TRACTABLE PROBLEMS IN OPTIMAL DECENTRALIZED CONTROL

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
AERONAUTICS AND ASTRONAUTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Michael Charles Rotkowitz
June 2005

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Sanjay Lall   Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Stephen Boyd
(Electrical Engineering)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Claire Tomlin

Approved for the University Committee on Graduate Studies.

Abstract

This thesis considers the problem of constructing optimal decentralized controllers. The problem is formulated as one of minimizing the closed-loop norm of a feedback system subject to constraints on the controller structure.

The notion of quadratic invariance of a constraint set with respect to a system is defined. It is shown that quadratic invariance is necessary and sufficient for the constraint set to be preserved under feedback. It is further shown that if the constraint set has this property, this allows the constrained minimum-norm problem to be solved via convex programming. These results are developed in a very general framework, and are shown to hold for continuous-time systems, discrete-time systems, or operators on Banach spaces, for stable or unstable plants, and for the minimization of any norm.

The utility of these results is then demonstrated on some specific constraint classes. An explicit test is derived for sparsity constraints on a controller to be quadratically invariant, and thus amenable to convex synthesis. Symmetric synthesis is also shown to be quadratically invariant.

The problem of control over networks with delays is then addressed as another constraint class. Multiple subsystems are considered, each with its own controller, such that the dynamics of each subsystem may affect those of other subsystems with some propagation delays, and the controllers may communicate with each other with some transmission delays. It is shown that if the communication delays are less than the propagation delays, then the associated constraints are quadratically invariant, and thus optimal controllers can be synthesized. We further show that this result still holds in the presence of computational delays.

This thesis unifies the few previous results on specific tractable decentralized control problems, identifies broad and useful classes of new solvable problems, and delineates the largest known class of convex problems in decentralized control.

Acknowledgments

It is an honor to be Sanjay Lall's first academic descendant.

Both the breadth and depth of his knowledge are truly remarkable, and I consider myself lucky to have been able to collaborate with him so closely for so many years. He has somehow maintained his high level of accessibility while our group burgeoned from just me to the thriving lab we have today. For the many times he has gone above and beyond the call of duty for me, I am forever grateful.

I would like to thank Stephen Boyd and Claire Tomlin for serving on my reading committee. Thanks to Stephen for training me to see problems in terms of their convexity or lack thereof, which put me in the right frame of mind to discover the results presented here when Sanjay introduced me to decentralized control. Thanks are also due to Matt West for serving on my defense committee on short notice, and to Ben Van Roy for serving as the chair of my defense committee when it looked like a chair might be impossible to find.

I would like to acknowledge a few professors from other institutions who have also been helpful. Thanks to Pablo Parillo (MIT) for asking some great questions when shown the preliminary version of these results, which ultimately helped lead to the broad generalizations presented here. Thanks to Geir Dullerud (Illinois) for many useful conversations, and for pointing out that symmetric synthesis would play a nice role in elucidating the necessity of quadratic invariance. Thanks to Anders Rantzer (Lund) for his continued interest in my work. The lattice examples of Section 5.3.4 were inspired by his questions on the topic. Thanks as well to Raff D'Andrea (Cornell) and Petros Voulgaris (Illinois) for many interesting conversations and for writing recommendations for me along with Geir.

Thank you to Been-Der Chen, Randy Cogill, Ritesh Madan, Matthew Peet, Prashant Srinivasan, and Ta-Chung Wang for making the Networked Systems and Control Lab a great place to work. Thanks especially to Randy Cogill for his contribution to Section 5.3 and to Been-Der Chen for his contribution to Lemma 5.

This work was supported financially by the Air Force Office of Scientific Research and the Stanford Graduate Fellowship program. Thanks to those at MathWorks who decided to establish a fellowship in honor of Cleve Moler, which I was fortunate enough to be awarded, and to Pat Cook for her knowledgeable and efficient administration of the SGF program. Several in the Aero/Astro department well-deserving of my gratitude include Sally Gressens, Lynn Kaiser, Ralph Levine, Dianne Le, Sherann Ellsworth, and Dana Parga. Thanks also to the statistics department for their flexibility and enabling me to get a masters there while completing this doctorate.

Thanks to Mark Peters for a decade of great friendship, including a much-welcomed return to Stanford. Thank you to Magdalena Jonikas for being a wonderful part of my life during the past year. And of course, thank you to my family for their constant support.

Finally, I would like to dedicate this thesis to Henry Grossbard.

Contents

List of Figures

Chapter 1

Introduction

Much of conventional controls analysis assumes that the controllers to be designed all have access to the same measurements. With the advent of complex systems, decentralized control has become increasingly important, where one has multiple controllers each with access to different information. Examples of such systems include flocks of aerial vehicles, autonomous automobiles on the freeway, the power distribution grid, spacecraft moving in formation, and paper machining.

In a standard controls framework, the decentralization of the system manifests itself as sparsity or delay constraints on the controller to be designed. Therefore a canonical problem one would like to solve in decentralized control is to minimize a norm of the closed-loop map subject to a subspace constraint as follows

$$\begin{aligned} \text{minimize} \quad & \|f(P,K)\| \\ \text{subject to} \quad & K \text{ stabilizes } P \\ & K \in S \end{aligned}$$

For a general linear time-invariant plant P and subspace S there is no known tractable algorithm for computing the optimal K. It has been known since 1968 [30] that even the simplest versions of this problem can be extremely difficult. In fact, certain cases have been shown to be intractable [15, 3]. However, there are also several special cases of this problem for which efficient algorithms have been found [2, 8, 10, 16, 26, 27]. This thesis unifies these cases and identifies a simple condition, which we call ***quadratic invariance***, under which the above problem may be recast as a convex optimization problem. The notion of quadratic invariance allows us to better understand this dichotomy between tractable and intractable optimal decentralized control problems. It further delineates the largest known class of decentralized problems for which optimal controllers may be efficiently synthesized.

Quadratic invariance is a simple algebraic condition relating the plant and the constraint set. The main results of this thesis hold for continuous-time systems, discrete-time systems, or operators on Banach spaces, for stable or unstable plants, and for the minimization of any norm.

In Chapter 2, we define quadratic invariance, and present some of its characteristics. In Chapter 3, we show that quadratic invariance is necessary and sufficient for the constraint set to be invariant under a linear fractional transformation (LFT), namely, the map from K to $K(I - GK)^{-1}$. This is first shown for operators over Banach spaces in Section 3.1, and then for causal operators on extended spaces in Section 3.2. This allows for convex synthesis of optimal controllers when the plant is stable.

In Chapter 4, we show that for possibly unstable plants, as long as a nominal controller exists which is both stable and stabilizing, this invariance implies that the information constraint is equivalent to affine constraints on the Youla parameter. Thus synthesizing optimal stabilizing controllers subject to quadratically invariant constraints is a convex optimization problem. We further show that this is still a convex optimization problem, even when such a nominal controller is unobtainable.

In Chapter 5 we apply these results to specific constraint classes. We first consider sparsity constraints in Section 5.1. We develop an explicit test for the quadratic invariance of sparsity constraints, and thus show that optimal synthesis subject to such constraints which pass the test may be cast as a convex optimization problem. As a consequence of the test, we show some examples of sparsity structures which are quadratically invariant, and also show that block diagonal constraints are never quadratically invariant unless the plant is block diagonal as well.

We show in Section 5.2 that optimal synthesis of a symmetric controller for a symmetric plant is also quadratically invariant and thus amenable to convex synthesis. This is important because this problem, while formerly known to be solvable, defied other efforts to classify tractable problems.

We then in Section 5.3 consider the problem of control over networks with delays. We find that optimizing the closed-loop norm may be formulated as a convex optimization problem when the controllers can communicate faster than the dynamics propagate. We further show that this result still holds in the presence of computational delay.

These results all hold for the minimization of an arbitrary norm. In Chapter 6 we show that if the norm of interest is the $\mathcal{H}_2$-norm, then the constrained convex optimization problem derived in Section 4 may be further reduced to an unconstrained convex optimization problem, and readily solved. We further show how similar techniques may be used to systematically find stabilizing decentralized controllers for quadratically invariant constraints. We then provide some numerical examples.

1.1 Prior Work

Decentralized control has been studied from many perspectives over the past half century, and there have been many striking results which illustrate the complexity of this problem. Important early work includes that of Radner [17], who developed sufficient conditions under which minimal quadratic cost for a linear system is achieved by a linear controller. An important example was presented in 1968 by Witsenhausen [30] where it was shown that for quadratic stochastic optimal control of a linear system, subject to a decentralized information constraint called *non-classical information*, a nonlinear controller can achieve greater performance than any linear controller. An additional consequence of the work of [13, 30] is to show that under such a non-classical information pattern the cost function is no longer convex in the controller variables, a fact which today has increasing importance.

With the difficulty of the general problem elucidated and the myth of ubiquitous linear optimality refuted, efforts followed to classify when linear controllers were indeed optimal, to discern when finding the optimal linear controller could be cast as a convex optimization problem, and to understand the complexity of decentralized control problems. In a later paper [31], Witsenhausen summarized several important results on decentralized control at that time, and gave sufficient conditions under which the problem could be reformulated so that the standard Linear-Quadratic-Gaussian (LQG) theory could be applied. Under these conditions, an optimal decentralized controller for a linear system could be chosen to be linear. Ho and Chu [10], in the framework of *team theory*, defined a more general class of information structures, called *partially nested*, for which they showed the optimal LQG controller to be linear. Roughly speaking, a plant-controller system is called partially nested if whenever the information of controller A is affected by the decision of a controller B, then A has access to all of the information that B has.

The study of the computational complexity of decentralized control problems has shown certain problems to be intractable. Blondel and Tsitsiklis [3] showed that the problem of finding a stabilizing decentralized static output feedback is NP-complete. This is also the case for a discrete variant of Witsenhausen's counterexample [15].

For particular information structures, the controller optimization problem may have a tractable solution, and in particular, it was shown by Voulgaris [26] that the so-called *one-step delay information sharing pattern* problem has this property. In [8] the LEQG problem is solved for this information pattern, and in [26] the $\mathcal{H}_2$, $\mathcal{H}_\infty$ and L_1 control synthesis problems are solved. A class of structured spatio-temporal systems has also been analyzed in [2], and shown to be reducible to a convex program. Several information structures are identified in [16] for which the problem of minimizing multiple objectives is reduced to a finite-dimensional convex optimization problem.

In this thesis we define a property called *quadratic invariance*, show that it is necessary and sufficient for the constraint set to be preserved under feedback, and that this allows optimal stabilizing decentralized controllers to be synthesized via convex programming. The tractable structures of [2, 8, 10, 16, 26, 27, 31] can all be shown to satisfy this property.

1.2 Preliminaries

Given topological vector spaces $\mathcal{X}, \mathcal{Y}$, let $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ denote the set of all maps $T : \mathcal{X} \to \mathcal{Y}$ such that T is linear and continuous. Note that if $\mathcal{X}, \mathcal{Y}$ are normed spaces, as in Section 3.1, then all such T are bounded, but that T may be unbounded in general. We abbreviate $\mathcal{L}(\mathcal{X}, \mathcal{X})$ with $\mathcal{L}(\mathcal{X})$.

Suppose that we have a generalized plant $P \in \mathcal{L}(\mathcal{W} \times \mathcal{U}, \mathcal{Z} \times \mathcal{Y})$ partitioned as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

so that $P_{11} : \mathcal{W} \to \mathcal{Z}$, $P_{12} : \mathcal{U} \to \mathcal{Z}$, $P_{21} : \mathcal{W} \to \mathcal{Y}$ and $P_{22} : \mathcal{U} \to \mathcal{Y}$. Suppose $K \in \mathcal{L}(\mathcal{Y},\mathcal{U})$. If $I - P_{22}K$ is invertible, define $f(P, K) \in \mathcal{L}(\mathcal{W}, \mathcal{Z})$ by

$$f(P, K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

The map $f(P, K)$ is called the (lower) ***linear fractional transformation*** (LFT) of P and K; we will also refer to this as the ***closed-loop map***. In the remainder of the thesis, we abbreviate our notation and define $G = P_{22}$, as we will refer to this block frequently.

Dual pairings

Given a linear vector space $\mathcal{X}$, let $\mathcal{X}^*$ denote the dual-space of $\mathcal{X}$, and let $\langle x, x^* \rangle$ denote the dual pairing of any $x \in \mathcal{X}$ and $x^* \in \mathcal{X}^*$. For $S \subseteq \mathcal{X}$ and $T \subseteq \mathcal{X}^*$ define

$$S^\perp = \Big\{ x^* \in \mathcal{X}^* \mid \langle x, x^* \rangle = 0 \text{ for all } x \in S \Big\}$$
$$^\perp T = \Big\{ x \in \mathcal{X} \mid \langle x, x^* \rangle = 0 \text{ for all } x^* \in T \Big\}$$

Kronecker products

Given $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{s\times q}$ let the *Kronecker product* of A and B be denoted by $A \otimes B$ and given by

$$A \otimes B = \begin{bmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & & \vdots \\ A_{m1}B & \cdots & A_{mn}B \end{bmatrix} \in \mathbb{C}^{ms\times nq}$$

Given $A \in \mathbb{C}^{m\times n}$, we may write A in term of its columns as

$$A = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}$$

and then associate a vector $\mathrm{vec}(A) \in \mathbb{C}^{mn}$ defined by

$$\mathrm{vec}(A) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Lemma 1. *Let $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{s\times q}, X \in \mathbb{C}^{n\times s}$. Then*

$$\mathrm{vec}(AXB) = (B^T \otimes A)\,\mathrm{vec}(X)$$

Proof. See, for example, [11]. ∎

Transfer functions

We use the following standard notation. Denote the imaginary axis by

$$j\mathbb{R} = \{z \in \mathbb{C} \mid \Re(z) = 0\}$$

and the closed right half of the complex plane by

$$\mathbb{C}_+ = \{z \in \mathbb{C} \mid \Re(z) \geq 0\}$$

A rational function $G : j\mathbb{R} \to \mathbb{C}$ is called ***real-rational*** if the coefficients of its numerator and denominator polynomials are real. Similarly, a matrix-valued function $G : j\mathbb{R} \to \mathbb{C}^{m\times n}$ is called real-rational if G_{ij} is real-rational for all i, j. It is called ***proper*** if

$$\lim_{\omega\to\infty} G(j\omega) \quad \text{exists and is finite,}$$

and it is called ***strictly proper*** if

$$\lim_{\omega\to\infty} G(j\omega) = 0.$$

Denote by $\mathcal{R}_p^{m\times n}$ the set of matrix-valued real-rational proper transfer matrices

$$\mathcal{R}_p^{m\times n} = \Big\{ G : j\mathbb{R} \to \mathbb{C}^{m\times n} \mid G \text{ proper, real-rational} \Big\}$$

and let $\mathcal{R}_{sp}^{m\times n}$ be

$$\mathcal{R}_{sp}^{m\times n} = \Big\{ G \in \mathcal{R}_p^{m\times n} \mid G \text{ strictly proper} \Big\}$$

Also let $\mathcal{RH}_\infty$ be the set of real-rational proper stable transfer matrices

$$\mathcal{RH}_\infty^{m\times n} = \Big\{ G \in \mathcal{R}_p^{m\times n} \mid G \text{ has no poles in } \mathbb{C}_+ \Big\}$$

where we have used the fact that functions in $\mathcal{RH}_\infty$ are determined by their values on $j\mathbb{R}$, and $\mathcal{RH}_\infty$ can thus be regarded as a subspace of $\mathcal{R}_p$. If $A \in \mathcal{R}_p^{n\times n}$ we say A is ***invertible*** if $\lim_{\omega\to\infty} A(j\omega)$ is an invertible matrix and $A(j\omega)$ is invertible for almost all $\omega \in \mathbb{R}$. Note that this is different from the definition of invertibility for the associated multiplication operator on L_2. If A is invertible we write $B = A^{-1}$ if $B(j\omega) = A(j\omega)^{-1}$ for almost all $\omega \in \mathbb{R}$. Note that, if $G \in \mathcal{R}_{sp}^{n_y\times n_u}$ then $I - GK$ is invertible for all $K \in \mathcal{R}_p^{n_u\times n_y}$. Both this fact and the inverse itself will be consistent with our definition of invertibility for operators on extended spaces.

When we consider transfer functions for discrete-time systems instead of continuous-time systems, $j\mathbb{R}$ is replaced with $\{z \in \mathbb{C} \mid |z| = 1\}$, the unit circle, $\mathbb{C}_+$ is replaced with $\{z \in \mathbb{C} \mid |z| \geq 1\}$, and $j\omega$ is replaced with $e^{j\omega}$. The above terms are then defined analogously.

Topology

Let $\mathcal{X}$ be a vector space and $\{\|\cdot\|_\alpha \mid \alpha \in I\}$ be a family of semi-norms on $\mathcal{X}$. The family is called ***sufficient*** if for all $x \in \mathcal{X}$ such that $x \neq 0$ there exists $\alpha \in I$ such that $\|x\|_\alpha \neq 0$. The topology generated by all open $\|\cdot\|_\alpha$-balls is called the topology generated by the family of semi-norms. Convergence in this topology is equivalent to convergence in every semi-norm, and continuity of a linear operator is equivalent to continuity in every semi-norm. See, for example, [35, 24].

Extended spaces

We introduce some notation for extended linear spaces. These spaces are utilized extensively in [7, 29]. The topologies developed here for operators on these spaces were first utilized in [20] and are used particularly in Section 3.2.

We define the truncation operator P_T for all $T \in \mathbb{R}_+$ on all functions $f : \mathbb{R}_+ \to \mathbb{R}$ such that $f_T = P_T f$ is given by

$$f_T(t) = \begin{cases} f(t) & \text{if } t \leq T \\ 0 & \text{if } t > T \end{cases}$$

and hereafter abbreviate $P_T f$ as f_T. We make use of the standard L_p Banach spaces equipped with the usual p-norm

$$L_p = \left\{ f : \mathbb{R}_+ \to \mathbb{R} \;\middle|\; \int_0^\infty (f(t))^p \text{ exists and is finite} \right\} \quad \|f\|_p = \left(\int_0^\infty (f(t))^p \, dt \right)^{\frac{1}{p}}$$

and the extended spaces

$$L_{pe} = \{ f : \mathbb{R}_+ \to \mathbb{R} \mid f_T \in L_p \text{ for all } T \in \mathbb{R}_+ \} \qquad \text{for all } p \geq 1$$

We let the topology on L_{2e} be generated by the sufficient family of semi-norms $\{\|\cdot\|_T \mid T \in \mathbb{R}_+\}$ where $\|f\|_T = \|P_T f\|_{L_2}$, and let the topology on $\mathcal{L}(L_{2e}^m, L_{2e}^n)$ be generated by the sufficient family of semi-norms $\{\|\cdot\|_T \mid T \in \mathbb{R}_+\}$ where $\|A\|_T = \|P_T A\|_{L_2^m \to L_2^n}$.

We use similar notation for discrete time. As is standard, we extend the discrete-time Banach spaces ℓ_p to the extended space

$$\ell_e = \{f : \mathbb{Z}_+ \to \mathbb{R} \mid f_T \in \ell_\infty \text{ for all } T \in \mathbb{Z}_+\}$$

Note that in discrete time, all extended spaces contain the same elements, since the common requirement is that the sequence is finite at any finite index. This motivates the abbreviated notation of ℓ_e.

We let the topology on ℓ_e be generated by the sufficient family of semi-norms $\{\|\cdot\|_T \mid T \in \mathbb{Z}_+\}$ where $\|f\|_T = \|P_T f\|_{\ell_2}$, and let the topology on $\mathcal{L}(\ell_e^m, \ell_e^n)$ be generated by the sufficient family of semi-norms $\{\|\cdot\|_T \mid T \in \mathbb{Z}_+\}$ where $\|A\|_T = \|P_T A\|_{\ell_2^m \to \ell_2^n}$.

When the dimensions are implied by context, we omit the superscripts of $\mathcal{R}_p^{m\times n}$, $\mathcal{R}_{sp}^{m\times n}, \mathcal{RH}_\infty^{m\times n}, L_{pe}^{m\times n}, \ell_e^{m\times n}$. We will indicate the restriction of an operator A to $L_2[0,T]$ or $\ell_e[0,T]$ by $A|_T$, and the restriction and truncation of an operator as $A_T = P_T A|_T$. Thus for every semi-norm in this thesis, one may write $\|A\|_T = \|A_T\|$. Given a set of operators S, we also denote $S_T = \{P_T A|_T \ ; \ A \in S\}$.

1.3 Problem Formulation

Given linear spaces $\mathcal{U}, \mathcal{W}, \mathcal{Y}, \mathcal{Z}$, generalized plant $P \in \mathcal{L}(\mathcal{W} \times \mathcal{U}, \mathcal{Z} \times \mathcal{Y})$, and a subspace of admissible controllers $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$, we would like to solve the following problem:

$$\begin{aligned} \text{minimize} \quad & \|f(P,K)\| \\ & K \in S \end{aligned} \tag{1.1}$$

Here $\|\cdot\|$ is any norm on $\mathcal{L}(\mathcal{W}, \mathcal{Z})$, chosen to encapsulate the control performance objectives, and S is a subspace of admissible controllers which encapsulates the decentralized nature of the system. The norm on $\mathcal{L}(\mathcal{W}, \mathcal{Z})$ may be either a deterministic measure of performance, such as the induced norm, or a stochastic measure of performance, such as the $\mathcal{H}_2$ norm. Many decentralized control problems may be formulated in this manner, and some examples are shown below. We call the subspace

S the ***information constraint***.

In Section 3.1, we consider Banach spaces and thus have the additional constraint that $(I-GK)$ be invertible. In Chapter 4 we consider extended spaces and thus have the additional constraint that K stabilize P.

This problem is made substantially more difficult in general by the constraint that K lie in the subspace S. Without this constraint, the problem may be solved by a simple change of variables, as discussed in Section 3.1.1. For specific norms, the problem may also be solved using a state-space approach. Note that the cost function $\|f(P,K)\|$ is in general a non-convex function of K. No computationally tractable approach is known for solving this problem for arbitrary P and S.

1.3.1 Some examples

Many standard centralized and decentralized control problems may be represented in the form of problem (1.1), for specific choices of P and S. Examples include the following.

Perfectly decentralized control

We would like to design n separate controllers $\{K_1,\dots,K_n\}$, with controller K_i connected to subsystem G_i of a coupled system, as in Figure 1.1.

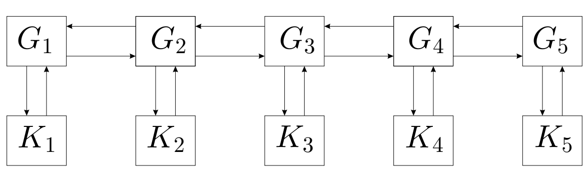

Figure 1.1: Perfectly decentralized control

When reformulated as a synthesis problem in the LFT framework above, the constraint set S is

$$S = \Big\{ K \in L(\mathcal{Y},\mathcal{U}) \;\Big|\; K = \operatorname{diag}(K_1,\dots,K_n) \Big\}$$

that is, S consists of those controllers that are *block-diagonal*. We will revisit this example in Section 5.1.3.

Delayed measurements

In this example we have n subsystems $\{G_1, \ldots, G_n\}$, each with its respective controller K_i, arranged so that subsystem i receives signals from controller i after a computational delay of c, controller i receives measurements from subsystem j with a transmission delay of $t|i-j|$, and subsystem i receives signals from subsystem $i+1$ delayed by propagation delay p. This is illustrated in Figure 1.2 for the case where $n = 3$.

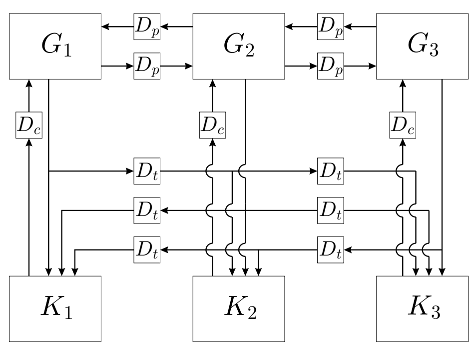

Figure 1.2: Distributed control with delays

When formulated as a synthesis problem in the LFT framework, the constraint set S may be defined as follows. Let $K \in S$ if and only if

$$K = \begin{bmatrix} D_c H_{11} & D_{t+c} H_{12} & \ldots & D_{(n-1)t+c} H_{1n} \\ D_{t+c} H_{21} & D_c H_{22} & \ldots & D_{(n-2)t+c} H_{2n} \\ \vdots & & & \vdots \\ D_{(n-1)t+c} H_{n1} & \ldots & & D_c H_{nn} \end{bmatrix}$$

for some $H_{ij} \in \mathcal{R}_p$ of appropriate spatial dimensions. The corresponding system G is given by

$$G = \begin{bmatrix} A_{11} & D_p A_{12} & \dots & D_{(n-1)p} A_{1n} \\ D_p A_{21} & A_{22} & \dots & D_{(n-2)p} A_{2n} \\ \vdots & & & \vdots \\ D_{(n-1)p} A_{n1} & \dots & & A_{nn} \end{bmatrix}$$

for some $A_{ij} \in \mathcal{R}_{sp}$.

We will consider a broadly generalized version of this example in Section 5.3, where, as an example of the utility of our approach, we provide conditions under which it may be solved via convex programming.

Chapter 2

Quadratic Invariance

We now turn to the main focus of this thesis, which is the characterization of constraint sets S that lead to tractable solutions for problem (1.1). The following property, first introduced in [18], will be shown to provide that characterization.

Definition 2. *Suppose $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, and $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$. The set S is called* ***quadratically invariant*** *under G if*

$$KGK \in S \qquad \text{for all } K \in S$$

Note that, given G, we can define a quadratic map $\Psi : \mathcal{L}(\mathcal{Y}, \mathcal{U}) \to \mathcal{L}(\mathcal{Y}, \mathcal{U})$ by $\Psi(K) = KGK$. Then a set S is quadratically invariant if and only if S is an invariant set of Ψ; that is $\Psi(S) \subseteq S$.

Definition 3. *Given a constraint set $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$, we define a complementary set $S^\star \subseteq \mathcal{L}(\mathcal{U}, \mathcal{Y})$ by*

$$S^\star = \Big\{ G \in \mathcal{L}(\mathcal{U}, \mathcal{Y}) \mid S \text{ is quadratically invariant under } G \Big\}$$

Theorem 4. *If S is a subspace, $S^\star$ is quadratically invariant under K for all $K \in S$.*

Proof. Suppose $K_1, K_2 \in S$ and $G \in S^\star$. First note that

$$K_1GK_2 + K_2GK_1 = (K_1 + K_2)G(K_1 + K_2) - K_1GK_1 - K_2GK_2$$

and since all terms on the right hand side of this equation are in S, we have $K_1GK_2 + K_2GK_1 \in S$. Then we have

$$\begin{aligned} &2K_2GK_1GK_2 = \\ &\quad (K_2 + K_1GK_2 + K_2GK_1)G(K_2 + K_1GK_2 + K_2GK_1) \\ &-(K_1GK_2 + K_2GK_1)G(K_1GK_2 + K_2GK_1) - K_2GK_2 \\ &+(K_1 - K_2GK_2)G(K_1 - K_2GK_2) - K_1GK_1 \end{aligned}$$

and since all terms on the right hand side of this equation are in S, we have $K_2GK_1GK_2 \in S$ for all $K_1, K_2 \in S$ and for all $G \in S^\star$. This implies $GK_1G \in S^\star$ for all $K_1 \in S$ and for all $G \in S^\star$, and the desired result follows. ∎

This tells us that the complementary set is quadratically invariant under any element of the constraint set, which will be very useful in proving the main result of Chapter 4.

We give another general lemma on quadratic invariance which will be useful throughout the remainder of the thesis.

Lemma 5. *Suppose $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, and $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$ is a subspace. If S is quadratically invariant under G, then*

$$K(GK)^n \in S \qquad \textit{for all } K \in S,\ n \in \mathbb{Z}_+$$

Proof. We prove this by induction. By assumption, given $K \in S$, we have that $KGK \in S$. For the inductive step, assume that $K(GK)^n \in S$ for some $n \in \mathbb{Z}_+$. Then

$$2K(GK)^{n+1} = (K + K(GK)^n)G(K + K(GK)^n) - KGK - K(GK)^{2n+1}$$

and since all terms on the right hand side of this equation are in S, we have $K(GK)^{n+1} \in S$. ∎

Chapter 3

Invariance Under Feedback

This chapter contains the main technical results of this thesis. We show that quadratic invariance is necessary and sufficient for the constraint set to be invariant under a linear fractional transformation, namely, the map from K to $K(I - GK)^{-1}$.

This is shown first in Section 3.1 for operators on Banach spaces, and it allows for the closed-loop minimum norm problem to be recast as a convex optimization problem. However, this framework does not allow for the possibility of unstable operators, and the result is subject to a technical condition. In Section 3.2, we turn our focus to causal operators on extended spaces. Making use of the topologies we defined in Section 1.2, we obtain a similar result which applies to possibly unstable operators and which is free from the strictures of technical conditions.

We define the map $h : \mathcal{L}(\mathcal{U}, \mathcal{Y}) \times \mathcal{L}(\mathcal{Y}, \mathcal{U}) \to \mathcal{L}(\mathcal{Y}, \mathcal{U})$ by

$$h(G, K) = -K(I - GK)^{-1}$$

for all G, K such that $I - GK$ is invertible. We will also make use of the notation $h_G(K) = h(G, K)$, which is then defined for all $K \in \mathcal{L}(\mathcal{Y}, \mathcal{U})$ such that $I - GK$ is invertible. Given $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, we note that h_G is an involution on this set, as a straightforward calculation shows that $h_G(h_G(K)) = K$.

3.1 Banach Spaces

In this section, we consider the case where $\mathcal{U}, \mathcal{W}, \mathcal{Y}, \mathcal{Z}$ are Banach spaces and thus P is a bounded linear operator. We then introduce a little more notation.

Given $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, we define the set $M \subseteq \mathcal{L}(\mathcal{Y},\mathcal{U})$ of controllers K such that $f(P, K)$ is well-defined by

$$M = \Big\{ K \in \mathcal{L}(\mathcal{Y},\mathcal{U}) \;\Big|\; (I - GK) \text{ is invertible} \Big\}$$

For any Banach space $\mathcal{X}$ and bounded linear operator $A \in \mathcal{L}(\mathcal{X})$ define the ***resolvent set*** $\rho(A)$ by $\rho(A) = \{\lambda \in \mathbb{C} \mid (\lambda I - A) \text{ is invertible}\}$ and the ***resolvent*** $R_A : \rho(A) \to \mathcal{L}(\mathcal{X})$ by $R_A(\lambda) = (\lambda I - A)^{-1}$ for all $\lambda \in \rho(A)$. We also define $\rho_{uc}(A)$ to be the unbounded connected component of $\rho(A)$.

Note that $1 \in \rho(GK)$ for all $K \in M$, and define the subset $N \subseteq M$ by

$$N = \Big\{ K \in \mathcal{L}(\mathcal{Y},\mathcal{U}) \;\Big|\; 1 \in \rho_{uc}(GK) \Big\}$$

We can now formally state the problem we are to address in this section as follows. Given Banach spaces $\mathcal{U}, \mathcal{W}, \mathcal{Y}, \mathcal{Z}$, generalized plant $P \in \mathcal{L}(\mathcal{W} \times \mathcal{U}, \mathcal{Z} \times \mathcal{Y})$, and a subspace of admissible controllers $S \subseteq \mathcal{L}(\mathcal{Y},\mathcal{U})$, we would like to solve the following problem

$$\begin{aligned} \text{minimize} \quad & \|f(P, K)\| \\ \text{subject to} \quad & K \in M \\ & K \in S \end{aligned} \tag{3.1}$$

3.1.1 Change of Variables

Letting $Q = h_G(K)$, we have

$$f(P, K) = P_{11} - P_{12} Q P_{21}$$

Hence we have the standard parametrization of all closed-loop maps which are achievable by bounded controllers K. This parametrization is related to the well-known *internal model principle* and Youla parametrization of stabilizing controllers. Now we can reformulate problem (3.1) as the following equivalent optimization problem.

$$\begin{aligned} \text{minimize} \quad & \|P_{11} - P_{12}QP_{21}\| \\ \text{subject to} \quad & Q \in M \\ & h_G(Q) \in S \end{aligned} \tag{3.2}$$

The closed-loop map is now affine in Q, and its norm is therefore a convex function of Q. If not for the last constraint, that is, the information constraint S, we could solve this problem to find Q, and then construct the optimal K for problem (3.1) via the transformation $K = h(Q)$.

Note that while we are not considering all $Q \in \mathcal{L}(\mathcal{Y},\mathcal{U})$, only those $Q \in M$, in many cases of practical interest M is dense in $\mathcal{L}(\mathcal{Y},\mathcal{U})$. We will further discuss eliminating this constraint in Section 3.1.3.

However, we see that the information constraint prevents this problem from being easily solved. Specifically, the set

$$\Big\{ Q \in \mathcal{L}(\mathcal{Y},\mathcal{U}) \;\Big|\; h_G(Q) \in S \Big\}$$

is not convex in general. The main thrust of this chapter is to seek conditions under which this last constraint may be converted to a convex constraint.

3.1.2 LFT Invariance

Before proving the main result of this section, we state the following preliminary lemmas regarding analyticity.

Lemma 6. *Suppose $D \subseteq \mathbb{C}$ is an open set, $\mathcal{X}$ is a Banach space, and $f : D \to \mathcal{X}$ is analytic. Suppose that $x \in D$, and $f(y) = 0$ for all y in an open neighborhood of x. Then $f(y) = 0$ for all y in the connected component of D containing x.*

Proof. See for example Theorem 3.7 in [6]. ■

Lemma 7. *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are Banach spaces, $D \subseteq \mathbb{C}$ is an open set, and $A : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator. Suppose $q : D \to \mathcal{X}$ is analytic, and $r : D \to \mathcal{Y}$ is given by $r = A \circ q$. Then r is analytic.*

Proof. This is a straightforward consequence of the definitions. ∎

Lemma 8. *Suppose $K \in \mathcal{L}(\mathcal{Y},\mathcal{U})$, $G \in \mathcal{L}(\mathcal{U},\mathcal{Y})$, and $\Gamma \in \mathcal{L}(\mathcal{Y},\mathcal{U})^*$. Define the function $q_\Gamma : \rho(GK) \to \mathbb{C}$ by*

$$q_\Gamma(\lambda) = \langle K R_{GK}(\lambda), \Gamma \rangle.$$

Then q_Γ is analytic.

Proof. Define the linear map $\gamma : \mathcal{L}(\mathcal{Y}) \to \mathbb{C}$ by

$$\gamma(G) = \langle KG, \Gamma \rangle \quad \text{for all } G \in \mathcal{L}(\mathcal{Y}).$$

Clearly γ is bounded, since

$$\|\gamma(G)\| \leq \|K\| \|\Gamma\| \|G\| \quad \text{for all } G \in \mathcal{L}(\mathcal{Y}).$$

Further $q_\Gamma = \gamma \circ R_{GK}$, and the resolvent is analytic, hence by Lemma 7 we have that q_Γ is analytic. ∎

The following lemma will be useful for proving the converse of our main results both in this section and in Section 3.2.

Lemma 9. *Suppose $\mathcal{U}, \mathcal{Y}$ are Banach spaces, $G \in \mathcal{L}(\mathcal{U},\mathcal{Y}), S \subseteq \mathcal{L}(\mathcal{Y},\mathcal{U})$ is a closed subspace, and S is not quadratically invariant under G. Then there exists $K \in S$ such that $I - GK$ is invertible and $h_G(K) \notin S$.*

Proof. There exists $K_0 \in S$ such that $K_0 G K_0 \notin S$. We will construct $K \in S$ such that $h_G(K) \notin S$. Without loss of generality we may assume $\|K_0\| = 1$. Choose $\Gamma \in S^\perp$ with $\|\Gamma\| = 1$ such that

$$\beta = \langle K_0 G K_0, \Gamma \rangle \in \mathbb{R} \quad \text{and} \quad \beta > 0,$$

and choose $\alpha \in \mathbb{R}$ such that

$$0 < \alpha < \frac{\beta}{\|G\|\big(\beta + \|G\|\big)}$$

Let $K = \alpha K_0$. Then $\|GK\| < 1$, $K \in S$, and

$$\langle K(I - GK)^{-1}, \Gamma\rangle = \sum_{i=0}^{\infty} \langle K(GK)^i, \Gamma\rangle$$

Thus

$$\begin{aligned}
\big|\langle K(I - GK)^{-1}, \Gamma\rangle\big| &= \left|\sum_{i=0}^{\infty} \langle K(GK)^i, \Gamma\rangle\right| \\
&= \left|\alpha^2\beta + \sum_{i=2}^{\infty} \langle K(GK)^i, \Gamma\rangle\right| \\
&\geq \alpha^2\beta - \alpha \sum_{i=2}^{\infty} \|G\|^i \alpha^i \\
&= \alpha^2 \left(\frac{\beta - \alpha\|G\|\big(\beta + \|G\|\big)}{1 - \alpha\|G\|}\right) \\
&> 0
\end{aligned}$$

Hence $K(I - GK)^{-1} \notin S$ as required. ■

Main result - Banach spaces. The following is the main result of this section. It states that given G, if a certain technical condition holds, then the constraint set S is quadratically invariant if and only if the information constraints on K are equivalent to the same *affine constraints* on the parameter $Q = h_G(K)$. In other words, subject to technical conditions, quadratic invariance is equivalent to invariance under feedback.

Theorem 10. *Suppose $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, and $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$ is a closed subspace. Further* *suppose* $N \cap S = M \cap S$. *Then*

$$S \text{ is quadratically invariant under } G \iff h_G(S \cap M) = S \cap M$$

Proof. ($\Longrightarrow$) Suppose $K \in S \cap M$. We first show that $h_G(K) \in S \cap M$. For any $\Gamma \in S^\perp$ define the function $q_\Gamma : \rho(GK) \to \mathbb{C}$ by

$$q_\Gamma(\lambda) = \langle K(\lambda I - GK)^{-1}, \Gamma \rangle.$$

For any λ such that $|\lambda| > \|GK\|$, the Neumann series expansion for R_{GK} gives

$$K(\lambda I - GK)^{-1} = \sum_{n=0}^{\infty} \lambda^{-(n+1)} K(GK)^n$$

By Lemma 5 we have $K(GK)^n \in S$ for all $n \in \mathbb{Z}_+$, and hence $K(\lambda I - GK)^{-1} \in S$ since S is a closed subspace. Thus,

$$q_\Gamma(\lambda) = 0 \quad \text{for all } \lambda \text{ such that } |\lambda| > \|GK\|$$

By Lemma 8, the function q_Γ is analytic, and since $\lambda \in \rho_{uc}(GK)$ for all $|\lambda| > \|GK\|$, by Lemma 6 we have

$$q_\Gamma(\lambda) = 0 \quad \text{for all } \lambda \in \rho_{uc}(GK).$$

It follows from $K \in N$ that $1 \in \rho_{uc}(GK)$, and therefore $q_\Gamma(1) = 0$. Hence

$$\langle K(I - GK)^{-1}, \Gamma \rangle = 0 \quad \text{for all } \Gamma \in S^\perp.$$

This implies

$$K(I - GK)^{-1} \in {}^\perp(S^\perp).$$

Since S is a closed subspace, we have ${}^\perp(S^\perp) = S$ (see for example [12], p. 118) and hence we have shown $K \in S \cap M \implies h_G(K) \in S$. Since h is a bijective involution on M, it follows that $h_G(S \cap M) = S \cap M$ which was the desired result.

($\Longleftarrow$) The converse follows immediately from Lemma 9. ∎

There are many cases of interest where the technical condition above is automatically satisfied, specifically any case where the plant is such that the resolvent set of GK is always connected. This includes the case where G is compact, such as for any problem where the Banach spaces are finite dimensional, as shown in the following corollary.

Corollary 11. *Suppose $G \in \mathcal{L}(\mathcal{U},\mathcal{Y})$ is compact and $S \subseteq \mathcal{L}(\mathcal{Y},\mathcal{U})$ is a closed subspace. Then*

$$S \text{ is quadratically invariant under } G \iff h_G(S \cap M) = S \cap M$$

Proof. This follows since if G is compact then GK is compact for any $K \in S$, and hence the spectrum of GK is countable, and so $N = M$. ■

3.1.3 Equivalent Problems

When the conditions of Theorem 10 are met, we have the following equivalent problem. K is optimal for problem (3.1) if and only if $K = h_G(Q)$ and Q is optimal for

$$\begin{aligned} \text{minimize} \quad & \|P_{11} - P_{12}QP_{21}\| \\ \text{subject to} \quad & Q \in M \\ & Q \in S \end{aligned} \tag{3.3}$$

We then consider further reducing the problem to the following convex optimization problem.

$$\begin{aligned} \text{minimize} \quad & \|P_{11} - P_{12}QP_{21}\| \\ \text{subject to} \quad & Q \in S \end{aligned} \tag{3.4}$$

If the minimum to this problem did occur for some $Q \notin M$, it would mean that the original problem approaches its optimum as $(I - GK)^{-1}$ blows up. Therefore the constraint that $Q \in M$ is unnecessary when the original problem is well-posed.

The solution procedure is then to solve problem (3.4) with convex programming, and recover the optimal controller for the original problem (3.1) as $K = h_G(Q)$.

3.1.4 Violation

The technical conditions of Theorem 10 are automatically satisfied in many cases of interest, such as in Corollary 11. However, the conditions cannot always be eradicated. We show an example where the resolvent set is not connected, the technical conditions thus fail, and the main result of this section no longer holds. This is meant to elucidate that the conditions are not for ease of proof, but are actually needed.

Consider the space of two-sided sequences

$$\ell_2 = \left\{ (\ldots, x_{-1}, x_0, x_1, \ldots) \;\middle|\; x_i \in \mathbb{R},\ \sum_{i=-\infty}^{\infty} x_i^2 < \infty \right\}.$$

Define the delay operator $D : \ell_2 \to \ell_2$ as $D(x)_i = x_{i-1}$. Let $\mathcal{Y} = \mathcal{U} = \ell_2$, let the plant be the identity $G = I$, and let S be the subspace of causal controllers

$$S = \{K \in \mathcal{L}(\ell_2) \mid P_T K P_T = P_T K \text{ for all } T \in \mathbb{Z}\}$$

such that S is clearly quadratically invariant under G. Now consider $K = 2D \in S$; we have

$$\begin{aligned}(I - GK)^{-1} &= -\frac{1}{2}D^{-1}\left(I - \frac{1}{2}D^{-1}\right)^{-1} \\ &= -\sum_{k=1}^{\infty} \frac{1}{2^k} D^{-k}\end{aligned}$$

and so $K \in M$. Also note that

$$\rho(GK) = \{\lambda \in \mathbb{C} \mid |\lambda| \neq 2\}$$

and hence $\rho_{uc}(GK) = \{\lambda \in \mathbb{C} \mid |\lambda| > 2\}$, which implies that $K \notin N$. Finally,

$$K(I - GK)^{-1} = -\sum_{k=0}^{\infty} \frac{1}{2^k} D^{-k} \notin S$$

So we have $G \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, $S \subseteq \mathcal{L}(\mathcal{Y}, \mathcal{U})$ is a closed subspace, and S is quadratically invariant under G, but $N \cap S \neq M \cap S$. We have then found a $K \in S \cap M$ such that $h_G(K) \notin S$, and so $h_G(S \cap M) \neq S \cap M$.

This shows that the technical conditions of Theorem 10 cannot be completely eradicated for arbitrary Banach spaces, and motivates us to find a different framework under which a similar result can be achieved without them.

3.2 Extended Spaces

In this section, we turn our focus to time-invariant causal operators on extended spaces. This framework allows us to both extend our main results to unbounded operators and to eliminate technical conditions from our assumptions.

From Lemma 5, and from the proof in the previous section, we see that if we could express $K(I-GK)^{-1}$ as $\sum_{n=0}^{\infty} K(GK)^n$, then this mapping would lie in S for all $K \in S$, provided that S was a closed subspace. Noting that K can be pulled outside of the sum due to continuity, even when it is unbounded, we thus seek conditions under which $\sum_{n=0}^{\infty}(GK)^n$ converges. We would like to be able to utilize this expansion not just for small K, as in the small gain theorem, but for arbitrarily large K as well. We consider the plant and controller as operators on extended spaces both because that will allow us to achieve this, and also so that unstable operators may be considered.

In Section 3.2.1, we develop conditions under which this Neumann series is guaranteed to converge in the topologies defined in Section 1.2. We then prove in Section 3.2.2 that under very broad assumptions quadratic invariance in necessary and sufficient for the constraint set to be preserved under feedback. These conditions include, but are not limited to, the case we are often interested in where $G \in \mathcal{R}_{sp}$ and $S \subseteq \mathcal{R}_p$.

3.2.1 Convergence of Neumann Series

We first analyze convergence of the *Neumann series*

$$(I - W)^{-1} = \sum_{n=0}^{\infty} W^n$$

where W is a general causal time-invariant linear operator on extended spaces. Note that while most of the results in this thesis have analogs in both continuous-time and discrete-time, the convergence proofs must be worked out separately. We first analyze the continuous-time case, and begin by providing a preliminary lemma which states that if a sequence of impulse responses converge in a particular sense, then their associated operators do as well.

Lemma 12. *Suppose $W_n \in \mathcal{L}(L_{2e}^m)$ is causal and time-invariant for all $n \in \mathbb{Z}_+$, $w^{(n)} \in L_{\infty e}$ is the impulse response of W_n, $a \in L_{\infty e}$ and $(w^{(n)})_T$ converges uniformly to a_T as $n \to \infty$ for all $T \in \mathbb{R}_+$. Then W_n converges to $A \in \mathcal{L}(L_{2e}^m)$, where A is given by $Au = a * u$.*

Proof. Given $u \in L_{2e}^m$ and $T \in \mathbb{R}_+$,

$$(a * u)_T = (a_T * u_T)_T$$

since $a(t) = 0$ and $u(t) = 0$ for $t < 0$. Hence $(a*u)_T \in L_2$, since $a_T \in L_1$ and $u_T \in L_2$, by Theorem 65 of [25]. Therefore, we can define $A \in \mathcal{L}(L_{2e}^m)$ by $Au = a * u$.

For any $n \in \mathbb{Z}_+$ and any $T \in \mathbb{R}_+$,

$$\begin{aligned} \|A - W_n\|_T^2 &= \sup_{u \in L_2, \|u\|_2 = 1} \|P_T A u - P_T W_n u\|_2^2 \\ &\leq \sup_{u \in L_2, \|u\|_2 = 1} \left\| \left(a_T - (w^{(n)})_T \right) * u \right\|_2^2 \\ &\leq \sup_{u \in L_2, \|u\|_2 = 1} \sum_{i=1}^{m} \sum_{j=1}^{m} \left\| \left(a_T - (w^{(n)})_T \right)_{ij} \right\|_1^2 \|u_j\|_2^2 \end{aligned}$$

and hence

$$\|A - W_n\|_T^2 \leq \sum_{i=1}^{m} \sum_{j=1}^{m} \left\| \left(a_T - (w^{(n)})_T \right)_{ij} \right\|_1^2$$

Since $(w_{ij}^{(n)})_T$ converges uniformly to a_T, for any $\epsilon > 0$ we can choose N such that for all $n \geq N$ and for all $i, j = 1, \ldots, m$, $\left| (a_{ij})_T(t) - (w_{ij}^{(n)})_T(t) \right| < \frac{\epsilon}{mT}$ for all $t \in [0, T]$

and thus $\|A - W_n\|_T < \epsilon$. So W_n converges to A in $\mathcal{L}(L_{2e}^m)$. ■

We can now prove convergence of the Neumann series under the given conditions by showing the convergence of impulse responses. The method for showing this is similar to that used for spatio-temporal systems in the appendix of [2].

Theorem 13. *Suppose $W \in \mathcal{L}(L_{2e}^m)$ is causal and time-invariant with impulse response matrix w such that $w \in L_{\infty e}$. Then $\sum_{n=0}^{\infty} W^n$ converges to an element $B \in \mathcal{L}(L_{2e}^m)$ such that $B = (I - W)^{-1}$.*

Proof. Let $q(T) = \sup_{t\in[0,T]} \|w(t)\| < \infty$ for all $T \in \mathbb{R}_+$, and let $w^{(n)}$ be the impulse response matrix of W^n. First we claim that $\|w^{(n)}(T)\| \le \frac{T^{n-1}}{(n-1)!}\, q(T)^n$ for all integers $n \ge 1$. This is true immediately for $n = 1$. For the inductive step,

$$\begin{aligned}
\left\|w^{(n+1)}(T)\right\| &= \left\|\int_{t=0}^{T} w(T-t)w^{(n)}(t)dt\right\| \\
&\le \int_{t=0}^{T} \|w(T-t)\| \cdot \|w^{(n)}(t)\|dt \\
&\le q(T)\int_{t=0}^{T} \|w^{(n)}(t)\|dt \\
&\le q(T)\int_{t=0}^{T} \frac{t^{n-1}}{(n-1)!} q(t)^n dt \\
&\le \frac{T^n}{n!}\, q(T)^{n+1}
\end{aligned}$$

Then $|w_{ij}^{(n)}(t)| \le \frac{T^{n-1}}{(n-1)!}\, q(T)^n$ for all $t \in [0,T]$,for all $n \ge 1$, and for all $i, j = 1, \ldots, m$. $\sum_{n=1}^{\infty} \frac{T^{n-1}}{(n-1)!}\, q(T)^n$ converges to $q(T)e^{Tq(T)}$, so by the Weierstrass M-test, $\sum_{n=1}^{\infty} \big(w_{ij}^{(n)}\big)_T$ converges uniformly and absolutely for all $i, j = 1, \ldots, m$.

Let $a = \sum_{n=1}^{\infty} w^{(n)}$. Then $a_{ij} \in L_{\infty e} \subseteq L_{1e}$ for all $i, j = 1, \ldots, m$, and we can define $A, B \in \mathcal{L}(L_{2e}^m)$ by $Au = a * u$ and $B = I + A$.

Then by Lemma 12, $\sum_{k=1}^{n} W^k$ converges to A in $\mathcal{L}(L_{2e}^m)$, and thus $\sum_{k=0}^{n} W^k$ converges to B in $\mathcal{L}(L_{2e}^m)$.

Lastly,

$$B(I - W) = (I - W)B = \sum_{n=0}^{\infty} W^n - \sum_{n=1}^{\infty} W^n = I$$

∎

A simple example of the utility of this result is as follows. Consider W represented by the transfer function $\frac{2}{s+1}$. Then $I - W = \frac{s-1}{s+1}$ is not invertible in $\mathcal{L}(L_2)$. However using the above theorem, the inverse in $\mathcal{L}(L_{2e})$ is given by $\sum_{n=0}^{\infty}(\frac{2}{s+1})^n = \frac{s+1}{s-1}$.

We now move on to analyze the discrete-time case. Let $r(\cdot)$ denote spectral radius.

Theorem 14. *Suppose $W \in \mathcal{L}(\ell_e^m)$ is causal and time-invariant with impulse response matrix w such that $w \in \ell_e$ and $r(w(0)) < 1$. Then $\sum_{n=0}^{\infty} W^n$ converges to an element $B \in \mathcal{L}(\ell_e^m)$ such that $B = (I - W)^{-1}$.*

Proof. We may represent $P_T W|_T$ with the block lower triangular Toeplitz matrix

$$W_T = \begin{bmatrix} w(0) & & \\ w(1) & \ddots & \\ \vdots & & \\ w(T) & \cdots & w(0) \end{bmatrix}$$

Since $w \in \ell_e$, $W_T \in \mathbb{R}^{mT \times mT}$. Then, $r(W_T) = r(w(0)) < 1$, which implies that $\sum_{n=0}^{\infty}(W_T)^n$ converges in $\mathbb{R}^{mT \times mT}$. Thus we can define $B \in \mathcal{L}(\ell_e^m)$ by $(Bu)_T = (\sum_{n=0}^{\infty}(W_T)^n)u_T$ for any $u \in \ell_e^m$ and any $T \in \mathbb{Z}_+$. It is then immediate that $\left\|B - \sum_{n=0}^{\infty} W^n\right\|_T \to 0$ as $n \to \infty$ for all T, and thus $\sum_{n=0}^{\infty} W^n$ converges to B in $\mathcal{L}(\ell_e^m)$.

Lastly,

$$B(I - W) = (I - W)B = \sum_{n=0}^{\infty} W^n - \sum_{n=1}^{\infty} W^n = I$$

∎

Note that while the conditions of Theorem 14 are necessary for convergence as well as sufficient, the conditions of Theorem 13 are not.

In particular, the above results imply the following corollary, which shows convergence of the Neumann series for strictly proper systems, possibly with delay. Note that the delay is redundant for discrete-time systems.

Corollary 15. *Suppose $W \in \mathcal{L}(L_{2e}^m)$ or $W \in \mathcal{L}(\ell_e^m)$ is given by $W_{ij} = D_{\tau_{ij}} G_{ij}$ where*

$\tau_{ij} \geq 0$ and $G_{ij} \in \mathcal{R}_{sp}$. Then $\sum_{n=0}^{\infty} W^n$ converges to an element $B \in \mathcal{L}(L_{2e}^m)$ such that $B = (I - W)^{-1}$.

3.2.2 LFT Invariance

In this section, we will show that for a very broad class of systems, quadratic invariance is necessary and sufficient for the information constraint S to be invariant under a linear-fractional transformation.

We first state a lemma which will help with the converse of our main result.

Lemma 16. *Suppose $S \subseteq \mathcal{L}(L_{2e}^m, L_{2e}^n)$ or $S \subseteq \mathcal{L}(\ell_e^m, \ell_e^n)$, and $C \notin S$. Then there exists T such that $C_T \notin S_T$.*

Proof. Suppose not. Then for every positive T, $C_T \in S_T$. Thus for every T, there exists $K \in S$ such that $P_T C|_T = P_T K|_T$, or $\|C - K\|_T = 0$. Since $\|A\|_T = 0$ only if $\|A\|_\tau = 0$ for all $\tau \leq T$, it follows that there exists $K \in S$ such that $\|C - K\|_T = 0$ for all T. But then $C - K = 0$, and so $C \in S$ and we have a contradiction. ■

We define a broad class of sets of controllers for which the closed-loop map will always be well-defined. Note that this includes the case which is often of interest where $G \in \mathcal{R}_{sp}$ and $S \subseteq \mathcal{R}_p$.

Definition 17. *We say that $S \subseteq \mathcal{L}(L_{2e}^{n_u}, L_{2e}^{n_y})$ is **inert** with respect to G if for all $K \in S$, $(gk)_{ij} \in L_{\infty e}$ for all $i, j = 1, \ldots, m$ where (gk) is the impulse response matrix of GK. We overload our notation and also define $S \subseteq \mathcal{L}(\ell_e^{n_u}, \ell_e^{n_y})$ to be **inert** if for all $K \in S$, $(gk)_{ij} \in \ell_e$ for all $i, j = 1, \ldots, m$ and $r((gk)(0)) < 1$ where (gk) is the discrete impulse response matrix of GK.*

Main result - extended spaces. The following theorem is the main result of this section. It states that quadratic invariance of the constraint set is necessary and sufficient for the set to be invariant under the LFT defined by h_G.

Theorem 18. *Suppose $G \in \mathcal{L}(L_{2e}^{n_u}, L_{2e}^{n_y})$ or $G \in \mathcal{L}(\ell_e^{n_u}, \ell_e^{n_y})$, and S is an inert closed subspace. Then*

$$S \text{ is quadratically invariant under } G \iff h_G(S) = S$$

Proof. ($\Longrightarrow$) Suppose $K \in S$. We first show that $h_G(K) \in S$.

$$K(I - GK)^{-1} = K \sum_{n=0}^{\infty} (GK)^n = \sum_{n=0}^{\infty} K(GK)^n$$

where the first equality follows from Theorems 13 and 14 and the second follows from the continuity of K.

By Lemma 5 we have $K(GK)^n \in S$ for all $n \in \mathbb{Z}_+$, and hence $K(I - GK)^{-1} \in S$ since S is a closed subspace.

So $K \in S \implies h_G(K) \in S$. Thus $h_G(S) \subseteq S$, and since h_G is involutive it follows that $h_G(S) = S$, which was the desired result.

($\Longleftarrow$) We now turn to the converse of this result. Suppose that S is not quadratically invariant under G. Then there exists $K \in S$ such that $KGK \notin S$, and thus by Lemma 16, there exists a finite T such that $P_T KGK|_T \notin S_T$. Since K and G are causal, we then have

$$K_T G_T K_T \notin S_T \qquad \text{where} \qquad K_T = P_T K P_T \in S_T \quad \text{and} \quad G_T = P_T G P_T$$

and thus S_T is not quadratically invariant under G_T. Then by Lemma 9 there exists $\tilde{K} \in S_T$ such that

$$\tilde{K}(I - G_T \tilde{K})^{-1} = \sum_{n=0}^{\infty} \tilde{K}(G_T \tilde{K})^n \notin S_T$$

By definition of S_T, there exists $K_0 \in S$ such that $\tilde{K} = P_T K_0|_T$. Then by causality of K_0 and G,

$$P_T \left(\sum_{n=0}^{\infty} K_0 (G K_0)^n \right) \Bigg|_T \notin S_T$$

and thus $h_G(K_0) = -\sum_{n=0}^{\infty} K_0 (G K_0)^n \notin S$. ∎

Chapter 4

Optimal Stabilizing Controllers

In this chapter we consider the problem of finding optimal stabilizing controllers subject to an information constraint. These results apply both to continuous-time and discrete-time systems. Note that throughout this section, the constraint set S is always inert, since $G \in \mathcal{R}_{sp}$ and $S \subseteq \mathcal{R}_p$.

The optimization problem we address is as follows. Given $P \in \mathcal{R}_p^{(n_z+n_y)\times(n_w+n_u)}$, and a subspace of admissible controllers $S \subseteq \mathcal{R}_p^{n_u \times n_y}$, we would like to solve:

$$\begin{aligned} \text{minimize} \quad & \|f(P,K)\| \\ \text{subject to} \quad & K \text{ stabilizes } P \\ & K \in S \end{aligned} \tag{4.1}$$

There have been several key results regarding controller parameterization and optimization which we will extend for decentralized control, relying heavily on our result from the previous chapter. The celebrated Youla parameterization [33] showed that given a coprime factorization of the plant, one may parameterize all stabilizing controllers. The set of closed-loop maps achievable with stabilizing controllers is then affine in this parameter, an important result which converts the problem of finding the optimal stabilizing controller to a convex optimization problem, given the factorization. Zames proposed a two-step compensation scheme [34] for strongly stabilizable plants, that is, plants which can be stabilized with a stable compensator.

In the first step one finds any controller which is both stable and stabilizing, and in the second one optimizes over a parameterized family of systems. This idea has been extended to nonlinear control [1], and in this chapter we show that it may be extended to decentralized control when the constraint set is quadratically invariant, as first shown in [19].

Our approach starts with a single nominal decentralized controller which is both stable and stabilizing, and uses it to parameterize all stabilizing decentralized controllers. The resulting parameterization expresses the closed-loop system as an affine function of a stable parameter, allowing the next step, optimization of closed-loop performance, to be achieved with convex programming. Techniques for finding an initial stabilizing controller for decentralized systems are discussed in detail in [23], and conditions for decentralized stabilizability were developed in [28].

In the final section of this chapter, we show that problem (4.1) can still be converted to a convex optimization problem when such a nominal controller is unobtainable.

4.1 Stabilization

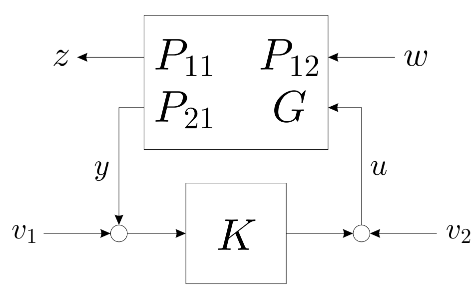

Figure 4.1: Linear fractional interconnection of P and K

We say that K ***stabilizes*** P if in Figure 4.1 the nine transfer matrices from (w, v_1, v_2) to (z, u, y) belong to $\mathcal{RH}_\infty$. We say that K ***stabilizes*** G if in the figure the four transfer matrices from (v_1, v_2) to (u, y) belong to $\mathcal{RH}_\infty$. P is called ***stabilizable*** if there exists $K \in \mathcal{R}_p^{n_u \times n_y}$ such that K stabilizes P, and it is called is called ***strongly stabilizable*** if there exists $K \in \mathcal{RH}_\infty^{n_u \times n_y}$ such that K stabilizes P. We denote by

$C_{\text{stab}} \subseteq \mathcal{R}_p^{n_u \times n_y}$ the set of controllers $K \in \mathcal{R}_p^{n_u \times n_y}$ which stabilize P. The following standard result relates stabilization of P with stabilization of G.

Theorem 19. *Suppose $G \in \mathcal{R}_{sp}^{n_y \times n_u}$ and $P \in \mathcal{R}_p^{(n_z+n_y)\times(n_w+n_u)}$, and suppose P is stabilizable. Then K stabilizes P if and only if K stabilizes G.*

Proof. See, for example, Chapter 4 of [9]. ■

4.2 Parameterization of Stabilizing Controllers

In this section, we review one well-known approach to solution of the feedback optimization problem (4.1) when the constraint that K lie in S is not present. In this case, one may use the following standard change of variables.

For a given system P, all controllers that stabilize the system may be parameterized using the well-known Youla parameterization [33], stated below.

Theorem 20. *Suppose we have a doubly coprime factorization of G over $\mathcal{RH}_\infty$, that is, $M_l, N_l, X_l, Y_l, M_r, N_r, X_r, Y_r \in \mathcal{RH}_\infty$ such that $G = N_r M_r^{-1} = M_l^{-1} N_l$ and*

$$\begin{bmatrix} X_l & -Y_l \\ -N_l & M_l \end{bmatrix} \begin{bmatrix} M_r & Y_r \\ N_r & X_r \end{bmatrix} = I$$

Then the set C_{stab} of all controllers in $\mathcal{R}_p$ which stabilize G is

$$C_{\text{stab}} = \left\{ (Y_r - M_r Q)(X_r - N_r Q)^{-1} \;\middle|\; X_r - N_r Q \text{ is invertible}, Q \in \mathcal{RH}_\infty \right\}$$

Furthermore, the set of all closed-loop maps achievable with stabilizing controllers is

$$\begin{aligned} &\left\{ f(P,K) \;\middle|\; K \in \mathcal{R}_p, K \text{ stabilizes } P \right\} \\ &\qquad = \left\{ T_1 - T_2 Q T_3 \;\middle|\; X_r - N_r Q \text{ is invertible}, Q \in \mathcal{RH}_\infty \right\} \end{aligned} \tag{4.2}$$

where $T_1, T_2, T_3 \in \mathcal{RH}_\infty$ *are given by*

$$\begin{aligned} T_1 &= P_{11} + P_{12} Y_r M_l P_{21} \\ T_2 &= P_{12} M_r \\ T_3 &= M_l P_{21} \end{aligned}$$

Proof. See, for example, Chapter 4 of [9]. ∎

This parameterization is particularly simple to construct in the case where we have a nominal stabilizing controller $K_{\text{nom}} \in \mathcal{RH}_\infty$; that is, a controller which is both stable and stabilizing.

Theorem 21. *Suppose* G *is strictly proper, and* $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty$*. Then all stabilizing controllers are given by*

$$C_{\text{stab}} = \Big\{ K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \;\Big|\; Q \in \mathcal{RH}_\infty \Big\}$$

and all closed-loop maps are given by (4.2) *where*

$$\begin{aligned} T_1 &= P_{11} + P_{12} K_{\text{nom}} (I - G K_{\text{nom}})^{-1} P_{21} \\ T_2 &= -P_{12} (I - K_{\text{nom}} G)^{-1} \\ T_3 &= (I - G K_{\text{nom}})^{-1} P_{21} \end{aligned} \tag{4.3}$$

Proof. A doubly coprime factorization for G over $\mathcal{RH}_\infty$ is given by

$$\begin{aligned} M_l &= (I - G K_{\text{nom}})^{-1} & M_r &= -(I - K_{\text{nom}} G)^{-1} \\ N_l &= G (I - K_{\text{nom}} G)^{-1} & N_r &= -G (I - K_{\text{nom}} G)^{-1} \\ X_l &= -I \quad Y_l = -K_{\text{nom}} & X_r &= I \quad Y_r = K_{\text{nom}} \end{aligned}$$

Then

$$\begin{aligned}(Y_r - M_r Q)(X_r - N_r Q)^{-1} \\ &= K_{\text{nom}} + Q(I + G(I - K_{\text{nom}}G)^{-1}Q)^{-1} \\ &= K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big)\end{aligned}$$

so all stabilizing controllers are given by

$$C_{\text{stab}} = \Big\{\, K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \;\Big|\; Q \in \mathcal{RH}_\infty \,\Big\}$$

The invertibility condition is met since G, and thus N_r, is strictly proper. ∎

This theorem tells us that if the plant is strongly stabilizable, that is, it can be stabilized by a stable controller, then given such a controller, we can parameterize the set of all stabilizing controllers. See [34] for a discussion of this, and [1] for an extension to nonlinear control. The parameterization above is very useful, since in the absence of the constraint $K \in S$, problem (4.1) can be reformulated as

$$\begin{aligned} \text{minimize} \quad & \|T_1 - T_2 Q T_3\| \\ \text{subject to} \quad & Q \in \mathcal{RH}_\infty \end{aligned} \tag{4.4}$$

The closed-loop map is now affine in Q, and its norm is therefore a convex function of Q. This problem is readily solvable by, for example, the techniques in [4]. After solving this problem to find Q, one may then construct the optimal K for problem (4.1) via $K = K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big)$.

4.3 Parameterization of Admissible Controllers

We now wish to extend the above result to parameterize all stabilizing controllers $K \in \mathcal{R}_p$ which also satisfy the information constraint $K \in S$. Applying the above change of variables to problem (4.1), we arrive at the following optimization problem.

$$\begin{aligned} \text{minimize} \quad & \|T_1 - T_2 Q T_3\| \\ \text{subject to} \quad & Q \in \mathcal{RH}_\infty \\ & K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \in S \end{aligned} \tag{4.5}$$

However, the set of all Q which satisfy this last constraint is not convex in general, and hence this problem is not easily solved. We thus develop conditions under which this set is in fact convex, so that the optimization problem (4.5) may be solved via convex programming. First we state a preliminary lemma.

Lemma 22. *Suppose $G \in \mathcal{R}_{sp}$, $S \subseteq \mathcal{R}_p$ is a closed subspace, and $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty \cap S$. Then S is quadratically invariant under $h(K_{\text{nom}}, G)$ if and only if*

$$S = \Big\{ K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \;\Big|\; Q \in S \Big\}$$

Proof. ($\Longrightarrow$) Suppose S is quadratically invariant under $h(K_{\text{nom}}, G)$, and further suppose there exists $Q \in S$ such that

$$K = K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big).$$

Since S is quadratically invariant under $h(K_{\text{nom}}, G)$ and S is an inert subspace, Theorem 18 implies that $h\big(h(K_{\text{nom}}, G), Q\big) \in S$, and since $K_{\text{nom}} \in S$ as well, $K \in S$.

Now suppose $K \in S$. Let

$$Q = h\big(h(K_{\text{nom}}, G), K_{\text{nom}} - K\big).$$

We know $K_{\text{nom}} - K \in S$, and since S is quadratically invariant under $h(K_{\text{nom}}, G)$, then by Theorem 18, we also have $Q \in S$.

($\Longleftarrow$) Now suppose S is not quadratically invariant under $h(K_{\text{nom}}, G)$. Then by Theorem 18 there exists $Q \in S$ such that $h\big(h(K_{\text{nom}}, G), Q\big) \notin S$, and thus $K = K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \notin S$. ∎

This lemma shows that if we can find a stable $K_{\text{nom}} \in S$ which is stabilizing, and if the condition that S is quadratically invariant under $h(K_{\text{nom}}, G)$ holds, then the set of all stabilizing admissible controllers can be easily parameterized with the same change of variables from Theorem 21. We now simplify this condition.

Main result - optimal stabilization. The following theorem is the main result of this chapter. It states that if the constraint set is quadratically invariant under the plant, then the information constraints on K are equivalent to *affine constraints* on the Youla parameter Q. Specifically, the constraint $K \in S$ is equivalent to the constraint $Q \in S$.

Theorem 23. *Suppose $G \in \mathcal{R}_{sp}$, $S \subseteq \mathcal{R}_p$ is a closed subspace, and $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty \cap S$. If S is quadratically invariant under G then*

$$S = \Big\{ K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \;\Big|\; Q \in S \Big\}$$

Proof. If S is quadratically invariant under G, then $G \in S^\star$. Further, by Theorem 4, $S^\star$ is quadratically invariant under K_{nom}, and then by Theorem 18, we have $h(K_{\text{nom}}, S^\star) = S^\star$. We then have $h(K_{\text{nom}}, G) \in S^\star$, and therefore S is quadratically invariant under $h(K_{\text{nom}}, G)$. By Lemma 22, this yields the desired result. ■

Remark 24. *When P is stable, we can choose $K_{\text{nom}} = 0$ and the parameterization reduces to $h_G(Q)$.*

Remark 25. *When $S = \mathcal{R}_p^{n_u \times n_y}$, which corresponds to centralized control, then the quadratic invariance condition is met and the result reduces to Theorem 21.*

4.4 Equivalent Convex Problem

When the constraint set is quadratically invariant under the plant, we now have the following equivalent problem. Suppose $G \in \mathcal{R}_{sp}^{n_y \times n_u}$ and $S \subseteq \mathcal{R}_p^{n_u \times n_y}$ is a closed subspace. Then K is optimal for problem (4.1) if and only if $K = K_{\text{nom}} -$

$h\big(h(K_{\text{nom}}, G), Q\big)$ and Q is optimal for

$$\begin{aligned} \text{minimize} \quad & \|T_1 - T_2 Q T_3\| \\ \text{subject to} \quad & Q \in \mathcal{RH}_\infty \\ & Q \in S \end{aligned} \tag{4.6}$$

where $T_1, T_2, T_3 \in \mathcal{RH}_\infty$ are given by equations (4.3). This problem may be solved via convex programming.

4.5 Convexity without Strong Stabilizability

Suppose that one cannot find a $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty \cap S$; that is, a controller with the admissible structure which is both stable and stabilizing. This may occur either because the plant is not strongly stabilizable, or simply because it is difficult to find. In this section we will show that problem (4.1) can still be reduced to a convex optimization problem, albeit one which is less straightforward to solve.

We will achieve this by bypassing the Youla parameterization, and using the change of variables previously associated with stable or bounded plants

$$R \;=\; h_G(K) \;=\; -K(I - GK)^{-1}$$

where R will be used instead of Q to elucidate that this *is not* a Youla parameter. The key observation is that internal stabilization is equivalent to an affine constraint in this parameter.

The constraint that K stabilize G, which is equivalent to the constraint that K stabilize P when the standard conditions of Theorem 19 hold, is defined as requiring that the maps from (v_1, v_2) to (u, y) in Figure 4.1 belong to $\mathcal{RH}_\infty$. This can be stated explicitly as

$$\begin{bmatrix} (I - KG)^{-1} & (I - KG)^{-1}K \\ G(I - KG)^{-1} & G(I - KG)^{-1}K \end{bmatrix} \in \mathcal{RH}_\infty$$

Making use of the relations

$$\begin{aligned}(I - GK)^{-1}G &= G(I - KG)^{-1} \\ (I - KG)^{-1} &= I + K(I - GK)^{-1}G\end{aligned}$$

we find that K stabilizes G if and only if

$$\begin{bmatrix} RG & R \\ G - GRG & GR \end{bmatrix} \in \mathcal{RH}_\infty \tag{4.7}$$

Suppose $G \in \mathcal{R}_{sp}^{n_y \times n_u}$ and $S \subseteq \mathcal{R}_p^{n_u \times n_y}$ is a quadratically invariant closed subspace. We may then use this result to transform the stabilization constraint of problem (4.1) and Theorem 18 to transform the information constraint to obtain the following equivalent problem. K is optimal for problem (4.1) if and only if $K = h_G(R)$ and R is optimal for

$$\begin{aligned}\text{minimize} \quad & \|P_{11} - P_{12}RP_{21}\| \\ \text{subject to} \quad & \begin{bmatrix} RG & R \\ G - GRG & GR \end{bmatrix} \in \mathcal{RH}_\infty \\ & R \in S\end{aligned} \tag{4.8}$$

This is a convex optimization problem. Its solution is discussed in Section 6.1.2.

Chapter 5

Specific Constraint Classes

In this section we apply these results to specific constraint classes. Armed with our findings on quadratic invariance, many useful results easily follow. In Section 5.1 we consider sparsity constraints. We develop a computational test for quadratic invariance of sparsity constraints, and show that norm minimization subject to such constraints that pass the test is a convex optimization problem. We also see an interesting negative result, that perfectly decentralized control is never quadratically invariant. In Section 5.2 we show that symmetric synthesis is quadratically invariant, and thus convex. In Section 5.3 we consider the problem of control over networks with delays. We show that if the controllers can communicate faster than the dynamics propagate along any link, then norm-optimal controllers may be found via convex programming.

5.1 Sparsity Constraints

Many problems in decentralized control can be expressed in the form of problem (1.1), where S is the set of controllers that satisfy a specified sparsity constraint. In this section, we provide a computational test for quadratic invariance when the subspace S is defined by block sparsity constraints. First we introduce some notation.

Suppose $A^{\mathrm{bin}} \in \{0,1\}^{m\times n}$ is a binary matrix. We define the subspace

$$\mathrm{Sparse}(A^{\mathrm{bin}}) = \Big\{ B \in \mathcal{R}_p \mid B_{ij}(j\omega) = 0 \text{ for all } i,j \\ \text{such that } A^{\mathrm{bin}}_{ij} = 0 \text{ for almost all } \omega \in \mathbb{R} \Big\}$$

Also, if $B \in \mathcal{R}_{sp}$, let $A^{\mathrm{bin}} = \mathrm{Pattern}(B)$ be the binary matrix given by

$$A^{\mathrm{bin}}_{ij} = \begin{cases} 0 & \text{if } B_{ij}(j\omega) = 0 \text{ for almost all } \omega \in \mathbb{R} \\ 1 & \text{otherwise} \end{cases}$$

Note that in this section, we assume that matrices of transfer functions are indexed by blocks, so that above, the dimensions of A^{bin} may be much smaller than those of B. Then, K^{bin}_{kl} determines whether controller k may use measurements from subsystem l, K_{kl} is the map from the outputs of subsystem l to the inputs of subsystem k, and G_{ij} represents the map from the inputs to subsystem j to the outputs of subsystem i.

5.1.1 Computational Test

We seek an explicit test for quadratic invariance of a constraint set defined by such a binary matrix. We first prove two preliminary lemmas.

Lemma 26. *Suppose* $S = \mathrm{Sparse}(K^{bin})$, *and let* $G^{bin} = \mathrm{Pattern}(G)$. *If* S *is quadratically invariant under* G, *then*

$$K_{ki} = 0 \text{ or } K_{jl} = 0 \text{ for all } (i,j,k,l) \text{ and } K \text{ such that } K^{bin}_{kl} = 0,\ G^{bin}_{ij} = 1,\ K \in S$$

Proof. Suppose there exists (i,j,k,l) and K such that

$$K^{\mathrm{bin}}_{kl} = 0,\ G^{\mathrm{bin}}_{ij} = 1,\ K \in S$$

but

$$K_{ki} \neq 0 \text{ and } K_{jl} \neq 0$$

Then we must have

$$K^{\text{bin}}_{ki} = 1,\ K^{\text{bin}}_{jl} = 1,\ i \neq l,\ j \neq k$$

Consider $K \in S$ such that

$$K_{ab} = 0 \text{ if } (a,b) \notin \{(k,i),(j,l)\}$$

Then

$$(KGK)_{kl} \;=\; \sum_r \sum_s K_{kr} G_{rs} K_{sl} \;=\; K_{ki} G_{ij} K_{jl}$$

Since $G_{ij} \neq 0$, we can easily choose K_{ki} and K_{jl} such that $(KGK)_{kl} \neq 0$. So $KGK \notin S$ and S is not quadratically invariant.

∎

Lemma 27. *Suppose* $S = \text{Sparse}(K^{bin})$*, and let* $G^{bin} = \text{Pattern}(G)$*. If*

$$K_{ki} = 0 \textit{ or } K_{jl} = 0 \textit{ for all } (i,j,k,l) \textit{ and } K \textit{ such that } K^{bin}_{kl} = 0,\ G^{bin}_{ij} = 1,\ K \in S$$

Then

$$K^{bin}_{ki} K^{bin}_{jl} = 0 \textit{ for all } (i,j,k,l) \textit{ such that } K^{bin}_{kl} = 0,\ G^{bin}_{ij} = 1$$

Proof. We show this by contradiction. Suppose there exists (i,j,k,l) such that

$$K^{\text{bin}}_{kl} = 0, \quad G^{\text{bin}}_{ij} = 1, \quad K^{\text{bin}}_{ki} K^{\text{bin}}_{jl} \neq 0.$$

Then

$$K^{\text{bin}}_{ki} \;=\; K^{\text{bin}}_{jl} \;=\; 1$$

and hence it must follow that there exists $K \in S$ such that $K_{ki} \neq 0$ and $K_{jl} \neq 0$. ∎

The following is the main result of this section. It provides a computational test for quadratic invariance when S is defined by sparsity constraints. It also equates quadratic invariance with a stronger condition.

Theorem 28. *Suppose* $S = \text{Sparse}(K^{bin})$*, and let* $G^{bin} = \text{Pattern}(G)$*. Then the following are equivalent:*

(i) S *is quadratically invariant under* G

(ii) $KGJ \in S$ *for all* $K, J \in S$

(iii) $K_{ki}^{bin}\ G_{ij}^{bin}\ K_{jl}^{bin}\ (1 - K_{kl}^{bin}) = 0$ *for all* $i, l = 1, \ldots, n_y$ *and* $j, k = 1, \ldots, n_u$

Proof. We will show that $(i) \implies (iii) \implies (ii) \implies (i)$. Suppose S is quadratically invariant under G. Then by Lemma 26,

$$K_{ki} = 0 \text{ or } K_{jl} = 0 \text{ for all } (i,j,k,l) \text{ and } K \text{ such that } K_{kl}^{\text{bin}} = 0;\ G_{ij}^{\text{bin}} = 1;\ K \in S$$

and by Lemma 27,

$$K_{ki}^{\text{bin}} K_{jl}^{\text{bin}} = 0 \text{ for all } (i,j,k,l) \text{ such that } K_{kl}^{\text{bin}} = 0,\ G_{ij}^{\text{bin}} = 1$$

which can be restated

$$K_{ki}^{\text{bin}}\ G_{ij}^{\text{bin}}\ K_{jl}^{\text{bin}}\ (1 - K_{kl}^{\text{bin}}) \ = \ 0$$

and which implies that

$$K_{ki} = 0 \text{ or } J_{jl} = 0 \text{ for all } (i,j,k,l),\ K,\ J \text{ such that } K_{kl}^{\text{bin}} = 0;\ G_{ij}^{\text{bin}} = 1;\ K, J \in S$$

which clearly implies

$$(KGJ)_{kl} \ = \ \sum_i \sum_j K_{ki} G_{ij} J_{jl} \ = \ 0 \text{ for all } (k,l), K, J \text{ such that } K_{kl}^{\text{bin}} = 0; K, J \in S$$

and thus

$$KGJ \in S \text{ for all } K, J \in S$$

which is a stronger condition than quadratic invariance and hence implies (i). ∎

This result shows us several things about sparsity constraints. In this case quadratic invariance is equivalent to another condition which is stronger in general. When G is symmetric, for example, the subspace consisting of symmetric K is quadratically invariant but does not satisfy condition (ii). Condition (iii), which gives us the computational test we desired, shows that quadratic invariance can be checked in time $O(n^4)$, where $n = \max\{n_u, n_y\}$. It also shows that, if S is defined by sparsity constraints, then S is quadratically invariant under G if and only if it is quadratically invariant under all systems with the same sparsity pattern.

5.1.2 Sparse Synthesis

The following theorem shows that for sparsity constraints, the test in Section 5.1 can be used to identify tractable decentralized control problems.

Theorem 29. *Suppose $G \in \mathcal{R}_{sp}$ and $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty \cap S$. Further suppose $G^{\text{bin}} = \text{Pattern}(G)$ and $S = \text{Sparse}(K^{\text{bin}})$ for some $K^{\text{bin}} \in \{0,1\}^{n_u \times n_y}$. If*

$$K^{bin}_{ki}\ G^{bin}_{ij}\ K^{bin}_{jl}\ (1 - K^{bin}_{kl}) = 0 \quad \textit{for all } i,l = 1,\dots,n_y \textit{ and } j,k = 1,\dots,n_u$$

then

$$S = \Big\{ K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \mid Q \in S \Big\}$$

Proof. Follows immediately from Theorems 23 and 28. ∎

5.1.3 Example Sparsity Patterns

We apply this test to examine a few specific sparsity structures.

Skyline structure

Consider matrices for which any non-zero entry must have non-zero entries below it, and more formally, define a matrix $A \in \mathbb{C}^{m \times n}$ to be a ***skyline matrix*** if for all $i = 2,\dots,m$ and all $j = 1,\dots,n$,

$$A_{i-1,j} = 0 \quad \text{if} \quad A_{i,j} = 0$$

An example is

$$K^{\text{bin}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Suppose G is lower triangular and K^{bin} is a lower triangular skyline matrix. Then $S = \text{Sparse}(K^{\text{bin}})$ is quadratically invariant under G. Some numerical examples with these structures are worked out in Section 6.2.

Plant and controller with the same structure

It is important to notice that G and S having the same sparsity structure does not imply that S is quadratically invariant under G. For example, consider

$$G = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

and let $S = \text{Sparse}(G)$. Then S is not quadratically invariant, as $G^3 \notin S$.

Perfectly Decentralized Control

We now show an interesting negative result. Let $n_u = n_y$, so that each subsystem has its own controller as in Figure 1.1.

Corollary 30. *Suppose there exists i, j, with $i \neq j$, such that $G_{ij} \neq 0$. Suppose K^{bin} is diagonal and $S = \text{Sparse}(K^{\text{bin}})$. Then S is not quadratically invariant under G.*

Proof. Let $G^{\text{bin}} = \text{Pattern}(G)$. Then

$$K^{\text{bin}}_{ii}\ G^{\text{bin}}_{ij}\ K^{\text{bin}}_{jj}\ (1 - K^{\text{bin}}_{ij}) = 1$$

The result then follows from Theorem 28. ■

It is important to note that the plant and controller do not have to be square to apply this result because of the block notation used in this section. This corollary tells us that if each subsystem has its own controller which may only use sensor information from its own subsystem, and any subsystem affects any other, then the system *is not* quadratically invariant. In other words, perfectly decentralized control is never quadratically invariant except for the trivial case where no subsystem affects any other.

5.2 Symmetric Constraints

The following shows that when the plant is symmetric, the methods introduced in this thesis could be used to find the optimal symmetric stabilizing controller. Symmetric synthesis is a well-studied problem, and there are many techniques that exploit its structure. Therefore, the methods in this thesis are possibly not the most efficient. However, it is important to note the quadratic invariance of this structure because it defied earlier attempts to classify solvable problems. This arises because a symmetric matrix multiplied by another, i.e. KG, is not guaranteed to yield a symmetric matrix, but a symmetric matrix left and right multiplied by the same symmetric matrix, i.e. KGK, will indeed.

Theorem 31. *Suppose*

$$\mathbb{H}^n = \left\{ A \in \mathbb{C}^{n\times n} \;\middle|\; A = A^* \right\}$$

and

$$S = \{K \in \mathcal{R}_p \mid K(j\omega) \in \mathbb{H}^n \text{ for almost all } \omega \in \mathbb{R}\}.$$

Further suppose $K_{\text{nom}} \in C_{\text{stab}} \cap \mathcal{RH}_\infty \cap S$ *and* $G \in \mathcal{R}_{sp}$ *with* $G(j\omega) \in \mathbb{H}^n$ *for almost all* $\omega \in \mathbb{R}$. *Then*

$$S = \Big\{ K_{\text{nom}} - h\big(h(K_{\text{nom}}, G), Q\big) \mid Q \in S \Big\}$$

Proof. Follows immediately from Theorem 23. ∎

5.3 Control over Networks with Delays

In this section we consider the problem of multiple subsystems, each with its own controller, such that the dynamics of each subsystem may effect those of other subsystems with some propagation delay, and the controllers may communicate with each other with some transmission delays. We again seek to synthesize linear controllers to minimize a closed-loop norm for the entire interconnected system. There is no known tractable solution for arbitrary propagation and transmission delays. This section uses the results from the previous chapters to find simple conditions on the delays such that this optimal control problem may be cast as a convex optimization problem.

We find that if the transmission delays satisfy the triangle inequality, and if the propagation delay between any pair of subsystems is at least as large as the transmission delay between those subsystems, then the problem is quadratically invariant. In other words, if data can be transmitted faster than dynamics propagate along any link, then optimal controller synthesis may be formulated as a convex optimization problem.

It is important to note the extreme generality of this framework and of this result. It holds for discrete-time systems and continuous-time systems. It holds for any norm that we wish to minimize. It does not assume that the dynamics of any subsystem are the same as those of any other, and they may all be completely different types of objects. Most importantly, the delay between any two subsystems is not assumed to have any relationship whatsoever to other delays in the system. They may be assigned independently for each link. Only in the final examples do we assume otherwise.

Prior Work

A vast amount of prior work on optimal control over networks assumes that the actions of any subsystem have no effect on the dynamics of other subsystems. For a few other specific structures, some of which are mentioned in Section 1.1, tractable methods have been found. One of the first problems of this nature to be studied was the one-step delayed information sharing problem. This problem assumes that each subsystem has a controller which can see its own output immediately, and can see outputs from all other subsystems after a delay of one time step. This problem has long been known to admit tractable solutions [31], and has also been studied more recently in an LFT framework [26]. An interesting class of spatio-temporal systems that allow for convex synthesis of optimal controllers was identified in [2], and named funnel causal systems. One of the tractable structures discussed in [16] involved evenly spaced subsystems which can pass measurements on at the same speed that the dynamics propagate, and [20] included a similar class of evenly spaced systems where the bound was found such that if the communication speed exceeded that bound the problem was amenable to convex synthesis.

These results are all unified and generalized by the simple conditions found in this section.

Section Outline

In Section 5.3.1, we develop some notation, define the propagation and transmission delays, explain why we may assume that the transmission delays satisfy the triangle inequality, and formulate the problem we wish to solve.

Section 5.3.2 contains the main result of the section, where we prove that if this triangle inequality is satisfied, and if the propagation delay associated with any pair of subsystems is at least as large as the associated transmission delay, then the information constraint is quadratically invariant, and thus, optimal control may be cast as a convex optimization problem.

In Section 5.3.3 we break these total transmission delays out into a pure transmission delay, representing the time it takes to communicate the information from one subsystem to another, and a computational delay, representing the time it takes to process the information before it is used by the controller. We find, somewhat surprisingly, that transmitting faster than the propagation of dynamics still guarantees convexity, and in fact, that the computational delay causes the condition to be relaxed.

We then consider a few more specific examples in Section 5.3.4. First is an example corresponding to a very general problem of the control of vehicles in formation. The vehicles may have arbitrary positions, their dynamics propagate at a constant speed, and they communicate their measurements at a constant speed. The optimal control problem is amenable to convex synthesis as long as the communication speed exceeds the propagation speed. Even though this itself is a broad generalization of previously identified tractable classes, it follows almost immediately from the results of this section. Conditions are then derived for convexity of optimal control over a lattice, for two different types of assumptions on the propagation of dynamics.

Finally in Section 5.3.5, we show how all of these results can be generalized to the case where the dynamics of each subsystem may effect those of other subsystems either with some propagation delay, or not at all, and any controller may communicate with any other either with some transmission delay, or not at all. In other words, delay constraints and sparsity constraints are combined.

5.3.1 Delays

We define Delay$(\cdot)$ to give the delay associated with a time-invariant causal operator

$$\text{Delay}(W) = \arg \inf_{\tau > 0} w(\tau) \neq 0 \quad \text{where } w \text{ is the impulse response of } W$$

Note that we then have the following inequalities for the delays of a composition or an addition of operators:

$$\begin{aligned} \text{Delay}(AB) &\geq \text{Delay}(A) + \text{Delay}(B) \\ \text{Delay}(A+B) &\geq \min\{\text{Delay}(A), \text{Delay}(B)\} \end{aligned}$$

Propagation Delays

Suppose there are n subsystems. For any pair of subsystems i and j we define the propagation delay p_{ij} as the amount of time before a controller action at subsystem j can affect an output at subsystem i as such

$$p_{ij} = \text{Delay}(G_{ij}) \qquad \text{for all } i, j \in 1, \ldots, n$$

Transmission Delays

For any pair of subsystems k and l we define the (total) transmission delay t_{kl} as the minimum amount of time before the controller of subsystem k may use outputs from subsystem l. Given these constraints, we can define the overall subspace of admissible controllers S such that $K \in S$ if and only if

$$\text{Delay}(K_{kl}) \;\geq\; t_{kl} \qquad \text{for all } k, l \in 1, \ldots, n$$

In Section 5.3.3 we will break these total transmission delays out into a pure transmission delay, representing the time it takes to communicate the information from one subsystem to another, and a computational delay, representing the time it takes to process the information before it is used by the controller.

Triangle inequality

For the main result of this section, we will assume that the triangle inequality holds amongst the transmission delays, that is,

$$t_{ki} + t_{ij} \;\geq\; t_{kj} \qquad \text{for all } k, i, j$$

This is typically a very reasonable assumption for the following reasons. t_{kj} is defined as the minimum amount of time before controller k can use outputs from subsystem j. So if there existed an i such that the inequality above failed, that would mean that controller k could receive that information more quickly if it came indirectly via controller i. We would thus reroute this information to go through i, t_{kj} would be reset to $t_{ki} + t_{ij}$, and the inequality would hold.

To put it another way, we could think of each subsystem as a node on a directed graph, with the initial distance from any node j to any node k as t_{kj}, the time it takes before controller k can directly use outputs from subsystem j. We then want to find the shortest overall time for any controller k to use outputs from any subsystem j, that is, the shortest path from node j to node k. So to find our final t_{kj}'s, we run Bellman-Ford or another shortest path algorithm on our initial graph [14], and the resulting delays are thus guaranteed to satisfy the triangle inequality.

5.3.2 Conditions for Convexity

Given a generalized plant P and transmission delays t_{kl} for each pair of subsystems, we define S as above, and then seek to solve problem (1.1). The delays associated with dynamics propagating from one subsystem to another are embedded in P. The subspace of admissible controllers, S, has been defined to encapsulate the constraints on how quickly information may be passed from one subsystem to another.

We first provide a necessary and sufficient condition for quadratic invariance in terms of these delays, which is derived fairly directly from our definitions.

Theorem 32. *Suppose that G and S are defined as above. S is quadratically invariant under G if and only if*

$$t_{ki} + p_{ij} + t_{jl} \;\geq\; t_{kl} \qquad \textit{for all } i, j, k, l \tag{5.1}$$

Proof. Given $K \in S$,

$$KGK \in S \iff \text{Delay}\big((KGK)_{kl}\big) \geq t_{kl} \text{ for all } k, l$$

We now seek conditions that cause this to hold.

$$(KGK)_{kl} = \sum_i \sum_j K_{ki} G_{ij} K_{jl}$$

and so for any k and l,

$$\begin{aligned} &\text{Delay}\big((KGK)_{kl}\big) \\ &\qquad \geq \min_{i,j}\{\text{Delay}(K_{ki}G_{ij}K_{jl})\} \\ &\qquad \geq \min_{i,j}\{\text{Delay}(K_{ki}) + \text{Delay}(G_{ij}) + \text{Delay}(K_{jl})\} \\ &\qquad \geq \min_{i,j}\{t_{ki} + p_{ij} + t_{jl}\} \end{aligned}$$

Thus S is quadratically invariant under G if

$$\min_{i,j}\{t_{ki} + p_{ij} + t_{jl}\} \;\geq\; t_{kl} \qquad \text{for all } k, l$$

which is equivalent to

$$t_{ki} + p_{ij} + t_{jl} \;\geq\; t_{kl} \qquad \text{for all } i, j, k, l$$

Now suppose that Condition (5.1) fails. Then there exists i, j, k, l such that

$$t_{ki} + p_{ij} + t_{jl} \;<\; t_{kl}$$

Consider K such that

$$K_{ab} = 0 \text{ if } (a, b) \notin \{(k, i), (j, l)\}$$

Then

$$(KGK)_{kl} \;=\; \sum_r \sum_s K_{kr} G_{rs} K_{sl} \;=\; K_{ki} G_{ij} K_{jl}$$

Since $\text{Delay}(G_{ij}) = p_{ij}$, we can easily choose K_{ki} and K_{jl} such that $\text{Delay}(K_{ki}) = t_{ki}$, $\text{Delay}(K_{jl}) = t_{jl}$, and

$$\text{Delay}\big((KGK)_{kl}\big) = t_{ki} + p_{ij} + t_{jl}$$

So $K \in S$ but $KGK \notin S$ and thus S is not quadratically invariant under G. ■

Main Result - networks. The following is the main result of this section. It states that if the transmission delays satisfy the triangle inequality, and if the propagation delay between any pair of subsystems is at least as large as the transmission delay between those subsystems, then the information constraint is quadratically invariant. In other words, if data can be transmitted faster than dynamics propagate along any link, then optimal controller synthesis may be cast as a convex optimization problem.

Theorem 33. *Suppose that G and S are defined as above, and that the transmission* *delays satisfy the triangle inequality. If*

$$p_{ij} \ \geq \ t_{ij} \qquad \textit{for all } i, j \tag{5.2}$$

then S is quadratically invariant under G.

Proof. Suppose Condition (5.2) holds. Then for all i, j, k, l we have

$$\begin{aligned} t_{ki} + p_{ij} + t_{jl} &\geq t_{ki} + t_{ij} + t_{jl} \\ &\geq t_{kl} \qquad \text{by the triangle inequality} \end{aligned}$$

and thus by Theorem 32, S is quadratically invariant under G. ■

Thus we have shown that the triangle inequality and Condition (5.2) are sufficient for quadratic invariance. The following remarks discuss assumptions under which they are necessary as well.

Remark 34. *If we assume that $t_{ii} = 0$ for all i, that is, that there is no delay before a subsystem's controller may use its own outputs, then we consider Condition* (5.1) *with $k = i, \ l = j$ and see that Condition* (5.2) *is necessary for quadratic invariance.*

Remark 35. *If we assume that $p_{ii} = 0$ for all i, that is, that there is no delay before a subsystem's controller actions affect its own dynamics, then we consider Condition* (5.1) *with $i = j$ and see that the triangle inequality is necessary for quadratic invariance.*

5.3.3 Computational Delays

In this section, we consider what happens when the controller of each subsystem has a computational delay c_i associated with it. The delay for controller i to use outputs from subsystem j, the total transmission delay, is then broken up into a pure transmission delay and this computational delay, as follows

$$t_{ij} = c_i + \tilde{t}_{ij}$$

If we were to assume that the triangle inequality held for the total transmission delays t_{ij} as before, then we would simply get the same results as in the previous section with the substitution above. In particular, we would find $p_{ij} \geq c_i + \tilde{t}_{ij}$ to be the condition for quadratic invariance. However, there are many cases where it makes sense to instead assume that the triangle inequality holds for the pure transmission delays $\tilde{t}_{ij}$, which is a stronger assumption. An example where such is clearly the case is provided in Section 5.3.4.

In this section we derive conditions for quadratic invariance when we can assume that the triangle inequality holds for the pure transmission delays $\tilde{t}_{ij}$, and get a surprising result.

As before, the propagation delays are defined as

$$p_{ij} \;=\; \text{Delay}(G_{ij}) \qquad \text{for all } i, j$$

and S is now defined such that $K \in S$ if and only if

$$\text{Delay}(K_{kl}) \;\geq\; c_k + \tilde{t}_{kl} \qquad \text{for all } k, l$$

Thus the necessary and sufficient condition for quadratic invariance from Theorem 32 becomes

$$c_k + \tilde{t}_{ki} + p_{ij} + c_j + \tilde{t}_{jl} \;\geq\; c_k + \tilde{t}_{kl} \qquad \text{for all } i, j, k, l$$

which reduces to

$$\tilde{t}_{ki} + p_{ij} + c_j + \tilde{t}_{jl} \;\geq\; \tilde{t}_{kl} \qquad \text{for all } i, j, k, l \tag{5.3}$$

The following theorem gives conditions under which the information constraint is quadratically invariant. It states that if the triangle inequality holds amongst the pure transmission delays, and if Condition (5.4) holds, then the information constraint is quadratically invariant. Surprisingly, we see that the computational delay now appears on the left side of the inequality. In other words, not only does transmitting data faster than dynamics propagate still allow for convex synthesis when we account for computational delay, but the condition is actually relaxed.

Theorem 36. *Suppose that G and S are defined as above, and that the pure transmission delays satisfy the triangle inequality. If*

$$p_{ij} + c_j \;\geq\; \tilde{t}_{ij} \qquad \text{for all } i, j \tag{5.4}$$

then S is quadratically invariant under G.

Proof. Suppose Condition (5.4) holds. Then for all i, j, k, l we have

$$\begin{aligned} \tilde{t}_{ki} + p_{ij} + c_j + \tilde{t}_{jl} \;&\geq\; \tilde{t}_{ki} + \tilde{t}_{ij} + \tilde{t}_{jl} \\ &\geq\; \tilde{t}_{kl} \qquad \text{by the triangle inequality} \end{aligned}$$

and thus Condition (5.3) holds and S is quadratically invariant under G. ∎

Thus we have shown that the triangle inequality and Condition (5.4) are sufficient for quadratic invariance. The following remark discusses an assumption under which the condition is necessary as well.

Remark 37. *If we assume that $\tilde{t}_{ii} = 0$ for all i, that is, that there is no additional delay before a subsystem's controller may use its own outputs, other than the computational delay, then we consider Condition (5.3) with $k = i$, $l = j$ and see that Condition (5.4) is necessary for quadratic invariance. Since the computational delay has been extracted, this is now a very reasonable assumption which is essentially true by definition.*

5.3.4 Network Examples

We consider here some special cases.

Vehicle Formation Example

We now consider an important special case, which corresponds to the problem of controlling multiple vehicles in a formation.

Suppose there are n subsystems (vehicles), with positions $x_1, \dots, x_n \in \mathbb{R}^d$. Typically, we'll have $d = 3$, but these results hold for arbitrary d.

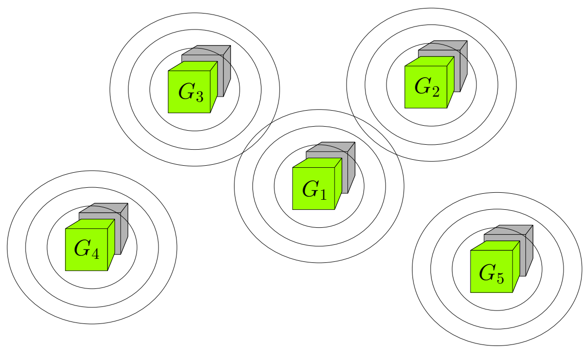

Figure 5.1: Communication and propagation in all directions

Let R represent the maximum distance between any two subsystems

$$R = \max_{i,j} \|x_i - x_j\|$$

For most applications of interest the appropriate norm throughout this section would be the Euclidean norm, but these results hold for arbitrary norm on $\mathbb{R}^d$.

We suppose that dynamics of all vehicles propagate at a constant speed, determined by the medium, such that the propagation delays are proportional to the distance between vehicles, as illustrated in Figure 5.1.

Let γ_p be the amount of time it takes dynamics to propagate one unit of distance, i.e., the inverse of the speed of propagation. For example, when considering formations of aerial vehicles, γ_p would equal the inverse of the speed of sound.

The system G is then such that

$$\text{Delay}(G_{ij}) \;=\; \gamma_p \|x_i - x_j\| \qquad \text{for all } i, j$$

We similarly suppose that data can be transmitted at a constant speed, such that the transmission delays are proportional to the distances between vehicles, such as if each vehicle could broadcast its information to the others. This is also illustrated in Figure 5.1.

Let γ_t be the amount of time it takes to transmit one unit of distance, i.e., the inverse of the speed of transmission. Let C be the computational delay at each vehicle. The set of admissible controllers is then defined such that $K \in S$ if and only if

$$\text{Delay}(K_{kl}) \;\geq\; C + \gamma_t \|x_k - x_l\| \qquad \text{for all } k, l$$

We can now apply Theorem 36 with

$$p_{ij} = \gamma_p \|x_i - x_j\|, \quad \tilde{t}_{ij} = \gamma_t \|x_i - x_j\|, \quad \text{and } c_i = C \quad \text{for all } i \text{ and } j$$

Clearly, $\tilde{t}_{ii} = 0$ for all i as in Remark 37, so the conditions of Theorem 36 are both necessary and sufficient for quadratic invariance.

Theorem 38. *Suppose that G and S are defined as above. S is quadratically invariant under G if and only if*

$$\gamma_p + \frac{C}{R} \;\geq\; \gamma_t$$

Proof. Since any norm satisfies the triangle inequality, the pure transmission delays clearly satisfy the triangle inequality, so applying Theorem 36, S is quadratically invariant under G if and only if

$$\gamma_p \|x_i - x_j\| + C \;\geq\; \gamma_t \|x_i - x_j\| \qquad \text{for all } i, j$$

which is equivalent to

$$\gamma_p + \frac{C}{R} \;\geq\; \gamma_t$$

■

Thus we see that, in the absence of computational delay, finding the minimum-norm controller may be reduced to a convex optimization problem when the speed of transmission is faster than the speed of propagation; that is, when $\gamma_p \geq \gamma_t$. We also see that this not only remains true in the presence of computational delay, but that we get a buffer relaxing the condition.

A similar result was previously achieved for a very specific case of vehicles equally spaced along a line [22]. This shows how the results of this section allow us to effortlessly generalize to the case considered in this subsection, where the vehicles have arbitrary positions in arbitrary dimensions. This is a crucial generalization for applications to realistic formation flight problems.

Two-Dimensional Lattice Example

In this subsection we will consider subsystems distributed in a lattice, and use these results to derive the conditions for convexity of the associated optimal decentralized control problem.

We first consider the case where the controllers can communicate along the edges of the lattice with a delay of t, and the dynamics similarly propagate along the edges with a delay of p, as illustrated in Figure 5.2.

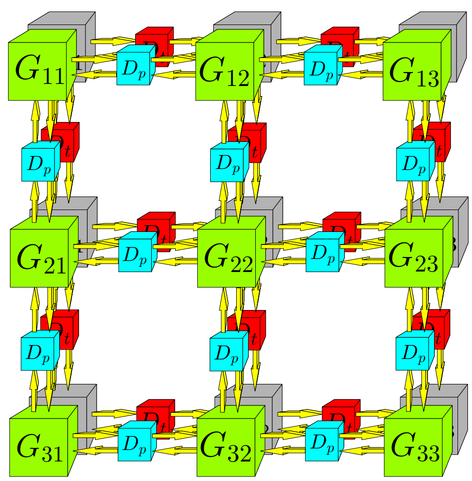

Figure 5.2: Two-dimensional lattice with dynamics propagating along edges

It is a straightforward consequence of this section that the optimal controllers may be synthesized with convex programming if

$$p \ \geq \ t$$

We now consider a more interesting variant, where the controllers again communicate only along the edges of the lattice, but now the dynamics propagate in all directions, as illustrated in Figure 5.3.

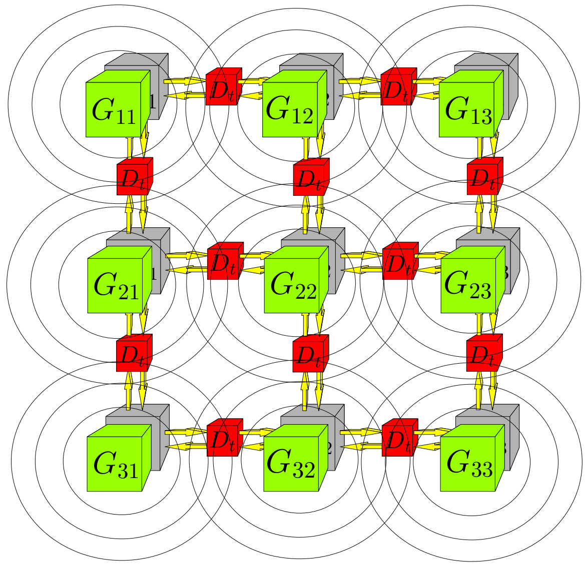

Figure 5.3: Two-dimensional lattice with dynamics propagating in all directions

Let γ_p be the amount of time it takes for the dynamics to propagate one unit of distance. Along a diagonal, for instance, between G_{11} and G_{22}, the propagation delay is $\gamma_p\sqrt{2}$ and the transmission delay is $2t$. The condition for convexity therefore becomes

$$\gamma_p \;\geq\; t\sqrt{2}$$

5.3.5 Combining Sparsity and Delay Constraints

In this section, we will discuss how sparsity constraints may be considered a special case of the framework analyzed in this section. We show how the two can be combined to handle the very general, realistic case of a network where some nodes are connected with delays as above and others are not connected at all. An explicit test for quadratic invariance in this case will be provided.

The key observation is that a sparsity constraint may be considered an infinite delay. We thus define an extended notion of propagation and transmission delays, where they are assigned to be sufficiently large when they do not exist, and then the results from the rest of this section may be applied to test for quadratic invariance and convexity.

Propagation Delays

We now consider a plant for which the controllers of certain subsystems may or may not have any effect on other subsystems, and when they do, there may be a propagation delay associated with that effect. First, let

$$G^{\text{bin}} \;=\; \text{Pattern}(G)$$

as in Section 5.1, so that $G^{\text{bin}}_{ij} = 0$ if subsystem i is not affected by inputs to subsystem j. We would then like to define the propagation delay p_{ij} to be extremely large if this is the case, as such

$$p_{ij} \;=\; \begin{cases} \text{Delay}(G_{ij}) & \text{if } \; G^{\text{bin}}_{ij} = 1 \\ H & \text{if } \; G^{\text{bin}}_{ij} = 0 \end{cases}$$

for some large H.

Transmission Delays

As in Section 5.1, we first assign a binary matrix K^{bin} such that $K^{\text{bin}}_{kl} = 0$ if controller k may never use outputs from subsystem l. For any other pair of subsystems k and l we define the (total) transmission delay t_{kl} as in the rest of this section; that is, as the minimum amount of time before the controller of subsystem k may use outputs from subsystem l. Given these constraints, we can define the overall subspace of admissible controllers S such that $K \in S$ if and only if

$$\begin{aligned} \text{Delay}(K_{kl}) \;&\geq\; t_{kl} & &\text{for all } k,l \quad \text{such that } K^{\text{bin}}_{kl} = 1 \\ K_{kl} \;&=\; 0 & &\text{for all } k,l \quad \text{such that } K^{\text{bin}}_{kl} = 0 \end{aligned}$$

We wish to assign a very large transmission delay to the latter case, and so define

$$t_{kl} = H \qquad \text{for all } k,l \quad \text{such that } K^{\text{bin}}_{kl} = 0$$

for the same large H as above.

Condition for Convexity

Given these extended definitions of propagation delays and transmission delays for a combination of sparsity and delay constraints, we can now test for quadratic invariance using Theorem 32.

These definitions of extended delays along with our definition of the constraint set S allow us to use this and the rest of the results of this section as long as H has been chosen large enough. Condition (5.1) is indeed necessary and sufficient for quadratic invariance as long as

$$H \;>\; 2\,\max\{t_{kl}\} + \max\{p_{ij}\}$$

where of course the first maximum is taken over all k, l such that $K^{\text{bin}}_{kl} = 1$ and the second is taken over all i, j such that $G^{\text{bin}}_{ij} = 1$. This arises because Condition (5.1) must fail if $K^{\text{bin}}_{kl} = 0$, but $K^{\text{bin}}_{ki} = G^{\text{bin}}_{ij} = K^{\text{bin}}_{jl} = 1$.

Chapter 6

Computation of Optimal Controllers

This chapter considers the computation of solutions to the convex optimization problems that have been identified and formulated throughout this thesis. Solution procedures are given in Section 6.1 and numerical examples are provided in Section 6.2.

6.1 Solution Procedure

We show in this section that if we wish to minimize the $\mathcal{H}_2$-norm, one further change of variables can be used in which the information constraint is eliminated.

In Section 6.1.1 problem (4.6) is then converted to an unconstrained problem which may be readily solved. We focus on sparsity constraints, as in [21], but the vectorization techniques in this section are easily applied to the other constraint classes of Chapter 5 as well. A similar method was used for symmetric constraints in [32].

In Section 6.1.2 we address the convex but more complex problem that was derived in Section 4.5 for when a nominal stabilizing controller is not available. A solution procedure is provided, and the implications for systematically finding stabilizing decentralized controllers are discussed.

6.1.1 Removal of Information Constraint

For ease of presentation, we now make a slight change of notation from Section 5.1. We no longer assume that the plant and controller are divided into blocks, so that K^{bin}_{kl} now determines whether the kl index of the controller may be non-zero, rather than determining whether controller k may use information from subsystem l, and G_{ij} similarly represents the ij index of the plant. K^{bin} therefore has the same dimension as the controller itself. n_u and n_y represent the total number of inputs and outputs, respectively.

Let

$$a = \sum_{i=1}^{n_u} \sum_{j=1}^{n_y} K^{\text{bin}}_{ij}$$

such that a represents the number of admissible controls, that is, the number of indices for which K is not constrained to be zero.

The following theorem gives the equivalent unconstrained problem.

Theorem 39. *Suppose x is an optimal solution to*

$$\begin{aligned} \textit{minimize} \quad & \|b + Ax\|_2 \\ \textit{subject to} \quad & x \in \mathcal{RH}_\infty \end{aligned} \tag{6.1}$$

where $D \in \mathbb{R}^{n_u n_y \times a}$ is a matrix whose columns form an orthonormal basis for $\text{vec}(S)$, *and*

$$b = \text{vec}(T_1), \qquad A = -(T_3^T \otimes T_2)D.$$

Then $Q = \text{vec}^{-1}(Dx)$ is optimal for (4.6) *and the optimal values are equivalent.*

Proof. We know that

$$Q \in \mathcal{RH}_\infty^{n_u \times n_y} \cap S \iff \text{vec}(Q) = Dx \quad \text{for some } x \in \mathcal{RH}_\infty^{a \times 1}$$

Since

$$\begin{aligned}
&\|T_1 - T_2 Q T_3\|_2 \\
&\quad = \|\mathrm{vec}(T_1 - T_2 Q T_3)\|_2 \quad \text{by definition of the } \mathcal{H}_2\text{-norm} \\
&\quad = \|\mathrm{vec}(T_1) - (T_3^{\mathrm{T}} \otimes T_2)\,\mathrm{vec}(Q)\|_2 \quad \text{by Lemma 1} \\
&\quad = \|\mathrm{vec}(T_1) - (T_3^{\mathrm{T}} \otimes T_2) D x\|_2 \\
&\quad = \|b + Ax\|_2
\end{aligned}$$

we have the desired result. ■

Therefore, we can find the optimal x for problem (6.1) using many available tools for unconstrained $\mathcal{H}_2$-synthesis, with

$$P_{11} = b \qquad P_{12} = A \qquad P_{21} = 1 \qquad P_{22} = 0^{1\times a}$$

then find the optimal Q for problem (4.6) as $Q = \mathrm{vec}^{-1}(Dx)$, and finally, find the optimal K for problem (4.1) as $K = K_{\mathrm{nom}} - h\big(h(K_{\mathrm{nom}}, G), Q\big)$.

6.1.2 Solution without Strong Stabilizability

We show in this section that vectorization can similarly be used to eliminate the information constraint when a nominal stable and stabilizing controller can not be found, as in Section 4.5. The resulting problem is not immediately amenable to standard software, as in the previous section, but methods for obtaining its solution are discussed.

Let $D \in \mathbb{R}^{n_u n_y \times a}$ be a matrix whose columns form an orthonormal basis for $\mathrm{vec}(S)$, as in the previous section, and now let

$$f = \mathrm{vec}(P_{11}), \qquad\qquad E = -(P_{21}^{\mathrm{T}} \otimes P_{12}) D,$$

$$d = \begin{bmatrix} 0 \\ \mathrm{vec}(G) \\ 0 \end{bmatrix}, \qquad C = \begin{bmatrix} (G^T \otimes I)D \\ -(G^T \otimes G)D \\ (I \otimes G)D \end{bmatrix}$$

Using similar arguments as in the previous section, we find that we may solve the following equivalent problem. Suppose x is an optimal solution to

$$\begin{aligned} \text{minimize} \quad & \|f + Ex\|_2 \\ \text{subject to} \quad & d + Cx \in \mathcal{RH}_\infty \\ & x \in \mathcal{RH}_\infty \end{aligned} \tag{6.2}$$

Then $R = \mathrm{vec}^{-1}(Dx)$ is optimal for (4.8) and the optimal values are equivalent. The optimal K for problem (4.1) could then be recovered as $K = h_G(R)$.

Remark 40. *While A, b of problem* (6.1) *are stable, C, d, E, f of problem* (6.2) *may very well be unstable. Notice also that $G \in \mathcal{R}_{sp}$ implies $C, d \in \mathcal{R}_{sp}$.*

Remark 41. *The last constraint comes from the upper right-hand block of Condition* (4.7)*, and the others come from the rest of that condition.*

Remark 42. *The relaxed problem*

$$\begin{aligned} \textit{minimize} \quad & \|f + Ex\|_2 \\ \textit{subject to} \quad & x \in \mathcal{RH}_\infty \end{aligned} \tag{6.3}$$

can be solved with standard software as described in the previous subsection, and gives a lower bound on the solution. If the result is such that the entire constraint of problem (6.2) *is satisfied, then the optimal value has been achieved.*

Remark 43. *For any $\mu > 0$ the following problem may be solved in the same standard* *manner*

$$\begin{aligned} \textit{minimize} \quad & \left\| \begin{bmatrix} f \\ \mu d \end{bmatrix} + \begin{bmatrix} E \\ \mu C \end{bmatrix} x \right\|_2 \\ \textit{subject to} \quad & x \in \mathcal{RH}_\infty \end{aligned} \tag{6.4}$$

and then the optimal value of x *as well as the optimal value of the objective function will approach those of problem* (6.2) *as* μ *approaches* 0 *from above.*

A reasonable solution procedure for problem (6.2) would then be to first solve the relaxed problem of Remark 42, and test whether $d+Cx \in \mathcal{RH}_\infty$ for the optimal value. If so, we are done and can recover the optimal K. If not, then solve problem (6.4) for values of μ which decrease and approach 0. This procedure in no way requires a controller that is both stable and stabilizing, so it is most useful when the plant is actually not strongly stabilizable, and thus no such controller exists.

Alternatively, as long as P is stabilizable by some $K \in S$, the solution to problem (6.4) for any $\mu > 0$ results in an x such that $\|d + Cx\|_2$ is finite. Thus $R = \mathrm{vec}^{-1}(Dx)$ satisfies Condition (4.7), and $K = h_G(R)$ is both stabilizing and lies in S. If it is also stable, we have then found a $K_{\mathrm{nom}} \in C_{\mathrm{stab}} \cap \mathcal{RH}_\infty \cap S$, and the procedures from the rest of this thesis may be used to find the optimal decentralized controller. This is ideal for the case where the plant is strongly stabilizable, but a stabilizing controller is difficult to find with other methods.

The techniques discussed here and in Section 4.5 involve not only finding optimal decentralized controllers, but also develop explicit procedures for first finding a stabilizing decentralized controller when one is not available otherwise. As there are no known systematic methods of finding stabilizing controllers for most quadratically invariant problems, this is an extremely important development, and an exciting avenue for future research.

6.2 Numerical Examples

We apply our results to some specific numerical examples, first for sparsity constraints, and then for delay constraints.

6.2.1 Sparsity Examples

Consider an unstable lower triangular plant

$$G(s) = \begin{bmatrix} \frac{1}{s+1} & 0 & 0 & 0 & 0 \\ \frac{1}{s+1} & \frac{1}{s-1} & 0 & 0 & 0 \\ \frac{1}{s+1} & \frac{1}{s-1} & \frac{1}{s+1} & 0 & 0 \\ \frac{1}{s+1} & \frac{1}{s-1} & \frac{1}{s+1} & \frac{1}{s+1} & 0 \\ \frac{1}{s+1} & \frac{1}{s-1} & \frac{1}{s+1} & \frac{1}{s+1} & \frac{1}{s-1} \end{bmatrix}$$

with P given by

$$P_{11} = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} \qquad P_{12} = \begin{bmatrix} G \\ I \end{bmatrix} \qquad P_{21} = \begin{bmatrix} G & I \end{bmatrix}$$

and a sequence of sparsity constraints $K_1^{\text{bin}}, \ldots, K_6^{\text{bin}}$

$$K_1^{\text{bin}} = \begin{bmatrix} 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&1 \end{bmatrix} \quad K_2^{\text{bin}} = \begin{bmatrix} 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 1&1&0&0&1 \end{bmatrix} \quad K_3^{\text{bin}} = \begin{bmatrix} 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 1&1&0&0&0 \\ 1&1&0&0&1 \end{bmatrix}$$

$$K_4^{\text{bin}} = \begin{bmatrix} 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 1&1&0&0&0 \\ 1&1&1&0&1 \end{bmatrix} \quad K_5^{\text{bin}} = \begin{bmatrix} 0&0&0&0&0 \\ 0&1&0&0&0 \\ 0&1&0&0&0 \\ 1&1&1&0&0 \\ 1&1&1&0&1 \end{bmatrix} \quad K_6^{\text{bin}} = \begin{bmatrix} 1&0&0&0&0 \\ 1&1&0&0&0 \\ 1&1&1&0&0 \\ 1&1&1&1&0 \\ 1&1&1&1&1 \end{bmatrix}$$

defining a sequence of information constraints $S_i = \text{Sparse}(K_i^{\text{bin}})$ such that each subsequent constraint is less restrictive, and such that each is quadratically invariant under G. We also use S_7 as the set of controllers with no sparsity constraints; i.e., the centralized case. A stable and stabilizing controller which lies in the subspace defined by any of these sparsity constraints is given by

$$K_{\text{nom}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{-6}{s+3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-6}{s+3} \end{bmatrix}$$

We can then find T_1, T_2, T_3 as in (4.3), and then find the stabilizing controller that minimizes the closed-loop norm subject to the sparsity constraints by solving problem (6.1), as outlined in Section 6. The graph in Figure 6.1 shows the resulting minimum $\mathcal{H}_2$-norms for the six sparsity constraints as well as for a centralized controller.

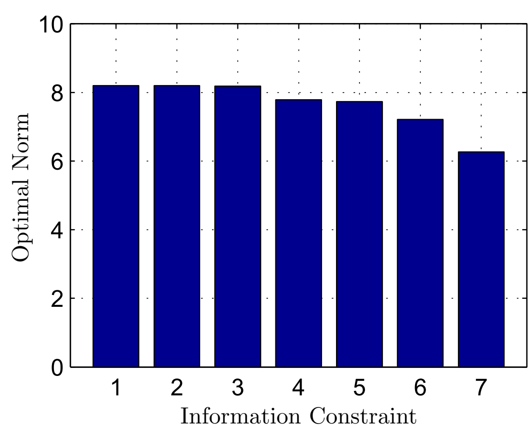

Figure 6.1: Sensitivity of optimal performance to sparsity constraints

6.2.2 Delay Examples

We consider a distributed control example like that of Figure 1.2. Let the plant be such that

$$G_{ij}(s) \;=\; 0.5^{|i-j|}\,\frac{1}{s+1}$$

so that the effect of inputs falls off for more distant subsystems.

Let the propagation delay be 0.5 seconds. We seek to minimize the closed-loop $\mathcal{H}_2$-norm where the rest of the generalized plant is given as

$$P_{11} = \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} \qquad P_{12} = \begin{bmatrix} G \\ \eta I \end{bmatrix} \qquad P_{21} = \begin{bmatrix} G & I \end{bmatrix} \qquad \eta = 1e-3$$

From the results of Section 5.3, and Section 5.3.4 in particular, we know that this problem is quadratically invariant for any transmission delays up to 0.5 seconds, and for any computational delays.

We first fix the computational delay at 0.1 seconds, and observe how the optimal performance varies as the transmission delay decreases from 0.5 to 0 seconds. This is shown in Figure 6.2, where the delay is indicated in tenths of a second, and the performance of the optimal centralized controller is shown for comparison, which corresponds to both a transmission delay and a computational delay of 0.

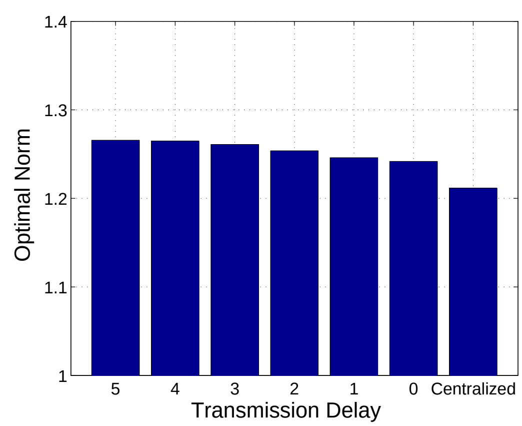

Figure 6.2: Sensitivity of optimal performance to transmission delay

We then fix the transmission delay at 0.2 seconds, and observe how the optimal performance varies as the computational delay decreases from 0.5 to 0 seconds. This is shown in Figure 6.3, where the delay is also indicated in tenths of a second, and the performance of the optimal centralized controller is again shown for comparison.

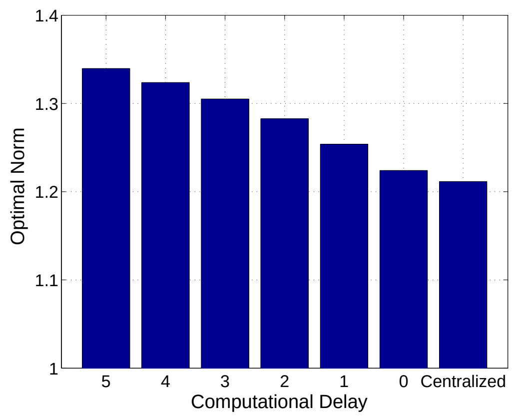

Figure 6.3: Sensitivity of optimal performance to computational delay

It is worth noting that even though the plants considered in this section are fairly simple, the sparsity and delay constraints are such that nearly all of these problems would have previously been considered to be intractable.

Chapter 7

Conclusion

We defined the notion of quadratic invariance of a constraint set with respect to a plant. We showed in Theorems 10 and 18 that quadratic invariance is necessary and sufficient for the constraint set to be preserved under feedback, for operators on Banach spaces and extended spaces, respectively. In Theorem 23, we then proved that quadratic invariance allows us to choose a controller parameterization such that the information constraint is equivalent to an affine constraint on the Youla parameter. Thus synthesizing optimal decentralized controllers becomes a convex optimization problem.

We then applied this to some specific constraint classes. We provided a test for sparsity constraints to be quadratically invariant, and thus amenable to convex synthesis. We noted that symmetric synthesis is included in this classification. We showed in Theorem 33 that for control over networks with delays, optimal controllers may be synthesized in this manner if the communication delays are less than the propagation delays. We further showed that this result still holds in the presence of computational delays.

We thus characterized a broad and useful class of tractable decentralized control problems, unifying many previous results regarding specific structures, and identifying many new ones.

Bibliography

[1] V. Anantharam and C. A. Desoer. On the stabilization of nonlinear systems. *IEEE Transactions on Automatic Control*, 29(6):569–572, 1984.

[2] B. Bamieh and P. G. Voulgaris. Optimal distributed control with distributed delayed measurements. In *Proceedings of the IFAC World Congress*, 2002.

[3] V. D. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.

[4] S. Boyd and C. Barratt. *Linear Controller Design: Limits of Performance*. Prentice-Hall, 1991.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003.

[6] J. B. Conway. *Functions of One Complex Variable I*. Springer Verlag, 1978.

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

[8] C. Fan, J. L. Speyer, and C. R. Jaensch. Centralized and decentralized solutions of the linear-exponential-gaussian problem. *IEEE Transactions on Automatic Control*, 39(10):1986–2003, 1994.

[9] B. Francis. A course in $\mathcal{H}_\infty$ control theory. In *Lecture notes in control and information sciences*, volume 88. Springer-Verlag, 1987.

[10] Y-C. Ho and K. C. Chu. Team decision theory and information structures in optimal control problems – Part I. *IEEE Transactions on Automatic Control*, 17(1):15–22, 1972.

[11] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[12] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley and Sons, Inc., 1969.

[13] S. Mitter and A. Sahai. Information and control: Witsenhausen revisited. *Lecture Notes in Control and Information Sciences*, 241:281–293, 1999.

[14] C. H. Papadimitriou and K. Steiglitz. *Combinatorial Optimization*. Prentice-Hall, Englewood Cliffs, NJ, 1982.

[15] C. H. Papadimitriou and J. N. Tsitsiklis. Intractable problems in control theory. *SIAM Journal on Control and Optimization*, 24:639–654, 1986.

[16] X. Qi, M. Salapaka, P.G. Voulgaris, and M. Khammash. Structured optimal and robust control with multiple criteria: A convex solution. *IEEE Transactions on Automatic Control*, 49(10):1623–1640, 2004.

[17] R. Radner. Team decision problems. *Annals of mathematical statistics*, 33:857–881, 1962.

[18] M. Rotkowitz and S. Lall. Decentralized control information structures preserved under feedback. In *Proceedings of the IEEE Conference on Decision and Control*, pages 569–575, 2002.

[19] M. Rotkowitz and S. Lall. Decentralized control of unstable systems and quadratically invariant information constraints. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2865–2871, 2003.

[20] M. Rotkowitz and S. Lall. Decentralized control subject to communication and propagation delays. In *Proceedings of the IEEE Conference on Decision and Control*, pages 4990–4997, 2004.

[21] M. Rotkowitz and S. Lall. On computation of optimal controllers subject to quadratically invariant sparsity constraints. In *Proceedings of the American Control Conference*, pages 5659–5664, 2004.

[22] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, to appear.

[23] D. D. Siljak. *Decentralized control of complex systems.* Academic Press, Boston, 1994.

[24] A. Taylor and D. Lay. *Introduction to Functional Analysis.* John Wiley and Sons Press, 1980.

[25] E. C. Titchmarsh. *Introduction to the theory of Fourier integrals.* Oxford University Press, 1937.

[26] P. G. Voulgaris. Control under structural constraints: An input-output approach. In *Lecture notes in control and information sciences*, pages 287–305, 1999.

[27] P. G. Voulgaris. A convex characterization of classes of problems in control with specific interaction and communication structures. In *Proceedings of the American Control Conference*, 2001.

[28] S. Wang and E. J. Davison. On the stabilization of decentralized control systems. *IEEE Transactions on Automatic Control*, 18(5):473–478, 1973.

[29] J. C. Willems. *The Analysis of Feedback Systems.* The M.I.T. Press, 1971.

[30] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal of Control*, 6(1):131–147, 1968.

[31] H. S. Witsenhausen. Separation of estimation and control for discrete time systems. *Proceedings of the IEEE*, 59(11):1557–1566, 1971.

[32] G.-H. Yang and L. Qiu. Optimal symmetric $\mathcal{H}_2$ controllers for systems with collocated sensors and actuators. *IEEE Transactions on Automatic Control*, 47(12):2121–2125, 2002.

[33] D.C. Youla, H.A. Jabr, and J.J. Bonjiorno Jr. Modern Wiener-Hopf design of optimal controllers: part II. *IEEE Transactions on Automatic Control*, 21:319–338, 1976.

[34] G. Zames. Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses. *IEEE Transactions on Automatic Control*, 26(2):301–320, 1981.

[35] R. Zimmer. *Essential Results of Functional Analysis.* The University of Chicago Press, 1990.