theta}) X^{-\frac{1}{2}}\right)^{-1} \| \leq \frac{1}{1 - \gamma}, \tag{3.30}$$

which leads to a uniform bound on $\| (\mathbf{I} - \boldsymbol{\Phi}(\mathbf{M} \star \boldsymbol{\theta}))^{-1} \|$. This, together with (3.18) and (3.19), gives a uniform bound on $\|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})^{-1}\|$. ∎

Remarks:

- Condition (3.10) plays the role of the test (2.40) for purely LTI uncertainty, and is of a similar complexity, although there is a price paid for the inclusion of structured LTV blocks, in terms of the size of the problem and the number of frequency variables in which to search.

- If there is only one LTV block, the condition is the same as if all the blocks were LTI, which can also be shown directly in terms of the small gain theorem for LTI systems.

- If there is no LTI uncertainty, the previous theorem gives a μ test for LTV perturbations. The augmented μ condition (3.10) is therefore equivalent to condition (2.42) in terms of the μ upper bound. In particular, if M is a constant matrix with no dynamics, the frequency sweep disappears and we obtain $\hat{\mu}_{\boldsymbol{\Phi}}(M) = \mu_{\boldsymbol{\Phi}}(M)$. This is precisely the lifting result from Bercovici et al. [7], except that the size of the augmentation is the number d of blocks, rather than the dimension of the matrix.

3.1.4 State-Space μ Condition

Theorem 3.2 provides a necessary and sufficient test for robust stability in terms of a μ-condition (3.10) which involves a search in d frequency variables. Using a stable state-space realization $M(\lambda) = \lambda I \star S$ as in (2.21), a state-space test in terms of μ of a single constant matrix can be derived. This is analogous to Condition (2.41) for LTI uncertainty.

Corollary 3.4 *The system has uniform robust stability if and only if* $\mu_{\tilde{\Delta}_S}(\tilde{S}) < 1$, *where* $\tilde{S} = \mathrm{diag}[S, \ldots, S]$, *and*

$$\tilde{\Delta}_S = \begin{bmatrix} \Delta_S^{11} & \cdots & \Delta_S^{1d} \\ \vdots & \cdots & \vdots \\ \Delta_S^{d1} & \cdots & \Delta_S^{dd} \end{bmatrix}, \tag{3.31}$$

$$\begin{cases} \Delta_S^{ii} = \mathrm{diag}[\lambda^i I,\ \Phi^{ii},\ \theta^i],\ \Phi^{ii} \in \Phi,\ i = 1 \ldots d \\ \Delta_S^{ij} = \mathrm{diag}[0,\ \Phi^{ij},\ 0],\ \Phi^{ij} \in \Phi,\ i \neq j,\ i,j = 1 \ldots d \end{cases} . \tag{3.32}$$

Proof: Since $M(\lambda) = \lambda I \star S$, we can apply the main loop theorem (Lemma 2.1) successively to obtain

$$\mu_{\tilde{\Delta}_S}(\tilde{S}) < 1 \Longleftrightarrow \max_{\lambda_1, \ldots, \lambda_d \in \mathbb{D}} \mu_{\tilde{\Delta}}(\tilde{M}(\lambda_1, \ldots, \lambda_d)) < 1, \tag{3.33}$$

where $\tilde{M}(\lambda_1, \ldots, \lambda_d) := \mathrm{diag}[M(\lambda_1), \ldots, M(\lambda_d)]$. The maximum modulus-like property [55] of complex μ implies that the maximum in (3.33) occurs for λ_k at the boundary of the disk, therefore the left hand side of (3.33) coincides with (c) in Theorem 3.2. ■

3.2 Convex Tests for the Mixed LTV/LTI Problem

The conditions of Section 3.1 are necessary and sufficient and reduce the LTV/LTI robust stability problem to the computation of a finite dimensional object. As remarked in Chapter 2, however, exact μ computation is hard, so practical use of these conditions will involve employing bounds such as those in standard software packages such as μ-Tools [5].

In particular, we analyze in this section the upper bounds for this problem (sufficient conditions for robust stability) which lead to convex optimization. For this purpose, the μ conditions obtained in Section 3.1 can be bounded by use of the constant matrix upper bound $\hat{\mu}$. We focus here on the upper bound $\hat{\mu}_{\hat{\Delta}}(\tilde{M})$ over the frequencies $\omega_1, \ldots, \omega_d$, which follows from (3.10). An examination of the augmented structure $\hat{\Delta}$ shows that the corresponding commuting matrices $\hat{X}$ are of the form

$$\tilde{X} = \mathrm{diag}[X^{\Phi}, X_1^{\theta}, X^{\Phi}, X_2^{\theta}, \ldots, X^{\Phi}, X_d^{\theta}], \tag{3.34}$$

where $X^{\Phi} \in \mathbb{X}^{\Phi}$, $X_k^{\theta} \in \mathbb{X}^{\theta}$.

An alternative way of deriving a sufficient condition is to directly apply scaled small-gain conditions to the original problem. In the case of mixed LTV/LTI analysis as in (3.2), the natural scaling set is of the mixed form where the portion X^{θ} which corresponds to the LTI blocks θ is allowed to vary in frequency, and the portion X^{Φ} corresponding to the LTV blocks is constant. Define

$$\boldsymbol{X_m} = \left\{ X(e^{j\omega}) = \mathrm{diag}[X^{\Phi},\ X^{\theta}(e^{j\omega})], X^{\Phi} \in \mathbb{X}^{\Phi}, X^{\theta}(e^{j\omega}) \in \mathbb{X}^{\theta} \text{ for each } \omega \right\}. \tag{3.35}$$

Without loss of generality the functions in $\boldsymbol{X_m}$ are assumed to be continuous over frequency. The notation $\boldsymbol{X_m^P}$ refers to the subset of positive elements in $\boldsymbol{X_m}$.

We now show that the two approaches are equivalent.

Proposition 3.5 *Given the system* (3.2), *the following are equivalent:*

$$\max_{\omega_1,\ldots,\omega_d} \hat{\mu}_{\tilde{\Delta}}(\tilde{M}(\omega_1,\ldots,\omega_d)) \ < \ 1, \tag{3.36}$$

$$\exists\, X(e^{j\omega}) \in \boldsymbol{X_m^P} : \ \max_{\omega} \bar{\sigma}\left(X(e^{j\omega})M(e^{j\omega})X(e^{j\omega})^{-1}\right) \ < \ 1. \tag{3.37}$$

Proof: If an X satisfying (3.37) is found, then for any choice of frequencies $\omega_1, \ldots, \omega_d$, setting

$$\tilde{X} = \mathrm{diag}[X^{\Phi}, X^{\theta}(e^{j\omega_1}), \ldots, X^{\Phi}, X^{\theta}(e^{j\omega_d})]$$

will result in $\bar{\sigma}\left(\tilde{X}\tilde{M}(\omega_1,\ldots,\omega_d)\tilde{X}^{-1}\right) < 1$, and therefore $\hat{\mu}_{\tilde{\Delta}}(\tilde{M}) < 1$, implying (3.36). For the converse, (3.36) implies that

$$\max_{\omega_1,\ldots,\omega_d} \inf_{\tilde{X}} \bar{\sigma}(\tilde{X}\tilde{M}(\omega_1,\ldots,\omega_d)\tilde{X}^{-1}) < 1. \tag{3.38}$$

For fixed $\omega_1, \ldots, \omega_d$, clearly there exists $\epsilon(\omega_1, \ldots, \omega_d) > 0$ such that

$$\inf_{\frac{1}{\epsilon}I \geq \tilde{X} \geq \epsilon I} \bar{\sigma}(\tilde{X}\tilde{M}(\omega_1,\ldots,\omega_d)\tilde{X}^{-1}) < 1. \tag{3.39}$$

Since the $\omega_1, \ldots, \omega_d$ vary in a compact set, and $\tilde{M}$ is continuous, it follows that a fixed ϵ can be found satisfying (3.39) across $\omega_1, \ldots, \omega_d$. Now assume that (3.37) does not hold. This means equivalently that the LMI conditions across frequency

$$X(e^{j\omega}) > 0, \quad M^*(e^{j\omega})X(e^{j\omega})M(e^{j\omega}) - X(e^{j\omega}) < 0 \tag{3.40}$$

do not have a solution $X \in \boldsymbol{X}_{\boldsymbol{m}}^{\boldsymbol{P}}$, i.e. they cannot be satisfied with the X^{Φ} part constant across frequency. As a consequence, the family of sets

$$\mathcal{X}_{\omega}^{\Phi} := \left\{ X^{\Phi} \in \mathbb{X}^{\Phi} : \exists X = \mathrm{diag}[X^{\Phi}, X^{\theta}(\omega)],\ \epsilon I \leq X \leq \frac{1}{\epsilon},\ M(e^{j\omega})^* X M(e^{j\omega}) - X < 0 \right\} \tag{3.41}$$

have empty intersection for ω ranging in $[-\pi, \pi]$. These sets are convex and compact, and by normalizing the last block of X^{Φ} to I, they are in a $d-1$ dimensional space. From Lemma A.1, there exist d such sets with empty intersection.

Therefore there exist d frequencies $\omega_1, \ldots, \omega_d$ such that $\bar{\sigma}\left(X_k M(e^{j\omega_k}) X_k^{-1}\right) < 1$ cannot be satisfied with matrices $\epsilon I \leq X_k \leq \frac{1}{\epsilon}$ which share a common X^{Φ} part. Equivalently, given (3.34), we have that $\bar{\sigma}\left(\tilde{X}\tilde{M}(\omega_1, \ldots, \omega_d)\tilde{X}^{-1}\right) < 1$ cannot be satisfied with $\epsilon I \leq \tilde{X} \leq \frac{1}{\epsilon}$ for these frequencies, which is in direct contradiction with the fact that ϵ satisfies (3.39) for all frequencies $\omega_1, \ldots, \omega_d$.

■

As a corollary of Proposition 3.5 and Theorem 3.2, condition (3.37) is sufficient for robust stability under mixed LTV/LTI perturbations; this could also be shown via standard small-gain arguments.

From a computational point of view, (3.37) has the advantage of involving a search over only one frequency variable. A direct approach would be to grid the frequency axis and convert (3.37) to an LMI condition[2]. Note, however, that the common scale X^{Φ} introduces a coupling in the problem, so one is left with a large LMI condition, with size growing with the number of frequency grid-points.

In comparison, (3.36) tells us that in fact, d frequency values suffice, and we must only solve a coupled LMI problem of this size. However, since these frequencies are not known a priori, one has to grid a d dimensional space of frequencies. Therefore (3.36) reduces the size of the coupled LMI at the expense of more gridding. For low values of d, this alternative may be convenient.

Regarding the conservatism of these conditions, the results of [64] in relation to condition (2.43) suggest that (3.37) becomes necessary if the LTI perturbations are replaced by arbitrarily slowly varying uncertainty.

Theorem 3.6 *The conditions in Proposition 3.5 are satisfied if and only if there exists $\nu > 0$ such that the system (3.2) is robustly stable for $\boldsymbol{\Delta} = \mathrm{diag}[\boldsymbol{\Phi}, \boldsymbol{\theta}]$, $\boldsymbol{\Phi} \in \boldsymbol{B}_{\boldsymbol{\Phi}^{LTV}}$, $\boldsymbol{\theta}$ structured operator in $\boldsymbol{B}^{\nu}$.*

[2]Another approach would be to parametrize $X^{\theta}(\omega)$ by basis functions.

A proof of this result can be obtained through an essentially routine (though quite involved) extension of the techniques of [64], as was done in [61]. A more elegant proof will follow as an application of the techniques in Chapters 5 and 6; we therefore postpone this topic until then.

3.3 Combination with Real Parametric Uncertainty

In this section we take a further step in the analysis under combined uncertainty structures; in addition to LTV and LTI blocks, we include real parametric perturbations. In the robustness analysis setup of Figure 3.2, $\mathbf{H}$ is an LTI stable system, and the combined LTV/LTI/parametric uncertainty is denoted by $\mathbf{\Psi}$. The name $\mathbf{\Delta}$ is reserved to represent jointly the LTV/LTI portions of $\mathbf{\Psi}$, to provide consistency with the previous sections. Therefore $\mathbf{\Delta}$ is defined as in (3.2). The additional structured perturbation $\boldsymbol{\varrho}$ consists of real parametric blocks (e.g. $\varrho = diag[\varrho_1 I, \ldots, \varrho_n I]$, $\varrho_i \in \mathbb{R}$).

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{\Phi} & 0 & 0 \\ 0 & \boldsymbol{\theta} & 0 \\ 0 & 0 & \boldsymbol{\varrho} \end{bmatrix} = \begin{bmatrix} \mathbf{\Delta} & 0 \\ 0 & \boldsymbol{\varrho} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_\Phi & \mathbf{H}_{\Phi\theta} & \mathbf{H}_{\Phi\varrho} \\ \mathbf{H}_{\theta\Phi} & \mathbf{H}_\theta & \mathbf{H}_{\theta\varrho} \\ \mathbf{H}_{\varrho\Phi} & \mathbf{H}_{\varrho\theta} & \mathbf{H}_\varrho \end{bmatrix} = \begin{bmatrix} \mathbf{H}_\Delta & \mathbf{H}_{\Delta\varrho} \\ \mathbf{H}_{\varrho\Delta} & \mathbf{H}_\varrho \end{bmatrix}. \tag{3.42}$$

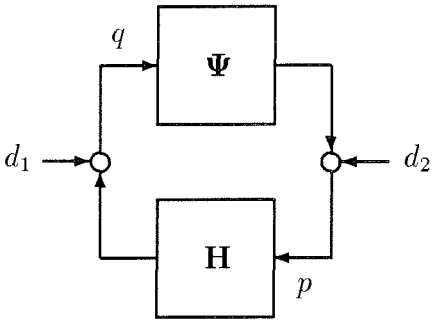


Figure 3.2: Robust stability under LTV/LTI/parametric uncertainty

We wish to obtain a necessary and sufficient condition for robust stability in this class, extending the results in Section 3.1. Since real parameters are a special case of time invariant uncertainty, at first sight it would appear that Theorem 3.2 applies directly, yielding a mixed-μ condition on $\tilde{H}(\omega_1, \ldots, \omega_d) = \text{diag}[H(e^{j\omega_1}), \ldots, H(e^{j\omega_d})]$, analogous to (3.10) but with the copies of ϱ in the augmented structure constrained to be real.

This augmentation would not capture, however, an additional property of real parametric uncertainty: in addition to taking real (rather than complex) values, a real parameter has no dynamics and therefore is constrained to be *constant across frequency*[3]. This suggests a modification of the augmented uncertainty structure for the case of real parameters, where they are forced to be constant across the augmentation. Consider the structure

$$\tilde{\Psi} = \left[\begin{array}{ccc} \Psi^{11} & \dots & \Psi^{1d} \\ \vdots & \dots & \vdots \\ \Psi^{d1} & \dots & \Psi^{dd} \end{array} \right], \left\{ \begin{array}{l} \Psi^{ii} = \mathrm{diag}[\Phi^{ii}, \theta^i, \varrho], i = 1 \dots d \\ \Psi^{ij} = \mathrm{diag}[\Phi^{ij}, \ 0, \ 0], \ i \neq j \end{array} \right. , \tag{3.43}$$

where $\Phi^{ij} \in \Phi$ and $\theta^i \in \Theta$ as before, but we constrain the copies of ϱ to be repeated across the augmentation. The $d = 2$ case is depicted in Figure 3.3 (a).

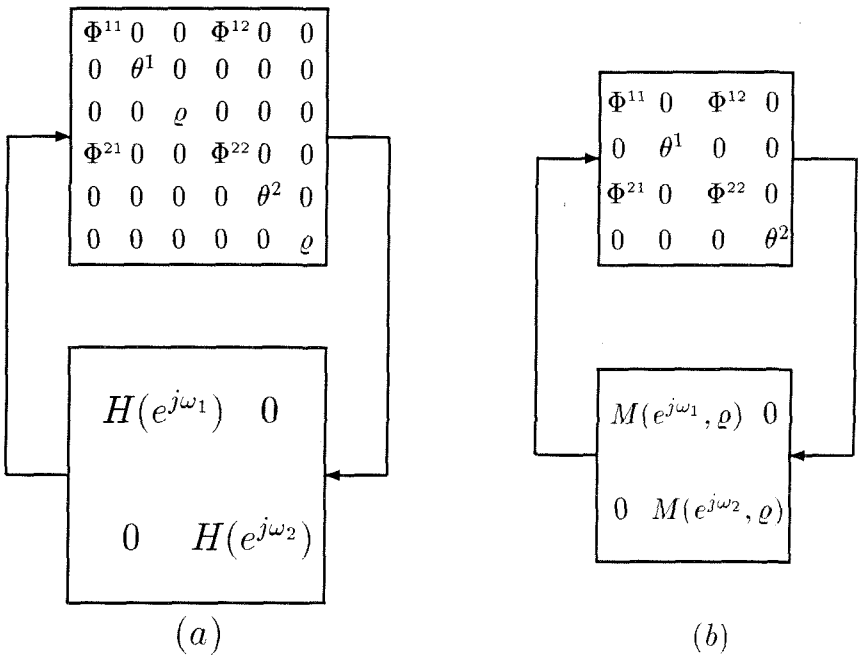


Figure 3.3: Augmented representations for the real parametric case.

It will also be useful to close the loop on the real parametric part, which will reduce the problem to the situation of Section 3.1. Assume that the real μ condition

$$\max_{\omega} \mu_{\varrho}(H_{\varrho}(e^{j\omega})) < 1 \tag{3.44}$$

holds, then for any fixed $\varrho \in B_{\varrho}$ we can define $\mathbf{M} = \mathbf{H} \star \boldsymbol{\varrho}$, with transfer function

$$M(e^{j\omega}, \varrho) := H(e^{j\omega}) \star \varrho = H_{\Delta} + H_{\Delta\varrho}\varrho(I - H_{\varrho}\varrho)^{-1}H_{\varrho\Delta}, \tag{3.45}$$

[3]This fact does not come into play in standard mixed μ with LTI/parametric uncertainty, where only one frequency is involved in the destabilization.

which is a stable time invariant system. We also define as in (3.3) the augmentation $\tilde{M}(\omega_1,\ldots,\omega_d,\varrho) := \mathrm{diag}[M(e^{j\omega_1},\varrho)\ldots M(e^{j\omega_d},\varrho)]$ for given $\omega_1,\ldots,\omega_d$, which can be obtained by LFT between the matrix $\tilde{H}(\omega_1,\ldots,\omega_d)$ and the ϱ portion of $\hat{\Psi}$. The uncertainty structure corresponding to $\tilde{M}$ is $\tilde{\Delta}$ as in (3.5), which is depicted in Figure 3.3 (*b*) for the case $d = 2$. We now provide the extension of Theorem 3.2.

Theorem 3.7 *In reference to the system* (3.42)*, the following are equivalent:*

(a) The system is uniformly robustly stable.

(b) Condition (3.44) holds, and with ω_k *varying in* $[-\pi,\pi]$, $\varrho \in B_\varrho$,

$$\max_{\omega_1,\ldots,\omega_d,\ \varrho} \mu_{\tilde{\Delta}}(\tilde{M}(\omega_1,\ldots,\omega_d,\varrho)) < 1. \tag{3.46}$$

(c) With ω_k *varying in* $[-\pi,\pi]$,

$$\max_{\omega_1,\ldots,\omega_d} \mu_{\tilde{\Psi}}(\tilde{H}(\omega_1,\ldots,\omega_d)) < 1, \tag{3.47}$$

where $\mu_{\tilde{\Psi}}$ *is the mixed (real/complex) structured singular value with respect to the structure* (3.43).

Proof: The equivalence of (*b*) and (*c*) is a direct application of Lemma 2.1.

Also, (*a*) implies condition (3.44), otherwise the real parameters ϱ would destabilize by themselves. Given (3.44), we observe that for fixed $\varrho \in B_\varrho$,

$$\mathbf{I} - \mathbf{\Psi H} = \begin{bmatrix} \mathbf{I} & -\mathbf{\Delta H}_{\Delta\varrho}(\mathbf{I} - \boldsymbol{\varrho}\mathbf{H}_\varrho)^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}(\mathbf{H} \star \boldsymbol{\varrho}) & \mathbf{0} \\ -\boldsymbol{\varrho}\mathbf{H}_{\varrho\Delta} & \mathbf{I} - \boldsymbol{\varrho}\mathbf{H}_\varrho \end{bmatrix}. \tag{3.48}$$

It follows that the (uniform) robust stability of system (3.42) is equivalent to the fact that for every ϱ in B_ϱ, the system $(\mathbf{M}, \mathbf{\Delta})$ is (uniformly) robustly stable. Since $\mathbf{\Delta}$ has the mixed LTV/LTI structure of Theorem 3.2, this is equivalent to condition (3.46) in (*b*). ■

The previous result has reduced the robust stability problem under LTV, LTI and parametric uncertainties to a mixed-μ condition across d-frequencies. We remark the following:

- In analogous manner to Corollary 3.4, a state-space condition can be derived from (3.47), which is equivalent to a single mixed μ problem.

- As usual, practical computation of a μ-condition such as (3.47) must be approached by upper and lower bounds, and possibly branch and bound techniques. Upper bounds will have the form given in (2.39); analogously to the situation in Section 3.2, we have the choice of writing these conditions in the original problem or in the augmented

problem. In this case, however, these are not equivalent, and the augmentation provides in principle a tighter bound. The reason for this is that the augmentation introduces repetition in the real uncertainty, increasing the freedom of the scaling matrices X^{ϱ}, G^{ϱ} corresponding to these blocks. This provides a way of imposing the condition that ϱ is constant across frequency, which is not imposed by an upper bound in the original problem.

3.4 Examples

In this section we illustrate the results of this chapter with a series of examples. These have been deliberately constructed so that there is a direct way to answer and interpret the robust stability question, thus providing more insight into the conditions given in the previous sections.

3.4.1 A System with LTV Uncertainty

We consider the interconnection of Figure (3.4), where $\mathbf{F}_1$, $\mathbf{F}_2$ are LTI single input/output systems, and $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ are uncertain perturbations with $\|\mathbf{\Phi}_i\| \leq 1$.

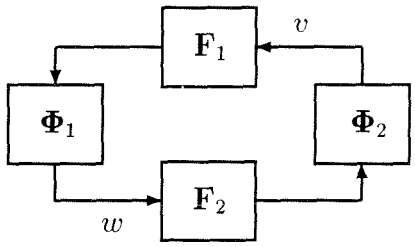


Figure 3.4: Example of analysis of LTV uncertainty

- If $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ were LTI perturbations, then they would commute with $\mathbf{F}_1$, $\mathbf{F}_2$ and from the small-gain theorem the robust stability condition would be

$$\|F_1 F_2\|_\infty < 1. \tag{3.49}$$

- From now on we consider the more interesting case when $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ are LTV perturbations. It turns out that in this simple configuration, the necessary and sufficient condition for robust stability is

$$\|F_1\|_\infty \|F_2\|_\infty < 1. \tag{3.50}$$

This condition is in general stricter than (3.49), since the two transfer functions $F_1(e^{j\omega})$ and $F_2(e^{j\omega})$ need not achieve their peak gain at the same frequency.

The sufficiency of (3.50) is clear from small-gain. To explain why it is also necessary, let us particularize in the example

$$F_1(\lambda) = \frac{0.1}{1 - 0.9\lambda}, \qquad F_2(\lambda) = \frac{-0.1}{1 + 0.9\lambda}. \tag{3.51}$$

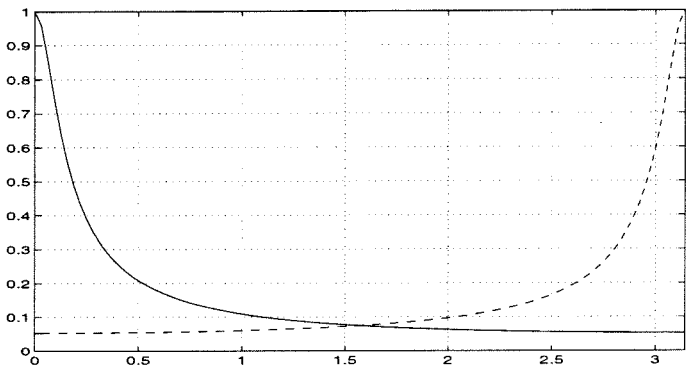


Figure 3.5: Transfer functions $F_1(e^{j\omega})$ and $F_2(e^{j\omega})$.

Here both systems have $\mathcal{H}_\infty$ norm equal to one, achieved respectively at frequencies $0, \pi$. Their magnitude frequency response plots are depicted in Figure 3.5 below. Condition (3.50) is therefore not satisfied. Let us show that the system can be destabilized by contractive LTV operators. In this case it suffices to consider the time-varying gains $\Phi_1(t) = \Phi_2(t) = (-1)^t$.

In reference to Figure 3.4, consider a constant signal $v(t) \equiv 1$ at the input of $\mathbf{F}_1$. Since the transfer function $F_1(e^{j\omega})$ is 1 at $\omega = 0$, the steady state output will be the same signal. The time-varying gain $\Phi_1(t)$ modulates this signal to $w(t) = (-1)^t$, which has all its frequency content at $\omega = \pi$, where $F_2(e^{j\omega})$ has value 1. This implies the steady state output of $\mathbf{F}_2$ is $w(t)$, which is demodulated back to $v(t)$ by $\Phi_2(t)$. We have a steady state signal in the loop, which violates robust stability. This informal argument can be formalized and extended to arbitrary $\mathbf{F}_1$, $\mathbf{F}_2$, and it illustrates strongly the "frequency shifting" properties of LTV perturbations.

We should recover the same answer if we do the analysis using the results in Section 3.1. For this purpose, we first rearrange Figure 3.4 to an $\mathbf{M}$-$\mathbf{\Delta}$ setup, where $\mathbf{\Delta}$ contains in this case only the LTV portion

$$\mathbf{\Delta} = \mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_2 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{0} & \mathbf{F}_1 \\ \mathbf{F}_2 & \mathbf{0} \end{bmatrix}. \tag{3.52}$$

Since $d = 2$, we must compute the augmentation of Theorem 3.2 over two frequency variables ω_1, ω_2. Figure 3.6 contains the resulting function $\mu_{\hat{\Delta}}(\hat{M}(\omega_1, \omega_2))$. computed using the

software package μ-Tools [5]. We find that the maximum is 1, achieved when the pair of frequencies is $(0, \pi)$ or $(\pi, 0)$, which is consistent with the previous analysis. Similar results can be obtained using the state-space condition in Corollary 3.4.

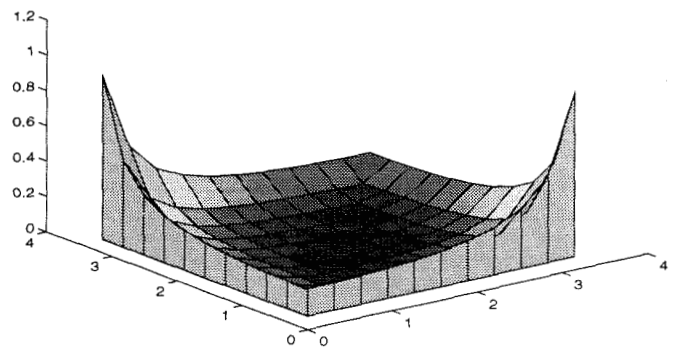


Figure 3.6: $\mu_{\tilde{\Delta}}(\tilde{M})$ as a function of ω_1, ω_2

3.4.2 Including LTI Uncertainty

We now modify the previous setup to include LTI uncertainty; F_1, F_2 are replaced by

$$F_1 = \frac{0.1}{1-(0.6+0.3\theta_1)\lambda}, \qquad F_2 = \frac{-0.1}{1+(0.6-0.3\theta_2)\lambda}, \qquad \|\theta_i\| \leq 1, \tag{3.53}$$

which are LFTs $F_i = T_i \star \theta_i$ on the LTI perturbations θ_i, as depicted in Figure 3.7.

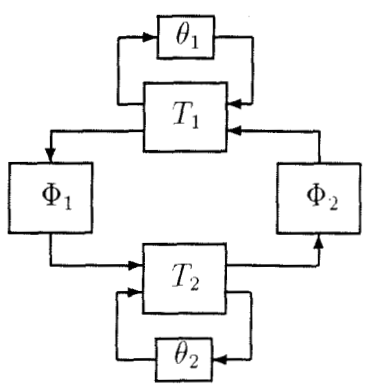


Figure 3.7: System with LTV and LTI uncertainty

From (3.50), it is clear that the worst-case θ_i are those which achieve $\max_{\theta_i} \|F_i\|_\infty$. The values in (3.53) have been chosen so that the worst-case perturbations are $\theta_1 = 1$,

$\theta_2 = -1$, which produce F_i as in (3.51). From the previous analysis the smallest destabilizing perturbation is of norm 1: $\theta_1 = 1$, $\theta_2 = -1$, $\Phi_1 = \Phi_2 = (-1)^t$.

These results are verified when we do the analysis of Theorem 3.2. After rearranging Figure 3.7 in the standard setup, and performing the augmentation, the μ computation gives a value $\max_{\omega_1,\omega_2} \mu_{\tilde{\Delta}}(\tilde{M}(\omega_1,\omega_2)) = 1$, achieved at $\omega_1 = 0$, $\omega_2 = \pi$.

If the θ_i in (3.53) correspond to real parameters instead of LTI perturbations, we still obtain the same answer from the robust stability analysis, since the worst-case values obtained above happen to be real.

3.4.3 Real Parametric vs LTI Perturbations

To produce an example where real and LTI perturbations give a different answer, one can use the same structure as in (3.53), but impose $\theta_1 = \theta_2 = \theta$ (repeated perturbation). In other words,

$$F_1 = \frac{0.1}{1-(0.6+0.3\theta)\lambda}, \qquad F_2 = \frac{-0.1}{1+(0.6-0.3\theta)\lambda}. \tag{3.54}$$

If θ is LTI, the repetition does not alter the results in Section 3.4.2, since an LTI perturbation can take the values 1 at $\omega = 0$ and -1 at $\omega = \pi$. An example is $\theta(e^{j\omega}) = e^{j\omega}$ ($\boldsymbol{\theta}$ is the delay operator $\boldsymbol{\lambda}$), which turns the F_i into second order systems with $\|F_1\|_\infty = \|F_2\|_\infty = 1$.

The same answer is obtained from the augmented μ test (3.10) which gives

$$\max_{\omega_1,\omega_2} \mu_{\tilde{\Delta}}(\tilde{M}(\omega_1,\omega_2)) = 1, \tag{3.55}$$

achieved at $\omega_1 = 0$, $\omega_2 = \pi$, and a destabilizing perturbation with $\theta^1 = 1$, $\theta^2 = -1$, as expected.

We now change θ for a real parameter $\varrho \in [-1,1]$, repeated in F_1, F_2. Since it is constant across frequency, it cannot maximize both $\|F_i\|_\infty$ simultaneously. In fact, straightforward calculations show that

$$\|F_1\|_\infty \|F_2\|_\infty = \frac{1}{16-9\varrho^2} \tag{3.56}$$

which has a maximum of $\frac{1}{7}$ for $\varrho \in [-1,1]$. This implies by (3.50) that the overall system is stable for $\|\Phi_i\| \leq 1$, $\varrho \in [-1,1]$.

Proceeding by augmentation, (3.47) gives a value

$$\max_{\omega_1,\omega_2} \mu_{\tilde{\Psi}}(\tilde{H}(\omega_1,\omega_2)) = 0.7906 < 1. \tag{3.57}$$

which yields the same stability conclusion by Theorem 3.7. In this case we can show directly that the smallest destabilizing perturbation for $\varrho \in \mathbb{R}$ and Φ_i LTV, has norm $\frac{1}{0.7906} = 1.2649$.

For this purpose, choose $\varrho \in [-\gamma, \gamma]$ and $\|\Phi_i\| \leq \gamma$. From (3.50), the full system will be stable as long as

$$\frac{1}{16 - 9\gamma^2} = \max_{\varrho \in [-\gamma,\gamma]} (\|F_1\|_\infty \|F_2\|_\infty) < \frac{1}{\gamma^2}, \tag{3.58}$$

which is equivalent to $\gamma < \sqrt{8/5} = 1.2649$, as expected.

To recapitulate, in this chapter we have shown that a combination of different classes of uncertain perturbations (LTV, LTI, parametric) can be analyzed for robust stability or $\mathcal{H}_\infty$ performance with the same mathematical tools as non-mixed problems. Structured singular value conditions were obtained, applicable to any combination of these uncertainty classes. Also, we remark that all the frequency domain conditions which were obtained can be extended directly to the continuous time case.

These new results raise a number of computational questions which are open for future research, regarding the most efficient approach for practical problems of this sort. In regard to the convex upper bounds which lead to coupled LMI problems across frequency, two alternatives (3.36) and (3.37) have been discussed and should be further explored. In relation to lower bound computation, the repeated structure of the augmented systems may be exploited in the algorithms.

Chapter 4

A Set-Based Methodology for White Noise Modeling

4.1 Motivation

In Chapter 2 we motivated the need for descriptions of uncertainty associated with engineering models. Essentially two approaches are available to describe uncertainty: one is to consider a set of allowable perturbations and perform worst-case analysis over this set; the other is to assign a probability distribution to the uncertainty and perform analysis in the average.

Most of the uncertainty descriptions presented in Section 2.1.2 were in terms of sets. This is the the simplest, most natural characterization for systematic errors due to undermodeling, which is a dominant factor in most control engineering problems; also, the issue of stability provides an incentive to take the worst-case point of view.

In robust control theory, the methodology based on sets is also applied to disturbance signals, by modeling them in terms of a ball in a signal space (e.g. l_2, l_∞), which motivates the $\mathcal{H}_\infty$ or $\mathcal{L}_1$ criteria for worst-case disturbance rejection. The main motivation for these disturbance models is mathematical convenience, since these performance specifications are easily combined with set descriptions of system uncertainty to analyze robust performance, as explained in Section 2.3.1.

This approach for disturbance modeling is pessimistic, however, since the worst-case signals which dominate the designs (e.g., sinusoids for $\mathcal{H}_\infty$ performance) are very unlikely to arise as empirical disturbances. These usually exhibit broadband spectral characteristics (*white noise*, or some filtered version), especially when they describe the cumulative macroscopic effect of very high dimensional fluctuations at the microscopic level. The statistics of these phenomena have been very accurately modeled by the theory of stochastic processes. The systematic study of the properties of dynamical systems under stochastic noise, pursued

by *stochastic* control theory, often leads to tractable results, the most notable being the solution of the optimal $\mathcal{H}_2$ (or LQG) control problem (see, e.g., [3]). The main limitation to its applicability is that noise is rarely the prevailing source of uncertainty, and the others do not fit easily into a stochastic description.

The desirable design specification, from both the performance and uncertainty points of view, appears to be in most cases Robust $\mathcal{H}_2$ performance: rejection of white signals in the worst-case over a set of plants. The historical difficulties (see Chapter 6) in obtaining a mathematical solution to this problem stem from the combination of two disparate methodologies: average-case analysis based on probability, and worst-case analysis based on functional analytic tools.

Another example of the difficulty of combining these frameworks is the relation between robust control and system identification. In mainstream system identification (see [45]), models are obtained in the style of time series analysis, by fitting parameters to data and explaining the errors in terms of random noise; the resulting models are, consequently, difficult to relate to those employed for robust control design. A more unified approach was sought recently by posing the identification problem in a worst-case setting; work in this area has, however, once again used a pessimistic view of disturbances, resulting in worst-case identification with weak consistency properties ([39, 85]) and high computational complexity ([16, 64]).

In this chapter we propose a new methodology for white noise modeling, aimed at resolving these difficulties.

The starting point is the following question: how does one decide whether a signal can be accurately modeled as a stochastic white noise trajectory? Deciding this from experimental data leads to a statistical hypothesis test on a finite length signal. In other words, one will accept a signal as white if it belongs to a certain set. The main idea of our formulation is to take this set as the *definition* of white noise, and carry out the subsequent analysis in a worst-case setting.

For this approach to be successful, these sets should:

- Exclude non-white signals (e.g. sinusoids) which are responsible for the conservatism of the $\mathcal{H}_\infty$ and $\mathcal{L}_1$ performance measures.
- Include likely instances of white noise. Here stochastic noise will be used as a guidance for the choice of a *typical set*, but not for average case analysis.
- Have simple enough descriptions to allow for tractable worst-case analysis.

With these objectives in mind, we will first analyze in Section 4.2 the case of finite horizon signals, and present set descriptions of white signals based on both the time and the

frequency domain points of view. Various tools from statistics and time series analysis are applied and developed to obtain a compromise between the stated objectives. In Section 4.3 we demonstrate the application of these ideas to worst-case system identification.

With the insight gained in Section 4.2, the case of infinite horizon signals is considered in Section 4.4. Two different settings are discussed, and one alternative which is most useful for robust control applications is developed in more detail.

Some Additional Notation

The following are some notational conventions and facts which are relevant only for the material in this chapter. $\mathbf{H}$ denotes an LTI system in $\mathcal{L}_c(l_2)$; most of this chapter will consider for simplicity scalar signals and single input/single output (SISO) systems; the multivariable case is considered in Section 4.4.3. In the SISO case we will assume that the impulse response $h(t)$ is in l_1, i.e.

$$\sum_{t=0}^{\infty} |h(t)| < \infty. \tag{4.1}$$

We now introduce the autocorrelation sequence of $\mathbf{H}$,

$$r_{\mathbf{H}}(\tau) := \sum_{t=0}^{\infty} h(t+\tau)h(t) \tag{4.2}$$

which is well defined, and itself a sequence in l_1, i.e. $\sum_{\tau=-\infty}^{\infty} |r_{\mathbf{H}}(\tau)| < \infty$. If $H(e^{j\omega})$ is the transfer function of $\mathbf{H}$, the spectrum $s_{\mathbf{H}}(\omega) := |H(e^{j\omega})|^2$ coincides with the Fourier transform of $r_{\mathbf{H}}(\tau)$. Also, the $\mathcal{H}_2$ norm of $\mathbf{H}$ satisfies

$$\|\mathbf{H}\|_2^2 = r_{\mathbf{H}}(0) = \sum_{t=0}^{\infty} |h(t)|^2 = \frac{1}{2\pi} \int_0^{2\pi} s_{\mathbf{H}}(\omega) d\omega. \tag{4.3}$$

For some of the frequency domain bounds obtained in this paper, we will further assume that $s_{\mathbf{H}}(\omega)$ is a function of bounded variation in $[0, 2\pi]$. This class is denoted $BV[0, 2\pi]$; see Appendix A for more details, including the definition of the total variation TV of a function. The time domain condition $\sum |\tau\, r_{\mathbf{H}}(\tau)| < \infty$ is sufficient for $s_{\mathbf{H}}(\omega) \in BV[0, 2\pi]$.

Finally, in the stochastic material presented, $\mathcal{P}$ denotes probability, and "I.I.D." is synonymous with independent, identically distributed.

4.2 The Finite Horizon Case

A reasonable starting point for white noise modeling is the case of a scalar valued, finite horizon, discrete time sequence $v(0), \ldots, v(N-1)$ of length N. The infinite horizon version will be considered in Section 4.4, which also covers the extension to vector-valued signals.

To analyze the response of a system with memory over a finite horizon, some convention must be made on the "past" values of the input signals. The two simplest choices are either to assume the system is initially at rest, or that it is in periodic steady state of period N. We will adopt the latter, since it leads to a more tractable spectral theory: the sequence $v(0), \ldots, v(N-1)$ will be identified with the periodic signal $v(t)$ of period N. This procedure is justified for analyzing stable systems with time constants which are small compared to N, so that the system is not sensitive to long range correlations in the input signals; this will be a standing assumption in this section.

The discrete Fourier transform (DFT) $V(k)$, $k = 0, \ldots, N-1$ of the sequence $v(t)$ is defined by the relations

$$V(k) = \sum_{t=0}^{N-1} v(t) e^{-j\frac{2\pi}{N}kt}, \qquad v(t) = \frac{1}{N} \sum_{k=0}^{N-1} V(k) e^{j\frac{2\pi}{N}kt}. \tag{4.4}$$

The circular autocorrelation sequence of v (*correlogram*) is given by

$$r_v(\tau) = \sum_{t=0}^{N-1} v(t+\tau) v(t), \quad \tau = 0 \cdots N-1, \tag{4.5}$$

and the sequence power spectrum (*periodogram*) by $s_v(k) = |V(k)|^2$, $k = 0 \cdots N-1$. The sequences $r_v(\tau)$ and $s_v(k)$ form a DFT pair. For an N-periodic signal $v(t)$, we will use as norm the energy over the period, $\|v\|^2 = r_v(0) = \frac{1}{N} \sum_{k=0}^{N-1} s_v(k)$.

The following relations follow immediately from the definitions.

Lemma 4.1 *Let* $\mathbf{H}$ *be an LTI, SISO system, with* $h(t) \in l_1$. *If* $v(t)$ *is an* N*-periodic input to* $\mathbf{H}$, *and* $y = \mathbf{H}v$ *is the corresponding steady state (periodic) output, then*

$$(i) \quad r_y(\tau) = \sum_{t=-\infty}^{\infty} r_{\mathbf{H}}(t) r_v(t-\tau), \tag{4.6}$$

$$(ii) \quad s_y(k) = s_{\mathbf{H}}(\tfrac{2\pi k}{N}) s_v(k). \tag{4.7}$$

4.2.1 White Noise Descriptions in the Time Domain

We wish to characterize white signals among sequences of length N; when faced with the problem of deciding whether an empirical signal is a sample of white noise, a statistician will perform a hypothesis test in terms of some statistic. A common choice (see [6, 45]) is the sample correlogram, which should approximate a delta function, the expected correlation for white noise. In other words a scalar signal is $v(t)$ categorized as white if $r_v(\tau)$ is small,

relative to $\|v\|^2 = r_v(0)$, for τ in a certain range (e.g. $1 \leq \tau \leq T$). For example, one can choose to specify that the correlogram (normalized to $r_v(0) = 1$), must fall inside a band around zero, of width γ, as depicted in Figure 4.1.

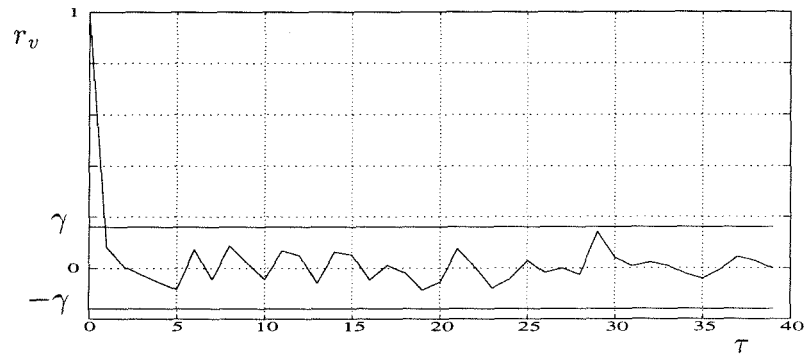


Figure 4.1: Correlogram of a pseudorandom sequence

From the classical statistical point of view, the choice of γ is associated to a level of significance of the test, which in turn depends on some stochastic model. But regardless of the reasoning behind this choice, ultimately the "whiteness" of the signal is decided in terms of whether it belongs or not to a parametrized set. This motivates the following:

Definition 5 *The set of signals* $v = (v(0), \ldots, v(N-1))$ *of length* N *which are white in the time domain sense, with accuracy* γ*, up to lag* T*, is defined by*

$$W_{N,\gamma,T} := \left\{ v \in \mathbb{R}^N : |r_v(\tau)| \leq \gamma r_v(0), \quad \tau = 1, \ldots, T \right\}. \tag{4.8}$$

Remark:

Other statistical tests can also be applied to the correlogram. The most commonly used choice in the statistical literature is the χ^2 test obtained from a sum of the squares of a fixed number of correlogram values [6]. The choice used in Definition 5 is preferable in our context, since it involves quadratic constraints on the signal v, which lead to more tractable worst-case analysis. Moreover, this test provides a very tight characterization of stochastic noise, as shown below.

The response of an LTI system to signals in such sets will now be analyzed from a worst-case perspective. The worst gain of the system under signals in $W_{N,\gamma,T}$ (a seminorm on systems) will be denoted

$$\|\mathbf{H}\|_{W_{N,\gamma,T}} := \sup \left\{ \frac{\|y\|}{\|v\|}, \; y = \mathbf{H}v, \; v \in W_{N,\gamma,T}, \|v\| \neq 0 \right\}. \tag{4.9}$$

Theorem 4.2 *Suppose the conditions of Lemma 4.1 hold, and* $v \in W_{N,\gamma,T}$. *Then*

$$\left|\frac{r_y(\tau)}{r_v(0)} - r_{\mathbf{H}}(\tau)\right| \leq \gamma \sum_{\substack{t=\tau-T \\ t\neq\tau}}^{\tau+T} |r_{\mathbf{H}}(t)| + \sum_{|t-\tau|>T} |r_{\mathbf{H}}(t)|. \tag{4.10}$$

Furthermore,

$$\left| \|\mathbf{H}\|^2_{W_{N,\gamma,T}} - \|\mathbf{H}\|_2^2 \right| \leq \gamma \sum_{\substack{t=-T \\ t\neq 0}}^{T} |r_{\mathbf{H}}(t)| + \sum_{|t|>T} |r_{\mathbf{H}}(t)|, \tag{4.11}$$

and for $\mathbf{H}$ *of finite impulse response* $h(t)$ *supported in* $[0, T]$,

$$\|H\|_2^2 \leq \|H\|^2_{W_{N,\gamma,T}} \leq \|H\|_2^2(1-\gamma) + \gamma \sum_{\tau=-T}^{T} |r_{\mathbf{H}}(\tau)|. \tag{4.12}$$

Proof: Equation (4.10) follows immediately from Lemma 1, and the definition of $W_{N,\gamma,T}$. Applying (4.10) at $\tau = 0$ gives (4.11). The upper bound in (4.12) follows from (4.11), the lower bound from the fact that the delta function is always a signal in the set $W_{N,\gamma,T}$.

■

The role of γ in this worst-case approach is to parametrize the freedom allowed in the disturbance signal.

Concentrating momentarily on the finite impulse response (FIR) case, it follows from (4.12) that for $\gamma = 0$, the worst-case gain is the $\mathcal{H}_2$ norm of the system. As γ is increased, more input signals are allowed, and for $\gamma = 1$ there are no restrictions in the input signals (since $|r_v(\tau)| \leq r_v(0)$ holds always), and therefore the induced norm approaches for large N the $\mathcal{H}_\infty$ norm of the system, which in the FIR case is equal to

$$\sup_{\omega} \left(r_{\mathbf{H}}(0) + 2 \sum_{\tau=1}^{T} r_{\mathbf{H}}(\tau) \cos \omega\tau \right)^{\frac{1}{2}}. \tag{4.13}$$

Note that the bound (4.12) is conservative for $\gamma = 1$; it is shown in [58] that it is exact for $\gamma < \frac{1}{T}$ and large N.

In the general, infinite impulse response case of (4.11), the parameter T also plays a role; for the constraints to be effective T must be at least comparable to the time constants of the system; the choice $T = N - 1$ gives the smallest sets, which are further studied below.

At this point in the discussion it would seem that if the purpose is to analyze white noise rejection, we could use the most restrictive choices ($\gamma = 0$, $T = N - 1$) which correspond to ideally white sequences. If these tools are to be used in a finite horizon setting, however, one

cannot expect that a realistic signal would have exactly zero autocorrelations; this would mean trading pessimistic disturbance modeling for an overly optimistic alternative.

There is no absolute answer as to what is a "realistic" white signal, but the strongest motivation for these disturbances comes from high dimensional fluctuations (e.g. particle agitation). These have been classically modeled as stochastic processes, or could also be interpreted in the context of deterministic chaos (see [72]). In any event, stochastic noise is known to provide a good model, regardless of whether the probability measure is due to chance or is the ergodic measure of a chaotic system. Therefore, a natural requirement for a realistic white noise set $W_{N,\gamma,T}$ is that it should have large probability for stochastic white signals. In the statistical language, this refers to the level of significance of the hypothesis test for white noise. We will analyze this asymptotically, when the length of N of the data record goes to infinity and γ, T are functions of N.

Theorem 4.3 *For each N let $v_N = (v(0), \ldots, v(N-1))$ be a vector of I.I.D. random variables, with zero mean and finite variance, and $\gamma_N > 0$.*

1. *If T is fixed, and $\gamma_N \sqrt{N} \overset{N\to\infty}{\longrightarrow} \infty$, then $\mathcal{P}(v_N \in W_{N,\gamma,T}) \overset{N\to\infty}{\longrightarrow} 1$.*
2. *If the $v(t)$ are bounded, and $\gamma_N \sqrt{\frac{N}{\log(N)}} \overset{N\to\infty}{\longrightarrow} \infty$, then $\mathcal{P}(v_N \in W_{N,\gamma,N-1}) \overset{N\to\infty}{\longrightarrow} 1$.*
3. *If the $v(t)$ are Gaussian, and $\gamma_N \frac{\sqrt{N}}{\log(N)^{\frac{3}{2}}} \overset{N\to\infty}{\longrightarrow} \infty$, then $\mathcal{P}(v_N \in W_{N,\gamma,N-1}) \overset{N\to\infty}{\longrightarrow} 1$.*

The previous theorem provides a very tight "typical set" for stochastic white noise: we argue that for many purposes, we can now ignore the probability measure and perform worst-case analysis over this set. One such case is disturbance rejection: by choosing $\gamma_N \overset{N\to\infty}{\longrightarrow} 0$ at a sufficiently slow rate, we find that the set $W_{N,\gamma,N-1}$ has asymptotically probability 1 and also $\|\mathbf{H}\|_{W_{N,\gamma,N-1}} \overset{N\to\infty}{\longrightarrow} \|\mathbf{H}\|_2$. We have therefore reinterpreted the $\mathcal{H}_2$ norm (asymptotically) as the worst-case gain over a typical set, rather than the average gain. Another situation where the probabilistic assumption can be replaced by a typical set is in the context of system identification, as will be discussed in Section 4.3.

This approach to disturbance modeling based on sets can, of course, also be applied when there is no stochastic model to begin with. What matters is the *statistical* information, which may be obtained directly from empirical correlograms, not the generating mechanism. In fact the bounds of Theorem 4.2 can be used even if the correlograms do not agree with the γ levels of stochastic noise.

Proof of Theorem 4.3

These statements fall in the realm of probability theory, and involve substantially different tools from the rest of the material in this thesis. Part 1 follows in fact from well known

results on asymptotic normality of the correlogram. We are not aware, however, of a proof of Parts 2 and 3, since such a test has apparently not been of interest to statisticians. Since these consistency results are instrumental in the application of these methods to system identification considered in Section 4.3, we will provide a proof below; this proof is, however, quite involved and not essential to the comprehension of the subsequent material.

Part 1: For the case of a fixed time lag τ, the distribution of the autocorrelation $r_v(\tau)$ has been extensively studied in the statistical literature [6, 2]; exact expressions for the distribution of $r_v(\tau)/r_v(0)$ when $v(t)$ is Gaussian are obtained in [2], and it follows that $\sqrt{N}\frac{r_v(\tau)}{r_v(0)}$ is asymptotically normal $\mathcal{N}(0,1)$. Since $\gamma\sqrt{N} \to \infty$, and T is fixed,

$$\mathcal{P}\left(v \notin W_{N,\gamma,T}\right) \leq \sum_{\tau=1}^{T} \mathcal{P}\left(\left|\sqrt{N}\frac{r_v(\tau)}{r_v(0)}\right| > \gamma\sqrt{N}\right) \overset{N\to\infty}{\longrightarrow} 0. \tag{4.14}$$

■

In parts 2, 3 of the theorem, the number of correlation constraints grows with the sample size, and the argument with the normal approximation cannot be used: even though each $r_v(\tau)$ for fixed τ is asymptotically normal, the joint distribution of $(r_v(1), \ldots r_v(N-1))$ is defined on a space of increasing dimension, where no global averaging occurs. Our proof relies on a Hoeffding inequality for sums of bounded random variables, [40]:

Theorem 4.4 (Hoeffding) *Let $z_0, \ldots z_{N-1}$ be I.I.D. random variables, of mean μ and bounded by $a \leq z_n \leq b$, define $\bar{z} = \frac{1}{N}\sum_{n=0}^{N-1} z_n$. Then for $\epsilon > 0$,*

$$\mathcal{P}\left(\bar{z} - \mu > \epsilon\right) \leq e^{\frac{-2N\epsilon^2}{(b-a)^2}}. \tag{4.15}$$

We want to apply this inequality to the sum $r_v(\tau) = \sum_{n=0}^{N-1} v(t)v((t+\tau)\mathrm{mod}N)$, with $v(0)\ldots v(N-1)$ I.I.D. The terms in this sum do not satisfy the independence requirement of Theorem 4.4, but their dependence is very slight, so the sum can be reduced to *three* sums of independent variables, as shown in the following sequence of Lemmas[1].

Lemma 4.5 *Let $\{a_1, \ldots, a_N\}$ be a permutation of $\{1, \ldots, N\}$. Then the set of ordered pairs $S = \{(1, a_1), \ldots, (N, a_N)\}$ can be partitioned into three disjoint sets S_1, S_2, S_3, of respective cardinality N_1, N_2, N_3, such that:*

1. *No two pairs which fall in a single S_i have a common element of $\{1, \ldots, N\}$ (i.e., if $(n, a_n), (m, a_m) \in S_i$, $n \neq m$, then $n \neq a_m$ and $m \neq a_n$).*
2. *$N_i \geq \frac{N}{5}$, $i = 1, 2, 3$.*

[1]The author acknowledges Geir Dullerud for pointing out this more elegant proof.

Proof: We perform the classification by induction. For a given n, assume that the pairs $(1, a_1), \ldots, (n, a_n)$ have been classified in disjoint sets $S_1^{(n)}$, $S_2^{(n)}$, $S_3^{(n)}$ which satisfy condition 1. Now consider a new pair $(n+1, a_{n+1})$. Since there are at most two pairs in S with an element in common with $(n+1, a_{n+1})$, at least one of the *three* $S_i^{(n)}$ will have none of these pairs and therefore condition 1 is maintained if $(n+1, a_{n+1})$ is added to it. This implies by induction that it is possible to partition S into sets S_1, S_2, S_3 satisfying condition 1.

Now consider their cardinalities N_1, N_2, N_3. Assume that $2N_i < N_j$ for some i, j. Since there are $2N_i$ elements in the pairs of S_i, and S_j has more *pairs*, then at least one pair in S_j shares no elements with those of S_i. Therefore this pair can be moved to S_i, maintaining condition 1. Repeating this procedure will lead to a partition S_1, S_2, S_3 satisfying condition 1, and in addition $2N_i \geq N_j \ \forall i, j$. If N_1 is, for example, the minimum of the N_i, then $N = N_1 + N_2 + N_3 \leq N_1 + 2N_1 + 2N_1 = 5N_1$ which implies condition 2 is satisfied.

■

Lemma 4.6 *Let $N \geq 3$, and $v(0), v(1), \ldots, v(N-1)$ be independent identically distributed random variables. Fix $1 \leq \tau < N$. Then $r_v(\tau)$ can be expressed as $r_v(\tau) = \Sigma_1 + \Sigma_2 + \Sigma_3$, where each Σ_i is the sum of N_i I.I.D. random variables, and $N_i \geq \frac{N}{5}$.*

Proof: For the permutation $\{a_1, \ldots, a_N\}$ given by the circular shift $a_n = (n+\tau) \mathrm{mod} N$, perform the classification into sets S_1, S_2, S_3 of Lemma 4.5. Then for each i choose

$$\Sigma_i := \sum_{(n, a_n) \in S_i} v(n) v(a_n). \tag{4.16}$$

By construction of the sets S_i, the terms in the sum Σ_i are I.I.D.

■

Now we return to the rest of Theorem 4.3.

Part 2: Assume $v(0), \ldots, v(N-1)$ are bounded random variables. $|v(t)| \leq K$. Pick $1 \leq \tau < N$. From Lemma 4.6, $r_v(\tau) = \Sigma_1 + \Sigma_2 + \Sigma_3$, where each Σ_i is the sum of N_i I.I.D. random variables, with zero mean and bounded in $[-K^2, K^2]$. Invoking Hoeffding's inequality and $N_i \geq \frac{N}{5}$, we have

$$\mathcal{P}\left(\frac{r_v(\tau)}{N} > \epsilon\right) \leq \sum_{i=1}^{3} \mathcal{P}\left(\frac{\Sigma_i}{N_i} > \epsilon\right) \leq \sum_{i=1}^{3} e^{\frac{-N_i \epsilon^2}{2K^4}} \leq 3e^{\frac{-N\epsilon^2}{10K^4}}. \tag{4.17}$$

The same argument can be employed to bound $\mathcal{P}\left(-\frac{r_v(\tau)}{N} > \epsilon\right)$. for each value of τ. This implies

$$\mathcal{P}\left(\max_{1 \leq \tau < N} \frac{|r_v(\tau)|}{N} > \epsilon\right) \leq 6Ne^{\frac{-N\epsilon^2}{10K^4}} = 6e^{\log(N)\left(1 - \frac{N\epsilon^2}{10K^4 \log(N)}\right)}. \tag{4.18}$$

Now choose $0 < \rho < E(v(t)^2)$. The complement of $W_{N,\gamma,N-1}$ can be written as

$$W^C_{N,\gamma,N-1} = \left\{ \max_{1\le\tau<N} \frac{|r_v(\tau)|}{r_v(0)} > \gamma \right\} \subset \left\{ \max_{1\le\tau<N} \frac{|r_v(\tau)|}{N} > \gamma\rho \right\} \cup \left\{ \frac{1}{N}\sum_{t=0}^{N-1} v(t)^2 < \rho \right\}. \quad (4.19)$$

The probability of the first set is bounded by (4.18), setting $\epsilon = \gamma\rho$. The probability of the second set can be bounded by another use of the Hoeffding inequality, applied to the bounded IID random variables $v(t)^2$. We arrive at

$$\mathcal{P}\left(W^C_{N,\gamma,N-1}\right) \le 6\epsilon^{\log(N)\left(1-\frac{N\gamma^2\rho^2}{10K^4\log(N)}\right)} + \epsilon^{\frac{-2N(E(v^2)-\rho)^2}{K^4}}. \quad (4.20)$$

The second term clearly goes to to zero as $N\to\infty$, and the same happens with the first term since by hypothesis $\frac{N\gamma_N^2}{\log(N)} \overset{N\to\infty}{\longrightarrow} \infty$.

Part 3: Assume $v(0),\ldots,v(N-1)$ are Gaussian random variables, $v(t) \sim \mathcal{N}(0,1)$. Choosing $K(N) = \sqrt{2\log(N)}$, define the random variables $u(t), t = 0,\ldots,N-1$ by truncation:

$$u(t) = \begin{cases} v(t) & \text{if } |v(t)| \le K(N) \\ 0 & \text{otherwise} \end{cases}, \quad (4.21)$$

$$\mathcal{P}\left(v(t) \ne u(t)\right) = \mathcal{P}\left(|v(t)| > K(N)\right) \le \frac{C}{K(N)} e^{\frac{-K(N)^2}{2}} = \frac{C}{N\sqrt{2\log(N)}}. \quad (4.22)$$

The inequality in (4.22) follows from a standard bound to the tail of the normal distribution (C is a constant). Denoting $u = (u(0),\ldots,u(N-1))$, $v = (v(0),\ldots,v(N-1))$, we conclude that

$$\mathcal{P}\left(u \ne v\right) \le N\mathcal{P}\left(u(t) \ne v(t)\right) \le \frac{C}{\sqrt{2\log(N)}} \overset{N\to\infty}{\longrightarrow} 0. \quad (4.23)$$

Observing that

$$\mathcal{P}\left(v \notin W_{N,\gamma,N-1}\right) \le \mathcal{P}\left(u \notin W_{N,\gamma,N-1}\right) + \mathcal{P}\left(u \ne v\right), \quad (4.24)$$

it remains to show that $\mathcal{P}\left(u \notin W_{N,\gamma,N-1}\right)$ also vanishes as $N\to\infty$. Since the variables $u(t)$ are bounded by $K(N)$, (4.20) gives

$$\mathcal{P}\left(u \notin W_{N,\gamma,N-1}\right) \le 6\epsilon^{\log(N)\left(1-\frac{N\gamma^2\rho^2}{10K^4\log(N)}\right)} + \epsilon^{\frac{-2N(E(u^2)-\rho)^2}{K^4}}. \quad (4.25)$$

The second term clearly has limit 0 as $N\to\infty$. The first term also goes to 0, since by hypothesis $\frac{N\gamma^2\rho^2}{K^4\log(N)} = \frac{\rho^2}{4}\frac{N\gamma^2}{\log(N)^3}$ goes to infinity.

■

4.2.2 Frequency Domain Descriptions

As the name implies, a white signal has flat distribution of energy across frequency, which in the finite horizon case would correspond to a flat periodogram, the DFT of a delta-function correlogram. The "raw" periodogram is typically very erratic, however, as demonstrated in Figure 4.2. This fact has been recognized for a long time (see, e.g. [6, 13]) in the statistical spectral analysis literature; correspondingly, the standard methods for power spectrum estimation are based on smoothing the periodogram, by some form of local averaging that reduces the fluctuations. This smoothing is most commonly done by convolution of the periodogram with a window function; an abundant literature (see [42]) has studied shapes and properties of these windows.

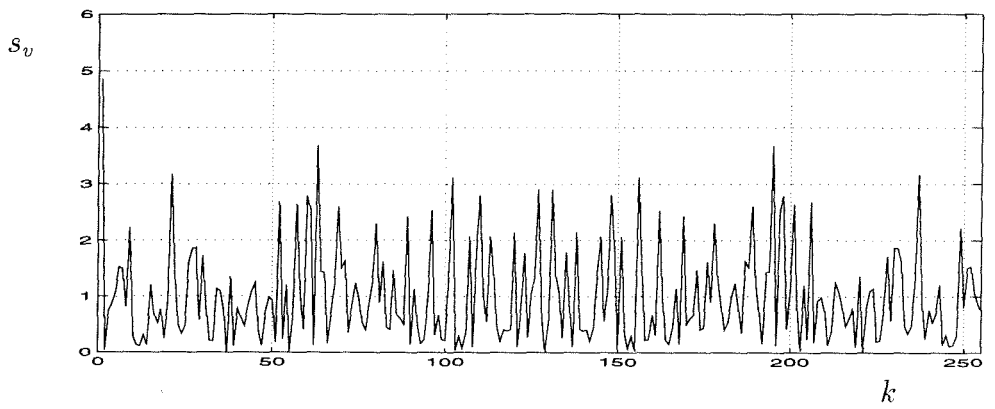


Figure 4.2: Periodogram of a pseudorandom sequence

In our context, we are interested in defining a set of typical periodograms, which is a hypothesis testing problem. Of course, the image of $W_{N,\gamma,T}$ under the DFT is such a set, but it does not have a simple description in terms of the frequency domain coordinates. We will therefore use a different characterization for the frequency domain which relies entirely on periodogram properties. One alternative is to specify that a "windowed" version of the periodogram be flat (this was pursued in [58]) but it is preferable to have a test which does not depend on a choice of window.

A very convenient alternative is provided by the Bartlett cumulative periodogram test (see [6, 42]), which consists of accumulating the periodogram and comparing the result to a linear function. Figure 4.3 contains the result of the accumulation process on the periodogram of Figure 4.2. As we see, the fluctuations have been smoothed by this integration and the result approximates a linear function in a *uniform* sense; this is the essence of definition which follows.

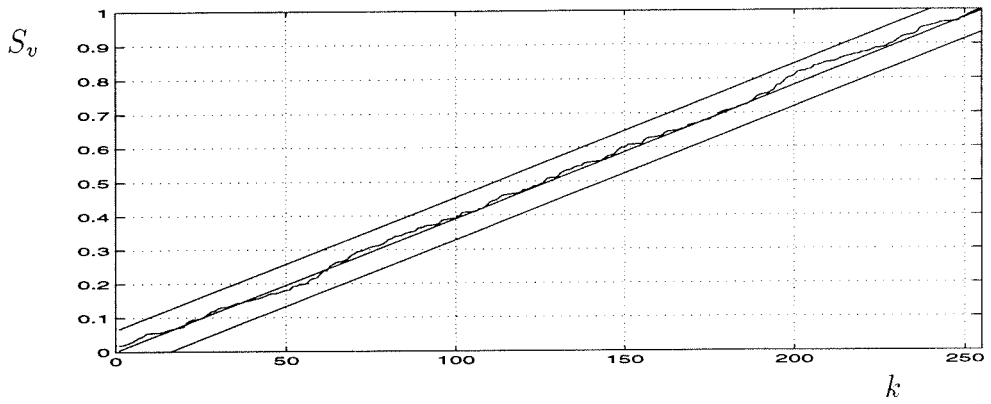


Figure 4.3: Cumulative periodogram and bounds for $\hat{W}_{N,\eta}$

Definition 6 *The set of white signals of length N in the frequency domain sense, with accuracy η is defined by*

$$\hat{W}_{N,\eta} = \left\{ v \in \mathbb{R}^N : \left| \frac{1}{N} \sum_{k=0}^{m-1} s_v(k) - \frac{m}{N} \|v\|^2 \right| \leq \eta \, \|v\|^2, \;\; 1 \leq m \leq N \right\}. \tag{4.26}$$

We will now analyze the frequency domain definition and exhibit properties which parallel those in the time domain. The worst-case induced norm of a system $\mathbf{H}$ under signals in the set $\hat{W}_{N,\eta}$ will be denoted $\|\mathbf{H}\|_{\hat{W}_{N,\eta}}$.

Theorem 4.7 *Consider a stable LTI system $\mathbf{H}$, with $s_{\mathbf{H}}(\omega) \in BV[0, 2\pi]$. Then*

$$\left| \|\mathbf{H}\|_2^2 - \|\mathbf{H}\|_{\hat{W}_{N,\eta}}^2 \right| \leq (\frac{1}{N} + \eta) TV(s_{\mathbf{H}}). \tag{4.27}$$

Proof: Fix $v \in \hat{W}_{N,\eta}$, $\|v\|^2 = 1$.

Define $\Gamma(k)$ by $\Gamma(0) = 0$, $\Gamma(m) := \frac{1}{N} \sum_{k=0}^{m-1} s_v(k)$, $1 \leq m \leq N$. Note that $\Gamma(N) = 1$. Let $y = \mathbf{H}v$, and for simplicity denote $s_{\mathbf{H}}(k)$ in place of $s_{\mathbf{H}}(\frac{2\pi k}{N})$. From (4.7) we have

$$\begin{aligned} \|y\|^2 &= \frac{1}{N} \sum_{k=0}^{N-1} s_{\mathbf{H}}(k) s_v(k) = \sum_{k=0}^{N-1} s_{\mathbf{H}}(k)(\Gamma(k+1) - \Gamma(k)) = \\ &= s_{\mathbf{H}}(N-1) + \sum_{k=1}^{N-1} (s_{\mathbf{H}}(k-1) - s_{\mathbf{H}}(k)) \Gamma(k). \end{aligned} \tag{4.28}$$

Similar calculations show that

$$\frac{1}{N} \sum_{k=0}^{N-1} s_{\mathbf{H}}(k) = s_{\mathbf{H}}(N-1) + \sum_{k=1}^{N-1} (s_{\mathbf{H}}(k-1) - s_{\mathbf{H}}(k)) \frac{k}{N}. \tag{4.29}$$

From (4.28), (4.29) we obtain (note that $v \in \hat{W}_{N,\eta}$ implies $|\Gamma(k) - \frac{k}{N}| \leq \eta$)

$$\left| \|y\|^2 - \frac{1}{N} \sum_{k=0}^{N-1} s_{\mathbf{H}}(k) \right| \leq \sum_{k=1}^{N-1} |s_{\mathbf{H}}(k-1) - s_{\mathbf{H}}(k)| \, |\Gamma(k) - \frac{k}{N}| \leq \eta \, TV(s_{\mathbf{H}}). \tag{4.30}$$

Also, by bounding the difference between the integral $\|\mathbf{H}\|_2^2 = \int_0^{2\pi} s_{\mathbf{H}}(\omega) \frac{d\omega}{2\pi}$ and a step function approximation, it follows that

$$\left| \|\mathbf{H}\|_2^2 - \frac{1}{N} \sum_{k=0}^{N-1} s_{\mathbf{H}}(k) \right| \leq \frac{1}{N} \; TV(s_{\mathbf{H}}), \tag{4.31}$$

which together with (4.30) leads to (4.27). ■

In reference to the properties of the set $\hat{W}_{N,\eta}$ in the case of stochastic noise, these have been studied in the statistical literature. We state the following result:

Theorem 4.8 *Let $v(0), \ldots, v(N-1), \ldots$ be I.I.D., zero mean, Gaussian random variables. If $\eta_N \sqrt{N} \overset{N \to \infty}{\longrightarrow} \infty$, then*

$$\mathcal{P}\left((v(0), \ldots, v(N-1)) \in \hat{W}_{N,\eta} \right) \overset{N \to \infty}{\longrightarrow} 1. \tag{4.32}$$

Remarks on the Proof:

The fact that a uniform bound is being applied to the cumulative periodogram means that we are imposing a number of constraints of the order of the sample size, as in Theorem 4.3 parts 2, 3; this again precludes simple arguments based on averaging.

The key observation, which led Bartlett [6] to propose this test, is to notice that the stochastic properties of the cumulative periodogram are similar to those used for tests on empirical distribution functions. The maximum deviation between an empirical distribution and the true distribution function forms the basis of the Kolmogorov-Smirnov test (see [12]), which has well known asymptotic properties. The connection with the cumulative periodogram can be seen as follows: in the case of Gaussian white noise, the periodogram values are independent and exponentially distributed (see [13]), which implies ([12], Prop. 13.15) that the normalized cumulative periodogram values $\frac{S_m}{S_N} = \frac{1}{N\|v\|^2} \sum_{k=0}^{m-1} s_v(k)$ have the same joint distribution as an ordered sample of uniform $(0,1)$ variables. From these arguments it follows that

$$\sqrt{N} \sup_{1 \leq m < N} \left| \frac{1}{N\|v\|^2} \sum_{k=0}^{m-1} s_v(k) - \frac{m}{N} \right|$$

converges in law to a fixed distribution. Since $\frac{1}{\eta_N\sqrt{N}} \overset{N\to\infty}{\longrightarrow} 0$, then

$$\frac{1}{\eta_N} \sup_{1\le m<N} \left| \frac{1}{N\|v\|^2} \sum_{k=0}^{m-1} s_v(k) - \frac{m}{N} \right| \xrightarrow{\mathcal{P}} 0,$$

which proves the theorem.

An additional remark is that although this proof is valid for Gaussian noise, there is indication in [6] that the asymptotic properties are insensitive to the noise distribution.

These asymptotic properties show that the frequency domain definition is adequate from the point of view of the objectives stated in Section 4.1: provided $\eta_N \overset{N\to\infty}{\longrightarrow} 0$, $\eta_N\sqrt{N} \overset{N\to\infty}{\longrightarrow} \infty$ the worst case disturbance rejection measure approaches the $\mathcal{H}_2$-norm of the system, while the class of signals contains asymptotically all typical instances of stochastic white noise. Thus the families of time and frequency domain sets have asymptotically the same properties, although they are different for any fixed N.

4.3 Application to Worst-Case System Identification

In the previous section we have provided set descriptions of finite horizon white signals aimed at worst-case analysis, and have shown that this procedure is sound and gives results which are consistent with the alternative stochastic setting. A first application of this approach is to worst-case system identification.

The classical literature on system identification (see [45] and references therein) characterizes model errors as due to stochastic noise; system identification in this setting is a special case of an estimation problem in statistical inference. From this perspective, the main requirement for an identification scheme is that if the true system is in the model class, the estimates are *consistent*, i.e. they converge to the true values in a stochastic sense, as the length of the experiment goes to infinity.

In contrast, robust control theory has relied on error models based on sets, e.g. a ball of systems in some norm. The desire to make identification and robust control more compatible has stimulated a research direction (see e.g. [39, 85, 16, 64]) which treats the system identification problem from a worst-case point of view, and seeks "hard" bounds on the identification error. In this formulation noise plays the role of an adversary; if, as is standard in robust control, it is allowed to vary over a large class (e.g. a ball in l_∞), then consistency of the estimates can no longer be ensured.

We now discuss these issues in the simple situation of a SISO model structure

$$y = h * u + d. \tag{4.33}$$

where the impulse response $h = (h(0), \ldots, h(T-1))$ is finite, and d is noise. Given data for y, u of length N, the problem is to estimate the system h. The equations in (4.33) can also be written in matrix form as

$$y = \mathcal{U}h + d, \tag{4.34}$$

where y, h, d are column vectors and $\mathcal{U}$ denotes the $N \times T$ Toeplitz matrix with first column u. The 2-norm will be used for signals here; the input is normalized to $\|u\|_2^2 = N$. To simplify the analysis, assume that the experiment was started at time $-(T-1)$, with values of u which are N-periodic.

In the classical theory, d is assumed to be stochastic white noise: IID random variables, with zero mean, variance σ^2. In this linear regression problem the minimum variance estimate for h is given by the least squares solution

$$\hat{h} = (\mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^* y, \tag{4.35}$$

where invertibility of $\mathcal{U}^*\mathcal{U}$ (*persistence of excitation*) is assumed. The estimator (4.35) is unbiased, and its covariance matrix $\sigma^2(\mathcal{U}^*\mathcal{U})^{-1}$ will converge to zero as $N \to \infty$, under stationarity assumptions in u. This implies that in the stochastic sense, the estimator will be consistent.

For worst-case identification, we first follow the usual approach which is to restrict d only in norm; suppose $\|d\|_2^2 \leq \varrho^2 N$ (noise to signal ratio ϱ). Since there is a linear relation (4.34) between h and d, the set of h values compatible with the data and the constraint $\|d\|_2^2 \leq \varrho^2 N$ will be an ellipsoid. It follows that if one wishes to minimize the maximum error in the 2-norm in h, the optimal choice is the center of the ellipsoid, which once again corresponds to the least squares solution (4.35). Assuming now for simplicity that u is purely white (i.e. $\mathcal{U}^*\mathcal{U} = NI$, this is also the optimal choice) the worst-case estimation error

$$\|\hat{h} - h\|_2 = \|(\mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^* d\|_2 = \frac{1}{N}\|\mathcal{U}^* d\|_2 \tag{4.36}$$

has a value of ϱ, corresponding, for example, to $d = \varrho u$.

We therefore find that although both points of view lead in this case to the same optimal estimate, they attach to it a different interpretation. In particular, consistency is lost in the worst-case setting: the estimation error cannot be made smaller than ϱ, no matter how long the data record is. The same was found in [39, 85] for other system norms. The reason for this pessimistic interpretation is that the noise, which plays an adversarial role, is allowed to vary in a class where it can "conspire" to have a high correlation with the input. This suggests that the desirable consistency interpretation can be recovered if the disturbance is constrained in the style of this paper to have low cross correlation with u.

One way of doing this was studied recently by Venkatesh and Dahleh [86]: the input is chosen to be periodic of period T (this allows for persistence of excitation of order T), and the set $W_{N,\gamma,N-1}$ is used[2] to restrict the disturbance d. The main observation from [86] is that in this case (assuming N is a multiple of T)

$$\|\mathcal{U}^* d\|_2^2 = \sum_{\tau=0}^{N-1} r_d(\tau) r_u^T(\tau), \tag{4.37}$$

where $r_d(\tau)$ is the correlogram of d (length N) and $r_u^T(\tau)$ is the correlogram for u of length T, repeated periodically. For a purely white u, we would have

$$r_u^T(\tau) = \begin{cases} T \;\text{ for } \tau = kT,\; k = 0, \ldots, \frac{N}{T} - 1 \\ 0 \;\; \tau \neq kT,\;\; 1 \leq \tau \leq N-1 \end{cases}. \tag{4.38}$$

Imposing that $d \in W_{N,\gamma,N-1}$, (4.36), (4.37) and (4.38) give

$$\|\hat{h} - h\|_2^2 \leq \frac{T}{N^2}\left(r_d(0) + \sum_{k=1}^{N/T-1} r_d(kT)\right) \leq \frac{T}{N^2}\varrho^2 N\left(1 + \gamma(\frac{N}{T} - 1)\right) \leq \varrho^2\left(\gamma + \frac{T}{N}\right). \tag{4.39}$$

We now consider another way to constrain the identification problem, which is to directly impose low correlation between u and d. For example, we can impose that $(\varrho u, d)$ is a white signal in the set $W_{N,\gamma,T}^2$, which is the multivariable version of $W_{N,\gamma,T}$. This set is defined analogously to (4.63) (see Section 4.4), and in particular imposes the cross correlation constraints

$$\varrho|\langle \boldsymbol{\lambda}^\tau u, d\rangle| = |\langle \boldsymbol{\lambda}^\tau \varrho u, d\rangle| \leq \gamma(\|\varrho u\|^2 + \|d\|^2) = 2\gamma\varrho^2 N, \quad \tau = 0, \ldots, T. \tag{4.40}$$

Since the elements of $\mathcal{U}^* d$ are $\langle \boldsymbol{\lambda}^\tau u, d\rangle$, $\tau = 0, \ldots, T-1$, these bounds can be applied to (4.36) giving

$$\|\hat{h} - h\|_2 = \frac{1}{N}\|\mathcal{U}^* d\|_2 \leq \frac{1}{N} 2\gamma\varrho N\sqrt{T} = 2\gamma\varrho\sqrt{T}. \tag{4.41}$$

In both cases ((4.39) and (4.41)) if $\gamma_N \to 0$ as $N \to \infty$, we obtain the consistency property $\hat{h}_N \overset{N\to\infty}{\longrightarrow} h$.

By choosing an appropriate decay rate for γ (e.g. $\gamma = \frac{1}{N^\alpha}, \alpha < \frac{1}{2}$), the chosen disturbance set has high probability from the stochastic viewpoint. Theorem 4.3 applies to the case $d \in W_{N,\gamma,N-1}$; a similar argument can be used for the multivariable case, or applied only to the constraints (4.40). Therefore, our class of disturbances is still rich enough to accommodate

[2] In [86] a variation of this set is used; it leads nevertheless to similar bounds as those given here.

classical identification. In addition, the errors in (4.39) and (4.41) will decay to zero in polynomial time, in contrast to the complexity results of [16, 64].

As those in [85, 16, 64], these results for FIR identification are mainly of conceptual value, and contribute to understand the properties of the identification problem from a worst-case perspective. In this simple case we have proposed no new identification algorithm, and have only discussed different interpretations of the least squares solution. We extract, however, important practical guidelines as to how a more general identification problem should be posed when worst-case guarantees are sought, as in cases of identification involving noise and set descriptions of unmodeled dynamics. To avoid conservatism the disturbance must be constrained explicitly, and correlation constraints are an adequate tool for this. These more general problems are a topic of future research.

4.4 The Infinite Horizon Case

The role of infinite horizon signals in mathematical modeling is that of an abstraction to describe the behavior of signals and systems over a long, but unspecified horizon; the chosen mathematical framework must extend naturally the finite horizon properties and lead to tractable analysis.

Two frameworks arise naturally for the study of deterministic spectral analysis: bounded power signals and bounded energy (l_2) signals.

4.4.1 Bounded Power Signals

There is a long historical tradition in a non-stochastic theory of white noise, going as far back as Wiener (see [88]), who considered ergodicity properties to build a spectral theory of stationary signals, motivated by problems of communication theory where noise is combined with deterministic signals. These ergodicity properties can be motivated in the theory of stochastic processes [12], as well as in the context of deterministic chaos [72].

For disturbance rejection problems, this approach was followed in Zhou et al. [99], who considered the class of bounded power signals (see also [98]), defined for discrete time by

$$\mathcal{BP} = \left\{ v(t) : r_v(\tau) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{t=-N}^{N} v(t+\tau)v(t) \text{ exists for each } \tau \right\}. \qquad (4.42)$$

The function $r_v(\tau)$ is the autocorrelation of the signal, and the power norm $\|v\|_P = (r_v(0))^{\frac{1}{2}}$ plays the role of a seminorm (with some restrictions, see below). The positive definite nature of the function $r_v(\tau)$ implies from Bochner's theorem (see [12]) that there exists a spectral distribution function $S_v(\omega), \omega \in [0, 2\pi]$ such that $r_v(\tau)$ is recovered from

the Stieltjes integral (see Appendix A)

$$r_v(\tau) = \frac{1}{2\pi} \int_0^{2\pi} e^{j\omega\tau} dS_v(\omega). \tag{4.43}$$

Equivalently, there exists a positive spectral measure which is the Fourier transform of $r_v(\tau)$; this allows for periodic effects, which correspond to atoms of this measure, or discontinuities in S_v. It also includes the case of an absolutely continuous spectrum, with the corresponding spectral density $s_v(\omega) = \frac{d}{d\omega} S_v(\omega)$.

The set of white signals in the class $\mathcal{BP}$ is defined by

$$W_0 = \{v \in \mathcal{BP} : r_v(\tau) = \|v\|_P^2 \delta(\tau)\} = \{v \in \mathcal{BP} : S_v(\omega) = \omega \|v\|_P^2\}, \tag{4.44}$$

where $\delta(\tau)$ is the Dirac function (1 at $\tau = 0$, 0 elsewhere). The signals in W_0 are exactly white, and furthermore W_0 is "typical" in the stochastic context:

Proposition 4.9 *Let $v = (v(0), \ldots, v(t), \ldots)$ be a sequence of I.I.D. random variables, with zero mean and finite variance. Then $\mathcal{P}(v \in W_0) = 1$.*

Proof: For a fixed $\tau \neq 0$, referring to [12] (proposition 6.31), we find that the random process $z(t) = v(t)v(t+\tau)$ is ergodic, so with probability 1,

$$\lim_{N \to \infty} \frac{1}{2N+1} \sum_{t=-N}^{N} v(t+\tau)v(t) = E[v(t+\tau)v(t)] = 0. \tag{4.45}$$

Therefore W_0 has probability 1 (countable intersection of probability 1 sets).

■

We now analyze the worst-case response of a system to signals in W_0. For an LTI single input/output system $\mathbf{H}$, we write the following input-output relationships:

$$(i) \quad r_y(\tau) = \sum_{t=-\infty}^{\infty} r_{\mathbf{H}}(t) r_v(t-\tau), \tag{4.46}$$

$$(ii) \quad dS_y(\omega) = |H(e^{j\omega})|^2 dS_v(\omega). \tag{4.47}$$

If these are satisfied, then the gain of $\mathbf{H}$ in power is $\|\mathbf{H}\|_2^2$ for every signal in W_0. Also, the gain in power under unrestricted inputs is the $\mathcal{H}_\infty$ norm of $\mathbf{H}$. This would make the bounded power class an appealing setting from the point of view of the objectives set forth at the beginning of the chapter; a similar approach was used in [99, 98].

The definition of the class $\mathcal{BP}$, however, raises a number of theoretical difficulties:

- In the first place, the relationships (4.46-4.47), which are easily proved for FIR systems, raise a number of mathematical problems for general LTI systems, since they involve changes of order in limiting processes. A partial result stated in [99] is that if $v \in \mathcal{BP}$, $v \in l_\infty$, and $\mathbf{H}$ is exponentially stable, $\mathbf{H}v$ is in $\mathcal{BP}$ and (4.46-4.47) hold.

- These mathematical difficulties are even more dominant if one wishes to use this class together with LTV operators, as motivated in Chapter 2.

- The continuous time generalization has greater problems, as noted in [99], since ideally white signals with flat spectrum would escape the bounded power class.

- Most importantly, there is little mathematical structure to the class $\mathcal{BP}$. In particular, it is not a vector space (not being closed under addition, see [48]), which greatly restricts the applicability of functional analytic tools. Consequently, it is not a seminormed space, and there is no inner product to go along with the quadratic structure of the definitions.

Although there may be ways around these mathematical difficulties, we argue that for the study of disturbance rejection problems, the theory is best developed in the space l_2 of bounded energy signals, which has the structure of Hilbert space. At first it may seem unnatural to consider white noise signals which decay to zero as time goes to infinity, rather than being stationary. We remark, however, the following:

- The asymptotic behavior should never be the determining factor in any sensible engineering model. From a practical point of view, the response of a system is adequately described by a long enough truncation in time. The problem with the $\mathcal{BP}$ class is precisely this: only the asymptotic behavior matters, and any finite horizon signal can, for example, be the truncation of a white noise signal.

- All the mathematical difficulties with the class $\mathcal{BP}$ arise because of this asymptotic behavior. It seems that the infinite horizon abstraction should help, rather than hinder, the treatment of phenomena over long time intervals.

- Actually, the same considerations apply to standard $\mathcal{H}_\infty$ theory. While the $\mathcal{H}_\infty$ norm is most naturally motivated [99, 98] by the gain in power for bounded power inputs, since this class includes sinusoids, most technical results on $\mathcal{H}_\infty$ are obtained by using l_2 as a signal space, which does not contain these signals, but captures the same system properties since it contains signals of arbitrarily narrow bandwidth.

4.4.2 l_2 Setting

For the reasons explained above, we now discuss set characterization of white noise in l_2 space. For l_2 scalar sequences, the autocorrelation is defined by $r_v(\tau) = \langle v, \boldsymbol{\Lambda}^\tau v \rangle$. The corresponding spectral measure as in (4.43) is absolutely continuous; the spectral density is defined by $s_v(\omega) = \frac{d}{d\omega} S_v(\omega) = |v(e^{j\omega})|^2$, where $v(e^{j\omega})$ is the Fourier transform of $v(t)$. In the case of l_2 signals conditions (4.46-4.47) are satisfied, for example, for $h(t) \in l_1$.

The set of white, scalar l_2 signals can be defined as

$$W_0 = \left\{ v \in l_2 : r_v(\tau) = \|v\|_2^2 \delta(\tau) \right\} = \left\{ v \in l_2 : |v(e^{j\omega})|^2 = \|v\|_2^2, \text{ a.e. in } [0, 2\pi] \right\}. \quad (4.48)$$

We will find it useful to introduce classes of approximately white signals in l_2; the following definitions are inspired on the finite horizon case. For the time domain version,

$$W_{\gamma,T} := \{ v \in l_2 : |r_v(\tau)| \leq \gamma r_v(0) \;\; \tau = 1, \ldots, T \}. \quad (4.49)$$

In the frequency domain, Definition 6 extends by considering the difference

$$F_v(s) := \int_0^s |v(\omega)|^2 \frac{dw}{2\pi} - \frac{s}{2\pi} \|v\|_2^2 \quad (4.50)$$

between the cumulative spectrum and a linear function. Define

$$\hat{W}_\eta = \{ v \in l_2 : \sup_{s \in [0,2\pi]} |F_v(s)| < \eta \}, \quad (4.51)$$

which imposes a uniform bound on this difference. The width η of this band could also be normalized by a factor $\|v\|_2^2$, as in Definition 6; we have not done this in (4.51) to streamline some of the proofs in Chapter 6, but this change is not essential.

As in Section 4.2, the worst-case gain of a system $\mathbf{H}$ under signals in $W_{\gamma,T}$ is denoted by

$$\|\mathbf{H}\|_{W_{\gamma,T}} := \sup \{ \|\mathbf{H}v\|_2 : v \in W_{\gamma,T}, \|v\|_2 \leq 1 \}, \quad (4.52)$$

and analogously for $\|\mathbf{H}\|_{W_\eta}$. We now provide bounds for these quantities, in the case of an LTI system $\mathbf{H}$. As a consequence of (4.46), we find that

$$\|\mathbf{H}\|_2^2 \leq \|\mathbf{H}\|_{W_{\gamma,T}}^2 \leq \|\mathbf{H}\|_2^2 + \gamma \sum_{\tau=-T}^{T} |r_{\mathbf{H}}(\tau)| + \sum_{|t|>T} |r_{\mathbf{H}}(t)|. \quad (4.53)$$

For the frequency domain bound we assume as in Section 4.2.2 that $s_{\mathbf{H}}(\omega)$ is of bounded variation.

Lemma 4.10 *Let $Y(\omega) \in BV[0, 2\pi]$. If $v \in W_\eta$, then*

$$\left| \int_0^{2\pi} Y(\omega)|v(e^{j\omega})|^2 \frac{d\omega}{2\pi} - \|v\|_2^2 \int_0^{2\pi} Y(\omega) \frac{d\omega}{2\pi} \right| \leq \eta \, TV(Y). \tag{4.54}$$

Proof: Defining $F_v(s)$ as in (4.50) ($F_v(0) = F_v(2\pi) = 0$), an integration by parts (see (A.5)) yields

$$\int_0^{2\pi} Y(\omega)(|v(e^{j\omega})|^2 - \|v\|_2^2) \frac{d\omega}{2\pi} = -\int_0^{2\pi} F_v(\omega) dY(\omega). \tag{4.55}$$

Since $v \in W_\eta$, then $\|F_v(\omega)\|_\infty = \sup_\omega |F_v(\omega)| < \eta$, so (A.3) implies that the right hand side of (4.55) can be bounded by $\eta \, TV(Y)$.

■

Picking $Y(\omega) = |H(e^{j\omega})|^2$ we conclude that

$$\|\mathbf{H}\|_2^2 \leq \|\mathbf{H}\|_{W_\eta}^2 \leq \|\mathbf{H}\|_2^2 + \eta \, TV(|H|^2). \tag{4.56}$$

As a consequence of these bounds, the system $\mathcal{H}_2$ norm can be motivated as the gain under signals in W_0, or equivalently by the limit norms

$$\lim_{\substack{\gamma \to 0+ \\ T \to \infty}} \|\mathbf{H}\|_{W_{\gamma,T}}, \qquad \lim_{\eta \to 0+} \|\mathbf{H}\|_{W_\eta}. \tag{4.57}$$

4.4.3 Multivariable Extension

This section explains how the previous methodology can be extended to deal with vector valued white signals. We only consider infinite horizon l_2 signals, which demonstrates all the necessary extensions; the same ideas could be applied in a finite horizon setting. For $v(t) \in l_2(\mathbb{R}^m)$, treated as a column vector, the matrix autocorrelation is given by

$$R_v(\tau) = \sum_{t=-\infty}^{\infty} v(t+\tau)v'(t). \tag{4.58}$$

A spectral (matrix) distribution function $S_v(\omega)$ is defined as before, verifying a matrix version of (4.43). In this l_2 case,

$$\frac{dS_v(\omega)}{d\omega} = s_v(\omega) = v(e^{j\omega})v^*(e^{j\omega}). \tag{4.59}$$

The 2-norm of the signal verifies

$$\|v\|_2^2 = \text{trace}(R_v(0)) = \frac{1}{2\pi} \int_0^{2\pi} \text{trace}(s_v(e^{j\omega})) d\omega. \tag{4.60}$$

If $\mathbf{H}$ is an LTI system with inputs $v \in l_2^m$ and outputs $y \in l_2^p$, then under mild assumptions

$$R_y(\tau) = \sum_{t=0}^{\infty}\sum_{s=0}^{\infty} h(s)R_v(\tau+t-s)h'(t), \tag{4.61}$$

$$s_y(\omega) = H(e^{j\omega})s_v(\omega)H(e^{j\omega})^*. \tag{4.62}$$

Vector valued white signals correspond, ideally, to a spectrum $s_v(\omega) = I$, or an autocorrelation $R_v(\tau) = \delta(\tau)I$, where I is the identity matrix. This means that the components of the vector are individually white signals, and are mutually uncorrelated.

For l_2 signals, however, the spectrum (4.59) is always a rank 1 matrix, so there is no multivariable extension, inside l_2, of the set W_0 given in (4.48)[3].

This difficulty disappears if one considers approximately white signals, extending the definitions (4.49) and (4.51). To measure distances between matrices we will use the matrix norm $\|A\|_\infty = \max_{i,j}|A_{i,j}|$; this choice is convenient, but not essential.

For the time domain case, define analogously to (4.49) the set

$$W_{\gamma,T}^m := \left\{ v(t) \in l_2^m : \left\| \frac{1}{\|v\|_2^2} R_v(\tau) - \delta(\tau)\frac{1}{m}I \right\|_\infty \le \gamma, \quad |\tau| \le T \right\}. \tag{4.63}$$

To interpret the normalizing factor $\frac{1}{m}$ in (4.63), note that if $R_v(0) = \frac{1}{m}I$, then $\|v\|_2^2 = 1$ from (4.60). For the corresponding definition of $\|\mathbf{H}\|_{W_{\gamma,T}^m}$, we scale the input norm by this factor,

$$\|\mathbf{H}\|_{W_{\gamma,T}^m} := \sup\{\|\mathbf{H}v\| : v \in W_{\gamma,T}^m, \frac{1}{m}\|v\|_2^2 \le 1\}. \tag{4.64}$$

This normalization is equivalent to assigning norm 1 to the ideal unit spectrum $s_v(\omega) = I$, and is convenient to match the standard definitions for the $\mathcal{H}_2$ norm. With this definition, a bound similar to (4.53) can be obtained from (4.61), leading to

$$\lim_{\substack{\gamma\to 0 \\ T\to\infty}} \|\mathbf{H}\|_{W_{\gamma,T}^m} = \|\mathbf{H}\|_2. \tag{4.65}$$

For the frequency domain case, consider

$$F_v(s) = \int_0^s v(\omega)v(\omega)^* \frac{dw}{2\pi} - \frac{s}{2\pi}\frac{\|v\|^2}{m} I_m \tag{4.66}$$

and define the set

$$\hat{W}_\eta^m := \left\{ v \in l_2(R^n) : \sup_{s\in[0,2\pi]} \|F_v(s)\|_\infty < \eta \right\}, \tag{4.67}$$

with the corresponding system norm

$$\|\mathbf{H}\|_{\hat{W}_\eta^m} := \sup\{\|\mathbf{H}f\| : v \in \hat{W}_\eta^m, \frac{1}{m}\|v\|_2^2 \le 1\}. \tag{4.68}$$

The following is an extension of Lemma 4.10

[3]This set can be defined in the class $\mathcal{BP}$, but this advantage is not very significant.

Lemma 4.11 *Let $Y(\omega)$ be a matrix function with components in $BV[0,2\pi]$. If $v \in \hat{W}_\eta^m$, then*

$$\left| \int_0^{2\pi} v(\omega)^* Y(\omega) v(\omega) \frac{d\omega}{2\pi} - \frac{\|v\|_2^2}{m} \int_0^{2\pi} \mathrm{trace}(Y(\omega)) \frac{d\omega}{2\pi} \right| \leq \eta \sum_{i,j} \left(|Y_{i,j}(2\pi)| + TV(Y_{i,j}) \right). \tag{4.69}$$

Using $Y(\omega) = H(e^{j\omega})^* H(e^{j\omega})$, Lemma 4.11 leads for an LTI $\mathbf{H}$ to

$$\|\mathbf{H}\|_{\hat{W}_\eta^m} \stackrel{\eta \to 0}{\longrightarrow} \|\mathbf{H}\|_2. \tag{4.70}$$

Remark:

If one is interested in analyzing white noise rejection for LTV systems, the $\mathcal{H}_2$ norm as in (2.23) is no longer meaningful, but it is natural give a definition based on this approach. In fact, (4.68) can be applied to any $\mathbf{H} \in \mathcal{L}(l_2)$, and (4.70) can be taken as a definition of $\|\mathbf{H}\|_2$. This system measure (a seminorm) captures the response to signals of flat spectrum, the interesting object from the point of view of applications, and extends the LTI definition. Alternatively, (4.64) and (4.65) could be used.

Conclusion

We have presented a non-standard approach for the characterization of white signals in discrete time. This methodology is motivated by the desire to combine white noise with the methods of robust control theory. So far we have shown that the approach is sensible and provides answers which agree with the more established methods.

The main pending question at this point is the tractability of the resulting worst-case analysis problems. This issue will only be settled in the subsequent chapters, but its consideration has in fact affected the choice of these definitions among many others which were explored. In particular, we have opted for signal sets based on Integral Quadratic Constraints (IQCs), which are the topic of the next chapter. Although the l_2 definitions could have been introduced without relation to the finite horizon theory, we will find that the choices which yield the most powerful analysis, are precisely those like $\hat{W}_\eta$ which are well grounded in the finite horizon statistical theory.

A natural question is to what extent these methods can be applied to continuous time white noise, which is known to be a complicated mathematical object. For disturbance rejection problems, the approach of frequency domain-based sets over l_2 can be extended to continuous time $\mathcal{L}_2$ signals, by constraining the cumulative spectrum to accuracy η only over a finite band B. By taking limits in the induced system norm with $\eta \to 0$ and $B \to \infty$, the response to continuous time white noise is adequately described. This approach will be developed in detail in the future.

Chapter 5

Integral Quadratic Constraints

In the decades of the 1960s and 1970s a large body of results was developed to analyze the stability of feedback configurations based on input-output descriptions of components, often referred to as "absolute stability theory". In this category fall the small-gain theorem, the passivity theorem, and numerous extensions based on multipliers (see [20, 89] and references therein), all of which characterize the possibly nonlinear components in terms of some constraint between the external signals.

In particular, the most powerful results arise when these signal constraints are *quadratic* in nature, as in the case of the passivity theorem. These Integral Quadratic Constraints (IQCs) can be written in the general form

$$\langle z, \mathbf{\Pi} z \rangle = \int_0^{2\pi} z(e^{j\omega})^* \Pi(e^{j\omega}) z(e^{j\omega})\, d\omega \quad \geq 0 \tag{5.1}$$

in the case of discrete time, where z is a vector signal in l_2 and $\mathbf{\Pi}$ is an LTI, self adjoint operator in $\mathcal{L}(l_2)$, with transfer function $\Pi(e^{j\omega}) = \Pi(e^{j\omega})^*$.

In the case of LTI nominal systems subject to nonlinearities or uncertainties which can be described by IQCs, powerful frequency domain and state-space tools are available to analyze stability. The central role of IQC's was promoted in the Russian school by Yakubovich [93, 94], who also proposed the use of the so-called S-procedure (see Section 5.1.1) to analyze systems subject to several nonlinearities or uncertainties. From this approach a complete framework for robustness analysis has developed, which to some degree parallels the "western" development of robust control theory summarized in Chapter 2. encompassing parametric, time-varying and nonlinear perturbations. For a recent survey see [52].

In particular, IQC-based methods are attractive because they result computationally in convex optimization problems. Generally speaking, it can be said that any robustness analysis problem which admits a convex characterization, corresponds to uncertainty described by IQCs. One recent example supporting the previous statement is given by the results

of Megretski [50, 49] which use IQC techniques to exactly characterize the constant scales condition (2.42) for robustness analysis under structured perturbations.

In this chapter we will give an overview of these results, and extend IQC techniques in a number of directions, including the characterization of slowly varying perturbations, and the use of IQCs in set descriptions of white noise. These characterizations, and some properties which are obtained at the end of the chapter, will provide a natural framework for the robust $\mathcal{H}_2$ performance problem which is the subject of Chapter 6.

5.1 IQCs and LTV Operators

In this section we establish a correspondence between IQCs and linear time-varying operators, which implies an equivalence between two ways of thinking about system uncertainty: one is to write an LFT in terms of an LTV perturbation as in Section 2.1.3, the other is to write an IQC involving the corresponding signals. Once again we remark, as in Section 2.1.2, that the restriction to linear operators may be relaxed: contractive nonlinear operators can also be described by an IQC.

We will show that scalar IQCs on l_2 such as (5.1) are in direct correspondence with *non-causal*, contractive operators on $\mathcal{L}(l_2)$. This makes IQCs a natural tool for extended robustness analysis questions where the perturbations are allowed to be non-causal. The material in Appendix B shows, however, that for the standard cases considered the causality issue can be dealt with separately. For this reason we will ignore questions of causality in this chapter.

5.1.1 IQCs for Arbitrary Operators

The simplest case of IQC is obtained by choosing

$$\Pi = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \quad z = \begin{bmatrix} q \\ p \end{bmatrix}, \tag{5.2}$$

which makes (5.1) equivalent to the norm inequality $\|q\|_2^2 \geq \|p\|_2^2$. We now represent this constraint in terms of a normalized operator perturbation:

Lemma 5.1 *Let $p \in l_2^m, q \in l_2^n$. The following are equivalent:*

$$(i)\ \|q\|_2^2 - \|p\|_2^2 \geq 0, \tag{5.3}$$

$$(ii)\ \exists \mathbf{\Delta} \in \mathcal{L}(l_2^n, l_2^m), \|\mathbf{\Delta}\| \leq 1 :\ \mathbf{\Delta} q = p. \tag{5.4}$$

Proof: $(ii) \Rightarrow (i)$ is immediate, for $(i) \Rightarrow (ii)$ choose

$$\mathbf{\Delta} : u \mapsto \frac{\langle q, u \rangle}{\|q\|^2} p. \tag{5.5}$$

■

Lemma 5.1 provides a translation between two types of specifications on a pair of system variables p, q: one is to impose that they are related by a contractive non-causal LTV operator, the other one is to specify that they satisfy the IQC (5.3).

This is a first example of the duality between the language of robust control as was presented in Chapter 2, and the IQC formulation. Although (5.1) appears to be more general, we will show in Chapter 7 that a complete duality can be obtained with a moderate extension of the standard robust control framework.

S-Procedure and Robustness Analysis

The IQC formulation can be used to address the question of robustness analysis under structured, time-varying uncertainty; the method which we now describe is due to Megretski and Treil [50]. Consider the robust stability analysis problem of Figure 5.1, which is a special case of Figure 2.9 for the case where the uncertainty is comprised of full blocks. The last block could also be used to describe $\mathcal{H}_\infty$ performance as in Figure 2.10.

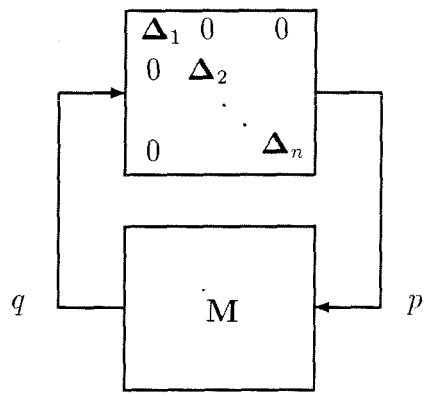


Figure 5.1: Robust stability analysis under structured uncertainty

We partition the signal p as $p = \mathrm{col}(p_1, \ldots, p_n)$ in accordance with the blocks $\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n$, and similarly for $q = \mathbf{M}p$. Consider the following set of quadratic forms:

$$\sigma_i(p) = \|(\mathbf{M}p)_i\|^2 - \|p_i\|^2, \quad i = 1 \ldots n. \tag{5.6}$$

If the σ_i are all non-negative at a certain $p \neq 0$, then from Lemma 5.1 there exist $\mathbf{\Delta}_i \in \mathcal{L}(l_2)$, $\|\mathbf{\Delta}_i\| \leq 1$, $i = 1, \ldots, n$, satisfying $\mathbf{\Delta}_i q_i = p_i$. Setting $\mathbf{\Delta} = \text{diag}[\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n]$ gives $(\mathbf{I} - \mathbf{\Delta}\mathbf{M})p = 0$, so $\mathbf{I} - \mathbf{\Delta}\mathbf{M}$ is not invertible. If, as convened, we ignore the issue of causality of $\mathbf{\Delta}$, this would violate robust stability.

The previous reasoning indicates that robustness analysis reduces to studying a simultaneous sign condition on a finite number of quadratic forms $\sigma_1, \ldots, \sigma_n$. The main method of analysis is to convert this condition to a single quadratic inequality by means of multipliers. This is known as the "S-procedure" [94], and in this particular case Megretski and Treil [50] have shown that the S-procedure is "lossless". Specifically, they show that given shift invariant quadratic forms $\sigma_1, \ldots, \sigma_n$ on l_2, the following are equivalent:

1. There does not exist $p \in l_2$ such that $\sigma_i > 0 \;\; i = 1, \ldots, n$.

2. There exist $x_i \geq 0$, $i = 1, \ldots, n$, not all zero such that $x_1\sigma_1 + \ldots + x_n\sigma_n \leq 0$.

In the case of (5.6), simple manipulations reduce the second condition to

$$\sum_{i=1}^{n} x_i \sigma_i(p) = \langle p, (\mathbf{M}^* X \mathbf{M} - X)p \rangle \leq 0, \tag{5.7}$$

where $X = \text{diag}[x_1 I, \ldots, x_n I] \geq 0$ is a scaling matrix of the form $\mathbb{X}$ considered in (2.35), corresponding to the uncertainty structure of Figure 5.1.

Now (5.7) is equivalent to condition $M(e^{j\omega})^* X M(e^{j\omega}) - X \leq 0$ across frequency, which (almost) corresponds to the LMI version of condition (2.42). Thus the previous argument has provided a sketch of the proof of the necessity of (2.42) for robust stability in the class $\mathbf{\Delta}^{LTV}$. Some refinements are required to obtain strict inequalities in the LMIs, and to take care of the causality issue; these will follow as a special case of the more general results of Chapter 6 and Appendix B.

Extension to $\delta\mathbf{I}$ Perturbations

The first extension to the previous method for robustness analysis, is to provide an IQC characterization of $\boldsymbol{\delta}\mathbf{I}$ where $\boldsymbol{\delta}$ is an arbitrary operator in $\mathcal{L}(l_2^1)$. It is now shown, following an idea of Doyle, that these operators are characterized by *matrix-valued* IQCs.

Lemma 5.2 *Let $p, q \in l_2^d$. The following are equivalent:*

$$(i) \quad \int_0^{2\pi} q(e^{j\omega})q(e^{j\omega})^* - p(e^{j\omega})p(e^{j\omega})^* \, d\omega \;\geq 0, \tag{5.8}$$

$$(ii) \quad \forall \zeta \in \mathbb{C}^d, \|\zeta^* q\|_2 \geq \|\zeta^* p\|_2 \,, \tag{5.9}$$

$$(iii) \quad \exists \boldsymbol{\delta} \in \mathcal{L}(l_2), \|\boldsymbol{\delta}\| \leq 1 : \;\; \boldsymbol{\delta}\mathbf{I} q = p. \tag{5.10}$$

Proof: $(i) \Leftrightarrow (ii)$ is immediate. Note that (i) is a matrix inequality.

$(iii) \Rightarrow (ii)$ follows from the linearity of δ. Let us show $(ii) \Rightarrow (iii)$.

If $\hat{q}_1, \ldots \hat{q}_r$, is an orthonormal basis of the subspace of l_2 spanned by the coordinates $q_1, \ldots q_d$ of q, we write $\hat{q} = Rq$, where R is an invertible matrix, $\hat{q} = \mathrm{col}(\hat{q}_1, \ldots \hat{q}_r, 0, \ldots 0)$. Let $\hat{p} = Rp$, then (ii) implies $\hat{p} = \mathrm{col}(\hat{p}_1, \ldots \hat{p}_r, 0, \ldots 0)$, and

$$\|\zeta_1 \hat{q}_1 + \ldots + \zeta_r \hat{q}_r\| \geq \|\zeta_1 \hat{p}_1 + \ldots + \zeta_r \hat{p}_r\| \quad \forall \zeta_1, \ldots \zeta_r \in \mathbb{C}. \tag{5.11}$$

Now define

$$\boldsymbol{\delta} : u \mapsto \sum_{i=1}^{r} \langle \hat{q}_i, u \rangle \hat{p}_i. \tag{5.12}$$

Then $\boldsymbol{\delta} : \hat{q}_i \mapsto \hat{p}_i, i = 1 \ldots r$, so $\boldsymbol{\delta}\mathbf{I}\, \hat{q} = \hat{p}$ which implies $\boldsymbol{\delta}\mathbf{I}\, q = p$.

Also, by (5.11) and the Bessel inequality, $\|\boldsymbol{\delta} u\| \leq \| \sum_{i=1}^{r} \langle \hat{q}_i, u \rangle \hat{q}_i \| \leq \|u\|$, so $\|\boldsymbol{\delta}\| \leq 1$.

■

The above characterization can be used to show the necessity of the constant scales condition (2.42) for the case where $\boldsymbol{\Delta}$ contains $\boldsymbol{\delta}\mathbf{I}$ blocks. To do this, the quadratic form σ_i in (5.6) must be replaced by the matrix valued quadratic function

$$\Sigma_i(p) = \int_0^{2\pi} \left[(Mp)_i(e^{j\omega})(Mp)_i^*(e^{j\omega}) - p_i(e^{j\omega}) p_i^*(e^{j\omega}) \right] \frac{d\omega}{2\pi}. \tag{5.13}$$

The corresponding S-procedure argument is given in Chapter 8.

5.1.2 IQCs for Slowly Varying Operators

The previous characterization for LTV operators in terms of IQCs is closely related to the availability of convex robustness analysis conditions such as (2.42).

If we look for a similar characterization of LTI uncertainty, and once again ignore the issue of causality, we find that for q, p in l_2,

$$\exists\, \boldsymbol{\Delta} \in \mathcal{L}(l_2),\ \mathrm{LTI},\ \|\boldsymbol{\Delta}\| \leq 1 :\ \boldsymbol{\Delta} q = p \Longleftrightarrow |q(e^{j\omega})|^2 \geq |p(e^{j\omega})|^2 \ a.e.,\ \omega \in [0, 2\pi]. \tag{5.14}$$

Unfortunately, this quadratic constraint is not an IQC, since there is no integration over frequency. The strong properties of IQCs to be presented in Section 5.3, depend fundamentally on this integration.

One can think of approximating the ideal constraint (5.14) by an IQC, by integrating $|q(e^{j\omega})|^2 - |p(e^{j\omega})|^2$ over a small frequency interval. It turns out that such constraint characterizes the class $\boldsymbol{B}^{\boldsymbol{\nu}}$ of slowly varying perturbations, in the sense introduced in (2.19).

To show this, we begin by considering signals which are supported in a frequency interval, and satisfy a norm inequality. The following lemma is closely related to Lemma 3.1 for sinusoidal signals.

Lemma 5.3 *Let p, q be signals in $l_2(\mathbb{Z})$ with Fourier transforms supported in a common interval $[s, s+h]$, and $\|q\|^2 \geq \|p\|^2$. Setting $\nu = 2\sin(\frac{h}{2})$, there exists $\boldsymbol{\Delta} \in \boldsymbol{B}^{\boldsymbol{\nu}}$, $\|\boldsymbol{\Delta}\| \leq 1$, such that $\boldsymbol{\Delta} q = p$.*

Proof: Without loss of generality, $\|q\| = 1$. We use the same operator $\boldsymbol{\Delta}$ from Lemma 5.1, $\boldsymbol{\Delta} : u \mapsto \langle q, u \rangle p$, which satisfies $\boldsymbol{\Delta} q = p$ and $\|\boldsymbol{\Delta}\| \leq 1$. We now calculate $\|\boldsymbol{\lambda}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\lambda}\|$.

$$\|(\boldsymbol{\lambda}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\lambda})u\|_2^2 = \| \langle q, u \rangle \boldsymbol{\lambda} p - \langle q, \boldsymbol{\lambda} u \rangle p\|_2^2 = \frac{1}{2\pi} \int_s^{s+h} | \langle q, u \rangle e^{j\omega} - \langle q, \boldsymbol{\lambda} u \rangle |^2 |p(e^{j\omega})|^2 d\omega. \tag{5.15}$$

The first factor inside the integral can be bounded by

$$| \langle q, (e^{j\omega} - \boldsymbol{\lambda})u \rangle | \leq \int_s^{s+h} |q(e^{j\bar{\omega}})| \, |e^{j\omega} - e^{j\bar{\omega}}| |u(e^{j\bar{\omega}})| \frac{d\bar{\omega}}{2\pi}. \tag{5.16}$$

It is easily verified that for $\omega, \bar{\omega} \in [s, s+h]$, $|e^{j\omega} - e^{j\bar{\omega}}| \leq 2\sin\frac{h}{2} = \nu$. Substitution back into (5.16), and the Cauchy-Schwarz inequality give

$$| \langle q, (e^{j\omega} - \boldsymbol{\lambda})u \rangle | \leq \nu \int_s^{s+h} |q(e^{j\bar{\omega}})| |u(e^{j\bar{\omega}})| \frac{d\bar{\omega}}{2\pi} \leq \nu \|q\|_2 \|u\|_2 = \nu \|u\|_2. \tag{5.17}$$

Now (5.17), (5.15) and $\|p\| \leq 1$ lead to

$$\|\boldsymbol{\lambda}\boldsymbol{\Delta} - \boldsymbol{\Delta}\boldsymbol{\lambda}\| \leq \nu. \tag{5.18}$$

■

We consider now the general case of signals which are not necessarily narrow-band, but satisfy an energy inequality at every band of a certain width h.

Lemma 5.4 *Let h be fixed, $\nu = 2\sin(\frac{h}{2})$. If the signals $p, q \in l_2(\mathbb{Z})$, satisfy*

$$\int_s^{s+h} (|q(e^{j\omega})|^2 - |p(e^{j\omega})|^2) d\omega \geq 0 \tag{5.19}$$

for every $s \in [0, 2\pi]$, then there exists $\boldsymbol{\Delta} \in \boldsymbol{B}^{\boldsymbol{\nu}}$, $\|\boldsymbol{\Delta}\| \leq 1$, such that $\boldsymbol{\Delta} q = p$.

Proof: The idea, illustrated in Figure 5.2, is to decompose p and q in terms of a filter bank of enough resolution, apply Lemma 5.3 to each pair of components, and assemble together the resulting operators.

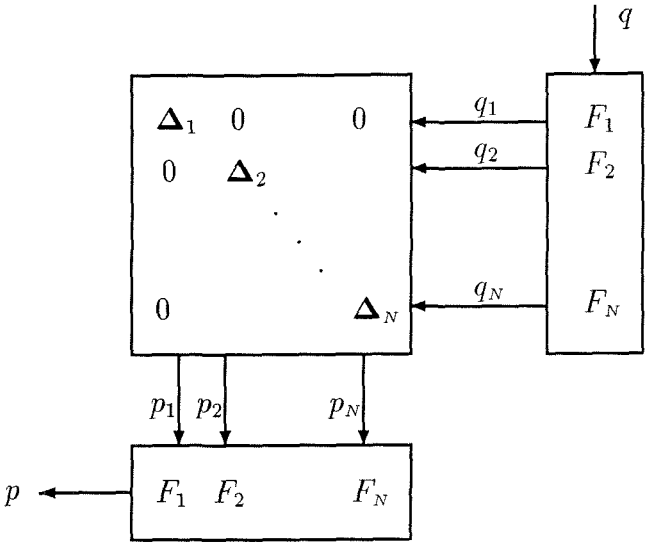


Figure 5.2: Illustration to the proof of Lemma 5.4

To make this precise, first choose an integer $N > \frac{2\pi}{h}$, and define $\mathbf{F}_1, \ldots, \mathbf{F}_N$ to be ideal bandpass filters, i.e. $F_i(e^{j\omega})$ is the indicator function of the interval $[\frac{2\pi(i-1)}{N}, \frac{2\pi i}{N}]$. Assuming for simplicity that the q and p are of the same vector dimension, define

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1\mathbf{I} \\ \vdots \\ \mathbf{F}_N\mathbf{I} \end{bmatrix}, \qquad \begin{bmatrix} q_1 \\ \vdots \\ q_N \end{bmatrix} = \mathbf{F}q, \qquad \begin{bmatrix} p_1 \\ \vdots \\ p_N \end{bmatrix} = \mathbf{F}p. \tag{5.20}$$

By definition, $\mathbf{F}$ is LTI and $\mathbf{F}^*\mathbf{F} = \mathbf{I}$. This implies $\|\mathbf{F}\| = \|\mathbf{F}^*\| = 1$.

Each pair of signals q_i, p_i have Fourier supports in a common interval of length h, and by hypothesis $\|q_i\| \geq \|p_i\|$. Applying Lemma 5.3 we can find a contractive operator $\mathbf{\Delta}_i \in \boldsymbol{B}^{\boldsymbol{\nu}}$, with $\mathbf{\Delta}_i q_i = p_i$. Define $\bar{\mathbf{\Delta}} = \mathrm{diag}[\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_N]$, and $\mathbf{\Delta} = \mathbf{F}^*\bar{\mathbf{\Delta}}\mathbf{F}$.

Then $\|\mathbf{\Delta}\| \leq 1$, and since $\mathbf{F}$ is LTI,

$$\|\boldsymbol{\lambda}\mathbf{F}^*\bar{\mathbf{\Delta}}\mathbf{F} - \mathbf{F}^*\bar{\mathbf{\Delta}}\mathbf{F}\boldsymbol{\lambda}\| = \|\mathbf{F}^*(\boldsymbol{\lambda}\bar{\mathbf{\Delta}} - \bar{\mathbf{\Delta}}\boldsymbol{\lambda})\mathbf{F}\| \leq \|\boldsymbol{\lambda}\bar{\mathbf{\Delta}} - \bar{\mathbf{\Delta}}\boldsymbol{\lambda}\| \leq \nu. \tag{5.21}$$

■

As a consequence of Lemma 5.4, we have characterized slowly varying perturbations in terms of a family of IQCs parametrized by a frequency variable $s \in [0, 2\pi]$. This suggests an extension to the approach of Section 5.1.1 for the analysis of robust stability in the configuration of Figure 5.1, when the uncertainty is assumed to be in $\boldsymbol{B}^{\boldsymbol{\nu}}$.

Given $p \in l_2$, we associate with it, for each i, the function $\varphi_i(p)$ of frequency

$$[\varphi_i(p)](s) = \int_s^{s+h} \left(|(Mp)_i(e^{j\omega})|^2 - |p_i(e^{j\omega})|^2 \right) \frac{d\omega}{2\pi}. \tag{5.22}$$

Instead of the scalar valued IQCs of (5.6), we now have "function-valued IQCs" given by the maps $\varphi_i : l_2 \rightarrow C_R[0, 2\pi]$ which assign to every $p \in l_2$ a continuous real function (5.22) of the frequency variable s.

In an analogous way to Section 5.1.1, the robustness analysis problem will reduce to finding a signal $p \neq 0$ which satisfies

$$[\varphi_i(p)](s) \geq 0, \quad i = 1, \ldots, n. \tag{5.23}$$

If such a signal is found, then a structured perturbation $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}_{NC}}$ can be constructed from Lemma 5.4, with $(\mathbf{I} - \mathbf{\Delta M})p = 0$.

From here on an "S-procedure"-type argument may be performed on the inequalities (5.23), and leads to an alternative proof of the result of Poolla and Tikku [64] regarding the necessity of the frequency-dependent scales condition (2.43) for robust stability over the class $\boldsymbol{B}^{\boldsymbol{\nu}}$. This proof is postponed until Chapter 6 where the extension to Robust $\mathcal{H}_2$ performance is also considered.

Remark: For perturbations of the form $\boldsymbol{\delta}\mathbf{I}$, and $\boldsymbol{\delta} \in \boldsymbol{B}^{\boldsymbol{\nu}}$, a natural combination of the ideas in Lemmas 5.2 and 5.4 gives a characterization of the form

$$\int_s^{s+h} q(e^{j\omega})q(e^{j\omega})^* - p(e^{j\omega})p(e^{j\omega})^* \, d\omega \; \geq 0, \quad s \in [0, 2\pi]. \tag{5.24}$$

5.2 IQCs and Set Descriptions of White Signals

In Section 5.1 IQCs were employed to describe a relationship between two signals; they can also be used to specify properties of individual signals. An important example of this second use is for the characterization of white noise.

The descriptions of white signals presented in Chapter 4 were based on the autocorrelation and the spectrum, which depend quadratically on the signal. This makes it possible to define the sets so that they correspond to IQCs, as is now discussed.

5.2.1 Autocorrelation Constraints

We first consider the definition of the set $W_{\gamma,T}$ in terms of a finite number of autocorrelation constraints. It was already pointed out by Megretski in [51] that IQCs could be used to constrain the signal autocorrelation. In fact, the set $W_{\gamma,T}$ over l_2 is defined in (4.49) for scalar signals, by the inequalities

$$\gamma r_v(0) \geq |r_v(\tau)| \quad \tau = 1, \ldots, T. \tag{5.25}$$

These can be rewritten, for real signals, in terms of the scalar IQCs

$$\sigma_\tau^{\pm}(v) := \|v \pm \boldsymbol{\lambda}^\tau v\|_2^2 - 2(1-\gamma)\|v\|_2^2 \geq 0. \tag{5.26}$$

This observation allows for the constraints to be combined with IQCs characterizing uncertainty, such as the σ_i introduced in (5.6) for the configuration of Figure 5.1. We can then pose a robust performance analysis problem where the disturbance signal is constrained to be in the set $W_{\gamma,T}$, as the problem of finding $p \neq 0$ verifying simultaneously

$$\begin{aligned} \sigma_i(p) = \|(\mathbf{M}p)_i\|^2 - \|p_i\|^2 &\geq 0, \quad i = 1, \ldots, n, \\ \sigma_\tau^{\pm}(p_n) &\geq 0, \quad \tau = 1, \ldots, T. \end{aligned} \tag{5.27}$$

To interpret (5.27) assume that the disturbance signal is p_n, and the IQC σ_n is a performance specification (equivalently, $\boldsymbol{\Delta}_n$ is included as the block $\boldsymbol{\Delta}_P$ in Figure 2.10). The constraints (5.27) test whether the performance can be violated using contractive $\boldsymbol{\Delta}_i$ and a disturbance signal $p_n \in W_{\gamma,T}$.

Once again, an S-procedure argument can be applied to the IQCs (5.27), leading to conditions with multipliers. We will not develop this approach here, since the problem of robust performance analysis under a finite number of additional IQCs in the disturbances will be reconsidered in Chapter 8 from a different (although equivalent) point of view. The application to robustness analysis over $W_{\gamma,T}$ will be included at that point as an example.

5.2.2 Frequency Domain Constraints

From the point of view of robust performance problems under white noise disturbances, we will find in Chapter 6 that more powerful results can be obtained from the frequency domain sets $\hat{W}_\eta$, which do not correspond to a finite number of scalar IQCs.

The definition (4.51) of $\hat{W}_\eta$ can, however, be represented by a constraint on a quadratic function which takes values on a function space, similarly to the situation of Section 5.1.2. For the case of scalar signals, consider the function $\rho : l_2 \rightarrow C_R[0, 2\pi]$, which associates to each signal v the function $F_v(s)$ from (4.50). Specifically, we have

$$[\rho(v)](s) := F_v(s) = \int_0^s |v(e^{j\omega})|^2 \frac{dw}{2\pi} - \frac{s}{2\pi}\|v\|_2^2, \tag{5.28}$$

which yields from (4.51) the characterization

$$v \in \hat{W}_\eta \Longleftrightarrow \rho(v) \in B(0, \eta). \tag{5.29}$$

In (5.29), $B(0, \eta)$ is the ball centered in 0 of radius η in the space $C_R[0, 2\pi]$, with the standard supremum norm.

In the case of multivariable noise signals $v \in l_2^m$, (4.66) indicates the natural definition of the map ρ, as

$$[\rho(v)](s) := F_v(s) = \int_0^s v(e^{j\omega})v(e^{j\omega})^* \frac{dw}{2\pi} - \frac{s}{2\pi m}\|v\|_2^2 I_m. \tag{5.30}$$

In this case ρ takes values in the space of continuous, hermitian matrix-valued functions defined on the interval $[0, 2\pi]$.

These characterizations are used in Chapter 6 for robustness analysis over the set $\hat{W}_\eta$.

5.3 Properties of Shift-Invariant Quadratic Functions

We have introduced several quadratic maps on l_2, which can be scalar-valued (5.6)-(5.26), matrix-valued (5.13), or function-valued (5.22)-(5.28). The key common property to all these definitions is that these maps are *shift invariant*, i.e. they take the same value on a signal p as on $\boldsymbol{\lambda} p$. This fact, together with the behavior of l_2 signals under arbitrary long shifts, will provide a basic property which is fundamental for future results. We now state this property for the scalar case.

Lemma 5.5 *Let $\boldsymbol{\Pi}$ be an LTI operator in $\mathcal{L}(l_2)$, and $\sigma(z) = \langle z, \boldsymbol{\Pi} z\rangle$. If $z, f \in l_2$, then*

$$\sigma(z + \boldsymbol{\lambda}^k f) \overset{k\to\infty}{\longrightarrow} \sigma(z) + \sigma(f). \tag{5.31}$$

Proof: From the time invariance of $\boldsymbol{\Pi}$, we have $\boldsymbol{\Pi}\boldsymbol{\lambda}^k = \boldsymbol{\lambda}^k\boldsymbol{\Pi}$ which implies

$$\left\langle z + \boldsymbol{\lambda}^k f, \boldsymbol{\Pi}(z + \boldsymbol{\lambda}^k f)\right\rangle = \langle z, \boldsymbol{\Pi} z\rangle + \langle f, \boldsymbol{\Pi} f\rangle + \left\langle \boldsymbol{\lambda}^k f, \boldsymbol{\Pi} z\right\rangle + \left\langle z, \boldsymbol{\lambda}^k \boldsymbol{\Pi} f\right\rangle. \tag{5.32}$$

The last two terms converge to 0 with k, since they are the inner products of a fixed l_2 signal with another one which is being shifted to infinity. This completes the proof.

■

The simplest case of Lemma 5.5 (for $\boldsymbol{\Pi} = \mathbf{I}$) is

$$\|z + \boldsymbol{\lambda}^k f\|_2^2 \overset{k\to\infty}{\longrightarrow} \|z\|_2^2 + \|f\|_2^2. \tag{5.33}$$

More interesting examples for which (5.31) holds are for the σ_i of (5.6) and the $\sigma_\tau^\pm$ of (5.26). Also, an analogous property to (5.31) holds for the matrix-valued quadratic functions Σ_i introduced in (5.13). The easiest way to show this is to observe that each entry in the matrix Σ_i is a scalar quadratic function of the form considered in Lemma 5.5.

We will require a similar property for the function-valued quadratic maps considered in (5.22) and (5.28). This property is covered by the following lemma.

Lemma 5.6 *Let* $\mathbf{\Pi}$ *be an LTI, self adjoint operator in* $\mathcal{L}(l_2)$, *and* $\beta : l_2 \rightarrow C_R[0, 2\pi]$ *be defined by*

$$[\beta(z)](s) = \int_0^s z(e^{j\omega})^* \Pi(e^{j\omega}) z(e^{j\omega}) d\omega. \tag{5.34}$$

If $z, f \in l_2$, *then*

$$\beta(z + \boldsymbol{\lambda}^k f) \stackrel{k\rightarrow\infty}{\longrightarrow} \beta(z) + \beta(f), \tag{5.35}$$

where (5.35) *means convergence in the topology of* $C_R[0, 2\pi]$ *(uniform convergence).*

Proof: Starting from (5.34), some algebra gives

$$\left[\beta(z + \boldsymbol{\lambda}^k f) - \beta(z) - \beta(f)\right](s) = 2Re \int_0^s [z^*(e^{j\omega}) \Pi(e^{j\omega}) e^{j\omega k} f(e^{j\omega})] d\omega. \tag{5.36}$$

It therefore suffices to show that the sequence of functions

$$\beta_k(s) := \int_0^s z^*(e^{j\omega})\ \Pi(e^{j\omega}) e^{j\omega k} f(e^{j\omega}) d\omega \tag{5.37}$$

converges uniformly to 0 as $k \rightarrow \infty$. Pointwise convergence follows from the fact that $\beta_k(s) = \left\langle \mathbf{H}_s z, \boldsymbol{\lambda}^k \mathbf{\Pi} f \right\rangle$, where $\mathbf{H}_s$ is the ideal filter $H_s(e^{j\omega}) = 1_{[0,s]}$. If convergence were not uniform, we could find $\epsilon > 0$, a subsequence k_j and points s_{k_j} with $|\beta_{k_j}(s_{k_j})| \geq \epsilon$. By compactness, taking a partial subsequence we can assume $s_{k_j} \stackrel{j\rightarrow\infty}{\longrightarrow} s_0$. Now

$$\beta_{k_j}(s_{k_j}) = \beta_{k_j}(s_0) + \int_{s_0}^{s_{k_j}} z^*(e^{j\omega})\ e^{j\omega k_j} \Pi(e^{j\omega}) f(e^{j\omega}) d\omega, \tag{5.38}$$

$$\Longrightarrow 0 < \epsilon \leq |\beta_{k_j}(s_{k_j})| \leq |\beta_{k_j}(s_0)| + \left| \int_{s_0}^{s_{k_j}} |z(e^{j\omega})| |\Pi(e^{j\omega}) f(e^{j\omega})| d\omega \right|. \tag{5.39}$$

The right-hand side of (5.39) converges to 0 from the pointwise convergence of β_{k_j}, and the fact that since z and $\mathbf{\Pi} f$ are l_2 functions, $|z(e^{j\omega})||\Pi(e^{j\omega}) f(e^{j\omega})| \in \mathcal{L}_1[0, 2\pi]$ from Cauchy-Schwarz. This is a contradiction. ■

As a consequence of Lemma 5.6. we find that if φ_i is defined as in (5.22), then

$$\varphi_i(z + \boldsymbol{\lambda}^k f) \stackrel{k\rightarrow\infty}{\longrightarrow} \varphi_i(z) + \varphi_i(f) \tag{5.40}$$

in the uniform topology. In fact such φ_i is defined as the difference between two quadratic functions of the general form of Lemma 5.6.

Similarly, for the white noise constraint $\rho(z)$ in (5.28). Lemmas 5.5 and 5.6 give

$$\rho(z + \boldsymbol{\lambda}^k f) \stackrel{k\rightarrow\infty}{\longrightarrow} \rho(z) + \rho(f) \tag{5.41}$$

in the space $C_R[0, 2\pi]$.

Remark: The previous properties can routinely be extended to the sum

$$z^{(0)} + \boldsymbol{\lambda}^k z^{(1)} + \cdots + \boldsymbol{\lambda}^{k(N-1)} z^{(N-1)} \tag{5.42}$$

of relatively shifted versions of N signals $z^{(0)}, \ldots, z^{(N-1)} \in l_2$. In particular, we state the properties

$$\| \sum_{r=0}^{N-1} \boldsymbol{\lambda}^{kr} z^{(r)} \|_2^2 \overset{k\to\infty}{\longrightarrow} \sum_{r=0}^{N-1} \|z^{(r)}\|^2, \tag{5.43}$$

$$\Lambda \left(\sum_{r=0}^{N-1} \boldsymbol{\lambda}^{kr} z^{(r)} \right) \overset{k\to\infty}{\longrightarrow} \sum_{r=0}^{N-1} \Lambda(z^{(r)}), \tag{5.44}$$

where in (5.44), the function Λ can be chosen to stand for σ_i as in (5.6), Σ_i as in (5.13), φ_i as in (5.22), or ρ as in (5.28-5.30), and the convergence is in the topology of the corresponding range space.

Chapter 6

Necessary and Sufficient Conditions for Robust $\mathcal{H}_2$ Performance

The results of this chapter show that $\mathcal{H}_2$ (LQG) performance specifications can be combined with structured uncertainty in the system, yielding robustness analysis conditions of the same nature and computational complexity as the corresponding conditions for $\mathcal{H}_\infty$ performance. These conditions are convex feasibility tests in terms of Linear Matrix Inequalities, and are proven to be necessary and sufficient under the same conditions as in the $\mathcal{H}_\infty$ case.

With these results, the tools of robust control can be viewed as coming full circle to treat the problem where it all began: guaranteeing margins for LQG regulators.

6.1 Historical Perspective

The advent of modern control in the 1960s brought a substantial transformation in control theory, with state-space methods and optimal control offering the promise of tractable, systematic methods for multivariable control design. This era was epitomized by the solution of the LQG control problem (see, for example, [3]), which provides an elegant, easily computable method for a well-motivated multivariable control design problem: optimizing the rejection of white noise disturbances for a closed loop system. It became increasingly clear in the late 1970s that modern control unfortunately provided limited tools to further treat model uncertainty, a fundamental requirement for a practical feedback theory and an issue which was often better addressed by the otherwise more primitive frequency domain techniques of classical control [41].

While LQ state feedback was shown to provide stability margin guarantees, further research [73] led to a counterexample showing that full LQG controllers had none [22]. This motivated efforts to reconcile LQG with classical methods [23], with some initial success in providing a robust LQG-based methodology [14]. The most popular development was

LQG/LTR [24, 1, 79], a multivariable version of classical loopshaping using LQG machinery. The problem of adding plant uncertainty directly to LQG remained unsolved, however, and ultimately these efforts pointed in other directions [24], particularly toward (structured) singular values and related methods ([28],[80]).

At about the same time as the critique of LQG robustness was becoming widely accepted, the new performance paradigm of $\mathcal{H}_\infty$ was being put forth [97]. It had close ties to the frequency domain and allowed singular value robustness conditions to be treated directly. More importantly, it allowed for the first time a very natural and relatively transparent analysis of robust performance ([25],[55]). While $\mathcal{H}_\infty$ soon replaced LQG (now referred to as $\mathcal{H}_2$) as the centerpiece of modern control, and research on $\mathcal{H}_\infty$ flourished in the 1980s, several developments helped bring $\mathcal{H}_2$ back into the picture.

The main weakness of $\mathcal{H}_\infty$ is that modeling signals using weighted norm balls ignores important structure, typically expressed in terms of spectral or correlation properties, as discussed in Section 4.1. The recognition of this conservatism has led to a resurgence of interest in the $\mathcal{H}_2$ performance measure. The desirable design specification, from both the performance and uncertainty points of view, appears to be in most cases Robust $\mathcal{H}_2$ performance: rejection of white signals in the worst-case over a set of plants. (The $\mathcal{L}_1$ theory is an attractive alternative to $\mathcal{H}_\infty$, but still suffers from a similar conservatism [15]).

Renewed interest in $\mathcal{H}_2$ performance was also stimulated by the striking fact that the most powerful computational solutions for the $\mathcal{H}_\infty$ control problem ([37],[26]) relied on the same state-space tools as LQG. This led to combining both performance measures (mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control, see e.g. [8, 44, 99, 27, 62]), and to various upper bounds for the $\mathcal{H}_2$ cost over a set of plants (e.g., [81, 63, 99, 32, 82]). In spite of these developments, the robust $\mathcal{H}_2$ problem lagged substantially behind $\mathcal{H}_\infty$ (or $\mathcal{L}_1$), where a sophisticated set of tools is available for the analysis of robust performance under structured uncertainty (see e.g. [55, 95, 5, 15]), including several results that exactly analyze robust performance with structured uncertainty in terms of computationally attractive convex conditions ([55, 43, 76, 50, 64]). No such results have previously been available for robust $\mathcal{H}_2$ performance.

In this chapter we provide the final step in the return of the $\mathcal{H}_2$ performance paradigm, casting it on an equal footing with $\mathcal{H}_\infty$. We present a convex condition for robust $\mathcal{H}_2$ performance analysis under structured uncertainty, of a very similar nature to the the corresponding condition for robust $\mathcal{H}_\infty$ performance, and with analogous properties. Computationally, it reduces to a Linear Matrix Inequality (LMI) over frequency which can be handled with analogous tools as in the $\mathcal{H}_\infty$ case. From a theoretical point of view, the condition is shown to be necessary and sufficient under the same assumptions for the uncertainty as in the corresponding $\mathcal{H}_\infty$ conditions.

Organization of this chapter

Section 6.2 contains a review of the convex conditions for robust $\mathcal{H}_\infty$ performance. Section 6.3 presents the robust $\mathcal{H}_2$ performance condition and states the corresponding theoretical results. Section 6.4 contains the proofs of sufficiency of these conditions, and Section 6.5 the necessity proofs, which rely on the IQC formulation of Chapter 5. In Section 6.6 we remark on the computational properties of this test. Section 6.7 compares these results to the previous work in the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ literature. Some remarks on robust $\mathcal{H}_2$ synthesis are given in Section 6.8. An example is given in Section 6.9.

6.2 Review of Robust $\mathcal{H}_\infty$ Performance Tests

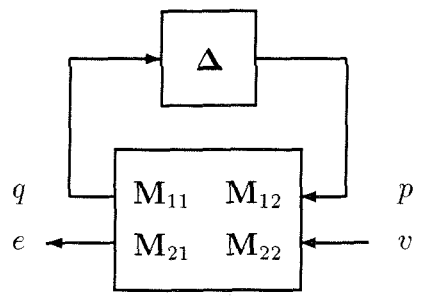


Figure 6.1: Robustness analysis configuration

To facilitate the comparison with the robust $\mathcal{H}_2$ performance conditions to be presented in Section 6.3, we give here a summary of the convex conditions for robust $\mathcal{H}_\infty$ performance analysis. These were already stated in Section 2.3.3 in the context of robust stability analysis; in this section we explicitly state them as robust $\mathcal{H}_\infty$ performance tests for the system of Figure 2.7, which is reproduced in Figure 6.1, and denoted by $(\mathbf{M}, \mathbf{\Delta})$.

As in Chapter 2, $\mathbf{\Delta}$ is an uncertainty operator with the structure (2.24); $\mathbb{X}$ denotes the corresponding set of commuting matrices, defined in (2.35), and $\mathbb{X}^P$ the positive subset. The set of frequency functions $X(e^{j\omega})$ which take values in the set $\mathbb{X}$ is denoted, as in (2.45), by $\boldsymbol{X}$ (analogously, $\boldsymbol{X^P}$ corresponding to $\mathbb{X}^P$).

The convex tests for robust $\mathcal{H}_\infty$ performance can be summarized as follows:

Condition 1 *There exists a function* $X(e^{j\omega}) \in \boldsymbol{X^P}$ *such that for all* $\omega \in [0, 2\pi]$,

$$M(e^{j\omega})^* \begin{bmatrix} X(e^{j\omega}) & 0 \\ 0 & I \end{bmatrix} M(e^{j\omega}) - \begin{bmatrix} X(e^{j\omega}) & 0 \\ 0 & I \end{bmatrix} < 0. \qquad (6.1)$$

Condition 1 is in fact the LMI version of the test (2.44) which was presented in Section 2.3.3. In this case the last block in the scaling matrix is normalized to be the identity matrix; this block corresponds to the performance block $\mathbf{\Delta}_P$ which appears if we convert $\mathcal{H}_\infty$ performance to robust stability as in Figure 2.10.

The inequality in (6.1) has been stated pointwise in frequency, which is appropriate since we are assuming a finite dimensional LTI system $\mathbf{M}$. In that case $M(e^{j\omega})$ is continuous, and without loss of generality $X(e^{j\omega})$ can be chosen continuous, or even rational (see [64], or Lemma 6.7 below). The inequality (6.1) can be interpreted equivalently in the operator sense,

$$\mathbf{H} < 0 \iff \exists \epsilon > 0 : \langle z, \mathbf{H}z \rangle < \epsilon \|z\|^2 \ \ \forall z \in l_2.$$

We now summarize the known results for this test, in reference to Definition 3 of robust $\mathcal{H}_\infty$ performance. For the general case of frequency dependent scales, we can state the following result:

Proposition 6.1 *If Condition 1 holds for a function $X(e^{j\omega}) \in \boldsymbol{X}^{\boldsymbol{P}}$, then the system $(\mathbf{M}, \mathbf{\Delta})$ has robust $\mathcal{H}_\infty$ performance for $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{LTI}}}$.*

Condition 1 is in general conservative for LTI uncertainty; the exact tests involving μ were given in Chapter 2. It remains, however, as an attractive condition since exact computation of μ is not tractable. Also, computational experience with lower bounds such as those used in [5] shows evidence that the two tests are not far apart, at least for full block structures.

Another indication that Condition 1 has small conservatism in the LTI case is the fact that it is exact for perturbations with an arbitrary small amount of time variation, in the sense of (2.19), as shown in [64]:

Proposition 6.2 *There exists $\nu > 0$ such that the system $(\mathbf{M}, \mathbf{\Delta})$ has robust $\mathcal{H}_\infty$ performance for $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{\nu}}}$ if and only if there exists a function $X(e^{j\omega}) \in \boldsymbol{X}^{\boldsymbol{P}}$ satisfying Condition 1.*

Finally, if we allow an unrestricted time variation, robustness analysis is obtained from Condition 1 by imposing $X(e^{j\omega})$ to be a constant function, as shown by [76, 50]:

Proposition 6.3 *The system $(\mathbf{M}, \mathbf{\Delta})$ has robust $\mathcal{H}_\infty$ performance with $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{LTV}}}$ if and only if there exists a constant matrix $X(e^{j\omega}) \equiv X \in \mathbb{X}$ satisfying Condition 1.*

6.3 Robust $\mathcal{H}_2$ Performance Tests

In this section we provide conditions for Robust $\mathcal{H}_2$ performance for the uncertain system of Figure 6.1, which are analogous to those presented in Section 6.2 for Robust $\mathcal{H}_\infty$ performance. In the sequel, m is the dimension of the input signal v. We now state an analysis test, which is a convex feasibility condition on the unknowns X, Y.

Condition 2 *There exists $X(e^{j\omega}) \in \boldsymbol{X}^{\boldsymbol{P}}$, and a matrix function $Y(e^{j\omega}) = Y^*(e^{j\omega})$ taking values in $\mathbb{C}^{m\times m}$, such that*

$$M(e^{j\omega})^* \begin{bmatrix} X(e^{j\omega}) & 0 \\ 0 & I \end{bmatrix} M(e^{j\omega}) - \begin{bmatrix} X(e^{j\omega}) & 0 \\ 0 & Y(e^{j\omega}) \end{bmatrix} < 0 \tag{6.2}$$

holds for all $\omega \in [0, 2\pi]$, and

$$\int_0^{2\pi} \mathrm{trace}(Y(e^{j\omega}))\frac{d\omega}{2\pi} < 1. \tag{6.3}$$

This condition is in fact very similar to Condition 1 for Robust $\mathcal{H}_\infty$ performance. The only addition is the incorporation of the function $Y(e^{j\omega})$. Heuristically, for $m = 1$, $Y(e^{j\omega})$ allows for the gain to be larger than 1 at some frequencies, provided that it is compensated at other frequencies by keeping the total effect $\int Y(e^{j\omega})d\omega$ less than 1; this imposes an "average over frequency" performance which corresponds to the $\mathcal{H}_2$ norm.

Remark:

To make Condition 2 precise we must specify the classes of functions $X(e^{j\omega})$, $Y(e^{j\omega})$ which are being considered.

- We will concentrate on the case where $\mathbf{M}$ is finite dimensional LTI; here the functions can be chosen to be rational without loss of generality, as shown in Lemma 6.7 below.
- The necessity proofs given in Section 6.5 are valid also for infinite dimensional $\mathbf{M}$, and yield functions X, Y of bounded variation (in $BV[0, 2\pi]$, see Appendix A). In this case (6.2) must be interpreted as an operator inequality.

The main result of this chapter is that this test answers the robust $\mathcal{H}_2$ performance problem. As in the case of Condition 1, different results can be stated in accordance with the nature of the perturbations in $\boldsymbol{\Delta}$, which exactly parallel Propositions 6.1, 6.2 and 6.3.

For the first one involving LTI perturbations, the LFT $\mathbf{\Delta} \star \mathbf{M}$ is an LTI system so the result can be stated using the standard definition (2.23) of the $\mathcal{H}_2$ norm, and proved with elementary frequency domain tools.

Theorem 6.4 *Suppose Condition 2 holds for matrix functions $X(e^{j\omega})$,$Y(e^{j\omega})$. If $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTI}}$, then the system is robustly stable and*

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTI}}} \|\mathbf{\Delta} \star \mathbf{M}\|_2 < 1. \tag{6.4}$$

We now show that Condition 2 has the same necessity properties as Condition 1 for the $\mathcal{H}_\infty$ case. To state the following results for which include non-LTI systems, we adopt the approach of (4.70) and give the following definition:

Definition 7 *The uncertain system $(\mathbf{M}, \mathbf{\Delta})$ with input $v \in l_2^m$ has robust $\mathcal{H}_2$ performance if it is robustly stable, and there exists $\eta > 0$ such that*

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}}} \|\mathbf{\Delta} \star \mathbf{M}\|_{\hat{W}_\eta^m} < 1. \tag{6.5}$$

The following necessary and sufficient conditions exactly characterize Condition 2 for both the frequency dependent and constant scales cases.

Theorem 6.5 *There exists $\nu > 0$ such that the system $(\mathbf{M}, \mathbf{\Delta})$ has robust $\mathcal{H}_2$ performance for $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^\nu}$ if and only if there exist bounded variation functions $X(e^{j\omega}) \in \boldsymbol{X}^{\boldsymbol{P}}$, $Y(e^{j\omega})$ satisfying Condition 2.*

Theorem 6.6 *The system $(\mathbf{M}, \mathbf{\Delta})$ has robust $\mathcal{H}_2$ performance for $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$, if and only if there exists a constant matrix $X \in \mathbb{X}$, and a bounded variation function $Y(e^{j\omega})$, satisfying Condition 2.*

Discussion

- These results are exact counterparts of Propositions 6.1, 6.2 and 6.3. In fact, Condition 2 can be regarded as a summary of tractable exact conditions for robustness analysis. Setting the blocks in X to be either constant or frequency-varying selects between LTV or LTI (slowly-varying) uncertainty. Selecting either a constant or frequency varying Y chooses between maximum over frequency or average over frequency performance (roughly, $\mathcal{H}_\infty$ or $\mathcal{H}_2$ performance; in rigor, for the multivariable case the norms also differ spatially, so for $\mathcal{H}_\infty$ the trace must be removed from (6.3)).

- Of these choices, it appears that the most natural "pairings" are $\mathcal{H}_\infty$ performance with LTV uncertainty, and $\mathcal{H}_2$ performance with LTI/slowly varying uncertainty.

  In the former, one has no information on the frequency content of the disturbance, or on the frequency domain properties of the uncertainty, which is allowed to transfer energy between frequencies (see Chapter 3). Correspondingly, the necessary and sufficient condition involves only constant scalings.

  In the latter, we have information on the spectral content of the disturbance, and also we are narrowly specifying the uncertainty in the frequency domain; the corresponding test involves frequency dependent scalings.

- Still, there may be situations which call for the other combinations of performance specifications and uncertainty. Also, combined uncertainty structures as those in Chapter 3 can be characterized by the corresponding combination of constant and frequency dependent X scales, and combinations of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance can be studied (see Section 6.7) by including Y terms only for the signals which are assumed white. For any of these combinations, we can state an exact characterization of the robust performance problem for which Condition 2 is necessary and sufficient, using the methods developed in the rest of this chapter.

- Theorem 6.5 gives indication that there is mild conservatism involved in using Condition 2 for LTI uncertainty, analogously to Proposition 6.2 for $\mathcal{H}_\infty$ performance.

  In the $\mathcal{H}_\infty$ case there was also supporting empirical evidence with computation of lower bounds for the LTI case based on μ, which is not available for $\mathcal{H}_2$. In fact, the restriction on *causality* of the LTI perturbations will provide an additional gap for $\mathcal{H}_2$ performance. To see this, consider the case of unstructured uncertainty, where the $\mathcal{H}_\infty$ conditions are known to be exact [55]. In the $\mathcal{H}_2$ case, with scalar inputs, it is easy to show that Condition 2 is exact for non-causal LTI perturbations, by simply choosing $\mathbf{\Delta}(e^{j\omega})$ to produce the worst gain at every frequency; this interpolation is in general only possible with non-causal $\mathbf{\Delta}$. The gap due to causality has not been quantified in general, but the results of [99] (see Section 6.7) suggest that it is not significant.

  In any case, the only necessary conditions available for the $\mathcal{H}_\infty$ and the $\mathcal{H}_2$ frequency dependent scales tests are Proposition 6.2 and Theorem 6.5, both indicating that these gaps are a modest price to pay for a convex characterization.

6.4 Sufficiency

In this section we prove the sufficiency of Condition 2 for robust $\mathcal{H}_2$ performance in the different cases of Theorems 6.4, 6.5 and 6.6.

A useful fact for the case of finite dimensional $\mathbf{M}$ is that without loss of generality, the frequency dependent scales may be assumed to be rational functions:

Lemma 6.7 *Suppose Condition 2 is satisfied by bounded variation functions $\bar{X}(e^{j\omega})$, $\bar{Y}(e^{j\omega})$. If $M(\lambda)$ is rational, there exist rational functions $D(\lambda)$, $E(\lambda)$ such that:*

- *D, D^{-1}, E, E^{-1} are in $\mathcal{RH}_\infty$.*
- *$D(\lambda)$ has the structure $\mathbb{X}$.*
- *Condition 2 is satisfied with $X(e^{j\omega}) = D(e^{j\omega})^* D(e^{j\omega})$, $Y(e^{j\omega}) = E(e^{j\omega})^* E(e^{j\omega})$.*

Proof: First, a topological argument implies that the scales $X(e^{j\omega})$, $Y(e^{j\omega})$ can be chosen to be continuous on $\mathbb{T}$. A sketch of this argument is the following:

1. Construct a finite open covering $\{\mathcal{B}_i\}$ of $\mathbb{T}$ such that in each open set $\mathcal{B}_i$, (6.2) is satisfied by a constant $\bar{X}_i > 0$, $\bar{Y}_i$. This is possible by continuity of $M(e^{j\omega})$, and compactness of $\mathbb{T}$. Also, the $\mathcal{B}_i$ can be chosen to be frequency intervals with a bound on their maximum length, and on the amount of overlap.

2. Consider a continuous partition of the unity $\{f_i(e^{j\omega})\}$ for $\mathbb{T}$, subordinate to the $\mathcal{B}_i$. This means that $f_i \geq 0$, continuous, $\mathrm{supp}(f_i) \subset \mathcal{B}_i$, and $\sum_i f_i(e^{j\omega}) \equiv 1$. Then choose $X(e^{j\omega}) := \sum_i f_i(e^{j\omega})\bar{X}_i$ and $Y(e^{j\omega}) := \sum_i f_i(e^{j\omega})\bar{Y}_i$ which satisfy (6.2) by convexity.

3. By the choice of the $\mathcal{B}_i$, $\int \mathrm{trace}(Y) = \sum_i \mathrm{trace}(\bar{Y}_i) \int f_i$ approximates a Riemann sum for the integral $\int \mathrm{trace}(\bar{Y})$; if $\mathcal{B}_i$ is fine enough, Y will satisfy (6.3).

Next, approximate the continuous scales uniformly on $\mathbb{T}$ by rational functions; due to Mergelyan's theorem [71] this can be done to an arbitrary precision, and therefore rational $X(e^{j\omega}) > 0$, $Y(e^{j\omega})$ are found satisfying (6.2) and (6.3). Note that from (6.2), $Y(e^{j\omega}) > 0$.

Finally, perform spectral factorizations (see [98]) to obtain D, E satisfying the requirements above (the factorization of X is done block by block to preserve the structure $\mathbb{X}$).

■

For the sufficiency proofs to follow we assume always that X is rational and factored as in Lemma 6.7. $\mathbf{D}$ is the system with transfer function $D(\lambda)$. In the case of Theorem 6.6, X is constant so we take $D = X^{\frac{1}{2}}$.

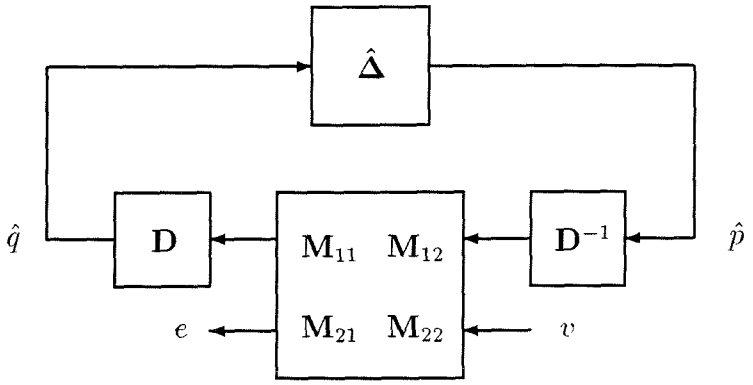


Figure 6.2: Equivalent system with $\mathbf{D}$ scaling

The robustness analysis setup of Figure 6.1 is equivalent to the configuration of Figure 6.2, where $\hat{\mathbf{\Delta}} = \mathbf{D}\mathbf{\Delta}\mathbf{D}^{-1}$. Incorporating the $\mathbf{D}$, $\mathbf{D}^{-1}$ together with $\mathbf{M}$, define

$$\hat{\mathbf{M}} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{M} \begin{bmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{6.6}$$

Clearly, robust stability for $(\mathbf{M}, \mathbf{\Delta})$ and $(\hat{\mathbf{M}}, \hat{\mathbf{\Delta}})$ are equivalent, and the same happens with performance since

$$\mathbf{\Delta} \star \mathbf{M} = \hat{\mathbf{\Delta}} \star \hat{\mathbf{M}}. \tag{6.7}$$

Also, (6.2) implies that for some $\epsilon > 0$, and all ω,

$$\hat{M}(e^{j\omega})^* \hat{M}(e^{j\omega}) - \begin{bmatrix} I & 0 \\ 0 & Y(e^{j\omega}) \end{bmatrix} < -\epsilon I. \tag{6.8}$$

6.4.1 Proof of Theorem 6.4

In this case, $\mathbf{D}$ and $\mathbf{\Delta}$ are both LTI operators which commute, so $\hat{\mathbf{\Delta}} = \mathbf{\Delta}$. The first block of the inequality (6.8) gives $\|\hat{\mathbf{M}}_{\mathbf{11}}\|_\infty < 1$, which implies from small gain the robust stability of $(\hat{\mathbf{M}}, \hat{\mathbf{\Delta}})$, hence of $(\mathbf{M}, \mathbf{\Delta})$.

To analyze the performance, consider a fixed frequency and the signal denominations of Figure 6.2. Then (6.8) gives

$$|e|^2 + |\hat{q}|^2 \leq (1-\epsilon)|\hat{p}|^2 + v^* Y(e^{j\omega}) v - \epsilon |v|^2. \tag{6.9}$$

Since $\Delta(e^{j\omega}) = \hat{\Delta}(e^{j\omega})$ is LTI, contractive we have $|\hat{p}|^2 \leq |\hat{q}|^2$, which leads to

$$v^*(\Delta \star M)(e^{j\omega})^*(\Delta \star M)(e^{j\omega})v = |e|^2 \leq v^* Y(e^{j\omega}) v - \epsilon |v|^2. \tag{6.10}$$

Since this holds for any v, we have

$$(\Delta \star M)(e^{j\omega})^*(\Delta \star M)(e^{j\omega}) \leq Y(e^{j\omega}) - \epsilon I \tag{6.11}$$

for each frequency. Computing the trace and integrating gives from (2.23) and (6.3)

$$\|\mathbf{\Delta} \star \mathbf{M}\|_2^2 \leq \int_0^{2\pi} \text{trace}(Y(e^{j\omega}))\frac{d\omega}{2\pi} - m\epsilon < 1. \tag{6.12}$$

■

6.4.2 Proof of Sufficiency for Theorem 6.6

In this case the scaling $D = X^{\frac{1}{2}}$ is a constant matrix, and from its spatial structure it commutes with the class $\boldsymbol{\Delta}^{\boldsymbol{LTV}}$, so once again $\hat{\mathbf{\Delta}} = \mathbf{\Delta}$. Then the small gain theorem implies robust stability, and we can write (6.9), which can be integrated across frequency to give

$$\|e\|^2 + \|\hat{q}\|^2 \leq (1-\epsilon)\|\hat{p}\|^2 + \int_0^{2\pi} v(e^{j\omega})^* Y(e^{j\omega}) v(e^{j\omega})\frac{d\omega}{2\pi} - \epsilon\|v\|^2. \tag{6.13}$$

Since $\|\hat{\mathbf{\Delta}}\| = \|\mathbf{\Delta}\| \leq 1$, then $\|\hat{p}\| \leq \|\hat{q}\|$, leading to

$$\|(\mathbf{\Delta} \star \mathbf{M})v\|^2 \leq \int_0^{2\pi} v(e^{j\omega})^* Y(e^{j\omega}) v(e^{j\omega})\frac{d\omega}{2\pi} - \epsilon\|v\|^2. \tag{6.14}$$

Fix $\eta > 0$; for $v \in \hat{W}_\eta^m$, we invoke Lemma 4.11 to bound

$$\int_0^{2\pi} v(e^{j\omega})^* Y(e^{j\omega}) v(e^{j\omega})\frac{d\omega}{2\pi} \leq \eta\, C(Y) + \frac{\|v\|^2}{m}\int_0^{2\pi} \text{trace}(Y(e^{j\omega}))\frac{d\omega}{2\pi}, \tag{6.15}$$

where $C(Y)$ depends only on Y. Substituting (6.15) into (6.14) and using (6.3) we arrive, for $\frac{\|v\|^2}{m} \leq 1$, to

$$\|(\mathbf{\Delta} \star \mathbf{M})v\|^2 \leq 1 - m\epsilon + \eta\, C(Y). \tag{6.16}$$

By choosing small enough η, the right hand side of (6.16) can be made less than 1; the bound holds for all $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{LTV}}}$, $v \in \hat{W}_\eta^m$, $\frac{\|v\|^2}{m} \leq 1$. Then from (4.68) we have

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{LTV}}}} \|(\mathbf{\Delta} \star \mathbf{M})\|_{\hat{W}_\eta^m} < 1. \tag{6.17}$$

■

6.4.3 Proof of Sufficiency for Theorem 6.5

If the perturbations are in $\boldsymbol{\Delta}^{\boldsymbol{\nu}}$ they do not commute with general LTI scalings $D(\lambda)$, so $\hat{\boldsymbol{\Delta}}$ is different from $\boldsymbol{\Delta}$. This necessitates a modification to the above proofs, which is analogous to the one provided in [64] for robust $\mathcal{H}_\infty$ performance.

Start with $D(\lambda) \in \mathcal{RH}_\infty$ from Lemma 6.7, and write the Taylor expansion

$$D(\lambda) = \sum_{t=0}^{\infty} D_t \lambda^t, \tag{6.18}$$

where the impulse response D_t converges exponentially to zero. In fact, substituting the frequency variable λ by the delay operator $\boldsymbol{\lambda}$, (6.18) provides a series expansion for the corresponding operator $\mathbf{D}$. Given $\boldsymbol{\Delta} \in \boldsymbol{\Delta}^{\boldsymbol{\nu}}$, we have

$$\boldsymbol{\Delta}\mathbf{D} - \mathbf{D}\boldsymbol{\Delta} = \sum_{t=0}^{\infty} D_t(\boldsymbol{\Delta}\boldsymbol{\lambda}^t - \boldsymbol{\lambda}^t\boldsymbol{\Delta}), \tag{6.19}$$

where we use the fact that the constant matrix $D_t \in \mathbb{X}$ commutes with $\boldsymbol{\Delta}$. It is simple to show that if $\boldsymbol{\Delta} \in \boldsymbol{B}^{\boldsymbol{\nu}}$, then $\|\boldsymbol{\Delta}\boldsymbol{\lambda}^t - \boldsymbol{\lambda}^t\boldsymbol{\Delta}\| < \nu t$, which gives the bound

$$\|\boldsymbol{\Delta}\mathbf{D} - \mathbf{D}\boldsymbol{\Delta}\| \leq \sum_{t=0}^{\infty} \bar{\sigma}(D_t)\, \|\boldsymbol{\Delta}\boldsymbol{\lambda}^t - \boldsymbol{\lambda}^t\boldsymbol{\Delta}\| \leq \nu \sum_{t=0}^{\infty} t\, \bar{\sigma}(D_t), \tag{6.20}$$

where the series converges due to the exponential decay of D_t. Consequently

$$\|\boldsymbol{\Delta} - \hat{\boldsymbol{\Delta}}\| = \|(\boldsymbol{\Delta}\mathbf{D} - \mathbf{D}\boldsymbol{\Delta})\mathbf{D}^{-1}\| \leq \nu \|\mathbf{D}^{-1}\| \sum_{t=0}^{\infty} t\, \bar{\sigma}(D_t), \tag{6.21}$$

which implies that for $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{\nu}}}$,

$$\|\hat{\boldsymbol{\Delta}}\| \leq 1 + \nu\, \kappa(D), \tag{6.22}$$

with $\kappa(D)$ depending only on D.

We now return to the sufficiency proof. Since $\|\hat{\mathbf{M}}_{\mathbf{11}}\|_\infty < 1$ from (6.8), then (6.22) implies that $(\hat{\mathbf{M}}, \hat{\boldsymbol{\Delta}})$ is robustly stable for sufficiently small ν, using small gain; therefore $(\mathbf{M}, \boldsymbol{\Delta})$ is robustly stable for this value of ν. Also, from (6.8) we derive (6.13) as in the previous proof; in this case (6.22) gives the bound $\|\hat{p}\| \leq (1 + \kappa\nu)\|\hat{q}\|$ for $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{\nu}}}$, which leads to

$$\|e\|^2 + \|\hat{q}\|^2 \leq (1-\epsilon)(1+\kappa\nu)^2\|\hat{q}\|^2 + \int_0^{2\pi} v(e^{j\omega})^* Y(e^{j\omega}) v(e^{j\omega}) \frac{d\omega}{2\pi} - \epsilon\|v\|^2. \tag{6.23}$$

By choosing ν small enough so that $(1-\epsilon)(1+\kappa\nu)^2 \leq 1$, (6.14) follows; from here on the proof continues exactly as in the case of Theorem 6.6, leading to

$$\sup_{\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{\nu}}}} \|(\boldsymbol{\Delta} \star \mathbf{M})\|_{\hat{\mathbb{W}}_\eta^m} < 1. \tag{6.24}$$

■

6.5 Necessity

The converse implications for Theorems 6.5 and 6.6 are based on the IQC formulation of Chapter 5, and require an extension of the "S-procedure losslessness" argument which was introduced in Section 5.1.1 for robust $\mathcal{H}_\infty$ performance.

For the purpose of $\mathcal{H}_2$ performance analysis, the signal sets $\hat{W}_\eta$ will be characterized by function-valued IQCs; correspondingly, the proofs below will rely on an infinite-dimensional convex analysis argument.

For simplicity we will first prove both Theorems in the situation of scalar inputs $v \in l_2^1$, and for uncertainty $\mathbf{\Delta} = \mathrm{diag}[\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n]$ consisting only of full blocks. At the end of the section we will explain the modifications required for the cases of multivariable noise and $\boldsymbol{\delta}\mathbf{I}$ perturbations.

6.5.1 Proof of Necessity in Theorem 6.6

In reference to Figure 6.1, let $z = col(z_1, \ldots, z_{n+1})$ be the vector of all inputs to the $\mathbf{M}$ system, where $z_1 \ldots z_n$ partition p in correspondence with the blocks $\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_n$, and $z_{n+1} = v$. Analogously $(\mathbf{M}z)_i, i = 1 \ldots n+1$ denotes the partition of the output of $\mathbf{M}$. Now consider the quadratic forms on l_2

$$\sigma_i(z) = \|(\mathbf{M}z)_i\|^2 - \|z_i\|^2, \quad i = 1 \ldots n+1. \tag{6.25}$$

From Chapter 5 we know that for $i = 1, \ldots, n$ the IQC $\sigma_i \geq 0$ represents the LTV perturbation $\mathbf{\Delta}_i$, if we ignore the restriction on causality. Also, the constraint $\sigma_{n+1} \geq 0$ tests the performance specification by comparing $\|e\|^2$ and $\|v\|^2$. If these were the only IQCs considered, we would be in the situation of Section 5.1.1 for robust $\mathcal{H}_\infty$ performance analysis. By adding the function-valued constraint

$$\rho(z_{n+1}) \in B(0, \eta), \tag{6.26}$$

where $\rho : l_2 \rightarrow C_R[0, 2\pi]$ is defined in (5.28), we impose that the disturbance $v = z_{n+1}$ is in the set $\hat{W}_\eta$. We now put together all the quadratic functions in

$$\Lambda(z) := (\sigma_1(z), \ldots, \sigma_{n+1}(z), \rho(z_{n+1})) \tag{6.27}$$

which maps l_2 into the real Banach space $\mathbb{V} = \mathbb{R}^{n+1} \oplus C_R[0, 2\pi]$.

If for some $z \neq 0$ we found that

$$\Lambda(z) \in \mathcal{K}_0 := \{(r_1, \ldots, r_{n+1}, g) : \ r_i \geq 0, \ \|g(s)\|_\infty < \eta\}. \tag{6.28}$$

then all the IQCs would be satisfied simultaneously and a contractive, non-causal, structured $\mathbf{\Delta}$ could be found such that the closed loop gain is at least 1 for a certain signal $v \in \hat{W}_\eta$, implying $\|\mathbf{\Delta} \star \mathbf{M}\|_{\hat{W}_\eta} \geq 1$. This reasoning leads to the following statement.

Proposition 6.8 *Suppose* $(\mathbf{M}, \mathbf{\Delta})$ *has robust* $\mathcal{H}_2$ *performance for* $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta LTV}}$. *For the fixed* η *given in Definition 7 and* $\epsilon > 0$, *define*

$$\mathcal{K}_\epsilon := \{(r_1, \ldots, r_{n+1}, g) : \; r_i > -\epsilon^2, \; \|g(s)\|_\infty < \eta\} \subset \mathbb{V}. \tag{6.29}$$

Also, with Λ *as in (6.27), define*

$$\nabla = \{\Lambda(z) : z \in l_2, \|z\| = 1\} \subset \mathbb{V}. \tag{6.30}$$

Then there exists $\epsilon > 0$ *such that* $\nabla \cap \mathcal{K}_\epsilon = \emptyset$.

Proof: See Appendix B. The main strengthening required to the previous argument is to include the causality of the perturbations, which is quite involved. Also, we have replaced the set $\mathcal{K}_0$ by the slightly larger set $\mathcal{K}_\epsilon$.

Proposition 6.8 reduces robust performance to a geometric separation condition in the space $\mathbb{V}$. To bring in techniques from convex analysis (see Appendix A), we note that $\mathcal{K}_\epsilon$ is open and convex in $\mathbb{V}$, and that

Lemma 6.9 *The closure* $\bar{\nabla}$ *of* ∇ *is convex in* $\mathbb{V}$.

Proof (Lemma): Consider $\Lambda_0 \in \text{co}(\nabla)$ (convex hull of ∇, see (A.6)):

$$\Lambda_0 = \sum_{r=0}^{N-1} \alpha_r \Lambda(z^{(r)}), \;\; \alpha_r \geq 0, \;\; \sum_{r=0}^{N-1} \alpha_r = 1, \;\; \|z^{(r)}\| = 1. \tag{6.31}$$

Define $f^k = \sum_{r=0}^{N-1} \sqrt{\alpha_r} \boldsymbol{\lambda}^{kr} z^{(r)}$. Recalling the property (5.44) which applies to every component of Λ, and the quadratic nature of Λ, it follows that

$$\Lambda(f^k) \stackrel{k\to\infty}{\longrightarrow} \sum_{r=0}^{N-1} \Lambda(\sqrt{\alpha_r} z^{(r)}) = \Lambda_0, \tag{6.32}$$

where the convergence is in the topology of the space $\mathbb{V}$. Also, $\|f^k\|^2 \stackrel{k\to\infty}{\longrightarrow} \sum_{r=0}^{N-1} \alpha_r = 1$ from (5.43), therefore $\Lambda(\frac{f^k}{\|f^k\|}) \stackrel{k\to\infty}{\longrightarrow} \Lambda_0$, so $\Lambda_0 \in \bar{\nabla}$.

We have shown $\text{co}(\nabla) \subset \bar{\nabla}$. This implies that $\text{co}(\bar{\nabla}) \subset \overline{\text{co}(\nabla)} \subset \bar{\nabla}$, so $\bar{\nabla}$ is convex. ∎

We are now in a position to apply the geometric version of the Hahn-Banach theorem stated in Theorem A.4, by choosing $\mathcal{K}_1 = \bar{\nabla}$, $\mathcal{K}_2 = \mathcal{K}_\epsilon$. Noting that $\bar{\nabla} \cap \mathcal{K}_\epsilon = \emptyset$ from Proposition 6.8 since $\mathcal{K}_\epsilon$ is open, we obtain the corresponding $\Gamma \in \mathbb{V}^*$, $\Gamma \neq 0$, $\alpha \in \mathbb{R}$. The structure of $\mathbb{V}$ and the Riesz representation theorem (see Appendix A) imply that Γ can be represented by $(x_1, \ldots, x_{n+1}, \Psi)$, where $x_i \in \mathbb{R}$, $\Psi \in BV[0, 2\pi]$. Then (A.9) yields

$$\sum_{i=1}^{n+1} x_i \sigma_i(z) + \int_0^{2\pi} \rho(z_{n+1}) d\Psi \leq \alpha, \quad \forall z \in l_2, \ \|z\| = 1, \tag{6.33}$$

$$\sum_{i=1}^{n+1} x_i r_i + \int_0^{2\pi} g d\Psi \geq \alpha, \quad \text{for} \begin{cases} r_i > -\epsilon^2 \\ \|g\|_\infty < \eta \end{cases}. \tag{6.34}$$

Concentrating on (6.34), we conclude that $x_i \geq 0$, $i = 1 \ldots n+1$; also, since $\mathcal{K}_\epsilon$ contains a ball of 0, and $\Gamma \neq 0$, then $\alpha < 0$. Now turning to (6.33), since $\sigma_i(z)$ are bounded functions it is possible to perturb the x_i to make them strictly positive, with (6.33) still holding for a new value $\alpha < 0$. Similarly, x_{n+1} can be normalized to 1.

It only remains to rewrite (6.33) using the definitions of σ_i, ρ. In the first place,

$$\sum_{i=1}^{n+1} x_i \sigma_i(z) = \left\langle z, \left(\mathbf{M}^* \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{M} - \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right) z \right\rangle \tag{6.35}$$

follows from (6.25), where $X = \mathrm{diag}[x_1 I, \ldots, x_n I] > 0$ is a constant scaling in $\mathbb{X}^P$.

Secondly, we treat the term corresponding to $\rho(z_{n+1}) = F_{z_{n+1}}$ from (5.28); an integration by parts (note $F_{z_{n+1}}(0) = F_{z_{n+1}}(2\pi) = 0$) gives

$$\int_0^{2\pi} F_{z_{n+1}}(s) d\Psi(s) = -\int_0^{2\pi} \Psi(s) F'_{z_{n+1}}(s) ds = -\int_0^{2\pi} \Psi(s) (|z_{n+1}(e^{js})|^2 - \|z_{n+1}\|^2) \frac{ds}{2\pi}. \tag{6.36}$$

Defining $Y(e^{j\omega}) = 1 + \Psi(\omega) - \int_0^{2\pi} \Psi(\omega) \frac{d\omega}{2\pi}$, the right hand side of (6.36) is equal to

$$-\int_0^{2\pi} (Y(e^{j\omega}) - 1) \, |z_{n+1}(e^{j\omega})|^2 \frac{d\omega}{2\pi} = \left\langle z, \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \mathbf{Y} \end{bmatrix} z \right\rangle. \tag{6.37}$$

Combining (6.35) and (6.37) into (6.33), we obtain

$$\left\langle z, \left(\mathbf{M}^* \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{M} - \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y} \end{bmatrix} \right) z \right\rangle \leq \alpha < 0, \quad \forall z \in l_2, \ \|z\| = 1. \tag{6.38}$$

This implies (6.2) holds. Finally, from the definition of $Y(e^{j\omega})$, we have $\int_0^{2\pi} Y(e^{j\omega}) \frac{d\omega}{2\pi} = 1$, and a small perturbation in Y will preserve (6.2) and yield

$$\int_0^{2\pi} Y(e^{j\omega}) \frac{d\omega}{2\pi} < 1, \tag{6.39}$$

which is (6.3) for this scalar case. ∎

6.5.2 Proof of the Necessity for Theorem 6.5

We are given $\nu > 0$ such that the system has robust $\mathcal{H}_2$ performance over $\boldsymbol{\Delta}^{\nu}$.

The main modification to the argument in Theorem 6.6 is to replace the σ_i of (6.25), suitable for arbitrary LTV operators, with the function-valued quadratic maps $\varphi_i : l_2 \to C_R[0, 2\pi]$ introduced in (5.22) to characterize the class $\boldsymbol{B}^{\nu}$. We write

$$[\varphi_i(z)](s) = \int_s^{s+h} \left(|(Mz)_i(e^{j\omega})|^2 - |z_i(e^{j\omega})|^2 \right) \frac{d\omega}{2\pi}, \quad i = 1, \ldots, n, \tag{6.40}$$

where $h > 0$ is defined such that $2 \sin \frac{h}{2} < \nu$.

The quadratic form $\sigma_{n+1}(z) = \|(\mathbf{M}z)_{n+1}\|^2 - \|z_{n+1}\|^2$ for performance, and the constraint $\rho(z_{n+1})$ for "whiteness" remain unchanged from Theorem 6.6. Correspondingly, define

$$\Lambda(z) := (\varphi_1(z), \ldots, \varphi_n(z), \sigma_{n+1}(z), \rho(z_{n+1})) \tag{6.41}$$

which maps l_2 into the real Banach space $\mathbb{V} = (C_R[0, 2\pi])^n \oplus \mathbb{R} \oplus C_R[0, 2\pi]$.

The following proposition, analogous to Proposition 6.8, is proved in Appendix B.

Proposition 6.10 *Suppose* $(\mathbf{M}, \boldsymbol{\Delta})$ *has robust* $\mathcal{H}_2$ *performance for* $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}}$. *For the fixed* η *given in Definition 7 and* $\epsilon > 0$, *define*

$$\mathcal{K}_\epsilon := \{(r_1(s), \ldots, r_n(s), r_{n+1}, g(s)) : \; r_i > -\epsilon^2, \; \|g(s)\|_\infty < \eta\} \subset \mathbb{V}. \tag{6.42}$$

Also, with Λ *as in (6.41) define*

$$\nabla = \{\Lambda(z) : z \in l_2, \|z\| = 1\} \subset \mathbb{V}. \tag{6.43}$$

Then there exists $\epsilon > 0$ *such that* $\nabla \cap \mathcal{K}_\epsilon = \emptyset$.

In (6.42) the constraints $r_i(s) > -\epsilon^2$ for $i = 1, \ldots, n$ are uniform bounds for $s \in [0, 2\pi]$. This makes $\mathcal{K}_\epsilon$ open and convex in the topology of $\mathbb{V}$.

Also, the closure of the set ∇ in (6.43) is convex. This follows from the same proof as Lemma 6.9, relying on the basic property (5.44) which was proved in Chapter 5 for any of the shift invariant quadratic maps.

Once again, we apply Theorem A.4 and obtain the corresponding functional $\Gamma \in \mathbb{V}^*$, $\Gamma \neq 0$, and $\alpha \in \mathbb{R}$. In this case the functionals have the form $(\xi_1, \ldots, \xi_n, x_{n+1}, \Psi)$ where $\xi_i \in BV[0, 2\pi]$, $x_{n+1} \in \mathbb{R}$, and $\Psi \in BV[0, 2\pi]$. We have the following:

$$\sum_{i=1}^{n} \int_0^{2\pi} \varphi_i(z) d\xi_i + x_{n+1}\sigma_{n+1}(z) + \int_0^{2\pi} \rho(z_{n+1}) d\Psi \;\leq\; \alpha, \;\; \forall z \in l_2, \;\; \|z\| = 1, \tag{6.44}$$

$$\sum_{i=1}^{n} \int_0^{2\pi} r_i d\xi_i + x_{n+1} r_{n+1} + \int_0^{2\pi} g d\Psi \;\geq\; \alpha, \;\; \text{for} \begin{cases} r_i(s) > -\epsilon^2, \; i \leq n \\ r_{n+1} > -\epsilon^2 \\ \|g\|_\infty < \eta \end{cases}. \tag{6.45}$$

Concentrating on (6.45), we conclude that $x_{n+1} \geq 0$, and that ξ_i are non-decreasing functions ($d\xi_i$ are non-negative measures) for $i = 1, \ldots, n$. Without loss of generality we can choose $\xi_i(0) = 0$, which makes $\xi_i(\omega) \geq 0$.

Also, since $\mathcal{K}_\epsilon$ contains a ball of 0, and $\Gamma \neq 0$, then $\alpha < 0$. This allows x_{n+1} to be perturbed and made strictly positive, and subsequently normalized to 1, satisfying (6.44) for scaled ξ_i, Ψ, and some $\alpha < 0$. To interpret this condition, we concentrate on one of the terms

$$\int_0^{2\pi} \varphi_i(z) d\xi_i = \int_0^{2\pi} \left[\int_s^{s+h} \left(|(Mz)_i(e^{j\omega})|^2 - |z_i(e^{j\omega})|^2 \right) \frac{d\omega}{2\pi} \right] d\xi_i(s). \tag{6.46}$$

It is a routine exercise to transform (6.46) by an integration by parts into

$$\int_0^{2\pi} \varphi_i(z) d\xi_i = \int_0^{2\pi} x_i(e^{j\omega}) \left(|(Mz)_i(e^{j\omega})|^2 - |z_i(e^{j\omega})|^2 \right) \frac{d\omega}{2\pi}, \tag{6.47}$$

where the function $x_i(e^{j\omega})$ is defined as

$$x_i(e^{j\omega}) = \begin{cases} \xi_i(2\pi) - \xi_i(2\pi + \omega - h) + \xi_i(\omega) & \text{for} \quad 0 \leq \omega \leq h \\ \xi_i(\omega) - \xi_i(\omega - h) & \text{for} \quad h \leq \omega \leq 2\pi \end{cases}. \tag{6.48}$$

Since ξ_i is non-decreasing, and $\xi_i \geq 0$, $x_i(e^{j\omega})$ is a non-negative function of frequency. Defining the frequency dependent scaling $X(e^{j\omega}) = \text{diag}[x_1(e^{j\omega})I, \ldots, x_n(e^{j\omega})I] \in \boldsymbol{X}$ we obtain from (6.47)

$$\sum_{i=1}^{n} \int_0^{2\pi} \varphi_i(z) d\xi_i + \sigma_{n+1}(z) = \left\langle z, \left(\mathbf{M}^* \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{M} - \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \right) z \right\rangle \tag{6.49}$$

which is analogous to (6.35) except that the scale $X(e^{j\omega})$ is frequency dependent. From here onwards the proof proceeds exactly as in Theorem 6.6 to take care of the term $\int_0^{2\pi} \rho(z_{n+1}) d\Psi$, and leads to the inequalities (6.38) and (6.39).

We have ensured $X(e^{j\omega}) \geq 0$, but strict inequality can be obtained at the end from continuity since $\alpha < 0$ in (6.38).

■

Remarks:

- If the whiteness constraints ρ are not included in the problem, this exact argument shows the necessity of Condition 1, with frequency dependent scales, for robust $\mathcal{H}_\infty$ performance in $\boldsymbol{\Delta}^{\nu}$. This is an alternative proof to the one given in [64].

- If a combination of arbitrary LTV blocks and $\boldsymbol{B}^{\nu}$ blocks is present, the proof follows exactly in the same way, by including the corresponding combination of σ and φ constraints. Thus we prove the necessity of the corresponding "mixed scales" condition for robust $\mathcal{H}_\infty$ performance (Theorem 3.6) or for robust $\mathcal{H}_2$ performance.

6.5.3 Extensions to the proofs

We now indicate briefly how the previous proofs must be modified to allow for multivariable white noise signals, or for $\boldsymbol{\delta}\mathbf{I}$ perturbations. We explain the changes in the simpler setup of Theorem 6.6, but analogous changes apply to Theorem 6.5.

Multivariable noise

If $v \in l_2^m$, then from (4.68) the performance quadratic constraint is changed to

$$\sigma_{n+1}(z) = \|(\mathbf{M}z)_{n+1}\|^2 - \frac{1}{m}\|z_{n+1}\|^2. \tag{6.50}$$

This amounts to a renormalization in the input, which is included for the reasons discussed in Section 4.4.3. Then (6.35) turns into

$$\sum_{i=1}^{n+1} x_i \sigma_i(z) = \left\langle z, \left(\mathbf{M}^* \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{M} - \begin{bmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \frac{1}{m}\mathbf{I} \end{bmatrix} \right) z \right\rangle. \tag{6.51}$$

Also, (5.30) indicates the definition for the whiteness constraint $\rho(z_{n+1})$:

$$[\rho(z_{n+1})](s) = F_{z_{n+1}}(s) = \int_0^s z_{n+1}(e^{j\omega}) z_{n+1}(e^{j\omega})^* \frac{d\omega}{2\pi} - \frac{s}{2\pi m}\|z_{n+1}\|^2 I_m.$$

ρ now takes values in the space of continuous, hermitian matrix-valued functions; the dual of this space can be identified with the space of hermitian, bounded variation matrix functions Ψ on $[0, 2\pi]$ (up to a constant matrix, for convenience we choose $\Psi(2\pi) = 0$), with the convention

$$\Gamma_\Psi(g) = \int_0^{2\pi} \mathrm{trace}(g(s) d\Psi(s)) := \sum_{i,j} \int_0^{2\pi} g_{i,j}(s) d\Psi_{i,j}(s).$$

The multivariable extension of (6.36) is

$$\int_0^{2\pi} \text{trace}(F_{z_{n+1}}(s)d\Psi(s)) \;=\; -\int_0^{2\pi} \text{trace}(\Psi(s)F'_{z_{n+1}}(s))ds \;=$$
$$-\int_0^{2\pi} z_{n+1}(e^{j\omega})^*\Psi(\omega)z_{n+1}(e^{j\omega})\frac{d\omega}{2\pi} + \frac{\|z_{n+1}\|^2}{m}\int_0^{2\pi} \text{trace}(\Psi(\omega))\frac{d\omega}{2\pi}. \tag{6.52}$$

Now define

$$Y(e^{j\omega}) = \frac{I}{m} + \Psi(\omega) - \frac{I}{m}\int_0^{2\pi} \text{trace}(\Psi(\omega))\frac{d\omega}{2\pi}. \tag{6.53}$$

This makes the final expression obtained in (6.52) equal to

$$-\int_0^{2\pi} z_{n+1}(e^{j\omega})^*\left(Y(e^{j\omega}) - \frac{I}{m}\right)z_{n+1}(e^{j\omega})\frac{d\omega}{2\pi} = \left\langle z, \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{m}\mathbf{I} - \mathbf{Y} \end{bmatrix} z \right\rangle. \tag{6.54}$$

Combining it with (6.51) leads to (6.2). Also, from the definition of Y we have

$$\int_0^{2\pi} \text{trace}(Y(e^{j\omega}))\frac{d\omega}{2\pi} = 1, \tag{6.55}$$

so (6.3) follows by perturbation.

δI operators in $\mathbf{\Delta}$

If the i-th block of $\mathbf{\Delta}$ is δI_{n_i}, then the scalar quadratic function σ_i must be replaced as in (5.13) by a matrix-valued function

$$\Sigma_i(z) = \int_0^{2\pi} \left[(Mz)_i(e^{j\omega})(Mz)_i^*(e^{j\omega}) - z_i(e^{j\omega})z_i^*(e^{j\omega})\right]\frac{d\omega}{2\pi}$$

which takes values in the space of hermitian $n_i \times n_i$ matrices. The functionals in this space are of the form $\Gamma_{X_i}(A) = \text{trace}(X_iA)$, where X_i is a full, hermitian matrix. The argument then proceeds in a similar fashion, X_i becoming a sub-block of the scaling matrix $X \in \mathbb{X}$.

More details on this case are given in the proof of Theorem 8.4.

6.6 Computational Issues

A test has been developed in the previous sections which characterizes robust $\mathcal{H}_2$ performance analysis of an uncertain system. This test is an infinite dimensional convex feasibility condition on the unknowns X and Y, specified as a Linear Matrix Inequality (LMI) across the frequency axis, of a similar complexity as Condition 1 for robust $\mathcal{H}_\infty$ performance.

In Section 2.3.3 we mentioned two standard approaches for handling the infinite dimensionality of these conditions: gridding the frequency axis, or parametrizing the scales in

terms of a basis of rational functions. Both approaches can indeed be applied to Condition 2, and involve minor modifications to their counterparts for Condition 1. We demonstrate this by commenting on the gridding approach for this problem: Condition 2 is approximated by considering frequency points $0 = \omega_0 < \omega_1 < \cdots < \omega_N = 2\pi$, and leads to the LMI problem

$$M(e^{j\omega_i})^* \begin{bmatrix} X_i & 0 \\ 0 & I \end{bmatrix} M(e^{j\omega_i}) - \begin{bmatrix} X_i & 0 \\ 0 & Y_i \end{bmatrix} < 0, \quad i = 1 \ldots N, \tag{6.56}$$

$$X_i > 0, \quad i = 1 \ldots N, \tag{6.57}$$

$$\frac{1}{2\pi} \sum_{i=1}^{N} \text{trace}(Y_i)(\omega_i - \omega_{i-1}) < 1, \tag{6.58}$$

where the unknowns Y_i are hermitian matrices and the X_i structured matrices.

For the LTV test, $X_i \equiv X$ is constant across the ω_i, which makes conditions (6.56-6.58) intrinsically coupled across frequency. For the LTI/slowly varying test, we use different variables X_i, $i = 1 \ldots N$. Although (6.58) still involves all frequency points, the following strategy can be used to decouple the problem across frequency:

- For each fixed frequency point ω_i, pose the problem:

  Minimize $\text{trace}(Y_i)$
  subject to

$$M(e^{j\omega_i})^* \begin{bmatrix} X_i & 0 \\ 0 & I \end{bmatrix} M(e^{j\omega_i}) - \begin{bmatrix} X_i & 0 \\ 0 & Y_i \end{bmatrix} < 0, \tag{6.59}$$

$$X_i > 0. \tag{6.60}$$

  The problem of minimizing a linear function of the unknowns, subject to an LMI constraint, falls in the class of eigenvalue problems (EVPs) considered in [10], and can be computed efficiently.

- Given all the solutions $Y_1 \ldots Y_N$, compute $\frac{1}{2\pi} \sum_{i=1}^{N} \text{trace}(Y_i)(\omega_i - \omega_{i-1})$, and compare the answer to 1. More directly, this sum will provide an approximation to the square of the worst-case $\mathcal{H}_2$ norm of the system; this follows from the fact that to test if the worst-case $\mathcal{H}_2$ norm is less than γ, it suffices to change 1 for γ^2 in (6.3) or (6.58).

6.7 Connections to Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Performance

In this section we relate these results to earlier work in the so-called mixed $\mathcal{H}_2/\mathcal{H}_\infty$ problem. There are many versions of this problem in the literature (a few are [8, 44, 99, 62]), all of which attempt to get a handle on robust $\mathcal{H}_2$ performance by studying first the situation where there is no uncertainty, but the performance specification is a combination of the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms.

A mixed $\mathcal{H}_2/\mathcal{H}_\infty$ performance problem can in fact be cast naturally in our setting, and leads to an analysis test which looks exactly like Condition 2, except that the scaling matrix $X(e^{j\omega})$ does not appear and is fixed to be the identity.

Proposition 6.11 *Consider a system* $\mathbf{M} = [\mathbf{M}_1 \ \mathbf{M}_2]$ *where the input* z *is partitioned in the vectors* z_1, $z_2 \in \mathbb{C}^m$. *The following are equivalent:*

$$\text{(i)}: \quad \exists \eta > 0: \ \sup \left\{ \|\mathbf{M}z\|^2 : \|z_1\|^2 + \frac{1}{m}\|z_2\|^2 \leq 1, \ z_2 \in \hat{W}_\eta \right\} < 1; \tag{6.61}$$

$$\text{(ii)}: \quad \textit{There exists } Y(e^{j\omega}) = Y(e^{j\omega})^* \in \mathbb{C}^{m\times m}:$$

$$M(e^{j\omega})^* M(e^{j\omega}) - \begin{bmatrix} I & 0 \\ 0 & Y(e^{j\omega}) \end{bmatrix} < 0 \quad \forall\, \omega \in [0, 2\pi], \tag{6.62}$$

$$\int_0^{2\pi} \text{trace}(Y(e^{j\omega}))\frac{d\omega}{2\pi} < 1. \tag{6.63}$$

This result is proved along the same lines as Theorem 6.6. only that since there is no uncertainty, one only considers the whiteness constraint $\rho(z) = F_{z_2}$, and the performance quadratic constraint $\sigma(z) = \|Mz\|^2 - \|z_1\|^2 - \frac{1}{m}\|z_2\|^2$, with $\|z_2\|^2$ weighted by $\frac{1}{m}$ for the reasons explained in Section 4.4.3.

Conditions (6.62) and (6.63) are therefore interpreted as a mixed performance problem where a portion of the input signal is constrained to be white. Various problems like this were considered in [99] for continuous time, with different assumptions on the relationship between z_1 and z_2. It is shown in [99] that the performance costs are not substantially different for these alternatives, and subsequently the attention is concentrated on the case where z_1 is restricted to be causally dependent on z_2. State-space methods for both analysis and synthesis for this alternative are given in [99] and the sequel paper [27] (a stochastic version appears in [62]).

Our condition, in contrast, corresponds to the case where there is no such causality restriction, which is only treated summarily in [99]. While at the level of the mixed performance problem it is not obvious which alternative to prefer, the version considered here has

advantages from the point of the robust $\mathcal{H}_2$ problem, since it is allows for the inclusion of X scales in a convex condition with the strong interpretation given in Theorems 6.5 and 6.6.

Remarks:

- These causality issues are strongly connected to the remarks of Section 6.3 regarding the causality of LTI perturbations. A possible conclusion from the results in [99] is that gaps due to this are not very significant, although this issue warrants further investigation.

- Analogously to the case treated in detail in [99], the mixed conditions (6.62-6.63) reduce to a finite dimensional test if $M(e^{j\omega})$ is rational. In fact, a Schur complement operation and some algebra shows that (6.62) is equivalent to

$$\|M_1\|_\infty < 1, \tag{6.64}$$

$$M_2^*(I - M_1 M_1^*)^{-1} M_2 < Y, \quad \forall\, \omega \in [0, 2\pi]. \tag{6.65}$$

  This implies that (6.62-6.63) can be tested by first checking (6.64), and then imposing that $\|N^{-1}M_2\|_2 < 1$, where N is the inversely stable spectral factor (see, e.g., [98]) satisfying $I - M_1 M_1^* = NN^*$. Both these operations can be computed efficiently by the same state-space techniques used in the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ literature.

6.8 Robust $\mathcal{H}_2$ Synthesis

Having obtained conditions for robust $\mathcal{H}_2$ performance analysis under structured uncertainty, it is natural to consider the problem of controller synthesis. If the nominal system $\mathbf{M}$ is obtained as the closed loop $\mathbf{G} \star \mathbf{K}$ in a feedback configuration, the problem is to design $\mathbf{K}$ such that $\mathbf{M}$ satisfies the robust $\mathcal{H}_2$ performance conditions.

It is unlikely that a tractable global solution will appear for this problem, since none is known for the case of $\mathcal{H}_\infty$ performance. As discussed in Section 2.4, the only general method for robust $\mathcal{H}_\infty$ synthesis is the so-called "D-K" iteration, where an an analysis step (Condition 1) is alternated with $\mathcal{H}_\infty$ synthesis.

Such iteration schemes can easily be extended to robust $\mathcal{H}_2$ synthesis, as is now described. Assume that the functions $X(e^{j\omega})$ and $Y(e^{j\omega})$ are as in Lemma 6.7: $X = D^*D$, $Y = E^*E$ with D, E, D^{-1}, E^{-1} in $\mathcal{RH}_\infty$, and satisfying Condition 2. In practice, they could be obtained by fitting frequency points, or using basis functions. It follows that (6.2) can be rewritten as

$$\left\| \begin{bmatrix} D & 0 \\ 0 & I \end{bmatrix} M \begin{bmatrix} D^{-1} & 0 \\ 0 & E^{-1} \end{bmatrix} \right\|_\infty < 1, \tag{6.66}$$

which leads to the following iteration procedure:

- For fixed D, E, reduce the norm in (6.66) by $\mathcal{H}_\infty$ synthesis.
- For a fixed controller, solve the analysis problem for D, E.

As for the standard D-K iteration, each step in this "(D,E)-K" iteration can be shown to improve the robust performance cost, but there is no reason to expect convergence to the global optimum.

The previous iteration was based in $\mathcal{H}_\infty$ synthesis. An alternative is suggested by the discussion in Section 6.7, where at the synthesis step one only includes the D scales with the plant as in (6.6), and employs the design schemes for the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ problem. As remarked before, the techniques in [27] correspond to a slightly different mixed problem, and it is not clear whether they extend to the situation of Proposition 6.11. The approximate method presented in [19] could also be used.

Therefore a number of issues remain open for future research, regarding this second iteration and the comparison between the two alternatives for practical problems.

6.9 Example

We consider the standard single input/output feedback system of Figure 6.3, where the plant $\mathbf{H}$ is subject to multiplicative, unstructured uncertainty $\mathbf{\Delta}$, weighted by $\mathbf{W}$. We wish to analyze the sensitivity of the tracking error e to a white disturbance appearing in v (which could be due, for example, to sensor noise).

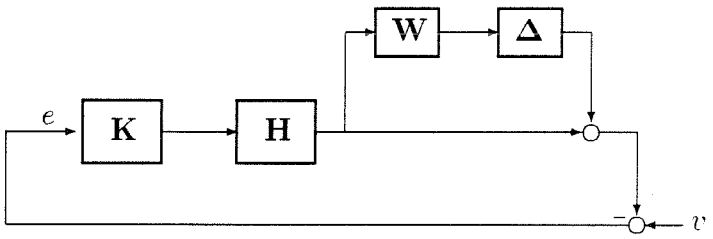


Figure 6.3: Rejection of sensor noise

Assuming $\mathbf{K}$ is nominally stabilizing, we redraw the uncertain system in the standard configuration of Figure 6.1, with

$$\mathbf{M} = \begin{bmatrix} \mathbf{WT} & \mathbf{WT} \\ \mathbf{S} & \mathbf{S} \end{bmatrix}, \tag{6.67}$$

where $S = \frac{1}{1+HK}$ and $T = \frac{HK}{1+HK}$ are the nominal sensitivity and complementary sensitivity functions (see, e.g. [21]).

Suppose now that $\mathbf{\Delta}$ is LTI uncertainty, and we wish to compute the worst-case $\mathcal{H}_2$ norm of the system. For this purpose we use frequency dependent scales $x(e^{j\omega}) > 0$, and at each frequency compute the minimum of $y(e^{j\omega})$, subject to the LMI (6.2), which in this case will have the form

$$\begin{bmatrix} x(|WT|^2 - 1) + |S|^2 & x|WT|^2 + |S|^2 \\ x|WT|^2 + |S|^2 & x|WT|^2 + |S|^2 - y \end{bmatrix} < 0. \tag{6.68}$$

We have omitted the dependence on the frequency variable $e^{j\omega}$. For (6.68) to be feasible, we must have $|WT| < 1$ at each frequency, which is in fact the robust stability test; also, a Schur complement operation reduces (6.68) to

$$x > \frac{|S|^2}{1 - |WT|^2}, \qquad y > \frac{(x|WT|^2 + |S|^2)x}{x(1 - |WT|^2) - |S|^2}. \tag{6.69}$$

Now if y is minimized subject to the constraints (6.69), some algebra leads to the value

$$y_{min} = \left(\frac{|S|}{1 - |WT|}\right)^2. \tag{6.70}$$

Therefore the integral in (6.3) has a minimum value

$$\int_0^{2\pi} y_{min}(e^{j\omega}) \frac{d\omega}{2\pi} = \left\| \frac{S}{1 - |WT|} \right\|_2^2, \tag{6.71}$$

which is an upper bound for the worst-case $\mathcal{H}_2$ norm for LTI uncertainty, and is exact for arbitrarily slowly-varying uncertainty, by Theorems 6.4 and 6.5.

In this particular problem, there is a direct way to corroborate this value and show that indeed it is exact for LTI (non-causal) uncertainty. To see this, write the sensitivity function for the uncertain system,

$$S_\Delta = \frac{1}{1 + HK(1 + W\Delta)} = \frac{S}{1 + WT\Delta}. \tag{6.72}$$

Assuming that $\|WT\|_\infty < 1$, we have

$$S_{max}(e^{j\omega}) = \max_{|\Delta(e^{j\omega})| \leq 1} |S_\Delta(e^{j\omega})| = \frac{|S(e^{j\omega})|}{1 - |WT(e^{j\omega})|}. \tag{6.73}$$

This fact has been typically used (see [21]) to show that the worst case $\mathcal{H}_\infty$ norm of the system is $\left\| \frac{S}{1-|WT|} \right\|_\infty$. By choosing the maximizing $\Delta(e^{j\omega})$ at each frequency (an element of $\mathcal{L}_\infty$) we arrive at the value (6.71) for the worst-case $\mathcal{H}_2$ norm.

Of course, this simple alternative method is not available for general multivariable systems and structured uncertainty; the adequate generalization is given by Condition 2.

Chapter 7

Implicit Uncertain Systems I: Motivation and Definitions

In the predominant viewpoint in systems and control theory, a system is an input-output (I/O) entity, where the external variables are clearly separated in two groups, and a cause-effect relationship is established between them. This approach entails a "signal flow" conception, adequate for systems which are deliberately built to match the I/O philosophy, such as computers and amplifiers. For the standard automatic control configuration where the controller is one of these devices (typically a computer), it is very natural to adopt the input-output point of view, and represent the feedback configuration by a cascade of blocks as in the diagram **G-K** of Figure 2.6. If one considers control theory to be confined to the study of such configurations, then the I/O point of view is completely satisfactory.

Control engineering is, however, a much broader discipline which addresses the general issues of modeling, system identification, system design, simulation, and optimization of complex systems from the point of view of their dynamical properties. In this larger perspective the computer control problem is a small piece which often has secondary technological importance, since very commonly dynamic performance is mainly determined by the design of the "hardware" **G**, and few degrees of freedom are left for a subsequent design of **K**. From this broader point of view the main challenge on the theoretical level is to provide a more unified theory, in which a common language of mathematical and computational machinery is used to perform the previous range of activities.

For these more general problems the input-output point of view often appears artificial, and sometimes inconvenient. As discussed in Section 7.1, when performing modeling of components in a large system it is more natural to specify the model in terms of a set of *equations* between variables which are a priori on an equal footing. The recognition of this fact has led Willems [91] to propose an alternative formulation for system theory where

the central concept is the *behavior*, a set of allowed signal trajectories, and no input-output partition is a priori established between the variables. The corresponding theory of finite dimensional linear systems has been extensively developed ([91], [92]).

This philosophy is even more natural for nonlinear systems, since many nonlinearities (e.g., hysteresis) do not fit the I/O concept, or can only be solved locally into I/O maps. For this reason, much of the theory of absolute stability referred to in Chapter 5 was formulated in terms of relations, rather than I/O maps (see, e.g. [96]).

The same considerations apply, correspondingly, to uncertain systems; if the relationship between the variables is not precisely known, the cause-effect point of view is itself suspect, and it is more natural to think of a relationship between variables. For example, descriptions of uncertainty based on IQCs such as (5.1) are completely symmetrical and do not require an I/O interpretation.

In Section 7.2 we will follow this point of view and introduce uncertain systems in implicit form, which generalize the LFT formulation for uncertainty. We show that in addition to the standard theory, implicit descriptions allow for a complete mapping of general IQCs, in the sense that the signals satisfying the IQC are parametrized as solutions to an uncertain equation. This extension of the LFT framework reduces a large class of robustness analysis problems to special instances of the general question

> $\mathcal{Q}$: *Given a family of equations parametrized by uncertainty, do there exist values of the uncertainty in a given class such that the equations have a solution?*

Stating $\mathcal{Q}$ at this general level has the advantage of highlighting the connection between robustness analysis and system identification. In fact, when the equations come from substitution of experimental data into a model which includes uncertainty (e.g. noise), $\mathcal{Q}$ amounts to the model validation question "is the model consistent with the data?". Also if we consider a model involving parameters, $\mathcal{Q}$ is closely related to the problem of finding parameter values which agree with the data, which is the system identification question.

The connection is, however, more concrete: it will be shown in Section 7.3, that, in particular, the implicit version of the LFT formulation provides a natural setting for posing model validation and system identification problems in the linear case.

Having established the motivation for implicit uncertain systems, in Section 7.4 we will develop a natural framework in which questions such as $\mathcal{Q}$ can be addressed, and reduce the general problem to a canonical case which will be studied in detail in Chapter 8.

7.1 Modeling and Implicit Descriptions

In order for a theoretical framework to be a useful tool for the modeling of complex systems, it must provide a transparent procedure for building large models based on subsystem models. In this way one "zooms in" on each particular component, obtains a mathematical description, and these are subsequently combined. One could think of creating a library of component models, and automating the combination stage by a routine which takes the subsystem models plus a specification of the interconnection, and outputs the full system model.

The only candidate method for this which has been considered so far is the block diagram representation, discussed in Chapter 2. This is in fact the method of choice in some commercial software products for modeling and simulation.

The block diagram representation models an interconnection as a cascade of input-output maps; the restriction to this type of interconnection poses a number of difficulties as a modeling technique. The most basic one is that an input-output representation must be chosen *a priori* at the component modeling stage, and at this level it is not always clear which variables should be regarded as inputs or outputs. For example, the model $v = Ri$ for a resistor can be thought as voltage driven or current driven, and the adequate choice will depend on its location in the full circuit. The need to commit to an I/O partition makes this method unsuitable for modularizing the modeling.

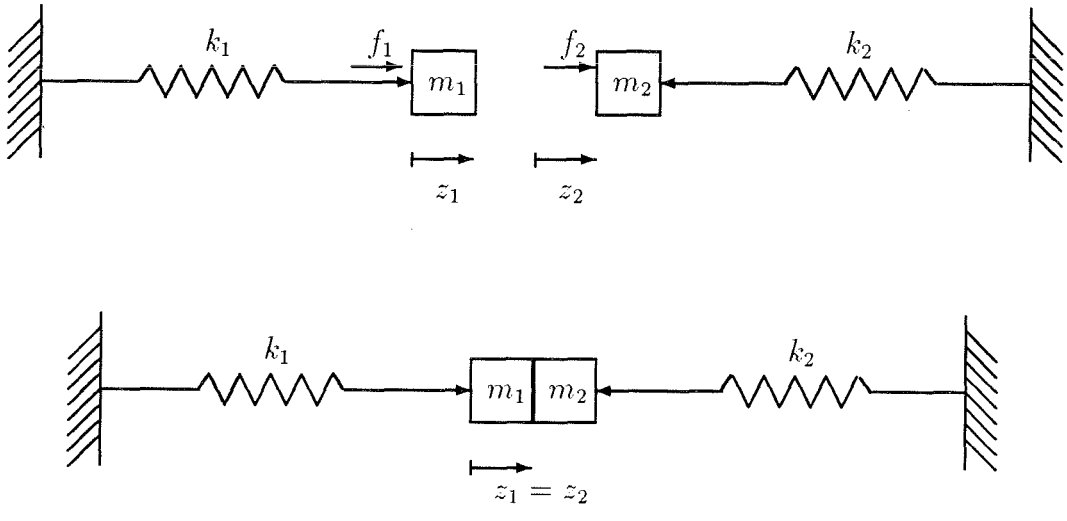


Figure 7.1: Interconnection of two mass-spring systems

As a further example, consider the interconnection of two mass-spring systems depicted

in Figure 7.1. Each corresponds to the differential equation

$$m_i \frac{d^2 z_i}{dt} + k_i z_i = f_i, \qquad i = 1, 2, \tag{7.1}$$

which is typically modeled as an I/O system with input f_i, and output z_i. Now if we wish to model the interconnection of Figure 7.1, where the two masses are bolted together, this would call for identifying the inputs, and the outputs of the two systems, which is not a cascade interconnection. A cascade interconnection could be obtained by changing the signal flow direction in one of the systems, say the first, and considering z_1 as an input and f_1 as an output. We observe, however, that

- Once again this goes against modularization; the component model must be changed before interconnection.
- The "cause-effect" concept which one typically associates with input-output models is lost; there is no justification for breaking the symmetry as to which is the cause, and which the effect, in each system.
- The new model involves differentiation, a feature which is typically considered undesirable in cascade interconnections.

Of course, there is no problem with the given interconnection, and a full system model can easily be obtained by adding to (7.1) the interconnection equations

$$z_1 = z_2, \tag{7.2}$$

$$f_1 + f_2 = 0. \tag{7.3}$$

The difficulties arise only if we insist on building models by cascade interconnection of input-output subsystems. If, as is typical in engineering modeling, one maintains the subsystem models at the level of equations, this difficulty disappears.

Similar considerations apply to the issue of feedback. Although conceptually the signal flow point of view is very appealing to understand feedback (hence its name), it is only strictly appropriate when the feedback interconnection is "buffered", as for example in the case of a computer connected to the system via A/D and D/A converters.

Other feedback interconnections, like those which are commonplace in electronic amplifiers, do not convert transparently to the **G-K** diagram of Figure 2.6. Although the signal flow concept is still used as a heuristic or approximate tool in the designs. for more precise modeling a more convenient approach is to write down and manipulate the corresponding equations.

These issues are even more predominant in the case of nonlinear systems. Often, nonlinear models are derived form physical laws such as mass or energy balances which are expressed as equations between the variables and have global validity. Obtaining an input-output representation involves solving nonlinear equations, which is not a tractable problem and leads to input-output descriptions which are only valid locally.

From this discussion it appears that a theory of systems which attempts to have a general impact in technology beyond the restricted case of the **G-K** configuration, should rely on a more flexible modeling technique than that provided by cascades of input-output systems. It seems natural, in its place, to formulate a theory based on the most commonly used modeling technique, namely representations in terms of implicit equations. A formal definition follows.

Definition 8 *An implicit system* $(\mathbb{W}, \mathbb{E}, \mathbf{R})$ *is defined by two vector spaces, the* **variable** *space* $\mathbb{W}$ *and the* **equation** *space* $\mathbb{E}$*, and an* **equation operator** $\mathbf{R} : \mathbb{W} \to \mathbb{E}$*. The* **behavior** *of the implicit system is the set* $\mathcal{B} = \mathrm{Ker}(\mathbf{R}) = \{w \in \mathbb{W} : \mathbf{R}w = 0\}$*. The system is called linear if* $\mathbf{R}$ *is a linear map.*

The definition above is closely related to the *behavioral* approach to system theory, introduced by Willems [91]. In this type of formulation, all variables in a system are a priori on an equal footing, without a distinction between inputs and outputs. The system laws are constraints in the possible values of these variables, which define a set: the behavior. For further discussion of the features of this modeling paradigm, see [91].

In Definition 8, we give priority to descriptions of the behavior in terms of implicit equations. These will play a central role in the analysis. Interconnections of subsystems in this setting reduce to superimposing equations.

An important special case of Definition 8 is the class of *dynamical* implicit systems, where the sets $\mathbb{W}$ and $\mathbb{E}$ are vector-valued signal spaces. As an example, if $R(\lambda)$ is a polynomial matrix, the differential equations $R(\frac{d}{dt})w = 0$ define an implicit system, where $\mathbf{R}$ is the differential operator $R(\frac{d}{dt})$, and $\mathbb{W}$, $\mathbb{E}$ can be chosen as spaces of smooth functions, or alternatively distribution spaces. The choice of the defining elements $\mathbb{W}$, $\mathbb{E}$, $\mathbf{R}$ depends on the type of analysis to be performed. In this chapter we use linear $\mathbf{R}$, discrete time and:

- Finite dimensional spaces of signals will be used for problems involving data.
- For stability issues, the extended spaces $\mathbb{W} = l_{2e}^q$, $\mathbb{E} = l_{2e}^p$ are used. In this case for the map $\mathbf{R}$ to be well defined we will require it to be causal.
- To formulate quantitative performance specifications, we use the l_2 space $\mathbb{W} = l_2^q$, $\mathbb{E} = l_2^p$, with $\mathbf{R} \in \mathcal{L}(l_2)$.

7.2 Implicit Uncertain Systems

We now incorporate into the implicit paradigm set descriptions of uncertainty in the style of robust control. In the mass-spring example of Figure 7.1, one could for example have the situation where there is uncertainty in the mass, e.g. represented by $m_i = m_{i_0} + \delta_{m_i}$, which would translate directly into uncertainty in the equation operator $\mathbf{R}$ corresponding to (7.1). This motivates the inclusion of uncertainty in implicit descriptions by using an equation map $\mathbf{R}(\mathbf{\Delta})$, where $\mathbf{\Delta}$ is an uncertainty operator.

Based on the experience of standard robust control, it is clear that a very flexible parametrization is obtained by choosing $\mathbf{\Delta}$ to be a structured operator as in (2.24), and the equation map to be a be a Linear Fractional Transformation $\mathbf{\Delta} \star \mathbf{H}$, as depicted in Figure 7.2, where $\mathbf{H} = \left[\begin{smallmatrix} \mathbf{M} & \mathbf{P} \\ \mathbf{N} & \mathbf{Q} \end{smallmatrix}\right]$ is an LTI map. This kind of representation was first considered in [17].

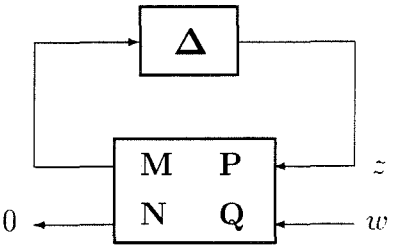


Figure 7.2: Implicit LFT system

A remark regarding Figure 7.2 is that it contains remnants of the "signal-flow" approach, since the perturbation $\mathbf{\Delta}$ is depicted as an input-output map. This is done to highlight the connection with the standard LFT paradigm, but from an implicit point of view the system in Figure 1 is simply characterized by the equations $(\mathbf{\Delta} \star \mathbf{H})w = 0$, or by an "internal" description where the z signals are included in the variable space $\mathbb{W}$, writing

$$\mathcal{R}(\mathbf{\Delta}, \mathbf{H}) \begin{bmatrix} z \\ w \end{bmatrix} = 0, \quad \text{where} \quad \mathcal{R}(\mathbf{\Delta}, \mathbf{H}) := \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}\mathbf{M} & -\mathbf{\Delta}\mathbf{P} \\ \mathbf{N} & \mathbf{Q} \end{bmatrix}. \tag{7.4}$$

The parameterization $\mathcal{R}(\mathbf{\Delta}, \mathbf{H})$ in (7.4) is affine in the parameter $\mathbf{\Delta}$. This simple form allows for the representation of a rich variety of uncertain systems. A standard input-output LFT uncertain system can be easily converted to this implicit form (see (7.8) below).

An implicit LFT representation can also be used to represent general Integral Quadratic Constraints; we recall from Lemma 5.1 that special classes of IQC's were in direct correspondence with non-causal operators in an LFT representation. The following construction shows that in fact Lemma 5.1 covers all cases of scalar valued IQCs, of the general form $\langle z, \Pi z \rangle \geq 0$.

For this purpose, given $\Pi(e^{j\omega}) = \Pi(e^{j\omega})^* \in \mathcal{L}_\infty$ one can obtain a representation

$$\Pi = V^*V - U^*U, \tag{7.5}$$

where $U(e^{j\omega})$, $V(e^{j\omega})$ can be chosen in $\mathcal{H}_\infty$ ($\mathcal{RH}_\infty$ if $\Pi(e^{j\omega})$ is rational). This is done, for example, by choosing $U = kI$ such that $U^*U + \Pi > 0$, and then using a spectral factorization (see [98]) to obtain $U^*U + \Pi = V^*V$.

The factorization (7.5) reduces (5.1) to $\|\mathbf{V}z\|_2^2 \geq \|\mathbf{U}z\|_2^2$. From Lemma 5.1, for each z satisfying (5.1) there exists an operator $\mathbf{\Delta}_C$, $\|\mathbf{\Delta}_C\| \leq 1$ such that $\mathbf{U}z = \mathbf{\Delta}_C\mathbf{V}z$. So the set of $z \in l_2$ satisfying (5.1) can be described as the union over $\mathbf{\Delta}_C, \|\mathbf{\Delta}_C\| \leq 1$ of the behavior of the uncertain implicit system

$$(\mathbf{U} - \mathbf{\Delta}_C\mathbf{V})z = 0. \tag{7.6}$$

Remarks:

- Although $\mathrm{Ker}(U - \mathbf{\Delta}_C V)$ is a linear subspace for each $\mathbf{\Delta}_C$, the union of the parameterized behaviors describes a more complicated set given by (5.1).
- $\{\|\mathbf{\Delta}_C\| \leq 1\}$ includes arbitrary time-varying *non-causal* operators. In this respect, implicit systems obtained from this procedure are a priori considered with $\mathbb{W} = l_2$.
- A finite number of IQCs can be given a representation (7.6), where $\mathbf{\Delta}_C$ is now a structured uncertainty operator.
- Analogously, the set $\bigcup_{\|\boldsymbol{\delta}_C\| \leq 1} \mathrm{Ker}(\mathbf{U} - \boldsymbol{\delta}_C\mathbf{I}\,\mathbf{V})$ for scalar $\boldsymbol{\delta}_C$ can be shown by means of Lemma 5.2 to correspond to the matrix-valued constraint

$$\int_{-\pi}^{\pi} V(e^{j\omega})z(e^{j\omega})z(e^{j\omega})^*V(e^{j\omega})^* - U(e^{j\omega})z(e^{j\omega})z(e^{j\omega})^*U(e^{j\omega})^*\, d\omega \;\geq 0. \tag{7.7}$$

We have provided a complete mapping of IQCs into uncertain equations. As a consequence, an uncertain system and additional constraints on the signals can be represented simultaneously by an implicit LFT system. To illustrate how these implicit descriptions might be used for robust performance analysis, consider the uncertain I/O system of Figure 7.3, where it is known that the disturbance input v satisfies a finite number of restrictions in terms of IQCs as in (5.1). We want to determine whether there exist signals v in the allowed class, and perturbations $\mathbf{\Delta}_u$ such that the system l_2 gain is 1 or larger. This last requirement is captured by the performance IQC $\|y\|^2 - \|v\|^2 \geq 0$.

The implicit equations for the system, the IQCs on v and the performance IQC are captured respectively by

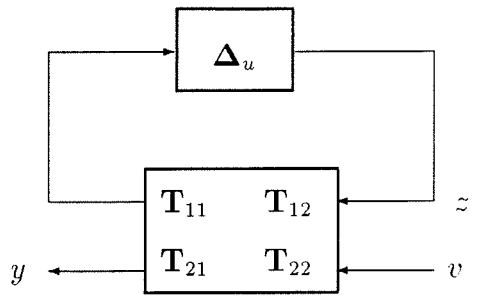


Figure 7.3: Input/Output LFT system

$$\begin{bmatrix} \mathbf{I} - \boldsymbol{\Delta}_u \mathbf{T}_{11} & \mathbf{0} & -\boldsymbol{\Delta}_u \mathbf{T}_{12} \\ \mathbf{T}_{21} & -\mathbf{I} & \mathbf{T}_{22} \end{bmatrix} \begin{bmatrix} z \\ y \\ v \end{bmatrix} = 0, \tag{7.8}$$

$$(\mathbf{U} - \boldsymbol{\Delta}_C \mathbf{V})\, v = 0, \tag{7.9}$$

$$\begin{bmatrix} \boldsymbol{\Delta}_P & -\mathbf{I} \end{bmatrix} \begin{bmatrix} y \\ v \end{bmatrix} = 0, \tag{7.10}$$

where $\boldsymbol{\Delta}_C$ and $\boldsymbol{\Delta}_P$ are norm bounded LTV operators, and $\boldsymbol{\Delta}_C$ is in general structured. The superposition of (7.8), (7.9) and (7.10) gives an implicit description for the robust performance analysis problem, which reduces to the question:

> $\mathcal{Q}$: "Does there exist a perturbation $(\boldsymbol{\Delta}_u, \boldsymbol{\Delta}_C, \boldsymbol{\Delta}_P)$ such that (7.8-7.10) have non-trivial l_2 solutions?"

In $\mathcal{Q}$ the input/output partition has been eliminated from the problem, and the analysis is posed in terms of equations and solutions, rather than maps and gains; this allows for a natural incorporation of the constraints (7.9). Questions of this type will be analyzed in the sequel.

Remark: The transformation from the constrained I/O problem to the implicit version is mainly a notational change, and any result which is obtained from this new version can also be obtained with the original objects. The advantage of this new form is that it provides a canonical representation for these over-constrained problems, for which results can be derived and then applied to a variety of special cases. Also , this generality is helpful in establishing connections with other problems, as is done in the following section.

7.3 Model Validation and System Identification

As discussed in the Section 7.1, the implicit formulation is a natural setting for modeling systems from first principles, since physical laws are typically specified as equations between the system variables. The other main method of obtaining models is "black-box" system identification from experimental data. We will argue in this section that system identification is conceptually a similar problem to robustness analysis, and that implicit LFT representations provide a way to formulate both of them in a consistent language.

The connection comes from the question of model validation, which inquires if a model is consistent with data. Take for example the model structure (see [45])

$$y = G(\boldsymbol{\lambda}, \theta)u + H(\boldsymbol{\lambda}, \theta)d, \tag{7.11}$$

where $\boldsymbol{\lambda}$ is the shift operator, θ is a vector of parameters, G and H are discrete time systems, and d is a disturbance. Assume the parameter θ has been chosen, and we are given data u and y. The model validation question is to check whether the model can account for the data with a plausible (e.g. small enough) instance of d. This question clearly fits into the general form $\mathcal{Q}$ stated in the introduction: given the equations, the parameters and the data, we must check whether there exist values of the uncertainty (in this case d) in a given set (e.g. $\|d\| \leq \gamma$) which verify the equations.

If the parameters are also taken to be uncertain, then $\mathcal{Q}$ amounts to looking for values of the parameters and noise which satisfy the equations, which is the type of question asked in system identification. Also, one can imagine a generalized model validation or identification setup where in addition to noise and parameters, there is some representation for unmodeled dynamics in the style of Chapter 2. This is especially relevant to establish closer connections between identification and robust control, which remains as a fundamental challenge. Once again, the basic problem can be cast in the form $\mathcal{Q}$, where now the uncertainty includes noise, unmodeled dynamics, and possibly parameters.

Making the description of $\mathcal{Q}$ more concrete, we will now show that in particular implicit LFT systems as those considered in Section 7.2 provide an attractive setup for a general class of model validation/system identification-type questions (henceforth denoted MV/ID). The following development is closely related to some recent work in [78, 53] where model validation problems where related to extensions of the structured singular value μ. In addition to providing a conceptual framework in which to relate robustness analysis and identification, this formulation suggests that computational tools developed to deal with the implicit robustness analysis formulation could be applied to system identification.

7.3.1 Model Validation as Implicit Analysis

Figure 7.4 shows a generic input-output MV/ID structure, where all the elements in the diagram are constant vectors and matrices. This setup is directly applicable to static problems, where there is no explicit time variable. If, however, we are dealing with a dynamic model over a finite time horizon, a static representation can be derived by stringing out the time horizon, and representing signals by vectors and linear systems by matrices (see [83]); also, this static approach can be used for model validation in the frequency domain, where the data comes as the response of the system at a number of frequency points (see [53]). From now on we only refer to the static question.

In Figure 7.4, d is a vector of unknown inputs (disturbances), constrained by $\|d\| \leq 1$; Δ is a structured matrix in a certain normalized class B_Δ; u and y are the measured inputs and outputs. The MV/ID problem is to find values of Δ and d consistent with Figure 7.4.

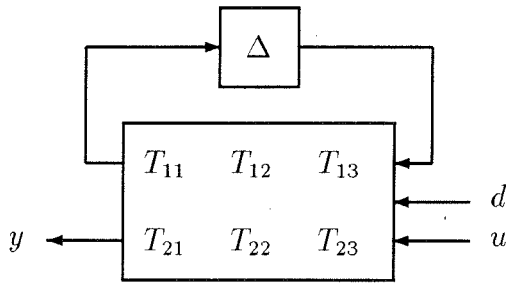


Figure 7.4: A standard input-output MV/ID setup.

This LFT structure captures a rich variety of linear identification problems. As a simple example, consider the standard linear regression

$$y = \mathcal{U}\theta + d, \tag{7.12}$$

where $\mathcal{U}$ and y are known, θ is a vector of unknown parameters and d is a vector of unknown errors. These equations are of the form of Figure 7.4, with

$$\Delta = \theta, \quad u = 1, \quad T = \begin{bmatrix} 0 & 0 & 1 \\ \mathcal{U} & I & 0 \end{bmatrix}. \tag{7.13}$$

In Figure 7.5, the equations of Figure 7.4 are represented in implicit form, with u and y combined into the vector v which includes all the known data:

$$v = \begin{bmatrix} u \\ y \end{bmatrix}, \quad \tilde{T}_{13} = \begin{bmatrix} T_{13} & 0 \end{bmatrix}, \quad \tilde{T}_{23} = \begin{bmatrix} T_{23} & -I \end{bmatrix}. \tag{7.14}$$

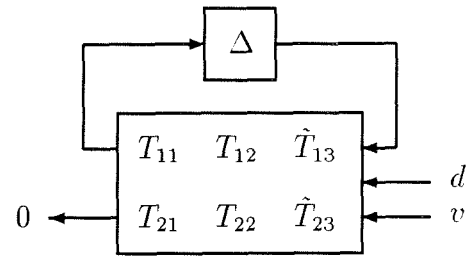


Figure 7.5: A standard MV/ID setup in implicit form.

Note that the input-output partition has been eliminated from the model. In fact, it could well be that we wish to validate some model based on observations of a system where this distinction is not available. Then we would arrive directly at Figure 7.5. For example, in the linear regression above, the "input" in (7.13) is an artifice of the representation.

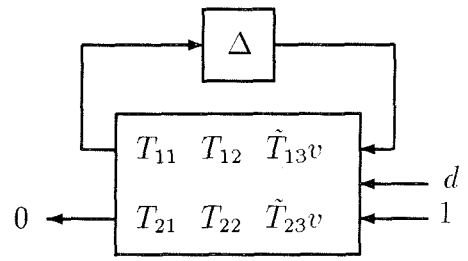


Figure 7.6: The MV/ID setup with data inside the matrix.

We can now incorporate the data v into the matrix by considering a fictitious scalar "input" of value 1. This results in Figure 7.6.

The representation has up to now two different sources of uncertainty: Δ and d. This distinction disappears and the constraint $\|d\| \leq 1$ is included in the problem by introducing the uncertainty block $\Delta_d = d$, $\|\Delta_d\| \leq 1$, and writing $d = \Delta_d 1$. This is shown in Figure 7.7; Δ_d is just a new name for d which reflects its location in the diagram. The representation of Figure 7.7 corresponds to a constant matrix version of the implicit LFT system (7.4).

The MV/ID question is now of the form : "Does there exist a nontrivial signal satisfying the implicit equations of Figure 7.7 for $\Delta \in B_\Delta$?" The fact that one signal is constrained to be 1 is irrelevant since everything can be normalized by linearity.

This question corresponds exactly to the robustness analysis question $\mathcal{Q}$ posed in Section 7.2, so the same theoretical and computational methods can be applied to both situations.

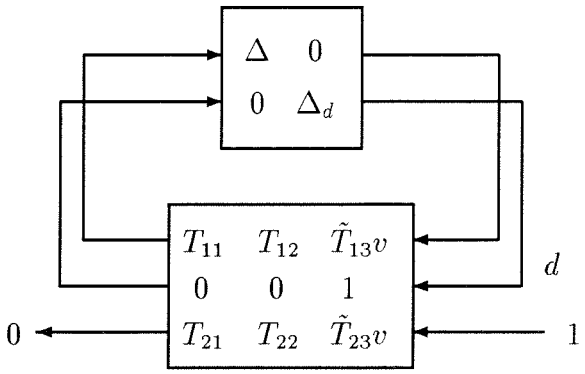


Figure 7.7: MV/ID as an implicit analysis problem.

7.4 Stability in Implicit Systems

As in the case of I/O systems, many important issues in the analysis of an implicit system hinge around the question of stability. In this section we introduce this concept and develop the basic framework which will be further analyzed in Chapter 8.

7.4.1 Stability and l_2-Stability

We first introduce the stability notion for an implicit system with no uncertainty

$$\mathbf{R}w = 0. \tag{7.15}$$

Informally, the system (7.15) is stable if the equations disallow any unbounded behaviors w (e.g., in $l_{2e} \backslash l_2$). This requires, in particular, that there are no degrees of freedom left in w, so there must be at least as many equations as variables; in the language of [91], the system is autonomous. We wish, however, to allow the possibility of over-constrained problems, with more equations than unknowns, such as the example considered in (7.8-7.10). These requirements are met by the following definition.

Definition 9 *Let* $(\mathbb{W}, \mathbb{E}, \mathbf{R})$ *be an implicit linear system over* l_{2e}*:* $\mathbb{W} = l_{2e}^q$, $\mathbb{E} = l_{2e}^p$, $\mathbf{R} : \mathbb{W} \to \mathbb{E}$ *causal map . The system is said to be well posed if* $\mathbf{R}$ *has a causal left inverse* $\mathbf{L} : \mathbb{E} \to \mathbb{W}$*. The system is stable if it is well posed and in addition,* L *is bounded on the range of* $\mathbf{R}$*, i.e.* $\exists \gamma > 0$ *such that for every* $e = \mathbf{R}w$*, and every* T*,* $\|\mathbf{P}^{\mathbf{T}} w\| \leq \gamma \|\mathbf{P}^{\mathbf{T}} e\|$.

Remarks

- The restriction to causal maps is included so that objects are well defined over the extended spaces.

- We use the *equation error* e as a way of testing the sensitivity of the equations, in analogous way as injected disturbances were used for the standard I/O configuration of Figure 2.8.

- In reference to Figure 2.8, leaving out the injected disturbances an implicit representation is given by

$$\begin{bmatrix} \mathbf{I} & -\mathbf{H} \\ -\mathbf{K} & \mathbf{I} \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} = 0, \tag{7.16}$$

  which is of the general form (7.15). Applying Definition 9, stability implies that $\begin{bmatrix} \mathbf{I} & -\mathbf{H} \\ -\mathbf{K} & \mathbf{I} \end{bmatrix}$ has a causal, bounded left inverse on $\mathcal{L}_c(l_{2e})$. This is slightly weaker than Definition 1, where full invertibility was required; in Definition 9, the operator $\mathbf{R}$ is not required to be *onto*, in other words the equation errors need not be free to span the whole of l_{2e}.

  The reason for this weaker definition in the case of implicit systems is to be able to extend the notion to systems which are over-constrained, with more equations than variables, such as the example considered in (7.8-7.10). In these cases, the operator will not be onto in general, and this should not be required. The equation errors are not physical noises, which should be included in the w variables; they just provide a means of testing sensitivity of the equations. For the case of $\mathbf{H}$, $\mathbf{K}$ LTI in (7.16), the definition is equivalent to the standard one, as will follow from Proposition 7.3.

- As an immediate consequence of the definition, if $e \in l_2$ and $e = \mathbf{R}w$ for a stable system $\mathbf{R}$, then w must be in l_2.

Definition 9 can also be interpreted in terms of the infinite matrix representation (2.12) of the equation operator $\mathbf{R}$, and its truncations R_T as in (2.14).

Proposition 7.1 *The implicit system* $(\mathbb{W}, \mathbb{E}, \mathbf{R})$ *over* l_{2e} *is stable if and only if*

$$\inf_{T \geq 0} \underline{\sigma}(R_T) > 0, \tag{7.17}$$

where $\underline{\sigma}(\cdot)$ *denotes minimum singular value of a matrix,* $\underline{\sigma}(A) = \min_{|\xi|=1} |A\xi|$.

Proof: From Definition 9, stability implies $|R_T\xi| \geq \frac{1}{\gamma}|\xi|$ from which (7.17) follows.

Conversely, let $\epsilon = \inf_{T\geq 0} \underline{\sigma}(R_T) > 0$. Then a causal left inverse $\mathbf{L}$ can be constructed by induction over T: given a lower triangular left inverse L_T of R_T, consider the matrix $R_{T+1} = \begin{bmatrix} R_T & 0 \\ \bar{r} & r_{T+1,T+1} \end{bmatrix}$. Since $\underline{\sigma}(R_{T+1}) > 0$, $r_{T+1,T+1}$ has left inverse $l_{T+1,T+1}$, then

$$L_{T+1} = \begin{bmatrix} L_T & 0 \\ -l_{T+1,T+1}\bar{r}L_T & l_{T+1,T+1} \end{bmatrix}$$

is a left inverse of R_{T+1}. This procedure produces a causal operator $\mathbf{L} : \mathbb{E} \rightarrow \mathbb{W}$. If $e = \mathbf{R}w$, then $\mathbf{P}^T e = R_T \mathbf{P}^T w$ therefore $\|\mathbf{P}^T e\| \geq \epsilon \|\mathbf{P}^T \mathbf{L} e\|$ so $\mathbf{L}$ is bounded on the range of $\mathbf{R}$.

■

In Definition 9 we restricted the attention to causal equation maps $\mathbf{R}$, which is the natural setting to view these equations over the extended spaces l_{2e}. The representation of IQCs in terms of uncertain equations discussed in Section 7.2 calls, however, for introducing non-causal maps. For this purpose one must restrict the analysis *a priori* to the space l_2 where these maps are well defined. In this setting, we wish to provide a definition which captures the analysis question $\mathcal{Q}$ discussed in Section 7.2. It turns out that such definition is closely related to Definition 9, with the modification that the causality restriction is removed, and the maps are restricted to l_2. Due to this analogy we will name this notion "l_2-stability", although the terminology is not necessarily very descriptive.

Definition 10 *Let $(\mathbb{W}, \mathbb{E}, \mathbf{R})$ be an implicit linear system over l_2. The system is l_2-stable if $\mathbf{R}$ is left invertible in $\mathcal{L}(l_2)$, i.e. $\exists \mathbf{L} \in \mathcal{L}(l_2)$ such that $\mathbf{LR} = \mathbf{I}$.*

An equivalent characterization is

Proposition 7.2 *$\mathbf{R}w = 0$ is l_2-stable if and only if*

$$\inf\{\|\mathbf{R}w\|_2 : w \in l_2, \|w\|_2 = 1\} > 0. \tag{7.18}$$

Proof: For the necessity, note that since $\mathbf{LR} = \mathbf{I}$, $\|\mathbf{L}\|\|\mathbf{R}w\| \geq \|w\|$. For the sufficiency, condition (7.18) implies that $\mathbf{R}$ is injective, and its range is closed. An application of the open mapping theorem (see [71]) implies that $\mathbf{R}$ has a bounded inverse on its range, which can be extended to a bounded operator on $\mathbb{E}$, resulting in a left inverse for $\mathbf{R}$.

■

Interpreting the definition, l_2-stability implies that the l_2 behavior of the system is the trivial space, and that this property is not "sensitive" to equation error: an arbitrary small equation error e does not allow solutions w, $\|w\| = 1$, to the equation $\mathbf{R}w = e$.

To illustrate these definitions, we will consider the case of an "autoregressive" system (see [91]) defined by linear time-invariant equations of the form $\mathbf{R} = R(\boldsymbol{\lambda})$, where $R(\lambda)$ is a $p \times q$ polynomial matrix.

Proposition 7.3 *Let $R(\lambda)$ be a $p \times q$ polynomial matrix. Then*

(i) The system $R(\boldsymbol{\lambda})w = 0$ is stable if and only if $R(\lambda)$ has full column rank for all λ in the closed unit disk $\overline{\mathbb{D}}$.

(ii) The system $R(\boldsymbol{\lambda})w = 0$ is l_2-stable if and only if $R(\lambda)$ has full column rank for all λ in the unit circle $\mathbb{T}$.

Proof:

(i) If $R(0)$ has a kernel, then $R(\boldsymbol{\lambda})$ cannot have a causal left inverse, so well posedness fails. If $R(\lambda_0)\bar{w} = 0$, $0 \neq |\lambda_0| \leq 1$ then the signal $w(t) = (\frac{1}{\lambda_0})^t \bar{w}$, $t > 0$ is in $l_{2e} \backslash l_2$, but $R(\boldsymbol{\lambda})w \in l_2$, violating stability.

If $R(\lambda)$ has full column rank over the closed unit disk, then the theory of coprime factorizations over the stable ring (see e.g. [87]) implies that $R(\lambda)$ has a left inverse in $\mathcal{RH}_\infty$. This gives a causal left inverse which is bounded (over all l_{2e}) so stability is satisfied.

(ii) If $R(e^{j\omega_0})\bar{w} = 0$, we can construct l_2 signals $w^{(k)}$ with norm 1 and $\|R(\boldsymbol{\lambda})w^{(k)}\| \stackrel{k\to\infty}{\longrightarrow} 0$, violating l_2-stability. Conversely if $R(e^{j\omega})$ has full column rank, $\min_{e^{j\omega}\in\mathbb{T}} \underline{\sigma}(R(e^{j\omega})) > 0$, so there is a left inverse $L(e^{j\omega}) \in \mathcal{L}_\infty$ for $R(e^{j\omega})$, i.e. a bounded left inverse $\mathbf{L}$ to $R(\boldsymbol{\lambda})$.

■

Remark: The same argument carries through if $R(\lambda)$ is a rational, rather than polynomial matrix, with no poles on the unit disk (respectively the unit circle).

Example: Consider $R(\lambda) = 1 - 2\lambda$. Setting $\mathbb{W} = \mathbb{E} = l_{2e}$, the implicit system is not stable, since the signal $w(t) = 2^t, t \geq 0$, which is in $l_{2e} \backslash l_2$ gives $e = \mathbf{R}w \in l_2$.
The system with $\mathbb{W} = \mathbb{E} = l_2$ is l_2-stable, however, since $\inf_{\|w\|=1} \|\mathbf{R}w\| = 1$.

7.4.2 Robust Stability and l_2-Stability

In this section we refer to the implicit uncertain system of (7.4). When the notion of stability over l_{2e} is being considered, $\mathbf{H}$ is assumed to be a causal, LTI operator in $\mathcal{L}_c(l_{2e})$, and the uncertainty $\boldsymbol{\Delta}$ is in the structured unit ball $\boldsymbol{B_\Delta}$ as in Section 2.2.3. The implicit system (7.4) is said to be robustly stable if it is stable for every $\boldsymbol{\Delta} \in \boldsymbol{B_\Delta}$.

For the notion of l_2-stability we will allow non-causal LTI maps $\mathbf{H}$, and perturbations over the class $\boldsymbol{B_{\Delta_{NC}}}$, without the restriction on causality. The implicit system (7.4) is robustly l_2-stable if it is l_2-stable for every $\boldsymbol{\Delta} \in \boldsymbol{B_{\Delta_{NC}}}$.

The constrained I/O robust performance problem posed in Section 7.2 converts to robust l_2-stability of the corresponding implicit system: a negative answer to $\mathcal{Q}$ implies a trivial l_2-kernel for every $\mathbf{\Delta}$. Left invertibility is technically slightly stronger, but conceptually the same type of condition.

As an initial step in the analysis of robust stability (or l_2-stability) of implicit systems, the configuration of (7.4) will be further reduced to a canonical case.

Proposition 7.4 *The implicit system* (7.4) *is robustly* (l_2) *stable if and only if*
(i) $\mathbf{Q}$ *is* (l_2) *stable, with bounded left inverse* $\mathbf{L}$.
(ii) The implicit system $\begin{bmatrix} \mathbf{I}-\mathbf{\Delta}\hat{\mathbf{M}} \\ \hat{\mathbf{N}} \end{bmatrix} z = 0$, *is robustly* ($l_2$) *stable, where*

$$\hat{\mathbf{M}} = \mathbf{M} - \mathbf{PLN}, \quad \hat{\mathbf{N}} = \mathbf{N} - \mathbf{QLN}. \tag{7.19}$$

Proof: Consider first the case of l_2-stability.

[Necessity]

If (7.4) is robust l_2-stable, then in particular $\mathcal{R}(\mathbf{0},\mathbf{H}) = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{N} & \mathbf{Q} \end{bmatrix}$ has a bounded left inverse, therefore $\mathbf{Q}$ has bounded left inverse $\mathbf{L}$. We now write the identity

$$\mathcal{R}(\mathbf{\Delta},\mathbf{H})\,\mathbf{S} = \mathbf{T}\,\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}}), \tag{7.20}$$

which holds for any $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}$, with the invertible operators $\mathbf{S}$ and $\mathbf{T}$ defined as

$$\mathbf{S} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{LN} & \mathbf{I} \end{bmatrix}, \ \mathbf{T} = \begin{bmatrix} \mathbf{I} & -\mathbf{\Delta PL} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \ \mathbf{S}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{LN} & \mathbf{I} \end{bmatrix}, \ \mathbf{T}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{\Delta PL} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{7.21}$$

and

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{M}} & \mathbf{0} \\ \hat{\mathbf{N}} & \mathbf{Q} \end{bmatrix}, \ \mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}}) = \begin{bmatrix} \mathbf{I}-\mathbf{\Delta}\hat{\mathbf{M}} & \mathbf{0} \\ \hat{\mathbf{N}} & \mathbf{Q} \end{bmatrix}. \tag{7.22}$$

By hypothesis, $\mathcal{R}(\mathbf{\Delta},\mathbf{H})$ is left invertible, hence so is $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$ from (7.20), which implies the left invertibility of $\mathcal{R}_1(\mathbf{\Delta},\hat{\mathbf{H}}) := \begin{bmatrix} \mathbf{I}-\mathbf{\Delta}\hat{\mathbf{M}} \\ \hat{\mathbf{N}} \end{bmatrix}$.

[Sufficiency]

If $\mathbf{\Upsilon}(\mathbf{\Delta}) = [\,\mathbf{\Upsilon}_1 \ \mathbf{\Upsilon}_2\,]$ is a bounded left inverse for $\mathcal{R}_1(\mathbf{\Delta},\hat{\mathbf{H}})$ then (note that $\mathbf{L}\hat{\mathbf{N}} = \mathbf{0}$ by definition of $\hat{\mathbf{N}}$)

$$\begin{bmatrix} \mathbf{\Upsilon}_1 & \mathbf{\Upsilon}_2(\mathbf{I}-\mathbf{QL}) \\ \mathbf{0} & \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{I}-\mathbf{\Delta}\hat{\mathbf{M}} & \mathbf{0} \\ \hat{\mathbf{N}} & \mathbf{Q} \end{bmatrix} = \mathbf{I}, \tag{7.23}$$

therefore $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$ is left invertible, and so is $\mathcal{R}(\mathbf{\Delta},\mathbf{H})$ by (7.20).

Now turning to the notion of stability over l_{2e} :

[Necessity]:

By hypothesis $\mathcal{R}(\mathbf{0},\mathbf{H}) = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{N} & \mathbf{Q} \end{bmatrix}$ has a causal left inverse, which is bounded over its range. Since it is an LTI map in $\mathcal{L}_c(l_{2e})$, from Proposition 7.3 we know that the left inverse can be chosen in $\mathcal{H}_\infty$ and therefore bounded over the whole of l_{2e}, not just the range. Let $\mathbf{L}$ be the corresponding left inverse for $\mathbf{Q}$. Then $\mathbf{S}$, $\mathbf{T}\,\mathbf{S}^{-1}$, $\mathbf{T}^{-1}$, defined as in (7.21) are maps in $\mathcal{L}_c(l_{2e})$. From here on the argument follows analogously to the l_2 case.

[Sufficiency]:

Fix $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}$, let $[\,\mathbf{\Upsilon}_1 \;\; \mathbf{\Upsilon}_2\,]$ be a causal left inverse of $\mathcal{R}_1(\mathbf{\Delta},\hat{\mathbf{H}})$, bounded on its range, and construct a causal left inverse for $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$ as in (7.23). The lower portion $[\,\mathbf{0} \;\; \mathbf{L}\,]$ is bounded on l_{2e} as discussed above. The top portion is

$$\begin{bmatrix} \mathbf{\Upsilon}_1 & \mathbf{\Upsilon}_2(\mathbf{I}-\mathbf{QL}) \end{bmatrix} = \begin{bmatrix} \mathbf{\Upsilon}_1 & \mathbf{\Upsilon}_2 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}-\mathbf{QL}) \end{bmatrix}, \tag{7.24}$$

where $\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}-\mathbf{QL}) \end{bmatrix}$ is bounded and maps a vector in the range of $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$, to the range of $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$ where $[\,\mathbf{\Upsilon}_1 \;\; \mathbf{\Upsilon}_2\,]$ is bounded, therefore the top portion is bounded on the range of $\mathcal{R}(\mathbf{\Delta},\hat{\mathbf{H}})$.

■

The previous result shows that for robustness analysis, it suffices to consider the canonical implicit system

$$\begin{bmatrix} \mathbf{I}-\mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix} z = 0, \tag{7.25}$$

where $\hat{\mathbf{M}},\hat{\mathbf{N}}$ are renamed as $\mathbf{M},\mathbf{N}$ for simplicity. This case will be analyzed in Chapter 8.

It is useful to compare this setup with the question of robust stability in standard robust control, which specifies the invertibility of $(\mathbf{I}-\mathbf{\Delta M})$. The main difference is that (7.25) allows for additional constraints defined by the $\mathbf{N}$ equations. A problem with more equations than variables such as the one considered in (7.8-7.10) will result in the presence of these additional equations. A problem where the $\mathbf{N}$ equations do not appear will be termed the "unconstrained" case. For example, if the IQCs (7.9) are eliminated from (7.8-7.10), the problem can be reduced to the equations

$$\left(I - \begin{bmatrix} \mathbf{\Delta}_u & \mathbf{0} \\ \mathbf{0} & \mathbf{\Delta}_P \end{bmatrix} \mathbf{T} \right) \begin{bmatrix} z \\ v \end{bmatrix} = 0 \tag{7.26}$$

which are in the standard form $(\mathbf{I}-\mathbf{\Delta M})z = 0$.

In an unconstrained problem, our definition of robust stability imposes left invertibility of $(\mathbf{I} - \mathbf{\Delta M})$ over $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}}$. It can be shown (see [60]) that in this case full invertibility holds, so there is no difference with the standard definition.

Remark:

The notions of stability and l_2-stability and this canonical representation have been developed in the context of dynamical systems. The results in Section 7.3 have also motivated the appearance of representations such as (7.4), where the defining elements are constant matrices. The analysis question $\mathcal{Q}$ for these constant matrix representations is to test for the existence of nontrivial solutions, which corresponds exactly to left invertibility. Correspondingly, the development of the canonical representation (7.25) also applies to these constant matrix analysis problems, which will be further discussed in Chapter 8.

Chapter 8

Implicit Uncertain Systems II: Analysis

In Chapter 7 we introduced a framework for analysis of systems in implicit form. This approach was motivated from first principles modeling where implicit equations are the natural object, but was shown to have advantages for a more general class of problems. In particular, it was shown that an implicit version of the LFT formulation provides a standard setup in which to cast a variety of analysis questions involving constraints, including robustness analysis under IQCs and a model validation problem.

It was further shown that the analysis of these implicit uncertain systems could be reduced to studying the canonical case

$$\begin{bmatrix} \mathbf{I} - \mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix} z = 0, \tag{8.1}$$

where $\mathbf{M}$ and $\mathbf{N}$ are LTI maps, and $\mathbf{\Delta}$ a structured operator. The robustness analysis question is to test for the existence of solutions to (8.1), or, more technically, to test the left invertibility of the operator $\begin{bmatrix} \mathbf{I} - \mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix}$. This question is a generalization of the standard "unconstrained" case where the $\mathbf{N}$ equations do not appear.

In this chapter we consider this analysis problem in a number of cases. We begin in Section 8.1 by studying constant matrix problems; these lead to a generalization of the structured singular value μ; the upper bound theory is developed in detail. In Section 8.2 we analyze LTV perturbation structures, and prove convex conditions extending those in the standard theory. Section 8.3 discusses the case where $\mathbf{M}$, $\mathbf{N}$ are described in state-space. Section 8.4 contains examples to demonstrate the application of this theory in two directions: robust performance analysis under constrained disturbances, and a review of the least squares identification problem from this perspective.

8.1 A Structured Singular Value for Implicit Systems

In many important cases, robustness analysis can be conducted in a constant matrix representation. In the implicit framework, these have the form

$$\begin{bmatrix} I - \Delta M \\ N \end{bmatrix} z = 0, \quad \Delta \in \mathit{\Delta} \subset \mathbb{C}^{n \times n}, \tag{8.2}$$

where $M \in \mathbb{C}^{n \times n}$, $N \in \mathbb{C}^{p \times n}$, and the structure $\mathit{\Delta}$ is still of the form (2.24), but with blocks which are constant, complex matrices rather than dynamical operators. $\mathit{\Delta}$ could also have blocks restricted to be real. By analogy with the dynamic case, we will say that the implicit system (8.2) is stable if

$$\text{Ker} \begin{bmatrix} I - \Delta M \\ N \end{bmatrix} = 0 \quad \forall \Delta \in B_{\mathit{\Delta}}. \tag{8.3}$$

Equivalently, stability implies left invertibility of $\begin{bmatrix} I - \Delta M \\ N \end{bmatrix}$ over $B_{\mathit{\Delta}}$. Condition (8.3) strongly resembles the PBH test in standard system theory. In fact, for the special case $\Delta = \delta I$, stability is equivalent to detectability of the pair (N, M).

Standard robust stability analysis for constant matrices is provided by the structured singular value μ [28, 55], as reviewed in Section 2.3.2. In the constrained case of (8.2), the natural extension is given by the following:

Definition 11 *The structured singular value $\mu_{\mathit{\Delta}}(N, M)$ of the implicit system* (8.2) *is defined as follows:*

If $\text{Ker} \begin{bmatrix} I - \Delta M \\ N \end{bmatrix} = 0 \; \forall \Delta \in \mathit{\Delta}$, *define* $\mu_{\mathit{\Delta}}(N, M) := 0$, *otherwise*

$$\mu_{\mathit{\Delta}}(N, M) := \frac{1}{\min \left\{ \bar{\sigma}(\Delta) : \Delta \in \mathit{\Delta}, \text{Ker} \begin{bmatrix} I - \Delta M \\ N \end{bmatrix} \neq 0 \right\}}. \tag{8.4}$$

A restatement of this definition is to say that (8.2) is stable if and only if $\mu_{\mathit{\Delta}}(N, M) < 1$.

Before considering the corresponding theory for "implicit μ" we remark on which problems will reduce to this representation.

- A first simple example, analogous to the standard theory, is that if the uncertainty in (8.1) is LTI, then the robust stability tests reduce to a constant matrix analysis across frequency. In particular, a direct application of Proposition 7.3 implies that the condition $\mu_{\mathit{\Delta}}(N(\lambda), M(\lambda)) < 1$ over the unit circle $|\lambda| = 1$ characterizes robust l_2 stability,

and the same condition over the unit disk $|\lambda| \leq 1$ characterizes robust stability. Note that in the general case (dynamic $\mathbf{N}$), the "maximum-modulus" property no longer holds; the maximum of $\mu_{\Delta}(N(\lambda), M(\lambda))$ need not occur on the circle.

- The previous case has not been very well motivated, since we argued for the consideration of over-constrained problems based on the representation of IQCs, which yield LTV blocks in $\mathbf{\Delta}$. If, in the situation of Figure 7.3, the uncertainty $\mathbf{\Delta}_u$ is LTI, then the implicit description (7.8-7.10) will have a *combined* LTI/LTV uncertainty structure, exactly as those considered in Chapter 3. Most of the theory of Chapter 3 can, however, be extended to the implicit case, reducing these combined problems to implicit μ analysis problems in augmented uncertainty structures, thus providing an application of the previous definition.

- The model validation question considered in Section 7.3 reduces directly to a constant matrix problem. The corresponding test $\mu_{\Delta}(N, M) < 1$ is equivalent to model invalidation.

8.1.1 Bounds for Implicit μ

Definition 11 translates the analysis problem to the computation of $\mu_{\Delta}(N, M)$; as in the standard case, exact computation is in general hard and one must rely on upper and lower bounds. We will only comment briefly here on the lower bound problem, and develop in detail the upper bound theory.

The lower bounds algorithms for the standard unconstrained case (no N equations) are based on the fact that $\mu_{\Delta}(M) = \max_{\Delta \in B_{\Delta}}(\rho(\Delta M))$, where $\rho(\cdot)$ denotes spectral radius. Algorithms which resemble the power iteration for spectral radius have been developed [55, 95], with good performance on typical problems.

For the constrained case, only eigenvalues with eigenvectors in the kernel of N are relevant. In the following, it will be convenient to parameterize this kernel by a matrix $N_{\perp}^{*}$, whose columns form a basis for the kernel of N. This leads to

$$\mathrm{Ker} \begin{bmatrix} I - \Delta M \\ N \end{bmatrix} = 0 \iff \mathrm{Ker}[N_{\perp}^{*} - \Delta M N_{\perp}^{*}] = 0. \tag{8.5}$$

Denoting $\rho^{g}(A, B) = \max\{|\beta| : \mathrm{Ker}(\beta A - B) \neq 0\}$ (maximum modulus of a generalized eigenvalue of A, B, the notation is not standard), we have

$$\mu_{\Delta}(N, M) = \max_{\Delta \in B_{\Delta}} \rho^{g}(N_{\perp}^{*}, \Delta M N_{\perp}^{*}). \tag{8.6}$$

These observations will presumably lead to an extension of the standard μ lower bound algorithms, although work in this problem is still premature, and it is not obvious that these algorithms would inherit the typical performance of the standard μ lower bound. Some initial work on a special case of this extension is documented in [53].

We will now consider the upper bounds for this version of the structured singular value. The following theory strongly parallels that of [55] for the standard case. $\mathbb{X}$ denotes the set of scaling matrices defined in (2.35), $\mathbb{X}^H$, $\mathbb{X}^P$ the subsets of hermitian and positive matrices, respectively. $\mathbb{X}^H$ is in fact a real vector space, and

$$\langle X, \hat{X} \rangle = \sum_{i=1}^{L} \mathrm{trace}(X_i \hat{X}_i) + \sum_{j=1}^{F} x_{L+j} \hat{x}_{L+j} \tag{8.7}$$

defines an inner product in $\mathbb{X}^H$. $\mathbb{X}^P$ and its closure $\overline{\mathbb{X}^P}$ are convex in $\mathbb{X}^H$.

Lemma 8.1 *For fixed $\beta > 0$, the following are equivalent:*

$$(i) \quad \exists X \in \mathbb{X}^P : \quad M^*XM - \beta^2 X - N^*N \;<\; 0; \tag{8.8}$$

$$(ii) \quad \exists X \in \mathbb{X}^P : \quad N_\perp (M^*XM - \beta^2 X) N_\perp^* \;<\; 0. \tag{8.9}$$

Proof: Since $NN_\perp^* = 0$, (i) implies (ii).

If (ii) holds, there exists $X \in \mathbb{X}^P$ such that $M^*XM - \beta^2 X < 0$ on the kernel of N. By continuity, there exists $\epsilon > 0$ such that $\langle (M^*XM - \beta^2 X)v, v \rangle < 0$ for all v, $|v| = 1$, $|Nv|^2 \leq \epsilon$. Now choose $\eta > 0$ such that $\lambda_{max}(M^*\eta XM - \beta^2 \eta X) < \epsilon$.

This gives $\langle (M^*\eta XM - \beta^2 \eta X - N^*N)v, v \rangle < 0$ for all $v \neq 0$, so ηX solves (i).

■

The previous conditions are both LMIs (strictly speaking, (i) is affine rather than linear). While version (i) is more directly related to robustness analysis tests, (ii) is of lower dimensionality and therefore preferable from a computational point of view. We define the upper bound for μ,

$$\hat{\mu}_\Delta(N, M) := \inf\{\beta > 0 : \ (8.8) \text{ is satisfied}\}. \tag{8.10}$$

The fact that $\mu_\Delta(N, M) \leq \hat{\mu}_\Delta(N, M)$ is a consequence of Theorem 8.2 below.

Remarks:

- LMI (8.9) for $\beta = 1$ has appeared in previous work [46] on stabilization of input-output LFT systems, where it characterizes the so-called Q-detectability of the pair (N,M) (this reinforces the connection with the PBH test mentioned earlier). It is shown in [46] that it is equivalent to the existence of an output injection matrix L such that $\inf_{X \in \mathbb{X}^P} \bar{\sigma}\,(X(M + LN)X^{-1}) < 1$.

- If the structure includes real δI blocks (corresponding to parametric uncertainty), the upper bound can be improved in the same manner as in (2.39) for the standard case.

For the analysis of the upper bound we introduce a set ∇^0 which is closely related with the corresponding method for standard μ (see [55]) and to the IQC methods to be used in Section 8.2. For $\zeta \in \mathbb{C}^n$, let

$$\begin{aligned}
&\Sigma_i^0(\zeta) = (M\zeta)_i(M\zeta)_i^* - \zeta_i\zeta_i^*, \quad i = 1 \dots L, \\
&\sigma_{L+j}^0(\zeta) = (M\zeta)_{L+j}^*(M\zeta)_{L+j} - \zeta_{L+j}^*\zeta_{L+j}, \quad j = 1 \dots F, \\
&\Lambda^0(\zeta) = \mathrm{diag}\left[\Sigma_1^0(\zeta), \dots, \Sigma_L^0(\zeta), \sigma_{L+1}^0(\zeta) I_{m_1}, \dots, \sigma_{L+F}^0(\zeta) I_{m_F}\right].
\end{aligned} \tag{8.11}$$

Define $\nabla^0 = \{\Lambda^0(\zeta) : N\zeta = 0,\ |\zeta| = 1\} \subset \mathbb{X}^H$; $\mathrm{co}(\nabla^0)$ denotes its convex hull.

Theorem 8.2

1. *$\mu_\Delta(N, M) < 1$ if and only if $\nabla^0 \cap \overline{\mathbb{X}^P} = \emptyset$.*
2. *The following are equivalent:*

$$\begin{aligned}
&(i) \quad \hat{\mu}_\Delta(N, M) < 1, \\
&(ii) \quad \exists X \in \mathbb{X}^P : M^*XM - X - N^*N < 0, \\
&(iii) \quad \mathrm{co}(\nabla^0) \cap \overline{\mathbb{X}^P} = \emptyset.
\end{aligned} \tag{8.12}$$

Proof:

1. Assume $\mu_\Delta(N, M) \geq 1$. Then there exists $\Delta \in B_\Delta$ such that (8.2) has nontrivial solutions, say ζ of norm 1. Then $N\zeta = 0$, $\Delta M\zeta = \zeta$, which gives

$$\begin{aligned}
\delta_i I(M\zeta)_i = \zeta_i \quad &\Rightarrow \quad (M\zeta)_i(M\zeta)_i^* \geq \zeta_i\zeta_i^* \quad \Rightarrow \quad \Sigma_i^0(\zeta) \geq 0, \quad i = 1 \dots L, \\
\Delta_{L+j}(M\zeta)_{L+j} = \zeta_{L+j} \quad &\Rightarrow \quad |(M\zeta)_{L+j}|^2 \geq |\zeta_{L+j}|^2 \quad \Rightarrow \quad \sigma_{L+j}^0(\zeta) \geq 0, \quad j = 1 \dots F.
\end{aligned} \tag{8.13}$$

   Therefore the matrix $\Lambda^0(\zeta)$ is in $\nabla^0 \cap \overline{\mathbb{X}^P}$. The converse follows similarly.

2. The equivalence of (i) and (ii) is just the definition of $\hat{\mu}_\Delta(N, M)$.

   Let $X > 0$ solve (8.12). For any $\zeta \in \mathbb{C}^n$, $N\zeta = 0$, some algebra shows that

$$\left\langle X, \Lambda^0(\zeta) \right\rangle = \sum_{i=1}^{L} \mathrm{trace}(X_i\Sigma_i^0(\zeta)) + \sum_{j=1}^{F} x_{L+j}\sigma_{L+j}^0(\zeta) = \zeta^*(M^*XM - X - N^*N)\zeta < 0. \tag{8.14}$$

   Also, $\langle X, Y \rangle \geq 0$ for all $Y \in \overline{\mathbb{X}^P}$. Therefore the hyperplane $\langle X, Y \rangle = 0$ in $\mathbb{X}^H$ strictly separates the sets ∇^0 and $\overline{\mathbb{X}^P}$, which implies their respective convex hulls $\mathrm{co}(\nabla^0)$ and $\overline{\mathbb{X}^P}$ are disjoint, proving (iii). Conversely, if $\mathrm{co}(\nabla^0)$, $\overline{\mathbb{X}^P}$ are disjoint, a strictly separating hyperplane can be found from Lemma A.3 leading back to (ii). ∎

8.1.2 μ-Simple Structures in the Implicit Case

The upper bound for implicit μ will be strict in general; equivalently, LMI (8.12) is not a necessary test for stability. In a similar manner as in the standard case [55], we now pose the question as to which special Δ structures give equality of μ and $\hat{\mu}$.

Definition 12 *The structure Δ is μ-simple in the implicit case if $\mu_\Delta(N,M) = \hat{\mu}_\Delta(N,M)$ for* any *matrices M, N.*

Theorem 8.3 *The following structures are μ-simple in the implicit case:*
(i) $\Delta = \{\delta I : \delta \in \mathbb{C}\}$;
(ii) $\Delta = \{\mathrm{diag}[\Delta_1, \ldots \Delta_F] : \Delta_i \in \mathbb{R}^{m_i \times m_i}\}$, with $F \leq 2$, for M, N real;
(iii) $\Delta = \{\mathrm{diag}[\Delta_1, \ldots, \Delta_F] : \Delta_i \in \mathbb{C}^{m_i \times m_i}\}$, with $F \leq 3$.

Proof:

(i) In the case $\Delta = \delta I$, if $\mu_\Delta(N,M) < 1$ then (N,M) is detectable in the usual system theoretic sense, so there exists an output injection L such that $\rho(M+LN) < 1$. From Lyapunov theory this implies there exists $X > 0$ such that

$$(M+LN)^* X (M+LN) - X < 0. \tag{8.15}$$

Multiplying on the left and right by $N_\perp$, $N_\perp^*$ gives $N_\perp(M^* X M - X)N_\perp^* < 0$ which implies $\hat{\mu}_\Delta(N,M) < 1$.

(ii) The only nontrivial case is $F = 2$. Let M, N, $\Delta = \mathrm{diag}[\Delta_1, \Delta_2]$ be *real* matrices. To analyze this case we must consider a real version of the ∇^0 set, with the same definition as before but with $\zeta \in \mathbb{R}^n$. Consider the $n \times r$ matrix $N_\perp'$ parameterizing the kernel of N ($\zeta = N_\perp' v$), and assume it is isometric. Then ∇^0 can be rewritten as

$$\nabla^0 = \{\Lambda(v) = [\sigma_1(v), \sigma_2(v)], v \in \mathbb{R}^r, |v| = 1\}. \tag{8.16}$$

where $\sigma_j(v) = v' H_j v$, and $H_j = (MN_\perp')_j'(MN_\perp')_j - (N_\perp)_j(N_\perp')_j$ is a real, symmetric matrix for $j = 1, 2$. To prove that this structure is μ-simple is equivalent, by Theorem 8.2 to the fact that

$$\nabla^0 \cap \overline{\mathbb{X}^F} = \emptyset \Longrightarrow \mathrm{co}(\nabla^0) \cap \overline{\mathbb{X}^F} = \emptyset \tag{8.17}$$

for any M, N, or equivalently for any symmetric H_1, H_2. In this case $\overline{\mathbb{X}^F} = (\mathbb{R}^+)^2$, the closed first quadrant in $\mathbb{R}^2$. We have therefore restated the problem as a geometric condition on the range of two real quadratic forms, which is equivalent to an "S-procedure losslessness" theorem from Yakubovich [94]; since this literature is not easily accessed we include a proof which is based on some modifications to the parallel results of [28].

Let $P = \Lambda(v_P)$, $Q = \Lambda(v_Q)$, be two distinct points in ∇^0 (v_P, $v_Q \in \mathbb{R}^r$, $|v_P| = |v_Q| = 1$). Define

$$V = \begin{bmatrix} v_P' \\ v_Q' \end{bmatrix} [\, v_P \;\; v_Q \,] > 0, \tag{8.18}$$

$$\hat{H}_i = V^{-\frac{1}{2}} \begin{bmatrix} v_P' \\ v_Q' \end{bmatrix} H_i [\, v_P \;\; v_Q \,] V^{-\frac{1}{2}}, \qquad j = 1, 2, \tag{8.19}$$

$$\mathcal{E} = \left\{ [\, \eta' \hat{H}_1 \eta \;,\; \eta' \hat{H}_2 \eta \,], \eta \in \mathbb{R}^2, |\eta| = 1 \right\}. \tag{8.20}$$

Then

- $\mathcal{E} \subset \nabla^0$. This follows from the fact that if $|\eta| = 1$, $|[\, v_P \;\; v_Q \,] V^{-\frac{1}{2}} \eta| = 1$ from (8.18).
- $P, Q \in E$. For P set $\eta_P = V^{\frac{1}{2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which verifies $|\eta_P|^2 = [\, 1 \;\; 0 \,] V \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |v_P|^2 = 1$, and analogously for Q.
- $\mathcal{E}$ is an ellipse in $\mathbb{R}^2$ (which may degenerate to a segment).

  Parametrize $\eta = (\cos(\theta), \sin(\theta))$, $\theta \in [-\pi, \pi]$. If $\hat{H}_j = \begin{bmatrix} a_j & b_j \\ b_j & c_j \end{bmatrix}$, then

$$\eta' \hat{H}_j \eta = \frac{a_j + c_j}{2} + \left[\tfrac{a_j - c_j}{2} \;\; b_j \right] \begin{bmatrix} \cos(2\theta) \\ \sin(2\theta) \end{bmatrix} \quad j = 1, 2. \tag{8.21}$$

  This implies that $\mathcal{E}$ is the image of the unit circle by an affine map, an ellipse.

We have shown that given two points in ∇^0, there exists an ellipse $\mathcal{E} \subset \nabla^0$ through those points. Now we return to (8.17). If $\mathrm{co}(\nabla^0) \cap (\mathbb{R}^+)^2 \neq \emptyset$, since ∇^0 is bounded and $(\mathbb{R}^+)^2$ is a cone, there exists a point in the *boundary* of $\mathrm{co}(\nabla^0)$ which falls in the first quadrant. Using Lemma A.2 there exist two points P,Q in ∇^0 such that the segment PQ intersects the first quadrant. But then the corresponding ellipse $\mathcal{E}$ will intersect $(\mathbb{R}^+)^2$; (the geometric picture is given in Figure 8.1(ii)). This implies $\nabla^0 \cap (\mathbb{R}^+)^2 \neq \emptyset$.

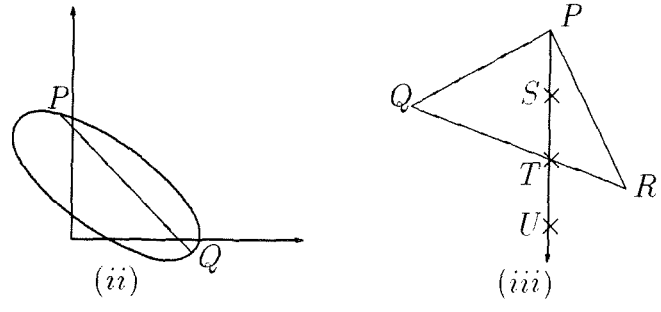


Figure 8.1: Illustration to the proof of Theorem 8.3

(iii).

We consider the case $F = 3$, the others follow similarly. The same procedure as in (ii) yields

$$\nabla^0 = \left\{\Lambda(v) = [\sigma_1(v), \sigma_2(v), \sigma_3(v)] \in \mathbb{R}^3, v \in \mathbb{C}^r, |v| = 1\right\}, \tag{8.22}$$

where $\sigma_j(v) = v^* H_j v$, and H_j are complex hermitian forms in $\mathbb{C}^r$. Similarly, we must show the geometric result

$$\nabla^0 \cap (\mathbb{R}^+)^3 = \emptyset \Longrightarrow \text{co}(\nabla^0) \cap (\mathbb{R}^+)^3 = \emptyset \tag{8.23}$$

for any H_1, H_2, H_3. Once again, this result appears in the "S-procedure" formulation [33]. The following proof is based on [28]. In particular, it is shown in [28] (analogously to (8.21)) that for the case $r = 2$, the set ∇^0 is the image of the unit sphere in $\mathbb{R}^3$ by an affine map $g : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. This gives an ellipsoid $\mathcal{E}$ (with no interior) in $\mathbb{R}^3$, which could also degenerate to a projection of such an ellipsoid in a lower dimensional subspace.

Given two distinct points $P = \Lambda(v_P)$, $Q = \Lambda(v_Q)$ in ∇^0, an analogous construction as the one given in (8.18-8.20) (with analogous proof) shows that there is such an ellipsoid $\mathcal{E} \subset \nabla^0$ through the two points.

Assume now that $\text{co}(\nabla^0) \cap (\mathbb{R}^+)^3 \neq \emptyset$. Picking a point in the boundary of $\text{co}(\nabla^0)$, Lemma A.2 implies that there are 3 points P,Q,R in ∇^0 such that some convex combination $S = \alpha P + \beta Q + \gamma R$ falls in $(\mathbb{R}^+)^3$. Geometrically, the triangle PQR intersects the positive "octant" at S.

Claim: S lies in a segment between 2 points in ∇^0.

This is obvious if P,Q, R are aligned or if any of α,β, γ is 0. If not, consider the following reasoning, illustrated in Figure 8.1. Write

$$S = \alpha P + \beta Q + \gamma R = \alpha P + (\beta + \gamma)\frac{1}{\beta + \gamma}(\beta Q + \gamma R) = \alpha P + (\beta + \gamma)T, \tag{8.24}$$

where T lies in the segment QR. Now consider the ellipsoid $\mathcal{E} \subset \nabla^0$ through Q and R. If it degenerates to 1 or 2 dimensions, then $T \in \mathcal{E} \subset \nabla^0$ and the claim is proved. If not, T is interior to the ellipsoid $\mathcal{E}$. The half line starting at P, through T must "exit" the ellipsoid a point $U \in \mathcal{E} \subset \nabla^0$ such that T is in the segment PU. Therefore S in the segment PU, and $P, U \in \nabla^0$, proving the claim.

To finish the proof, we have found two points in ∇^0 such that the segment between them intersects $(\mathbb{R}^+)^3$. The corresponding ellipsoid $\mathcal{E} \subset \nabla^0$ between these points must clearly also intersect $(\mathbb{R}^+)^3$. Therefore $\nabla^0 \cap (\mathbb{R}^+)^3 \neq \emptyset$.

■

In reference to structures with only full blocks, we have shown that the situation is analogous to the standard case of [55]: the upper bound is exact for a maximum of 3 complex full blocks or 2 real full blocks. The only notable difference in the implicit case is the fact that the structure $\Delta = \{\text{diag}[\delta_1 I, \Delta_2]\}$ is no longer μ-simple, as shown in the following example.

Example:

Let $\Delta = \text{diag}\,[\delta_1 I_2, \Delta_2]$, $\delta_1 \in \mathbb{C}$, $\Delta_2 \in \mathbb{C}^{2\times 2}$,

$$N_\perp^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 4 & 0 \\ 0 & 4 \end{bmatrix}, \quad MN_\perp^* = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 3 & 2 \\ 2 & 2 \end{bmatrix}.$$

The top half of $N_\perp^* - \Delta M N_\perp^*$ is $\begin{bmatrix} 1-2\delta_1 & 0 \\ 0 & 1-3\delta_1 \end{bmatrix}$, so the kernel is nontrivial only for $\delta_1 = 1/2$ or $\delta_1 = 1/3$. In the first case, the kernel must be the span of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ therefore $\begin{bmatrix} 4 \\ 0 \end{bmatrix} = \Delta_2 \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. This can be achieved with $\bar{\sigma}\,(\Delta_2)$ of at least $4/\sqrt{13}$.

A similar argument with $\delta_1 = 1/3$, shows that for a nontrivial kernel, $\bar{\sigma}\,(\Delta_2) \geq \sqrt{2}$. The first perturbation is smaller so $\mu_\Delta(N, M) = \sqrt{13}/4 < 1$.

For the LMI, write $X = \text{diag}\,[X_0, I_2]$, with $X_0 = \begin{bmatrix} x & y \\ y^* & z \end{bmatrix}$. Some algebra gives

$$N_\perp(M^* X M - X)N_\perp^* = \begin{bmatrix} 3(x-1) & 5(y+2) \\ 5(y^*+2) & 8(z-1) \end{bmatrix}. \tag{8.25}$$

For (8.25) to be negative definite, and $X > 0$, we must have

$$0 < x < 1; \quad 0 < z < 1: \quad |y|^2 < xz; \quad |y+2|^2 < \frac{24}{25}(1-x)(1-z). \tag{8.26}$$

This implies $|y| < 1$, $|y+2| < 1$ which is impossible, so there is no solution to LMI (8.9) with $\beta = 1$. Consequently, $\mu_\Delta(N, M) < \hat{\mu}_\Delta(N, M)$.

We have developed the basic theory of the structured singular value for implicit analysis. Given the fact that this problem is NP-hard, the main open challenge is the development of heuristics to compute the answer in typical cases, avoiding unacceptable growth rates with problem size. This requirement is particularly important in the case of problems involving data, as those discussed in Section 7.3.

8.2 Implicit Analysis with LTV Uncertainty

In this section we return to dynamic problems, and study the analysis configuration (8.1) for the case where $\mathbf{M}$, $\mathbf{N}$ are LTI maps, and $\mathbf{\Delta}$ is a structured, LTV operator. Given the relationship between IQCs and LTV operators, this is a natural situation for an over-constrained analysis problem. As in Chapter 7, we address separately the conditions for stability and l_2 stability; in Section 8.2.3 we discuss how both cases apply to a constrained robust performance problem.

8.2.1 A Necessary and Sufficient Condition for Robust l_2-Stability

We first consider analysis in the l_2 setting; $\mathbf{M}$,$\mathbf{N}$, $\mathbf{\Delta}$ are in $\mathcal{L}(l_2)$; $\mathbf{M}$, $\mathbf{N}$ are LTI, and $\mathbf{\Delta}$ varies in the set

$$\boldsymbol{B}_{\boldsymbol{\Delta}_{\boldsymbol{NC}}^{\boldsymbol{LTV}}} := \{\mathbf{\Delta} \in \mathcal{L}(l_2), \mathbf{\Delta} = \mathrm{diag}\,[\delta_1 \mathbf{I}_{n_1}, \ldots, \delta_L \mathbf{I}_{n_L}, \mathbf{\Delta}_{L+1}, \ldots, \mathbf{\Delta}_{L+F}]\}\,, \tag{8.27}$$

where we include non-causal structured operators. $\mathbb{X}$ is the set of constant scales as in (2.35). The following result is an extension of the constant scales condition (2.42) for standard robust stability analysis.

Theorem 8.4 *The system (8.1) is robustly l_2-stable if and only if*

$$\exists X \in \mathbb{X}^P : \quad \mathbf{M}^* X \mathbf{M} - X - \mathbf{N}^* \mathbf{N} < 0. \tag{8.28}$$

Remark: The inequality in (8.28) is interpreted in the sense of operators on l_2, i.e. $\Pi < 0$ if $\langle z, \Pi z \rangle \leq -\epsilon \|z\|^2$ for some $\epsilon > 0$. For the case where $\mathbf{M}$, $\mathbf{N}$ are finite dimensional, (8.28) reduces to an the LMI across frequency

$$\exists X \in \mathbb{X}^P : \quad M(e^{j\omega})^* X M(e^{j\omega}) - X - N(e^{j\omega})^* N(e^{j\omega}) < 0 \quad \forall\, \omega \in [0, 2\pi]. \tag{8.29}$$

Proof:

[Sufficiency]:

Fix $X > 0$ which solves (8.28). Without loss of generality it can be assumed that $X = I$. This results from the fact that invertible operations yield

$$\begin{bmatrix} X^{\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}\mathbf{M} \\ \mathbf{N} \end{bmatrix} X^{\frac{1}{2}} = \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}\bar{\mathbf{M}} \\ \bar{\mathbf{N}} \end{bmatrix}, \tag{8.30}$$

where $\bar{\mathbf{M}} := X^{\frac{1}{2}} \mathbf{M} X^{-\frac{1}{2}}$, $\bar{\mathbf{N}} := \mathbf{N} X^{-\frac{1}{2}}$ verify (8.28) with $X = I$. The notion of negative definiteness allows by continuity to find $\gamma^2 < 1$ such that

$$\mathbf{M}^* \mathbf{M} - \gamma^2 \mathbf{I} - \mathbf{N}^* \mathbf{N} < 0. \tag{8.31}$$

Let $e, z \in l_2$, $\|z\| = 1$, and $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}}$ satisfy

$$\begin{bmatrix} \mathbf{I} - \mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix} z = e = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}. \tag{8.32}$$

From $\|\mathbf{\Delta}\| \leq 1$ and (8.31) we have

$$\|\mathbf{\Delta M} z\|^2 \leq \|\mathbf{M}z\|^2 = \langle \mathbf{M}^*\mathbf{M}z, z\rangle \leq \gamma^2\langle z, z\rangle + \langle \mathbf{N}^*\mathbf{N}z, z\rangle = \gamma^2 + \|e_2\|^2. \tag{8.33}$$

Therefore $\|\mathbf{\Delta M}z\| \leq \gamma + \|e_2\|$, and $\|e_1\| = \|z - \mathbf{\Delta M}z\| \geq (1-\gamma) - \|e_2\|$. This leads to $\|e\| \geq \frac{1-\gamma}{\sqrt{2}}$, which proves robust l_2-stability by Proposition 7.2.

[Necessity:]

The proof is based on an S-procedure losslessness argument similar to those used in Chapters 5 and 6. Some minor modifications are required to account for the additional constraints $\mathbf{N}$. We will also use this opportunity to explain in more detail the proof for the case where $\mathbf{\Delta}$ includes $\boldsymbol{\delta}\mathbf{I}$ operators, which was only mentioned briefly in Chapter 6.

Analogously to (5.6), and (5.13), we characterize the uncertainty by the following IQCs:

$$\begin{aligned}
&\Sigma_i(z) = \int_0^{2\pi} \left[(Mz)_i(e^{j\omega})(Mz)_i^*(e^{j\omega}) - z_i(e^{j\omega})z_i^*(e^{j\omega})\right] \frac{d\omega}{2\pi}, \quad i = 1 \ldots L; \\
&\sigma_{L+j}(z) = \|(\mathbf{M}z)_{L+j}\|^2 - \|z_{L+j}\|^2, \quad j = 1 \ldots F; \\
&\Lambda(z) = \operatorname{diag}\left[\Sigma_1(z), \ldots, \Sigma_L(z), \sigma_{L+1}(z)I_{m_1}, \ldots, \sigma_{L+F}(z)I_{m_F}\right].
\end{aligned} \tag{8.34}$$

The Σ_i characterize the $\boldsymbol{\delta}\mathbf{I}$ blocks, and the σ_{L+j} the full blocks. As usual, the IQCs are collected in Λ which takes values in $\mathbb{X}^H$. Define a subset of $\mathbb{X}^H$,

$$\nabla^\epsilon = \{\Lambda(z) : \|z\| = 1, \|\mathbf{N}z\| < \epsilon\}. \tag{8.35}$$

We make the following statements:

Lemma 8.5 *The closure $\overline{\nabla^\epsilon}$ of ∇^ϵ is convex and compact in $\mathbb{X}^H$.*

Proposition 8.6 *If the system (8.1) is robustly l_2-stable, there exists $\epsilon > 0$ such that*

$$\overline{\nabla^\epsilon} \cap \overline{\mathbb{X}^P} = \emptyset. \tag{8.36}$$

Proof: Lemma 8.5 follows from the methods in Chapter 5, in an analogous way to Lemma 6.9. Proposition 8.6 is covered in Appendix B.

$\overline{\nabla^\epsilon}$ and $\overline{\mathbb{X}^P}$ are disjoint convex sets in the real, finite dimensional vector space $\mathbb{X}^H$, where we use the inner product (8.7). $\overline{\nabla^\epsilon}$ is compact and $\overline{\mathbb{X}^P}$ is closed. From Lemma A.3 there exists $X \in \mathbb{X}^H$, α, β in $\mathbb{R}$ such that

$$\langle X, \Lambda \rangle \leq \alpha < \beta \leq \langle X, \hat{X} \rangle \quad \forall\, \Lambda \in \overline{\nabla^\epsilon},\ \ \hat{X} \in \overline{\mathbb{X}^P}. \tag{8.37}$$

Since $\overline{\mathbb{X}^P}$ is a cone, β can be chosen to be 0. It is easy to show that

$$\langle X, \hat{X} \rangle \geq 0 \ \ \forall \hat{X} \in \mathbb{X}^H,\ \hat{X} \geq 0 \implies X \geq 0. \tag{8.38}$$

A small perturbation of X ensures $X > 0$ ($X \in \mathbb{X}^P$), and by continuity and compactness of $\overline{\nabla^\epsilon}$ we can modify α to achieve for the new X

$$\langle X, \Lambda \rangle \leq \alpha < 0 \quad \forall \Lambda \in \overline{\nabla^\epsilon}. \tag{8.39}$$

Furthermore, by scaling down X (and α) we can ensure (8.39) holds and

$$\|\mathbf{M}^* X \mathbf{M} - X\| = \gamma^2 < \epsilon^2. \tag{8.40}$$

Let $\Psi = \mathbf{M}^* X \mathbf{M} - X - \mathbf{N}^* \mathbf{N}$. Now for any $z = l_2$, $\|z\| = 1$,

$$\langle X, \Lambda(z) \rangle = \sum_{i=1}^{L} \mathrm{trace}(X_i \Sigma_i(z)) + \sum_{j=1}^{F} x_{L+j} \sigma_{L+j}(z) = \langle z, (\mathbf{M}^* X \mathbf{M} - X) z \rangle. \tag{8.41}$$

If $\|\mathbf{N}z\| < \epsilon$, then $\Lambda(z) \in \nabla^\epsilon$ so from (8.39) we have

$$\langle z, \Psi z \rangle \leq \langle z, (\mathbf{M}^* X \mathbf{M} - X) z \rangle = \langle X, \Lambda(z) \rangle \leq \alpha < 0. \tag{8.42}$$

If $\|\mathbf{N}z\| \geq \epsilon$, then from (8.40) we have

$$\langle z, \Psi z \rangle \leq \langle z, (\mathbf{M}^* X \mathbf{M} - X) z \rangle - \|\mathbf{N}z\|^2 \leq \gamma^2 - \epsilon^2 < 0. \tag{8.43}$$

We conclude that $\langle z, \Psi z \rangle \leq \max(\alpha, \gamma^2 - \epsilon^2) < 0$ for $\|z\| = 1$, which implies $\Psi < 0$.

■

8.2.2 Necessary, Sufficient Conditions in the l_{2e} Setting

The motivation provided in Chapter 7 for over-constrained problems refers mainly to the l_2 setting. As seen in Chapter 2, however, in the standard case the condition $M^* X M - X < 0$ is necessary and sufficient for robust stability in the sense of l_{2e}. It therefore seems natural to explore this issue in the constrained case of (7.25). We first consider necessity:

Theorem 8.7 *Let* $\mathbf{M}$, $\mathbf{N}$ *be LTI in* $\mathcal{L}_c(l_{2e})$, $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$. *If the system (8.1) is robustly stable then (8.28) holds.*

The proof can be derived exactly as in Theorem 8.4. The only consideration is that in the proof of Proposition 8.6, causality of perturbations must be imposed; given the results of Appendix B, this is of no consequence.

Regarding the issue of sufficiency of condition (8.28), we consider the following example.

Example:

Let $\mathbf{M} = 2$, $\mathbf{N} = 1 - 2\boldsymbol{\lambda}$. $\Delta \in \mathcal{L}_c(l_{2e})$. For any $X > 0$,

$$\langle (\mathbf{M}^* X \mathbf{M} - X - \mathbf{N}^* \mathbf{N}) v, v \rangle = 3X\|v\|^2 - \|(1 - 2\boldsymbol{\lambda})v\|^2 \leq (3X - 1)\|v\|^2, \tag{8.44}$$

so for $0 < X < \frac{1}{3}$, (8.28) is satisfied.

But the perturbation $\Delta = \boldsymbol{\lambda}$ gives $\begin{bmatrix} \mathbf{I} - \boldsymbol{\Delta}\mathbf{M} \\ \mathbf{N} \end{bmatrix} = \begin{bmatrix} 1 - 2\boldsymbol{\lambda} \\ 1 - 2\boldsymbol{\lambda} \end{bmatrix}$ which is unstable.

This example shows that in general, condition (8.28), (or (8.29)) is *not* sufficient for robust stability, even for time invariant perturbations: (8.28) does not provide information of the behavior outside l_2. One could think of strengthening condition (8.29) to include frequency points inside the disk; this would eliminate the previous counterexample, and in fact guarantee robust stability with LTI perturbations, but it is still not sufficient for the LTV case. A counterexample for this is the system $\mathbf{M} = 4\boldsymbol{\lambda} - 8\boldsymbol{\lambda}^2$, $\mathbf{N} = 1 - 4\boldsymbol{\lambda} + 4\boldsymbol{\lambda}^2$; we omit the rather lengthy verification. This difficulty can be related to a situation in [90], involving a frequency domain condition over the right half plane.

We can state, however, a partial result which is applicable in many cases.

Theorem 8.8 *Let* $\mathbf{M}$ *be* LTI *in* $\mathcal{L}_c(l_{2e})$, *and* $\mathbf{N}$ *be a static map. If (8.28) holds, then system (8.1) is robustly stable over* $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$.

Proof:

Analogously to Theorem 8.4, assume without loss of generality that $X = I$, and let γ satisfy (8.31). Let T be fixed, z, e satisfy (8.32), with truncation $\mathbf{P}^{\mathbf{T}} z$ of norm 1. From causality

$$\mathbf{P}^{\mathbf{T}} \begin{bmatrix} \mathbf{I} - \boldsymbol{\Delta}\mathbf{M} \\ \mathbf{N} \end{bmatrix} \mathbf{P}^{\mathbf{T}} z = \mathbf{P}^{\mathbf{T}} e = \begin{bmatrix} \mathbf{P}^{\mathbf{T}} e_1 \\ \mathbf{P}^{\mathbf{T}} e_2 \end{bmatrix}. \tag{8.45}$$

Since $\mathbf{P}^{\mathbf{T}} z \in l_2$, from (8.31) we obtain

$$\|\mathbf{M}\mathbf{P}^{\mathbf{T}} z\|^2 < \gamma^2 \|\mathbf{P}^{\mathbf{T}} z\|^2 + \|\mathbf{N}\mathbf{P}^{\mathbf{T}} z\|^2. \tag{8.46}$$

Since $\mathbf{N}$ is static, $\mathbf{N}\mathbf{P}^{\mathbf{T}}z = \mathbf{P}^{\mathbf{T}}e_2$; also $\|\mathbf{P}^{\mathbf{T}}\boldsymbol{\Delta}\mathbf{M}\mathbf{P}^{\mathbf{T}}z\|^2 \leq \|\mathbf{M}\mathbf{P}^{\mathbf{T}}z\|^2$, leading to

$$\|\mathbf{P}^{\mathbf{T}}e_1\| = \|\mathbf{P}^{\mathbf{T}}(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})\mathbf{P}^{\mathbf{T}}z\| \geq 1 - \gamma - \|\mathbf{P}^{\mathbf{T}}e_2\|. \tag{8.47}$$

We conclude that $\|\mathbf{P}^{\mathbf{T}}e\| \geq \epsilon$ for a fixed $\epsilon > 0$, independent of T, $\boldsymbol{\Delta}$, which implies robust stability from Proposition 7.1. ■

8.2.3 Application to Constrained Robust Performance Analysis

To conclude this section we review the problem considered in Section 7.2, of robust performance analysis under constrained disturbances. Referring to (7.8-7.10) and Figure 7.3 we assume that $\mathbf{T}$, $\mathbf{U}$ and $\mathbf{V}$ are causal, LTI, stable maps. We recall from section 7.2 that without loss of generality in the analysis, $\mathbf{U}$ can be chosen to be a static map $k\mathbf{I}$. Also the perturbations $\boldsymbol{\Delta}_P$, $\boldsymbol{\Delta}_C$ vary in the class of arbitrary time-varying operators.

The reduction procedure from Proposition 7.4 yields a system of the form (8.1), where

$$\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\Delta}_u & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Delta}_C \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \mathbf{0} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{V} & \mathbf{0} \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} \mathbf{0} & \mathbf{U} & -\mathbf{I} \end{bmatrix}. \tag{8.48}$$

If $\mathbf{U}$ is static, we observe that $\mathbf{N}$ is static, as in the assumption of Theorem 8.8. Therefore if $\boldsymbol{\Delta}_u$ is LTV uncertainty, condition (8.28) is necessary and sufficient for

- robust l_2-stability in the case $\boldsymbol{\Delta}_u$, $\boldsymbol{\Delta}_P$, $\boldsymbol{\Delta}_C$ in $\mathcal{L}(l_2)$,
- robust stability in the case $\boldsymbol{\Delta}_u$, $\boldsymbol{\Delta}_P$, $\boldsymbol{\Delta}_C$ in $\mathcal{L}_c(l_{2e})$,

If we review the problem statement that let to equations (7.8-7.10), the case we are interested in testing for is a hybrid of the two cases considered above. $\boldsymbol{\Delta}_u$ will typically be a causal perturbation, and it must be guaranteed in the first place that the z variable does not "blow up" (it remains in l_2 if v and the equation errors are in l_2). Once this is known, the analysis can be restricted to l_2 and the (possibly non-causal) perturbations $\boldsymbol{\Delta}_P$, $\boldsymbol{\Delta}_C$ can be considered, casting the robust performance analysis as a robust l_2 stability question, as argued in section 7.2.

This hybrid question is also answered by the test (8.28); in fact, since the first block of $\mathbf{N}$ is zero, the upper portion of $\mathbf{M}^*X\mathbf{M} - X - \mathbf{N}^*\mathbf{N}$ is $\mathbf{T}_{11}^*X_u\mathbf{T}_{11} - X_u$, which provides the standard robust stability test in the z variables, in addition to the robust l_2 stability test on all the variables.

8.3 Analysis of State-Space Systems

In the previous sections the LTI systems $\mathbf{M}$, $\mathbf{N}$ were treated as operators and the convex analysis conditions are infinite dimensional, involving a search in the frequency domain. We now consider the case when $\mathbf{M}$, $\mathbf{N}$ are finite dimensional, causal LTI systems, described by the joint state-space realization

$$\left[\begin{array}{c}\mathbf{M}\\ \mathbf{N}\end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C_M & D_M \\ C_N & D_N \end{array}\right]. \tag{8.49}$$

The state equations are of the form $x = \boldsymbol{\lambda}(Ax + Bz)$, with $\boldsymbol{\lambda}$ the delay operator, and x the state. By adding these equations in implicit form to (8.1) we obtain

$$\begin{aligned}\left(\mathbf{I} - \left[\begin{array}{cc}\boldsymbol{\lambda}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Delta}\end{array}\right]\left[\begin{array}{cc} A & B \\ C_M & D_M\end{array}\right]\right)\left[\begin{array}{c} x \\ z\end{array}\right] &= 0,\\ \left[\begin{array}{cc} C_N & D_N\end{array}\right]\left[\begin{array}{c} x \\ z\end{array}\right] &= 0,\end{aligned} \tag{8.50}$$

which will be represented by

$$\left[\begin{array}{c} I - \boldsymbol{\Delta}_s M_S \\ N_S\end{array}\right]\left[\begin{array}{c} x \\ z\end{array}\right] = 0, \tag{8.51}$$

with M_S, N_S constant matrices, and the delay-uncertainty operator $\boldsymbol{\Delta}_s := \mathrm{diag}[\boldsymbol{\lambda}, \boldsymbol{\Delta}]$. Note that (8.51) has once more the form (8.1). In this section we consider causal operators $\boldsymbol{\Delta}$ and the corresponding notion of robust stability.

If $\boldsymbol{\Delta}$ is LTI, then the robust stability test is an implicit μ condition $\mu_{\boldsymbol{\Delta}_s}(N_S, M_S) < 1$; a more useful test is, however, the corresponding upper bound

$$M_S^* X_S M_S - X_S - N_S^* N_S < 0, \tag{8.52}$$

where $X_S = \mathrm{diag}[X_0, X]$ is defined to commute with Δ_S, as in (2.46) for the standard case. A remark is that this upper bound is conservative for $\mu_{\boldsymbol{\Delta}_s}(N_S, M_S) < 1$ even for an unstructured $\boldsymbol{\Delta}$ (recall the characterization of μ-simple structures in Section 8.1).

Condition (8.52) can be related to LTV $\boldsymbol{\Delta}$: if follows from Theorem 8.8 that (8.52) is sufficient condition for robust stability of (8.51), for any $\boldsymbol{\Delta}_s := \mathrm{diag}[\boldsymbol{\lambda}, \boldsymbol{\Delta}]$, and $\boldsymbol{\Delta}$ in $\boldsymbol{B\Delta}_{LTV}$.

It is not clear, however, whether the condition is necessary, as it is in the unconstrained case. A special case where this holds is when the constraints N_S do not involve the state

variables x: consider the state space implicit system of Figure 8.2(a), which corresponds to the case $C_N = 0$ in (8.50). The corresponding system (8.1), depicted in Figure 8.2(b), has static $\mathbf{N}$ constraints. As remarked in Section 8.2.3, a general robust performance problem with a finite number of IQCs in the disturbance variables can be cast in this special form.

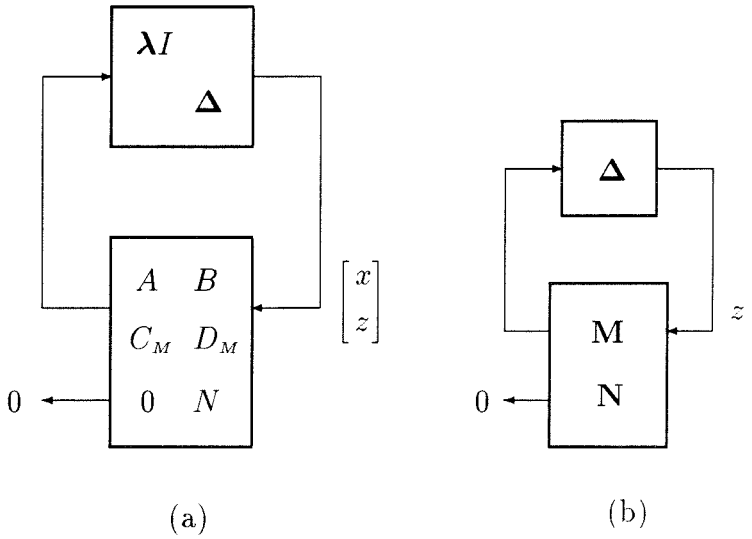


Figure 8.2: State space formulation of the robust stability problem

Theorem 8.9 *In the notation of (8.50), assume that $C_N = 0$ and that $\mathbf{\Delta}$ varies in $\boldsymbol{B}_{\mathbf{\Delta}^{LTV}}$. The following are equivalent:*

(i) (8.52) *holds.*

(ii) *The implicit system* (8.51) *of Figure 8.2 (a) is robustly stable.*

(iii) $\rho(A) < 1$, *and the implicit system* (8.1) *of Figure 8.2 (b) is robustly stable.*

(iv) $\rho(A) < 1$, *and* $\exists X \in \mathbb{X}^P : \mathbf{M}^* X \mathbf{M} - X - \mathbf{N}^* \mathbf{N} < 0.$ (8.53)

Proof: $(i) \Rightarrow (ii)$ This follows from Theorem 8.8, as discussed before.

$(ii) \Rightarrow (iii)$. From (ii), $\begin{bmatrix} \mathbf{I} - \mathbf{\Delta}_s M_s \\ N_s \end{bmatrix}$ has a left inverse in $\mathcal{L}_e(l_{2\epsilon})$ for every $\mathbf{\Delta}_s = \mathrm{diag}[\lambda I, \mathbf{\Delta}]$, $\mathbf{\Delta} \in \boldsymbol{B}_{\mathbf{\Delta}^{LTV}}$. Setting $\mathbf{\Delta}_s = \mathrm{diag}[\lambda I, 0]$ implies that $(I - \lambda A)$ has a left inverse in $\mathcal{L}_e(l_{2\epsilon})$, so $\rho(A) < 1$. Now for any fixed $\mathbf{\Delta}_s$, the identity

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{\Delta} C_M (\mathbf{I} - \lambda A)^{-1} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}_s M_s \\ N_s \end{bmatrix} \begin{bmatrix} \mathbf{I} & (\mathbf{I} - \lambda A)^{-1} \lambda B \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} - \lambda A & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \mathbf{\Delta M} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \tag{8.54}$$

implies that the right hand side of (8.54) has a left inverse in $\mathcal{L}_c(l_{2e})$. So $\begin{bmatrix} I - \mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix}$ has a left inverse in $\mathcal{L}_c(l_{2e})$.

$(iii) \Rightarrow (iv)$ This is a direct application of Theorem 8.7.

$(iv) \Rightarrow (i)$.

Let the columns of $N_\perp^*$ form a basis for the kernel of N; (iv) leads to

$$N_\perp M(e^{j\omega})^* X M(e^{j\omega}) N_\perp^* < N_\perp X N_\perp^* \quad \forall \omega \in [-\pi, \pi]. \tag{8.55}$$

Since $N_\perp X N_\perp^* > 0$, pre and post multiplication by $(N_\perp X N_\perp^*)^{-\frac{1}{2}}$ gives $\|\bar{M}(\lambda)\|_\infty < 1$, where $\bar{M}(\lambda) = X^{\frac{1}{2}} M(\lambda) N_\perp^* (N_\perp X N_\perp^*)^{-\frac{1}{2}}$. We will now change notation, redefining $N_\perp^* (N_\perp X N_\perp^*)^{-\frac{1}{2}}$ as $N_\perp^*$, since it still spans the same column space. So $N_\perp X N_\perp^* = I$, and $\bar{M}(\lambda) = X^{\frac{1}{2}} M(\lambda) N_\perp^*$, which has state space realization

$$\left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C}_M & \bar{D}_M \end{array}\right] = \left[\begin{array}{c|c} A & B N_\perp^* \\ \hline X^{\frac{1}{2}} C_M & X^{\frac{1}{2}} D_M N_\perp^* \end{array}\right]. \tag{8.56}$$

It is well known (see, e.g. [55]) that the $\mathcal{H}_\infty$ condition $\rho(\bar{A}) < 1$, $\|\bar{M}(\boldsymbol{\lambda})\|_\infty < 1$ implies the existence of a solution $X_0 > 0$ to the LMI

$$\begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C}_M & \bar{D}_M \end{bmatrix}^* \begin{bmatrix} X_0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \bar{A} & \bar{B} \\ \bar{C}_M & \bar{D}_M \end{bmatrix} - \begin{bmatrix} X_0 & 0 \\ 0 & I \end{bmatrix} < 0. \tag{8.57}$$

Substituting the expressions for $\bar{A}, \bar{B}, \bar{C}_M, \bar{D}_M$, and using $N_\perp X N_\perp^* = I$, (8.57) leads to

$$\begin{bmatrix} I & 0 \\ 0 & N_\perp \end{bmatrix}^* \left(M_S^* \begin{bmatrix} X_0 & 0 \\ 0 & X \end{bmatrix} M_S - \begin{bmatrix} X_0 & 0 \\ 0 & X \end{bmatrix} \right) \begin{bmatrix} I & 0 \\ 0 & N_\perp^* \end{bmatrix} < 0.$$

Since $N_S = [0 \;\; N]$, $N_{S\perp} = \begin{bmatrix} I & 0 \\ 0 & N_\perp \end{bmatrix}$, so setting $X_S = \mathrm{diag}[X_0 \; X]$ gives

$$N_{S\perp} (M_S^* X_S M_S - X_S) N_{S\perp}^* < 0, \tag{8.58}$$

which implies (i) from Lemma 8.1.

■

To conclude this section we develop state-space conditions for the problem of robust performance analysis under constraints in the disturbance, considered in Section 7.2. For convenience, the problem is represented in Figure 8.3, and summarized as follows.

We are given an uncertain input-output system given as an LFT between an LTI map $\mathbf{T}$ and a structured uncertainty operator $\mathbf{\Delta}_u$. The question is to test whether the worst-case l_2 gain from v to y in the presence of uncertainty $\mathbf{\Delta}_u$ is less than β, when the input signal

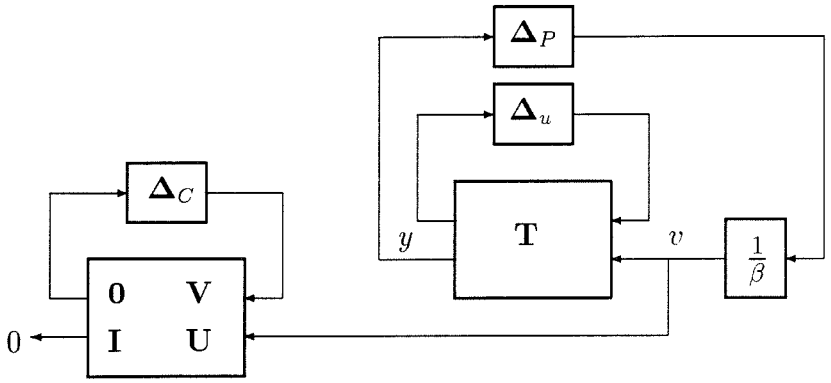


Figure 8.3: Robust performance analysis with disturbance IQCs

v is forced to satisfy a finite number of IQCs represented in implicit form by the equations (7.9), and shown pictorially in the left of Figure 8.3. The performance condition is captured by the block $\mathbf{\Delta}_P$ as in (7.10).

In Section 8.2.3 we reconsidered this problem and showed that if $\mathbf{\Delta}_u$ varied in the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$, it reduced to a convex test (8.28), with the definitions of (8.48). Also, in this case the constraints $\mathbf{N}$ can be chosen to be static. This means that they are in the conditions of Theorem 8.9, and therefore the problem is equivalent to an LMI (8.52) in state-space. To obtain the matrices M_S, N_S we write the state-space realizations

$$T(\lambda) = \left[\begin{array}{c|cc} A_T & B_{T1} & B_{T2} \\ \hline C_{T1} & D_{T11} & D_{T12} \\ C_{T2} & D_{T21} & D_{T22} \end{array}\right], \qquad V(\lambda) = \left[\begin{array}{c|c} A_V & B_V \\ \hline C_V & D_V \end{array}\right].$$

The corresponding equations (8.51) have $\mathbf{\Delta}_s = \mathrm{diag}[\boldsymbol{\lambda}\mathbf{I}, \boldsymbol{\lambda}\mathbf{I}, \mathbf{\Delta}_u, \mathbf{\Delta}_P, \mathbf{\Delta}_C]$,

$$\begin{aligned} M_S &= \left[\begin{array}{ccccc} A_T & 0 & B_{T1} & \frac{1}{\beta}B_{T2} & 0 \\ 0 & A_V & 0 & \frac{1}{\beta}B_V & 0 \\ C_{T1} & 0 & D_{T11} & \frac{1}{\beta}D_{T12} & 0 \\ C_{T2} & 0 & D_{T12} & \frac{1}{\beta}D_{T22} & 0 \\ 0 & C_V & 0 & \frac{1}{\beta}D_V & 0 \end{array}\right], \\ N_S &= \left[\begin{array}{ccccc} 0 & 0 & 0 & \frac{1}{\beta}U & -I \end{array}\right]. \end{aligned} \tag{8.59}$$

Let β_{opt} be the infimum of the values of the parameter β such that LMI (8.52) is feasible; this is a measure of the worst-case gain under $\mathbf{\Delta}_u \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$ and v satisfying the IQCs.

8.4 Application Examples

In this section the implicit analysis methods are applied to two different problems: least squares system identification, and robust performance analysis under white noise.

For both cases there are better solutions available: the standard least squares solution, and the robust $\mathcal{H}_2$ performance conditions of Chapter 6. The objective here is to demonstrate the potential of the implicit analysis framework, which has much more generality and applies in particular to these two problems of very different nature.

8.4.1 The Least Squares Identification Problem

A canonical example of a computationally easy problem in system identification is the least squares problem: in the linear regression setup $y = \mathcal{U}\theta + d$ of (7.12) we wish to find the values of θ which satisfy the equations and minimize the 2 norm of the vector d. Assuming $\mathcal{U}$ is full column rank, this problem has an explicit solution $\hat{\theta} = (\mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^*y$, which gives a minimum $\|d\|^2$ equal to

$$\gamma_0^2 = y^*(I - \mathcal{U}(\mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^*)y. \tag{8.60}$$

We will now reformulate this problem as an implicit analysis question, as was suggested in Section 7.3. The purpose of this is to try the implicit analysis methodology in a simple case, and to show that the simplicity is not entirely hidden in the new formulation.

We first specialize the MV/ID setup of Figure 7.7 to the linear regression case. This is done by following the steps in Section 7.3, and yields the diagram on Figure 8.4.

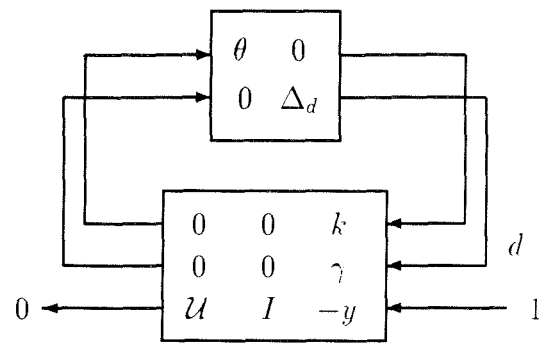


Figure 8.4: The least squares problem.

The only modification to the setup in Figure 7.7 is that scalings k, γ are added to fix the allowable sizes of θ, d respectively. For the least squares problem, we will later on attempt to minimize γ and let $k \to \infty$ to make θ unrestricted.

In reference to the discussion in Section 7.3, the corresponding analysis question $\mathcal{Q}$ is to test whether there exist non-trivial solutions to the equations of Figure 8.4. For $y \neq 0$ (assume $|y| = 1$), the reduction method of Proposition 7.4 applies and produces a constant matrix problem in the canonical form (8.2), with

$$M = \begin{bmatrix} ky^*\mathcal{U} & ky^* \\ \gamma y^*\mathcal{U} & \gamma y^* \end{bmatrix}, \quad N = (I - yy^*)\begin{bmatrix} \mathcal{U} & I \end{bmatrix}. \tag{8.61}$$

There will exist solutions θ and d, with $|\theta| \leq k$, $|d| \leq \gamma$ satisfying the equations of Figure 8.4, if and only if $\mu_\Delta(N, M) \geq 1$. In this case the block structure Δ consists of two real blocks, and M, N are real. This means we are in one of the μ-simple cases of Theorem 8.3, and the question is equivalent to $\hat{\mu}_\Delta(N, M) \geq 1$, which can be tested by the LMI (8.9), for $\beta = 1$. This fact is already indication that our problem remains tractable in the new formulation, although a solution based on LMIs will be less efficient than the least squares solution.

To make the point clearer, we can show explicitly that the LMI approach gives the least squares solution as $k \to \infty$. Consider the LMI (8.9) with $\beta = 1$; the X scaling in this case consists of two scalar parameters, one for each of the full blocks. Let us call the first one $x > 0$, and we can normalize the second one to 1. Some algebra gives $N_\perp^* = \begin{bmatrix} I & 0 \\ -\mathcal{U} & y \end{bmatrix}$,

$$N_\perp(M^*XM - X)N_\perp^* = \begin{bmatrix} -(xI + \mathcal{U}^*\mathcal{U}) & -\mathcal{U}^*y \\ -y^*\mathcal{U} & k^2x + \gamma^2 - 1 \end{bmatrix}. \tag{8.62}$$

A Schur complement operation reduces (8.9) to

$$k^2x + \gamma^2 - 1 + y^*\mathcal{U}(xI + \mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^*y < 0. \tag{8.63}$$

As $k \to \infty$, the unknown x will have to go to zero if (8.63) is to be satisfied; this implies that $(xI + \mathcal{U}^*\mathcal{U})^{-1} \to (\mathcal{U}^*\mathcal{U})^{-1}$ and the LMI will be feasible for large k if and only if

$$\gamma^2 < 1 - y^*\mathcal{U}(\mathcal{U}^*\mathcal{U})^{-1}\mathcal{U}^*y = \gamma_0^2, \tag{8.64}$$

where γ_0 corresponds to (8.60). Recapitulating,

- For $\gamma < \gamma_0$ the LMI is feasible and therefore $\hat{\mu}_\Delta(N, M) = \mu_\Delta(N, M) < 1$, which in turn implies that the answer to $\mathcal{Q}$ is negative (no solutions with $\|d\| \leq \gamma$).
- For $\gamma > \gamma_0$ and large enough k the LMI is not feasible, $\hat{\mu}_\Delta(N, M) = \mu_\Delta(N, M) \geq 1$ and therefore there exist solutions to the MV/ID problem.

So again we find that γ_0 from (8.60) is the minimum norm for d. We have not shown how to solve for θ and d, but this information can also be obtained from the LMI approach.

Of course, we are not advocating this method for a least squares problem; this is a method suitable for a large class of problems. We have simply shown that it remains tractable in this simple case.

8.4.2 Robust Performance with Disturbances in $W_{\gamma,T}$

In Section 6.9 we considered an application of the Robust $\mathcal{H}_2$ performance conditions of Chapter 6 to a problem of rejection of sensor noise, which we reproduce below. In this diagram v is scalar white noise, and we wish to quantify the gain from v to e (sensitivity) in the worst case over the unstructured perturbation $\mathbf{\Delta}_u$.

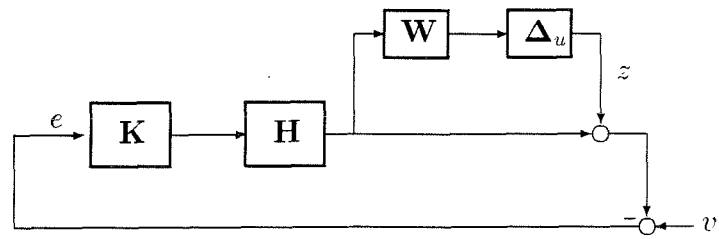


Figure 8.5: Rejection of sensor noise

The conditions of Chapter 6, which were specifically developed for the robust $\mathcal{H}_2$ performance question, are of course the best way to address this problem. We will use, however, this problem to demonstrate the more general tools for analysis under disturbance constraints which were developed in this chapter. In particular, the white noise disturbance will be described by means of the sets $W_{\gamma,T}$ based on autocorrelation constraints. It was shown in Section 5.2 that for scalar noise these are scalar IQCs, therefore they can be converted to the implicit description (7.9)[1]. This means that robust performance analysis under signals in $W_{\gamma,T}$ falls in the general class of problems for which state space conditions were derived in Section 8.3; these conditions correspond to the case of LTV uncertainty.

The procedure for robust performance analysis is the following:

- For a fixed T, consider the IQCs $\sigma_\tau^{\pm}$ of (5.26) (in what follows we choose $\gamma = 0$). Obtain a representation of the form $(\mathbf{U} - \mathbf{\Delta}_c\mathbf{V})v = 0$, with $\mathbf{U}$ constant, and find a state space realization for $V(\lambda)$.

- Using a state space realization for the generalized plant compute the matrices M_S and N_S of (8.59), for a given β.

- Check for feasibility of the LMI (8.52), and minimize over β.

This procedure was applied to the numerical example $K = 2$, $H(\lambda) = \frac{1}{1-3.3\lambda+\lambda^2}$, and $W = 0.25$. These values were chosen so that the uncertainty affects the sensitivity in a significant way; this is exhibited in Figure 8.6, where the lower curve indicates the nominal sensitivity

[1] For multivariable noise, an analogous description with δI blocks is available, see [59].

function S, and the upper curve is the worst-case sensitivity function S_{max} obtained in (6.73). The values

$$\|S_{max}\|_\infty = 6.50, \qquad \|S_{max}\|_2 = 2.39, \tag{8.65}$$

correspond to the worst-case $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms of the system under (possibly non-causal) LTI uncertainty, as shown in Section 6.9.

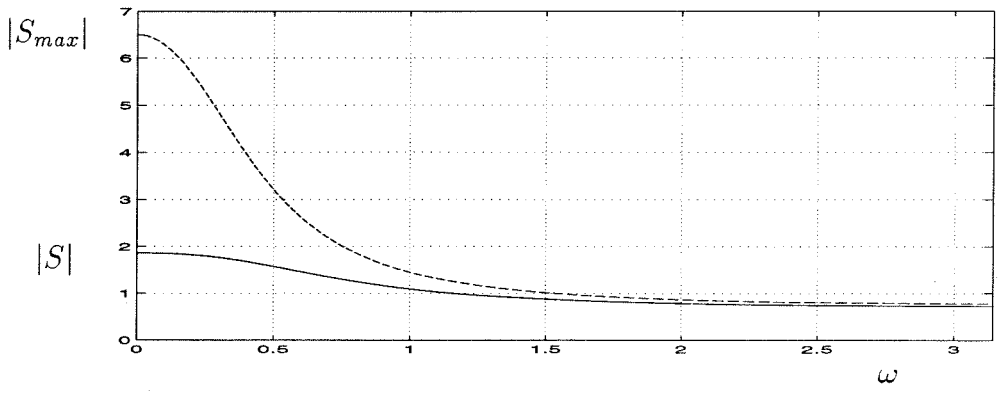


Figure 8.6: Nominal and worst-case sensitivity functions (magnitude)

We now apply the procedure described above for analysis over $W_{0,T}$ and obtain the value β_{opt} as a function of T. Asymptotically, as the number of constraints increases, β_{opt} converges down to a robust $\mathcal{H}_2$ performance measure over the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$. The plot of $\beta_{opt}(T)$ (obtained using the LMI Control Toolbox [35]) is given in Figure 8.7.

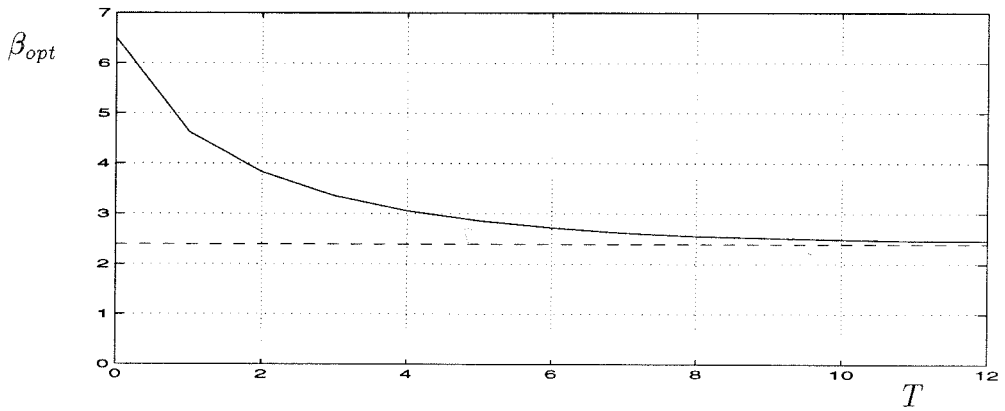


Figure 8.7: Worst-case induced norm of uncertain system over $W_{0,T}$

For $T = 0$ (no constraints) β_{opt} corresponds to the worst-case $\mathcal{H}_\infty$ norm under $\boldsymbol{\Delta}_u \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$. As T increases, we approach the worst-case $\mathcal{H}_2$ norm. Both values appear to coincide in this example with those in (8.65) for LTI uncertainty. This is not, however, a general fact; other examples (see [59]) exhibit a gap between LTI and LTV uncertainty.

Chapter 9

Concluding Remarks

To conclude, we first summarize the contributions of this thesis:

- In Chapter 3 we provided non-conservative tests for the evaluation of robust stability and $\mathcal{H}_\infty$ performance in the case of combined time varying, time invariant and parametric uncertainty structures. These are structured singular value conditions obtained by frequency augmentation. We also analyzed the convex upper bounds for these combined problems.

- The methodology of Chapter 4 for white noise modeling based on sets provides a useful technique for situations involving a combination of "hard bounds" and white noise, by describing the latter in terms of a typical set. We have shown its potential for questions of white noise rejection and worst-case system identification, but this line of thinking may prove useful in other engineering contexts.

- The most significant result is Condition 2 in Chapter 6, for analysis of Robust $\mathcal{H}_2$ performance. This result to some extent completes the picture of linear robust control theory, since we now have convex conditions for the evaluation of robust performance with all the disturbance rejection criteria ($\mathcal{H}_2$, $\mathcal{H}_\infty$, $\mathcal{L}_1$). In fact, as remarked in Chapter 6, Condition 2 provides a summary of the analysis problems involving l_2 signal norms which lead to a convex characterization, including $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance, or combinations thereof.

- The implicit analysis framework developed in Chapters 7 and 8, connects robustness analysis more directly with first principles models which take naturally an implicit form. It has also been shown to be a flexible tool for the consideration of various analysis questions involving equations subject to uncertainty, and signal constraints which can be reparametrized in this way. In particular, it provides a common language for the formulation of robustness analysis and system identification questions.

The following are some directions for future research:

- Regarding the results of Chapter 6, we remark that although the theory was developed for discrete-time systems, Condition 2 extends verbatim to the continuous time case, with analogous results. This extension requires a generalization of the set descriptions of white noise for continuous time, and of the notions of slowly varying uncertainty as suggested in [64]; the details will be reported in the future.

- A future research question is the study of state-space methods for the robust $\mathcal{H}_2$ problem. In particular, in Section 6.7 we remarked that that the mixed performance problem admits a state space evaluation. This may lead to a single LMI test for Condition 2 in the case of constant scales, as will be investigated in the future. Also, controller synthesis for this version of the mixed problem has not been considered and would be instrumental in developing iteration schemes for robust $\mathcal{H}_2$ synthesis.

- A reasonable follow-up on the results of this thesis is a complete development of computational tools for these convex conditions, at the level of existing software tools [5] for the $\mathcal{H}_\infty$ performance measure. Once these are in place, the theory can be tried out in engineering problems and develop experience on the judicious use of these tools.

- Other problems have led to μ conditions which are in general NP-hard; the most important example of this are the problems reduced to implicit μ, which includes the model validation question of Section 7.3. The possibility of the implicit formulation to provide a practical (and not only conceptual) unification of tools for robustness analysis and identification, rests on the development of heuristic algorithms, such as the μ lower bound, which should be able to compute the answer in "typical" instances of the problem. These issues must be pursued at the computational level.

Many signs appear to indicate that linear robust control is a mature field, conceptually well developed, and where the main theoretical results have been obtained. This thesis has contributed to complete the picture in robustness analysis. The robust synthesis problem still does not have a global solution for any of the criteria, but is likely to remain this way.

One of the fundamental open challenges lies in the extensions of the robustness theory to the case of nonlinear systems, which by its nature will involve very different tools from those described in this thesis. Perhaps a lesson can be carried through, however, to this open area, namely that these problems should be approached by a creative combination, without rigid boundaries, between diverse mathematical techniques.

Appendix A

Mathematical Complements

The following mathematical facts are collected here for ease of reference. This material can be found, for example, in [47], [77], [70].

A.1 Bounded Variation Functions, Stieltjes Integrals and the Riesz Theorem

First we introduce the space $BV[a,b]$ of real-valued functions of bounded variation in the interval $[a,b] \subset \mathbb{R}$. A function $\Psi(t)$ is of bounded variation if

$$TV(\Psi) := \sup \sum_{i=1}^{N} |\Psi(t_i) - \Psi(t_{i-1})| < \infty, \tag{A.1}$$

where the supremum is taken over partitions $(t_0, \ldots, t_N)$ of $[a,b]$. $TV(\Psi)$ is called the total variation of Ψ.

We will also use the space $C_R[a,b]$ of continuous, real-valued functions on $[a,b]$, with the norm $\|g\|_\infty := \sup_{t\in[a,b]} |g(t)|$.

Given $\Psi \in BV[a,b]$ and $g \in C_R[a,b]$, we introduce the Stieltjes integral

$$\int_a^b g(t)d\Psi(t). \tag{A.2}$$

In the case considered, which is the only one needed in this thesis, the integral (A.2) can be defined (see [77]) as a limit of Riemann sums analogously to standard calculus. A more abstract version is to consider Ψ to be the distribution function of a regular, signed measure μ on $[a,b]$, $\Psi(t) = \mu([a,t])$, and interpret (A.2) as a Lebesgue integral.

A basic property is that

$$\left| \int_a^b g(t)d\Psi(t) \right| \leq \|g\|_\infty TV(\Psi). \tag{A.3}$$

This bound implies that the map $\Gamma_\Psi : C_R[a,b] \to \mathbb{R}$ given by

$$\Gamma_\Psi(g) = \int_a^b g(t)d\Psi(t) \tag{A.4}$$

defines a bounded linear functional on the space $\mathbb{C}_R[a,b]$. The Riesz representation theorem states that *every* functional in the dual space $C_R[a,b]^*$ is of this form.

We will also use the formula of integration by parts for the Stieltjes integral,

$$\int_a^b g(t)d\Psi(t) = g(b)\Psi(b) - g(a)\Psi(a) - \int_a^b \Psi(t)dg(t), \tag{A.5}$$

which holds, for example, for $\Psi \in BV[a,b]$, $g \in C_R[a,b]$. Furthermore, if g has an integrable derivative $g'(t)$, the integral on the right can be written as $\int_a^b \Psi(t)g'(t)dt$.

A.2 Convex Analysis Results

Given a set $\mathcal{K}$ in a vector space $\mathbb{V}$, its convex hull is defined as the set

$$\mathrm{co}(\mathcal{K}) = \left\{ \sum_{k=1}^{N} \alpha_k v_k : \; N \in \mathbb{Z}^+, \alpha_k \geq 0, \sum_{k=1}^{N} \alpha_k = 1, \; v_k \in \mathcal{K} \right\} \tag{A.6}$$

of convex combinations of elements in $\mathcal{K}$. The set $\mathcal{K}$ is convex if $\mathrm{co}(\mathcal{K}) = \mathcal{K}$.

We first state some lemmas from convex analysis in finite dimensional space.

Lemma A.1 (Helly) *Let $\{\mathcal{K}_\omega\}_{\omega\in\Omega}$ be family of convex closed sets in $\mathbb{R}^d$, of which at least one is bounded. If $\bigcap_{\omega\in\Omega} \mathcal{K}_\omega = \emptyset$ then there exist $d+1$ sets $\mathcal{K}_{\omega_i}, i = 1 \ldots d+1$ with empty intersection.*

Lemma A.2 *If $\mathcal{K} \subset \mathbb{R}^d$, every point in $\mathrm{co}(\mathcal{K})$ is a convex combination of $d+1$ points in $\mathcal{K}$; for $\mathcal{K}$ compact, every point in the boundary of $\mathrm{co}(\mathcal{K})$ is a convex combination of d points in $\mathcal{K}$.*

Lemma A.3 *Let $\mathcal{K}_1$,$\mathcal{K}_2$ be disjoint convex sets in $\mathbb{R}^d$. where $\mathcal{K}_1$ is compact and $\mathcal{K}_2$ is closed. Then there exists a vector $x \in \mathbb{R}^d$, and α, β in $\mathbb{R}$ such that*

$$\langle x, k_1 \rangle \leq \alpha < \beta \leq \langle x, k_2 \rangle \;\; \forall k_1 \in \mathcal{K}_1, \; k_2 \in \mathcal{K}_2. \tag{A.7}$$

Remarks on the Proof: Lemmas A.1 and A.3 can be found in [70]. The first part of Lemma A.2 is a classical result of Caratheodory (see [70]), and implies that for every $v \in \text{co}(\mathcal{K})$, there exists a *simplex* of the form

$$S(v_1, \ldots, v_{d+1}) = \left\{ \sum_{k=1}^{d+1} \alpha_k v_k : \alpha_k \geq 0, \sum_{k=1}^{d+1} \alpha_k = 1 \right\} \tag{A.8}$$

with vertices $v_k \in \mathcal{K}$, which contains v.

The refinement for points in the boundary can be shown as follows. First note that if $\mathcal{K} \subset \mathbb{R}^d$ is compact, $\text{co}(\mathcal{K})$ is compact so it contains its boundary. For a point v in this boundary, pick the corresponding simplex from (A.8). If the v_k are in a lower dimensional hyperplane, then d points will suffice to generate v. If not, then every point in $S(v_1, \ldots, v_{d+1})$ corresponding to $\alpha_k > 0 \ \forall k$ will be interior to $S(v_1, \ldots, v_{d+1}) \subset \text{co}(\mathcal{K})$. Since v is in the boundary of $\text{co}(\mathcal{K})$, one of the α_k's must be 0 and a convex combination of d points will suffice, completing Lemma A.2.

■

In regard to the hyperplane separation result of Lemma A.3, we emphasize the *strict* separation obtained in this finite dimensional version, which is used in several of the proofs.

We also use an infinite dimensional hyperplane separation result, which is a geometric version of the Hahn-Banach theorem, taken from [47]:

Theorem A.4 *Let $\mathcal{K}_1$, $\mathcal{K}_2$ be convex sets in a real normed space $\mathbb{V}$, such that $\mathcal{K}_2$ has non-empty interior, and $\mathcal{K}_1$ contains no interior points of $\mathcal{K}_2$. Then there exists a bounded functional $\Gamma \in \mathbb{V}^*$, $\Gamma \neq 0$, and a real number α such that*

$$\Gamma(k_1) \leq \alpha \leq \Gamma(k_2), \quad \textit{for all } k_1 \in \mathcal{K}_1, \ k_2 \in \mathcal{K}_2. \tag{A.9}$$

In this case the separation in non-strict; note also the technical assumption on non-empty interior for the set $\mathcal{K}_2$.

Appendix B

Destabilizing Perturbations and the Issue of Causality

In this chapter we provide the proofs of some technical results in the thesis which require a construction of a perturbation which destabilizes, or violates robust performance. In particular, we pay attention to the possibility of achieving this with a *causal* perturbation, an issue which is non-trivial since many of the results are based on IQCs, which as shown in Chapter 5 correspond directly to non-causal operators.

Before embarking in the very technical material which follows, it is useful to discuss the reasons for the causality requirement.

The most straightforward answer is that choosing a causal $\mathbf{\Delta}$ in a well-posed LFT interconnection ensures the causality of the input-output model, which is a requirement in order for these to be physically realizable. Note, however, that there are many other considerations which make a model "realistic", other than causality; in fact the perturbations $\mathbf{\Delta}$ constructed in this Appendix are quite complicated and involve long memory shifts, aspects which do not arise naturally in physical models.

Why, then, insist on causality of $\mathbf{\Delta}$? The most important reason is that uncertainty descriptions are used to assess questions of stability. These are only meaningful when there is a chosen direction in time, and signals are not known a priori to have bounded norm. Causal operators are the only mathematical setting to accommodate time directionality and the extended spaces l_{2e} of unbounded signals. Also, the LFT formulation relies on "pulling out the $\mathbf{\Delta}$" and treating it as a system. Although one may choose to ignore other considerations which make a model realistic and cover them by a larger set, causality of $\mathbf{\Delta}$ appears to be a basic restriction for a sound theory of robust stability in the LFT framework.

The material in this Appendix should be viewed, accordingly, as completing a consistent mathematical framework for the results based on the IQC formulation.

We will show that for the classes of uncertainty which correspond directly to IQCs, such as $\boldsymbol{\Delta}^{LTV}$ and $\boldsymbol{\Delta}^{\nu}$, the robust performance tests remain necessary under the causality restriction: if robust performance can be violated with non-causal $\boldsymbol{\Delta}$, it also fails under causal $\boldsymbol{\Delta}$. The development is based on a construction due to Shamma [76], which is extended to a larger class of problems.

For simplicity we will work in the space $l_2(\mathbb{Z})$ with bidirectional time, where the shift operator $\boldsymbol{\lambda}$ is unitary, $\boldsymbol{\lambda}^{-1}$ being the backward shift. An interval of integers is denoted by $[t_1, t_2) = \{t \in \mathbb{Z} : t_1 \leq t < t_2\}$; $\mathbf{P}^{[t_1,t_2)}$ is the projection operator into the time interval $[t_1, t_2)$.

Sections B.1 and B.2 introduce some basic tools. These are applied to the issue of uniform robust stability in Section B.3, then to the proofs of Chapter 6 in Section B.4, and to the proofs in Chapter 8 in Section B.5.

B.1 Operators with Finite Support

An operator $\boldsymbol{\Delta} \in \mathcal{L}(l_2)$ is supported on the interval $[t_1, t_2)$ if

$$\boldsymbol{\Delta} = \mathbf{P}^{[t_1,t_2)} \boldsymbol{\Delta} \mathbf{P}^{[t_1,t_2)}. \tag{B.1}$$

Equivalently, the infinite matrix representation as in (2.12) is only nonzero for entries (k, t) such that $k, t \in [t_1, t_2)$. The operator is therefore effectively represented by a finite matrix on that interval. No causality assumptions are made in this section.

Theorem B.1 *Given an operator* $\mathbf{M} \in \mathcal{L}(l_2)$, *and* $\epsilon > 0$, *suppose there exists* $z \in l_2$, $\|z\|_2 = 1$, *and* $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}_{NC}^{LTV}}$ *such that* $\|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})z\|_2 < \epsilon$.

Then there exists an operator $\bar{\boldsymbol{\Delta}} \in \boldsymbol{B}_{\boldsymbol{\Delta}_{NC}^{LTV}}$ *supported on a finite interval* $[t_1, t_2)$, *and* $\bar{z}$, *supported on* $[t_1, t_2)$, *such that* $\|\bar{z}\|_2 = 1$, $\|(\mathbf{I} - \bar{\boldsymbol{\Delta}}\mathbf{M})\bar{z}\|_2 < \epsilon$. *Furthermore:*

1. *If* $\|\boldsymbol{\lambda}\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_i\boldsymbol{\lambda}\| < \nu$ *for some block* $\boldsymbol{\Delta}_i$ *in* $\boldsymbol{\Delta}$, *then the corresponding* $\bar{\boldsymbol{\Delta}}_i$ *can be chosen to satisfy* $\|\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i - \bar{\boldsymbol{\Delta}}_i\boldsymbol{\lambda}\| < \nu$.
2. *If* $\|\mathbf{N}z\|_2 < \epsilon$ *for* $\mathbf{N} \in \mathcal{L}(l_2)$, *then* $\bar{z}$ *can be chosen to satisfy* $\|\mathbf{N}\bar{z}\|_2 < \epsilon$.
3. *If a portion* z_n *of the signal* z *is in* $\hat{W}_\eta^{rm}$, *then* $\bar{z}_n$ *can be chosen in* $\hat{W}_\eta^{rm}$.

Proof: The natural approach is to proceed by truncation of z and $\boldsymbol{\Delta}$ to a finite interval. This will work except that the truncation does not preserve the $\boldsymbol{B}^{\nu}$ property (heuristically, it imposes an abrupt time-variation). To obtain a truncation which preserves slow time variation, one must multiply signals by a time-varying gain which slowly decreases to zero. For each T, consider a real-valued function $a^T(t)$ such that

- a^T has finite support.
- $a^T(t) = 1$ for $t \in [-T, T]$, and $|a^T(t)| \leq 1 \; \forall \, t \in \mathbb{Z}$.
- $\sup_t |a^T(t+1) - a^T(t)| = \alpha$.

Introduce the operator $\mathbf{A}^T : l_2 \to l_2$ of multiplication by a^T, i.e. $\mathbf{A}^T : v(t) \mapsto a^T(t)v(t)$. Then $\|\mathbf{A}^T\| = 1$, and

$$[(\boldsymbol{\lambda}\mathbf{A}^T - \mathbf{A}^T\boldsymbol{\lambda})v](t) = \left(a^T(t-1) - a^T(t)\right) v(t-1), \tag{B.2}$$

which implies that $\|\boldsymbol{\lambda}\mathbf{A}^T - \mathbf{A}^T\boldsymbol{\lambda}\| = \sup_t |a^T(t+1) - a^T(t)| = \alpha$.

Now define $\bar{\boldsymbol{\Delta}}^T = \mathbf{A}^T\boldsymbol{\Delta}\mathbf{A}^T$, of finite support. The identity

$$\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i^T - \bar{\boldsymbol{\Delta}}_i^T\boldsymbol{\lambda} = (\boldsymbol{\lambda}\mathbf{A}^T - \mathbf{A}^T\boldsymbol{\lambda})\boldsymbol{\Delta}_i\mathbf{A}^T + \mathbf{A}^T(\boldsymbol{\lambda}\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_i\boldsymbol{\lambda})\mathbf{A}^T + \mathbf{A}^T\boldsymbol{\Delta}_i(\boldsymbol{\lambda}\mathbf{A}^T - \mathbf{A}^T\boldsymbol{\lambda})$$

which holds for any sub-block, gives

$$\|\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i^T - \bar{\boldsymbol{\Delta}}_i^T\boldsymbol{\lambda}\| \leq \|\boldsymbol{\lambda}\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_i\boldsymbol{\lambda}\| + 2\alpha. \tag{B.3}$$

Hence $\bar{\boldsymbol{\Delta}}^T$ provides a way to truncate $\boldsymbol{\Delta}$ at a moderate cost in terms of time variation.

It follows directly from the definition of $\mathbf{A}^T$ and $\bar{\boldsymbol{\Delta}}^T$ that

$$\mathbf{A}^T z \overset{T\to\infty}{\longrightarrow} z, \qquad \bar{\boldsymbol{\Delta}}^T\mathbf{M}\mathbf{A}^T z \overset{T\to\infty}{\longrightarrow} \boldsymbol{\Delta}\mathbf{M}z, \tag{B.4}$$

with convergence in the sense of l_2. Therefore $\|\mathbf{A}^T z\|_2 \overset{T\to\infty}{\longrightarrow} 1$, and

$$\left\|\left(\mathbf{I} - \bar{\boldsymbol{\Delta}}^T\mathbf{M}\right)\mathbf{A}^T z\right\|_2 \overset{T\to\infty}{\longrightarrow} \|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})z\|_2 < \epsilon. \tag{B.5}$$

This means that the condition $\|(\mathbf{I} - \bar{\boldsymbol{\Delta}}\mathbf{M})\bar{z}\|_2 < \epsilon$ will be satisfied if we choose the finitely supported $\bar{\boldsymbol{\Delta}} = \bar{\boldsymbol{\Delta}}^T$ and $\bar{z} = \bar{z}^T := \frac{\mathbf{A}^T z}{\|\mathbf{A}^T z\|_2}$, for large enough T.

We now show that the remaining properties can also be satisfied.

1. If $\|\boldsymbol{\lambda}\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_i\boldsymbol{\lambda}\| < \nu$, then for small enough α we have $\|\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i - \bar{\boldsymbol{\Delta}}_i\boldsymbol{\lambda}\| < \nu$ from (B.3).
2. Since $\bar{z}^T \overset{T\to\infty}{\longrightarrow} z$ in l_2, then $\|\mathbf{N}\bar{z}^T\|_2 < \epsilon$ for large enough T.
3. Since $\bar{z}_n^T \overset{T\to\infty}{\longrightarrow} z_n$ in l_2, the third property will follow if we can show that whenever $v^T \overset{T\to\infty}{\longrightarrow} v$ in l_2^m, $F_{v^T} \overset{T\to\infty}{\longrightarrow} F_v$ in the sense of uniform convergence of functions, where

$$F_v(s) := \int_0^s v(\omega)v(\omega)^* \frac{dw}{2\pi} - \frac{s}{2\pi}\frac{\|v\|_2^2}{m} I_m, \tag{B.6}$$

as defined in (4.66). We now prove it for the scalar case $m = 1$, writing

$$[F_{v^T} - F_v](s) = \left\langle \mathbf{H}_s v^T, v^T \right\rangle - \left\langle \mathbf{H}_s v, v \right\rangle + \frac{s}{2\pi} \left(\|v\|_2^2 - \|v^T\|_2^2 \right), \tag{B.7}$$

where $\mathbf{H}_s$ is the ideal filter $H_s(e^{j\omega}) = 1_{[0,s]}$. The last term in (B.7) converges uniformly to 0, and the first two terms can be bounded by writing

$$\left| \left\langle \mathbf{H}_s v^T, v^T - v \right\rangle + \left\langle \mathbf{H}_s (v^T - v), v \right\rangle \right| \leq \|v^T\|_2 \|v^T - v\|_2 + \|v^T - v\|_2 \|v\|_2$$

which converges to zero. ■

B.2 Causal Operators

The previous result provides a general truncation scheme, but still does not address causality. To consider this question, we restrict the attention to causal LTI systems $\mathbf{M}$.

Theorem B.2 *Given a causal,* LTI *system* $\mathbf{M} \in \mathcal{L}_c(l_2)$, *and* $\epsilon > 0$, *suppose there exists* $z \in l_2$, $\|z\|_2 = 1$, *and* $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}_{\boldsymbol{NC}}^{\boldsymbol{LTV}}}$ *such that* $\|(\mathbf{I} - \mathbf{\Delta M})z\|_2 < \epsilon$.

Then there exists a causal operator $\hat{\mathbf{\Delta}} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\boldsymbol{LTV}}}$ *supported on a finite interval* $[0, T)$, *and* $\hat{z}$, *supported on* $[0, T)$, *such that* $\|\hat{z}\|_2 = 1$, $\|(\mathbf{I} - \hat{\mathbf{\Delta}}\mathbf{M})\hat{z}\|_2 < \epsilon$. *Furthermore:*

1. *If* $\|\boldsymbol{\lambda}\mathbf{\Delta}_i - \mathbf{\Delta}_i\boldsymbol{\lambda}\| < \nu$ *for some block* $\mathbf{\Delta}_i$ *in* $\mathbf{\Delta}$, *then the corresponding* $\hat{\mathbf{\Delta}}_i$ *can be chosen to satisfy* $\|\boldsymbol{\lambda}\hat{\mathbf{\Delta}}_i - \hat{\mathbf{\Delta}}_i\boldsymbol{\lambda}\| < \nu$.
2. *If* $\|\mathbf{N}z\|_2 < \epsilon$ *for an* LTI *system* $\mathbf{N} \in \mathcal{L}(l_2)$, *then* $\hat{z}$ *can be chosen to satisfy* $\|\mathbf{N}\hat{z}\|_2 < \epsilon$.
3. *If a portion* z_n *of the signal* z *is in* $\hat{W}_\eta^m$, *then* $\hat{z}_n$ *can be chosen in* $\hat{W}_\eta^m$.

Proof:

The first step is to apply Theorem B.1 and obtain the finite support $\bar{\mathbf{\Delta}}$ and $\bar{z}$; since $\mathbf{M}$ is LTI we can assume the support is in $[0, t_0)$. Now we choose an integer N such that

$$\frac{1}{N} + \|(I - \bar{\mathbf{\Delta}}\mathbf{M})\bar{z}\|_2^2 < \epsilon^2. \tag{B.8}$$

For $k > t_0$, define

$$\hat{z}^k := \frac{1}{N^{\frac{1}{2}}} \sum_{r=0}^{N-1} \boldsymbol{\lambda}^{kr} \bar{z}, \tag{B.9}$$

$$\hat{\mathbf{\Delta}}^k := \sum_{r=1}^{N-1} \boldsymbol{\lambda}^{kr} \bar{\mathbf{\Delta}} \boldsymbol{\lambda}^{-k(r-1)}. \tag{B.10}$$

The signal $\hat{z}^k$ is obtained by adding N relatively shifted copies of $\bar{z}$, in a similar fashion to (5.42), with a normalizing factor to obtain $\|\hat{z}^k\| = 1$.

The definition of $\hat{\mathbf{\Delta}}^k$ is best interpreted with the infinite matrix representation of Figure B.1. The matrix $\bar{\mathbf{\Delta}}_k$ is the truncation of the infinite matrix for $\bar{\mathbf{\Delta}}$, to the interval $[0, k)$ which contains its support. Thus $\hat{\mathbf{\Delta}}^k$ is made of $N-1$ shifted copies $\boldsymbol{\lambda}^{k(r-1)}\bar{\mathbf{\Delta}}\boldsymbol{\lambda}^{-k(r-1)}$ of $\bar{\mathbf{\Delta}}$, with an extra shift $\boldsymbol{\lambda}^k$ ensuring that the result is causal (lower triangular matrix).

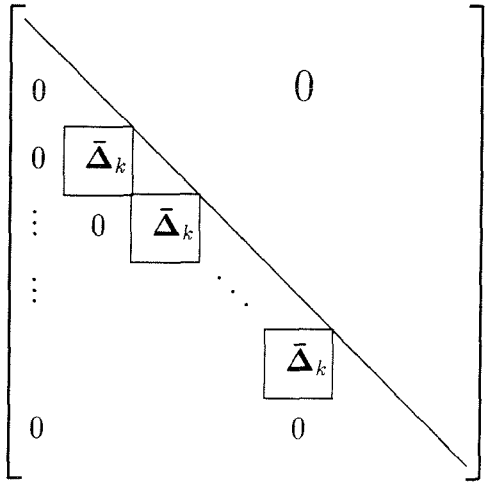


Figure B.1: Infinite matrix representation of the operator $\hat{\mathbf{\Delta}}^k$

We will now show that for sufficiently large k, $\hat{\mathbf{\Delta}}^k$ and $\hat{z}^k$ satisfy the required conditions. We start by computing

$$\hat{\mathbf{\Delta}}^k\,\mathbf{M}\,\hat{z}^k = \frac{1}{N^{\frac{1}{2}}}\sum_{r=1}^{N-1}\sum_{l=0}^{N-1}\boldsymbol{\lambda}^{kr}\bar{\mathbf{\Delta}}\boldsymbol{\lambda}^{-k(r-1)}\boldsymbol{\lambda}^{kl}\mathbf{M}\bar{z} = \frac{1}{N^{\frac{1}{2}}}\sum_{l=0}^{N-2}\sum_{r=l+1}^{N-1}\boldsymbol{\lambda}^{kr}\bar{\mathbf{\Delta}}\boldsymbol{\lambda}^{k(l+1-r)}\mathbf{M}\bar{z}. \quad \text{(B.11)}$$

where the first equality uses the time invariance of $\mathbf{M}$. For the second equality, observe that $\mathbf{M}\bar{z}$ is supported on $[0,\infty)$, therefore for $r \le l$, $\boldsymbol{\lambda}^{k(l+1-r)}\mathbf{M}\bar{z}$ is supported on $[k,\infty)$, and $\bar{\mathbf{\Delta}}$ in $[0, t_0)$ which implies $\bar{\mathbf{\Delta}}\boldsymbol{\lambda}^{k(l+1-r)}\mathbf{M}\bar{z} = 0$ for these terms. We now rewrite (B.11) as

$$\hat{\mathbf{\Delta}}^k\,\mathbf{M}\,\hat{z}^k = \frac{1}{N^{\frac{1}{2}}}\left(\sum_{l=0}^{N-2}\boldsymbol{\lambda}^{k(l+1)}\bar{\mathbf{\Delta}}\mathbf{M}\bar{z} + \sum_{l=0}^{N-3}\sum_{r=l+2}^{N-1}\boldsymbol{\lambda}^{kr}\bar{\mathbf{\Delta}}\boldsymbol{\lambda}^{k(l+1-r)}\mathbf{M}\bar{z}\right). \quad \text{(B.12)}$$

Denote the second sum in (B.12) by S_k. For each term in this sum, $\boldsymbol{\lambda}^{k(l+1-r)}$ is a backward shift of at least k. Since $\bar{\mathbf{\Delta}}$ is supported on $t \ge 0$, it only operates effectively on the "tail" signal $\mathbf{P}^{[k,\infty)}\mathbf{M}\bar{z}$, which converges to 0 as $k \to \infty$. Since the number of terms is fixed, we conclude that $S_k \overset{k\to\infty}{\longrightarrow} 0$ in l_2. Now (B.12) leads to

$$(\mathbf{I} - \hat{\mathbf{\Delta}}^k\,\mathbf{M})\hat{z}^k = \frac{1}{N^{\frac{1}{2}}}\left(\sum_{r=0}^{N-1}\boldsymbol{\lambda}^{kr}\bar{z} - \sum_{l=0}^{N-2}\boldsymbol{\lambda}^{k(l+1)}\bar{\mathbf{\Delta}}\mathbf{M}\bar{z} - S_k\right) \quad \text{(B.13)}$$

$$= \frac{1}{N^{\frac{1}{2}}}\left(\bar{z} + \sum_{r=1}^{N-1} \boldsymbol{\lambda}^{kr}(\mathbf{I} - \bar{\boldsymbol{\Delta}}\mathbf{M})\bar{z} - S_k\right). \tag{B.14}$$

The support of $\bar{\boldsymbol{\Delta}}$ and $\bar{z}$ in $[0, k)$ implies that if we leave out S_k, the remaining terms in (B.14) are mutually orthogonal, so they have collectively a norm

$$\sqrt{\frac{1}{N} + \frac{N-1}{N}\|(\mathbf{I} - \bar{\boldsymbol{\Delta}}\mathbf{M})\bar{z}\|_2^2}$$

which is less than ϵ because of (B.8). Since $S_k \overset{k\to\infty}{\longrightarrow} 0$, we now that for sufficiently large k,

$$\|(\mathbf{I} - \hat{\boldsymbol{\Delta}}^k\,\mathbf{M})\hat{z}^k\| < \epsilon. \tag{B.15}$$

It remains to show that the additional properties stated in Theorem B.1 can also be satisfied. Properties 2 and 3 follow immediately from the results in Chapter 5; by construction of $\hat{z}^k$, we find from (5.44) that

$$\|\mathbf{N}\hat{z}^k\|_2^2 \overset{k\to\infty}{\longrightarrow} \|\mathbf{N}\bar{z}\|_2^2, \tag{B.16}$$

$$\rho(\hat{z}_n^k) \overset{k\to\infty}{\longrightarrow} \rho(\bar{z}_n), \tag{B.17}$$

where $\rho(\cdot)$ is defined in (5.30). Finally, for the i-th block we write

$$\boldsymbol{\lambda}\hat{\boldsymbol{\Delta}}_i^k - \hat{\boldsymbol{\Delta}}_i^k\boldsymbol{\lambda} = \sum_{r=1}^{N-1} \boldsymbol{\lambda}^{kr}\left(\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i - \bar{\boldsymbol{\Delta}}_i\boldsymbol{\lambda}\right)\boldsymbol{\lambda}^{-k(r-1)}. \tag{B.18}$$

Denote $\boldsymbol{\Upsilon}_i = \boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i - \bar{\boldsymbol{\Delta}}_i\boldsymbol{\lambda}$; this operator has its support in $[0.\, t_0+1) \subset [0, k)$, which means that $\boldsymbol{\Upsilon}_i = \mathbf{P}^{[0,k)}\boldsymbol{\Upsilon}_i\mathbf{P}^{[0,k)}$. Then for $v \in l_2$,

$$(\boldsymbol{\lambda}\hat{\boldsymbol{\Delta}}_i^k - \hat{\boldsymbol{\Delta}}_i^k\boldsymbol{\lambda})v = \sum_{r=1}^{N-1} \boldsymbol{\lambda}^{kr}\mathbf{P}^{[0,k)}\boldsymbol{\Upsilon}_i\mathbf{P}^{[0,k)}\boldsymbol{\lambda}^{-k(r-1)}v = \sum_{r=1}^{N-1} \mathbf{P}^{[kr,k(r+1))}\boldsymbol{\lambda}^{kr}\boldsymbol{\Upsilon}_i\boldsymbol{\lambda}^{-k(r-1)}\mathbf{P}^{[k(r-1),kr)}v. \tag{B.19}$$

Due to the orthogonality of the projections on disjoint intervals. we have

$$\begin{aligned}\left\|(\boldsymbol{\lambda}\hat{\boldsymbol{\Delta}}_i^k - \hat{\boldsymbol{\Delta}}_i^k\boldsymbol{\lambda})v\right\|_2^2 &= \left\|\sum_{r=1}^{N-1} \mathbf{P}^{[kr,k(r+1))}\boldsymbol{\lambda}^{kr}\boldsymbol{\Upsilon}_i\boldsymbol{\lambda}^{-k(r-1)}\mathbf{P}^{[k(r-1),kr)}v\right\|_2^2 \\ &= \sum_{r=1}^{N-1} \|\boldsymbol{\Upsilon}_i\boldsymbol{\lambda}^{-k(r-1)}\mathbf{P}^{[k(r-1),kr)}v\|_2^2 \\ &\leq \|\boldsymbol{\Upsilon}_i\|^2 \sum_{r=1}^{N-1} \|\mathbf{P}^{[k(r-1),kr)}v\|_2^2 \leq \|\boldsymbol{\Upsilon}_i\|^2\|v\|_2^2. \end{aligned} \tag{B.20}$$

We conclude that $\|\boldsymbol{\lambda}\hat{\boldsymbol{\Delta}}_i^k - \hat{\boldsymbol{\Delta}}_i^k\boldsymbol{\lambda}\| \leq \|\boldsymbol{\Upsilon}_i\| = \|\boldsymbol{\lambda}\bar{\boldsymbol{\Delta}}_i - \bar{\boldsymbol{\Delta}}_i\boldsymbol{\lambda}\|$. so if a prescribed rate of variation ν holds for $\bar{\boldsymbol{\Delta}}_i$, the same holds for $\hat{\boldsymbol{\Delta}}_i^k$.

■

B.3 A Uniformity Property for Robust Stability

In Definition 2 we introduced the concept of uniform robust stability, which slightly reinforces the definition of robust stability. In reference to an $\mathbf{M}$-$\mathbf{\Delta}$ configuration such as Figure 2.9, with $\mathbf{M}$, $\mathbf{\Delta}$ in $\mathcal{L}_c(l_2)$, this property means that in addition to $\mathbf{I}-\mathbf{\Delta M}$ being invertible for each $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}}$, there is a uniform bound on the inverse norm across $\boldsymbol{B}_{\boldsymbol{\Delta}}$.

For robust stability problems in the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTI}}$, it follows from μ-analysis that this uniformity requirement is automatically guaranteed. In this section we show that the same happens for the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$, and essentially also for $\boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}}$.

We start with the following result, which is stated in the more general setting of the implicit formulation of Chapter 8.

Theorem B.3 *Given* LTI *systems* $\mathbf{M}, \mathbf{N} \in \mathcal{L}(l_2)$, *suppose there exists a sequence of operators* $\mathbf{\Delta}^k \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}_{NC}}$, *and a sequence of signals* $z^k \in l_2$, $\|z^k\| = 1$, *such that*

$$\left\| \begin{bmatrix} \mathbf{I}-\mathbf{\Delta}^k\mathbf{M} \\ \mathbf{N} \end{bmatrix} z^k \right\| \xrightarrow{k\to\infty} 0. \tag{B.21}$$

Then there exists a unique $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}_{NC}}$, *and a sequence of signals* $\tilde{z}^k \in l_2$ *such that*

$$\left\| \begin{bmatrix} \mathbf{I}-\mathbf{\Delta M} \\ \mathbf{N} \end{bmatrix} \tilde{z}^k \right\| \xrightarrow{k\to\infty} 0. \tag{B.22}$$

Furthermore:

1. *For* $\mathbf{M}$ *causal,* $\mathbf{\Delta}$ *can be chosen to be causal.*
2. *If* $\|\boldsymbol{\lambda}\mathbf{\Delta}_i^k - \mathbf{\Delta}_i^k\boldsymbol{\lambda}\| < \nu$ *for the* i*-th block* $\mathbf{\Delta}_i^k$ *in* $\mathbf{\Delta}^k$, *then the corresponding* $\mathbf{\Delta}_i$ *can be chosen to satisfy* $\|\boldsymbol{\lambda}\mathbf{\Delta}_i - \mathbf{\Delta}_i\boldsymbol{\lambda}\| < \nu$.

Proof:

The following construction was suggested by Dullerud [30]. First note that applying Theorem B.1, we can assume $\mathbf{\Delta}^k$, z^k have finite support, say $[0, T_k)$. If, in addition, $\mathbf{M}$ is causal, $\mathbf{\Delta}^k$ can be assumed causal from Theorem B.2.

By enlarging the definition of T_k, we can ensure $\|\mathbf{P}^{[T_k,\infty)}\mathbf{M}z^k\|_2 \xrightarrow{k\to\infty} 0$. Let τ_k be defined by $\tau_0 = 0$, $\tau_{k+1} = \tau_k + T_k + 1$ for $k > 0$. Introduce the operator

$$\mathbf{\Delta} = \sum_{k=0}^{\infty} \boldsymbol{\lambda}^{\tau_k}\mathbf{\Delta}^k\boldsymbol{\lambda}^{-\tau_k} \tag{B.23}$$

which has the infinite matrix representation of Figure B.2, where $\mathbf{\Delta}^k_{T_k}$ denotes the truncated matrix as in (2.14) corresponding to $\mathbf{\Delta}^k$. Note that if $\mathbf{\Delta}^k$ is causal, so is $\mathbf{\Delta}$.

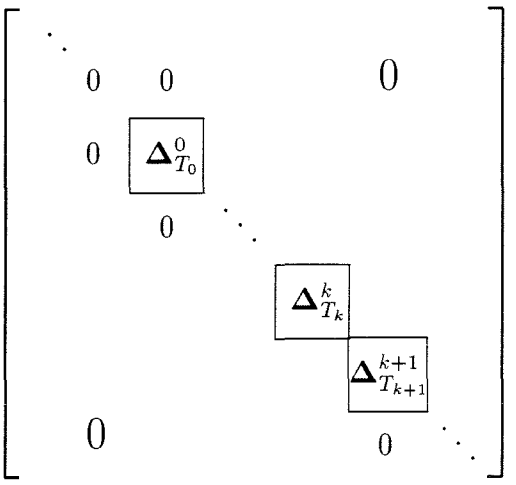


Figure B.2: Infinite matrix representation of the operator $\boldsymbol{\Delta}$

An analogous reasoning to the one which led to (B.20) implies that if $\|\boldsymbol{\lambda}\boldsymbol{\Delta}_i^k - \boldsymbol{\Delta}_i^k\boldsymbol{\lambda}\| < \nu$ for every k, then $\|\boldsymbol{\lambda}\boldsymbol{\Delta}_i - \boldsymbol{\Delta}_i\boldsymbol{\lambda}\| < \nu$. Similarly, it follows that since $\|\boldsymbol{\Delta}^k\| \leq 1$, $\|\boldsymbol{\Delta}\| \leq 1$. Defining the sequence $\tilde{z}^k := \boldsymbol{\lambda}^{\tau_k} z^k$, it only remains to show that (B.22) holds. We write

$$\mathbf{M}\tilde{z}^k = \boldsymbol{\lambda}^{\tau_k}\mathbf{M}z^k = \boldsymbol{\lambda}^{\tau_k}\mathbf{P}^{[0,T_k)}\mathbf{M}z^k + e^k, \tag{B.24}$$

where $e^k = \boldsymbol{\lambda}^{\tau_k}\mathbf{P}^{[T_k,\infty)}\mathbf{M}z^k \overset{k\to\infty}{\longrightarrow} 0$ from the choice of the T_k above. This leads to

$$\boldsymbol{\Delta}\mathbf{M}\tilde{z}^k = \boldsymbol{\Delta}e^k + \sum_{l=0}^{\infty}\boldsymbol{\lambda}^{\tau_l}\boldsymbol{\Delta}^l\boldsymbol{\lambda}^{-\tau_l}\boldsymbol{\lambda}^{\tau_k}\mathbf{P}^{[0,T_k)}\mathbf{M}z^k = \boldsymbol{\Delta}e^k + \boldsymbol{\lambda}^{\tau_k}\boldsymbol{\Delta}^k\mathbf{M}z^k. \tag{B.25}$$

Consequently

$$\|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})\tilde{z}^k\|_2 = \|\boldsymbol{\lambda}^{\tau_k}(\mathbf{I} - \boldsymbol{\Delta}^k\mathbf{M})z^k - \boldsymbol{\Delta}e^k\|_2 \tag{B.26}$$

converges to zero as $k\to\infty$, using (B.21) and $\|e^k\|_2 \overset{k\to\infty}{\longrightarrow} 0$. Also, since $\mathbf{N}$ is LTI we have $\|\mathbf{N}\tilde{z}^k\| = \|\mathbf{N}z^k\| \overset{k\to\infty}{\longrightarrow} 0$. Therefore (B.22) is satisfied.

■

The result is now applied to the question of uniform robust stability in the class $\boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$.

Corollary B.4 *Suppose* $\mathbf{M}$ *is a causal* LTI *system in* $\mathcal{L}_c(l_2)$ *, and* $\mathbf{I} - \boldsymbol{\Delta}\mathbf{M}$ *is invertible in* $\mathcal{L}_c(l_2)$ *for each* $\boldsymbol{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$*. Then*

$$\sup_{\boldsymbol{\Delta}\in\, \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}} \left\|(\mathbf{I} - \boldsymbol{\Delta}\mathbf{M})^{-1}\right\| < \infty. \tag{B.27}$$

Proof: If (B.27) does not hold, then for any $\epsilon > 0$ we can find $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$ such that $\|(\mathbf{I} - \mathbf{\Delta M})^{-1}\| > 1/\epsilon$. Therefore there exists $z \in l_2$ such that $\|z\|_2 = 1$, $\|(\mathbf{I} - \mathbf{\Delta M})z\|_2 < \epsilon$.

Since this holds for any ϵ, there exist sequences z^k, $\|z^k\|_2 = 1$, $\mathbf{\Delta}^k \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$ such that

$$\|(\mathbf{I} - \mathbf{\Delta}^k \mathbf{M})z^k\|_2 \overset{k\to\infty}{\longrightarrow} 0. \tag{B.28}$$

Applying now Theorem B.3, we find a single, causal $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}$ and signals $\tilde{z}^k$, $\|\tilde{z}^k\| = 1$ satisfying $\|(\mathbf{I} - \mathbf{\Delta M})\tilde{z}^k\|_2 \overset{k\to\infty}{\longrightarrow} 0$. This contradicts the hypothesis.

■

Turning now to the case of $\boldsymbol{\Delta}^{\nu}$, we show that if robust stability holds for a certain class, uniform robust stability holds for any class of smaller time-variation.

Corollary B.5 *Suppose* $\mathbf{M}$ *is a causal* LTI *system in* $\mathcal{L}_c(l_2)$, *and* $\mathbf{I} - \mathbf{\Delta M}$ *is invertible in* $\mathcal{L}_c(l_2)$ *for each* $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}}$. *If* $\overline{\nu} < \nu$, *then*

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\overline{\nu}}}} \|(\mathbf{I} - \mathbf{\Delta M})^{-1}\| < \infty. \tag{B.29}$$

Proof: If (B.29) does not hold, then for any $\epsilon > 0$ we find $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\overline{\nu}}}$ and $z \in l_2$ such that $\|z\|_2 = 1$, $\|(\mathbf{I} - \mathbf{\Delta M})z\|_2 < \epsilon$. This gives a sequence $\mathbf{\Delta}^k$, z^k satisfying (B.28). Additionally, $\|\mathbf{\lambda}\mathbf{\Delta}^k - \mathbf{\Delta}^k\mathbf{\lambda}\| < \nu$.

Now Theorem B.3 gives $\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{\nu}}$, $\tilde{z}^k$, $\|\tilde{z}^k\| = 1$ satisfying $\|(\mathbf{I} - \mathbf{\Delta M})\tilde{z}^k\|_2 \overset{k\to\infty}{\longrightarrow} 0$. This contradicts the hypothesis.

■

B.4 Proofs for Chapter 6

This section contains the proof of Propositions 6.8 and 6.10. The notation is taken from in Chapter 6.

B.4.1 Proof of Proposition 6.8

By hypothesis the system is robustly stable: using Corollary B.4 we can write the bound

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}} \|(\mathbf{I} - \mathbf{\Delta M}_{11})^{-1}\| = \beta < \infty. \tag{B.30}$$

We also know by hypothesis that

$$\sup_{\mathbf{\Delta} \in \boldsymbol{B}_{\boldsymbol{\Delta}^{LTV}}} \|\mathbf{\Delta} \star \mathbf{M}\|_{W_\eta} = \gamma < 1. \tag{B.31}$$

Given $\epsilon > 0$, suppose $\nabla \cap \mathcal{K}_\epsilon \neq \emptyset$. Then there exists $z = \mathrm{col}(p, v)$, $\|z\| = 1$, such that $\Lambda(z) \in \mathcal{K}_\epsilon$, with Λ defined in (6.27). This implies that $\rho(z_{n+1}) \in B(0, \eta)$, so $v \in W_\eta$, and

$$\sigma_i(z) = \|(\mathbf{M}z)_i\|_2^2 - \|z_i\|_2^2 > -\epsilon^2, \qquad i = 1 \ldots n+1. \tag{B.32}$$

We are now in the situation of Lemma 5.1, except for the term ϵ^2; it is shown analogously that there exist contractive operators $\mathbf{\Delta}_i \in \mathcal{L}(l_2)$ such that

$$\|z_i - \mathbf{\Delta}_i(\mathbf{M}z)_i\|_2 \leq \epsilon, \qquad i = 1 \ldots n+1. \tag{B.33}$$

Therefore we obtain a structured operator $\mathbf{\Delta} = \mathrm{diag}[\mathbf{\Delta}_1, \ldots, \mathbf{\Delta}_{n+1}] \in \boldsymbol{B}_{\boldsymbol{\Delta}_{NC}^{LTV}}$ such that $\|(\mathbf{I} - \mathbf{\Delta}\mathbf{M})z\| = O(\epsilon)$. From Theorem B.2, a causal operator $\hat{\mathbf{\Delta}} = \mathrm{diag}[\hat{\mathbf{\Delta}}_1, \ldots, \hat{\mathbf{\Delta}}_{n+1}]$ and a signal $\hat{z}$, $\|\hat{z}\|_2 = 1$ are found satisfying

$$\|(\mathbf{I} - \hat{\mathbf{\Delta}}\mathbf{M})\hat{z}\|_2 = O(\epsilon). \tag{B.34}$$

We now isolate the first n components from the $(n+1)$-th in (B.34), using the notation $\hat{z} = \mathrm{col}(\hat{p}, \hat{v})$ and $\mathbf{M}\hat{z} = \mathrm{col}(\hat{q}, \hat{e})$. Note that from Theorem B.2, $\hat{v}$ can be chosen in $\hat{W}_\eta$. The last component in (B.34) gives

$$\|\hat{v} - \hat{\mathbf{\Delta}}_{n+1}\hat{e}\|_2 = O(\epsilon), \tag{B.35}$$

and the first n components give

$$\hat{p}_i = \hat{\mathbf{\Delta}}_i\hat{q}_i + \hat{d}_i, \quad i = 1, \ldots, n, \tag{B.36}$$

where the signal $\hat{d}$, of norm $O(\epsilon)$ can be viewed as a disturbance injected at the interconnection, as depicted in Figure B.3.

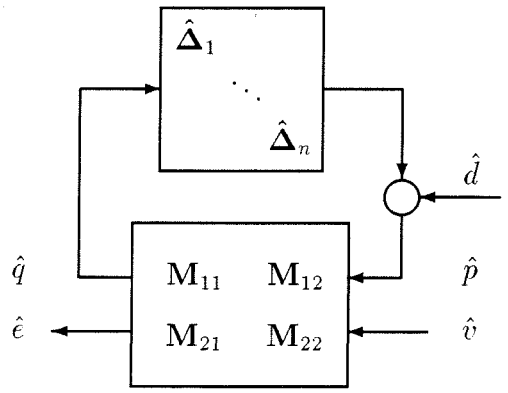


Figure B.3: Uncertain system with injected $\hat{d}$

We now state the following bounds:

$$\|\hat{v}\|_2 - O(\epsilon) \leq \|\hat{e}\|_2 \leq \beta\|\mathbf{M}_{21}\|\|\hat{d}\|_2 + \gamma\|\hat{v}\|_2. \tag{B.37}$$

The lower bound is a direct consequence of (B.35). The upper bound is obtained by writing $\hat{e}$ as the superposition of the contributions of the inputs $\hat{d}$ and $\hat{v} \in \hat{W}_\eta$ in Figure B.3, and using the bounds (B.30) and (B.31).

If $\mathcal{K}_\epsilon \cap \nabla \neq \emptyset$ holds for every ϵ, take $\epsilon_k \stackrel{k\to\infty}{\longrightarrow} 0$, and generate the sequences $\hat{\mathbf{\Delta}}^k$, $\hat{z}^k = \mathrm{col}(\hat{p}^k, \hat{v}^k)$, and $\hat{d}^k$ with $\|\hat{z}^k\|_2 = 1$, $\|\hat{d}^k\|_2 = O(\epsilon_k)$. From the bounds (B.37) and $\gamma < 1$ it follows that we must have $\|\hat{v}^k\|_2 \stackrel{k\to\infty}{\longrightarrow} 0$; also, $\|\hat{d}^k\|_2 \stackrel{k\to\infty}{\longrightarrow} 0$. Now from (B.30) the gain from $(\hat{d}, \hat{v})$ to $\hat{z}$ can be uniformly bounded across $\hat{\mathbf{\Delta}}$, which implies that $\|\hat{z}^k\|_2 \stackrel{k\to\infty}{\longrightarrow} 0$, which is a contradiction.

Therefore there must exist $\epsilon > 0$ such that $\mathcal{K}_\epsilon \cap \nabla = \emptyset$.

■

B.4.2 Proof of Proposition 6.10

This proof requires very minor modifications of the previous one. By hypothesis we have robust stability under the class $\boldsymbol{B}_{\boldsymbol{\Delta}^\nu}$. Choosing $\bar{\nu}$ such that $2\sin\frac{h}{2} < \bar{\nu} < \nu$, Corollary B.5 implies that

$$\sup_{\mathbf{\Delta}\in\boldsymbol{B}_{\boldsymbol{\Delta}^{\bar{\nu}}}} \left\|(\mathbf{I} - \mathbf{\Delta}\mathbf{M}_{11})^{-1}\right\| = \beta < \infty. \tag{B.38}$$

We also know by hypothesis that

$$\sup_{\mathbf{\Delta}\in\,\boldsymbol{B}_{\boldsymbol{\Delta}^\nu}} \|\mathbf{\Delta} \star \mathbf{M}\|_{\hat{W}_\eta} = \gamma < 1. \tag{B.39}$$

With ∇, $\mathcal{K}_\epsilon$ defined in (6.43-6.42), suppose $\nabla \cap \mathcal{K}_\epsilon \neq \emptyset$. Then there exists $z = \mathrm{col}(p, v)$, $\|z\|_2 = 1$, such that $\Lambda(z) \in \mathcal{K}_\epsilon$, with Λ defined in (6.41). This implies that $\rho(z_{n+1}) \in B(0, \eta)$, so $v \in W_\eta$, and

$$[\varphi_i(z)](s) = \int_s^{s+h} \left(|(Mz)_i(e^{j\omega})|^2 - |z_i(e^{j\omega})|^2\right) \frac{d\omega}{2\pi} > -\epsilon^2, \quad i = 1, \ldots, n, \tag{B.40}$$

$$\sigma_{n+1}(z) = \|(\mathbf{M}z)_{n+1}\|_2^2 - \|z_{n+1}\|_2^2 > -\epsilon^2. \tag{B.41}$$

We can now apply Lemmas 5.4 and 5.1 to obtain contractive operators, $\mathbf{\Delta}_i$, satisfying $\|\boldsymbol{\lambda}\mathbf{\Delta}_i - \mathbf{\Delta}_i\boldsymbol{\lambda}\| \leq 2\sin\frac{h}{2}$ for $i = 1, \ldots, n$, and $\mathbf{\Delta}_{n+1} \in \mathcal{L}(l_2)$ such that

$$\|z_i - \mathbf{\Delta}_i(\mathbf{M}z)_i\|_2 \leq \epsilon, \quad i = 1 \ldots n+1. \tag{B.42}$$

Once again we can obtain from Theorem B.2 a causal operator $\hat{\mathbf{\Delta}} = \text{diag}[\hat{\mathbf{\Delta}}_1, \ldots, \hat{\mathbf{\Delta}}_{n+1}]$ and a signal $\hat{z}$, $\|\hat{z}\|_2 = 1$ satisfying

$$\|(\mathbf{I} - \hat{\mathbf{\Delta}}\mathbf{M})\hat{z}\|_2 = O(\epsilon). \tag{B.43}$$

Note that from Theorem B.2 we can ensure that the blocks $\hat{\mathbf{\Delta}}_1, \ldots, \hat{\mathbf{\Delta}}_n$ are in $\boldsymbol{B}^{\bar{\nu}}$, and the signal $\hat{v}$ is in $\hat{W}_\eta$. From here on the proof carries on exactly as in the case of Proposition 6.8.

■

B.5 Proof of Proposition 8.6

For a given $\epsilon > 0$, assume that $\overline{\nabla^\epsilon} \cap \overline{\mathbb{X}^P} \neq \emptyset$. Choose $z \in l_2$, such that $\|z\| = 1$, $\|\mathbf{N}z\| < \epsilon$, and $\Lambda(z) + \epsilon^2 I \geq 0$. This implies that

$$\Sigma_i(z) = \int_0^{2\pi} \left[(Mz)_i(e^{j\omega})(Mz)_i^*(e^{j\omega}) - z_i(e^{j\omega})z_i^*(e^{j\omega})\right] \frac{d\omega}{2\pi} > -\epsilon^2 I, \quad i = 1 \ldots L, \tag{B.44}$$

$$\sigma_{L+j}(z) = \|(\mathbf{M}z)_{L+j}\|^2 - \|z_{L+j}\|^2 > -\epsilon^2, \quad j = 1 \ldots F. \tag{B.45}$$

We can now apply Lemmas 5.2 and 5.1 to obtain contractive operators $\boldsymbol{\delta}_i\mathbf{I}$, $i = 1, \ldots, L$, and $\mathbf{\Delta}_{L+j}$, such that

$$\|z_i - \boldsymbol{\delta}_i\mathbf{I}(\mathbf{M}z)_i\|_2 \leq O(\epsilon), \quad i = 1 \ldots L, \tag{B.46}$$

$$\|z_{L+j} - \mathbf{\Delta}_{L+j}(\mathbf{M}z)_{L+j}\|_2 \leq \epsilon, \quad j = 1 \ldots F. \tag{B.47}$$

Constructing $\mathbf{\Delta} = \text{diag}[\boldsymbol{\delta}_1\mathbf{I}, \ldots, \boldsymbol{\delta}_L\mathbf{I}, \mathbf{\Delta}_{L+1}, \ldots, \boldsymbol{\delta}_i\mathbf{\Delta}_{L+F}]$ leads to $\left\| \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}\mathbf{M} \\ \mathbf{N} \end{bmatrix} z \right\| = O(\epsilon)$. Since this holds for any ϵ, there exist sequences of signals z^k, $\|z^k\| = 1$ and perturbations $\mathbf{\Delta}^k$ such that

$$\left\| \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}^k\mathbf{M} \\ \mathbf{N} \end{bmatrix} z^k \right\| \xrightarrow{k\to\infty} 0. \tag{B.48}$$

Now we invoke Theorem B.3 to find $\mathbf{\Delta}$, and $\hat{z}^k$ $\|\hat{z}^k\| = 1$ such that

$$\left\| \begin{bmatrix} \mathbf{I} - \mathbf{\Delta}\mathbf{M} \\ \mathbf{N} \end{bmatrix} z^k \right\| \xrightarrow{k\to\infty} 0, \tag{B.49}$$

which contradicts robust l_2-stability.

■

Remark: The same technique can be used for the proof of Theorem 8.7, since from Theorem B.3, $\mathbf{\Delta}$ can be chosen to be causal when $\mathbf{M}$ is causal.

Bibliography

[1] Athans M., *A Tutorial on the LQG/LTR Method*, Proceedings 1986 ACC, Seattle, WA, pp. 1289-1296.

[2] Anderson R.L., Anderson T.W., *Distribution of the Circular Serial Coefficient for Residuals From a Fitted Fourier Series*, Ann. Math. Statist., 21, pp. 59-81, 1956.

[3] Anderson B., Moore J.B., *Optimal Control: Linear Quadratic Methods*, Prentice Hall, 1990.

[4] Balakrishnan V., Huang Y., Packard A., and Doyle J., *Linear Matrix Inequalities in Analysis with Multipliers*, Proceedings 1994 ACC, Baltimore, MD., pp. 1228-1232.

[5] Balas G., Doyle J., Glover K., Packard A., and Smith R., *The μ Analysis and Synthesis Toolbox,* MathWorks and MUSYN, 1991.

[6] Bartlett M.S., *An Introduction to Stochastic Processes*, Cambridge University Press, 1955.

[7] Bercovici H., Foias C., Tannenbaum A., *Structured Interpolation Theory*, Operator Theory Adv. and App. , 47, pp. 195-220, 1990.

[8] Bernstein D.S, Haddad W.H., *LQG Control with an $\mathcal{H}_\infty$ Performance Bound: A Riccatti Equation Approach*, IEEE Trans. A.C., Vol 34(3), pp. 293-305, 1989.

[9] Billingsley, P., *Convergence of Probability Measures*, J.Wiley & Sons, New York, 1968.

[10] Boyd, S.P., El Ghaoui L., E. Feron E., and Balakrishnan V., *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, Pennsylvania, 1994.

[11] Braatz, R., Young, P., Doyle, J., Morari, M., *Computational Complexity of μ Calculation*, IEEE Trans. A. C., vol 39, pp. 1000-1002, 1994.

[12] Breiman L., *Probability*, Addison-Wesley, 1968.

[13] Brillinger D., *Time Series Data Analysis and Theory*, Mc Graw-Hill, 1981.

[14] Cunningham T.B., Shaner D.A., and Doyle J.C., *State Reconstruction for Flight Control Reversion Modes*, Proceedings 1977 CDC.

[15] Dahleh M.A., Diaz-Bobillo I., *Control of Uncertain Systems: A Linear Programming Approach*, Prentice-Hall, 1995.

[16] Dahleh M.A, Theodosopoulos T.V., Tsitsiklis J., *The Sample Complexity of Worst-Case Identification of FIR Linear Systems*, Sys. & Control Letters, 20, pp. 157-166, 1993.

[17] D'Andrea R., Paganini F., *Interconnection of Uncertain Behavioral Systems for Robust Control,* Proceedings 1993 CDC, San Antonio, Texas, pp. 3642-3647.

[18] D'Andrea R., $\mathcal{H}_\infty$ *Optimization with Spatial Constraints*, to appear in Proceedings 1995 CDC, Orlando, FA.

[19] D'Andrea R., *LMI Approach to Mixed Performance Objective Controllers: Application to Robust* $\mathcal{H}_2$ *Synthesis*, CDS Tech. Rep. CDS95-020, Calif. Inst. of Technology, 1995.

[20] Desoer C.A., Vidyasagar M., *Feedback Systems: Input-Output Properties*, Academic Press, 1975.

[21] Doyle J., Francis B.A., Tannenbaum A., *Feedback Control Theory*, Macmillan, 1992.

[22] Doyle J., *Guaranteed Margins for LQG Regulators*, IEEE Trans. A.C., Vol 23(4), pp. 756-757, 1978.

[23] Doyle, J.C., and Stein G., *Robustness with Observers*, IEEE Trans. A.C., Aug. 1979.

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

[25] Doyle, J.C., Wall J.E., Stein G., *Performance and Robustness Analysis for Structured Uncertainty*, Proceedings 1982 CDC, pp. 629–636.

[26] Doyle J., Glover K., Khargonekar P., Francis B., *State-Space Solutions to Standard* $\mathcal{H}_2$ *and* $\mathcal{H}_\infty$ *Problems*, IEEE Trans. A.C., Vol 34(8), pp. 831-847, 1989.

[27] Doyle J., Zhou K., Glover K., Bodenheimer B., *Mixed* $\mathcal{H}_2$ *and* $\mathcal{H}_\infty$ *Performance Objectives II: Optimal Control*, IEEE Trans. A.C., Vol 39(8), pp. 1575-1587, 1994.

[28] Doyle J., *Analysis of Feedback Systems with Structured Uncertainty*, IEE Proceedings, 129, 242-250, 1982.

[29] Dullerud G.E., *Control of Uncertain Sampled-Data Systems*, Systems & Control Series, Birkhauser, Boston, 1995.

[30] Dullerud G., *Personal Communication*, 1994.

[31] Fan M., Tits A., Doyle J., *Robustness in the Presence of Mixed Parametric Uncertainty and Unmodeled Dynamics*, IEEE Trans. A. C., vol 36, pp. 25-38, 1991.

[32] Feron E., *Analysis of Robust $\mathcal{H}_2$ Performance with Multipliers*, Proceedings 1994 CDC, pp. 2015-2020.

[33] Fradkov A., Yakubovich V.A., *The S-Procedure and Duality Theorems for Nonconvex Problems of Quadratic Programming*, Vestnik Leningrad University (1) 1973, pp. 81-87 (in Russian).

[34] Francis B.A., *A Course in $\mathcal{H}_\infty$ Control Theory*, Lecture Notes in Control and Information Sciences, No. 88, Springer-Verlag, 1987.

[35] Gahinet, P., Nemirovskii A., Laub A., Chilali M., *The LMI Control Toolbox*, The MathWorks Inc., Beta-Release, 1994.

[36] Garey M.R., Johnson D.S., *Computers and Intractability: A Guide to the Theory of NP Completeness*, W. H. Freeman, 1979.

[37] Glover K., Doyle J., *A State Space Approach to $\mathcal{H}_\infty$ Optimal Control*, Springer-Verlag Lecture Notes in Control and Information Sciences, No. 135, 1989.

[38] Glover, K., *All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and their $\mathcal{L}_\infty$-Error Bounds*, Int. J. Control, Vol.39, pp. 1115-1193, 1984.

[39] Helmicki A.J., Jacobson C.A., Nett C. N, *Control Oriented System Identification: A Worst-Case/Deterministic Approach in $\mathcal{H}_\infty$*, IEEE Trans. A. C., Vol 36 (10), pp. 1163-1176, 1991.

[40] Hoeffding W., *Probability Inequalities for Sums of Bounded Random Variables*, American Statistical Association Journal, 59, pp. 13-30, March 1963.

[41] Horowitz I.M., Shaked U., *Superiority of Transfer Function over State Variable Methods*, IEEE Trans. A.C., Feb 1975, pp. 84-97.

[42] Jenkins G., Watts D., *Spectral Analysis and its Applications*, Holden-Day, 1968.

[43] Khammash M., Pearson J.B., *Performance Robustness of Discrete-Time Systems with Structured Uncertainty*, IEEE Trans A.C., vol AC-36, 4, pp. 398-412, 1991.

[44] Khargonekar P., Rotea M., *Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Control: A Convex Optimization Approach*, IEEE Trans. A.C., Vol 36(7), pp. 824-837, 1991.

[45] Ljung, L., *System Identification Theory for the User*, Prentice-Hall, 1987.

[46] Lu W.M., Zhou K., Doyle J.C., *Stabilization of LFT Systems,* Proceedings 1991 CDC, Brighton, England, pp. 1239-1244.

[47] Luenberger D., *Optimization by Vector Space Methods*, Wiley 1968.

[48] Mari J., *A Counterexample in Power Signals Space*, preprint, submitted to IEEE TAC.

[49] Megretski A., *Necessary and Sufficient Conditions of Stability: A Multiloop Generalization of the Circle Criterion*, IEEE Trans. A.C., Vol 38(5), 1993.

[50] Megretski A., Treil S., *Power Distribution Inequalities in Optimization and Robustness of Uncertain Systems*, Journal of Mathematical Systems. Estimation and Control, Vol 3, No.3, pp. 301-319, 1993.

[51] Megretski A., *S-Procedure in Optimal Non-Stochastic Filtering,* Tech. Rep. Code Trita/Mat-92-0015, Royal Inst. of Tech., Sweden, 1992.

[52] Megretski A., Rantzer A., *System Analysis via Integral Quadratic Constraints*, Tech Report ISRN LUTFD2/TFRT–7531–SE, Lund Institute of Technology, Sweden, 1995.

[53] Newlin M., *Model Validation, Control, and Computation*, Ph.D. Thesis, California Institute of Technology, 1995.

[54] Packard A., Doyle J.C., Balas G., *"Linear, Multivariable Robust Control with a μ-Perspective,"* Journal of Dynamic Systems Measurement and Control-Transactions of the ASME, Vol. 115, pp. 426-438, 1993.

[55] Packard A., Doyle J.C., *The Complex Structured Singular Value*, Automatica, Vol. 29, No. 1, pp. 71-109, 1993.

[56] Packard A., Fan M., Doyle J., *A Power Method for the Structured Singular Value*, Proceedings 1988 IEEE CDC., pp. 2132-2137.

[57] Packard A., Pandey P, *Continuity Properties of the Real/Complex Structured Singular Value,* IEEE Trans. A.C., vol 38, pp. 415-428, 1993.

[58] Paganini F., *Set Descriptions of White Noise and Worst Case Induced Norms*, Proceedings 1993 CDC, San Antonio, Texas, pp. 3658-3663.

[59] Paganini F., D'Andrea R., and Doyle J., *Behavioral Approach to Robustness Analysis*, Proceedings 1994 ACC, Baltimore, MD., pp. 2782-2786.

[60] Paganini F., Doyle J., *Analysis of Implicit Uncertain Systems, Parts I and II*, CDS Tech. Rep. CDS94-018, California Institute of Technology, 1994, submitted to IEEE Trans A.C.

[61] Paganini F., *Analysis of Systems with Combined Time Invariant/Time Varying Structured Uncertainty*, Proceedings 1995 ACC, Seattle, WA., pp. 3878-3882.

[62] Peters M.A., and Stoorvogel A.A., *Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Control in a Stochastic Framework*, Lin. Alg. & Apps., Vol 205-206, pp. 971-996, 1994.

[63] Petersen I., McFarlane D.,Rotea M., *Optimal Guaranteed Cost Control of Discrete-Time Uncertain Linear Systems*, 1993 IFAC World Congress, Sydney, pp. 407-410.

[64] Poolla K., Tikku A., *Robust Performance against Time-Varying Structured Perturbations*, IEEE Trans. A.C., Vol 40(9), pp. 1589-1602, 1995.

[65] Poolla K., Khargonekar P., Tikku A., Krause J., Nagpal K., *A Time-Domain Approach to Model Validation*, IEEE Trans. A.C., 39(5), pp. 951-950, 1994.

[66] Poolla K., Tikku A., *On the Time Complexity of Worst-Case System Identification*, IEEE Trans AC., Vol 39(5), pp. 944-950, 1994.

[67] Rantzer A., *Uncertainties with Bounded Rates of Variation*, Proceedings 1993 ACC, San Francisco, CA, pp. 29-30.

[68] Rantzer A., Megretski A., *A Convex Parametrization of Robustly Stabilizing Controllers*, IEEE Trans. A.C., Vol 39(9), pp. 1802-1808, 1994.

[69] Redheffer, R., *On a Certain Linear Fractional Transformation*, J. Math. Phys., 39, pp. 269-286, 1960.

[70] Rockafellar R. T., *Convex Analysis*, Princeton University Press, 1970.

[71] Rudin W., *Real and Complex Analysis*, McGraw Hill, 1987.

[72] Ruelle D., *Chaotic Evolution and Strange Attractors*, Cambridge University Press, 1989.

[73] Safonov M.G., *Stability and Robustness of Multivariable Feedback Systems*, MIT Press, 1980.

[74] Safonov M., Athans M., *A Multiloop Generalization of the Circle Criterion for Stability Margin Analysis*, IEEE Trans A.C., Vol 26(2), pp. 415-422, 1981.

[75] Sandberg, I.W. *A Frequency Domain Condition for the Stability of Feedback Systems Containing a Single Time-Varying Nonlinear Element*, Bell Sys. Tech. J., Vol 43(3), pp. 1601-1608, 1964.

[76] Shamma, J., *Robust Stability with Time Varying Structured Uncertainty*, IEEE Trans. A.C., Vol 39, 4, pp. 714-724, 1994.

[77] Smirnov V.I., *A Course of Higher Mathematics Vol V*, Pergamon Press, 1964.

[78] Smith, R., Doyle, J., *Model Validation: A Connection Between Robust Control and Identification* , IEEE Trans A.C., Vol 37(6), pp. 942-952, 1992.

[79] Stein G., Athans M., *The LQG/LTR Procedure for Multivariable Feedback Control Design*, IEEE Trans. A.C., Vol AC-32(2), pp. 105-114, 1987.

[80] Stein G., Doyle J.C., *Beyond Singular Values and Loopshapes*, AIAA Journal of Guidance and Control, January, 1991.

[81] Stoorvogel A.A., *The Robust $\mathcal{H}_2$ Control Problem: A Worst-Case Design*, IEEE Trans. A.C., Vol 38(9), pp. 1358-1370, 1993.

[82] Sznaier M., *An (Almost) Exact Solution to General SISO Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Problems via Convex Optimization*, Proceedings 1993 ACC, San Francisco, CA, pp. 250-254.

[83] Tierno J., *A Computational Approach to Nonlinear System Analysis*, Ph.D. Thesis, California Institute of Technology, 1995.

[84] Toker O., Ozbay H., *On the NP-Hardness of the Purely Complex μ Computation, Analysis/Synthesis, and Some Related Problems in Multidimensional Systems*, Proceedings 1995 ACC, pp. 447-451.

[85] Tse D., Dahleh M.A., Tsitsiklis J., *Optimal Asymptotic Identification Under Bounded Disturbances*, IEEE Trans A.C., Vol 38 (8), pp. 1176-1190, 1993.

[86] Venkatesh, S. R., Dahleh M.A., *Classical System Identification in a Deterministic Setting*, to appear in Proceedings 1995 CDC.

[87] Vidyasagar M., *Control System Synthesis: A Factorization Approach*, MIT Press, 1985.

[88] Wiener N., *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, Wiley, New York, 1950.

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

[90] Willems J.C., *On the Existence of a Nonpositive Solution to the Ricatti Equation*, IEEE Trans A.C., Vol 19, pp. 592-593, 1974.

[91] Willems J.C., *Paradigms and Puzzles in the Theory of Dynamical Systems*, IEEE Trans. A.C., Vol. 36, pp. 259-294, 1991.

[92] Willems J.C., *Feedback in a Behavioral Setting*, Systems, Models and Feedback: Theory and Applications, pp. 179-191, 1992.

[93] Yakubovich, V.A., *Frequency Conditions for the Absolute Stability of Control Systems with Several Nonlinear or Linear Nonstationary Units*. Autom. Telemech., pp. 5-30, 1967.

[94] Yakubovich, V. A., *S-Procedure in Nonlinear Control Theory*, Vestnik Leningrad University, pp. 62-77, 1971 (in Russian) (English translation in Vestnik Leningrad Univ. 4:73-93, 1977).

[95] Young P., *Robustness with Parametric and Dynamic Uncertainty*, Ph.D. Thesis, California Institute of Technology, 1993.

[96] Zames G., *On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems. Part I: Conditions Derived Using Concepts of Loop Gain, Conicity, and Positivity*, IEEE Trans. A.C., Vol AC-11(2), pp. 228:238, 1966.

[97] Zames G., *Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses*. IEEE Trans. A.C., Vol AC-26(2), pp. 301-320, 1981.

[98] Zhou K., with Doyle J.C., Glover K., *Robust and Optimal Control*, Prentice Hall, 1995.

[99] Zhou K., Glover K., Bodenheimer B., Doyle J., *Mixed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Performance Objectives I: Robust Performance Analysis*, IEEE Trans. A.C., Vol 39(8), pp. 1564-1574, 1994.