Model Reduction and Minimality for Uncertain Systems

Thesis by

Carolyn L. Beck

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

California Institute of Technology
Pasadena, California
1997
Submitted January 12, 1996

To My Parents

Acknowledgements

I am deeply grateful to my advisor, Professor John C. Doyle, for providing me with his perspective and his insight during the course of this research. He has always encouraged me to push myself further, and faster, and has done so with humor and warmth.

It has been a pleasure to interact with the other members of the Caltech EE Controls Group during the past six years, all of whom have contributed to creating an environment which has been both challenging and enjoyable. In particular, I thank Matt Newlin for many helpful and interesting discussions, and Pascale Bendotti for both providing the power plant applications for my research and for collaborating on the implementation.

I was fortunate to have spent Michaelmas term of 1992 at Cambridge University with the CUED group: my sincere thanks to Professor Keith Glover for this opportunity. My thanks also to my committee, Professors Richard Murray, Jerrold Marsden and Yaser Abu-Mostafa of Caltech, and Professor Roy Smith of UC Santa Barbara, who also gave me invaluable encouragement early on during my graduate work at Caltech.

Throughout my term at Caltech I have been very fortunate to have had the support of family and friends. I thank my parents and my siblings for always having provided a caring and supportive atmosphere "at home", and for helping me keep it all in perspective. I also thank Laura, Sean, Kelli, Carolyn, Colleen, Michael, Jaimee and Adam for all the fun and for letting me be the true monster that I am. I am extremely grateful to my friends Carey and Becky for being there, and for sharing long runs, coffee breaks and movies. Additionally, my thanks to the members of the Caltech Cross Country and Track teams of the past 6 years, especially to Coach Jim O'Brien for making the track a wonderful escape for me from the office, and for also being a true friend.

Finally, an acknowledgement of my deepest gratitude to Geir, not only for his encouragement and advice during the past few years of this work, but more importantly for his companionship and unconditional support. *HS.*

Abstract

The emphasis of this thesis is on the development of systematic methods for reducing the size and complexity of uncertain system models. Given a model for a large complex system, the objective of these methods is to find a simplified model which accurately describes the physical system, thus facilitating subsequent control design and analysis.

Model reduction methods and realization theory are presented for uncertain systems represented by Linear Fractional Transformations (LFTs) on a block diagonal uncertainty structure. A complete generalization of balanced realizations, balanced Gramians and balanced truncation model reduction with guaranteed error bounds is given, which is based on computing solutions to a pair of Linear Matrix Inequalities (LMIs). A necessary and sufficient condition for exact reducibility of uncertain systems, the converse of minimality, is also presented. This condition further generalizes the role of controllability and observability Gramians, and is expressed in terms of singular solutions to the same LMIs. These reduction methods provide a systematic means for both uncertainty simplification and state order reduction in the case of uncertain systems, but also may be interpreted as state order reduction for multi-dimensional systems.

LFTs also provide a convenient way of obtaining realizations for systems described by rational functions of several noncommuting indeterminates. Such functions arise naturally in robust control when studying systems with structured uncertainty, but also may be viewed as a particular type of description for a formal power series. This thesis establishes connections between minimal LFT realizations and minimal linear representations of formal power series, which have been studied extensively in a variety of disciplines, including nonlinear system realization theory. The result is a fairly complete development of minimal realization theory for LFT systems.

General LMI problems and solutions are discussed with the aim of providing sufficient background and references for the construction of computational procedures to reduce uncertain systems. A simple algorithm for computing balanced reduced models of uncertain systems is presented, followed by a discussion of the application of this procedure to a pressurized water reactor for a nuclear power plant.

Table of Contents

List of Figures

Notation and Symbols

$\in$	belongs to
$\subset$	subset
$\cup$	set union
$\cap$	set intersection
$\times$	Cartesian product
$\emptyset$	null set
$\circ$	composition of operators
$\to$	tends to, maps to
$\mapsto$	maps to
$>$ $(\geq)$	greater (or equal) value, or more positive (semi) definite
$<$ $(\leq)$	lesser (or equal) value, or more negative (semi) definite
$\mathbb{Z}, \mathbb{Z}^+$	integers, non-negative integers
$\mathbb{R}$, $\mathbb{R}^+$	real numbers, non-negative real numbers
$\mathbb{C}$	complex numbers
$\mathbb{R}^{n\times m}, \mathbb{C}^{n\times m}$	ring of $n \times m$ real and complex matrices
$\lvert\alpha\rvert$	absolute value of $\alpha \in \mathbb{C}$
$\lvert v\rvert_2$	Euclidean norm of $v \in \mathbb{C}^n$
$\lVert v\rVert$	norm of $v \in \mathcal{V}$
$< u, v >$	inner product of $u, v \in \mathcal{V}$
I_n	$n \times n$ identity matrix
$\overline{\sigma}(A)$	maximum singular value of the matrix A
$\lambda_{min}(A)$, $\lambda_i(A)$	minimum eigenvalue, i^{th}eigenvalue of A
A^*	complex conjugate transpose or adjoint of A
l_2^n	square summable sequences mapping $\mathbb{Z}^+$ to $\mathbb{C}^n$ or $\mathbb{R}^n$
$\mathcal{L}(\mathcal{V})$	bounded linear operators on $\mathcal{V}$
$\boldsymbol{\lambda}$	delay operator on l_2^n
$\mathrm{spec}(T)$	spectrum of the operator T
$\mathrm{Im}\ T$	Image of T
$\mathrm{Ker}\ T$	Kernel of T
$\lVert T\rVert_{l_2\to l_2}$	l_2 induced norm of T
$\lVert G\rVert_\infty$	H_∞ norm of the system operator G
$\mu_\Delta(M)$	structured singular value of an operator M
$\mathrm{diag}[\alpha_1, \ldots, \alpha_n]$	$n \times n$ diagonal operator with elements $\alpha_1, \ldots, \alpha_2$

Acronyms

FPS	Formal Power Series
LFT	Linear Fractional Transformation
LMI	Linear Matrix Inequality
LPV	Linear Parameter Varying
LTI	Linear Time Invariant
LTV	Linear Time Varying
I/O	Input/Output
SISO	Single-Input Single-Output
MIMO	Multi-Input Multi-Output
SIMO	Single-Input Multi-Output
MR	Minimal Rank
BTMR	Balanced Truncation Model Reduction
PWR	Pressurized Water Reactor
RS	Robust Stability
RP	Robust Performance

Chapter 1

Introduction

Model based control methods are commonly used in the design of large, complex systems. Specifically, a mathematical model of the system is constructed, utilizing, for example, first principles analysis and experimental data, which is then used for subsequent control system design and analysis. For the purposes of feedback control highly accurate models are desired. However, such accuracy often requires that complicated high-order models be used, which in turn lead to more difficult control design problems from both an engineering and a computational perspective. The emphasis of this thesis is on the development of methods for reducing the size and complexity of the model while retaining the essential features of the system description. The main goal of these methods is to find a simplified system model which describes the physical system accurately enough so that controllers designed based on this simplified model perform well when implemented on the real system. Directly related to the topic of model reduction are the realization theory concepts of minimality and its converse reducibility, which are also addressed in detail in this thesis.

A fundamental limitation in achieving desired system performance via any control design process is the inherent uncertainty in modelling the dynamics of the system under consideration. This uncertainty arises during the modelling process, which requires making a number of assumptions, estimations and simplifications; for example, uncertainty is often attributed to unmodelled dynamics such as nonlinearities and disturbances, and to incomplete knowledge of exact values for many of the system parameters. The effects of model uncertainty in feedback control may be substantial, particularly for high performance systems, since many control strategies attempt to utilize all system information present in the model in order to meet demanding performance requirements. If the uncertainty in a system model is not adequately accounted for, the control strategy chosen may rely on exploiting system dynamics which are not actually present. On the other hand, overestimating uncertainty in the system model may result in designs which are too conservative, giving poor system performance.

One approach for reconciling these requirements is to design controllers that perform well on a set of models, rather than on a single model. The model set is defined using a nominal model which is considered to be perturbed by a prescribed uncertainty set; that is, the model itself explicitly includes an uncertainty description. By appropriately defining and structuring the uncertainty set, a model set is constructed which covers a range of possible system behavior, without allowing for too many unlikely or impossible models. These models and the systems they represent are referred to as *uncertain systems.*

There has been much research activity on model reduction methods in recent years, however, previous reduction methods have addressed only reduction of the state dimension of the model (that is, the nominal model) and fail to address the issue of reducing the uncertainty description. In notable contrast to such methods, this thesis presents a systematic model reduction method to reduce both the state dimension and the uncertainty description, providing a greater reduction in the overall size and complexity of the model. Furthermore, related realization theory for uncertain systems, including an explicit method to determine the existence of, and compute, minimal order equivalent realizations for uncertain system models is addressed. Both the model reduction methods and the realization theory developed in this research are applicable to multi-dimensional system realizations, and include the standard one-dimensional (1D) results as the simplest case.

1.1 Historical Overview

The development of earlier theory relevant to this research proceeded along two somewhat separate paths: one related to the *robustness* framework originally proposed by Zames in 1966 [73], and the other to the state-space realization theory developed mainly in the '60s by Gilbert [31], Zadeh and Desoer [72], Kalman [42], Rosenbrock [60] and others. The intention of this section is not to give a comprehensive review of these areas, but to note key ideas and results leading to the research described in this text.

In [73], Zames introduced the small gain theorem, which provides an exact robust stability test for systems perturbed by *unstructured* dynamic uncertainty. This test is said to be robust in that it holds when the nominal model is subjected to all allowable values of the uncertainty. These exact results for unstructured uncertainty give sufficient conditions for robust stability of systems with respect to *structured* uncertainty. However, for structured uncertainty, these results are often conservative. As a result, the notion of rearranging the uncertainty into block diagonal form and using structured scaling matrices to reduce conservativeness in the tests was suggested in the early '80s by Doyle [20], and Safonov [62]. We consider the framework developed by Doyle and coworkers for modelling systems with structured uncertainty, that of dynamic perturbations to a nominal system which enter in a linear fractional manner; see [21], [26], [55], [71] and the references therein for further details.

More recently, synthesis methods have been developed which provide systematic techniques to construct controllers for systems subject to structured uncertainty, and for which robust stability and performance are guaranteed (see for example [34], [71], [54]). These controllers have at least the same state dimensions and uncertainty set complexity as the original system model. Moreover, the synthesis of these controllers and the subsequent system analysis often rely on complicated computational solutions which become increasingly difficult to implement as the model size and complexity grows. Thus, the need for reducing both the nominal model and uncertainty description has become apparent.

A number of methods for reducing the state dimension of models were proposed in the '80s; examples include the balanced truncation model reduction method and its additive H_∞ norm error bound, and the optimal Hankel norm model reduction method and its Hankel norm error bound. These are state-space methods, and rely to a large extent on the notion of finding *balanced realizations* for systems. The use of balanced realizations was first proposed by Moore [51] as a means of better analyzing realizations for reducibility based on the comparative controllability and observability of the system states. This was intended as a more computable alternative to the problem of finding minimal state-space realizations, originally put forth by Kalman [42] and Gilbert [31]. Thus, from its inception, the notion of balanced model reduction has been intertwined with the notions of minimality, controllability and observability, and solutions to state-space Lyapunov equations. Specifically, when the controllability and observability Gramians, the solutions to the Lyapunov equations, are equivalent and diagonal the associated state-space model is said to be balanced. The states corresponding to the small-valued elements of the balanced Gramian are both weakly controllable and weakly observable and can be truncated with relatively little resulting error. The guaranteed a priori error bounds for the balanced model reduction method were found independently by Enns [24] and Glover [33]; the corresponding bounds for discrete-time systems were presented by Hinrichsen and Pritchard [36].

The work in this thesis builds on the balanced truncation method for 1D systems, generalizing these techniques and related realization theory to the linear fractional transformation (LFT) setting. The LFT models and results discussed herein are applicable to uncertain systems, multi-dimensional systems, or formal power series; we will focus mainly on the representation of uncertain systems in this setting but will include some discussion of and results for the latter two. The main results in this thesis include a necessary and sufficient condition for the exact reducibility of LFT systems, leading naturally to a notion of minimality for these systems. Furthermore, systematic model reduction methods with guaranteed a priori upper error bounds are given for uncertain and multi-dimensional systems models.

1.2 Outline of the Thesis

Chapter 2: Mathematical Preliminaries

In Chapter 2, basic definitions and facts from linear analysis and abstract algebra that are used in the main text and appendices are stated. In Section 2.1, a general discussion of normed spaces is given to provide background on the structure of the signal space with which we are mainly concerned; bounded linear operators acting on such spaces are also discussed. Topics of discussion include induced norms, convergence of sequences and operators, and operator spectra. Section 2.2 contains definitions for the basic algebraic structures of groups, rings and modules, along with one useful result on module homomorphisms; this latter section provides reference material for the formal power series discussions in Chapter 6. Relevant matrix identities and factorizations are presented in Section 2.3.

Chapter 3: Standard State-Space Realization Theory

Chapter 3 contains an overview of standard realization theory for one-dimensional (1D) systems. The emphasis of this thesis is on the reducibilty of models, which naturally leads to consideration of the concepts of minimality, controllability and observability, system Lyapunov equations and the use of similarity transformations. Before presenting our results on the reducibility of uncertain system models, we review these definitions and results from standard state-space realization theory in order to more readily note analogues and differences.

Chapter 4: Realization Theory for Uncertain Systems

One of the main results of this thesis, a necessary and sufficient condition for the reducibility of uncertain systems, is presented in Chapter 4. Prior to stating and proving this result, we introduce the paradigm we use for modelling uncertain systems: the Linear Fractional Transformation (LFT) framework. The main objects in these models are a constant realization matrix and a structured uncertainty set, with the focus being on uncertainty sets that have repeated scalar structures. We define a restricted class of similarity transformations for these LFT realizations, and then generalize the notions of stability, Lyapunov equations, Gramians, and balanced realizations. The necessary and sufficient reducibility condition is then stated and proved, and minimality is discussed. Comments on related realization theory topics such as controllability and observability are then given.

Chapter 5: Model Reduction of Uncertain Systems

We review the balanced truncation model reduction method and related error bounds in Chapter 5, first presenting the standard results for 1D systems, followed by the extension of these methods to uncertain systems. New model reduction error bounds for uncertain systems that are tighter than the original balanced truncation bounds are then presented. In order to quantitatively evaluate these model reduction error bounds, we start by defining an induced 2-norm for uncertain systems modelled by LFTs.

Chapter 6: LFTs and Formal Power Series

In Chapter 6 we discuss connections between the notion of minimality we present for LFTs in Chapter 4, and the notion of minimal representations for formal power series (FPS), developed mainly in the '70s in the context of nonlinear system realization theory. If we consider LFT realizations where the only structure we assume for the uncertainty set is the spatial structure of repeated scalar blocks, then the resulting LFT systems may also be viewed as a representation of rational functions in multiple noncommuting indeterminates, that is, as a particular realization of a FPS. The form of the FPS representations and the definition of minimality used differ from those used for the LFT representations we consider; we show that given a minimal FPS representation or a minimal LFT realization, the other (minimal) form can be directly computed.

Chapter 7: Computational Methods for Model Reduction and Applications

In Chapters 4 and 5 we present the solution to the model reduction problem in the form of two linear matrix inequalities (LMIs). In this chapter, we discuss general LMI problems and solutions, providing background and references on standard convex optimization methods and interior point methods which can be used for constructing algorithms to reduce uncertain systems. We then present one simple suboptimal procedure for solving the Lyapunov inequalities we consider, followed by a discussion of a "proof of concept" application to a pressurized water reactor for a nuclear power plant. The pressurized water reactor of this plant is described, controller designs based on full and reduced models of the reactor are discussed, and performance comparisons are given.

Chapter 2

Mathematical Preliminaries

In this chapter, we present definitions and facts from linear analysis and abstract algebra to be used as a reference in this thesis.

2.1 Normed Spaces and Bounded Linear Operators

We first define the basic underlying structure for the system spaces we consider in this thesis. A very brief presentation is made here; for more details see [9].

Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ (or $\mathbb{C}$). A **norm** is a mapping from $\mathcal{V}$ into $\mathbb{R}^+$, denoted by $\|\cdot\|$ and satisfying

(i) $\|x\| = 0$ if and only if $x = 0$;
(ii) $\|\alpha x\| = |\alpha| \, \|x\|$ for all $x \in \mathcal{V}$ and scalar α;
(iii) $\|x + y\| \leq \|x\| + \|y\|$.

The pair $(\mathcal{V}, \|\cdot\|)$ is referred to as a **normed space**.

An **inner product** is a mapping from $\mathcal{V} \times \mathcal{V}$ into $\mathbb{C}$, denoted by $\langle\cdot,\cdot\rangle$, and such that for all $x, y, z \in \mathcal{V}$ and scalars α and β the following are satisfied:

(i) $\langle x, x\rangle \geq 0$, with equality if and only if $x = 0$;
(ii) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$;
(iii) $\langle y, x\rangle = \langle x, y\rangle^*$.

An **inner product space** is defined by the pair $(\mathcal{V}, \langle\cdot,\cdot\rangle)$. The inner product $\langle\cdot,\cdot\rangle$ on $\mathcal{V}$ satisfies the *Cauchy-Schwartz inequality:*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Note that an inner product on a vector space $\mathcal{V}$ defines a norm on $\mathcal{V}$, $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$.

The systems we will consider in this thesis have their inputs and outputs in the space of all square summable sequences, l_2^n, that is, the space consisting of all sequences $x =$

$(x(1), x(2), \ldots)$ mapping $\mathbb{Z}^+$ to $\mathbb{R}^n$ such that

$$\|x\|_2 = \left(\sum_{k=1}^{\infty} |x(k)|_2^2 \right)^{\frac{1}{2}} < \infty,$$

where $|\cdot|_2$ denotes the Euclidean norm; note that this normed space is also an inner product space.

A sequence $\{x(k)\}_1^\infty$ is said to converge to x, in a normed space $\mathcal{V}$, if $\|x(k) - x\| \to 0$ as $k \to \infty$. If $\|x(k) - x(l)\| \to 0$ as $k \to \infty$ and $l \to \infty$ independently, then $\{x(k)\}_1^\infty$ is a Cauchy sequence. A normed space $\mathcal{V}$ is **complete** if every Cauchy sequence is convergent to an element of $\mathcal{V}$. A complete normed space is called a *Banach Space*; a complete inner product space is called a *Hilbert space*. The space l_2, defined above, is the canonical example of a Hilbert space.

A subset $\mathcal{W} \subset \mathcal{V}$ is **closed** if every convergent sequence in $\mathcal{W}$ converges to an element of $\mathcal{W}$, and is **bounded** if there is a constant $K \geq 0$ such that $\|w\| \leq K$ for every $w \in \mathcal{W}$. Let $\mathcal{W} \subset \mathbb{C}^{n \times n}$, then the Bolzano-Weierstrass theorem states that the following are equivalent (see [48] for details) :

(i) $\mathcal{W}$ is closed and bounded.

(ii) Every sequence in $\mathcal{W}$ has a subsequence which converges to an element of $\mathcal{W}$.

Let $\mathcal{V}$ and $\mathcal{W}$ be normed spaces over the same field. A **linear operator** from $\mathcal{V}$ to $\mathcal{W}$ is a map $T : \mathcal{V} \to \mathcal{W}$ such that

$$T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2),$$

for all $x_1, x_2 \in \mathcal{V}$ and scalars α and β. Furthermore, $T : \mathcal{V} \to \mathcal{W}$ is a **bounded linear operator** if

$$\|T\|_{\mathcal{V} \to \mathcal{W}} := \sup_{x \in \mathcal{V}, x \neq 0} \frac{\|Tx\|_{\mathcal{W}}}{\|x\|_{\mathcal{V}}} < \infty,$$

where $\|T\|_{\mathcal{V} \to \mathcal{W}}$ denotes the **induced norm** of T, and $\|\cdot\|_{\mathcal{V}}$, $\|\cdot\|_{\mathcal{W}}$ denote the respective norms of the spaces $\mathcal{V}$ and $\mathcal{W}$. The space of all such bounded linear operators is denoted by $\mathcal{L}(\mathcal{V}, \mathcal{W})$; when $\mathcal{V}$ and $\mathcal{W}$ are the same, this space is denoted by $\mathcal{L}(\mathcal{V})$. A sequence $T_k \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ is said to converge to T in the induced norm topology if $\|T_k - T\|_{\mathcal{V} \to \mathcal{W}} \to 0$ as $k \to \infty$.

The **image** of an operator T is Im $T = \{Tx : x \in \mathcal{V}\}$, and the **kernel** of T is Ker $T = \{x : Tx = 0\}$. T is a vector space *isomorphism* if and only if Ker $T = \{0\}$ and Im $T = \mathcal{W}$.

Let $\mathcal{V}$ be a Banach Space. An operator $T \in \mathcal{L}(\mathcal{V})$ is invertible, or nonsingular, if there exists an element $T^{-1} \in \mathcal{L}(\mathcal{V})$ such that $TT^{-1} = T^{-1}T = I$. The **spectrum** of T is defined as the set

$$\text{spec}(T) := \{\lambda \in \mathbb{C} : (\lambda I - T) \text{ is not invertible in } \mathcal{L}(\mathcal{V})\}.$$

Lemma 2.1 *The spectrum* $spec(T)$ *is a closed subset of the disc* $\{\lambda \in \mathbb{C} : |\lambda| < \|T\|\}$.

Details for Lemma 2.1 can be found in [9] (Chapter 12, Corollary 4).

2.2 Algebraic Concepts

In this section we state a few basic definitions from abstract algebra. This section should be used for reference while reading Chapter 6, namely for the discussions on formal power series, and the associated technical results in Appendix D. We present only those facts from the theory of groups, rings and modules to which we will refer specifically in the sequel; the reader is referred to [37] for more details.

Groups

A **binary operation** on a nonempty set G is a function $G \times G \to G$, which we will denote by the product notation, that is $(a, b) \mapsto ab$.

Definition 2.2 *A* **semigroup** *is a nonempty set* G *together with a binary operation on* G *which is*

(i) *associative:* $a(bc) = (ab)c$ *for all* $a, b, c \in G$;

a **monoid** *is a semigroup* G *which contains an*

(ii) *identity element* $e \in G$ *such that* $ae = ea = a$ *for all* $a \in G$;

a **group** *is a monoid* G *such that*

(iii) *for every* $a \in G$ *there exists an inverse element* $a^{-1} \in G$ *such that* $a^{-1}a = aa^{-1} = e$.

A semigroup G *is said to be* **abelian** *or* **commutative** *if its binary operation is*

(iv) *commutative:* $ab = ba$ *for all* $a, b \in G$.

Our main interest will be in monoids.

Rings

Definition 2.3 *A* **ring** *is a nonempty set* R *together with two binary operations, denoted by addition* $(+)$ *and multiplication such that:*

(i) $(R, +)$ *is an abelian group with identity element* 0;

(ii) $(ab)c = a(bc)$ *for all* $a, b, c \in R$;

(iii) $a(b + c) = ab + ac$ *and* $(a + b)c = ac + bc$.

(iv) If in addition, $ab = ba$ for all $a, b \in R$, then R is said to be a **commutative ring.**

(v) Furthermore, if R contains an element 1_R such that $1_R a = a 1_R = a$ for all $a \in R$, then R is said to be a **ring with identity.**

A **semiring** is a ring without subtraction, that is, a ring with no additive inverse. For example, the set of subsets of a monoid is equipped with the structure of a semiring, with addition being defined as the union of subsets. Rings and fields (defined below) are also semirings.

Definition 2.4 *An element c in a ring R with identity is said to be a* **unit** *if there exists $d \in R$ such that $dc = cd = 1_R$. A ring D with identity 1_D in which every nonzero element is a unit is called a* **division ring**. *A* **field** *is a commutative division ring.*

Obvious examples of fields include the real and complex numbers, $\mathbb{R}$ and $\mathbb{C}$.

Definition 2.5 *Let R be a ring, and S_R a nonempty subset of R that is closed under addition and multiplication in R. If S_R is itself a ring under these operations then S_R is called a* **subring** *of R.*

Modules and Homomorphisms

Modules over a ring are a generalization of abelian groups, which are actually modules over $\mathbb{Z}$.

Definition 2.6 *Let R be a ring. A (right) R-***module** *is an additive abelian group A together with a function $A \times R \to A$, $(a, r) \mapsto ar$, such that for all $r, s \in R$ and $a, b \in A$:*

(i) $(a + b)r = ar + br$.

(ii) $a(r + s) = ar + as$.

(iii) $(ar)s = a(rs)$.

(iv) If R has an identity element 1_R and $a1_R = a$ for all $a \in A$, then A is said to be a (right) **unitary** *R-***module.**

If R is a division ring, then a (right) unitary R-module is called a (left) **vector space.**

Obvious analogous definitions can be made for left R-modules via a function $A \times R \to A$.

If R is a commutative ring, then every left R-module A can be given the structure of a right R-module by defining $ar = ra$ for $r \in R$ and $a \in A$. Thus, every module over a commutative ring can be assumed to be both a left and a right module, without loss of generality.

Definition 2.7 *Let A and B be modules over a ring R. A function $f : A \to B$ is a (right)* R-**module homomorphism** *if for all $a, c \in A$ and $r \in R$*

$$f(a+c) = f(a) + f(c) \quad \text{and} \quad f(ar) = f(a)r.$$

f is an R-**module isomorphism** *(respectively* **epimorphism, monomorphism***) if it is bijective (respectively surjective, injective).*

We also note the following result; see [37] for details (Chapter IV, Theorem 1.7).

Theorem 2.8 *If R is a ring and $f : A \to B$ is an R-module homomorphism, then there exists a unique R-module homomorphism $g : A/Kerf \to B$ such that $g(a + Ker\, f) = g(a)$ for all $a \in A$, $Im\, g = Im\, f$ and $Ker\, g = \emptyset$. g is an R-module isomorphism if and only if f is an R-module epimorphism. In particular, $A/Ker\, f$ is isomorphic to $Im\, f$.*

Pictorially, this result can be represented by the commutative diagram shown in Figure 2.1.

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[d] & B \\
A/\text{Ker f} \arrow[ur, "g"'] &
\end{tikzcd}$$

Figure 2.1: R-module Isomorphisms

2.3 Matrix Identities and Factorizations

We use the following matrix inverse formulas and decompositions throughout the thesis.

The Matrix Inversion Formula

In the following, assume A and C are nonsingular $n \times n$ and $m \times m$ matrices, respectively, then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}.$$

Schur Complement Inverse Formula

Assume U is nonsingular, and $\Gamma = X - WU^{-1}V$ is nonsingular. Then

$$\begin{bmatrix} U & V \\ W & X \end{bmatrix}^{-1} = \begin{bmatrix} U^{-1} + U^{-1}V\Gamma^{-1}WU^{-1} & -U^{-1}V\Gamma^{-1} \\ -\Gamma^{-1}WU^{-1} & \Gamma^{-1} \end{bmatrix}.$$

Γ is known as the Schur complement of U.

Singular Value Decomposition

Given a matrix $A \in \mathbb{C}^{n\times m}$ with $\text{rank}(A) = r \leq \min\{n, m\}$, then we can factor A as follows:

$$A = U\Sigma V$$

where $U \in \mathbb{C}^{n}$, $U^*U = I$, $V \in \mathbb{C}^{m\times m}$, $VV^* = I$, and $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{n\times m}$ with $\Sigma_r = \text{diag}[\sigma_1, \ldots, \sigma_r]$. The $\sigma_i \in \mathbb{R}^+$, typically ordered by descending value, are called the *singular values* of A, and are equivalent to the square root of the eigenvalues of AA^*. The maximum singular value is denoted by $\overline{\sigma}(A)$, and is also the induced 2-norm of A.

Cholesky Factorization

Any matrix $A \geq 0$ may be written in the form

$$A = LL^*$$

where L is a lower triangular matrix with non-negative diagonal entries.

Chapter 3

Standard State-Space Realization Theory

In order to facilitate the discussion of realization theory and reducibility of uncertain systems, we begin with a brief review of standard realization theory, that is, realization theory for one-dimensional (1D) systems for which uncertainty descriptions are not included in the models. We state a number of results, which are now considered standard, without proof. More complete details may be found in many texts on the subject, see for example [41]; a geometric perspective for the concepts discussed in this chapter is given in [70].

We consider finite dimensional, linear time-invariant discrete time systems described by state-space equations of the form

$$\begin{aligned} x(k) &= A\boldsymbol{\lambda}x(k) + Bu(k) \\ y(k) &= C\boldsymbol{\lambda}x(k) + Du(k), \end{aligned} \tag{3.1}$$

where $\boldsymbol{\lambda}$ is the delay operator, $x(k) \in \mathbb{R}^n$ represents the state at time k, and u and y represent the input and output sequences, respectively; at any time k, $u(k) \in \mathbb{R}^m$ and $y(k) \in \mathbb{R}^q$. $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{q\times n}$, and $D \in \mathbb{R}^{q\times m}$ are the realization matrices. Throughout this thesis we will denote this realization by

$$M := \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

For the 1D case, we denote the system operator by $G := \boldsymbol{\lambda}I \star M = D + C\boldsymbol{\lambda}(I - A\boldsymbol{\lambda})^{-1}B$; that is, G is the map that takes the input signal u to the output signal $y := Gu$.

Remark 3.1 *We assume throughout this thesis that the realization matrices are real matrices as this is commonly true; however, the results we present also hold for complex valued realization matrices.*

3.1 Equivalent Realizations and Similarity Transformations

Given a realization M, one may obtain an equivalent realization by applying a similarity transformation. Specifically, we say two realizations,

$$M_1 = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix},$$

are **equivalent** if $G_1 = G_2$. Given a realization M_1, we can form another equivalent realization by a coordinate transformation of the state variables. For example, suppose $T \in \mathbb{C}^{n\times n}$, nonsingular, and

$$\hat{x}(k) = Tx(k), \tag{3.2}$$

then

$$\begin{aligned} \hat{x}(k) &= TA_1T^{-1}\boldsymbol{\lambda}\hat{x}(k) + TB_1u(k) \\ \text{and } y(k) &= C_1T^{-1}\boldsymbol{\lambda}\hat{x}(k) + Du(k). \end{aligned} \tag{3.3}$$

Denoting $A_2 = TA_1T^{-1}$, $B_2 = TB_1$, and $C_2 = C_1T^{-1}$, it is straightforward to verify that $G_1 = G_2$.

Two matrices related as $A_2 = TA_1T^{-1}$ are said to be **similar**, thus state variable transformations like that given in (3.2) and (3.3) are commonly referred to as *similarity transformations*. By applying a properly chosen similarity transformation to a given state-space realization, we are often able to find an equivalent realization with a more desirable structure, for example a balanced realization; this is discussed in Section 3.4.

3.2 Controllability, Observability and Minimality

The notions of controllability and observability of a state-space realization are central in standard realization theory. Of particular interest is the fact that easily implemented tests for controllability and observability can be used directly to determine if a given system realization is minimal. One of the underlying results used in standard realization theory is the Cayley-Hamilton theorem, which states that every square matrix A satisfies it's own *characteristic equation* $a(\lambda)$. That is, if $A \in \mathbb{R}^{n\times n}$ has characteristic polynomial $a(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$, then $a(A) = 0$. In other words, A^n can always be written as a linear combination of the lower order matrix terms $A^{n-1}, \ldots, A, I$.

Briefly, a state $\overline{x}$ is **reachable** if there exists a finite time k, and a sequence of inputs $u(0), u(1), \ldots, u(k)$ such that the initial state $x(0)$ can be transferred to $\overline{x}$ at time k, that is, $x(k) = \overline{x}$. Let $\mathcal{R}_0$ denote the set of all states that are reachable from $x(0) = 0$. Then $\mathcal{R}_0$ is a linear subspace of $\mathbb{R}^n$, and using the state-space equations of (3.1) and the Cayley-Hamilton theorem, we see that $\mathcal{R}_0$ can be written in terms of the realization matrices as

$$\mathcal{R}_0 = \operatorname{Im} B + A(\operatorname{Im} B) + \cdots + A^{n-1}(\operatorname{Im} B).$$

A realization M is said to be **controllable** if $\mathcal{R}_0 = \mathbb{R}^n$, that is, if the initial state can be transferred to any fixed $\overline{x} \in \mathbb{R}^n$ via a finite input sequence. Defining the *controllability matrix* Γ by

$$\Gamma = [B \; AB \; A^2B \cdots A^{n-1}B] \tag{3.4}$$

leads to the following result.

Theorem 3.2 *A realization M is controllable if and only if rank$(\Gamma) = n$.*

Similarly, a realization M is **observable** if, given an output sequence $y(0), \ldots, y(k)$, the initial state $x(0)$ can be uniquely determined, under the assumption $u = 0$. Again, we can use the state-space equations of (3.1) to see that a realization is observable if

$$\mathcal{N}_0 = \bigcap_{k=1}^{n} \operatorname{Ker}\left(CA^{k-1}\right) = 0.$$

$\mathcal{N}_0$ is referred to as the unobservable subspace. Defining the *observability matrix* $\mathcal{O}$ by

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{3.5}$$

gives the following *dual* result to Theorem 3.2.

Theorem 3.3 *A realization M is observable if and only if rank$(\mathcal{O}) = n$.*

Clearly, similarity transformations do not affect controllability and observability.

One of the important applications, theoretically, of the tests available for controllability and observability is as an explicit test for *minimality*, the definition of which is fairly obvious, but is given for completeness in the following:

Definition 3.4 *A realization M of the system defined by G is* **minimal** *if there is no equivalent realization for G,*

$$M_1 = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix},$$

such that dim$(A_1) <$ dim(A).

If a realization is minimal, it is also *irreducible*, that is we cannot reduce the size of the realization matrix M without incurring error. The following result, first established by Kalman [42], is used to provide for a direct test of minimality.

Theorem 3.5 *A realization M is minimal if and only if it is both controllable and observable.*

Furthermore, an important fact is that all minimal realizations for a given transfer function are related by similarity transformation.

Duality

Note that the controllability and observability matrices are duals in the sense that controllability (observability) of the realization M is equivalent to observability (controllability) of the realization

$$M^T = \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix}.$$

The realizations M and M^T are referred to as *dual realizations.*

Decomposition Structures for Realizations

Given a realization M that is not minimal, we can always find a similarity transformation T to produce the following decomposition of the realization matrices [42], [31]:

$$\widehat{M} = \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} M \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & 0 \end{bmatrix} = \left[\begin{array}{cccc|c} \hat{A}_{co} & 0 & \hat{A}_{13} & 0 & \hat{B}_{co} \\ \hat{A}_{21} & \hat{A}_c & \hat{A}_{23} & \hat{A}_{24} & \hat{B}_c \\ 0 & 0 & \hat{A}_o & 0 & 0 \\ 0 & 0 & \hat{A}_{43} & \hat{A}_{44} & 0 \\ \hline \hat{C}_{co} & 0 & \hat{C}_o & 0 & 0 \end{array}\right], \quad (3.6)$$

where we have assumed $D = 0$ for simplicity of exposition. The subsystem defined by $\hat{A}_{co}$, $\hat{B}_{co}$, and $\hat{C}_{co}$ is both controllable and observable, and $G_{co} = G$, that is, the system operator is determined solely by the controllable and observable subsystem; alternatively, the subsystem defined by

$$M_{\overline{co}} = \left[\begin{array}{c|c} \hat{A}_{44} & 0 \\ \hline 0 & 0 \end{array}\right]$$

is neither controllable nor observable. The subsystem defined by

$$A_{c\bar{o}} = \begin{bmatrix} \hat{A}_{co} & 0 \\ \hat{A}_{21} & \hat{A}_c \end{bmatrix}, \; B_{c\bar{o}} = \begin{bmatrix} \hat{B}_{co} \\ \hat{B}_c \end{bmatrix} \text{ and } C_{c\bar{o}} = \begin{bmatrix} \hat{C}_{co} & 0 \end{bmatrix}$$

is controllable, but not observable; and similarly, that by

$$A_{\bar{c}o} = \begin{bmatrix} \hat{A}_{co} & \hat{A}_{13} \\ 0 & \hat{A}_o \end{bmatrix}, \; B_{\bar{c}o} = \begin{bmatrix} \hat{B}_{co} \\ 0 \end{bmatrix} \text{ and } C_{\bar{c}o} = \begin{bmatrix} \hat{C}_{co} & \hat{C}_o \end{bmatrix}$$

is observable, but not controllable. One approach for obtaining a similarity transformation that takes a given realization to one with the decomposition structure of (3.6) is discussed in Section 3.4.

3.3 Lyapunov Equations

Fundamental to our interests regarding realization theory and model reduction is the concept of system Lyapunov equations and their relation to stability, and to controllability and observability.

Stability

With regard to standard linear time-invariant (LTI) systems described by state-space realizations, there are two definitions commonly used for stability, *internal* and *external* stability. Informally, a system is said to be externally stable if a bounded input sequence u results in a bounded output y; this is also often referred to as input/output stability. Internal stability, also called asymptotic stability, refers to stability of the realization itself; as a result, we are more concerned with internal stability in this work. Additionally, if a system is internally stable it is also externally stable. We will refer to internal stability simply as stability in the sequel.

A realization M is said to be stable if the solution of

$$x(k) = A\boldsymbol{\lambda}x(k), \quad x(0) = x_0$$

tends to zero as $k \to \infty$, for any x_0. A well-known test for stability of a discrete-time LTI system realization, is

$$|\lambda_i(A)| < 1$$

for every $i = 1, \ldots, n$, where $\{\lambda_i(A)\}$ are the eigenvalues of A. An equivalent stability condition, based on a general method developed by Lyapunov in the 1890's for the study of both linear and nonlinear system stability, is given in the following.

Theorem 3.6 *(Lyapunov) A realization M is stable, if and only if for any matrix $Q > 0$, there exists a matrix $Y > 0$ satisfying*

$$A^*YA - Y + Q = 0.$$

The matrix equation given in Theorem 3.6 is one example of a *Lyapunov equation.*

Controllability and Observability Gramians

Given a stable realization M, two important Lyapunov equations in systems theory are the Lyapunov controllability equation,

$$AYA^* - Y + BB^* = 0, \tag{3.7}$$

and the Lyapunov observability equation,

$$A^* X A - X + C^* C = 0. \tag{3.8}$$

It is important to note that the unique solutions $Y \geq 0$ and $X \geq 0$ to equations (3.7) and (3.8) are equivalently defined as the Gramians

$$Y = \sum_{k=0}^{\infty} A^k B B^* (A^*)^k \text{ and } X = \sum_{k=0}^{\infty} (A^*)^k C^* C A^k. \tag{3.9}$$

The infinite sums of (3.9) are guaranteed to converge due to the stability assumption. Y and X are called the *controllability Gramian* and the *observability Gramian*, respectively. A standard result from Lyapunov theory is that M is controllable if and only if $Y > 0$, and M is observable if and only if $X > 0$. This can be seen by considering the rank conditions of Theorems 3.2 and 3.3, or the decomposition structure discussed in Section 3.2.

3.4 Balanced Realizations and Reducibility

Balanced realizations were first proposed by Moore [51] to better evaluate the model reduction problem, and the relation between lower order approximation and the minimal realization theory developed by Kalman (reviewed in Section 3.2). It is reasonable to assume that if we are given a minimal realization with large dimensions, we might want to reduce the size of the realization prior to completing system analysis or control synthesis. One obvious approach to this is to eliminate the *weakly controllable* and *weakly observable* states.

Consider the following realization:

$$M = \left[\begin{array}{cc|c} -1 & -32 & 1 \\ \frac{1}{2} & 2 & \frac{1}{4} \\ \hline -1 & 16 & 0 \end{array}\right].$$

For this system, the controllability Gramian is

$$Y = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{64} \end{bmatrix} > 0.$$

The smaller valued element of Y corresponds to a weakly controllable state. Eliminating this state from the model gives the lower order approximation,

$$M_r = \left[\begin{array}{c|c} -1 & 1 \\ \hline -1 & 0 \end{array}\right].$$

However, note that $\|G - G_r\|_\infty = 2$ (which is larger than $\|G\|_\infty$). The large relative error in this model approximation results from eliminating a state which is weakly controllable, but strongly observable, in fact the observability Gramian is

$$X = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 64 \end{bmatrix}.$$

This example is constructed based on consideration of the results of Moore [51], and those of Enns [24] and Glover [33], which show that for the purposes of reduction, we should use both Gramians to evaluate the relative controllability and observability of the different state variables. In particular, we use a coordinate transformation on the state-space so that the resulting Gramians are equivalent, that is, $\hat{Y} = \hat{X} = \Sigma$.

Definition 3.7 *Let M be a stable, minimal realization. Then M is a* **balanced realization** *if there exists a matrix $\Sigma = diag[\sigma_1, \dots, \sigma_n] \geq 0$ satisfying*

$$A\Sigma A^* - \Sigma + BB^* = 0, \ \textit{and}\ A^*\Sigma A - \Sigma + C^*C = 0. \tag{3.10}$$

The entries, σ_i, of Σ are called the **Hankel singular values** *of the system.*

We refer to such a realization as balanced to reflect the fact that all the states are equally controllable and observable, and therefore, if we now reduce the least controllable part of the state-space, we also reduce the least observable part.

Observe that if we transform the realization matrices A, B and C to $\hat{A} = TAT^{-1}$, $\hat{B} = TB$ and $\hat{C} = CT^{-1}$, then

$$\hat{Y} = \sum_{k=0}^{\infty} \hat{A}^k\hat{B}\hat{B}^*(\hat{A}^*)^k = T\sum_{k=0}^{\infty} A^kBB^*(A^*)^kT^* = TYT^* \tag{3.11}$$

and similarly

$$\hat{X} = T^{-1*}XT^{-1}, \tag{3.12}$$

where Y and X are the Gramians for the original realization. Thus, we would like to find a transformation T_{bal} such that $T_{bal}YT_{bal}^* = T_{bal}^{-1*}XT_{bal}^{-1}$. Transforming the original realization M by T_{bal} then gives a balanced realization. Given a minimal realization, a balancing transformation T_{bal} can constructed from $Y > 0$ and $X > 0$ as follows:

$$\begin{aligned} Y &= RR^* \\ U\Sigma^2U^* &= RXR^* \\ T_{bal} &= \Sigma U^*R^{-1*}. \end{aligned} \tag{3.13}$$

Computational issues involved in balancing realizations are discussed in more detail in Chapter 7. Error bounds resulting from the truncation of balanced realizations are given in Chapter 5.

If the realization, M, of a 1D system is not minimal, then there exists a similarity transformation, T, such that the controllability and observability Gramians are diagonal, and the controllable and observable subsystem is balanced. The following theorem is standard, so the proof is omitted.

Theorem 3.8 *For any stable system realization* $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *there exists* T *such that* $\widehat{M} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix}$ *has controllability and observability Gramians given by*

$$Y = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} \Sigma_1 & & & \\ & 0 & & \\ & & \Sigma_3 & \\ & & & 0 \end{bmatrix}$$

respectively, with Σ_1, Σ_2, Σ_3 *diagonal and positive definite.*

Since the uncontrollable and unobservable modes of any system realization are not present in the corresponding system transfer function, we can truncate the associated states, corresponding to the zeros in Y and X above, and obtain a minimal realization which has both Gramians equal to Σ_1. Such a system is **reducible** in the sense that there exists a lower order equivalent realization. Thus, for 1D system models with no uncertainty, singular controllability and observability Gramians indicate reducibility of a model, and lower order equivalent realizations are found using similarity transformations and truncations.

Specifically, one transforms a realization M to $\widehat{M}$ as in Theorem 3.8, and then partitions $\widehat{M}$ as follows:

$$\widehat{M} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} & \hat{B}_1 \\ \hat{A}_{21} & \hat{A}_{22} & \hat{B}_2 \\ \hat{C}_1 & \hat{C}_2 & D \end{bmatrix},$$

where $\dim(\hat{A}_{11}) = \dim(\Sigma_1)$. Finally, one truncates the blocks of $\widehat{M}$ corresponding to the unobservable and uncontrollable states, leaving

$$M_r = \begin{bmatrix} \hat{A}_{11} & \hat{B}_1 \\ \hat{C}_1 & D \end{bmatrix}.$$

This reduced realization results in an equivalent transfer function, that is, $D + C(zI - A)^{-1}B = D + \hat{C}_1(zI - \hat{A}_{11})^{-1}\hat{B}_1$.

3.5 Additional Realization Theory

A number of tests equivalent to Theorem 3.2 or Theorem 3.3 exist for determining whether a given realization is controllable or observable; these can be found in any text on the subject. However, a few remarks about the related properties of stabilizability and detectability are made here for completeness and later reference.

Stabilizability

Consider again the standard LTI state-space equations, namely

$$x(k) = A\boldsymbol{\lambda}x(k) + Bu(k)$$

and suppose we may choose

$$u(k) = Fx(k) + r(k)$$

where $r(k)$ is an external reference input, and $F : \mathbb{R}^n \to \mathbb{R}^m$ is a state feedback matrix. Note that the effect of introducing state feedback is to take the realization

$$M = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \text{ to } M_{sf} = \begin{bmatrix} A + BF & B \\ C & 0 \end{bmatrix}.$$

The following results are standard:

(i) If M is controllable, then M_{sf} is controllable.

(ii) Let Λ be any symmetric set of n complex numbers. Then M is controllable if and only if there exists a map $F : \mathbb{R}^n \to \mathbb{R}^m$ such that $\{\lambda_i(A + BF)\} = \Lambda$.

The result of (ii) leads us to the following.

Definition 3.9 *A realization M is* **stabilizable** *if there exists a map $F : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$|\lambda_i(A + BF)| < 1$$

for all $i = 1, \ldots, n$.

Thus, a realization is stabilizable if the uncontrollable subsystem is stable.

Detectability

A dual property to that of stabilizability is detectability. We now consider injecting some form of the output back into the system, that is,

$$x(k) = A\boldsymbol{\lambda}x(k) + Bu(k) + Ly(k)$$
$$y(k) = C\boldsymbol{\lambda}x(k)$$

where $L : \mathbb{R}^q \to \mathbb{R}^n$ is the output injection matrix, and

$$M_{oi} = \begin{bmatrix} A + LC & B \\ C & 0 \end{bmatrix}.$$

Then the obvious duals to (i) and (ii) hold, and we have the following notion of detectability.

Definition 3.10 *A realization M is* **detectable** *if there exists a map $L : \mathbb{R}^q \to \mathbb{R}^n$ such that*

$$|\lambda_i(A + LC)| < 1$$

for all $i = 1, \ldots, n$.

That is, a realization is detectable if the unobservable subsystem is stable.

In the following Chapter, we generalize the notions of similarity transformations and Lyapunov equations to multi-dimensional and uncertain system realizations. As a result, we are able to define generalized Gramians and balanced realizations for uncertain systems, which lead to a necessary and sufficient condition for exact reducibility; in Chapter 5, we discuss the formulation of error bounds for non-exact reduction of uncertain system models.

Chapter 4

Realization Theory for Uncertain Systems

The term *uncertain system* refers to a system described by perturbations on a nominal model, where the perturbations are represented in the form of structured uncertainty sets. This paradigm, used for evaluating the effects of modelling errors, was first proposed in [20], [62] and [21].

A convenient and general framework for representing and manipulating uncertain models is to use linear fractional transformations (LFTs) on the structured uncertainty sets. A comprehensive theory has been developed for such systems involving a great variety of assumptions on the uncertainty (see, for example, [55], [22], and the references therein). We give a brief review of the general LFT framework, followed by discussions on realization theory and related concepts for LFTs on repeated scalar uncertainty structures. The main result of this chapter is the necessary and sufficient reducibility condition stated and proved in Theorems 4.12 and 4.15. This result is based on reduced rank solutions to two LMIs that generalize Lyapunov equations.

4.1 Linear Fractional Transformations

The LFT paradigm, shown pictorially in Figure 4.1, allows for a mathematical representation of uncertainty in system models.

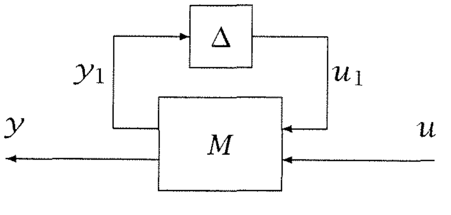

Figure 4.1: LFT/Uncertain System

In general, Δ represents uncertainty, or a dynamic element, and $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is a realization of the mapping from u to y which is given by the LFT

$$y = (\Delta \star M)u, \quad \Delta \in \boldsymbol{\Delta}$$

where

$$\Delta \star M := D + C\Delta(I - A\Delta)^{-1}B,$$

whenever the inverse is well-defined. Note that if we let $\Delta = \boldsymbol{\lambda} I$ then we recover the transfer function operator $\boldsymbol{\lambda} I \star M = D + C\boldsymbol{\lambda}(I - A\boldsymbol{\lambda})^{-1}B$ and a standard state-space realization with state x, input u and output y. There is both an extensive theory and a plethora of software available for manipulating state-space systems. By simply allowing the Δ block to represent more general system operators as described above, LFT systems provide a convenient framework for adding uncertainty in which essentially all of the major state-space results may then be generalized.

In much of the robust control literature, M represents the nominal system model, which is assumed to consist of a linear time-invariant transfer function for the plant plus system weighting functions on the inputs and outputs, and Δ represents the uncertainty. In the LFT models we consider, the transfer relations for the plant and weightings are also explicitly written as LFTs on the delay operator $\boldsymbol{\lambda}$; thus $\boldsymbol{\lambda}$ is included in the uncertainty strucure Δ. That is, we assume M is a constant matrix, and Δ represents both the system uncertainty and delay operators. In particular, we refer to M as the system realization matrix, and we assume Δ lies in a prescribed set $\boldsymbol{\Delta}$ defined by

$$\boldsymbol{\Delta} = \left\{\operatorname{diag}\left[\delta_1 I_{n_1}, \ldots, \delta_p I_{n_p}, \Delta_{m_1}, \ldots, \Delta_{m_f}\right] : \delta_i \in \mathcal{L}(l_2),\ \Delta_{m_i} \in \mathcal{L}(l_2^{m_i})\right\}, \tag{4.1}$$

where one of the δ_i represents $\boldsymbol{\lambda}$. We refer to the $\delta_i I_{n_i}$ as repeated scalar blocks, and to the Δ_j as full blocks.

The source of perturbations or uncertainty in the model may be due to unmodelled dynamics in the physical system, including nonlinearities and disturbances, or to parametric uncertainties, or any combination of the aforementioned. Because each perturbation source is likely to enter the real system at a different location, collecting these into one uncertainty block results in Δ having a diagonal block structure (see [74] and [23] for examples). Furthermore, the perturbations are often assumed to be norm-bounded operators, and frequently have additional structure, such as time-invariance or real parametric variance. For example, for analysis purposes we will often consider Δ which lie in a norm-bounded subset of $\boldsymbol{\Delta}$, that is,

$$\boldsymbol{B}_{\Delta} = \left\{\Delta \in \boldsymbol{\Delta} : \|\Delta\|_{l_2 \to l_2} \leq 1\right\}, \tag{4.2}$$

where $\|\cdot\|_{l_2-l_2}$ denotes the induced norm. We will denote these uncertain system models by the pair $(\boldsymbol{\Delta}, M)$.

We assume throughout the sequel that $y \in l_2^q$ and $u \in l_2^m$, although this assumption is not required for the results in this chapter. For notational convenience, dimensions will not be given in the sequel unless required for clarity.

Remark 4.1 *Redheffer was the first to consider the analysis of LFTs on structured sets [59], namely for the case of two full blocks (i.e., $p = 0$ and $f = 2$). As a result, the expression $(\Delta \star M)$ is sometimes referred to as the Redheffer star product.*

Repeated Scalar Uncertainty Structures

In this thesis, we focus on repeated scalar uncertainty sets, that is,

$$\boldsymbol{\Delta} = \begin{bmatrix} \delta_1 I_{n_1} & & & \\ & \delta_2 I_{n_2} & & \\ & & \ddots & \\ & & & \delta_p I_{n_p} \end{bmatrix}. \tag{4.3}$$

Many of the results we present are valid for the uncertainty structure of (4.1), with both repeated scalar and full uncertainty blocks; however, for the main results of this chapter (the reducibility results), the repeated scalar case is the more technically interesting case. Furthermore, for both the reducibility results of this chapter and the model reduction results of Chapter 5, the presentation of the full block case is notationally cumbersome. Thus the reducibility results of Section 4.4 are presented only for repeated scalar structures.

As we often consider inputs and outputs as signals in l_2, it is reasonable to view the δ_i as arbitrary time-varying operators on l_2, for example in the case of power systems, which can be modelled using non-periodic time-varying uncertainty. Alternatively, we may assume the δ_i represent real valued parametric uncertainty, for example, component tolerances, or we may assume the uncertainty block consists solely of multiple delay or shift operators corresponding to multiple transform variables; in the latter case, the LFT $(\Delta \star M)$ defines a multi-dimensional system.

Generally speaking, the more structure that is imposed on $\boldsymbol{\Delta}$, the more difficult computation for analysis and design becomes. If the only structure we assume for the uncertainty block is spatial (that is, repeated scalar block diagonal), then the LFT of a matrix M on $\boldsymbol{\Delta}$ reduces to a representation of rational functions in multiple noncommuting indeterminates. Such an LFT system may then be viewed as a particular realization of a formal power series (FPS) [8]. Connections between LFT and FPS realizations are discussed in Chapter 6.

For most of the results discussed in this chapter we take the most general assumptions for the uncertainty block, that is, we assume the δ_i are noncommuting variables, be they either completely abstract indeterminants in a power series or arbitrary time-varying operators on l_2. The results we obtain are then applicable to all of the aforementioned cases; to the more abstract settings as well as to systems with parametric uncertainty or multi-dimensional systems. Although this may lead to conservative conditions in the latter cases, results constructed for multi-dimensional realizations or parametric uncertainty would not be applicable in the more general settings.

4.2 Equivalent LFT Realizations and Allowable Transformations

Analogous to the standard state-space framework, given an LFT realization $(\boldsymbol{\Delta}, M)$, one way to obtain an equivalent realization is by applying a *structured* similarity transformation.

For repeated scalar uncertainty structures, we define equivalence as follows.

Definition 4.2 *Two realizations,*

$$\boldsymbol{\Delta}_1 = \left\{ \begin{bmatrix} \delta_1 I_{n_1} & & \\ & \ddots & \\ & & \delta_p I_{n_p} \end{bmatrix} : \delta_i \in \mathcal{L}(l_2) \right\}, \quad M_1 = \begin{bmatrix} A_1 & B_1 \\ C_1 & D \end{bmatrix}$$

$$\textit{and } \boldsymbol{\Delta}_2 = \left\{ \begin{bmatrix} \delta_1 I_{r_1} & & \\ & \ddots & \\ & & \delta_p I_{r_p} \end{bmatrix} : \delta_i \in \mathcal{L}(l_2) \right\}, \quad M_2 = \begin{bmatrix} A_2 & B_2 \\ C_2 & D \end{bmatrix}$$

are **equivalent** *if* $\Delta_1 \star M_1 = \Delta_2 \star M_2$ *for all* $\delta_i \in \mathcal{L}(l_2)$, $i = 1, \dots, p$.

This definition is quite easily generalized for uncertainty structures containing full blocks.

Similarity transformations are defined for LFT realizations exactly as in the standard case. However, in order for a transformed realization to be equivalent to the original realization, we require that the transformation commute with the uncertainty structure.

Definition 4.3 *Let* $n = \sum_{i=1}^{p} n_i + \sum_{j=1}^{f} m_j$. *The* **commutative matrix set** *for a given uncertainty set* $\boldsymbol{\Delta}$ *is denoted by* $\mathcal{T}$, *and defined by*

$$\mathcal{T} := \{T \in \mathbb{C}^{n \times n} : T\Delta = \Delta T, \ \ \textit{for all } \Delta \in \boldsymbol{\Delta}\}.$$

When $\boldsymbol{\Delta}$ is defined as in (4.1), the elements of $\mathcal{T}$ have the block diagonal structure $T = \operatorname{diag}[T_1, \dots, T_p, t_1 I_{m_1}, \dots, t_f I_{m_f}]$, where each $T_i \in \mathbb{C}^{n_i \times n_i}$ and $t_j \in \mathbb{C}$. We refer to a nonsingular element T in the set $\mathcal{T}$ as an *allowable transformation*. The following lemma results from a direct application of the definition of $(\Delta \star M)$, and holds for both repeated scalar and full block structures without any modifications required.

Lemma 4.4 *Given a LFT realization $(\mathbf{\Delta}, M)$, and any nonsingular $T \in \mathcal{T}$, denote*

$$\widehat{M} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix}.$$

Then $(\Delta \star M) = (\Delta \star \widehat{M})$ *for all* $\Delta \in \mathbf{\Delta}$.

Thus, allowable transformations lead to a change of coordinates for LFTs on structured uncertainty sets.

4.3 Lyapunov Inequalities and Balanced Realizations

In order to generalize the concepts of Lyapunov equations and Gramians to uncertain systems, we first discuss stability analysis of such systems.

Consider the system in Figure 4.1 with $\mathbf{\Delta}$ and M defined as in (4.1). We say such a system is stable when the map $(\Delta \star M)$ is well-defined for every $\Delta \in \mathbf{B}_\Delta$; precisely speaking, this is a robust l_2-stability condition which we will henceforth refer to simply as stability.

Definition 4.5 *Let* $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *be a constant matrix and* $\mathbf{\Delta} \subset \mathcal{L}(l_2)$. *The system defined by the pair* $(\mathbf{\Delta}, M)$ *is* **stable** *if* $(I - A\Delta)$ *is invertible in* $\mathcal{L}(l_2)$ *for each* $\Delta \in \mathbf{B}_\Delta$.

Stability Conditions for Uncertain Systems

If the only assumption placed on the uncertainty set $\mathbf{\Delta}$ is the spatial structure, that is, $\mathbf{\Delta}$ consists of full block and repeated scalar block (but otherwise arbitrary) linear operators on l_2, then a necessary and sufficient LMI stability condition exists, which was first given in [57], and is stated below in Theorem 4.6. A sketch of the proof is provided in Appendix A; details can be found in [57]. This condition extends the sufficient scaled small gain condition for robust stability, and recent results on the necessity of constant scalings for linear time-varying uncertainty obtained independently by Megretski [49, 50] and by Shamma [63] for full block diagonal uncertainty structures. Note that one of the δ_i may represent the delay operator $\boldsymbol{\lambda}$.

Theorem 4.6 *([57]) Given an uncertainty set,* $\mathbf{\Delta} \subset \mathcal{L}(l_2)$, *and a constant matrix* A,

$$(I - A\Delta) \text{ is invertible in } \mathcal{L}(l_2), \text{ for all } \Delta \in \mathbf{B}_\Delta$$

if and only if there exists a matrix $Y > 0$, $Y \in \mathcal{T}$ *such that*

$$AYA^* - Y < 0 \tag{4.4}$$

Proof. See Appendix A. ∎

Since this stability condition is defined in terms of A and $\boldsymbol{\Delta}$, we will sometimes say the matrix A is stable with respect to the Δ structure, meaning that the map $(\Delta \star M)$ is well-defined for all $\Delta \in \boldsymbol{B}_{\Delta}$.

Condition (4.4) directly extends the Lyapunov inequalities for stability of a standard state-space system. Using the LMI condition of (4.4), we can readily show the following stability condition for block structured realization matrices (see also Theorem 3.8 of [47]). This result is used in the sufficiency proof of the reducibility condition in Section 4.4, and again holds for both repeated scalar and full block structures without any modifications required.

Lemma 4.7 *Given a constant matrix A with an associated uncertainty structure Δ, where*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \text{ and } \Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix},$$

then there exists a matrix $Y > 0$, $Y \in \mathcal{T}$ satisfying $AYA^* - Y < 0$, if and only if there exists matrices $Y_1 > 0$, $Y_1 \in \mathcal{T}_1$ and $Y_2 > 0$, $Y_2 \in \mathcal{T}_2$ satisfying

$$A_{11}Y_1A_{11}^* - Y_1 < 0 \text{ and } A_{22}Y_2A_{22}^* - Y_2 < 0.$$

Proof. Necessity is immediate, thus we consider sufficiency; that is suppose we have $Y_1 \in \mathcal{T}_1$ and $Y_2 \in \mathcal{T}_2$, $Y_1,\ Y_2 > 0$, satisfying

$$A_{11}Y_1A_{11}^* - Y_1 < 0 \text{ and } A_{22}Y_2A_{22}^* - Y_2 < 0.$$

Let $\tilde{Y} = \operatorname{diag}[Y_1,\ Y_2]$. Consider the set of allowable transformations $T_\alpha = \operatorname{diag}[\alpha I,\ I]$, where $\alpha \in \mathbb{R}^+$, and denote $A_\alpha = T_\alpha A T_\alpha^{-1}$. Note that

$$A_0 = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix},$$

and therefore

$$A_0\tilde{Y}A_0^* - \tilde{Y} < 0.$$

Since the eigenvalues of $A_\alpha\tilde{Y}A_\alpha^* - \tilde{Y}$ are continuous functions of α, there exists a fixed $\overline{\alpha} > 0$ for which

$$A_{\overline{\alpha}}\tilde{Y}A_{\overline{\alpha}}^* - \tilde{Y} < 0.$$

Defining $Y = T_{\overline{\alpha}}^{-1}\tilde{Y}T_{\overline{\alpha}}^{-1}$ gives $AYA^* - Y < 0$. ∎

Remark: Stability Conditions for Multi-Dimensional Systems

For multi-dimensional systems, as well as systems that are modelled using constant parametric or LTI uncertainty, the stability condition given in Theorem 4.6 is sufficient, but not necessary. Alternatively, the structured singular value is both a necessary and sufficient stability criterion for these systems.

Definition 4.8 *The* **structured singular value**, *$\mu_{\boldsymbol{\Delta}}(M)$, of a matrix M with respect to a block structure $\boldsymbol{\Delta}$ is given by*

$$\mu_{\boldsymbol{\Delta}}(M) = \left(\min_{\Delta \in \boldsymbol{\Delta}} \{\overline{\sigma}(\Delta) : \; det(I - M\Delta) = 0\}\right)^{-1}. \tag{4.5}$$

Recent results in [14] and [67] have shown that the μ *recognition problem* - "is $\mu > c$?" for a given real positive scalar c - is NP-hard for $\boldsymbol{\Delta}$ sets containing real valued and/or LTI uncertainty blocks. It is generally accepted, although not proven, that a problem being NP-hard means it cannot be computed in polynomial time in the worst case. The difficulty in computing μ affects not only stability analysis, but also the computation of system norms and controller synthesis; upper and lower bounds for μ problems are computed instead. In this thesis, we use sufficient conditions for multi-dimensional systems, such as those in Theorem 4.6, which can be computed using convex optimization techniques. Furthermore, these conditions are sufficient for a larger class of uncertainties than is the structured singular value, and are both necessary and sufficient for LTV uncertainty.

Structured Gramians

An equivalent stability condition to that of (4.4) is the existence of a matrix $X > 0$, $X \in \mathcal{T}$ satisfying $A^*XA - X < 0$. By scaling Y and X by constant gains, we immediately obtain the following corollary to Theorem 4.6, generalizing the notion of Lyapunov equations for standard system realizations.

Corollary 4.9 *If $(\boldsymbol{\Delta}, M)$ is stable, then there exist $Y \geq 0$ and $X \geq 0$, both in $\mathcal{T}$, which satisfy the Lyapunov inequalities*

$$AYA^* - Y + BB^* \leq 0 \text{ and } A^*XA - X + C^*C \leq 0. \tag{4.6}$$

We refer to any matrices $Y \geq 0$ and $X \geq 0$ in $\mathcal{T}$ that satisfy (4.6) as *structured Gramians*, with the understanding that these are non-unique solutions to the inequalities of (4.6) and do not satisfy equations (3.9) as in the standard case. Inequalities are required rather than strict equalities as these Gramians must commute with the uncertainty structure; there is no guarantee that structured matrix solutions exist for the case of equalities.

Note that the LMIs in (4.4) and (4.6) are not affected by permutations of the realization matrices A, B, and C. For example, let Π be any matrix such that $\Pi\,\Pi^T = I$, and suppose we have a solution $Y \in \mathcal{T}$ to (4.4). Denote $\mathcal{A} = \Pi^T A \Pi$ and $\mathcal{Y} = \Pi^T P \Pi$. Then,

$$\mathcal{A}\mathcal{Y}\mathcal{A}^* - \mathcal{Y} = \Pi^T(AYA^* - Y)\Pi < 0.$$

Using such permutations, we can easily prove the following lemma.

Lemma 4.10 *Suppose*

$$A = \left[\begin{array}{cc|cc} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ \hline A_{31} & A_{32} & A_{33} & A_{34} \\ 0 & A_{42} & 0 & A_{44} \end{array}\right]$$

is stable with respect to the uncertainty structure

$$\Delta = \left[\begin{array}{c|c} \delta_1 I_{n_1} & 0 \\ \hline 0 & \delta_2 I_{n_2} \end{array}\right].$$

Then $\left[\begin{array}{c|c} A_{22} & A_{24} \\ \hline A_{42} & A_{44} \end{array}\right]$ *is stable with respect to the structure* $\Delta_r = \left[\begin{array}{c|c} \delta_1 I_{r_1} & 0 \\ \hline 0 & \delta_2 I_{r_2} \end{array}\right]$, *where* $r_i \le n_i$.

One permutation which leads to the result of Lemma 4.10 is

$$\Pi = \left[\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I \\ 0 & 0 & I & 0 \end{array}\right].$$

Similarly, we can show that $\left[\begin{array}{c|c} A_{11} & A_{13} \\ \hline A_{31} & A_{33} \end{array}\right]$ is stable with respect to the uncertainty structure $\Delta_r = \left[\begin{array}{c|c} \delta_1 I_{q_1} & 0 \\ \hline 0 & \delta_2 I_{q_2} \end{array}\right]$, where $q_i \le n_i$; additionally, $r_1 + q_1 = n_1$ and $r_2 + q_2 = n_2$.

Balanced Realizations

As in the standard case, balanced realizations for uncertain systems can be constructed by using a similarity transformation T to transform the realization M to $\widehat{M}$, and the structured Gramians Y and X to $\hat{Y}$ and $\hat{X}$, where $\hat{Y} = \hat{X} = \Sigma$. Naturally, we allow only similarity transformations that commute with the uncertainty structure, since they do not change the mapping from u to y, that is, $(\Delta \star M) = (\Delta \star \widehat{M})$ for all $\Delta \in \boldsymbol{\Delta}$. Furthermore, for any allowable transformation $T \in \mathcal{T}$, and structured Gramians Y and X, $\hat{Y} = TYT^* \in \mathcal{T}$ and $\hat{X} = (T^{-1})^* X T^{-1} \in \mathcal{T}$ are solutions to the Lyapunov inequalities for $\widehat{M}$.

Consider the system in Figure 4.1 where $\boldsymbol{\Delta}$ is specified as in (4.1); we define balanced realizations for uncertain systems as follows.

Definition 4.11 *Let $(\boldsymbol{\Delta}, M)$ be a stable, uncertain system realization. Then M is a* **balanced realization** *if there exists a diagonal matrix $\Sigma \geq 0$ such that*

$$A\Sigma A^* - \Sigma + BB^* \leq 0 \text{ and } A^*\Sigma A - \Sigma + C^*C \leq 0, \tag{4.7}$$

and where

$$\Sigma = diag[\Sigma_1, \cdots, \Sigma_p, \cdots, \Sigma_{p+f}] \tag{4.8}$$

with $\Sigma_i = diag[\sigma_{i1} I_{s_{i1}}, \cdots, \sigma_{it_i} I_{s_{it_i}}] \geq 0$; $\sigma_{i1} \geq \cdots \geq \sigma_{it_i}$ and the dimension of block Σ_i is $n_i = \sum_{j=1}^{t_i} s_{ij}$ for $i = 1, \ldots, p$; and $\Sigma_i = diag[\sigma_{i1} I_{s_{i1}}]$ with dimension $m_i = s_{i1}$ for $i = p+1, \ldots, p+f$.

Note that since we now consider Lyapunov *inequalities*, more than one balanced realization and accompanying balanced structured Gramian Σ are likely to exist. We will often refer to the elements σ_{ij} of Σ as *generalized singular values.*

The reducibility of an uncertain system realization can now be stated as a condition on the realization matrices A, B and C and solutions X and Y to the Lyapunov inequalities. Henceforth, we restrict the uncertainty set we consider to that containing only repeated scalar blocks.

4.4 A Necessary and Sufficient Reducibility Condition for Uncertain Systems

For standard 1D systems there is a well defined notion of minimality, or equivalently controllability and observability. In order to develop similar definitions for system models which incorporate uncertainty descriptions into the realizations, we first prove the following sufficient condition for exact reducibility, stated in Theorem 4.12. This condition provides the first step in the development of realization theory results for uncertain systems. Theorem 4.12 holds when the δ_i are defined by transform variables, norm-bounded real or complex perturbations, or time-varying operators on l_2, thus, this result is applicable to both multi-dimensional and uncertain system realizations.

Throughout this section we denote the *full* and *reduced* system realizations by

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \text{ and } M_r = \begin{bmatrix} A_r & B_r \\ C_r & D \end{bmatrix},$$

with corresponding repeated scalar uncertainty structures:

$$\mathbf{\Delta} = \left\{\text{diag}\left[\delta_1 I_{n_1}, \delta_2 I_{n_2}, \ldots, \delta_p I_{n_p}\right] : \ \delta_i \in \mathcal{L}(l_2)\right\}, \tag{4.9}$$

and

$$\mathbf{\Delta}_r = \mathcal{I}_r(\mathbf{\Delta}) = \left\{\text{diag }\left[\delta_1 I_{r_1}, \delta_2 I_{r_2}, \ldots, \delta_p I_{r_p}\right] : \delta_i \in \mathcal{L}(l_2)\right\} \tag{4.10}$$

where the notation $\mathcal{I}_r(\Delta)$ is used to emphasize that Δ_r represents a reduced or lower dimension copy of Δ, and is not an independent uncertainty structure; furthermore, $\dim(\Delta) > \dim(\Delta_r)$. The difference between the full and reduced realizations, the error system, is realized by

$$\tilde{E} = \begin{bmatrix} A & 0 & B \\ 0 & A_r & B_r \\ C & -C_r & 0 \end{bmatrix} \text{ and } \tilde{\mathbf{\Delta}} = \left\{\begin{bmatrix} \Delta & 0 \\ 0 & \Delta_r \end{bmatrix} : \Delta \in \mathbf{\Delta}, \Delta_r \in \mathbf{\Delta}_r\right\}.$$

Note that as a result of the dependence between $\mathbf{\Delta}$ and $\mathbf{\Delta}_r$, the commutative matrix set for $\tilde{\mathbf{\Delta}}$ includes matrices with the following block structure:

$$T = \begin{bmatrix} \text{diag}(T_i^n) & \text{diag}(T_i^{nr}) \\ \text{diag}(T_i^{rn}) & \text{diag}(T_i^r) \end{bmatrix}$$

where $\dim(T_i^n) = n_i \times n_i$, $\dim(T_i^{nr}) = n_i \times r_i$, $\dim(T_i^{rn}) = r_i \times n_i$ and $\dim(T_i^r) = r_i \times r_i$ for all $i = 1, \ldots, p$.

Theorem 4.12 *(Sufficiency): Given the stable system representation $(\mathbf{\Delta}, M)$, there exists a reduced representation, $(\mathbf{\Delta}_r, M_r)$, such that $(\tilde{\Delta} \star \tilde{E}) = 0$ for all $\tilde{\Delta} \in \mathbf{B}_{\tilde{\Delta}}$, if there exists singular $X \geq 0$ or $Y \geq 0$, both in $\mathcal{T}$, satisfying*

(i) $AYA^* - Y + BB^* \leq 0$

or

(ii) $A^*XA - X + C^*C \leq 0$.

Furthermore, max *$(\dim(\Delta) - \dim(\Delta)_r)$ is equal to the number of zero-valued eigenvalues of the product YX.*

Proof. Suppose there exists $Y \geq 0$ satisfying (i). (The proof for $X \geq 0$ satisfying (ii) is essentially the same, and therefore is not presented.)

Without loss of generality, we can assume $p = 2$, that is, that $\mathbf{\Delta} = \{\text{diag}[\delta_1 I_{n_1}, \delta_2 I_{n_2}] : \delta_i \in \mathcal{L}(l_2)\}$. The proof extends immediately to $p > 2$, either directly or by recursive application.

Suppose $Y = \begin{bmatrix} Y_1 & 0 \\ 0 & Y_2 \end{bmatrix}$ with $Y_1 > 0$ and $Y_2 \geq 0$, where Y_i has the same dimensions as I_{n_i}, $i = 1, 2$. (If $Y_1 \geq 0$, the proof is the essentially the same, but notationally more cumbersome). Furthermore, we can transform Y_2 to $\begin{bmatrix} \hat{Y}_2 & 0 \\ 0 & 0 \end{bmatrix}$ with $\hat{Y}_2 > 0$. We thus can assume Y_2 has this structure.

Partition the system matrices accordingly with respect to the structure of Y, that is,

$$A = \begin{bmatrix} A_1 & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} \quad \text{and } C = [C_1 \; C_2 \; C_3]$$

where A_1, B_1 and C_1 are dimensioned compatibly with Y_1; A_{22}, B_2 and C_2 are dimensioned compatibly with $\hat{Y}_2$; and A_{33}, B_3 and C_3 are dimensioned compatibly with the 0 submatrix of Y_2. Partition $\Delta \in \boldsymbol{\Delta}$ similarly so that $\Delta = \operatorname{diag}[\delta_1 I_{n_1}, \delta_2 I_{r_2}, \delta_2 I_{q_2}]$, where $r_2 + q_2 = n_2$, and r_2 is the dimension of $\hat{Y}_2$.

By assumption, $AYA^* - Y + BB^* \leq 0$. In particular,

$$[A_{31} \; A_{32} \; A_{33}] \begin{bmatrix} Y_1 & 0 & 0 \\ 0 & Y_2^1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} A_{31}^* \\ A_{32}^* \\ A_{33}^* \end{bmatrix} + B_3 B_3^* \leq 0$$

thus $A_{31} = 0$, $A_{32} = 0$, and $B_3 = 0$, since both $Y_1 > 0$ and $\hat{Y}_2 > 0$. Denote

$$\widetilde{A}_{11} = \begin{bmatrix} A_1 & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \; \widetilde{A}_{12} = \begin{bmatrix} A_{13} \\ A_{23} \end{bmatrix}, \; \widetilde{A}_{22} = A_{33}, \; \widetilde{B}_1 = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \; \widetilde{C}_1 = [C_1 \; C_2]$$

and $\widetilde{C}_2 = C_3$.

Let $M_r = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{B}_1 \\ \widetilde{C}_1 & D \end{bmatrix}$ and $\boldsymbol{\Delta}_r = \{\operatorname{diag}[\delta_1 I_{n_1}, \delta_2 I_{r_2}] : \delta_i \in \mathcal{L}(l_2)\}$. Construct the difference realization $(\widetilde{\Delta} \star \widetilde{E}) = (\Delta \star M) - (\Delta_r \star M_r)$, and use the similarity transformation

$$T_0 = \begin{bmatrix} I_{n_1+r_2} & 0 & -I_{n_1+r_2} \\ 0 & 0 & I_{n_1+r_2} \\ 0 & I_{q_2} & 0 \end{bmatrix}$$

so that the transformed difference system realization is

$$\widetilde{E} = \begin{bmatrix} \widetilde{A}_{11} & 0 & \widetilde{A}_{12} & 0 \\ 0 & \widetilde{A}_{11} & 0 & \widetilde{B}_1 \\ 0 & 0 & \widetilde{A}_{22} & 0 \\ \widetilde{C}_1 & 0 & \widetilde{C}_2 & 0 \end{bmatrix}$$

and

$$\widetilde{\boldsymbol{\Delta}} = \{\operatorname{diag}[\Delta_r, \Delta_r, \delta_2 I_{q_2}], \Delta_r \in \boldsymbol{\Delta}_r, \;\; \delta_2 \in \mathcal{L}(l_2)\}.$$

In order to show $(\widetilde{\Delta} \star \widetilde{E}) = (\Delta \star M) - (\Delta_r \star M_r) = 0$, we must first show both $\widetilde{A}_{11}$ and $\widetilde{A}_{22}$ are stable.

By assumption, $A = \begin{bmatrix} A_1 & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$ is stable with respect to the uncertainty set $\mathbf{\Delta}$, thus there exists

$$Q = \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & Q_{22} & Q_{23} \\ 0 & Q_{23}^* & Q_{33} \end{bmatrix} > 0 : \; A^*QA - Q < 0. \tag{4.11}$$

By extracting the upper left submatrix of (4.11) we see that $\widetilde{A}_{11}^*\widetilde{Q}\widetilde{A}_{11} - \widetilde{Q} < 0$, where $\widetilde{Q} = \text{diag}[Q_1, \; Q_{22}]$. Therefore, $\widetilde{A}_{11}$ is stable with respect to the uncertainty set $\mathbf{\Delta}_r$.

Similarly, we can show that $\widetilde{A}_{22}$ is stable with respect to the uncertainty structure $\delta_2 I_{q2}$ by considering the lower right submatrix of the matrix inequality $APA^* - P < 0$, where a solution

$$P = \begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_{22} & P_{23} \\ 0 & P_{23}^* & P_{33} \end{bmatrix} > 0$$

exists, also by stability of A.

Since both $\widetilde{A}_{11}$ and $\widetilde{A}_{22}$ are stable, $\widetilde{E}$ is stable with respect to the uncertainty set $\widetilde{\mathbf{\Delta}}$ by Lemma 4.7. Thus, it is then a routine calculation to show $\widetilde{\Delta} \star \widetilde{E} = 0$ for all $\widetilde{\Delta} \in \mathbf{B}_{\widetilde{\Delta}}$. ∎

By now proving that the existence of a singular LMI solution, X or Y, to the Lyapunov inequalities is necessary for a lower dimension realization to exist, we have a complete notion of reducibility for uncertain systems which is similar to that for 1D systems. The proof for the necessity condition is based on the following two lemmas, the first of which assumes the set $\mathbf{\Delta}$ consists of noncommutative elements δ_i. That is, this condition is not necessary for multi-dimensional systems, or systems with real or complex valued perturbations. Proofs for Lemmas 4.13 and 4.14 are given following the proof for Theorem 4.15.

Lemma 4.13 *Suppose the stable system realization $(\mathbf{\Delta}, M)$ is given, where $\mathbf{\Delta} \subset \mathcal{L}(l_2)$ is an arbitrary linear operator. If $(\Delta \star M) = 0$, for all $\Delta \in \mathbf{B}_\Delta$, then there exist $X \geq 0$ and $Y \geq 0$, both in $\mathcal{T}$, satisfying*

(i) $AYA^* - Y + BB^* \leq 0$

(ii) $A^*XA - X + C^*C \leq 0$ and

(iii) $XY = 0$.

Lemma 4.14 *Suppose $\widetilde{X} = \begin{bmatrix} X & X_1 \\ X_1^* & X_2 \end{bmatrix} \geq 0$ and $\widetilde{Y} = \begin{bmatrix} Y & Y_1 \\ Y_1^* & Y_2 \end{bmatrix} \geq 0$, where*

$$dim(X) = dim(Y) > dim(X_2) = dim(Y_2).$$

If $\widetilde{X}\widetilde{Y} = 0$, then either X or Y is singular.

We now state and prove the necessity condition.

Theorem 4.15 *(Necessity): Suppose the stable system realization* $(\mathbf{\Delta}, M)$ *is given. If there exists a reduced realization* $(\mathbf{\Delta}_r, M_r)$ *such that* $(\tilde{\Delta} \star \tilde{E}) = 0$*, for all* $\tilde{\Delta} \in \mathbf{B}_{\tilde{\Delta}}$*, then there exists singular* $X \geq 0$ *or* $Y \geq 0$*, both in* $\mathcal{T}$*, satisfying*

(i) $AYA^* - Y + BB^* \leq 0$

or

(ii) $A^*XA - X + C^*C \leq 0.$

Proof. Recall that the difference system $(\tilde{\Delta} \star \tilde{E}) = (\Delta \star M) - (\Delta_r \star M_r)$ is given by

$$\tilde{E} = \begin{bmatrix} A & 0 & B \\ 0 & A_r & -B_r \\ C & C_r & 0 \end{bmatrix}, \quad \tilde{\Delta} = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_r \end{bmatrix}.$$

By Lemma 4.13, if $(\tilde{\Delta} \star \tilde{E}) = 0$ then there exist $\tilde{X} \geq 0$ and $\tilde{Y} \geq 0$, both in $\tilde{\mathcal{T}}$ satisfying the Lyapunov inequalities for the uncertain system $(\tilde{\mathbf{\Delta}}, \tilde{E})$, and $\tilde{X}\tilde{Y} = 0$. Since $\tilde{X}$ and $\tilde{Y}$ commute with $\tilde{\Delta} \in \tilde{\mathbf{\Delta}}$, they have the structure

$$\tilde{X} = \begin{bmatrix} X & X_1 \\ X_1^* & X_2 \end{bmatrix} \text{ and } \tilde{Y} = \begin{bmatrix} Y & Y_1 \\ Y_1^* & Y_2 \end{bmatrix}$$

where X and Y commute with $\Delta \in \mathbf{\Delta}$. Then $X \geq 0$ and $Y \geq 0$ satisfy the Lyapunov inequalities for the uncertain system $(\mathbf{\Delta}, M)$, and by Lemma 4.14 either X or Y is singular. ■

Proofs: Lemma 4.13 and Lemma 4.14

The following proof relies on expanding the LFT $(\Delta \star M)$ as a formal power series; further discussion of formal power series representations and connections to LFT realizations is given in Chapter 6; this proof we developed for Lemma 4.13, which is based on using 2-norms of $(\Delta \star M)$ rather than a series expansion, is given in Appendix B. The original proof is much longer than that presented here, but leads more directly to an unobservable - uncontrollable type of decomposition structure.

Proof of Lemma 4.13:

Consider the series $S = D + \Delta \star M = \sum_{k=0}^{\infty} C\Delta(A\Delta)^k B$. We first partition the matrices A, B, and C accordingly with the Δ structure, that is,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix}; \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix}; \quad C = \begin{bmatrix} C_1 & \cdots & C_p \end{bmatrix}.$$

Then, expanding the series S gives

$$S = D + \sum_{i=1}^{p} C_i B_i \delta_i + \sum_{k=0}^{\infty} \sum_{i_0,\dots,i_k=1}^{p} C_{i_k} A_{i_k i_{k-1}} \cdots A_{i_1 i_0} B_{i_0} \delta_{i_k} \cdots \delta_{1_0},$$

where, by assumption, $S = 0$.

Since the δ_i are noncommuting, S is identically zero if and only if each coefficient in the series is zero, that is,

$$D = 0,\ C_i B_i = 0,\ C_i A_{ij} B_j = 0,\ C_i A_{ij} A_{jk} B_k = 0, \dots,$$

for every $i, j, k = 1, \dots, p$. We consider each set of terms separately and show that the given realization may be transformed to one having a particular decomposition structure, similar to the decomposition structure given for 1D systems in equation (3.6).

First, consider the $C_i B_i$ terms: $C_i B_i = 0$ if and only if there exists T_i nonsingular, such that

$$C_i T_i^{-1} = [0\ 0\ \tilde{C}]_i \text{ and } T_i B_i = \begin{bmatrix} \tilde{B} \\ 0 \\ 0 \end{bmatrix}_i,$$

where $\tilde{C}_i$ has full column rank, $\tilde{B}_i$ has full row rank, and the submatrices of C_i and B_i are equivalently partitioned, for each $i, j = 1, \dots, p$. We will henceforth absorb all such transformations and assume the realization matrices are already structured into zero and non-zero block submatrices. We then partition each A_{ij} accordingly with the partitions of C_i and B_j.

Consider the $C_i A_{ij} B_j$ terms:

$$C_i A_{ij} B_j = [0\ 0\ \tilde{C}]_i \begin{bmatrix} A^{11} & A^{12} & A^{13} \\ A^{21} & A^{22} & A^{23} \\ A^{31} & A^{32} & A^{33} \end{bmatrix}_{ij} \begin{bmatrix} \tilde{B} \\ 0 \\ 0 \end{bmatrix}_j = \tilde{C}_i A_{ij}^{31} \tilde{B}_j = 0.$$

Since $\tilde{C}_i$ and $\tilde{B}_j$ are both full rank, this implies that $A_{ij}^{31} = 0$.

Next consider the $C_i A_{ij} A_{jl} B_l$ terms:

$$C_i A_{ij} A_{jl} B_l = [0\ \tilde{C}_i A_{ij}^{32}\ \tilde{C}_i A_{ij}^{33}] \begin{bmatrix} A_{jl}^{11} \tilde{B}_l \\ A_{jl}^{21} \tilde{B}_l \\ 0 \end{bmatrix} = 0 \text{ if and only if } A_{ij}^{32} A_{jl}^{21} = 0. \tag{4.12}$$

As with the C_iB_i terms, we can transform A^{32}_{ij} to $[0\ 0\ \tilde{A}^{32}_{ij}]$ and A^{21}_{jl} to $\begin{bmatrix} \tilde{A}^{21}_{jl} \\ 0 \\ 0 \end{bmatrix}$ where $\tilde{A}^{32}_{ij}$ has full column rank, $\tilde{A}^{21}_{jl}$ has full row rank, and the submatrices of A^{32}_{ij} and A^{21}_{jl} are equivalently partitioned. We now have the following decomposition structure for each A_{ij}:

$$A_{ij} = \left[\begin{array}{ccccc} A^{11} & \vdots & A^{12} & \vdots & A^{13} \\ \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ \tilde{A}^{21} & & & & \\ 0 & \vdots & A^{22} & \vdots & A^{23} \\ 0 & & & & \\ \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ 0 & \vdots & 0\ \ 0\ \ \tilde{A}^{32} & \vdots & A^{33} \end{array}\right]. \tag{4.13}$$

Note that across each block row, indexed by for example i, the A^{21} blocks in the set of submatrices $\{A_{ij}\}_{j=1}^{p}$ are equivalently partitioned; similarly down each block column indexed by j, the A^{32} blocks in the set of submatrices $\{A_{ij}\}_{i=1}^{p}$ are also equivalently partitioned.

Evaluating the next few sets of series coefficients, $C_iA_{ij}A_{jl}A_{lw}B_w$, ..., we obtain the same decomposition structure for each of the A^{22}_{ij} blocks as that in (4.13). This process is repeated for a finite number of series coefficients, leading to a decomposition structure for each C_j, A_{ij}, B_i subsystem; this finite number depends on the number of variables, p, and the dimensions of the realization matrices. As an example, for $p = 2$, the resulting decomposition can be generally written as follows:

$$M = \left[\begin{array}{ccccc|c} \hat{A}^{11}_{11} & \hat{A}^{12}_{11} & \vdots & \hat{A}^{11}_{12} & \hat{A}^{12}_{12} & \hat{B}_1 \\ 0 & \hat{A}^{22}_{11} & \vdots & 0 & \hat{A}^{22}_{12} & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{A}^{11}_{21} & \hat{A}^{12}_{21} & \vdots & \hat{A}^{11}_{22} & \hat{A}^{12}_{22} & \hat{B}_2 \\ 0 & \hat{A}^{22}_{21} & \vdots & 0 & \hat{A}^{22}_{22} & 0 \\ \hline 0 & \hat{C}_1 & \vdots & 0 & \hat{C}_2 & 0 \end{array}\right]. \tag{4.14}$$

Note that the matrix partitions in (4.14) (denoted by $\hat{}$ s) do not necessarily correspond to the previous partitions (denoted by $\tilde{}$ s). Note also that the above partitions are constructed such that the $\hat{A}^{11}_{ii}$ and $\hat{A}^{22}_{ii}$ submatrices are square, and the lower left submatrix of each $\hat{A}_{ii}$ which is identically zero has the largest dimensions possible.

Applying Lemma 4.10, and carrying out the matrix multiplications, it is straightforward to see there exist structured singular matrices

$$\hat{X} = \mathrm{diag}[0, \hat{X}^{22}_1, 0, \hat{X}^{22}_2, \ldots, 0, \hat{X}^{22}_p] \geq 0 \text{ and } \hat{Y} = \mathrm{diag}[\hat{Y}^{11}_1, 0, \hat{Y}^{11}_2, 0, \ldots, \hat{Y}^{11}_p, 0] \geq 0$$

both in $\mathcal{T}$, satisfying

$$A\hat{Y}A^* - \hat{Y} \leq 0 \text{ and } A^*\hat{X}A - \hat{X} \leq 0.$$

Furthermore, scaling $\hat{Y}$ and $\hat{X}$ by constants, as necessary, gives $X \geq 0$ and $Y \geq 0$ both in $\mathcal{T}$, such that conditions (i) and (ii) are satisfied, and clearly $XY = 0$.

■

In the proof of Lemma 4.13, it is implicitly assumed that the uncertainty structure is $\boldsymbol{\Delta} = \left\{\text{diag}\left[\delta_1 I_{n_1}, \ldots, \delta_p I_{n_p}\right]\right\}$. The result can immediately be extended to uncertainty structures such as $\tilde{\boldsymbol{\Delta}} = \left\{\text{diag}\left[\delta_1 I_{n_1}, \ldots, \delta_p I_{n_p}, \delta_1 I_{r_1}, \ldots, \delta_p I_{r_p}\right]\right\}$ by permuting $\tilde{\boldsymbol{\Delta}}$ to $\hat{\boldsymbol{\Delta}} = \left\{\text{diag}\left[\delta_1 I_{n_1+r_1}, \ldots, \delta_p I_{n_p+r_p}\right]\right\}$, and also permuting $\tilde{E}$, Y, and X accordingly. We then can show that Lemma 4.13 holds for $(\tilde{\boldsymbol{\Delta}}, \tilde{E})$ with Y and X in $\tilde{\mathcal{T}}$ by permuting back to the original structures.

Proof of Lemma 4.14:

Let $\dim(X) = \dim(Y) = n \times n$ and $\dim(X_2) = \dim(Y_2) = k \times k$, where $k < n$. By contradiction, suppose both X and Y are nonsingular. Then XY is nonsingular and $\text{rank}(XY) =$ n.

Note that $\tilde{X}\tilde{Y} = 0$ implies that $XY = -X_1 Y_1^*$, thus $\text{rank}(X_1 Y_1^*) = n$, where $\dim(X_1) = n \times k$ and $\dim(Y_1^*) = k \times n$. However,

$$\text{rank}(X_1 Y_1^*) \leq \min\{\text{rank}(X_1),\ \text{rank}(Y_1^*)\} \leq k < n.$$

Thus X or Y is singular and $\text{rank}(XY) \leq k$. ■

Remark 4.16 *We may consider reducibility of realizations with uncertainty structures containing full blocks, that is, $\boldsymbol{\Delta}$ as defined in (4.1). However, the submatrices Y_j and X_j of Y and X, respectively, corresponding to the full blocks, Δ_j, are diagonal scaling blocks, for example, $Y_j = y_j I_{m_j}$, $y_j \in \mathbb{R}$. Clearly, if Y_j is singular, $y_j = 0$. In this case, in order for the Lyapunov inequalities to hold, entire sub-blocks of the realization matrices A, B and C will be zero and the result obvious.*

The results of Theorems 4.12 and 4.15 imply that, given an uncertain or multidimensional system representation, if structured singular solutions to either of a pair of LMIs can be found, then an equivalent lower dimension realization exists. Furthermore, if the uncertainty can be properly described by time-varying, or noncommuting, operators on l_2, then the existence of lower dimension realizations *requires* such singular LMI solutions. The development of computational methods for solving such LMI problems is a popular research area in the control community, and, in fact, many efficient convex optimization algorithms exist (see [12] and the references therein). The fact that we would like to find singular solutions to these LMIs complicates the computational requirements in that the complete set of constraints results in an optimization problem which is not convex. A suboptimal computational solution is presented in Chapter 7.

Decomposition Structures for Uncertain Realizations

Via the sufficiency proof (that is, the proof for Theorem 4.12), it is clear that the existence of a singular structured Gramian implies that an equivalent realization can be found which has a decomposition structure like that in (3.6). For example, consider the realization

$$\boldsymbol{\Delta} = \left\{\operatorname{diag}\left[\delta_1 I_{n_1}, \delta_2 I_{n_2}\right] : \ \delta_i \in \mathcal{L}(l_2)\right\}, \ M = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & 0 \end{array}\right],$$

and suppose structured Gramians $X = \begin{bmatrix} X_1 & \\ & X_2 \end{bmatrix} \geq 0$ and $Y = \begin{bmatrix} Y_1 & \\ & Y_2 \end{bmatrix} \geq 0$ are found where X_1, X_2, Y_1 and Y_2 are all singular. Then, as in the sufficiency proof, we can find an equivalent realization $\widehat{M}$, by allowable transformation, which has the decomposition structure:

$$\widehat{M} = \left[\begin{array}{cccccccccc|c}
\hat{A}_{11}^{co} & 0 & \hat{A}_{11}^{13} & 0 & \vdots & \hat{A}_{12}^{co} & 0 & \hat{A}_{12}^{13} & 0 & & \hat{B}_1^{co} \\
\hat{A}_{11}^{21} & \hat{A}_{11}^{c} & \hat{A}_{11}^{23} & \hat{A}_{11}^{24} & \vdots & \hat{A}_{12}^{21} & \hat{A}_{12}^{c} & \hat{A}_{12}^{23} & \hat{A}_{12}^{24} & & \hat{B}_1^{c} \\
0 & 0 & \hat{A}_{11}^{o} & 0 & \vdots & 0 & 0 & \hat{A}_{12}^{o} & 0 & & 0 \\
0 & 0 & \hat{A}_{11}^{43} & \hat{A}_{11}^{44} & \vdots & 0 & 0 & \hat{A}_{12}^{43} & \hat{A}_{12}^{44} & & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\
\hat{A}_{21}^{co} & 0 & \hat{A}_{21}^{13} & 0 & \vdots & \hat{A}_{22}^{co} & 0 & \hat{A}_{22}^{13} & 0 & & \hat{B}_2^{co} \\
\hat{A}_{21}^{21} & \hat{A}_{21}^{c} & \hat{A}_{21}^{23} & \hat{A}_{21}^{24} & \vdots & \hat{A}_{22}^{21} & \hat{A}_{22}^{c} & \hat{A}_{22}^{23} & \hat{A}_{22}^{24} & & \hat{B}_2^{c} \\
0 & 0 & \hat{A}_{21}^{o} & 0 & \vdots & 0 & 0 & \hat{A}_{22}^{o} & 0 & & 0 \\
0 & 0 & \hat{A}_{21}^{43} & \hat{A}_{21}^{44} & \vdots & 0 & 0 & \hat{A}_{22}^{43} & \hat{A}_{22}^{44} & & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\
\hat{C}_1^{co} & 0 & \hat{C}_1^{o} & 0 & \vdots & \hat{C}_2^{co} & 0 & \hat{C}_2^{o} & 0 & & 0
\end{array}\right]. \tag{4.15}$$

Note also that in the necessity proofs (that is, both the proof for Theorem 4.15 given in this chapter, and the alternative proof given in Appendix B), a decomposition structure is constructed for the error system $(\tilde{\boldsymbol{\Delta}}, \tilde{E})$; this suggests that iteratively solving for allowable transformations will provide a means to find a decomposition structure for any reducible realization, and hence to the computation of a *minimal* realization.

4.5 Minimality

One notable result which follows immediately from the proof for Theorem 4.15 is that all minimal realizations for an uncertain system may be obtained by allowable transformations and truncations, where we define minimal as follows:

Definition 4.17 *A realization* $(\boldsymbol{\Delta}, M)$ *is* **minimal** *if* $dim(\Delta)$ *is lowest among all equivalent realizations.*

The minimality result is stated in the following corollary.

Corollary 4.18 *Given a stable system realization* $(\boldsymbol{\Delta}, M)$*, all minimal realizations are found by similarity transformations,*

$$\widehat{M} = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix} : T \in \mathcal{T},$$

and truncations.

If we consider the simplest case for these LFT representations of uncertain systems, that is, there is no uncertainty and $\Delta = \lambda I$, we obtain the standard results, excepting the inequalities in the Lyapunov equations. At the other extreme, in the case of formal power series, we do not even have the operator structure for Δ, as the δ_i are simply noncommuting indeterminates. Thus stability and norms have no meaning, and indeed are artificial in the context of realization theory. There are many ways to extend the LFT machinery to this case, but the simplest way to remove the stability requirements for the A matrix is by scaling. Note that there will always be some value $\gamma > 0$ sufficiently small (for example, $\gamma < 1/\|A\|$) such that there exists a matrix $Y > 0$, $Y \in \mathcal{T}$ satisfying

$$\gamma^2 AYA^* - Y < 0. \tag{4.16}$$

The input/output map $(\Delta \star M)$ is then well-defined on l_2 for every

$$\Delta \in \mathbf{B}_\gamma \boldsymbol{\Delta} = \left\{\Delta \in \boldsymbol{B}_\Delta : \|\Delta\|_{l_2 \to l_2} \leq \gamma\right\},$$

and solutions $Y \geq 0$ and $X \geq 0$, both in $\mathcal{T}$ can be found satisfying the LMIs

$$\gamma^2 AYA^* - Y + BB^* \leq 0 \text{ and } \gamma^2 A^* XA - X + C^*C \leq 0.$$

If singular X and Y can be found, the realization is then reducible, as in Theorem 4.12, with respect to the uncertainty set $\mathbf{B}_\gamma\boldsymbol{\Delta}$, and further reducing the value of γ does not effect the existence of singular solutions to these LMIs. Note that this γ scaling illuminates the fact that the reducibility result may be viewed as a topological result, that is, that $(\Delta \star M \; - \; \Delta_r \star M_r)$ is the zero operator for all operators Δ in a neighborhood of zero if and only if there exist singular structured solutions to the Lyapunov inequalities.

4.6 Additional Realization Theory

Up to this point, we have given reducibility conditions for a given realization in terms of structured Gramians, without any discussion of controllability and observability, or a direct test for minimality. Naturally, we would like to develop generalizations of the standard controllability and observability matrices, and determine the relation these matrices would have not only to the structured Gramians, but also to a direct notion of minimality. In this section, we remark on these and related topics. We state the results that have been generalized and discuss the interpretations that can be given to these notions for uncertain systems.

Controllability and Observability Matrices

If we naively define the controllability and observability matrices, Γ and $\mathcal{O}$, as in (3.4) and (3.5), then we cannot directly generalize the relationships between the ranks of Γ and $\mathcal{O}$ to reducibility via singular structured Gramians, as the structure of the uncertainty is not taken into account. For example, suppose we take Γ to be defined as in 3.4, and denote $N = \sum_{i=1}^{p} n_i$. If there exists a singular structured Gramian $Y \geq 0$, then it is easy to see that $\operatorname{rank}(\Gamma) < N$ by considering the decomposition structure of (4.15). However, we cannot necessarily say the converse is true.

Consider the following example:

$$A = \begin{bmatrix} -0.11 & 0.04 \\ -0.48 & 0.17 \end{bmatrix}, \; B = \begin{bmatrix} 0.10 \\ 0.30 \end{bmatrix}, \text{ and } \boldsymbol{\Delta} = \left\{ \begin{bmatrix} \delta_1 & \\ & \delta_2 \end{bmatrix} : \delta_i \in \mathcal{L}(l_2) \right\}.$$

Then, $\Gamma = \begin{bmatrix} 0.100 & 0.001 \\ 0.300 & 0.003 \end{bmatrix}$ has rank 1, but there is no singular $Y = \begin{bmatrix} y_1 & \\ & y_2 \end{bmatrix} \geq 0$ satisfying $AYA^* - Y + BB^* \leq 0$.

As a more reasonable generalization of the realization theory results associated with the controllability and observability matrices, we propose the following constructions, which do account for the inherent structure of the system realization:

Definition 4.19 *Given an uncertain system realization $(\boldsymbol{\Delta}, M)$, where $\boldsymbol{\Delta}$ is structured as in (4.9), then the* **controllability matrix** *is defined by*

$$\Gamma = \begin{bmatrix} B_1 & A_{11}B_1 & \cdots & A_{1p}B_p & A_{11}^2 B_1 & \cdots & A_{11}A_{1p}B_p & A_{12}A_{21}B_1 & \cdots \\ B_2 & A_{21}B_1 & \cdots & A_{2p}B_p & A_{21}A_{11}B_1 & \cdots & A_{21}A_{1p}B_p & A_{22}A_{21}B_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ B_p & A_{p1}B_1 & \cdots & A_{pp}B_p & A_{p1}A_{11}B_1 & \cdots & A_{p1}A_{1p}B_p & A_{p2}A_{21}B_1 & \cdots \end{bmatrix}.$$

Furthermore, we denote the block rows by $\Gamma_i = [B_i \; A_{i1}B_1 \; \cdots]$.

The partitioning of this controllability matrix into block rows is similar to the partitioning proposed for 2-dimensional system controllability matrices when a relationship to minimality is desired (see [10] and the references therein). However the block elements of the above controllability matrix are different than in the 2D case (or multi-dimensional case), as the system variables represented by the δ_i in the LFT realizations we consider are assumed to be noncommuting. As a result, there is not addition of the terms in the matrices along block rows and columns.

We can now show that the following relation holds between singular structured Gramians and rank conditions on Γ.

Theorem 4.20 *Given a stable uncertain system realization* $(\boldsymbol{\Delta}, M)$, *where* $\boldsymbol{\Delta}$ *is defined as in (4.9), then there exists a singular* $Y \in \mathcal{T}$, $Y \geq 0$ *satisfying* $AYA^* - Y + BB^* \leq 0$ *if and only if rank*$(\Gamma_i) < n_i$ *for some* $i = 1, \ldots, p$.

Proof. Necessity is quite straightforward: If there exists a singular structured controllability Gramian Y, then, as in the proof of Theorem 4.12, using allowable transformations we can construct a decomposition structure similar to that in (4.15) from which it is easy to see that the block rows of the controllability matrix will have reduced rank.

Sufficiency can be shown using an approach similar to that taken for the proof of Lemma 4.13. For each $i = 1, \ldots, p$, denote the rank(Γ_i) by r_i, where we assume $r_i < n_i$ for at least one i. Then there exist nonsingular matrices $T_i \in \mathbb{C}^{n_i \times n_i}$ such that

$$T_i \Gamma_i = \begin{bmatrix} \tilde{\Gamma}_i \\ 0 \end{bmatrix}$$

where $\tilde{\Gamma}_i$ has r_i rows. This implies that

$$T_i B_i = \begin{bmatrix} \tilde{B}_i \\ 0 \end{bmatrix}, \text{ and } T_i A_{ij} B_j = T_i A_{ij} T_j^{-1} \begin{bmatrix} \tilde{B}_j \\ 0 \end{bmatrix}$$

has the form $\begin{bmatrix} \times \\ 0 \end{bmatrix}$, and as a result

$$T_i A_{ij} T_j^{-1} = \begin{bmatrix} \tilde{A}^{11} & \tilde{A}^{12} \\ 0 & \tilde{A}^{22} \end{bmatrix},$$

where the zero block has dimensions $(n_i - r_i) \times r_j$.

Using the allowable transformation $T = \text{diag}[T_1, T_2, \ldots, T_p]$ on the given realization matrix M yields an equivalent realization with an uncontrollable-like decomposition structure similar to that of (4.15). Following the same steps listed at the end of the proof of Lemma 4.13, we can construct a singular structured controllablity Gramian. ■

The obvious dual definition for the observability matrix can be given, along with the corresponding rank condition which can be proven using a similar argument:

Definition 4.21 *Given an uncertain system realization* $(\boldsymbol{\Delta}, M)$*, where* $\boldsymbol{\Delta}$ *is structured as in (4.9), then the* **observability matrix** *is defined by*

$$\mathcal{O} = \begin{bmatrix} C_1 & C_2 & \cdots & C_p \\ C_1A_{11} & C_1A_{12} & \cdots & C_1A_{1p} \\ \vdots & \vdots & \vdots & \vdots \\ C_pA_{p1} & C_pA_{p2} & \cdots & C_pA_{pp} \\ C_1A_{11}^2 & C_1A_{11}A_{12} & \cdots & C_1A_{11}A_{1p} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Furthermore, we denote the block columns by $\mathcal{O}_i = \begin{bmatrix} C_i \\ C_1A_{1i} \\ C_2A_{2i} \\ \vdots \end{bmatrix}$.

Theorem 4.22 *Given a stable uncertain system realization* $(\boldsymbol{\Delta}, M)$*, where* $\boldsymbol{\Delta}$ *is defined as in (4.9), then there exists a singular* $X \in \mathcal{T}, X \geq 0$ *satisfying* $A^*XA - X + C^*C \leq 0$ *if and only if* rank$(\mathcal{O}_i) < n_i$ *for some* $i = 1, \ldots, p$.

Moreover, we conjecture that these rank tests can be completed on finite dimension controllability and observability matrices, where the maximum dimensions are determined by the number of copies of each δ_i in the structure Δ. Finite dimension results exist for 2D system realizations [10], and for representations of power series over multiple noncommuting indeterminants [65].

Stabilizability and Detectability

For a standard state-space system, the use of static state-feedback or static output-injection for system stabilization is a well-known theoretical tool. In order to find state-feedback or output-injection matrices stabilizing a given realization, certain conditions must be satisfied, namely the realization must be stabilizable or detectable. If these conditions are met, a stabilizing feedback controller can be constructed from the state-feedback and a state-observer, and all stabilizing controllers can be parameterized. The extension of these concepts to uncertain systems have been completed by Lu, et al., [47]. We state the stabilizabililty and detectability results here, along with connections to PBH tests for uncertain systems, which have been developed by Paganini [57].

Consider the LFT system of Figure (4.1), where $\boldsymbol{\Delta}$ contains arbitrary time-varying structured operators. First note that if $(\boldsymbol{\Delta}, M)$ is stabilizable by a static state-feedback matrix F, then by Theorem 4.6 there exists a matrix $Y \in \mathcal{T}$, $Y > 0$ such that

$$(A + BF)Y(A + BF)^* - Y < 0.$$

If $\operatorname{rank}(B) = P < \dim(\Delta) = N$, and if we find a matrix $B_\perp \in \mathbb{R}^{N\times(N-P)}$ such that $B^* B_\perp = 0$ and $\operatorname{rank}(B_\perp) = N - P$, then an equivalent condition for static state-feedback stabilizability is

$$B_\perp^* AYA^* B_\perp - B_\perp^* Y B_\perp < 0.$$

Continuing this development leads to the following Theorem and Definition.

Theorem 4.23 *For the LFT system of Figure (4.1), where $\boldsymbol{\Delta}$ contains arbitrary time-varying structured operators, the following are equivalent:*

(i) There exists a static feedback matrix F such that $A + BF$ is stable.

(ii) There exists a matrix $Y \in \mathcal{T}$ with $Y = Y^* > 0$ such that

$$B_\perp^* AYA^* B_\perp - B_\perp^* Y B_\perp < 0. \tag{4.17}$$

(iii) There exists a matrix $Y \in \mathcal{T}$ with $Y = Y^* > 0$ such that

$$AYA^* - Y - BB^* < 0. \tag{4.18}$$

(iv) The map $[I - A\Delta \; B] : l_2^{n+m} \to l_2^n$ is surjective for all $\Delta \in \mathbf{B}_\Delta$.

Moreover, if $Y \in \mathcal{T}$ with $Y = Y^* > 0$ satisfies any of the above LMIs, then one such stabilizing static state-feedback matrix is

$$F = -(B^* Y^{-1} B)^{-1} B^* Y^{-1} A. \tag{4.19}$$

The equivalence of (i), (ii) and (iii) is shown in [47]; the equivalence with the PBH-like condition (iv) follows from the related result in [57]. This leads to the following definition (referred to as $\mathcal{Q}$-stabilizability in [47]):

Definition 4.24 *If any of the conditions $(i) - (iv)$ of Theorem 4.23 is satisfied, we say that the LFT system is* **stabilizable**.

In this case, the analogy with the standard case is complete. The dual notion of *detectability* can be characterized in the same manner [47].

Reachable and Observable Subspaces

Straightforward generalizations of reachable and observable subspaces for uncertain system realizations are presented by D'Andrea in [58], where the LFTs are now restricted to be causal operators. We summarize these results in the following.

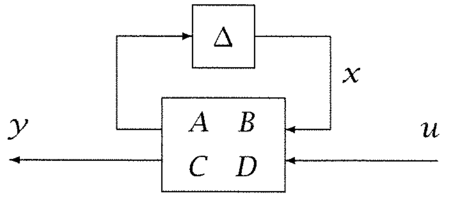

Figure 4.2: LFT Realization

Consider the LFT system in Figure 4.2, where $\Delta \in \boldsymbol{\Delta}$ is a causal linear operator with the structure defined by (4.9), and $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{R}^{n+q \times n+m}$. First consider the corresponding system equations:

$$\begin{aligned} x &= \Delta(Ax + Bu) \\ x(0) &= 0. \end{aligned} \tag{4.20}$$

The reachable set of M with respect to a fixed operator $\Delta \in \boldsymbol{\Delta}$ is denoted by $\mathcal{R}_{(\Delta,M)}$ and defined by:

$$\mathcal{R}_{(\Delta,M)} := \{\overline{x} \in \mathbb{R}^n | \exists u \in l_2, k \in \mathbb{Z}^+, \text{ such that } x(k) = \overline{x} \text{ satisfies} (4.20)\}. \tag{4.21}$$

Due to the possible time variation of Δ, $\mathcal{R}_{(\Delta,M)}$ may not be a vector space. The following definition is from [58].

Definition 4.25 *The* **reachable set** *of M with respect to the set $\boldsymbol{\Delta}$ is denoted by $\mathcal{R}_{(\boldsymbol{\Delta},M)}$ and defined by:*

$$\mathcal{R}_{(\boldsymbol{\Delta},M)} := \bigcup_{\Delta \in \boldsymbol{\Delta}} \mathcal{R}_{(\Delta,M)}. \tag{4.22}$$

As $\mathcal{R}_{(\boldsymbol{\Delta},M)}$ consists of all possible vectors that may be reached in finite time, it is straightforward to show that $\mathcal{R}_{(\boldsymbol{\Delta},M)}$ is a vector space, referred to as the **reachable subspace** of $(\boldsymbol{\Delta}, M)$. In [58] it is also shown that there exists a $\overline{\Delta} \in \boldsymbol{\Delta}$ such that $\mathcal{R}_{(\overline{\Delta},M)} = \mathcal{R}_{(\boldsymbol{\Delta},M)}$, and that an allowable transformation T_r exists which results in a decomposition of M into reachable and unreachable subsystems.

A similar development is presented in [58] for defining and constructing the *unobservable subspace* of $(\boldsymbol{\Delta}, M)$.

An alternate method to that presented in this thesis for constructing a minimal realization is given in [58] which based on decomposing the original realization into reachable and unobservable subspaces; the definition used for minimal is that each separate block dimension of Δ, that is each n_i, is lowest among all equivalent realizations. The resulting decomposition structure is the same as that in (4.15). This method may be applied to realizations that are not stable. However, the procedure relies on finding reduced rank matrices, which may be numerically ill-conditioned. Furthermore, no related error bounds can be computed when exact reductions cannot be found; thus, we prefer to solve the Lyapunov inequalities for structured Gramians, which allows us to directly compute reduced realizations with error bounds when exact reduction cannot be completed (see Chapter 5).

In this chapter, we have presented a fairly thorough treatment of the exact reduction of uncertain system realizations, via the existence of structured singular Gramians, and touched on related realization theory topics. To further complete the discussion on the reduction of uncertain system realizations, in the following chapter we present error bounds for the reduction of uncertain systems when *singular* structured Gramians cannot be found, that is, when the reduction is not exact.

Chapter 5

Model Reduction of Uncertain Systems

As discussed in Chapter 4, the most general way to interpret LFT models is to view the δ_i as noncommuting indeterminants. However, to quantitatively evaluate model reduction methods and their associated error bounds, we need a less abstract setting. In particular, we may consider one of the δ_i, say δ_1, as the delay operator in an uncertain discrete-time system. The remaining δ_i are then viewed as time-varying operators on l_2. Alternatively, we may view the δ_i as different transform variables in a multi-dimensional system, or assume the uncertainty itself has additional structure, such as time-invariance or real parametric variance. In these contexts, we may reduce a realization for an uncertain or multi-dimensional system without finding *singular* structured Gramians, and assess the difference between the full and reduced models using induced norms. In the case of uncertain systems, model reduction implies simplification of the uncertainty descriptions, whereas in the case of multi-dimensional systems, model reduction means state order reduction, and a system model that may be reduced without error as discussed in the preceding chapter is reducible or non-minimal as in the 1D case.

In this chapter, we review the balanced truncation model reduction (BTMR) method and the related error bounds, first presenting the standard results for 1D systems ([51], [24], [33], [36]), followed by the extension of these methods to uncertain systems, which were originally presented in [69]. Simplified proofs for the BTMR error bounds are provided in Appendix C. We then give a new model reduction result for uncertain systems that also relies on the solution of two LMIs, and gives tighter error bounds than the BTMR method. These new results, which were first noted in [4], are based on technical machinery presented in [54] and [56].

5.1 Norms for Uncertain Systems

In order to quantify the error resulting from the model reduction process, naturally, we use the induced 2-norm, as we are considering input and output signals, u and y in the space of summable sequences, l_2. For a 1D system with system operator denoted by G, so that $y = Gu$,

$$\|G\|_{l_2 \to l_2} = \sup_{\theta \in [0,2\pi]} \overline{\sigma}[G(e^{j\theta})] = \|G\|_\infty ;$$

this norm is commonly referred to as the H_∞ norm of the system. For an uncertain or multi-dimensional system represented by the LFT $(\Delta \star M)$, we use the *structured induced 2-norm*, (SI2-norm), which we define as follows:

Definition 5.1 *The* **SI2-norm** *of a stable system* $(\boldsymbol{\Delta}, M)$ *is given by*

$$\|\boldsymbol{\Delta} \star M\|_{SI2} = \sup_{\Delta \in \mathbf{B}_\Delta} \|\Delta \star M\|_{l_2 \to l_2} . \tag{5.1}$$

Clearly, the SI2-norm is a generalization of the H_∞ norm, and for 1D system realizations with no uncertainty these norms are identical.

The difference between two realizations, $(\boldsymbol{\Delta}_1, M_1)$ and $(\boldsymbol{\Delta}_2, M_2)$, is evaluated in the SI2-norm by forming the difference realization of $(\widetilde{\Delta} \star \widetilde{E}) = (\Delta_1 \star M_1) - (\Delta_2 \star M_2)$, that is, by setting

$$\widetilde{E} = \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ C_1 & -C_2 & D_1 - D_2 \end{bmatrix} \text{ and } \widetilde{\Delta} = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \tag{5.2}$$

and computing $\left\|\widetilde{\boldsymbol{\Delta}} \star \widetilde{E}\right\|_{SI2}$.

An equivalent formulation for the SI2-norm of a system when $\boldsymbol{\Delta}$ represents arbitrary time varying l_2 uncertainty is given in the following lemma. This formulation more readily allows for computation via software packages developed for solving LMIs ([30], [25]).

Lemma 5.2 *The SI2-norm of a stable system* $(\boldsymbol{\Delta}, M)$ *is equivalent to*

$$\inf\left\{\gamma : \textit{ there exists } T \textit{ such that } \overline{\sigma}\left(\begin{bmatrix} TAT^{-1} & \frac{1}{\gamma^{\frac{1}{2}}}TB \\ \frac{1}{\gamma^{\frac{1}{2}}}CT^{-1} & \frac{1}{\gamma}D \end{bmatrix}\right) < 1\right\} \tag{5.3}$$

where $T \in \mathcal{T}$ *and* $\gamma > 0$.

Proof. Let $\Delta_F \in \mathcal{L}(l_2^p)$ and $M_\gamma = \begin{bmatrix} A & \frac{1}{\gamma^{\frac{1}{2}}}B \\ \frac{1}{\gamma^{\frac{1}{2}}}C & \frac{1}{\gamma}D \end{bmatrix}$. Then for a stable system $(\boldsymbol{\Delta}, M)$,

$$\inf\left\{\gamma : \text{ there exists } T \text{ such that } \overline{\sigma}\left(\begin{bmatrix} TAT^{-1} & \frac{1}{\gamma^{\frac{1}{2}}}TB \\ \frac{1}{\gamma^{\frac{1}{2}}}CT^{-1} & \frac{1}{\gamma}D \end{bmatrix}\right) < 1\right\}$$

$$= \inf\left\{\gamma : \text{ there exists } Y > 0 \text{ such that } M_\gamma \begin{bmatrix} Y & 0 \\ 0 & I \end{bmatrix} M_\gamma^* - \begin{bmatrix} Y & 0 \\ 0 & I \end{bmatrix} < 0\right\}, \tag{5.4}$$

for Y and T both in $\mathcal{T}$ and $\gamma > 0$. Applying Theorem 4.6 and the stability assumption on $(\boldsymbol{\Delta}, M)$, we see that (5.4) is equivalent to

$$\inf\left\{\gamma : \left(I - M_\gamma \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_F \end{bmatrix}\right) \text{ is invertible in } \mathcal{L}(l_2) \text{ for all } \Delta_F \in \mathbf{B}_{\Delta_F}\right\}$$

$$= \inf\left\{\gamma : \begin{bmatrix} I - A\Delta & -\frac{1}{\gamma^{\frac{1}{2}}}B\Delta_F \\ 0 & I - (\frac{1}{\gamma}\Delta \star M)\Delta_F \end{bmatrix} \text{is invertible in } \mathcal{L}(l_2) \text{ for all } \Delta_F \in \mathbf{B}_{\Delta_F}\right\} \tag{5.5}$$

$$= \inf\left\{\gamma : (I - (\frac{1}{\gamma}\Delta \star M)\Delta_F) \text{ is invertible in } \mathcal{L}(l_2) \text{ for all } \Delta_F \in \mathbf{B}_{\Delta_F}\right\}$$

$$= \sup_{\Delta \in \mathbf{B}\boldsymbol{\Delta}} \|\Delta \star M\|_{l_2 \to l_2}.$$

■

Using Lemma 5.2, we can immediately prove the following result which generalizes a similar result for 1D systems (see Lemma C.2 in the Appendix C) to uncertain systems.

Lemma 5.3 *Suppose $(\boldsymbol{\Delta}, M)$ represents a stable uncertain system, then $\|\boldsymbol{\Delta} \star M\|_{SI2} \leq 1$ if and only if there is a realization $\widehat{M}$, where $(\Delta \star M) = (\Delta \star \widehat{M})$ for all $\Delta \in \boldsymbol{\Delta}$, such that $\overline{\sigma}(\widehat{M}) \leq 1$.*

We say a constant matrix X satisfying $\overline{\sigma}(X) \leq 1$ is *contractive*, and is *strictly contractive* if $\overline{\sigma}(X) < 1$. If we consider uncertainty structures which contain LTI or commuting operators, for example, real parametric uncertainty, LTI dynamic uncertainty, or when the $\boldsymbol{\Delta}$ set represents mulitple shift operators corresponding to different transform variables in a multi-dimensional system, then the expression on the right in (5.3) is an upper bound for the system SI2-norm. Therefore, the existence of a contractive realization is a sufficient condition for $\|\boldsymbol{\Delta} \star M\|_{SI2} \leq 1$ when the δ_i of the Δ block are LTI. This sufficiency is all that is needed for the balanced truncation model reduction bounds to hold for both multi-dimensional and uncertain systems.

5.2 Model Reduction Error Bounds for Stable Systems

We begin by stating the BTMR results, first given for 1D discrete time systems in [36], and for multi-dimensional and uncertain systems in [69]. The results for 1D systems are given first, followed by the generalized results for uncertain systems. Proofs are provided in Appendix C. We then state and prove a necessary and sufficient condition for satisfaction of tighter model reduction error bounds, which are also measured in the SI2-norm.

Balanced Truncation Model Reduction

Consider a stable discrete time system with the following realization

$$M = \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ C_1 & C_2 & D \end{bmatrix}, \ \Delta = \lambda I.$$

Suppose Y and X are two positive semi-definite symmetric matrices satisfying the following Lyapunov inequalities

$$AYA^* - Y + BB^* \leq 0 \tag{5.6}$$

$$A^*XA - X + C^*C \leq 0. \tag{5.7}$$

We consider Lyapunov *inequalities* in (5.6) and (5.7) in order to generalize the 1D continuous time results to uncertain systems in discrete time. Note that Lyapunov inequalities are also used in [36] for the development of BTMR error bounds for 1D discrete time systems. An example is given in [36] to illustrate how the use of solutions to the inequalities may result in better error bounds than the use of the true system Gramians.

The significance of these inequalities in the 1D case is that while the zero-valued eigenvalues of Y or X still have corresponding uncontrollable and/or unobservable states, the converse need not be true. This is most easily seen via a simple example: consider the system with realization

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.2 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ C = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

It is clear that

$$Y = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } X = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

satisfy the inequalities of (5.6) and (5.7), as the first state is unobservable and the second state is uncontrollable. However, $Y = X = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, for example, also satisfy the inequalities of (5.6) and (5.7). Thus, we can truncate states in the same manner suggested by Theorem 3.8, but the resulting system may not be minimal. Subsequently, we will assume the states corresponding to the zero-valued eigenvalues of Y and X have been truncated, and when we refer to balanced system realizations it will now be in a looser sense, that is with $Y > 0$ and $X > 0$ satisfying the Lyapunov (strict) inequalities and

$$Y = X = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

with

$$\begin{aligned} \Sigma_1 &= \operatorname{diag}(\sigma_1 I_{s_1}, \sigma_2 I_{s_2}, \ldots, \sigma_r I_{s_r}) > 0 \\ \Sigma_2 &= \operatorname{diag}(\sigma_{r+1} I_{s_{r+1}}, \ldots, \sigma_n I_{s_n}) > 0, \end{aligned} \tag{5.8}$$

where s_i denotes the multiplicity of σ_i. Note that the σ_i are not necessarily ordered, and are assumed to be distinct, although distinctness is not required (see Remark 5.6).

The BTMR results for 1D systems, given below, are separated into a lemma stating that the truncation of a stable, balanced realization is also stable and balanced, and a theorem stating the upper error bound results, measured in the H_∞ norm. The proof for the lemma can be found in [36]. A proof for the theorem that is more concise than that presented in [69] is given in Appendix C. This proof generalizes immediately to system representations which include uncertainty.

Lemma 5.4 *Suppose* $M = \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ C_1 & C_2 & D \end{bmatrix}$ *is a balanced, stable realization. Then the truncated system realization given by*

$$M_r = \begin{bmatrix} A_{11} & B_1 \\ C_1 & D \end{bmatrix}$$

is also balanced and stable.

Theorem 5.5 *Suppose* $M = \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ C_1 & C_2 & D \end{bmatrix}$ *is a balanced, stable realization for the system described by* G *with* $X = Y = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} > 0$*, as defined in (5.8). Let* $M_r = \begin{bmatrix} A_{11} & B_1 \\ C_1 & D \end{bmatrix}$ *denote the balanced, stable, truncated system realization for* G_r*. Then*

$$\|G - G_r\|_\infty \leq 2 \sum_{i=r+1}^{n} \sigma_i.$$

Proof. See Appendix C. ∎

Remark 5.6 *Distinctness of the* σ_i *in* Σ_1 *and* Σ_2 *is not required for stability of the truncated subsystem in the discrete time case, although it is in the continuous time case. However, there is no reason to truncate the system in the middle of a* $\sigma_i I_{s_i}$ *block as this will not improve the error bound.*

BTMR Error Bounds for Uncertain Systems

In order to derive the model reduction error bounds for balanced uncertain systems, we partition the realization matrices A, B, C, and the balanced Gramian Σ so as to separate the subblocks that will be truncated. We again assume that the realization submatrices corresponding to the zero eigenvalues of the structured Gramians have been truncated, as suggested by Theorem 4.12. We then consider solutions $Y > 0$ and $X > 0$ that yield strict inequalities for the system Lyapunov inequalities. Consider the block structure $\boldsymbol{\Delta} = \left\{\text{diag}\left[\delta_1 I_{n_1}, \ldots, \delta_p I_{n_p}\right] : \delta_i \in \mathcal{L}(l_2)\right\}$; with A, B and C partitioned conformally with this structure as

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & \cdots & C_p \end{bmatrix}.$$

We further partition each block of Σ by $\Sigma_i = \text{diag}[\hat{\Sigma}_{1i}, \Sigma_{2i}]$, for $i = 1, \ldots, p$, where the realization submatrices corresponding to Σ_{2i} will be truncated. Denote

$$\hat{\Sigma}_{1i} = \text{diag}[\sigma_{i1} I_{s_{i1}}, \ldots, \sigma_{ik_i} I_{s_{ik_i}}] > 0,$$

and

$$\Sigma_{2i} = \text{diag}[\sigma_{i(k_i+1)} I_{s_{i(k_i+1)}}, \ldots, \sigma_{it_i} I_{s_{it_i}}] > 0 \quad k_i \le t_i.$$

We then truncate both Σ_{2i} and the corresponding parameter matrices, for example, we truncate

$$A_{11} = \begin{bmatrix} \hat{A}_{11} & A_{11_{12}} \\ A_{11_{21}} & A_{11_{22}} \end{bmatrix}, \; B_1 = \begin{bmatrix} \hat{B}_1 \\ B_{1_2} \end{bmatrix} \text{ and } C_1 = \begin{bmatrix} \hat{C}_1 & C_{1_2} \end{bmatrix}$$

to $\hat{A}_{11}$, $\hat{B}_1$ and $\hat{C}_1$. Partitioning and truncating each A_{ij}, B_j and C_i, $i, j = 1, \ldots, p$ similarly results in the following truncated system,

$$\widehat{M} = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & D \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \cdots & \hat{A}_{1p} & \hat{B}_1 \\ \vdots & \ddots & \vdots & \vdots \\ \hat{A}_{p1} & \cdots & \hat{A}_{pp} & \hat{B}_p \\ \hat{C}_1 & \cdots & \hat{C}_p & D \end{bmatrix},$$

with uncertainty set $\hat{\boldsymbol{\Delta}} = \left\{\text{diag}[\delta_1 I_{\hat{n}_1}, \ldots, \delta_p I_{\hat{n}_p}]\right\}$ where $\hat{n}_i = \sum_{j=1}^{k_i} s_{ij}$.

As in the 1D case, truncating a balanced stable uncertain system realization results in a lower order realization which is balanced and stable; this is easily seen by considering the system Lyapunov inequalities, and is given in Appendix C for completeness.

Lemma 5.7 *Suppose $(\hat{\Delta}, \widehat{M})$ is the reduced model obtained from truncating the balanced stable system (Δ, M). Then $(\hat{\Delta}, \widehat{M})$ is also balanced and stable.*

Proof. See Appendix C. ■

We now state the BTMR error bound theorem for uncertain and multi-dimensional systems.

Theorem 5.8 *Suppose $(\hat{\mathbf{\Delta}}, \widehat{M})$ is the reduced model obtained from the balanced stable system $(\mathbf{\Delta}, M)$. Then*

$$\|(\mathbf{\Delta} \star M) - (\hat{\mathbf{\Delta}} \star \widehat{M})\|_{SI2} \leq 2 \sum_{i=1}^{p} \sum_{j=k_i+1}^{t_i} \sigma_{ij}. \tag{5.9}$$

Proof. See Appendix C. ■

Improved Error Bounds: LMI-Based Model Reduction

A tighter model reduction bound then that given in Theorem 5.8 can be achieved using the solutions to the system Lyapunov inequalities by utilizing machinery presented in [54, 56]; similar machinery is also given in [40] and [1]. The tighter error bound is derived from Lemma 5.2, and from the results of [54, 56], which are given below. Throughout this section we again refer to the realizations and uncertainty structures:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \ \mathbf{\Delta} = \left\{\text{diag}\left[\delta_1 I_{n_1}, \delta_2 I_{n_2}, \ldots, \delta_p I_{n_p}\right] : \delta_i \in \mathcal{L}(l_2)\right\},$$

$$M_r = \begin{bmatrix} A_r & B_r \\ C_r & D \end{bmatrix}, \ \mathbf{\Delta}_r = \mathcal{I}_r(\mathbf{\Delta}) = \left\{\text{diag}\left[\delta_1 I_{r_1}, \delta_2 I_{r_2}, \ldots, \delta_p I_{r_p}\right] : \delta_i \in \mathcal{L}(l_2)\right\},$$

and

$$\tilde{\mathbf{\Delta}} = \left\{\begin{bmatrix} \Delta & 0 \\ 0 & \Delta_r \end{bmatrix} : \Delta \in \mathbf{\Delta}, \Delta_r \in \mathbf{\Delta}_r\right\}.$$

We again assume one of the δ_i represents the delay operator on l_2, and the remaining δ_i are LTV operators on l_2. The commutative matrix sets corresponding to these uncertainty structures are denoted, as before, by $\mathcal{T}$, $\mathcal{T}_r$ and $\widetilde{\mathcal{T}}$, where the set $\widetilde{\mathcal{T}}$ for $\tilde{\mathbf{\Delta}}$ includes matrices with the following block structure:

$$T = \begin{bmatrix} \text{diag}(T_i^n) & \text{diag}(T_i^{nr}) \\ \text{diag}(T_i^{rn}) & \text{diag}(T_i^r) \end{bmatrix}$$

with $\dim(T_i^n) = n_i \times n_i$, $\dim(T_i^{nr}) = n_i \times r_i$, $\dim(T_i^{rn}) = r_i \times n_i$ and $\dim(T_i^r) = r_i \times r_i$ for all $i = 1, \ldots, p$. For the uncertain system representations $(\mathbf{\Delta}, M)$ and $(\mathbf{\Delta}_r, M_r)$, the difference system $(\tilde{\Delta} \star \tilde{E}) = (\Delta \star M) - (\Delta_r \star M_r)$ is formed as in (5.2).

Given a system representation $(\boldsymbol{\Delta}, M)$, we show that for any $\epsilon > 0$, a lower dimension realization $(\boldsymbol{\Delta}_r, M_r)$ exists such that the SI2-norm of the difference system $(\tilde{\boldsymbol{\Delta}}, \tilde{E})$ is bounded above by ϵ if and only if there exist solutions, X_ϵ and Y_ϵ, to the Lyapunov inequalities (5.6) and (5.7) that also satisfy a rank constraint.

Theorem 5.9 *Given a system realization $(\boldsymbol{\Delta}, M)$, there exists a reduced realization $(\boldsymbol{\Delta}_r, M_r)$ such that $\|\tilde{\boldsymbol{\Delta}} \star \tilde{E}\|_{SI2} \leq \epsilon$ if and only if there exists $X_\epsilon > 0$ and $Y_\epsilon > 0$, both in $\mathcal{T}$, satisfying*

(i) $AX_\epsilon A^* - X_\epsilon + BB^* < 0$

(ii) $A^* Y_\epsilon A - Y_\epsilon + C^* C < 0,$ *and*

(iii) $\lambda_{min}(X_\epsilon Y_\epsilon) = \epsilon^2$, *with multiplicity* $\sum_{i=1}^{p}(n_i - r_i)$

where $\epsilon > 0$.

The existence of solutions, $X_\epsilon > 0$ and $Y_\epsilon > 0$, to the Lyapunov inequalities is both necessary and sufficient for guaranteeing the bound of ϵ. Note that when $\epsilon = 0$, we then have the necessary and sufficient reducibility condition presented in Chapter 4. Thus, this error bound is tight in the sense that if we can find optimal solutions to the Lyapunov inequalities, that is to a pair of LMIs, we will be able to find a lower dimension approximation to the full model that results in the least possible error. Applying Theorem 5.9 recursively to a balanced realization results in error bounds for model reduction which are lower than those of Theorem 5.8 by a factor of two, however the resulting additive error bounds are not tight.

To prove Theorem 5.9, we use the following results, Lemmas 5.10 and 5.11, taken directly from [54, 56]; proofs may be found in [56]. These results are applied to the error system $(\tilde{\boldsymbol{\Delta}}, \tilde{E})$, in conjunction with Lemma 5.2, to construct the LMI conditions (i) and (ii), and the rank constraint (iii) of Theorem 5.9. We provide proofs for each of these lemmas for the sake of completeness. Let $R \in \mathbb{R}^{l\times l}$, $U \in \mathbb{R}^{l\times m}$ and $V \in \mathbb{R}^{q\times l}$, where $m, q \leq l$.

Lemma 5.10 *([54, 56]): Suppose $U_\perp \in \mathbb{R}^{l\times(l-m)}$ and $V_\perp \in \mathbb{R}^{(l-q)\times l}$ satisfy $U^* U_\perp = 0$, $VV_\perp^* = 0$, with $[U \quad U_\perp]$ and $\begin{bmatrix} V \\ V_\perp \end{bmatrix}$ invertible. Let $Z \in \mathbb{C}^{l\times l}$ be a given set of positive definite Hermitian matrices. Then*

$$\inf_{\substack{Q \in \mathbb{R}^{m\times q} \\ Z \in \mathcal{Z}}} \overline{\sigma}(Z^{\frac{1}{2}}(R + UQV)Z^{-\frac{1}{2}}) < 1$$

if and only if there exists $Z \in \mathcal{Z}$ such that

$$V_\perp(R^* Z R - Z)V_\perp^* < 0 \ \ \textit{and} \ \ U_\perp^*(RZ^{-1}R^* - Z^{-1})U_\perp < 0. \tag{5.10}$$

Proof. Fix $Z \in \mathcal{Z}$ and denote $f(Z) = \inf_{Q\in\mathbb{R}^{m\times p}} \overline{\sigma}(Z^{\frac{1}{2}}(R+UQV)Z^{-\frac{1}{2}})$. Clearly $f(Z) < 1$ if and only if

$$Z^{-\frac{1}{2}}(R+UQV)^* Z(R+UQV)Z^{-\frac{1}{2}} - I < 0, \tag{5.11}$$

or equivalently

$$Z^{\frac{1}{2}}(R+UQV)Z^{-1}(R+UQV)^* Z^{\frac{1}{2}} - I < 0. \tag{5.12}$$

Multiplying out the terms in (5.11) and (5.12) gives $f(Z) < 1$ if and only if

$$Z^{-\frac{1}{2}}(R^*ZR + R^*ZUQV + V^*Q^*U^*ZR + V^*QU^*UQV)Z^{-\frac{1}{2}} - I < 0$$

and

$$Z^{\frac{1}{2}}(RZ^{-1}R^* + RZ^{-1}V^*Q^*U^* + UQVZ^{-1}R^* + UQVZ^{-1}V^*Q^*U^*)Z^{\frac{1}{2}} - I < 0.$$

Now define $\widetilde{U} = Z^{\frac{1}{2}}U$ and $\widetilde{V} = VZ^{-\frac{1}{2}}$. Note that $(Z^{-\frac{1}{2}}U_\perp)^*\widetilde{U} = 0$, thus the columns of $Z^{-\frac{1}{2}}U_\perp$ span the space orthogonal to the range of $\widetilde{U}$; similarly, since $\widetilde{V}(V_\perp Z^{\frac{1}{2}}) = 0$, the rows of $V_\perp Z^{\frac{1}{2}}$ span the space orthogonal to the range of $\widetilde{V}$. Substituting and simplifying gives the final result. ■

The LMI conditions given in (5.10) are respectively convex in Z and Z^{-1}, but the two conditions together cannot be formulated as a convex constraint on either variable for a general set of matrices $\mathcal{Z}$. However, for the model reduction problem addressed in this thesis (as for the synthesis problem discussed in [54]) these two conditions can be reformulated into one jointly convex condition, coupled by a non-convex rank constraint.

Using the notation of [54] for our model reduction problem, we define

$$R = \begin{bmatrix} A & 0 & B \\ 0 & 0 & 0 \\ C & 0 & D \end{bmatrix}, \quad U = \begin{bmatrix} 0 & 0 \\ I_r & 0 \\ 0 & I_q \end{bmatrix}, \quad V = \begin{bmatrix} 0 & I_r & 0 \\ 0 & 0 & I_m \end{bmatrix}$$

and

$$Q = \begin{bmatrix} A_r & B_r \\ -C_r & -D_r \end{bmatrix}.$$

Note that $\widetilde{E} = R + UQV$. Furthermore, $U_\perp = [I\ 0\ 0]^T$ and $V_\perp = [I\ 0\ 0]$; we then accordingly define $\widetilde{U} = [U_{\perp,1}^T\ U_{\perp,3}^T]^T = [I\ 0]^T$ and $\widetilde{V} = [V_{\perp,1}\ V_{\perp,3}] = [I\ 0]$, and note that the dimensions m, q and l correspond to the row and column dimensions of M_r and the dimensions of R, respectively.

In applying Lemma 5.10, if we set $\mathcal{Z} := \left\{ \begin{bmatrix} \widetilde{T} & 0 \\ 0 & I \end{bmatrix} : \widetilde{T} \in \widetilde{\mathcal{T}}, \widetilde{T} = \widetilde{T}^* > 0 \right\}$, then instead of casting (5.10) as a pair of convex constraints on $Y = \widetilde{T}$ and $X = Y^{-1}$, we need only consider constraints on the $(1,1)$ blocks of X and Y, X_i^n and Y_i^n, due to the structure of R, U, V, and the set $\widetilde{\mathcal{T}}$. In particular the following lemma may be used to form one LMI constraint on X_i^n and Y_i^n.

Lemma 5.11 *([54, 56]): Suppose $X \in \mathbb{C}^{n\times n}$ and $Y \in \mathbb{C}^{n\times n}$ are given, with $X = X^* > 0$ and $Y = Y^* > 0$. Then there exist matrices $X_2 \in \mathbb{C}^{n\times r}$ and $X_3 \in \mathbb{C}^{r\times r}$, with $X_3 = X_3^* > 0$, such that*

$$\begin{bmatrix} X & X_2 \\ X_2^* & X_3 \end{bmatrix} > 0 \text{ and } \begin{bmatrix} X & X_2 \\ X_2^* & X_3 \end{bmatrix}^{-1} = \begin{bmatrix} Y & Y_{12} \\ Y_{21} & Y_3 \end{bmatrix},$$

if and only if

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \geq 0 \text{ and } \operatorname{rank}\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \leq n + r.$$

Proof. Sufficiency: Pre- and post-multiplying $\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \geq 0$ by $[I \;\; -Y^{-1}]$ and $\begin{bmatrix} I \\ -Y^{-1} \end{bmatrix}$, it is clear that $(X - Y^{-1}) \geq 0$ and $\operatorname{rank}(X - Y^{-1}) \leq r$. Defining X_2 via the (Cholesky) factorization $X_2X_2^* = (X - Y^{-1})$ and $X_3 = I$ completes the construction.

Necessity: Using the Schur complement formula,

$$\begin{bmatrix} X & X_2 \\ X_2^* & X_3 \end{bmatrix}^{-1} = \begin{bmatrix} X^{-1} + X^{-1}X_2(X_3 - X_2^*X^{-1}X_2)^{-1}X^{-1} & \times \\ \times & \times \end{bmatrix}.$$

Denoting the $(1,1)$ element of the inverse by Y, and using the matrix inversion formula gives $Y^{-1} = (X - X_2X_3^{-1}X_2^*)$, thus $X - Y^{-1} = X_2X_3^{-1}X_2^* \geq 0$ and $\operatorname{rank}(X - Y^{-1}) \leq r$. ■

For the reduction problems we consider, the dimensions n and r correspond to the dimensions of Δ and Δ_r; that is, $n = \sum_{i=1}^{p} n_i$ and $r = \sum_{i=1}^{p} r_i$. Lemma 5.12 follows directly from the preceeding lemmas, and is also a corollary to Theorem 6.3 in [54].

Lemma 5.12 *Suppose the realization $(\mathbf{\Delta}, M)$ is given, with $R, U, V, \tilde{U}$ and $\tilde{V}$ defined as above. Then there exists a realization $(\mathbf{\Delta}_r, M_r)$ and a matrix $Z \in \widetilde{\mathcal{T}}$, $Z > 0$ satisfying*

$$\overline{\sigma}\left(\begin{bmatrix} Z & 0 \\ 0 & I \end{bmatrix} \tilde{E} \begin{bmatrix} Z^{-1} & 0 \\ 0 & I \end{bmatrix}\right) < 1$$

if and only if there exist $X_i^n > 0$, $Y_i^n > 0$ for $i = 1, \ldots, p$ satisfying

(a) $\tilde{U}^T \left(M \begin{bmatrix} diag(X_i^n) & 0 \\ 0 & I \end{bmatrix} M^T - \begin{bmatrix} diag(X_i^n) & 0 \\ 0 & I \end{bmatrix}\right) \tilde{U} < 0$

(b) $\tilde{V} \left(M^T \begin{bmatrix} diag(Y_i^n) & 0 \\ 0 & I \end{bmatrix} M - \begin{bmatrix} diag(Y_i^n) & 0 \\ 0 & I \end{bmatrix}\right) \tilde{V}^T < 0$

(c) $\begin{bmatrix} X_i^n & I \\ I & Y_i^n \end{bmatrix} \geq 0$

If these conditions are feasible, then the dimensions of Δ_r and M_r are determined by defining $r_i = \text{rank}(X_i^n - (Y_i^n)^{-1})$ for each $i = 1, \ldots, p$.

We can now prove Theorem 5.9. For convenience, we denote the ϵ-scaled difference system realization by

$$\widetilde{E}_\epsilon = \begin{bmatrix} A & 0 & \frac{1}{\epsilon^{\frac{1}{2}}}B \\ 0 & A_r & \frac{1}{\epsilon^{\frac{1}{2}}}B_r \\ \frac{1}{\epsilon^{\frac{1}{2}}}C & -\frac{1}{\epsilon^{\frac{1}{2}}}C_r & \frac{1}{\epsilon}(D - D_r) \end{bmatrix}. \tag{5.13}$$

Proof of Theorem 5.9:

By Lemma 5.2,

$$\left\|\widetilde{\Delta} \star \widetilde{E}\right\|_{SI2} \leq \epsilon \text{ if and only if there exists } T \in \widetilde{\mathcal{T}} \text{ such that}$$

$$\overline{\sigma}\left(\begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \widetilde{E}_\epsilon \begin{bmatrix} T^{-1} & 0 \\ 0 & I \end{bmatrix}\right) < 1.$$

Now we need only apply Lemma 5.12 to $\widetilde{E}_\epsilon$ and multiply out the matrices in statements (a) and (b) of Lemma 5.12. Then $\left\|\widetilde{\Delta} \star \widetilde{E}\right\|_{SI2} \leq \epsilon$ if and only if there exist $X = \text{diag}(X_i^n) > 0$, and $Y = \text{diag}(Y_i^n) > 0$ for $i = 1, \ldots, p$ satisfying

$$AXA^* + \frac{1}{\epsilon}BB^* - X < 0, \quad A^*YA + \frac{1}{\epsilon}C^*C - Y < 0 \tag{5.14}$$

and

$$\begin{bmatrix} X_i^n & I \\ I & Y_i^n \end{bmatrix} \geq 0.$$

Multiplying the matrix inequalities in (5.14) by ϵ, and denoting $X_\epsilon = \epsilon X$ and $Y_\epsilon = \epsilon Y$ gives (i) and (ii). Additionally we have

$$\begin{bmatrix} X_{\epsilon i}^n & \epsilon I \\ \epsilon I & Y_{\epsilon i}^n \end{bmatrix} \geq 0, \text{ with } \text{rank}\left(\frac{1}{\epsilon^2}X_{\epsilon i}^n - (Y_{\epsilon i}^n)^{-1}\right) = r_i. \tag{5.15}$$

Condition (iii) is obtained by pre- and post-multiplying (5.15) by $[\frac{1}{\epsilon}I \quad -(Y_{\epsilon i}^n)^{-1}]$ and $\begin{bmatrix} \frac{1}{\epsilon}I \\ -(Y_{\epsilon i}^n)^{-1} \end{bmatrix}$ respectively, giving

$$\frac{1}{\epsilon^2}X_{\epsilon i}^n - (Y_{\epsilon i}^n)^{-1} \geq 0,$$

thus $X_{\epsilon i}^n Y_{\epsilon i}^n \geq \epsilon^2 I$. Applying the rank condition implies $\text{rank}(X_{\epsilon i}^n Y_{\epsilon i}^n - \epsilon^2 I) = r_i$, thus $\lambda_{min}(X_{\epsilon i}^n Y_{\epsilon i}^n) = \epsilon^2$ with multiplicity $n_i - r_i$, for all $i = 1, \ldots, p$. Since X_ϵ and Y_ϵ are block diagonal compositions of $X_{\epsilon i}^n$ and $Y_{\epsilon i}^n$ the result follows. ■

When the uncertainty structure, $\boldsymbol{\Delta}$, contains time-invariant operators δ_i, as in the case of multi-dimensional system representations, the existence of X_ϵ and Y_ϵ are sufficient to ensure $\left\|\widetilde{\boldsymbol{\Delta}} \star \widetilde{E}\right\|_{SI2} \leq \epsilon$. For 1D continuous time systems with no uncertainty, similar results have been obtained by Kavranoglu and Bettayeb [43] via an alternate method that requires simultaneously computing a pair of matrices B_0 and C_0 augmenting the system realization matrices B and C, and solutions X and Y to the *augmented* Lyapunov equations, such that $\lambda_{min}(XY) = \epsilon^2$ with multiplicity $n - r$. Also, following the work presented in [69], balanced truncation model reduction error bounds for continuous time uncertain systems were presented in [16].

In this chapter we have shown that, via the computation of structured Gramians, we are able to reduce uncertain system realizations with guarantee upper error bounds, where these bounds are given in an induced 2-norm. From the preceeding chapter, we know that if either of these structured Gramians is singular, the uncertain system may be reduced with zero error. We now need only address the actual computation of these structured Gramians, given a realization. Computational issues are discussed in Chapter 7, along with a discussion of one application of these methods. Prior to the discussion of computational issues, we first address the connections alluded to in the proof of Theorem 4.15 between uncertain system realizations in the LFT framework and representations of formal power series in multiple noncommuting indeterminants; we focus on the relations between minimal LFT realizations and minimal FPS representations.

Chapter 6

Linear Fractional Transformations and Formal Power Series

In this chapter, we relate the notion of minimality presented for LFT realizations in Chapter 4 to realization theory results for formal power series (FPS), recalling that a series expansion of the LFT $(\Delta \star M)$ was used in the necessity proof of our main reducibility result, Theorem 4.15. In particular, we consider the realization theory originating from the work of Fliess (see [28], [29] and the references therein), and Isidori [38], who used *recognizable* series to develop realization theory for nonlinear and bilinear systems. Sontag also used recognizable series for a realization theory of discrete time nonlinear systems [64]. The realization theory based on formal power series results in a different definition of minimality, which is related to the rank of the series Hankel matrix. We discuss the relationship between the notions of minimality for LFT and series realizations, and present a method for obtaining one type of minimal realization from the opposing type.

In order to relate LFTs and formal series, we first review standard notation and results for these series, focusing on the theoretical developments in this area, rather than algorithms and computational results.

6.1 Formal Power Series

Formal power series have long been used in many branches of mathematics, most notably in combinatorics and enumeration. First proposed by Schutzenberger [66] as a generalization of automata and formal languages, the application of FPS for bilinear and nonlinear system realization theory was developed extensively throughout the 70's and 80's (see, for example, [28], [29], [64]).

In the most general form, a formal power series S is a function $\mathcal{X}^{\dagger} \to \mathcal{R}$, defined by

$$S = \sum_{w \in \mathcal{X}^*} s(w)w,$$

where $\mathcal{X}^\dagger$ is the monoid generated by a nonempty set $\mathcal{X}$ of p indeterminants, and $\mathcal{R}$ is a semiring. An element $x \in \mathcal{X}$ is called a *letter*, and an element $w \in \mathcal{X}^\dagger$ a *word*, where the length of a word $w = x_1 x_2 \cdots x_n$ is n, and is denoted by $|w|$; the empty word is denoted by 1. The product operation of the monoid $\mathcal{X}^\dagger$ is the concatenation defined by $(x_1 \cdots x_n) \cdot (y_1 \cdots y_k) = x_1 \cdots x_n y_1 \cdots y_k$.

A coefficient of the series S is denoted by $s(w)$, and is the image by S of a word w, that is, $s(w) \in \mathcal{R}$. The support of S is defined by $\text{supp}(S) = \{w \in \mathcal{X}^\dagger |\ s(w) \neq 0\}$. Note that a polynomial is a formal series with finite support. The ring of polynomials in p indeterminants (defined by $\mathcal{X}$) over $\mathcal{R}$ is denoted by $\mathcal{R}\langle\mathcal{X}\rangle$, and similarly the ring of formal series in p indeterminants over $\mathcal{R}$ by $\mathcal{R}\langle\langle\mathcal{X}\rangle\rangle$; $\mathcal{R}\langle\mathcal{X}\rangle$ is a subsemiring of $\mathcal{R}\langle\langle\mathcal{X}\rangle\rangle$.

A formal series may also be represented in matrix form by an associated Hankel matrix.

Definition 6.1 *The* **Hankel matrix** *of a formal series S is the matrix $\mathcal{H}_S$ indexed by $\mathcal{X}^\dagger \times \mathcal{X}^\dagger$ defined by $\mathcal{H}_S(u, v) = s(u \cdot v)$, for all words $u, v \in \mathcal{X}^\dagger$.*

A standard result for formal series connects the *rank* of a series S with the rank of the matrix, $\mathcal{H}_S$. To define the rank of S, we must first define the kernel of S. Although we will lose some generality, at this point we assume $\mathcal{R}$ is a field, and we consider series with coefficients in $\mathcal{R}^{q\times m}$. This is the case most relevant to the remainder of this chapter. We first endow $\mathcal{R}^{q\times m}\langle\langle\mathcal{X}\rangle\rangle$ with the structure of a right $\mathcal{R}\langle\mathcal{X}\rangle$-module by allowing the operation of addition of series to be defined coefficient-wise, and the product of a series $S \in \mathcal{R}^{q\times m}\langle\langle\mathcal{X}\rangle\rangle$ and a polynomial $\rho \in \mathcal{R}\langle\mathcal{X}\rangle$, denoted by $S \cdot \rho$, to be defined as follows:

(i) $S \cdot 1 = S$

(ii) for all $x \in \mathcal{X}$ the series $S \cdot x$ is given by

$$(S \cdot x)(w) = s(wx), \text{for } w \in \mathcal{X}^\dagger \tag{6.1}$$

(iii) for all $\rho_1, \rho_2 \in \mathcal{R}\langle\mathcal{X}\rangle,\ \ \alpha_1, \alpha_2 \in \mathcal{R}$

$$S \cdot (\alpha_1\rho_1 + \alpha_2\rho_2) = \alpha_1(S \cdot \rho_1) + \alpha_2(S \cdot \rho_2)$$

We then define the kernel of S by

$$\text{Ker } S = \{\rho \in \mathcal{R}\langle\mathcal{X}\rangle | S \cdot \rho = 0\}.$$

Definition 6.2 *The* **rank** *of a formal series S is equal to the codimension of Ker S, that is rank(S) = dim$(\mathcal{R}\langle\mathcal{X}\rangle$/Ker $S)$.*

Note that by applying Theorem 2.8, we know that $(\mathcal{R}\langle\mathcal{X}\rangle/\text{Ker } S)$ is isomorphic to Im S as a right $\mathcal{R}\langle\mathcal{X}\rangle$-module; thus rank$(S)$ = dim(Im S), which leads to the following result:

Theorem 6.3 *([15], [28], [38]) The rank of a formal series is equal to the rank of its Hankel matrix.*

6.2 Recognizable Series and Nonlinear Realization Theory

A recognizable series is a formal series S for which there exist *linear representations* with the following form [66]:

$$\begin{aligned} &\text{there exists } N \in \mathbb{Z}^+, \quad \text{a morphism } f : \mathcal{X}^\dagger \to \mathcal{R}^{N\times N} \\ &\text{and matrices } h \in \mathcal{R}^{q\times N}, \ g \in \mathcal{R}^{N\times m} \text{ such that} \\ &s(w) = hf(w)g, \quad \text{for all } w \in \mathcal{X}^\dagger. \end{aligned} \tag{6.2}$$

These are often referred to as *rational* series: a formal series is called a rational series if it is an element of the smallest subset of $\mathcal{R}\langle\langle\mathcal{X}\rangle\rangle$ which is closed under a set of operations called the rational operations. Basically, a FPS is a rational series if it can be expressed as a finite number of sums, products and inversions of polynomials. Schutzenberger showed that rational and recognizable series are equivalent [66]. As we are interested in the linear representations of a formal series, we refer to these series as recognizable series.

Similarity amongst linear representations for a series is defined as for state-space system realizations, that is, via the existence of an invertible matrix T that is used to transform h, g and $f(w)$; that is, $h \mapsto hT^{-1}$, $g \mapsto Tg$, and $f(w) \mapsto Tf(w)T^{-1}$. Minimality for series realizations is defined as follows:

Definition 6.4 *A* **minimal linear representation** *of a series S is a representation with minimal dimension N among all its representations.*

The dimensions q and m are always fixed. Strictly speaking, the standard definition given for recognizable series has $q = 1$ and $m = 1$ in the above, which would give us a series corresponding to a single-input single-output (SISO) system mapping. The following definitions and results for recognizable series are basic to the remainder of this chapter, and were also originally developed for the SISO case, where, in particular, $\mathcal{R}$ is assumed to be a field. These results can be extended to the multi-input multi-output (MIMO) case, that is $q \geq 1$ and $m \geq 1$. The LFT realizations we consider throughout this thesis represent MIMO systems.

Theorem 6.5 *([66], [28], [15])*

(a) A given series S is recognizable if and only if the rank of its Hankel matrix is finite.

(b) The rank of a recognizable series is equal to the minimum of the dimensions of all linear representations of S.

(c) Two minimal linear representations of a series S are similar.

Part (c) of Theorem 6.5 leads to the following corollary, originally noted by Schutzenberger, also in [66].

Corollary 6.6 *Suppose $\{h, f, g\}$ and $\{\overline{h}, \overline{f}, \overline{g}\}$ are two linear representations of a series S, and assume the latter representation is reduced. Then there exists a third representation $\{\hat{h}, \hat{f}, \hat{g}\}$ similar to $\{h, f, g\}$ with the following block decomposition:*

$$\hat{h} = [\times, \overline{h}, 0], \ \hat{f} = \begin{bmatrix} f_1 & 0 & 0 \\ \times & \overline{f} & 0 \\ \times & \times & f_2 \end{bmatrix}, \ \hat{g} = \begin{bmatrix} 0 \\ \overline{g} \\ \times \end{bmatrix}.$$

Necessity of Theorem 6.5, part (a) is immediate for the MIMO case; sufficiency of part(a), and part(b) have been shown for the case of SIMO bilinear system mappings by Isidori [38], [39], and for MIMO state-affine nonlinear system mappings by Sontag [64]. The results of either Sontag or Isidori are applicable to the uncertain systems we consider. The extension of part(c) to the MIMO case is given in Appendix D; as a result the extension of the corollary follows.

We consider only recognizable, or finite rank series in the sequel. In keeping with the notation used for recognizable series in the development of nonlinear realization theory, to which we will draw the most explicit connections, we consider the set of p non-commuting variables, $\mathcal{X} = \{\delta_1, \ldots, \delta_p\}$, and the associated index set $\mathcal{I} = \{1, 2, \ldots, p\}$. Let $\mathcal{I}_k$ denote the set of all sequences of k elements $(i_k \ldots i_1)$ of $\mathcal{I}$, where the empty sequence is denoted by $\emptyset$. Define $\mathcal{I}^\dagger = \bigcup_{k \geq 0} \mathcal{I}_k$, where $\mathcal{I}^\dagger$ has the composition rule, $(i_k \ldots i_1)(j_l \ldots j_1) \mapsto (i_k \ldots i_1 j_l \ldots j_1)$, and $\mathcal{I}_0 = \emptyset$. Then, to each multi-index $(i_k \ldots i_0)$ we associate the word $(\delta_{i_k} \cdots \delta_{i_0})$. As in [38] and [64], we consider a formal power series in p noncommutative indeterminates with coefficients in the ring of real matrices, $\mathbb{R}^{q \times m}$, that is, the mapping $S : \mathcal{I}^\dagger \to \mathbb{R}^{q \times m}$ represented by the form

$$S = s(\emptyset) + \sum_{k=0}^{\infty} \sum_{i_0, \ldots, i_k = 0}^{p} s(i_k \cdots i_0) \delta_{i_k} \cdots \delta_{i_0}$$

where $s(i_k \cdots i_0)$ is the coefficient of the $(i_k, \ldots, i_0)$-th term. The Hankel matrix associated with this series is the infinite matrix whose elements are defined by

$$\mathcal{H}_S(i_k \ldots i_0, j_l \ldots j_0) = s(i_k \cdots i_0 j_l \cdots j_0).$$

Given a recognizable series in p indeterminants, constructive procedures exist for obtaining specific forms of minimal linear representations: Fliess presents a procedure for constructing a representation in the form of (6.2) corresponding to a nonlinear SISO map [28]; Isidori gives procedures to obtain SISO and SIMO bilinear system realizations in [38] and [39]; and Sontag gives a procedure for constructing a minimal MIMO state-affine realization from a discrete-time nonlinear input-output map [64]. Sontag's realization procedure can be modified to give matrices $\{H, F_i, G\}$, $i = 1, \dots p$, satisfying state equations of the form

$$\begin{aligned} x &= \textstyle\sum_{i=1}^{p} F_i \delta_i (x + Gu) \\ y &= Hx + HGu. \end{aligned}$$

These equations are relevant to the LFT systems we consider. We will henceforth refer to $\{H, F_i, G\}$ as *series realizations*, where the equivalent series is

$$S = H(I - \sum_{i=1}^{p} F_i \delta_i)^{-1} G = H \sum_{k=0}^{\infty} (\sum_{i=1}^{p} F_i \delta_i)^k G. \tag{6.3}$$

Alternatively, if one is given a series realization $\{H, F_i, G\}$, this realization is minimal if the matrices $\mathcal{O}_S$ and Γ_S defined by

$$\mathcal{O}_S(i_k \dots i_0) = HF_{i_k} \cdots F_{i_0} \text{ and } \Gamma_S(j_l \dots j_0) = F_{j_l} \cdots F j_0 G$$

are both full rank; furthermore, the rank test for $\mathcal{O}_S$ and Γ_S can be performed on finite matrices [65]. Note also that $\mathcal{H}_S = \mathcal{O}_S \Gamma_S$. (See also [10] for discussions of related results in 2-dimensional system realization theory).

6.3 Connections: Minimal LFT and Series Realizations

Given an uncertain system realization, $(\mathbf{\Delta}, M)$, with corresponding I/O behavior described by the LFT, $D + C\Delta(I - A\Delta)^{-1}B$, recall that we can form a power series by expanding the $(I - A\Delta)^{-1}$ term, giving

$$S = \Delta \star M = D + \sum_{k=0}^{\infty} C\Delta(A\Delta)^k B. \tag{6.4}$$

To obtain a relation between a LFT realization and a minimal series realization $\{H, F_i, G\}$ we compare the terms in the respective series defined by these realizations. We first partition the matrices A, B, and C conformally with the Δ structure, that is,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & \ddots & \vdots \\ A_{p1} & \cdots & A_{pp} \end{bmatrix}; \quad B = \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix}; \quad C = \begin{bmatrix} C_1 & \cdots & C_p \end{bmatrix}.$$

Then, expanding the series S defined in (6.4) gives

$$\Delta \star M = D + \sum_{i=1}^{p} C_i B_i \delta_i + \sum_{k=0}^{\infty} \sum_{i_0, \dots, i_k = 1}^{p} C_{i_k} A_{i_k i_{k-1}} \cdots A_{i_1 i_0} B_{i_0} \delta_{i_k} \cdots \delta_{1_0}.$$

As discussed in Section 6.2, we can compute a minimal realization $\{H, F_i, G\}$ for a given series S, such that

$$S = H \sum_{k=0}^{\infty} (\sum_{i=1}^{p} F_i \delta_i)^k G. \tag{6.5}$$

Suppose we start with this minimal series realization $\{H, F_i, G\}$ and factor $F_i = L_i R_i$, where L_i has full column rank and R_i has full row rank. This factorization is nonunique; we will henceforth refer to such factorizations of a matrix as *minimal rank* (or MR) factorizations. By equating terms in (6.4) and (6.5) we obtain an LFT representation from the series representation:

$$D = HG, \; C_i = HL_i, \; B_i = R_i G, \text{ and } A_{ij} = R_i L_j. \tag{6.6}$$

It is readily seen that $\dim(A_{ii}) = \operatorname{rank}(F_i)$.

Conversely, if we are given a minimal LFT realization, we may obtain a corresponding series realization, $\{H, F_i, G\}$, by computing a minimal rank factorization

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} R \\ H \end{bmatrix} [L \; G] \tag{6.7}$$

where $\begin{bmatrix} R \\ H \end{bmatrix}$ has full column rank, and $[L \; G]$ has full row rank, and R and L are partitioned into p submatrices, for example, $L = [L_1 \; L_2 \cdots L_p]$. Defining $F_i = L_i R_i$ gives the realization, $\{H, F_i, G\}$.

The conversion procedure of (6.7) is clearly reversible, up to an allowable similarity transformation which accounts for the nonuniqueness; however, the minimal dimension for LFT realizations is $\dim(A)$, and for series realizations is $\dim(F_i) = \operatorname{rank}(\mathcal{H}_S)$. These two dimensions are in general not equal, with the dimensions of LFT realizations being related to the $\operatorname{rank}(F_i)$ of the series realizations. Bounds can be constructed to relate these dimensions, but neither dimension is always greater than or equal to the other. For uncertain systems the dimension $\dim(A)$ is more natural since it measures the number of copies of the δ_i required to build an interconnection that realizes the series, a generalization of the number of delays or integrators needed to realize a 1D transfer function. Furthermore, the $\operatorname{rank}(\mathcal{H}_S)$ depends on the constant term $s(\emptyset) = HG = D$, whereas $\dim(A)$ clearly does not. We could allow an additional constant term in the series realization so that the $\operatorname{rank}(\mathcal{H}_S)$ is minimized, then the dimension of the "minimal" realization $\{H, F_i, G\}$ would also be minimized. However, a more interesting question is whether the above formulas transform one type of minimal realization into the other. The answer is affirmative.

Proposition 6.7 *Given a minimal LFT realization (Δ, M), the series realization obtained via a MR factorization is minimal.*

Proof. By Corollary 4.18, we know all minimal LFT realizations are found using similarity transformations and truncations, operations which necessarily do not increase rank. Thus, for a minimal LFT realization $\operatorname{rank}\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right)$ is minimized over all equivalent realizations.

Now, we compute $\{H, F_i, G\}$ from $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ using a MR factorization, where

$$\begin{aligned} \operatorname{rank}\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) &= \operatorname{rank}\left(\begin{bmatrix} R \\ H \end{bmatrix}\begin{bmatrix} L & G \end{bmatrix}\right) \\ &= \dim(L_i R_i) = \dim(F_i). \end{aligned}$$

■

We then have the following corollary, from Proposition 6.7 and Theorem 6.5:

Corollary 6.8 *Suppose we have a minimal LFT realization (Δ, M), and form the associated series S as in (6.4). Then*

$$\operatorname{rank}(H_S) = \operatorname{rank}\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right). \tag{6.8}$$

The converse result, that of obtaining a minimal LFT realization from a minimal series realization, can be shown in the same manner, using Corollary 6.6.

Proposition 6.9 *Given a minimal series realization $\{H, F_i, G\}$, the LFT realization obtained via a MR factorization is minimal.*

Proof. From the minimal series realization, we compute the LFT realization via the MR factorization

$$F_i = \begin{bmatrix} \\ L_i \\ \\ \end{bmatrix} \begin{bmatrix} R_i \end{bmatrix},$$

giving

$$A_{ij} = \begin{bmatrix} R_i \end{bmatrix} \begin{bmatrix} \\ L_j \\ \\ \end{bmatrix}.$$

Thus $\dim(A) = \sum_{i=1}^{p} \dim(A_{ii}) = \sum_{i=1}^{p} \operatorname{rank}(F_i)$, so $\dim(A)$ is minimal if the $\operatorname{rank}(F_i)$ is minimal for all $i = 1, \ldots, p$, which is easily shown: $\operatorname{Rank}(F_i)$ is not changed by similarity transformation, thus rank(F_i) is a constant over all minimal series realizations. Furthermore, by Corollary 6.6, rank(F_i) for all non-minimal realizations of a series is either equal or greater to that for minimal realizations of the same series. ■

In this chapter, a brief presentation of the relation between minimal LFT realizations and minimal FPS representations is given. These results could lead to alternate methods for computing the minimal realization for an uncertain system. Using the series formulation given in equation (6.4) for a (nonminimal) uncertain system, a minimal FPS representation may be constructed via the procedure proposed by Sontag in [64]. A minimal LFT realization could then be obtained using a MR factorization. Computationally, however, there still exist numerical problems in the construction procedures for series representations. Thus, at this point, implementation of reduction methods for uncertain systems, as for 1D systems, are completed in the most computationally robust manner by balancing the realization and reducing with the guaranteed error bounds of Chapter 5.

Chapter 7

Computational Methods for Model Reduction and Applications

The results stated in Theorems 4.12 and 4.15 indicate that the existence of structured singular solutions to either of two Lyapunov inequalities for an uncertain or multidimensional system realization ensures the existence of an equivalent lower dimension realization. Furthermore, if the system uncertainty can be properly described by time-varying operators on l_2, then the existence of lower dimension realizations *requires* such singular solutions.

The computation of solutions to these Lyapunov inequalities belongs to a large class of problems, widely known as Linear Matrix Inequality (LMI) problems. The development of computational methods for solving LMIs has progressed rapidly in the control community in recent years, and, in fact, convex optimization algorithms for LMIs may now be obtained commercially ([30], [25]). The fact that we would like to find singular or rank constrained solutions to the LMIs we consider complicates the computational requirements, resulting in non-convex optimization problems for which heuristics must be used; these heuristics frequently rely on LMI solution methods.

In this chapter, we first discuss general LMI problems and solutions, with the aim of providing sufficient background and references for those interested in constructing computational solutions for reducing uncertain systems. We then present a simple suboptimal procedure for solving the Lyapunov inequalities for balanced model reduction of uncertain systems, followed by a discussion of the application of this procedure to a pressurized water reactor for a nuclear power plant.

7.1 Linear Matrix Inequalities

General LMI problems can be succinctly described by the following decision problem:

$$\begin{aligned} &\text{Does there exist } X : A^*XA + XB + B^*X + C^*C \leq 0 \\ &\qquad\qquad \text{subject to: } X \in \mathcal{T} \end{aligned} \tag{7.1}$$

where $\mathcal{T}$ describes the feasible set for X, that is, the set of block diagonal structured Hermitian matrices, and A, B, and C are constant matrices having corresponding block partitions. Additional constraints on X are sometimes required, for example, in the model reduction problems we consider, X must be positive semi-definite; for examples of systems and control problems that can be formulated as LMIs see [12]. Denote $A(X) = A^*XA+XB+B^*X+C^*C$. One way of reformulating the above LMI decision problem as an optimization problem is

$$\min_X \lambda_{max}(A(X)) \tag{7.2}$$
$$\text{subject to: } X \in \mathcal{T}.$$

$\lambda_{max}(A(X))$ is referred to as the objective function, and the condition $X \in \mathcal{T}$ is a constraint. When $A(X)$ is an Hermitian matrix which is affine in X, $\lambda_{max}(A(X))$ belongs to the class of convex, non-differentiable functions, for which numerous optimization methods exist. For general LMI problems, applications of the theory of *self-concordant* functions and the use of such functions in interior-point methods have recently found widespread use (see [53] and the references therein; see also [52] and [68] as examples). We provide general but brief descriptions of convex optimization and descent methods, and interior-point methods.

Convex Optimization Methods

Descent methods are used to iteratively solve unconstrained optimization problems. Denoting the objective function by $f(x)$, the basic steps in a descent algorithm are as follows: start from an initial point, x_k, determine by a fixed rule a direction of movement; move in that direction to the point x_{k+1}, which gives a minimum value for $f(x_k)$ on that line. In equation form, each step of a descent method computes

$$x_{k+1} = x_k - c_k Q_k g_k$$

where g_k is the descent direction, and the specific algorithm used determines Q_k. The main difference between descent algorithms is the rule by which the descent direction is determined at each stage. If f is differentiable, the gradient of f is often used to determine the descent direction. For example, in Newton's method, g_k is the gradient of f, and Q_k is the inverse of the Hessian of f. A line search may be done to determine the value $c_k \in R$ which minimizes $f(x_k - c_k Q_k g_k)$.

For smooth objective functions, descent methods are relatively simple to apply since gradients may be used for descent directions. Additionally, many of these methods converge quadratically to a local minimum. However, for non-differentiable problems, the determination of descent directions can be quite complicated and computationally expensive; convergence is not always guaranteed. In particular, for LMI problems, determining descent directions may be as difficult as solving the original problem. However, the application of interior-point methods to LMI problems yields smooth subproblems which can be solved using standard descent methods. More detailed descriptions of descent methods can be found in most texts on optimization, including [27], [46], and [32].

For non-differentiable optimization problems, convex programming methods such as *cutting-plane methods* and *ellipsoid methods* are much simpler than most descent methods to implement. They guarantee convergence to a solution to within a prespecified tolerance of a global minimum; however, they often require a substantially greater number of iterations than descent methods, and hence are much slower. Convex programming algorithms rely on the ability to compute values of the function and a *subgradient* for a given objective, f. A subgradient of f at a point x is any $g \in R^n$ such that

$$f(z) \geq f(x) + g^T(z - x) \qquad \text{for all } z \in R^n. \tag{7.3}$$

For example, the gradient of a differentiable convex function is a subgradient. Convex optimization algorithms have previously been applied to LMI problems; see for example [11] and [13]. For in-depth coverage of convex programming methods see [32] or [27].

Interior-Point Methods in Convex Programming

In the past five years, the most common approach taken for solving LMIs has been via the use of interior-point methods. Interior-point methods are used to convert constrained convex optimization problems into unconstrained optimization problems, to which general algorithms such as Newton's method may be applied. These methods utilize a *penalty* function which is added to the objective function and greatly increases the objective function cost for violation of the constraints.

The basic idea behind interior-point methods is to convert the inequality constraint functions into a *barrier* function; that is, a function whose value approaches infinity near the border of the feasible region. For example, suppose we have a standard optimization problem with the objective function denoted by f and inequality constraint functions denoted by $h_i, 1 \leq i \leq m$. One example of a barrier function commonly used is the logarithmic barrier function,

$$\phi(x, r) = -r \sum_{i=1}^{m} \ln h_i(x)$$

where the parameter r determines the weight of the function. A new unconstrained optimization problem is then given by

$$\min_x P(x, r) = \min_x (f(x) + \phi(x, r)) \tag{7.4}$$

where r decreases as the iterations progress. Typically, a Newton type minimization method is then used. Under mild conditions ([46], [27]),

$$\lim_{r\to 0}(\min_x P(x,r)) = \min_x f(x).$$

This minimum is achieved in the interior of the feasible region, since near the boundary of the feasible set, $P(x,r)$ approaches infinity.

In [53] it is shown that for a large class of logarithmic barrier functions, the resulting unconstrained (LMI) optimization problems (for example, $P(x,r)$ above) solved via interior-point algorithms satisfy a desirable property called *self-concordance*. A function P is said to be *self-concordant with value a* if $|P'''(x)| \leq 2a^{-\frac{1}{2}}(P''(x))^{\frac{3}{2}}$ for any x in the feasible domain, where the derivatives are taken along any direction in the vector space in which the feasible domain is a subset. P is strongly self-concordant if P is self-concordant and the level sets of P are closed. If P is strongly self-concordant, then minimization of P via Newton-type methods will be quadratically convergent under easily tested conditions. Nesterov and Nemirovskii further prove that tight bounds exist on the number of iterations required to find the optimum point.

There are two main classes of interior-point methods used in convex programming: *path-following methods* and *potential reduction methods.* Path-following is the more traditional approach, relying on tracking the minimizers of P as the value of r decreases. Potential reduction methods generally allow for much larger step sizes in the minimizing iterations than do the path-following methods, and as a result are more computationally attractive. A number of potential reduction methods were developed in the '80s: the method of Karmarkar, the projective method of Nemirovskii, and the primal-dual method of Todd and Ye (see [53] for references). A projective method has been used in the LMI optimization routines of the MATLAB-based software package *The LMI Control Toolbox* [30], which we have used to develop a suboptimal algorithm for balanced model reduction of uncertain systems. These algorithms are described in the following section. Primal-dual algorithms for LMI optimization have also been developed [25]. A thorough discussion of interior-point methods is provided in [53].

7.2 A Suboptimal Algorithm for Balanced Model Reduction LMIs

Recall that the LMIs with which we are concerned are system Lyapunov inequalities constructed using the realization matrices from an uncertain system realization $(\boldsymbol{\Delta}, M)$:

$$AYA^* - Y + BB^* \leq 0 \text{ and } A^*XA - X + C^*C \leq 0,$$

where Y and X are required to be semi-definite matrices which commute with the Δ structure. Ideally, we would like to find minimum rank structured Gramians Y and X to the above LMIs, or structured Gramians for which we can balance and truncate our uncertain system representations with minimal error. Thus, we would like to solve the following optimization problem:

Given an uncertain system realization $(\boldsymbol{\Delta}, M)$ *find:*

$$\min \lambda, \ \max R : \ \det(YX - \lambda I_R) = 0$$

Subject to:

$$\begin{aligned} AYA^* - Y + BB^* &\leq 0 \\ A^*XA - X + C^*C &\leq 0 \\ Y, X \geq 0, \ \ Y, X &\in \mathcal{T} \end{aligned}$$

That is, we want to find solutions Y and X for which the product YX has a very small-valued minimum eigenvalue with high multipicity. Although feasible solutions, X and Y, may easily be computed using convex programming methods (in particular, using any of the recent LMI solvers [30, 25]), the optimization problem itself is a *reduced rank LMI problem* and thus presents a more difficult computational problem. Unfortunately, reduced rank LMI problems result in neither convex nor quasi-convex optimization problems, thus we cannot directly apply the existing LMI algorithms to obtain solutions. However, some of these LMI methods may be used in heuristic algorithms to obtain suboptimal solutions for reduced rank LMI problems [17], [18].

A Supobtimal Approach

Rather than focusing on the development of general algorithms to solve minimum rank LMI problems, we have instead used a straightforward alternative to obtain suboptimal solutions for model reduction. This suboptimal algorithm is given in the following, which we henceforth refer to as the *Trace* algorithm:

Given an uncertain system realization $(\boldsymbol{\Delta}, M)$ *find:*

$$\circ \min_{Y \in \mathcal{T}} \ \mathrm{Trace}(Y) : AYA^* - Y + BB^* \leq 0, \ \text{and} \ Y \geq 0$$

and

$$\circ \min_{X \in \mathcal{T}} \ \mathrm{Trace}(X) : A^*XA - X + C^*C \leq 0, \ \text{and} \ X \geq 0.$$

Note that for any Hermitian matrix, $\mathrm{Trace}(Y) = \sum_i \lambda_i(Y)$; additionally, $\mathrm{Trace}(Y)$ is a differentiable function of the elements of the matrix Y, thus the objective function in our Trace algorithm is computationally attractive. We also note the following:

Proposition 7.1 *Given two positive semi-definite matrices Y and X,*

$$\mathit{Trace}(YX) \leq \ \mathit{Trace}(Y) \ \mathit{Trace}(X).$$

Proof. Since the Trace function is a matrix inner product,

$$\mathrm{Trace}(YX) \leq (\,\mathrm{Trace}(Y^2))^{\frac{1}{2}}(\,\mathrm{Trace}(X^2))^{\frac{1}{2}} = (\sum_i \lambda_i(Y)^2)^{\frac{1}{2}}(\sum_i \lambda_i(X)^2)^{\frac{1}{2}},$$

by the Cauchy-Schwartz inequality. Furthermore, for any $v \in \mathbb{R}^n$, $\|v\|_2 \leq \|v\|_1$, thus,

$$\mathrm{Trace}(YX) \leq (\sum_i \lambda_i(Y))(\sum_i \lambda_i(X)).$$

■

Thus, separately minimizing the trace of Y and of X leads to solutions for which the value of $\sum_i \lambda_i(YX)$ is lowered, and hence the value of each $\lambda_i(YX)$ is also often lowered.

Preliminary tests of the Trace algorithm have been completed using LMI-LAB [30]: we have constructed 20 uncertain system realizations, each with 2 to 5 uncertainty variables and dimensions ranging from 5 to 15. These realizations are constructed to be reducible, that is, for each realization there exist singular structured matrices X and Y satisfying the associated Lyapunov inequalities. Evaluation of the Trace algorithm on the test realizations is based on the eigenvalues of the resulting LMI solutions X and Y. Specifically, we consider the ratio, denoted by η, of the largest "zero valued" eigenvalue to the smallest non-zero valued eigenvalue. So, for example, if $\mathrm{eig}(X) = \{1.000,\ 0.724,\ 0.711,\ 0.531,\ 1.000 \times 10^{-6}\}$, then $\eta = 5.31 \times 10^{-5}$. As we know a priori the dimensions which may be reduced with no error for each test case, we are then able to determing the "success" or "failure" of the Trace algorithm. The results based on three different criteria for η are given in Table 7.2.

η Upper Bound	% Models Reducible
5×10^{-5}	100%
1×10^{-6}	86%
1×10^{-10}	77%

Table 7.1: Preliminary Trace Algorithm Results

Recall that finding feasible (sub)optimal solutions to the Lyapunov inequalities is only the first step of the model reduction procedure. To complete the balancing and reducing, we must also find an allowable transformation that simultaneously diagonalizes Y and X.

Simultaneous Diagonalization Algorithms

In Chapter 3, one method for computing a balancing transformation is given in equation (3.13), for which it is assumed that Y and X are strictly positive definite. This assumption is standard; if either of the (structured) Gramians is singular, the realization may be reduced without error as shown in Theorem 4.12. This method extends directly to the *structured* solutions we obtain for the Lyapunov inequalities, resulting in an *allowable* balancing transformation, as required. However, this balancing procedure is invariably poorly conditioned if either Y or X has a large condition number - a situation which is often desirable for small reduction errors. The following modified procedure helps to alleviate this problem:

Given $Y > 0$ *and* $X > 0$:	
Compute Cholesky Factorizations:	$Y = RR^*$ and $X = Q^*Q$
Calculate Singular Values:	$QR = U\Sigma V^*$
Construct Balancing Transformations:	$T = U^*Q;\ T_{\text{inv}} = RV$
Balance:	$\hat{A} = TAT_{\text{inv}};\ \hat{B} = TB;\ \hat{C} = CT_{\text{inv}}$

Note that this is not a similarity transformation, strictly speaking, since $TT_{\text{inv}} = \Sigma$; also, $TYT^* = T_{\text{inv}}^* X T_{\text{inv}} = \Sigma^2$. However, the original realization (A, B, C), is similar to $(\Sigma^{-\frac{1}{2}}\hat{A}\Sigma^{-\frac{1}{2}}, \Sigma^{-\frac{1}{2}}\hat{B}, \hat{C}\Sigma^{-\frac{1}{2}})$. Thus, after reducing $(\hat{A}, \hat{B}, \hat{C})$, the resulting lower order realization is scaled by $\Sigma_r^{-\frac{1}{2}}$, where Σ_r is the reduced order Σ. This procedure, developed for standard 1D systems [34], is also immediately applicable to uncertain system realizations. Specifically, if $Y \in \mathcal{T}$ and $X \in \mathcal{T}$, then $T \in \mathcal{T}$ and $T_{\text{inv}} \in \mathcal{T}$, assuming the singular value decomposition is computed block by block (or does not sort the singular values), so that the block structure is left intact.

The development of balancing and simultaneous diagonalization algorithms that have reasonable numerical properties has also been addressed in the literature; for example, see [44], [61] and the references therein. Detailed accounting of the computational costs of two balancing algorithms is also given in [44].

7.3 Application: A Pressurized Water Reactor

In this application, we focus on identifying and controlling a model of a Pressurized Water Reactor (PWR) in a nuclear power plant. The PWR has parameters that vary slowly over time as the fuel ages, and as the power load on the plant changes. We model the PWR as an uncertain system using the LFT framework, where the activity of the fuel and the power load are considered uncertain parameters. Although the PWR is a continuous time system, the final model we use is obtained through standard system identification techniques and is in discrete time, that is, we use sampled input/output data. Thus, the model reduction methods of Chapter 5 may be applied to the uncertain (discrete time) system models for the PWR, resulting in simplified models for which μ-synthesis techniques are then used to design controllers. We first consider only variation of the plant dynamics with the fuel reactivity, as a test case for the combination of system identification, model reduction and μ-synthesis methods; the results obtained using the systematic model reduction techniques are compared to a control design completed for a PWR model which is simplified based on engineering knowledge of the system. We then consider variations in both the operating power and the age of the fuel, assuming bounded arbitrary time-varying perturbations. Comparisons are made of the full and reduced order closed-loop models. We will not emphasize the system identification and control design processes in this discussion; see [55] and the references therein for an explanation of μ-synthesis, and [5] for a more detailed description of the identification and control design for this project. This work is part of an ongoing collaboration with Pascale Bendotti, who is employed with the Direction des Etudes Recherches, Electricite de France.

We briefly review the operation of and control objectives for the PWR. State-space models for the PWR are then discussed and LFT models are derived. The model reduction results are given, and the simulated behavior of the closed-loop models described.

Operation of the Pressurized Water Reactor

The main objective in controlling a PWR is to provide the commanded power while respecting certain physical constraints. Consider the application depicted in Figure 7.1. This is the primary circuit, and our goal is to control this part of the reactor. The pressurized water in the primary circuit transmits the heat generated by the nuclear reaction to the steam generator. In the steam generator, water of the secondary circuit is converted to steam, which drives a turbo-alternator to generate electricity.

The PWR has an inner control loop which maintains constant pressure in the primary circuit. As a result, as the steam flow increases in the secondary circuit, the temperature in the primary circuit decreases. From a control standpoint, the required power corresponds to a specific steam flow that may be viewed as a measurable disturbance. Hence, one natural control objective is to track a temperature reference derived from the steam flow. The rate of the nuclear reaction is regulated by the control rods. The rods capture neutrons, slowing down the reaction; withdrawing them increases the reaction. Because the control rods enter the top of the reactor, the rate of reaction is always higher at the bottom of the reactor. The axial offset is defined as the difference in power generated between the top and bottom of the PWR. Safety specifications require minimizing the axial offset; this also increases the lifetime of the fuel and reduces operating costs.

To achieve these objectives the PWR has two independent sets of rods. The two control inputs are the *rates of motion* of the control rods, denoted u_1 and u_2. The *positions* of the control rods, denoted v_1 and v_2, respectively, are measurable.

New nuclear fuel is more active than older fuel and thus the plant dynamics may be quite different as the fuel ages. The resulting changes observed in the behavior of the

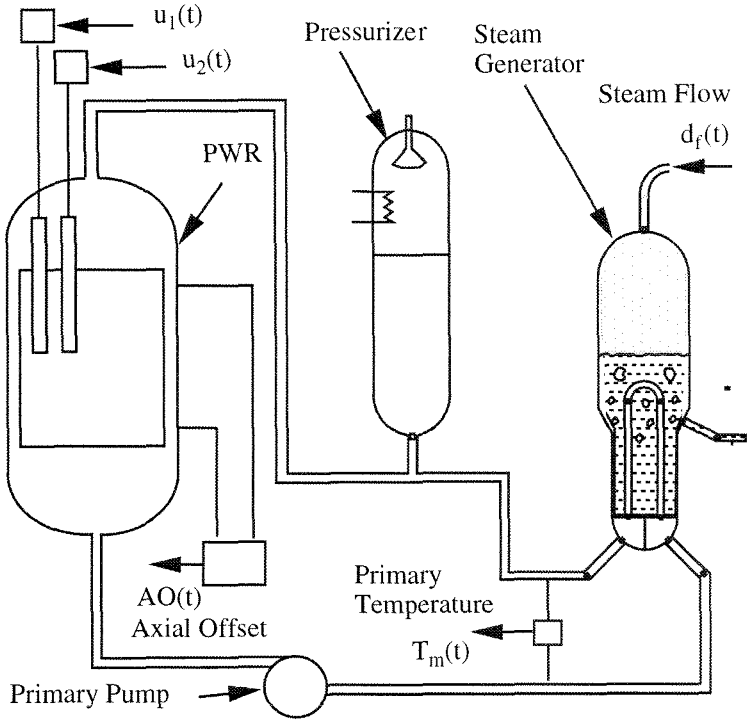

Figure 7.1: Primary Circuit and Steam Generator.

plant are described in the next section. These dynamics are also dependent on the power demand, since the load on the plant also modifies the operating point of the PWR.

Modelling the PWR

The physical system dynamics around an operating point are modelled by the transfer function:

$$y_s(t) = G_v(\boldsymbol{\lambda})v(t) + G_d(\boldsymbol{\lambda})d_f(t) \tag{7.5}$$

with

$$y_s(t) = \begin{pmatrix} T_m(t) \\ P_I(t) \\ AO(t) \end{pmatrix} \text{ and } v(t) = \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix}$$

where $\boldsymbol{\lambda}$ denotes the delay operator. The inputs for this model are the steam flow disturbance d_f and the positions of the two sets of rods; the outputs are the mean temperature T_m, the primary power P_I, and the axial offset AO, respectively (see Figure 7.1).

To quantify the change in behavior of the plant as the fuel gets older, and as the power load varies, a nonlinear simulator has been used which is based on first principles models. It includes models for the pressurizer, steam generator, and the turbine, but not the alternator. Models have been identified using the technique presented in [7]; the identification method consists of estimating coefficients for state-space realizations of the system at different operating points by minimizing a quadratic criterion via an iterative Gauss-Newton algorithm [45].

As an example of the variation of the system dynamics with changes in the system operating point, step-responses of the models are shown in Figure 7.2 for the fuel at half-life, and operating at both $0.5\mathcal{P}_n$ (dashed lines) and $0.99\mathcal{P}_n$ (mixed lines); these models are denoted by $\tilde{G}_{50}$, and $\tilde{G}_{99}$, respectively. For comparison purposes, step responses corresponding to models assuming the fuel is new, and operating at $0.5\mathcal{P}_n$ (denoted G_{50}), are plotted in the same Figure (solid lines).

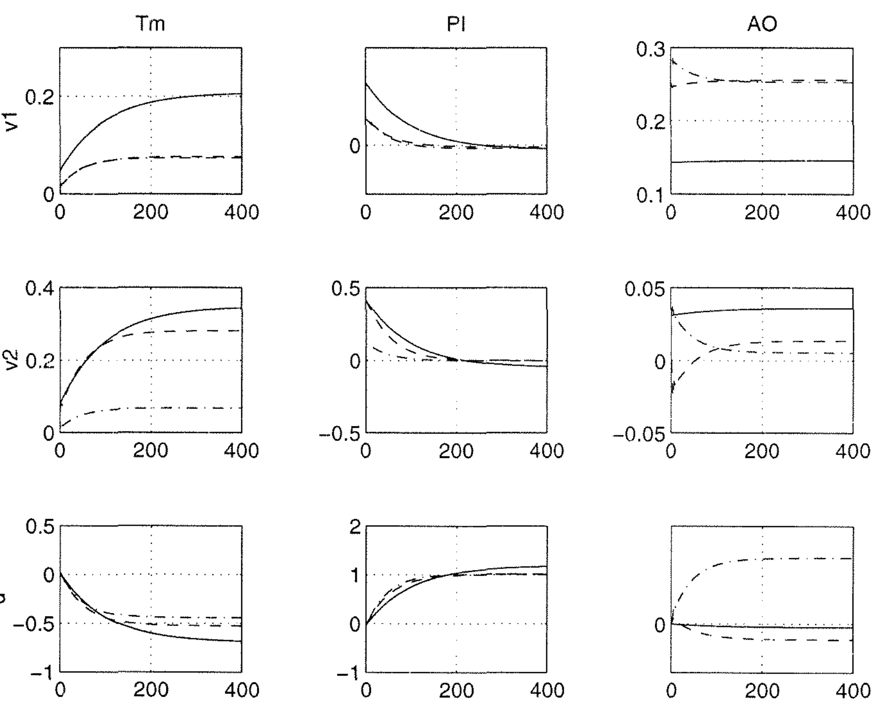

Figure 7.2: Step-responses of G_{50}, $\tilde{G}_{50}$ and $\tilde{G}_{99}$

PWR Model Variations with Fuel Age

As the fuel gets older it is less active, thus the control authority is decreased. However, the plant is more maneuverable, that is the plant has lower gains and shorter time constants for its dominant dynamics. Note in Figure 7.2 that the gains corresponding to the second control position, ν_2, do not change as much as those of the first control position, ν_1, with respect to the fuel age. This is because the static characteristic of the second set of rods is periodically re-identified, hence the corresponding gains are rescaled. At lower power, the second control input has more authority than the first, especially when the fuel is older, and using the second control results in a smaller axial offset. At high power, however, the second control input has less authority, so more control must come from the first control input.

In considering only the variation of the plant dynamics with the fuel age, we restrict the operating range, and only the behavior of the plant around $0.5\mathcal{P}_n$ is considered. However, previous work has shown that a robust controller designed at $0.5\mathcal{P}_n$ will maintain performance up to $0.9\mathcal{P}_n$ [7]. In order to simplify the notation, the subsystem G_V defined in (7.5) and obtained from the models G_{50} and $\tilde{G}_{50}$ will be denoted G and $\tilde{G}$, respectively.

If we consider state-space models like those discussed in Chapter 3 for the PWR, the coefficients in the realization matrices are not fixed, but vary slowly over the lifetime of the fuel, thus their exact values are uncertain at any fixed point in time. To model this uncertainty, we assume the coefficients in the realization may take any value in a fixed interval, that is we consider parametric uncertainty as follows:

$$p(\delta) = p_{nom} + p_{del}\delta \tag{7.6}$$

where p represents a vector of uncertain coefficients and δ is a vector of bounded, scalar, slowly time-varying parameters.

The uncertain system model with realization matrix

$$M(\delta) = \begin{bmatrix} A(\delta) & B(\delta) \\ C(\delta) & D(\delta) \end{bmatrix}$$

is rewritten as an LFT, $S(\delta, \boldsymbol{\lambda}) = \Delta \star M$, where the diagonal uncertainty structure $\boldsymbol{\Delta}$ is comprised of repeated entries of δ and the delay operator $\boldsymbol{\lambda}$, and M is a constant matrix. M has much larger dimensions than $M(\delta)$. In particular for the PWR, two uncertain sytems are considered. The first, denoted S_1, contains a subset of uncertain coefficients selected based on engineering knowledge of the physical system. The second system, denoted S_2, where all of the coefficients are uncertain, is subject to model reduction. The respective realization matrices $M_1(\delta)$ and $M_2(\delta)$ are as follows:

$$M_1(\delta) = \left[\begin{array}{c|cc} a(\delta) & b_{V_1}(\delta) & b_{V_2} \\ \hline c_{Tm}(\delta) & d_{Tm_1}(\delta) & d_{Tm_2}(\delta) \\ c_{P_I} & d_{PI_1}(\delta) & d_{PI_2} \\ c_{AO} & d_{AO_1}(\delta) & d_{AO_2} \end{array}\right] \quad \text{and}$$

$$M_2(\delta) = \left[\begin{array}{c|cc} a(\delta) & b_{V_1}(\delta) & b_{V_2}(\delta) \\ \hline c_{Tm}(\delta) & d_{Tm_1}(\delta) & d_{Tm_2}(\delta) \\ c_{P_I}(\delta) & d_{PI_1}(\delta) & d_{PI_2}(\delta) \\ c_{AO}(\delta) & d_{AO_1}(\delta) & d_{AO_2}(\delta) \end{array}\right].$$

The coefficients related solely to the steam flow demand, d_f, do not vary with the age of the fuel, thus only two inputs, the control rod positions ν_1 and ν_2, are listed above.

To derive the realization corresponding to S_1, each term of the two first-order models G and $\tilde{G}$ is compared. The greatest variation in responses is related to the first input (see plots on the first row in Figure 7.2), hence we model the coefficients in the second column of the realization $M_1(\delta)$ as being uncertain. To account for the variation in the dominant time constant and the temperature gain, the coefficients a and c_{Tm} are also modelled as being uncertain. Finally, the slight variation in the authority of ν_2 is reflected in the direct term d_{Tm_2}.

Placing the representation of S_1 into the LFT framework with a constant realization matrix M_1 results in a Δ_1 structure defined by

$$\boldsymbol{\Delta}_1 = \{\mathrm{diag}\,[\delta I_7, \;\; \boldsymbol{\lambda}] : |\delta| \le 1 \;\; \delta, \boldsymbol{\lambda} \in \mathcal{L}(l_2)\}.$$

The representation for S_2 is similarly placed in the LFT framework with M_2 constant and

$$\boldsymbol{\Delta}_2 = \{\mathrm{diag}\,[\delta I_{12}, \;\; \boldsymbol{\lambda}] : |\delta| \le 1; \delta, \boldsymbol{\lambda} \in \mathcal{L}(l_2)\}.$$

Model Reduction and Design Results: Variations with Fuel Age

The objective is to determine if S_2 is reducible, in particular with respect to the uncertainty block δI_{12}. Applying the model reduction results discussed in Chapter 5, we first obtain solutions $X > 0$ and $Y > 0$ to the Lyapunov inequalities using the Trace algorithm discussed in the preceeding section. The elements of the diagonalized matrix $\Sigma = Y = X$, which we refer to as the *generalized singular values* for the model, are given in Table 7.3. The realization is then balanced and truncated. We obtain a reduced realization for the

block	Singular Values							
δI_{12}	0.5709	0.0577	0.0512	0.0377	0.0048	0.0024	0.0021	$0.0020 I_5$
$\boldsymbol{\lambda}$	1.3657							

Table 7.2: Generalized Singular Values for the PWR

uncertain system S_2, with an uncertainty structure of

$$\boldsymbol{\Delta}_{2r} = \{\mathrm{diag}\,[\delta I_4, \;\; \boldsymbol{\lambda}] : |\delta| \le 1; \; \delta, \boldsymbol{\lambda} \in \mathcal{L}(l_2)\}.$$

The model reduction error bound based on the results of Chapter 5 is 1.13×10^{-2}. As a check on this predicted error bound, a magnitude response plot of

$$\sup_{|\delta|\le 1} \|(\boldsymbol{\Delta}_2 \star M_2) - (\boldsymbol{\Delta}_{2r} \star M_{2r})\|_{l_2 \to l_2}$$

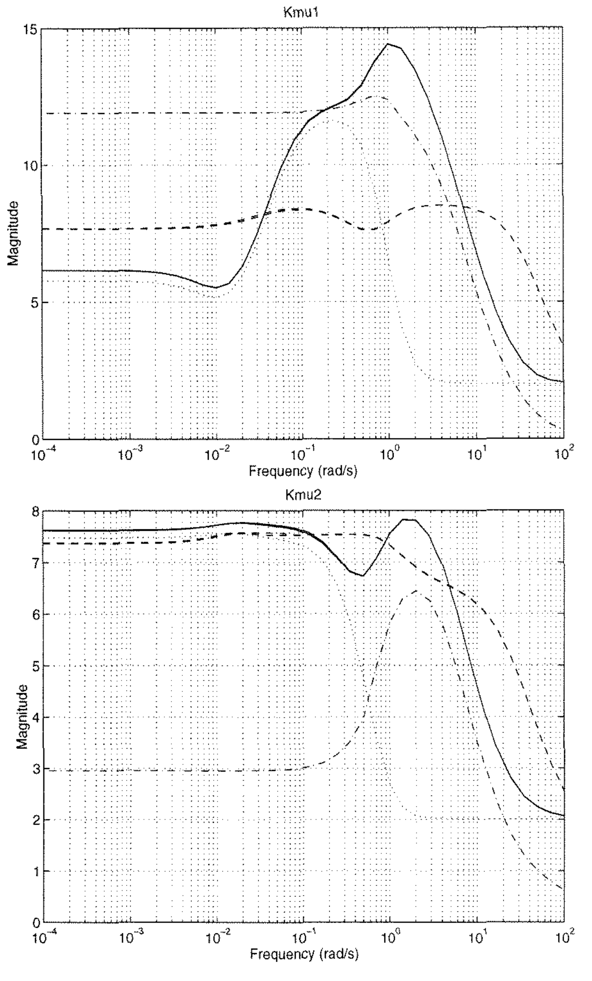

Figure 7.3: RS, RP and μ bounds: 1st and 2nd D-K iteration

is used to determine an observed error bound. From this plot, the maximum observed error is less than 2.5×10^{-4}. We denote the reduced realization by S_{2r}.

Using the models S_1 and S_{2r}, controllers were designed using complex μ-synthesis, which we denote by Kmu1 and Kmu2 (details can be found in [5]). The robust performance (RP) μ-bounds for the closed-loop systems are shown in Figure 7.3 (solid and dashed lines), where two iterations (of bounds computation) were completed; also shown are the nominal performance bounds (dotted lines) and the robust stability (RS) bounds (mixed lines). Basically, lower valued bounds indicate better system performance. Thus, from a μ-theory perspective, the results for the controller designed using S_{2r}, the reduced model, are better. As a result, we can trade robustness over a larger and more conservative uncertainty set for increased performance on a smaller but more appropriate uncertainty set.

The behavior of the closed-loop systems was also evaluated using the nonlinear simulator of the PWR. The nuclear fuel is assumed to be at half of its expected lifetime. Figure

7.4 shows the simulation results for Kmu1, and Figure 7.5 shows the same simulation responses for Kmu2. In each of the Figures, T_m (solid lines) is the mean temperature, and T_{ref} (dashed lines) is the reference temperature; P_I (solid lines) is the primary output power and d_f (dashed lines) is the steam demand; AO is the axial offset. The control signals are plotted in solid (u_1) and dashed (u_2) lines.

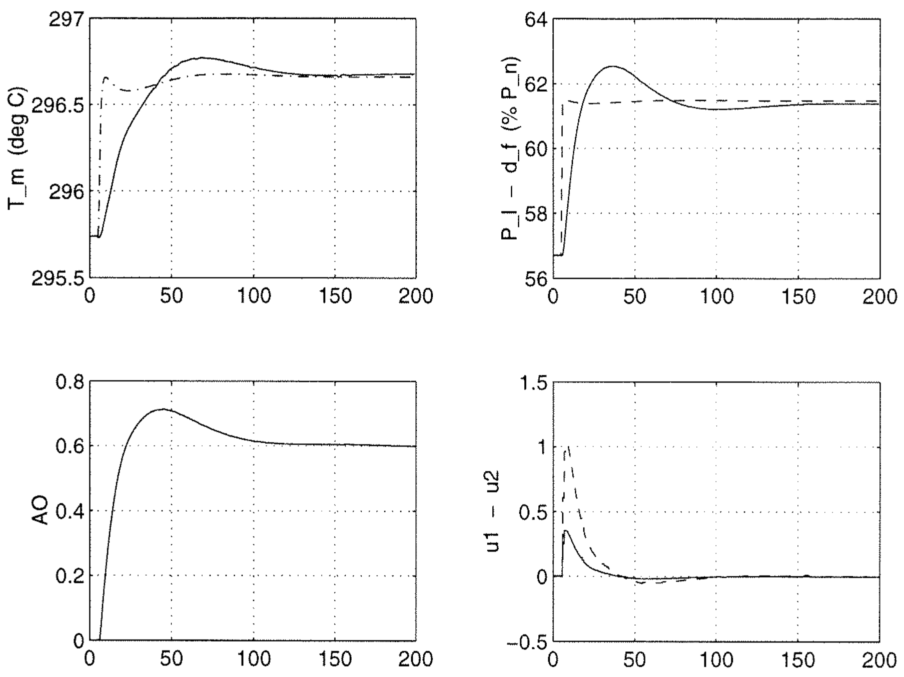

Figure 7.4: Closed-loop response: Controller Kmu1

The two control designs result in similar step response behavior, however, note that the closed-loop response with Kmu2 results in less axial offset than that for Kmu1. Both controllers result in minimal use of the control rods, as the controls do not change sign for the commanded step change; this lengthens the lifetime of the fuel. Overall the controller designed using the reduced model S_{2r} performs better than that using the model simplified based on engineering experience.

PWR Model Variations with Fuel Reactivity and Operating Power

We now consider simultaneous variations in the state-space coefficients with respect to the age of the fuel and the power demand. The state-space realization for the system is

$$M(\delta_1, \delta_2) = \left[\begin{array}{c|cc} a(\delta_1, \delta_2) & b_{v_1}(\delta_1, \delta_2) & b_{v_2}(\delta_1, \delta_2) \\ \hline c_{Tm}(\delta_1, \delta_2) & d_{Tm_1}(\delta_1, \delta_2) & d_{Tm_2}(\delta_1, \delta_2) \\ c_{P_I}(\delta_1, \delta_2) & d_{PI_1}(\delta_1, \delta_2) & d_{PI_2}(\delta_1, \delta_2) \\ c_{AO}(\delta_1, \delta_2) & d_{AO_1}(\delta_1, \delta_2)) & d_{AO_2}(\delta_1, \delta_2) \end{array}\right].$$

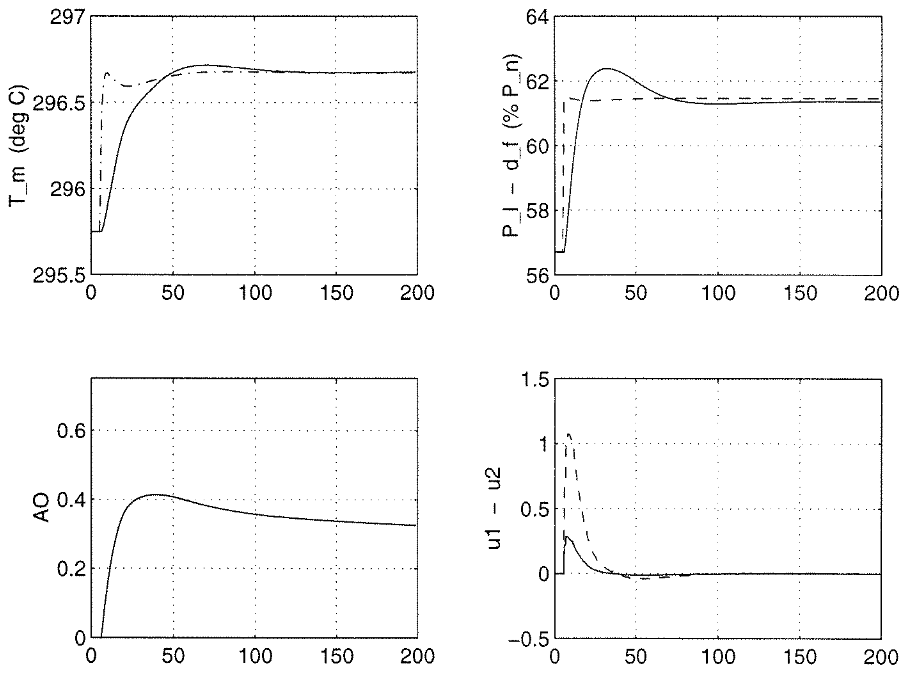

Figure 7.5: Closed-loop response: Controller Kmu2

The coefficients in $M(\delta_1, \delta_2)$ vary with the uncertainty in a more complicated manner than described before. The parameter dependence is evaluated over a range of operating power (δ_1) and fuel reactivity (δ_2). All coefficients of the selected first-order models are individually fit with rational functions of the form $e + f\delta_1(1 - g\delta_1)^{-1}h$ and $l + m\delta_2$, for which the coefficients are obtained using a least-squares technique. The complete parameterization is of the form

$$(l + m\delta_2)\frac{e + (fh - eg)\delta_1}{1 - g\delta_1} + (l' + m'\delta_2) \tag{7.7}$$

where $-1 \leq \delta_i \leq 1, i = 1, 2$. We can again rewrite this as an LFT, $S(\delta_1, \delta_2, \boldsymbol{\lambda}) = (\boldsymbol{\Delta} \star M)$, where the uncertainty structure for the *full* LFT realization is

$$\boldsymbol{\Delta} = \{\operatorname{diag}[\delta_1 I_{12},\ \delta_2 I_{12},\ \boldsymbol{\lambda}] : |\delta_i| \leq 1;\ \delta_i, \boldsymbol{\lambda} \in \mathcal{L}(l_2)\},$$

and M is a constant matrix. See [6] for more details on modelling and control synthesis for the PWR with respect to two uncertain parameters.

With two varying parameters, this model is too complicated to simplify using intuition about the physical operations of the system. Thus, we compare μ-synthesis controllers designed using the full model and a reduced model.

Model Reduction and Design Results: Variations with Fuel Age and Operating Power

Applying the model reduction method of Chapter 5, we obtain solutions $X > 0$ and $Y > 0$ to the Lyapunov inequalities, again via the Trace algorithem, and a diagonalized matrix $\Sigma = Y = X$ with elements shown in Table 7.3.

block	Singular Values							
$\delta_1 I_{12}$	1.0014	0.3054	0.1239	0.0812	0.0229	0.0058	$0.0055 I_5$	0.0051
$\delta_2 I_{12}$	1.2099	0.5362	0.1876	0.0402	0.0193	0.0074	$0.0055 I_5$	0.0037
$\boldsymbol{\lambda}$	1.9040							

Table 7.3: Generalized Singular Values for the PWR

The realization $(\boldsymbol{\Delta}, M)$ is then balanced and truncated. We obtain a reduced realization for the uncertain system, denoted by S_r, with an uncertainty structure of

$$\boldsymbol{\Delta}_r = \{\text{diag}\,[\delta_1 I_4,\ \delta_2 I_3,\ \boldsymbol{\lambda}] : |\delta_i| \leq 1;\ \delta_i, \boldsymbol{\lambda} \in \mathcal{L}(l_2)\}.$$

The model reduction error bound based on the results of Chapter 5 is 1.1×10^{-1}. As a check on this predicted error bound, magnitude response plots are used to determine an observed error bound. From these plots, the maximum observed error is less than 4.5×10^{-3}.

Controllers for both the full and reduced models were designed using μ-synthesis methods. RP and RS bounds have been compared, and nonlinear simulations run on the two closed-loop systems. There is no degradation in performance with the controller designed for the reduced model; the RP bounds are essentially the same. Nonlinear simulations of the closed-loop systems are shown in Figures 7.6 and 7.7. The nuclear fuel is assumed to be at half of its expected lifetime, and the load at maximum operating power. These responses, along with the μ-performance analysis, indicate that the controller designed using the reduced model performs almost identically to that designed using the full model.

A general discussion of methods for computing reduced uncertain system models has been given. One simple approach was implemented and applied to LFT models for the PWR circuit of a nuclear power plant. The results, to date, of this application show the model reduction techniques described in this thesis to be beneficial in the control design and analysis of uncertain systems.

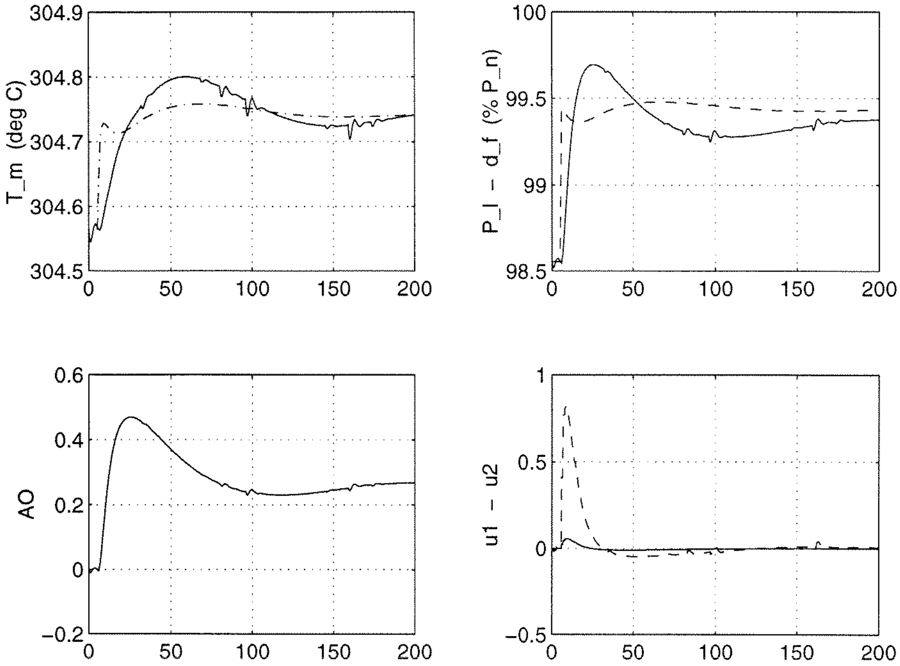

Figure 7.6: Closed-loop response: Full Order Controller Design

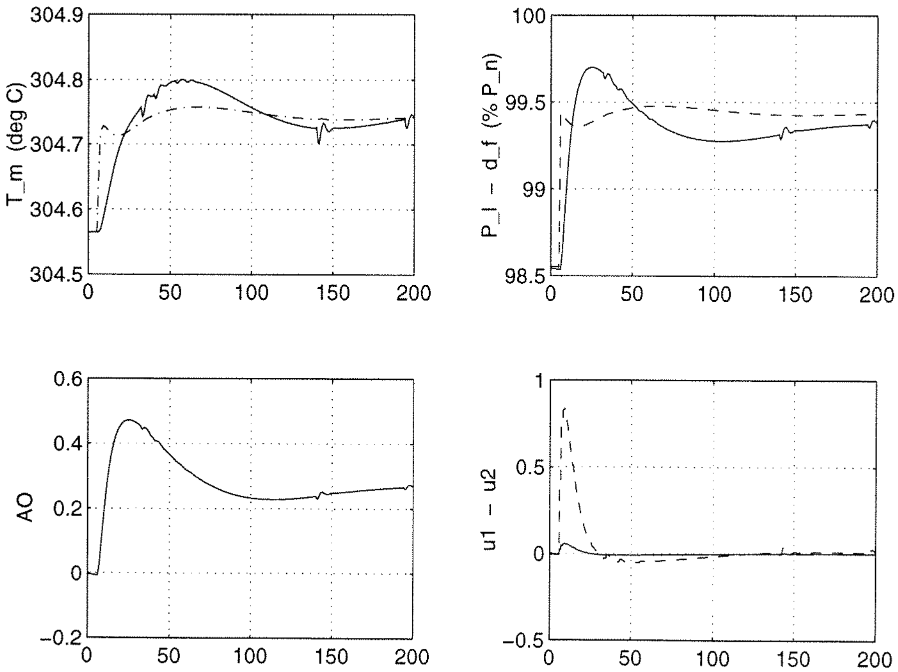

Figure 7.7: Closed-loop response: Reduced Order Controller Design

Chapter 8

Conclusions

We conclude with a brief summary of the work presented in this thesis, and a few remarks on future directions for related research.

8.1 Summary

The emphasis of this research has been on model reduction methods and the development of related realization theory for uncertain systems. These uncertain systems are generally modelled by a nominal LTI system and a set description of uncertainty or perturbations on the system. Using the LFT framework, we can rewrite these models as a LFT of a constant realization matrix on a block diagonal structured set containing both uncertainty and delay operators. In this setting, we can compute guaranteed error bounds for reducing the dimensions of both the state-space and the uncertainty description by generalizing the notions of balancing and truncating state-space realizations.

When the model reduction can be completed with zero error, the model is said to be reducible. Naturally, this has led us to consider related realization theory concepts for these uncertain system models, most obviously reducibility and minimality, and to a limited extent the construction of meaningful controllability and observabilty matrices in this setting. We have presented a necessary and sufficient reducibility condition for systems represented by LFTs on structured uncertainty sets, which is also sufficient for multi-dimensional state-space realizations, and which has allowed us to draw connections to nimimal representations of formal power series.

Practical aspects of this thesis include the implementation of the model reduction methods for uncertain systems in Matlab subroutines, and the subsequent application to a pressurized water reactor of a nuclear power plant.

8.2 Future Research

We note related future research topics.

- In this thesis, we have addressed the reduction of *stable* uncertain systems. Obviously, reduction methods and error bounds for *unstable* uncertain systems are needed. One immediate method would be to scale (or damp) the realization and uncertainty set, as discussed in Section 4.5, and apply the techniques discussed herein. Alternatively, a coprime factorization approach could be used; an initial attempt at this is discussed in [4].

- We have focused on balanced truncation model reduction techniques. Other methods such as singular perturbation approximation and frequency weighted techniques are also being developed.

- Realization theory for uncertain systems can be further developed; complete connections among all existing results have yet to be established.

- We presented one suboptimal algorithm for computing reduced realizations. Although this algorithm is relatively fast, and is easily implemented, the development of good heuristics for computing reduced rank solutions to general LMI problems would benefit not only the model reduction and reducibility results we have discussed in this thesis, but would also have more widespread application within the control community. In particular, the primal-dual potential reduction methods discussed in [68] may be of use for this problem type.

- Lower error bounds for model reduction of uncertain systems are currently being investigated.

Appendix A

Robust l_2-Stability Results

We provide a sketch of the proof for Theorem 4.6, the LMI condition for robust l_2 stability to block structured time-varying uncertainty. Sufficiency of this stability condition can be shown using established analysis results given in Chapter 2. The necessity condition is somewhat trickier: the following lemmas, due to Paganini [57], will be used in the proof.

Lemma A.1 *Let $z \in l_2^n$ and $v \in l_2^m$. The following are equivalent:*

$$(i)\ \|v\|_2^2 - \|z\|_2^2 \geq 0$$
$$(ii)\ \text{there exists } \Delta \in \mathcal{L}(l_2^n, l_2^m),\ \|\Delta\| \leq 1 :\ \Delta v = z$$

Lemma A.2 *Let $z, v \in l_2^d$. The following are equivalent:*

$$(i)\ \int_{-\pi}^{\pi} v(e^{j\omega})v(e^{j\omega})^* - z(e^{j\omega})z(e^{j\omega})^* d\omega \geq 0$$
$$(ii)\ \text{for all } \eta \in \mathbb{C}^d,\ \|\eta^* v\|_2 \geq \|\eta^* z\|_2$$
$$(iii)\ \text{there exists } \delta \in \mathcal{L}(l_2),\ \|\delta\| \leq 1 :\ \delta I_d v = z$$

Lemmas A.1 and A.2 are immediate, with the exception of Lemma A.2, $(ii) \Rightarrow (iii)$, which we do not use directly, and hence will not discuss in detail.

The uncertainty structure defined in (4.1), with both full and repeated scalar blocks is considered. Let $z \in l_2^n$, and consider the conformal partitioning of z,

$$z = [z_1^T \cdots z_p^T\ z_{p+1}^T \cdots z_{p+f}^T]^T.$$

An analogous partitioning of Az can be made. Consider the following functions of z and $\xi = Az$:

$$\Phi_i(z) = \sum_{k=-\infty}^{\infty} \xi_i(k)\xi_i(k)^* - z_i(k)z_i(k)^*,\ i = 1, \dots, p$$
$$\rho_{p+j}(x) = \left\|(Ax)_{p+j}\right\|^2 - \left\|z_{p+j}\right\|^2,\ j = 1, \dots, f$$
$$\Lambda(z) = \text{diag}[\Phi_1(z), \dots, \Phi_p(z), \rho_{p+1}(z)I_{m_1}, \dots, \rho_{p+f}(z)I_{m_f}].$$

Recall that the set of constant matrices that commute with Δ is denoted by $\mathcal{T}$. Now consider the following sets:

$$\mathcal{T}^+ = \{T \in \mathcal{T} : T > 0\},\ \overline{\mathcal{T}}^+ = \{T \in \mathcal{T} : T \geq 0\},\ \text{and}$$

$$\nabla = \{\Lambda(z) : \|z\| = 1\}.$$

Note that we can define an inner product space based on $\mathcal{T}$ by defining

$$\langle T, W\rangle = \sum_{i=1}^{p} \mathrm{Trace}(T_i W_i) + \sum_{j=1}^{f} t_{p+j} w_{p+j}.$$

Lemma A.3 *[57] ∇ is convex and compact in the inner product space $\mathcal{T}$.*

Proof of Theorem 4.6:

Sufficiency follows from a straightforward application of spectra properties of operators on Hilbert spaces. First note that there exists $Y > 0, Y \in \mathcal{T}$ such that $AYA^* - Y < 0$ if and only if $\overline{\sigma}(Y^{-\frac{1}{2}} A Y^{\frac{1}{2}}) < 1$. Therefore $\left\|Y^{-\frac{1}{2}} A Y^{\frac{1}{2}} \Delta\right\| \leq \left\|Y^{-\frac{1}{2}} A Y^{\frac{1}{2}}\right\| \|\Delta\| < 1$ for all $\Delta \in \boldsymbol{B}_\Delta$, and as a result

$$1 \notin \mathrm{spec}(Y^{-\frac{1}{2}} A Y^{\frac{1}{2}} \Delta), \text{ for any } \Delta \in \boldsymbol{B}_\Delta$$

by Lemma 2.1. Thus

$$(I - Y^{-\frac{1}{2}} A Y^{\frac{1}{2}} \Delta) = Y^{-\frac{1}{2}}(I - A\Delta)Y^{\frac{1}{2}} \text{ is invertible in } \mathcal{L}(l_2) \text{ for all } \Delta \in \boldsymbol{B}_\Delta.$$

The result is then obvious.

Necessity is shown via two main steps: first, if $(I - A\Delta)$ is invertible for all $\Delta \in \boldsymbol{B}_\Delta$, then $\nabla \cap \overline{\mathcal{T}}^+ = \emptyset$; and second, $\nabla \cap \overline{\mathcal{T}}^+ = \emptyset$ implies there exists $Y \in \overline{\mathcal{T}}^+ : AYA^* - Y < 0$.

Assume $\nabla \cap \overline{\mathcal{T}}^+ \neq \emptyset$, then there exists $z \in l_2 :\ \|z\| = 1$ and $\Lambda(z) \geq 0$. Let $\xi = Az$. Then $\left\|\xi_{p+j}\right\|^2 - \left\|z_{p+j}\right\|^2 \geq 0$, for $j = 1, \ldots, f$ and $\sum_k \xi_i(k)\xi_i(k)^* - z_i(k)z_i(k)^* \geq 0$, for $i = 1, \ldots, p$. By applying Lemmas A.1 and A.2, we can construct a $\Delta \in \boldsymbol{B}_\Delta$, LTV, such that $\Delta A z = z$, leading to a contradiction.

Now, since ∇ and $\overline{\mathcal{T}}^+$ are disjoint convex sets in the inner product space $\mathcal{T}$, and ∇ is compact and $\overline{\mathcal{T}}^+$ is a closed cone, we can use a separating hyperplane argument to find $Y \in \mathcal{T}$ and $\eta > 0$ such that

$$\langle Y, \Lambda\rangle \leq -\eta < 0 \leq \langle Y, T\rangle$$

for all $\Lambda \in \nabla$ and $T \in \overline{\mathcal{T}}^+$. Note that $\langle Y, T\rangle \geq 0$ for every $T \in \overline{\mathcal{T}}^+$ implies $Y \geq 0$. Using a small perturbation of Y to ensure $Y > 0$, then by continuity and compactness of ∇ we can scale η to $\overline{\eta}$ such that $\langle Y, \Lambda\rangle \leq -\overline{\eta} < 0$, for all $\Lambda \in \nabla$.

Then, for any $z \in l_2^n$, $\|z\| = 1$, $\xi = Az$,

$$\langle Y, \Lambda(z)\rangle = \sum_{i=1}^{p} \mathrm{Trace}(Y_i \Phi_i(z)) + \sum_{j=1}^{f} y_{p+j}\rho_{p+j}(z) = \langle (A^* Y A - Y)z, z\rangle \leq -\overline{\eta}.$$

■

Appendix B

Necessary and Sufficient Reducibility Condition: An Alternate Proof

An alternative proof for Lemma 4.13 is given which is based on evaluating reducibility of the system representation in the structured induced 2-norm. We use the form of the SI2-norm given in Lemma 5.2 to prove Lemma 4.13, thus again this condition is necessary for time-varying operator uncertainty.

The Bolzano-Weierstrass theorem (see [48] for example) is used without reference in this proof of Lemma 4.13. The following lemma is also used in this proof.

Lemma B.1 *Suppose we are given two sequences, $\sigma_k \in \mathbb{R}$ and $b_k \in \mathbb{R}^n$, such that $\lim_{k\to\infty} \sigma_k b_k = 0$, and $\lim_{k\to\infty} b_k = b_0$, where b_k and b_0 are bounded. Then*

$$\lim_{k\to\infty} \sigma_k b_k = \lim_{k\to\infty} \sigma_k b_0 = 0.$$

Proof. If $b_0 = 0$, the result is trivial.

If $b_0 \neq 0$, then for any $\epsilon > 0$, there exists N such that $\|b_0 - b_k\|_2 < \epsilon$ for all $k \geq N$. But $\|\sigma_k b_k\|_2 \geq \|\sigma_k\|(\|b_0\|_2 - \epsilon)$, implying $\lim_{k\to\infty} \sigma_k = 0$, thus $\lim_{k\to\infty} \sigma_k b_0 = 0$. ■

Proof of Lemma 4.13:

We first consider the special case $\Delta \in \boldsymbol{\Delta}$ where

$$\boldsymbol{\Delta} = \left\{\text{diag}\left[\delta_1, \delta_2, \ldots, \delta_p\right] : \delta_i \in \mathcal{L}(l_2)\right\}.$$

Thus, the matrix set $\mathcal{T}$ consists of strictly diagonal matrices, that is,

$$\mathcal{T} = \{\text{diag}[t_1, t_2, \ldots, t_p] : t_i \in \mathbb{R}\}.$$

By assumption, and Lemma 5.2, $\|\boldsymbol{\Delta} \star M\|_{SI2} =$

$$\inf\left\{\gamma : \text{ there exists } T \text{ such that } \overline{\sigma}\left(\begin{bmatrix} TAT^{-1} & \frac{1}{\gamma^{\frac{1}{2}}}TB \\ \frac{1}{\gamma^{\frac{1}{2}}}CT^{-1} & \frac{1}{\gamma}D \end{bmatrix}\right) < 1\right\} = 0,$$

which implies we can construct a sequence $y_k \in \mathbb{R}^+$, where $\lim_{k\to\infty} y_k = 0$, and an accompanying sequence of matrices, $T_k \in \mathcal{T}$, such that for each $y_k > 0$,

$$\overline{\sigma}\left(\begin{bmatrix} T_k A T_k^{-1} & \frac{1}{y_k^{\frac{1}{2}}} T_k B \\ \frac{1}{y_k^{\frac{1}{2}}} C T_k^{-1} & \frac{1}{y_k} D \end{bmatrix}\right) < 1.$$

Without loss of generality, we assume that

$$\lim_{k\to\infty} \overline{\sigma}(T_k A T_k^{-1}) \leq 1, \quad \lim_{k\to\infty} T_k B = 0,$$

$$\text{and } \lim_{k\to\infty} C T_k^{-1} = 0, \tag{B.1}$$

otherwise, we can construct a subsequence of T_k to satisfy these limits. If neither T_k nor T_k^{-1} converge to a finite limit, we similarly assume they converge in an extended sense by refining the sequences: $\lim_{k\to\infty} T_k = \hat{T}$, and $\lim_{k\to\infty} T_k^{-1} = \hat{T}^{-1}$, where

$$\hat{T} = \text{diag}[0,\ldots,0,\sigma_1,\ldots,\sigma_m,\infty,\ldots,\infty],$$

$$\hat{T}^{-1} = \text{diag}[\infty,\ldots,\infty,\sigma_1^{-1},\ldots,\sigma_m^{-1},0,\ldots,0]$$

and $m < p$.

Partitioning A, B and C correspondingly, and applying (B.1), we obtain an uncontrollable/unobservable decomposition, that is

$$\hat{T}B = \hat{T}\begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} = 0 \text{ implies } B_2 = 0 \text{ and } B_3 = 0;$$

$$C\hat{T}^{-1} = [C_1 \; C_2 \; C_3]\hat{T}^{-1} = 0 \text{ implies } C_1 = 0 \text{ and } C_2 = 0;$$

and

$$\hat{T}A\hat{T}^{-1}\hat{T}^{-1*}A^*\hat{T}^* \leq I \text{ implies } A_{21} = 0,\; A_{31} = 0 \text{ and } A_{32} = 0.$$

Thus we can write $M = \begin{bmatrix} A_{11} & A_{12} & A_{13} & B_1 \\ 0 & A_{22} & A_{23} & 0 \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & C_3 & 0 \end{bmatrix}$.

We know by stability that there exists a nonsingular $\tilde{T} \in \mathcal{T}$ such that $\overline{\sigma}(\tilde{T}A\tilde{T}^{-1}) < 1$. We then partition $\tilde{T}$ into three subblocks corresponding to the partitioning of A, that is, $\tilde{T} = \begin{bmatrix} \tilde{T}_1 & 0 & 0 \\ 0 & \tilde{T}_2 & 0 \\ 0 & 0 & \tilde{T}_3 \end{bmatrix}$ and define

$$\tilde{Y} = \begin{bmatrix} \tilde{T}_1^{-1}\tilde{T}_1^{-1*} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \tilde{X} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \tilde{T}_3^*\tilde{T}_3 \end{bmatrix}.$$

Since $A_{11}\widetilde{T}_1^{-1}\widetilde{T}_1^{-1*}A_{11}^* - \widetilde{T}_1^{-1}\widetilde{T}_1^{-1*} < 0$ and $A_{33}^*\widetilde{T}_3^*\widetilde{T}_3 A - \widetilde{T}_3^*\widetilde{T}_3 < 0$, by scaling $\widetilde{X}$ and $\widetilde{Y}$ we can then find X and Y which satisfy conditions $(i) - (iii)$.

A similar, but less general, argument is used to show that conditions $(i) - (iii)$ hold if either $\lim_{k_j \to \infty} T_{k_j}$ and/or $\lim_{k_j \to \infty} T_{k_j}^{-1}$ exist.

The proof for uncertainty structures with repeated δ_i is obtained by extending the above using fairly standard linear algebra and analysis techniques. In this case the commutative matrices, T_k, are block diagonal, rather than strictly diagonal. Using singular value decompositions we set $T_k = U_k \Sigma_k V_k^*$, where Σ_k is strictly diagonal and U_k and V_k are block diagonal unitary matrices. Since the maximum singular value is unitarily invariant, we drop the U_k terms and rewrite the limits in (B.1),

$$\lim_{k \to \infty} \overline{\sigma}(\Sigma_k V_k^* A V_k \Sigma_k^{-1}) \leq 1, \quad \lim_{k \to \infty} \Sigma_k V_k^* B = 0,$$

$$\text{and } \lim_{k \to \infty} C V_k \Sigma_k^{-1} = 0, \tag{B.2}$$

again, without loss of generality, assuming we've properly chosen the subsequence k so that these limits are satisfied.

Since $\boldsymbol{V} = \{V_k : V_k^* V_k = I\}$ is a closed bounded set, V_k has a convergent subsequence in $\boldsymbol{V}$. Similarly, $W_k = \Sigma_k V_k^* A V_k \Sigma_k^{-1}$ has a convergent subsequence with unity bounded maximum singular value. We take a series of subsequences denoted by $\hat{k}$ such that $\lim_{\hat{k} \to \infty} \gamma_{\hat{k}} = 0$, $\lim_{\hat{k} \to \infty} V_{\hat{k}} = V_0$ and $\lim_{\hat{k} \to \infty} W_{\hat{k}} = W_0$, where $V_0 \in \boldsymbol{V}$ and $W_0^* W_0 \leq I$. Denote $B_0 = V_0^* B$, $C_0 = C V_0$ and $A_0 = V_0^* A V_0$. Applying Lemma B.1, we can show that the limits in (B.2) can be written as

$$\lim_{\hat{k} \to \infty} \Sigma_{\hat{k}} B_0 = 0, \quad \lim_{\hat{k} \to \infty} C_0 \Sigma_{\hat{k}}^{-1} = 0 \tag{B.3}$$

and,

$$\lim_{\hat{k} \to \infty} W_{\hat{k}} = \lim_{\hat{k} \to \infty} \Sigma_{\hat{k}} A_{\hat{k}} \Sigma_{\hat{k}}^{-1} = W_0 \tag{B.4}$$

where $A_{\hat{k}} = V_{\hat{k}}^* A V_{\hat{k}}$, and $\lim_{\hat{k} \to \infty} A_{\hat{k}} = A_0$. We then apply methods of the preceeding proof to each δ_i block of $\Sigma_{\hat{k}}$ and A_0, B_0 and C_0 to obtain an uncontrollable/unobservable like decomposition. Finally, we use stability and permutations to obtain matrices, $X \geq 0$ and $Y \geq 0$, both in $\mathcal{T}$, satisfying $(i) - (iii)$.

In particular, consider the case $\boldsymbol{\Delta} = \{\text{diag}[\delta_1 I_{n_1},\ \delta_2 I_{n_2}]\}$. Suppose neither $\lim_{\hat{k} \to \infty} \Sigma_{\hat{k}}$ nor $\lim_{\hat{k} \to \infty} \Sigma_{\hat{k}}^{-1}$ exists. Then, as in the preceeding non-repeated case, we assume these limits exist in an extended sense, so that there is subsequence $\hat{k}$, and a permutation within each δ_i block such that $\lim_{\hat{k} \to \infty} \Sigma_{\hat{k}} \to \hat{\Sigma} = \text{diag}[\hat{\Sigma}_1, \hat{\Sigma}_2]$ where

$$\hat{\Sigma}_1 = \text{diag}[0, \ldots, 0, \sigma_{11}, \ldots, \sigma_{1m}, \infty, \ldots, \infty]$$

and

$$\hat{\Sigma}_2 = \text{diag}[0, \ldots, 0, \sigma_{21}, \ldots, \sigma_{2r}, \infty, \ldots, \infty]$$

with $\hat{\Sigma}^{-1}$ described analogously.

Permuting and partitioning A_0, B_0 and C_0 corresponding to $\hat{\Sigma}$, and using the subsequence for $\hat{\Sigma}$ in (B.3) and (B.4), leads to the following decompostion:

$$M_0 = \left[\begin{array}{ccc|ccc|c} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} & B_1 \\ 0 & A_{22} & A_{23} & 0 & A_{25} & A_{26} & 0 \\ 0 & 0 & A_{33} & 0 & 0 & A_{36} & 0 \\ \hline A_{41} & A_{42} & A_{43} & A_{44} & A_{45} & A_{46} & B_4 \\ 0 & A_{52} & A_{53} & 0 & A_{55} & A_{56} & 0 \\ 0 & 0 & A_{63} & 0 & 0 & A_{66} & 0 \\ \hline 0 & 0 & C_3 & 0 & 0 & C_6 & 0 \end{array}\right]$$

where M_0 represents the realization after transformation by V_0. The partitions of M_0 are shown which correspond to the partitioning of $\hat{\Sigma}$. For simplicity we have not changed the notation of the system matrices after permuting.

As A_0 is stable, the submatrices

$$\begin{bmatrix} A_{33} & A_{36} \\ A_{63} & A_{66} \end{bmatrix} \text{ and } \begin{bmatrix} A_{11} & A_{14} \\ A_{41} & A_{44} \end{bmatrix}$$

are stable by Lemma 4.7. By Lemma 4.9 there exist

$$\tilde{Y} = \begin{bmatrix} Y_1 & 0 \\ 0 & Y_4 \end{bmatrix} > 0 \text{ and } \tilde{X} = \begin{bmatrix} X_3 & 0 \\ 0 & X_6 \end{bmatrix} > 0$$

satisfying

$$\begin{bmatrix} A_{11} & A_{14} \\ A_{41} & A_{44} \end{bmatrix} \tilde{Y} \begin{bmatrix} A_{11} & A_{14} \\ A_{41} & A_{44} \end{bmatrix}^* - \tilde{Y} + \begin{bmatrix} B_1 \\ B_4 \end{bmatrix} [B_1^* \; B_4^*] < 0$$

and

$$\begin{bmatrix} A_{33} & A_{36} \\ A_{63} & A_{66} \end{bmatrix}^* \tilde{X} \begin{bmatrix} A_{33} & A_{36} \\ A_{63} & A_{66} \end{bmatrix} - \tilde{X} + \begin{bmatrix} C_3^* \\ C_6^* \end{bmatrix} [C_3 \; C_6] < 0.$$

Now, setting $Y_0 = \mathrm{diag}[Y_1, 0, 0, Y_4, 0, 0]$ and $X_0 = \mathrm{diag}[0, 0, X_3, 0, 0, X_6]$ gives

$$A_0 Y_0 A_0^* - Y_0 + B_0 B_0^* \leq 0, \quad A_0^* X_0 A_0 - X_0 + C_0^* C_0 \leq 0$$

and $X_0 Y_0 = 0$.

Finally, define $Y = V_0^* Y_0 V_0$ and $X = V_0 X_0 V_0^*$. Then $Y \geq 0$ and $X \geq 0$ are commutative and satisfy conditions $(i) - (iii)$. ■

Appendix C

Balanced Truncation Model Reduction: Proofs

In order to prove that the balanced truncation error bounds of Theorem 5.5 and Theorem 5.8 hold, we require a number of preliminary lemmas. We begin by discussing contractive matrices and associated results which are of general use for both 1D and uncertain systems. An initial version of these results were first published in [69].

C.1 Contractive Realizations

Definition C.1 *A matrix X is* **contractive** *if $\|X\| = \overline{\sigma}(X) \leq 1$, and* **strictly contractive** *if $\|X\| < 1$.*

If the matrix X is a realization matrix, the following lemma gives a well-known result ([13], [22]) on the relationship between the $\mathcal{H}_\infty$ norm of a transfer matrix and realizations of the transfer matrix.

Lemma C.2 *Suppose G represents a stable discrete time transfer matrix, then $\|G\|_\infty \leq 1 (< 1)$ if and only if there is a realization for G, denoted by $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, such that M is contractive (strictly contractive).*

A generalized version of this lemma for uncertain systems is given in Chapter 5. Lemma C.2 and the following lemma, which relates the contractiveness of a matrix to that of related submatrices, provide the main steps in the proof of the error bounds for balanced truncation model reduction.

Lemma C.3 *Suppose $U = \begin{bmatrix} U_{11} & U_{12} \\ Z & U_{22} \end{bmatrix}$ and $V = \begin{bmatrix} V_{11} & Z \\ V_{21} & V_{22} \end{bmatrix}$ are contractive (strictly contractive). Then*

$$M := \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}U_{11} & U_{12} \\ \frac{1}{\sqrt{2}}V_{11} & Z & \frac{1}{\sqrt{2}}U_{22} \\ V_{21} & \frac{1}{\sqrt{2}}V_{22} & 0 \end{bmatrix}$$

is also contractive (strictly contractive).

Proof. The result is easily proved by dilating M to the following matrix,

$$M_d := \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}U_{11} & U_{12} & \vdots & \frac{1}{\sqrt{2}}U_{11} \\ \frac{1}{\sqrt{2}}V_{11} & Z & \frac{1}{\sqrt{2}}V_{22} & \vdots & 0 \\ V_{21} & \frac{1}{\sqrt{2}}V_{22} & 0 & \vdots & -\frac{1}{\sqrt{2}}V_{22} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{2}}V_{11} & 0 & -\frac{1}{\sqrt{2}}U_{22} & \vdots & -Z \end{bmatrix},$$

and noting that $M_d^* M_d \leq I$. ■

Given a balanced system realization with $\Sigma_2 = I$, we can prove the following lemma.

Lemma C.4 *Given a balanced realization $(\boldsymbol{\Delta}, M)$ with $Y = X = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} > 0$, satisfying the Lyapunov inequalities, then*

$$\begin{bmatrix} \Sigma_1^{-\frac{1}{2}} A_{12} & \Sigma_1^{-\frac{1}{2}} A_{11} \Sigma_1^{\frac{1}{2}} & \Sigma_1^{-\frac{1}{2}} B_1 \\ A_{22} & A_{21}\Sigma_1^{\frac{1}{2}} & B_2 \end{bmatrix} \text{ and } \begin{bmatrix} A_{21}\Sigma_1^{-\frac{1}{2}} & A_{22} \\ \Sigma_1^{\frac{1}{2}} A_{11} \Sigma_1^{-\frac{1}{2}} & \Sigma_1^{\frac{1}{2}} A_{12} \\ C_1 \Sigma_1^{-\frac{1}{2}} & C_2 \end{bmatrix}$$

are contractive.

Proof. Rewriting equations (5.6) and (5.7) gives

$$[A\ B] \begin{bmatrix} Y & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^* \\ B^* \end{bmatrix} \leq Y \tag{C.1}$$

and

$$[A^*\ C^*] \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A \\ C \end{bmatrix} \leq X. \tag{C.2}$$

Premultiplying and postmultiplying (C.1) by $Y^{-\frac{1}{2}}$ and (C.2) by $X^{-\frac{1}{2}}$ shows that the matrices

$$\begin{bmatrix} Y^{-\frac{1}{2}} A Y^{\frac{1}{2}} & Y^{-\frac{1}{2}} B \end{bmatrix} \text{ and } \begin{bmatrix} X^{\frac{1}{2}} A X^{-\frac{1}{2}} \\ C X^{-\frac{1}{2}} \end{bmatrix} \tag{C.3}$$

are contractive. Substituting $\begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix}$ for Y and X in (C.3), and permuting the resulting matrices gives the desired result. ■

C.2 Balanced Trunction Model Reduction Error Bounds

Proof of Theorem 5.5:

The proof of Theorem 5.5 relies heavily on the preceeding lemmas. We assume that $\Sigma_2 = I$. In this case, we must show that

$$\|G - G_r\|_\infty \leq 2.$$

The final result follows from scaling and applying this result recursively.

By Lemma C.2, it suffices to show that there exists a realization for $\frac{1}{2}(G - G_r)$ which is contractive. One realization for $\frac{1}{2}(G - G_r)$ is given by

$$M = \begin{bmatrix} A_{11} & 0 & 0 & \frac{1}{\sqrt{2}}B_1 \\ 0 & A_{11} & A_{12} & \frac{1}{\sqrt{2}}B_1 \\ 0 & A_{21} & A_{22} & \frac{1}{\sqrt{2}}B_2 \\ -\frac{1}{\sqrt{2}}C_1 & \frac{1}{\sqrt{2}}C_1 & \frac{1}{\sqrt{2}}C_2 & 0 \end{bmatrix}.$$

Motivated by the results of Lemmas C.3 and C.4, we consider the following similarity transformation,

$$T = \begin{bmatrix} -\frac{1}{\sqrt{2}}\Sigma_1^{\frac{1}{2}} & \frac{1}{\sqrt{2}}\Sigma_1^{\frac{1}{2}} & 0 & 0 \\ 0 & 0 & I & 0 \\ \frac{1}{\sqrt{2}}\Sigma_1^{-\frac{1}{2}} & \frac{1}{\sqrt{2}}\Sigma_1^{-\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix},$$

giving

$$TMT^{-1} = \begin{bmatrix} \Sigma_1^{\frac{1}{2}}A_{11}\Sigma_1^{-\frac{1}{2}} & \frac{1}{\sqrt{2}}\Sigma_1^{\frac{1}{2}}A_{12} & 0 & 0 \\ \frac{1}{\sqrt{2}}A_{21}\Sigma_1^{-\frac{1}{2}} & A_{22} & \frac{1}{\sqrt{2}}A_{21}\Sigma_1^{\frac{1}{2}} & \frac{1}{\sqrt{2}}B_2 \\ 0 & \frac{1}{\sqrt{2}}\Sigma_1^{-\frac{1}{2}}A_{12} & \Sigma_1^{-\frac{1}{2}}A_{11}\Sigma_1^{\frac{1}{2}} & \Sigma_1^{-\frac{1}{2}}B_1 \\ C_1\Sigma_1^{-\frac{1}{2}} & \frac{1}{\sqrt{2}}C_2 & 0 & 0 \end{bmatrix}.$$

To prove the main result, we will show that TMT^{-1} is contractive, and hence M is contractive. Note that

$$TMT^{-1} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \widehat{M}$$

where

$$\widehat{M} = \begin{bmatrix} 0 & \vdots & \frac{1}{\sqrt{2}}\Sigma_1^{-\frac{1}{2}}A_{12} & \vdots & \Sigma_1^{-\frac{1}{2}}A_{11}\Sigma_1^{\frac{1}{2}} & \Sigma_1^{-\frac{1}{2}}B_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{2}}A_{21}\Sigma_1^{-\frac{1}{2}} & \vdots & A_{22} & \vdots & \frac{1}{\sqrt{2}}A_{21}\Sigma_1^{\frac{1}{2}} & \frac{1}{\sqrt{2}}B_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \Sigma_1^{\frac{1}{2}}A_{11}\Sigma_1^{-\frac{1}{2}} & \vdots & \frac{1}{\sqrt{2}}\Sigma_1^{\frac{1}{2}}A_{12} & \vdots & 0 & 0 \\ C_1\Sigma_1^{-\frac{1}{2}} & \vdots & \frac{1}{\sqrt{2}}C_2 & \vdots & 0 & 0 \end{bmatrix}.$$

Let

$$U_{11} = \Sigma_1^{-\frac{1}{2}}A_{12}, \quad U_{12} = [\Sigma_1^{-\frac{1}{2}}A_{11}\Sigma_1^{\frac{1}{2}} \ \ \Sigma_1^{-\frac{1}{2}}B_1], \quad U_{22} = [A_{21}\Sigma_1^{\frac{1}{2}} \ \ B_2],$$

$$V_{11} = A_{21}\Sigma_1^{-\frac{1}{2}}, \ V_{21} = \begin{bmatrix} \Sigma_1^{\frac{1}{2}}A_{11}\Sigma_1^{-\frac{1}{2}} \\ C_1\Sigma_1^{-\frac{1}{2}} \end{bmatrix}, \ V_{22} = \begin{bmatrix} \Sigma_1^{\frac{1}{2}}A_{12} \\ C_2 \end{bmatrix} \text{ and } Z = A_{22}.$$

Note that $U = \begin{bmatrix} U_{11} & U_{12} \\ Z & U_{22} \end{bmatrix}$ and $V = \begin{bmatrix} V_{11} & Z \\ V_{21} & V_{22} \end{bmatrix}$ are contractive by Lemma C.4. Applying Lemma C.3 shows that $\widehat{M}$ is contractive, and therefore M is contractive. Thus $\frac{1}{2}\|G - G_r\|_\infty \leq 1$ by Lemma C.2. ∎

Proof of Lemma 5.7

Suppose $(\boldsymbol{\Delta}, M)$ is stable and balanced, and

$$A = \begin{bmatrix} \hat{A} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \ B = \begin{bmatrix} \hat{B} \\ B_2 \end{bmatrix}, \text{ and } C = \begin{bmatrix} \hat{C} & C_2 \end{bmatrix}.$$

Then there exists

$$\Sigma = \begin{bmatrix} \hat{\Sigma} & \\ & \Sigma_2 \end{bmatrix} > 0, \text{ diagonal},$$

such that

$$A\Sigma A^* - \Sigma + BB^* < 0 \text{ and } A^*\Sigma A - \Sigma + C^*C < 0,$$

which implies

$$\hat{A}\hat{\Sigma}\hat{A}^* - \hat{\Sigma} + A_{12}\Sigma_2 A_{12}^* + \hat{B}\hat{B}^* < 0 \text{ and } \hat{A}^*\hat{\Sigma}\hat{A} - \hat{\Sigma} + A_{21}^*\Sigma_2 A_{21} + \hat{C}^*\hat{C} < 0.$$

Thus, $(\hat{\boldsymbol{\Delta}}, \widehat{M})$ is also stable and balanced. ∎

Proof of Theorem 5.8:

The proof for Theorem 5.8 follows from a strict generalization of the proof given for the 1D case.

By repeated permutations, scalings and truncations, we can apply the methods of Theorem 5.5 along with the sufficiency direction of Lemma 5.3 to obtain the stated bound. That is, we assume the system realization is reduced from

$$M = \begin{bmatrix} \widehat{A} & A_{12} & \widehat{B} \\ A_{21} & A_{22} & B_2 \\ \widehat{C} & C_2 & D \end{bmatrix}$$

to $\widehat{M} = \begin{bmatrix} \widehat{A} & \widehat{B} \\ \widehat{C} & D \end{bmatrix}$, with $\Sigma = \mathrm{diag}[\widehat{\Sigma}, I]$. The corresponding uncertainty structure is reduced from $\Delta = \mathrm{diag}[\delta_1 I_{n_1}, \cdots, \delta_p I_{n_p}]$ to $\widehat{\Delta} = \mathrm{diag}[\delta_1 I_{n_1}, \cdots, \delta_{p-1} I_{n_{p-1}}, \delta_p I_{\hat{n}_p}]$, where $\hat{n}_p = \sum_{j=1}^{(t_p-1)} s_{pj} < n_p$, that is, only the representation submatrices corresponding to the last uncertainty variable δ_p in Δ and the last singular value σ_{pt_p} in Σ_p are reduced. As in the 1D case, we assume $\sigma_{pt_p} = 1$ and subsequently show that $\frac{1}{2}\left\|(\boldsymbol{\Delta} \star M) - (\widehat{\boldsymbol{\Delta}} \star \widehat{M})\right\|_{SI2} \leq 1$.

Let

$$\widetilde{M} = \begin{bmatrix} \widehat{A} & 0 & 0 & \frac{1}{\sqrt{2}}\widehat{B} \\ 0 & \widehat{A} & A_{12} & \frac{1}{\sqrt{2}}\widehat{B} \\ 0 & A_{21} & A_{22} & \frac{1}{\sqrt{2}}B_2 \\ -\frac{1}{\sqrt{2}}\widehat{C} & \frac{1}{\sqrt{2}}\widehat{C} & \frac{1}{\sqrt{2}}C_2 & 0 \end{bmatrix}$$

and $\widetilde{\boldsymbol{\Delta}} = \mathrm{diag}[\widehat{\boldsymbol{\Delta}}, \boldsymbol{\Delta}]$, with corresponding commutative matrix set $\widetilde{\mathcal{T}}$. Using a similarity transformation T like that in the proof of Theorem 5.5 gives

$$T\widetilde{M}T^{-1} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix} M_I$$

where

$$M_I = \begin{bmatrix} 0 & \vdots & \frac{1}{\sqrt{2}}\widehat{\Sigma}^{-\frac{1}{2}}A_{12} & \vdots & \widehat{\Sigma}^{-\frac{1}{2}}\widehat{A}\widehat{\Sigma}^{\frac{1}{2}} & \widehat{\Sigma}^{-\frac{1}{2}}\widehat{B} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{2}}A_{21}\widehat{\Sigma}^{-\frac{1}{2}} & \vdots & A_{22} & \vdots & \frac{1}{\sqrt{2}}A_{21}\widehat{\Sigma}^{\frac{1}{2}} & \frac{1}{\sqrt{2}}B_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \widehat{\Sigma}^{\frac{1}{2}}\widehat{A}\widehat{\Sigma}^{-\frac{1}{2}} & \vdots & \frac{1}{\sqrt{2}}\widehat{\Sigma}^{\frac{1}{2}}A_{12} & \vdots & 0 & 0 \\ \widehat{C}\widehat{\Sigma}^{-\frac{1}{2}} & \vdots & \frac{1}{\sqrt{2}}C_2 & \vdots & 0 & 0 \end{bmatrix}.$$

Now defining U and V as in Theorem 5.5 and using Lemma C.4 and the fact that the Lyapunov inequalities are satisfied by assumption, gives us that U and V are contractive. Applying Lemma C.3 then implies that M_I is contractive, and hence $\widetilde{M}$ is contractive. Finally, from Lemma 5.3 we have that $\frac{1}{2}\left\|\widetilde{\boldsymbol{\Delta}} \star \widetilde{M}\right\|_{SI2} \leq 1$. ∎

Remark C.5 *Note that although the similarity transformation T used in the preceeding proof is not strictly an element of* $\widetilde{\mathcal{T}}$*, it can be written as the composition of an element of* $\widetilde{\mathcal{T}}$ *and a permutation, and thus is allowable.*

Appendix D

An Extended Proof for Recognizable Series Realizations

We provide a proof for the result given in Theorem 6.5, part(c), that two minimal linear representations for a series are similar, where we assume $\mathcal{R}$ is a field, and we consider series with coefficients in $\mathcal{R}^{q\times m}$. Recall that $\mathcal{R}^{q\times m}\langle\langle X\rangle\rangle$ has the structure of a right $\mathcal{R}\langle X\rangle$-module as described by the identities in equation (6.1). The presentation of these results combines both the notation and the logic of the proofs used in [38] and [8] for the (respective) "SIMO" and "SISO" cases.

We consider linear representations $\{h, f, g\}$ such that $s(w) = hf(w)g$ for all $w \in X^{\dagger}$, thus $f \cdot \rho$ for $\rho \in \mathcal{R}\langle X\rangle$ is defined analogously to $S \cdot \rho$. Furthermore, for the proof we assume, without loss in generality, that the q "outputs" of the series map are distinct (that is, the map defined by the series is injective, or h is full rank).

Theorem D.1 *Two minimal linear representations of the same series are similar.*

Proof. Suppose $\{h, f, g\}$ is a minimal linear representation of dimension N for a series S, and define

$$J = \{\rho \in \mathcal{R}\langle X\rangle | hf \cdot \rho = 0\}.$$

Then $\mathcal{R}\langle X\rangle/J$ is isomorphic to $h(\text{Im } f)$ by Theorem 2.8, thus

$$\dim(\mathcal{R}\langle X\rangle/J) = \dim(h(\text{Im } f)) \leq N$$

as an $\mathcal{R}$-vector space.

Clearly $J \subset \text{Ker } S$, thus

$$\dim(\mathcal{R}\langle X\rangle/J) \geq \dim(\mathcal{R}\langle X\rangle/\text{Ker } S) = \dim(\text{Im } S) = \text{rank}(\mathcal{H}_S) = N.$$

Thus $\dim(\mathcal{R}\langle X\rangle/J) = N$, $J = \text{Ker } S$, and the $\mathcal{R}\langle X\rangle$-modules Im S and $h(\text{Im } f)$ are isomorphic. Furthermore, using the distinctness assumption, we have that $h(\text{Im })f = \mathcal{R}^{q\times N}$.

As a result, there exists an $\mathcal{R}$-isomorphism $\zeta : \mathcal{R}^{q\times N} \to \operatorname{Im} S$ such that for any $\rho \in \mathcal{R}\langle X\rangle$ and $\nu \in \mathcal{R}^{q\times N}$,

$$\zeta(\nu f \cdot \rho) = \zeta(\nu) \cdot \rho \text{ and } \zeta(h) = S.$$

Now let φ be defined on $\operatorname{Im} S$ by $\varphi(\overline{S}) = \overline{s}(1)$. Then for $\nu = hf \cdot \rho$, we have $\varphi(\zeta(\nu)) = \varphi(\zeta(h) \cdot \rho) = \varphi(S \cdot \rho) = (S \cdot \rho)(1) = hf \cdot \rho g = \nu g$. Thus

$$\varphi \circ \zeta = g,$$

if g is defined by $\nu \mapsto \nu g$.

Now, suppose $\{\hat{h}, \hat{f}, \hat{g}\}$ is another minimal linear representation (of dimension N) for S. Then there exists an analogous isomorphism $\hat{\zeta}$. Therefore, there exists an isomorphism $\psi = \zeta^{-1} \circ \hat{\zeta} : \mathcal{R}^{q\times N} \to \mathcal{R}^{q\times N}$, satisfying

$$\psi(\nu \hat{f} \cdot \rho) = \psi(\nu) f \cdot \rho, \quad \psi(\hat{h}) = h, \text{ and } g \circ \psi = \hat{g}.$$

The final result is obtained by writing these relations in matrix form. ■

Bibliography

[1] Apkarian, P., and Gahinet, P., 1995, "A Convex Characterization of Gain-Scheduled H-infinity Controllers", *IEEE Trans. on Auto. Control*, Vol. 40, pp. 853-864.

[2] Beck, C., Doyle, J., and Glover, K., "Model Reduction of Multi-Dimensional and Uncertain Systems", to appear *IEEE Trans. on Auto. Control*, October 1996; also California Institute of Technology Technical Report CDS94-017.

[3] Beck, C., Doyle, J., "A Necessary and Sufficient Minimality Condition for Uncertain Systems", submitted to *IEEE Trans. on Auto. Control*; also California Institute of Technology Technical Report CDS94-019.

[4] Beck, C., and Doyle, J., 1993 "Model Reduction of Behavioural Systems", *Proceedings 32nd CDC*, San Antonio, Texas.

[5] Bendotti, P., and Beck, C., 1995, "Modelling and Control of a Pressurized Water Reactor Using μ-Synthesis Techniques", *Proceedings 1996 IFAC World Congress*, San Francisco, California.

[6] Bendotti, P., and Beck, C., 1996, "On the Role of LFT Model Reduction Methods in Robust Controller Synthesis for a Pressurized Water Reactor", report in preparation.

[7] Bendotti, P., and Bodenheimer, B., 1994, "Identification and $\mathcal{H}_\infty$ Control Design for a Pressurized Water Reactor," *Proceedings 33rd CDC*, Orlando, Florida.

[8] Berstel, J., and Reutenauer, C., 1988, "Rational Series and Their Languages", Monographs on Theoretical Computer Science, Springer-Verlag.

[9] Bollobas, B., 1990, *Linear Analysis*, Cambridge University Press, Cambridge, UK.

[10] Bose, N.K., 1982, *Applied Multidimensional Systems Theory*, Van Nostrand Reinhold Company, New York, NK.

[11] Boyd, S., and Barret, C., 1991, *Linear Controller Design*, Prentice Hall, New Jersey.

[12] Boyd, S,. El Ghaoui, L., Feron, E., and Balakrishnan, V., 1994 *Linear Matrix Inequalities in Systems and Control Theory*, SIAM.

[13] Boyd, S. and Yang, Q., 1988, *Simultaneous Quadratic Lyapunov Functions*, Stanford Research Report.

[14] Braatz, R., Young, P., Doyle, J., and Morari, M., 1994, "Computational Complexity of μ Calculation", *IEEE Trans. on Auto. Control* Vol. 39, pp. 1000-1002.

[15] Carlyle, J., and Paz, A., 1971, "Realizations by Stochastic Finite Automaton", *J. Comput. System Sci.*, 5, pp. 26-40.

[16] Chen, X., and Wen, J.T., 1994, "Model Reduction of Multidimensional Positive Real System", *Proceedings 33rd CDC*, Orlando, Florida.

[17] David, J., and DeMoor, B., 1993, "The Opposite of Analytic Centering for Solving Minimum Rank Problems in Control and Identification", *Proceedings 32nd CDC*, San Antonio, Texas.

[18] David, J., and DeMoor, B., 1994, "Designing Reduced Order Output Feedback Controllers using a Potential Reduction Method", *Proceedings 1994 ACC*, Baltimore, Maryland.

[19] Desai, U.B. and Pal, D., 1984, "A Transformation Approach to Stochastic Model Reduction", *IEEE Trans. on Auto. Control*, Vol. 29, pp. 1097-1100.

[20] Doyle, J. C., 1982, "Analysis of Feedback Systems with Structured Uncertainties", *IEE Proceedings, Part D*, Vol. 129, no. 6, pp. 242-250.

[21] Doyle, J. C., Wall, J. E.,, and Stein, G., 1982, "Performance and Robustness Analysis for Structured Uncertainty", *Proceedings 21st CDC.*

[22] Doyle, J., Packard, A., and Zhou, K., 1991, "LFTs, LMIs, and μ", *Proceedings 30th CDC*, Brighton, UK.

[23] Dullerud, G. E., 1995, *Control of Uncertain Sampled-Data Systems*, Birkhauser, NY.

[24] Enns, D.F., 1984, *Model Reduction for Control System Design*, Ph.D. Dissertation, Stanford University.

[25] El Ghaoui, L., Delebecque, F., and Nikoukah, R., 1995, *LMITOOL: A User-Friendly Interface for LMI Optimization.*

[26] Fan, M.K.H., Tits, A.L., and Doyle, J.C., 1991, "Robustness in the Presence of Mixed Parametric Uncertainty and Unmodelled Dynamics", *IEEE Trans. on Auto. Control*, Vol 36, pp. 25-38.

[27] Fiacco, A.V., and McCormick, G.P., 1990, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM Classics in Applied Mathematics, Pennsylvania.

[28] Fliess, M., 1974, "Matrices de Hankel", *J. Math. pures et appl.*, Vol. 53, pp. 197-222.

[29] Fliess, M., Lamnabhi, M., and Lamnabhi, F., 1983, "An Algebraic Approach to Nonlinear Functional Expansions", *IEEE Trans. on Circuits and Systems*, Vol. 30, No. 8, pp. 554-570.

[30] Gahinet, P., Nemirovskii, A., Laub, A. J., and Chilali, M., 1994, *The LMI Control Toolbox*, Beta-Release, The MathWorks Inc.

[31] Gilbert, E., 1963, "Controllability and Observability in Multivariable Control Systems," *SIAM J. Control*, Vol. 1, pp. 128-151.

[32] Gill, E.G., Murray, W., and Wright, M.H., 1981, *Practical Optimization*, Academic Press Inc., London.

[33] Glover, K., 1984, "All Optimal Hankel-norm Approximations of Linear Multivariable Systems and their L^∞-Error Bounds", *Int. J. Control.*, Vol.39, pp. 1115-1193.

[34] Balas, G., Doyle, J., Glover, K., Packard, A., and Smith, R., 1993, μ-Analysis and Synthesis Toolbox for MATLAB, The MathWorks Inc.

[35] Halmos, P.R., 1982, *A Hilbert Space Problem Book*, Springer-Verlag, New York.

[36] Hinrichsen, D. and Pritchard, A.J., 1990, "An Improved Error Estimate for Reduced-Order Models of Discrete-Time Systems", *IEEE Trans. on Auto. Control*, Vol. 35, pp. 317-320.

[37] Hungerford, T.W., 1974, *Algebra*, Graduate Texts in Mathematics, Springer-Verlag.

[38] Isidori, A., 1989, *Nonlinear Control Systems*, Springer-Verlag.

[39] Isidori, A., 1973, "Direct Construction of Minimal Bilinear Realizations from Nonlinear Input-Output Maps", *IEEE Trans. on Auto. Control*, Vol. 18, pp. 626-631.

[40] Iwasaki, T., and Skelton, R.E., 1994, "All Controllers for the General $\mathcal{H}_\infty$ Problem: LMI Existence Conditions and State Space Formulas", *Sys. and Cont. Let.*, Vol. 30, pp. 1307-1317.

[41] Kailath, T., 1980, *Linear Systems*, Prentice-Hall Inc., New Jersey.

[42] Kalman, R. E., 1963, "Mathematical Descriptions of Linear Systems," *SIAM J. Control*, Vol. 1, pp. 152-192.

[43] Kavranoglu, D., and Bettayeb, M, 1993, "Characterization of the Solution to the Optimal H_∞ Model Reduction Problem", *Sys. and Cont. Lett*, Vol. 20, pp. 99-107.

[44] Laub, A., Heath, M., Paige, C., and Ward, R., 1987, "Computation of System Balancing Transformations and Other Applications of Simultaneous Diagonalization Algorithms", *IEEE Trans. on Auto. Control*, Vol. 32, pp. 115-122.

[45] Ljung, L., 1987, *System Identification: Theory for the User*, Prentice Hall.

[46] Luenberger, D.G., 1989, *Linear and Nonlinear Programming*, Addison-Wesley, Massachusetts.

[47] Lu, W., Zhou, K., and Doyle, J.C., 1993, "Stabilization of Linear Systems with Structured Perturbations", California Institute of Technology Technical Report CDS93-014.

[48] Marsden, J., 1974, *Elementary Classical Analysis*, W. H. Freeman and Co., New York.

[49] Megretski, A., 1993, "Necessary and Sufficient Conditions of Stability: A Multiloop Generalization of the Circle Criterion", *IEEE Trans. on Auto. Control*, Vol. 38, pp. 753-756.

[50] Megretski, A., and Treil, S., 1990-91, "S-Procedure and Power Distribution Inequalities: A New Method in Optimization and Robustness of Uncertain Systems", Mittag-Leffler Institute, Rep. No. 1.

[51] B.C. Moore, 1981, "Principal Component Analysis of Linear Systems: Controllability, Observability and Model Reduction", *IEEE Trans. on Auto. Control*, Vol. 26, pp. 17-32.

[52] Nemirovskii, A., and Gahinet, P., 1993, "The Projective Method for Solving Linear Matrix Inequalities", *Proceedings 1994 ACC*, Baltimore, Maryland.

[53] Nesterov, Y., and Nemirovskii, A., 1994, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Studies in Applied Mathematics, Philadelphia.

[54] Packard, A., 1994, "Gain Scheduling via Linear Fraction Transformations", *Sys. and Cont. Let.*, Vol. 22, pp. 79-92.

[55] Packard, A., and Doyle, J. C., 1993, "The Complex Structured Singular Value", *Automatica*, Vol. 29, pp 71-110.

[56] Packard, A., Zhou, K., Pandey, P., Leonhardson, J., and Balas, G., 1992, " Optimal, Constant I/O Similarity Scaling for Full-Information and State-Feedback Control Problems", *Sys. and Contr. Let.*, Vol 19, pp. 271-280.

[57] Paganini, F., 1995, *Sets and Constraints in the Analysis of Uncertain Systems*, Ph. D. Dissertation, California Institute of Technology.

[58] D'Andrea, R., 1995, "Kalman Decomposition of Multi-D Systems and Minimality", preprint.

[59] Redheffer, R., 1959, "Inequalities for a Matrix Ricatti Equation", *Journal of Mathematics and Mechanics*, Vol. 8.

[60] Rosenbrock, H.H., 1970, *Multivariable and State-Space Theory*, Wiley, New York.

[61] Safonov, M.G., and Chiang, R.Y., 1988, "A Schur Method for Balanced Model Reduction", *Proceedings 1988 ACC*, Atlanta, Georgia.

[62] Safonov, M. G,. 1982, "Stability Margins of Diagonally Perturbed Multivariable Feedback Systems", *IEE Proceedings, Part D*, Vol. 129, no. 6, pp 251-256.

[63] Shamma, J., 1994, "Robust Stability with Time-Varying Structured Uncertainty", *IEEE Trans. on Auto. Control*, Vol. 39, pp. 714-724.

[64] Sontag, E., 1979 "Realization Theory of Discrete-Time Nonlinear Systems: Part I - the Bounded Case", *IEEE Trans. on Circuits and Systems*, Vol. 26, pp. 342-356.

[65] Sontag, E., Conversations, various times and places, 1994.

[66] Schutzenberger, M., 1961, "On the Definition of a Family of Automata", *Inform. Control*, Vol. 4.

[67] Toker, O., 1995, *Complexity Issues in System Theory and Solution Procedures for Certain Robust Control Problems*, Ph. D. Dissertation, The Ohio State University.

[68] Vandenberghe, L., and Boyd, S., 1995, "A Primal-Dual Potential Reduction Method for Problems Involving Matrix Inequalities", *Mathematical Programming*, Vol. 69, pp. 205-236.

[69] Wang, W., Doyle, J., Beck, C., and Glover, K., 1991, "Model Reduction of LFT Systems", *Proceedings, 30th CDC*, Brighton, UK.

[70] Wonham, W. M., 1985, *Linear Multivariable Control; A Geometric Approach*, Springer-Verlag, New York, NY, 3rd Edition.

[71] Young, P.M., 1993, *Robustness with Parametric and Dynamic Uncertainty*, Ph. D. Dissertation, California Institute of Technology.

[72] Zadeh, L.A., and Desoer, C.A., 1963, *Linear System Theory - A State-Space Approach*, McGraw-Hill, New York.

[73] Zames, G., 1966, "On the Input-Output Stability of Nonlinear Time-Varying Feedback Systems, Parts I and II", *IEEE Trans. on Auto. Control*, Vol. 11, pp. 228-238 and pp. 465-476.

[74] Zhou, K., with Doyle, J.C., and Glover, K., 1995, *Robust and Optimal Control*, Prentice-Hall, New Jersey.