Order Number 9408647

Design of control systems to meet l_∞ specifications

McDonald, James Stuart, Ph.D.

Rice University, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

RICE UNIVERSITY

Design of Control Systems to Meet l_∞ Specifications

by

James Stuart McDonald

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

J. Boyd Pearson, Chairman
J.S. Abercrombie Professor of Electrical
and Computer Engineering

Ian D. Walker
Assistant Professor of Electrical
and Computer Engineering

Richard A. Tapia
Professor of Mathematical Sciences

Houston, Texas

August, 1992

Design of Control Systems to Meet l_∞ Specifications

James Stuart McDonald

Abstract

Three control system design problems aimed at meeting specifications given in terms of the l_∞ norm, or peak magnitude, of disturbances and errors are solved.

The specification in the first problem requires simply that the error satisfy an l_∞ norm constraint for all time provided that the disturbance satisfies a corresponding constraint. This is the standard l_1 problem, and the results here include generalizations of many known results on this problem: existence of optimal compensators, FIR sub-optimal approximation, and super-optimal approximation.

The specifications in the remaining two problems are based on l_∞ measures of *weighted* disturbances and errors; these weights can be chosen such that constraining a weighted signal constrains, for example, its peak rate and/or acceleration.

One is an *incremental* weighted specification, requiring that the weighted error satisfy an l_∞ norm constraint *up until any given time* provided that the weighted disturbance satisfies a corresponding constraint *up until the same time*. The other is a weighted specification which requires that the weighted error satisfy an l_∞ norm constraint *for all time* provided that the weighted disturbance satisfies a corresponding constraint *for all time*. The associated design problems turn out to be distinct.

For each of the two weighted specifications an appropriate system norm (or gain with respect to the given weights) is defined and it is shown that it can be computed by solving a standard l_1 problem (in the incremental weighted case) or a very similar problem (in the weighted case). Results for each weighted design problem parallel those for the unweighted, or standard l_1, case: existence of optimal compensators, FIR sub-optimal approximation, and super-optimal approximation.

Contents

Chapter 1

Introduction

1.1 Background and Motivation

Control system design is concerned with modifying the behavior of a given system (the *plant*) in such a way as to make it more desirable in some sense. This is generally done by interconnecting the plant with a system (the *compensator*) specified by the designer; the aim is that the interconnection of plant and compensator have more desirable behavior than the plant alone.

A feedback interconnection is one in which some of the plant outputs (the *measured outputs*) are "fed back" through the compensator which, in turn, manipulates some of the plant inputs (the *control inputs*). It has long been known that if a feedback interconnection is used the compensator can be chosen such that the resulting *closed loop system* has many desirable properties; for this reason the study of control system design has been concerned mainly with the design of feedback compensators.

Even in this narrower meaning, many considerations are involved in control system design. The designer may or may not be free to choose which of the plant outputs to measure or which of the plant inputs to control; this choice, if available, can be the dominant consideration in the design process. Whatever the choice, physical considerations such as distance (a plant can be a physically large system such as an electrical power grid) may impose constraints on the structure of the compensator itself; it may necessarily be "distributed" or "de-centralized".

A particularly problematic aspect of the design process involves specifying "desirable" or even "optimal" behavior. This requires first determining the significant sources of the plant's undesirable behavior. These are often modelled as a set of exogeneous signals entering through a set of plant inputs (the *disturbance inputs*). Next, acceptable properties of the closed loop system must be described; this is often done in terms of a set of plant outputs (the *regulated outputs* or *error outputs*). If these things are done, a *specification*, or statement of desirable behavior, can be formulated. One possible class of specifications has the form

- The disturbance inputs belong to some class (of possible disturbances) implies that the error outputs belong to some class (of acceptable errors).

If this statement is true for the closed loop system then the compensator *satisfies* (or *meets*) *the specification.* Given such a specification, the associated *design problem* is to find, if possible, a compensator which satisfies it.

If it is assumed that the available measured outputs and control inputs have been chosen, that the structure of the compensator is otherwise unconstrained, and that significant disturbance inputs and error outputs have been prespecified then the feedback interconnection of plant $\mathcal{G}$ and compensator $\mathcal{C}$ in Figure 1.1 is the most general model of the situation. w and u are the disturbance and control inputs, respectively, to $\mathcal{G}$, and z and y are its error and measured outputs, respectively. The setting of Figure 1.1 has become, over the last decade or so, the standard problem setting for many design problems because of its generality. (As drawn in the figure, $\mathcal{G}$ is actually more often called the *generalized plant* since, historically, disturbance inputs and error outputs have not been considered a part of the "plant" model)

Formulating precise specifications and solving the associated design problems requires mathematical models of the signals and systems of Figure 1.1. If, for example, the possible disturbances can be modelled as a ball in a normed space $\boldsymbol{W}$ with norm $\|\cdot\|_W$ and the acceptable errors as a ball in another space $\boldsymbol{Z}$ with norm $\|\cdot\|_Z$ then a precise specification can be formulated:

- $w \in \boldsymbol{W}$ and $\|w\|_W \leq 1$ implies $z \in \boldsymbol{Z}$ and $\|z\|_Z \leq 1$.

If, in addition, $\mathcal{G}$ and $\mathcal{C}$ can be modelled as linear operators between appropriate spaces then $\mathcal{C}$ satisfies the specification if and only if the closed loop operator $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ from w to z is a bounded linear operator from $\boldsymbol{W}$ into $\boldsymbol{Z}$ and its induced norm is

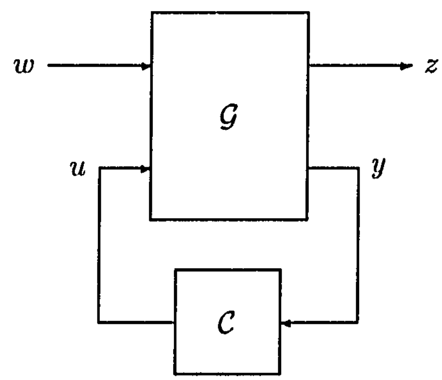

Figure 1.1: Standard Problem Setting

sufficiently small, i.e., $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\| \leq 1$. In this setting, a given $\mathcal{C}$ can be said to *achieve a performance level of* $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|$, and $\mathcal{C}$ can properly be called *optimal* (with respect to a given class of compensators under consideration) if it achieves the minimal performance level (over all compensators in that class).

Design problems associated with specifications of this type are called *disturbance rejection* problems, and were first formulated by Zames in [1]. He observed in particular that if $\boldsymbol{W} = \boldsymbol{Z} = L_2$ (the Lebesgue space) then, for a large class of continuous time linear time invariant $\mathcal{G}$ and $\mathcal{C}$, $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\| = \|\hat{T}_{zw}(\mathcal{G},\mathcal{C})\|_{H_\infty}$. (Here $\hat{T}_{zw}(\mathcal{G},\mathcal{C})$ denotes the transfer function of the closed loop system from w to z and $\|\cdot\|_{H_\infty}$ is the norm on the Hardy space $\boldsymbol{H}_\infty$.) Simply stated, a compensator satisfies an L_2 disturbance rejection specification if and only if the peak of its frequency response magnitude does not exceed 1.

The L_2 disturbance rejection problem is usually called the $\boldsymbol{H}_\infty$ problem for obvious reasons, and has been widely studied since its formulation in 1981. It has been largely solved [2], including the multivariable case in which all signals are vectors; it is known when optimal compensators exist, parametrizations of all compensators achieving performance no greater than some given level are available, bounds have been established on the necessary complexity of optimal compensators, etc. Moreover, extensive computational tools are widely available for its solution (available commercially, for example, as part of MatLab) and it has been applied in a number of practical settings.

If $\boldsymbol{W} = \boldsymbol{Z} = l_\infty$ (the classical sequence space) then the corresponding specification is

l_∞ Disturbance Rejection Specification (l_∞ DRS):

- $w \in l_\infty$ and $\|w\|_{l_\infty} \leq 1$ implies $z \in l_\infty$ and $\|z\|_{l_\infty} \leq 1$.

Vidyasagar made an observation in [3] similar to that of Zames; for a large class of discrete time linear time invariant $\mathcal{G}$ and $\mathcal{C}$, $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\| = \|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_1}$. (Here $T_{zw}(\mathcal{G},\mathcal{C})$ denotes the impulse response of the closed loop system from w to z and $\|\cdot\|_{l_1}$ the norm on the classical sequence space l_1.) Hence a compensator satisfies an l_∞ disturbance rejection specification if and only if the sum of the absolute values of the closed loop impulse response coefficients does not exceed 1.

The l_∞ disturbance rejection problem is, again for obvious reasons, usually called the l_1 problem. As it was formulated only in 1986, it has received less study and is less well understood. The key elements of its solution, however, were given by Dahleh and Pearson in a series of papers [4] [5] [6].

There has been an enormous amount of research into the $\boldsymbol{H}_\infty$ problem; only two landmark papers have been cited here. There has also been a significant amount of research into associated problems in which various system norms are minimized. Even for the two principal signal norms discussed here, there are many variations depending upon the classes of systems considered. There has been work, for example, on $\boldsymbol{H}_\infty$ for discrete time systems [7], continuous time L_1 [8], disturbance rejection with time-varying plants and/or compensators [9][10] or in hybrid systems (consisting of continuous time plant and discrete time compensator) [11] [12].

Any norm-based disturbance rejection problem can be thought of as a "worst-case" design problem; z is required to be in a certain normed ball *no matter what disturbance w actually occurs*, provided that it is also in a specified normed ball. There is no notion of probability; any possible disturbance is considered as likely to occur as any other and as a result design to satisfy disturbance rejection specifications is inherently "conservative". Nonetheless, the mathematical framework in which disturbance rejection problems are formulated has allowed sophisticated methods for analysis of compensator performance and synthesis of optimal compensators to be developed, and has led to new insights into fundamental limitations on achievable performance.

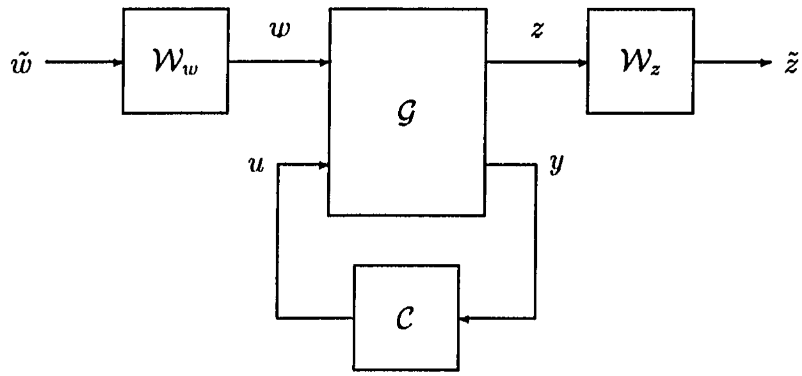

Figure 1.2: Cascade weighting

An appropriate goal for researchers in the area of control system design is thus to try to alleviate the conservatism of these methods without relinquishing the mathematical framework which makes theoretical results obtainable and, therefore, design possible. One way in which this can be done is to provide more flexibility in formulating specifications; some of the conservatism in design results from the fact that actual classes of possible disturbances and acceptable errors cannot be completely described in terms appropriate to the theory. If they can be more completely described then the specification can be made to more accurately reflect reality and the resulting designs will be less conservative.

There is a particularly simple approach which has been widely taken to broaden the class of specifications which can be addressed by the $\boldsymbol{H}_\infty$ and the l_1 theories. It involves the selection of weighting systems (*weights*) $\mathcal{W}_w$ and $\mathcal{W}_z$ and connecting them in cascade with w and z, respectively, as shown in Figure 1.2. The new signals $\tilde{w}$ and $\tilde{z}$ are "fictitious", as are the weights; the class of possible disturbances w is simply being modelled as a normed ball in the appropriate space passed through $\mathcal{W}_w$ and the class of acceptable errors as those which, after passing through $\mathcal{W}_z$, lie in a normed ball. The disturbance rejection specification corresponding to this *cascade weighting* scheme is

Cascade Weighted DRS:

- $w = \mathcal{W}_w\tilde{w}$ for some $\tilde{w} \in \boldsymbol{W}$ with $\|\tilde{w}\|_W \leq 1$ implies $\mathcal{W}_z z \in \boldsymbol{Z}$ and $\|\mathcal{W}_z z\|_Z \leq 1$.

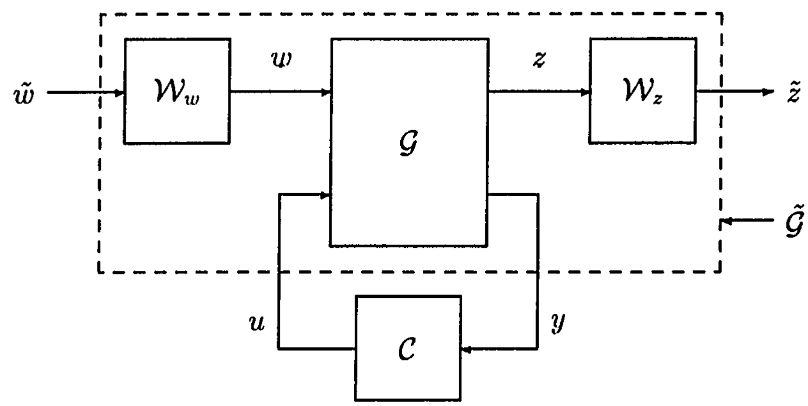

Figure 1.3: Equivalent generalized plant for cascade weighting

Under suitable assumptions on the weights, the L_2 and the l_∞ versions of the cascade weighted disturbance rejection problem are equivalent to the $\boldsymbol{H}_\infty$ and l_1 problems, respectively, for the generalized plant $\tilde{\mathcal{G}}$ depicted in Figure 1.3. This is an obvious advantage from a theoretical standpoint and, in the case of the $\boldsymbol{H}_\infty$ problem, has an appealing practical interpretation. Because the $\boldsymbol{H}_\infty$ norm is the maximum of the frequency response, $\boldsymbol{H}_\infty$ design is often viewed not as aiming to satisfy disturbance rejection specifications in terms of normed balls, but instead as direct frequency domain design. That is, specifications can be formulated in terms of the shape of the closed loop frequency response; in particular, specifying that its magnitude not exceed a given bound which is a function of frequency. Under this interpretation, cascade weights can be chosen so that $\|\hat{T}_{\tilde{z}\tilde{w}}(\tilde{\mathcal{G}},\mathcal{C})\|_{H_\infty} \leq 1$ is equivalent to $|\hat{T}_{\tilde{z}\tilde{w}}(\mathcal{G},\mathcal{C})|$ lying under any given curve (as a function of frequency). Hence $\boldsymbol{H}_\infty$ design can address such specifications precisely using cascade weights.

The situation is more complicated for l_∞. The error weight $\mathcal{W}_z$ has an appealing interpretation because of the definition of the l_∞ norm for vector signals (as the maximum l_∞ norm of any component signal). In particular, $\mathcal{W}_z$ can be chosen such that $\|\mathcal{W}_z z\|_{l_\infty} \leq 1$ if and only if $|z(k)| \leq 1$ and $|z(k) - z(k-1)| \leq 1$ for all k. Thus an error is acceptable if and only if both its magnitude *and* its rate of change are bounded by 1 for all time. In fact, bounds on n-th order differences of z for any desired n can be incorporated into the specification of the acceptable error class; 2nd order difference bounds, for example, specify limited accelerations or forces in a mechanical system.

The interpretation of $\mathcal{W}_w$ is problematic, however; it is not known how to choose $\mathcal{W}_w$ such that an l_∞ ball of $\tilde{w}$s produces precisely a class of magnitude and rate bounded disturbances w. Moreover, the frequency domain interpretation which gives cascade weights their appeal in the $\boldsymbol{H}_\infty$ case is absent here; there is no simple relationship between the l_1 or l_∞ norms and frequency response of a system or frequency content of a signal.

Motivated by the appealing interpretation of $\mathcal{W}_z$ in the l_∞ disturbance rejection problem, a weighted specification can be formulated as follows.

Weighted l_∞ DRS:

- $\mathcal{W}_w w \in l_\infty$ and $\|\mathcal{W}_w w\|_\infty \leq 1$ implies $\mathcal{W}_z z \in l_\infty$ and $\|\mathcal{W}_z z\|_\infty \leq 1$.

In this specification both the possible disturbances and the acceptable errors can be specified in the manner that errors can be in the cascade weighted specification. In particular, classes of possible disturbances consisting precisely of those satisfying bounds on magnitude and/or rate and/or additional n-th order differences can be specified.

A related specification with an appealing interpretation of its own is

Incremental Weighted l_∞ DRS:

- $\|\mathcal{P}_n \mathcal{W}_w w\|_\infty \leq 1$ implies $\|\mathcal{P}_n \mathcal{W}_z z\|_\infty \leq 1$ *for all* n.

where $\mathcal{P}_n$ denotes truncation at time n (setting to zero after time n). This is different from the disturbance rejection specifications above in that it does not define classes of possible disturbances and acceptable errors a priori, but has a temporal aspect. It requires that the weighted error satisfy a norm constraint *up until any given time* provided that the weighted disturbance satisfies a norm constraint *up until the same time.* Specifications similar to these have been considered in [13] and [14].

1.2 Scope

In this thesis the three l_∞ design problems, defined in general terms in Section 1.1, are considered:

- l_∞ Disturbance Rejection

- Weighted l_∞ Disturbance Rejection
- Incremental Weighted l_∞ Disturbance Rejection

For each, a detailed statement of the specification or class of specifications associated with the problem is given and a consistent set of results are obtained under minimal assumptions:

- The problem is formulated as minimization of a norm of the closed loop system $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ from disturbance to error over a well defined class of compensators.
- It is shown to be equivalent in a strong (though not exact) sense to a minimum distance problem in an infinite dimensional normed linear space.
- Existence of a minimizer for this problem is established.
- Computable approximate and/or exact solution methods are given, i.e.,
  - The minimum distance problem can in some cases be reduced to a finite dimensional problem and in these cases exact solution methods are given.
  - In all other cases, finite dimensional optimization problems which approximate it (both sub-optimally and, in the incremental case, super-optimally) are given.
- All optimization problems formulated are shown to be equivalent to (infinite or finite) linear programs. Detailed formulations are given for each linear program.

In the case of the unweighted l_∞ disturbance rejection problem, it is well known that the l_1 norm of the closed loop impulse response is the norm which should be minimized. In the two weighted problems, however, an appropriate norm is first defined and it is shown how to compute each.

Because all the problems arise in the same basic setting (that of Figure 1.1) this setting is also defined and explored in some detail:

- The class of exogeneous signals to be considered is defined.
- A general class of systems is defined to which the generalized plant is assumed to belong, and from which the compensator is to be chosen.

- The algebraic properties of such systems and their feedback interconnections are described.

1.3 Organization

The remainder of the thesis is organized as follows:

Chapter 2 gives a careful definition of the standard problem setting and the basic results on signals, systems, and feedback interconnections outlined in Section 1.2. Many of these results are algebraic in nature; the details are supplied in Appendix A. There notation is introduced, terms briefly defined, and some results on algebraic properties of sequence spaces given.

Chapters 3, 4, and 5 are each devoted to a separate design problem. Chapter 3 considers the l_∞ disturbance rejection problem, or l_1 problem. It covers all aspects of the problem outlined in Section 1.2, in that order: problem statement, equivalence to a minimum distance problem, existence of a minimizer, approximate and/or exact solution methods, and linear programming formulations of each. Appendices B and C contain supporting material used in all of these chapters; the general subject of each is normed linear spaces, duality, operators, and their adjoints. Appendix B defines notation, gives some simple duality results, and quotes some crucial facts concerning minimum distance problems. Appendix C is concerned specifically with normed spaces of sequences, their duality relations, and several classes of operators defined on them and their adjoints.

Chapters 4 and 5 consider the incremental weighted and the weighted l_∞ disturbance rejection problems, respectively. Each is organized identically to Chapter 3 except that each requires an additional preliminary section in which a system norm appropriate to the problem is defined and a method given for its computation. A second additional section follows in each chapter in which the respective problems are formulated in terms of these norms. The problems are considered in this order (incremental first) for technical reasons; the incremental problem is more similar mathematically to the unweighted problem, and more results are possible.

Each of chapters 2 through 5 has a brief introductory section which outlines the chapter in some detail, introduces necessary notation, and states all assumptions in effect throughout that chapter. (Some notation required for the appendices may be defined in the associated chapter(s) and vice versa.) Each concludes with a brief discussion.

Chapter 6 contains a summary and discussion of the thesis and some suggestions for future related research.

The remainder of this chapter consists of two sections: Section 1.4, which describes in general terms the contribution of the thesis, and Section 1.5, which defines some general notation and terminology.

1.4 Contribution of the Thesis

This section describes in very general terms related work and the contribution of this thesis; each chapter summary contains a more detailed discussion of related work with references, and the contribution of that chapter.

The objective of the thesis is to give both comprehensive and detailed solutions to the three design problems posed. All three problems are solved from specification through to the formulation of optimization problems which can be readily implemented with the information provided. This is motivated by the desire to resolve any ambiguity in the practical meaning of the theory, to identify key assumptions made along the path from specification to solution, and to allow the implementation of design algorithms where they do not presently exist.

Chapter 2 provides a common setting for all three problems and, while most results there are known in some form or another, there are some novel aspects to the approach taken.

The treatment of the l_∞ disturbance rejection problem, or l_1 problem, in Chapter 3 consists of work published by the author [15] and a large number of considerable improvements on that work. It represents the first treatment of the general (4 block or "bad rank") problem and contains generalized versions of most known design methods.

Both the incremental weighted and weighted problems of Chapters 4 and 5, respectively, are extremely little studied in spite of their apparent practical appeal. Only minor portions of the results here, which come close to reproducing those available for the standard l_1 problem, have appeared elsewhere. These chapters represent the most novel aspect of the thesis, and the most important from a theoretical viewpoint.

1.5 Notation

R the real numbers.

$\mathbf{Z}, \mathbf{Z}_+$ the integers and the non-negative integers, respectively.

C the complex numbers.

$\Re z, \Im z$ the real and imaginary parts, respectively, of $z \in \mathbf{C}$.

$\mathbf{D}, \overline{\mathbf{D}}$ the open and closed unit disks in **C**, respectively.

In the following definitions $\boldsymbol{X}$ is a set and m and n are positive integers.

$\boldsymbol{X}^{m\times n}$ the set of all $m \times n$ matrices with entries in $\boldsymbol{X}$. An element $X \in \boldsymbol{X}^{m\times n}$ is a *matrix over* $\boldsymbol{X}$.

$\boldsymbol{X}^n$ the set of all column vectors of dimension n with entries in $\boldsymbol{X}$, i.e., $\boldsymbol{X}^{n\times 1}$.

X_{ij} for $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$, the ij-th entry of $X \in \boldsymbol{X}^{m\times n}$.

$X_{i\cdot}, X_{\cdot j}$ for $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$, the i-th row and j-th column, respectively, of $X \in \boldsymbol{X}^{m\times n}$.

Some notational conventions are observed as closely as possible throughout:

- signals are denoted by lower case letters (e.g., x).
- both systems and maps are denoted by calligraphic letters (e.g., $\mathcal{H}$).
- impulse response matrices of systems are denoted by corresponding upper case letters (e.g., H).
- transfer function matrices of systems are denoted by corresponding hatted upper case letters (e.g., $\hat{H}$), where the z-transform is defined with z as the delay.

- Sets or spaces of the above objects are denoted by boldface letters in the corresponding font (e.g., $\boldsymbol{\mathcal{H}}$ denotes a set of systems or maps).

The major exception to the above convention is that boldface upper case letters are used to denote general sets and spaces (e.g., $\boldsymbol{X}$).

Note: Throughout the thesis a product (GH) of impulse response matrices or any matrices of sequences means convolution $(G * H)$.

Chapter 2

Problem Setting

In this chapter, the standard problem setting of Chapter 1 is more carefully defined. It is the setting in which all the design problems of later chapters are solved. The main results of the chapter are contained in three sections.

Section 2.1 defines the class of exogeneous signals to be considered; they are allowed to be completely arbitrary discrete time signals (sequences) except that they must be *suddenly applied.* That is, they must be zero prior to some finite time. This is not an uncommon assumption, but its formulation here differs from the usual. A large class of systems is defined on such signals; *any* map from the signal space into the signal space is considered a system. Linearity, time invariance, and causality are all defined and the set of linear time invariant systems have convolution representations; hence they are characterized by their impulse responses. Stability is defined as usual, in terms of boundedness as an operator.

Section 2.2 contains some results on the algebraic properties of signals and systems. As defined in Section 2.1, signals and linear time invariant systems (more precisely, their impulse responses) are identical algebraically; they form a field under convolution and pointwise addition. The causal linear time invariant systems form a subdomain of this field and, moreover, the fraction field corresponding to the causal systems is the field of linear time invariant systems (and of signals). The stable linear time invariant systems form and its causal subset form distinct subdomains.

Section 2.3 uses the algebraic structure and similarity of signals and systems to define and characterize well posedness for the feedback interconnection in the standard problem setting. Also defined and characterized are internal stability of the interconnection and stabilizability of the generalized plant. Under a mild assumption, a YJBK-type parametrization of all causal linear time invariant compensators which stabilize a given stabilizable causal linear time invariant generalized plant is derived.

The chapter concludes with a discussion in Section 2.4 of related work and some unique features of the results.

Appendix A is referenced frequently throughout, and Appendix C occasionally.

Notation

If $\boldsymbol{T}$ is any subset of $\mathbf{Z}$, a *sequence on* $\boldsymbol{T}$ is a map from $\boldsymbol{T}$ into $\mathbf{R}$. If x is a sequence on $\boldsymbol{T}$ it can be written in terms of its elements as $x = \{x(k)\}_{k \in \boldsymbol{T}}$. The *support* of a sequence x, written supp x, is the subset of $\boldsymbol{T}$ consisting of the indices of all non-zero elements of x. The following are sets of sequences which will be used throughout:

l the set of all sequences on $\mathbf{Z}_+$.

l_∞ the subset of l consisting of all magnitude-bounded sequences.

l_1 the subset of l consisting of all absolutely summable sequences.

$l(\mathbf{Z})$ the set of all sequences on $\mathbf{Z}$.

l_+ the subset of $l(\mathbf{Z})$ consisting of all right-supported sequences, i.e.,

$$l_+ := \{x \in l(\mathbf{Z}) : \exists k_x \in \mathbf{Z} \text{ such that supp } x \subset \{k_x, k_x + 1, \ldots\}\}$$

$l_{\infty+}$ the subset of l_+ consisting of all magnitude-bounded sequences.

l_{1+} the subset of l_+ consisting of all absolutely summable sequences.

2.1 Discrete Time Signals and Systems

In this section a class of suddenly applied discrete time signals is defined; they form a linear space under pointwise scalar multiplication and addition. Both scalar and vector signals are considered. The magnitude-bounded signals form a normed linear space.

Scalar systems are defined as maps from the scalar signal space into itself. Subsets of the set of systems are identified according to properties of linearity, time invariance, and causality; the linear time invariant ones have convolution representations.

System stability is defined as boundedness as an operator from the space of magnitude-bounded sequences into itself, and the linear time invariant stable systems are identified.

Definition 2.1.1 *A* signal of dimension n *is any element of* l_+^n *for some positive integer* n. *A* scalar signal *is a signal of dimension* 1.

Under this definition, the set of signals models *all* suddenly applied excitations, regardless of the time at which they occur. The usual model for such signals is l; Definition 2.1.1 is potentially more general, at least when considering time varying systems, and it results in a clearer delineation of the system properties of causality and time invariance. It also results in a convenient algebraic structure encompassing both signals and systems, described in Section 2.2.

It is easy to check that the set l_+ of all scalar signals is a real linear space with elementwise addition and scalar multiplication, i.e., given $x_1, x_2 \in l_+$,

$$x_1 + x_2 := \{x_1(k) + x_2(k)\}_{k \in \mathbf{Z}}$$

and, given $x \in l_+$ and $\alpha \in \mathbf{R}$,

$$\alpha x := \{\alpha x(k)\}_{k \in \mathbf{Z}} .$$

Hence, for any n, the set of all signals of dimension n is also a linear space, again with elementwise operations.

The set $l_{\infty+}^n$ for any n is a subspace of l_+^n which is normed under

$$\|x\|_{l_\infty} := \max\{\sup\{|x_j(k)| : k \in \mathbf{Z}\} : j \in \{1, \ldots, n\}\} .$$

The set $l_{\infty+}^n$ thus provides a model for the set of all suddenly applied excitations having finite *peak magnitude.* For $x \in l_{\infty+}^n$, $\|x\|_{l_\infty}$ is the maximum peak magnitude of any element x_j of x.

An important scalar signal is the *unit impulse* denoted by δ and defined

$$\delta(k) := \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

The following definition of a system allows convenient study of the responses of systems to such signals as well as cascade, parallel, and feedback interconnections of systems,

Definition 2.1.2 *A* system of dimension $m \times n$ *is a map from the space of signals of dimension* n *into the space of signals of dimension* m*, i.e.,* $\mathcal{H} : l_+^n \mapsto l_+^m$. $\Sigma^{n \mapsto m}$ *denotes the set of systems of dimension* $m \times n$.

A scalar system $\mathcal{H}$ *is a system of dimension* 1×1. *The notation* Σ *denotes the set of all scalar systems.*

Definition 2.1.2 ensures that the operations of scalar multiplication, addition, and multiplication are can be defined on the set of scalar systems (scalar multiplication and addition as usual for maps between linear spaces, and composition as multiplication). In practical terms, this means that cascade and parallel connections of systems are well defined, provided only that their dimensions are compatible. In fact, if $\mathcal{H}_s$ is a scalar system and $\mathcal{H} \in \Sigma^{n \mapsto m}$, the compositions $\mathcal{H}_s\mathcal{H}$ and $\mathcal{H}\mathcal{H}_s$ make sense if we interpret that $\mathcal{H}_s$ is applied to each element of its vector input signal.

The response of every scalar system to the input δ is well defined; hence every system $\mathcal{H} \in \Sigma$ has an associated *impulse response* denoted by $h := \mathcal{H}\delta \in l_+$. Every system $\mathcal{H} \in \Sigma^{n \mapsto m}$ has an associated impulse response *matrix*

$$H = \begin{bmatrix} H_{11} & \cdots & H_{1n} \\ \vdots & \ddots & \vdots \\ H_{m1} & \cdots & H_{mn} \end{bmatrix}$$

where $H_{ij} \in l_+$ is the response of the i-th output of $\mathcal{H}$ when the j-th input is δ and all other inputs are zero.

Since the space of signals of any dimension is linear, linearity for systems has a natural definition.

Definition 2.1.3 *A system* $\mathcal{H} \in \Sigma^{n \mapsto m}$ *is* linear *if*

$$\mathcal{H}(\alpha x_1 + \beta x_2) = \alpha\mathcal{H}(x_1) + \beta\mathcal{H}(x_2) \quad \forall \alpha, \beta \in \mathbf{R} \text{ and } x_1, x_2 \in l_+^n.$$

$\boldsymbol{\Sigma}_l^{n\mapsto m}$ *denotes the set of all linear systems of dimension* $m \times n$ *which have a kernel representation, i.e., for which there exists a doubly indexed sequence* $\{\boldsymbol{H}(k,l)\}_{(k,l)\in\mathbf{Z}\times\mathbf{Z}}$ *of matrices over* $\mathbf{R}$ *such that for every* $x \in l_+$

$$(\mathcal{H}x)_i(k) = \sum_{j=1}^{n}\sum_{l=-\infty}^{\infty} \boldsymbol{H}_{ij}(k,l)x_j(l), \quad i \in \{1,\ldots,m\}.$$

It is easy to see that every system with a kernel representation is linear, but $\boldsymbol{\Sigma}_l$ need not contain all linear systems. Two other system properties of interest can be defined in terms of two special scalar systems (the first is actually a family of systems indexed by $N \in \mathbf{Z}$). The *N-th truncation* $\mathcal{P}_N$ is defined, given $x \in l_+$,

$$(\mathcal{P}_N x)(k) := \begin{cases} x(k) & k \le N \\ 0 & k > N \end{cases}$$

and the *delay* $\mathcal{S}$ is defined, given $x \in l_+$,

$$\mathcal{S}x := \{x(k-1)\}_{k\in\mathbf{Z}}\,.$$

Note that $\mathcal{P}_N, \mathcal{S} \in \boldsymbol{\Sigma}_l$.

Definition 2.1.4 *A system* $\mathcal{H} \in \boldsymbol{\Sigma}^{n\mapsto m}$ *is*

- time invariant *if it commutes with the delay, i.e.,*

$$\mathcal{S}\mathcal{H} = \mathcal{H}\mathcal{S}.$$

- causal *if*

$$\mathcal{P}_N\mathcal{H}\mathcal{P}_N = \mathcal{P}_N\mathcal{H} \quad \forall N \in \mathbf{Z}.$$

All design problems will be solved given a causal linear time invariant, and compensators will be required to have the same properties. Accordingly we define

$$\begin{aligned} \boldsymbol{\Sigma}_{lti}^{n\mapsto m} &:= \{\mathcal{H} \in \boldsymbol{\Sigma}_l^{n\mapsto m} : \mathcal{H} \text{ is time invariant}\} \\ \boldsymbol{\Sigma}_{clti}^{n\mapsto m} &:= \{\mathcal{H} \in \boldsymbol{\Sigma}_{lti}^{n\mapsto m} : \mathcal{H} \text{ is causal}\} \end{aligned}$$

It is easy to check that the binary operation *convolution* denoted by "$*$" and defined, given $x_1, x_2 \in l_+$,

$$x_1 * x_2 := \left\{ \sum_{l=-\infty}^{\infty} x_1(k-l)x_2(l) \right\}_{k\in\mathbf{Z}}$$

is well defined since x_1 and x_2 are right-supported.

The following facts state that systems in Σ_{lti} have convolution representations in terms of their impulse response; they are immediate consequences of standard facts for scalar systems with kernel representations (see, e.g., [16]).

Fact 2.1.5 *If $\mathcal{H} \in \Sigma_l$ then $\mathcal{H} \in \Sigma_{lti}$ if and only if*

$$\mathcal{H}x = h * x \quad \forall x \in l_+^n$$

where $h \in l_+$ denotes the impulse response of $\mathcal{H}$.

Fact 2.1.6 *If $\mathcal{H} \in \Sigma_{lti}$ then $\mathcal{H} \in \Sigma_{clti}$ if and only if* supp $h \subset \mathbf{Z}_+$.

All design problems are l_∞ design problems, i.e., their specifications are in terms of l_∞ norms (peak magnitudes) of disturbances and errors. The following is therefore an appropriate definition of stability.

Definition 2.1.7 *A system $\mathcal{H} \in \Sigma^{n \mapsto m}$ is* stable *if*

- *$\mathcal{H}x \in l_{\infty+}^m$ for all $x \in l_{\infty+}^n$ and*
- *there exists $c < \infty$ such that*

$$\|\mathcal{H}x\|_{l_\infty} \leq c \, \|x\|_{l_\infty} \quad \forall x \in l_{\infty+}^n. \tag{2.1}$$

If $\mathcal{H}$ is stable the smallest c satisfying (2.1) is denoted by $\|\mathcal{H}\|_{l_\infty\text{-i}}$.

The next fact follows immediately from well known results; see for example [16]. For the definition of $\|\cdot\|_{l_1}$ see Fact C.2.1.

Fact 2.1.8 *$\mathcal{H} \in \Sigma_{lti}^{n \mapsto m}$ is stable if and only if $H \in l_{1+}^{m \times n}$. For a stable system $\mathcal{H} \in \Sigma_{lti}^{n \mapsto m}$,*

$$\|\mathcal{H}\|_{l_\infty\text{-i}} = \|H\|_{l_1}.$$

2.2 Algebraic Properties of Signals and Systems

Both Σ_{lti} and Σ_{clti} have useful and related algebraic structures, as do the subset of stable systems in each. Most of the algebraic terms used in this section are meant in their most standard sense; see Section A.1 for definitions in case of uncertainty.

Proposition A.2.1 shows that l_+ forms a ring under convolution and pointwise addition. The following proposition shows that the linear time invariant systems form a ring which is isomorphic to l_+; each system is identified by its impulse response, which lies in l_+.

Proposition 2.2.1 *When addition is defined as the usual addition for maps on a linear space and multiplication is defined as composition, Σ_{lti} forms a ring which is isomorphic to l_+ under the map $\phi : \Sigma_{lti} \mapsto l_+$ defined, given $\mathcal{H} \in \Sigma_{lti}$,*

$$\phi\mathcal{H} := h.$$

Proof: It has already been noted that the linear systems form a natural algebra; to show that Σ_{lti} is a ring it suffices to show that it is closed under addition and multiplication. Accordingly, let $\mathcal{H}_1, \mathcal{H}_2 \in \Sigma_{lti}$ be given and note that

$$\mathcal{S}(\mathcal{H}_1 + \mathcal{H}_2) = \mathcal{S}\mathcal{H}_1 + \mathcal{S}\mathcal{H}_2 = \mathcal{H}_1\mathcal{S} + \mathcal{H}_2\mathcal{S} = (\mathcal{H}_1 + \mathcal{H}_2)\mathcal{S}$$

since $\mathcal{S} \in \Sigma_l$ and

$$\mathcal{S}\mathcal{H}_1\mathcal{H}_2 = \mathcal{H}_1\mathcal{S}\mathcal{H}_2 = \mathcal{H}_1\mathcal{H}_2\mathcal{S}$$

since $\mathcal{H}_1$ and $\mathcal{H}_2$ are each time invariant. Hence both $\mathcal{H}_1 + \mathcal{H}_2$ and $\mathcal{H}_1\mathcal{H}_2$ are time invariant and Σ_{lti} is closed.

Next it is shown that this ring is isomorphic under the given map to l_+. ϕ is well defined and, by its definition, maps Σ_{lti} into l_+. Next we show it is a bijection. By Fact 2.1.6, $\mathcal{H}x = h * x = (\phi\mathcal{H}) * x$ for all $x \in l_+$ so that $\phi\mathcal{H}_1 = \phi\mathcal{H}_2 \Rightarrow \mathcal{H}_1 = \mathcal{H}_2$. To show that ϕ is onto, let $\tilde{h} \in l_+$ be given and define $\mathcal{H} \in \Sigma_{lti}$ by convolution with $\tilde{h}$, i.e.,

$$\mathcal{H}x := \tilde{h} * x \qquad \forall x \in l_+.$$

It is easy to check that $\mathcal{H}$ is well defined and that it is in Σ_{lti}. Moreover $\phi\mathcal{H} := h = \tilde{h}$, showing that ϕ is onto. It is also easy to check that ϕ is a homomorphism, i.e., that

$$\phi(\mathcal{H}_1 + \mathcal{H}_2) = \phi\mathcal{H}_1 + \phi\mathcal{H}_2 \quad \text{and} \quad \phi(\mathcal{H}_1\mathcal{H}_2) = (\phi\mathcal{H}_1)(\phi\mathcal{H}_2)$$

for any two systems $\mathcal{H}_1, \mathcal{H}_2 \in \Sigma_{lti}$, using the definitions of addition and multiplication in l_+ from Proposition A.2.1. □

As a consequence of Proposition 2.2.1, the ring Σ_{lti} inherits all of the algebraic structure of the ring l_+. In particular, Proposition A.2.1 shows that it is a field under convolution and pointwise addition. Moreover, since the set of all scalar signals is also l_+, signals and systems can be viewed as identical algebraically if systems are identified with their impulse responses. Every dimensionally sensible interconnection of systems with inputs and arbitrarily defined outputs can hence be described by a matrix equation over the field l_+; all "block diagram algebra" can be performed as though the signals and systems were vectors and matrices, respectively over $\mathbf{R}$.

Fact A.2.2 states that l also forms a ring under convolution (defined for sequences on $\mathbf{Z}_+$) and pointwise addition. Proposition A.2.3 shows that l is isomorphic to a subring of l_+. The next proposition shows that the causal systems form a ring which is isomorphic to l.

Proposition 2.2.2 *$\boldsymbol{\Sigma}_{clti}$ forms a subring of $\boldsymbol{\Sigma}_{lti}$ isomorphic to l under the map ϕ_c : $\boldsymbol{\Sigma}_{clti} \mapsto l$ defined, given $\mathcal{H} \in \boldsymbol{\Sigma}_{clti}$,*

$$\phi_c \mathcal{H} := \{h(k)\}_{k \in \mathbf{Z}_+} .$$

Proof: To show that $\boldsymbol{\Sigma}_{clti}$ forms a subring of $\boldsymbol{\Sigma}_{lti}$, it suffices to show that it is closed under addition and multiplication in $\boldsymbol{\Sigma}_{lti}$. Accordingly, let $N \in \mathbf{Z}$ be given and note that

$$\mathcal{P}_N(\mathcal{H}_1 + \mathcal{H}_2) = \mathcal{P}_N\mathcal{H}_1 + \mathcal{P}_N\mathcal{H}_2 = \mathcal{P}_N\mathcal{H}_1\mathcal{P}_N + \mathcal{P}_N\mathcal{H}_2\mathcal{P}_N = \mathcal{P}_N(\mathcal{H}_1 + \mathcal{H}_2)\mathcal{P}_N$$

since $\mathcal{P}_N \in \boldsymbol{\Sigma}_l$ and

$$\mathcal{P}_N\mathcal{H}_1\mathcal{H}_2 = \mathcal{P}_N\mathcal{H}_1\mathcal{P}_N\mathcal{H}_2 = \mathcal{P}_N\mathcal{H}_1\mathcal{H}_2\mathcal{P}_N$$

since $\mathcal{H}_1$ and $\mathcal{H}_2$ are each causal.

By Fact 2.1.6, $\mathcal{H} \in \boldsymbol{\Sigma}_{lti}$ is in $\boldsymbol{\Sigma}_{clti}$ if and only if supp $h \subset \mathbf{Z}_+$. Using this, the proof of Proposition 2.2.1 is easily adapted to show that ϕ_c is a bijection. That it is a homomorphism is also easily checked to complete the proof. $\square$

As a consequence of Proposition 2.2.2, $\boldsymbol{\Sigma}_{clti}$ inherits all of the algebraic structure of the ring l. In particular, Proposition A.2.6 shows that l_+ is the field of fractions $\boldsymbol{F}_l$ corresponding to l.

Remark 2.2.3 *In view of the preceding propositions, systems in $\boldsymbol{\Sigma}_{clti}^{n \mapsto m}$ and $\boldsymbol{\Sigma}_{lti}^{n \mapsto m}$ can be identified with matrices in $l^{m \times n}$ and $\boldsymbol{F}_l^{m \times n}$, respectively, and signals in l_+^n with vectors in $\boldsymbol{F}_l^n$.*

Among the other consequences of Proposition 2.2.2 are that $\boldsymbol{\Sigma}_{clti}$ is a proper Euclidean domain (Proposition A.2.4); hence systems in $\boldsymbol{\Sigma}_{clti}^{n \mapsto m}$ have Smith forms and systems in $\Sigma_{lti}^{n \mapsto m}$ have Smith-McMillan forms over Σ_{clti} [17, Appendix B]. Also, applying Corollary A.2.5 provides a test for invertibility of $\mathcal{H} \in \Sigma_{clti}$; $\mathcal{H}$ is invertible if and only if $(\phi_c\mathcal{H})(0) = h(0) \neq 0$.

Fact A.2.7 states that l_1 is a subring of l. The next proposition shows that the stable systems in Σ_{clti} form a subring of Σ_{clti} which is isomorphic to l_1.

Proposition 2.2.4 *The set of stable systems in Σ_{clti} forms a subring of Σ_{clti} isomorphic to l_1 under the map ϕ_c defined in Proposition 2.2.2.*

Proof: The map ϕ_c has already been shown to be a bijective homomorphism from all of Σ_{clti} to l. Fact 2.1.8 states that $\mathcal{H} \in \Sigma_{lti}$ is stable if and only if $h \in l_{1+}$. Hence $\mathcal{H} \in \Sigma_{clti}$ is stable if and only if $\phi_c\mathcal{H} \in l_1$, completing the proof. □

As a consequence of Proposition 2.2.4 and Fact A.2.7, the stable systems in Σ_{clti} form a Hermite domain.

2.3 A General Feedback System

Figure 2.1 depicts a the general feedback interconnection of Chapter 1 which is the setting for all the design problems. The only assumptions on $\mathcal{G}$ and/or $\mathcal{C}$ in effect throughout the chapter are linearity and causality, although most results require additional hypotheses. The dimensions of $\mathcal{G}$ are arbitrary and are dropped, as this should cause no confusion.

Although systems are defined in Section 2.1 in such a way that cascade and parallel interconnections are always well defined, issues of *well posedness* and *internal stability* arise when they are interconnected in feedback [16] [18]. In fact, the additional disturbance inputs v_y and v_u are injected at the input and output of $\mathcal{C}$ in Figure 2.1 solely for the purpose of defining internal stability.

After well posedness and internal stability for a given pair of systems $(\mathcal{G},\mathcal{C})$ are defined, simple tests are given for each. For a given $\mathcal{G}$, there need not exist any $\mathcal{C} \in \Sigma_{clti}$ which internally stabilizes the system, so the notion of *stabilizability* is defined and a test for it given. (Stabilizability is defined in terms of existence of a causal linear time invariant $\mathcal{C}$ because all the design problems impose this restriction.) For a given stabilizable $\mathcal{G} \in \Sigma_{clti}$ a YJBK-type parametrization of all $\mathcal{C}$s which internally stabilize the system is derived under a mild assumption.

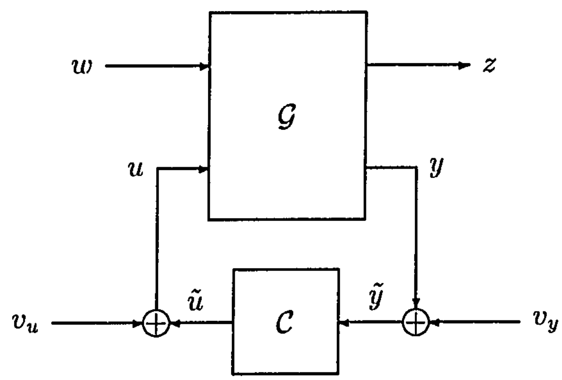

Figure 2.1: Standard problem setting with additional disturbances

The interconnection of Figure 2.1 can be described by the equations

$$\begin{bmatrix} z \\ y \end{bmatrix} = \mathcal{G} \begin{bmatrix} w \\ u \end{bmatrix} \tag{2.2}$$

$$u = v_u + \mathcal{C}y \tag{2.3}$$

and a relation $\Re(\mathcal{G},\mathcal{C})$ on the inputs and outputs in the figure defined as follows

$$\Re(\mathcal{G},\mathcal{C}) := \left\{ \left(\begin{bmatrix} w \\ v_u \\ v_y \end{bmatrix}, \begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} \right) \in l_+ \times l_+ : (2.2),\ (2.3) \text{ satisfied} \right\}. \tag{2.4}$$

In a feedback interconnection as in Figure 2.1 outputs need not exist or be uniquely specified given the inputs. Moreover, a well defined interconnection of two causal systems can define a non-causal system.

Definition 2.3.1 *The pair $(\mathcal{G},\mathcal{C})$ is* well posed *if the relation $\Re(\mathcal{G},\mathcal{C})$ defined in (2.4) describes a causal system $\mathcal{T}(\mathcal{G},\mathcal{C}) : l_+ \mapsto l_+$ as follows:*

$$\mathcal{T}(\mathcal{G},\mathcal{C}) \left(\begin{bmatrix} w \\ v_u \\ v_y \end{bmatrix} \right) := \begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} \text{ such that } \left(\begin{bmatrix} w \\ v_u \\ v_y \end{bmatrix}, \begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} \right) \in \Re(\mathcal{G},\mathcal{C}). \tag{2.5}$$

All design problems require that $\mathcal{C}$ be chosen to ensure not only that $(\mathcal{G},\mathcal{C})$ is well posed but also that $\mathcal{T}(\mathcal{G},\mathcal{C})$ is stable according to Definition 2.1.7.

Definition 2.3.2 *The pair $(\mathcal{G},\mathcal{C})$ is* stable *if it is well posed and $\mathcal{T}(\mathcal{G},\mathcal{C})$ defined by (2.5) is stable. If $(\mathcal{G},\mathcal{C})$ is stable then $\mathcal{C}$* stabilizes *$\mathcal{G}$.*

All design problems also require that $\mathcal{C}$ be chosen from $\boldsymbol{\Sigma}_{clti}$, and that it stabilize $\mathcal{G}$. Such a choice of $\mathcal{C}$ need not exist for every $\mathcal{G}$.

Definition 2.3.3 *$\mathcal{C}(\mathcal{G})$ denotes the set of compensators $\mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ which stabilize $\mathcal{G}$, i.e.,*

$$\mathcal{C}(\mathcal{G}) := \{\mathcal{C} \in \boldsymbol{\Sigma}_{clti} : (\mathcal{G}, \mathcal{C}) \text{ is stable}\}.$$

$\mathcal{G}$ is stabilizable *if $\mathcal{C}(\mathcal{G}) \neq \emptyset$.*

The remainder of this section consists of results concerning well posedness and stability. First, there is a simple algebraic test for well posedness of $(\mathcal{G}, \mathcal{C})$ when $\mathcal{G}, \mathcal{C} \in \boldsymbol{\Sigma}_{clti}$, and when such a pair is well posed, $\mathcal{T}(\mathcal{G}, \mathcal{C}) \in \boldsymbol{\Sigma}_{clti}$. ($\boldsymbol{U}(l)$ denotes the set of units of the ring l; see Corollary A.2.5.)

Note: The dependence of the matrices on the right in (2.7) on $\mathcal{G}$ and $\mathcal{C}$ has been dropped.

Proposition 2.3.4 *If $\mathcal{G}, \mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ then $(\mathcal{G}, \mathcal{C})$ is well posed if and only if*

$$\det(I - G_{yu}C) = \det(I - CG_{yu}) \in \boldsymbol{U}(l).$$

If $(\mathcal{G}, \mathcal{C})$ is well posed then $\mathcal{T}(\mathcal{G}, \mathcal{C}) \in \boldsymbol{\Sigma}_{clti}$ and

$$T(\mathcal{G}, \mathcal{C}) = \tag{2.6}$$

$$\begin{bmatrix} G_{zw} + G_{zu}C(I - G_{yu}C)^{-1}G_{yw} & G_{zu}(I - CG_{yu})^{-1} & G_{zu}C(I - G_{yu}C)^{-1} \\ C(I - G_{yu}C)^{-1}G_{yw} & (I - CG_{yu})^{-1} & C(I - G_{yu}C)^{-1} \\ (I - G_{yu}C)^{-1}G_{yw} & G_{yu}(I - CG_{yu})^{-1} & (I - G_{yu}C)^{-1} \end{bmatrix}$$

$$=: \begin{bmatrix} T_{zw} & T_{zv_u} & T_{zv_y} \\ T_{uw} & T_{uv_u} & T_{uv_y} \\ T_{\tilde{y}w} & T_{\tilde{y}v_u} & T_{\tilde{y}v_y} \end{bmatrix} \tag{2.7}$$

Proof: If $\mathcal{G}, \mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ then, in view of Remark 2.2.3, equations (2.2) and (2.3) defining the input/output pairs of Figure 2.1 can be rewritten as follows

$$\begin{bmatrix} I & -G_{zu} & 0 \\ 0 & I & -C \\ 0 & -G_{yu} & I \end{bmatrix} \begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} G_{zw} & 0 & 0 \\ 0 & I & 0 \\ G_{yw} & 0 & I \end{bmatrix} \begin{bmatrix} w \\ v_u \\ v_y \end{bmatrix} \tag{2.8}$$

where the impulse response matrices have elements in l, and the vector signals have elements in $l_+ = \boldsymbol{F}_l$.

Recall that, for matrices over any ring, $\det(I + D_1 D_2) = \det(I + D_2 D_1)$ so that $\det(I - G_{yu}C) = \det(I - CG_{yu}) =: \Delta$, and note that

$$\det \begin{bmatrix} I & -G_{zu} & 0 \\ 0 & I & -C \\ 0 & -G_{yu} & I \end{bmatrix} = \det(I - G_{yu}C) = \Delta.$$

If $\Delta \in \boldsymbol{U}(l)$ then the matrix on the left in (2.8) is invertible in l. An easy calculation shows that

$$\begin{bmatrix} I & -G_{zu} & 0 \\ 0 & I & -C \\ 0 & -G_{yu} & I \end{bmatrix}^{-1} \begin{bmatrix} G_{zw} & 0 & 0 \\ 0 & I & 0 \\ G_{yw} & 0 & I \end{bmatrix} = \begin{bmatrix} T_{zw} & T_{zv_u} & T_{zv_y} \\ T_{uw} & T_{uv_u} & T_{uv_y} \\ T_{\tilde{y}w} & T_{\tilde{y}v_u} & T_{\tilde{y}v_y} \end{bmatrix} \in l$$

so that the closed loop system is causal and the unique output given any exogeneous inputs w, v_u, and v_y in l_+ is given by

$$\begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} T_{zw} & T_{zv_u} & T_{zv_y} \\ T_{uw} & T_{uv_u} & T_{uv_y} \\ T_{\tilde{y}w} & T_{\tilde{y}v_u} & T_{\tilde{y}v_y} \end{bmatrix} \begin{bmatrix} w \\ v_u \\ v_y \end{bmatrix} \tag{2.9}$$

Thus $(\mathcal{G}, \mathcal{C})$ is well posed.

Now suppose $\Delta \notin \boldsymbol{U}(l)$. If $\Delta = 0$ then the matrix on the left in (2.8) is singular and hence there exist sequences z, u, and $\tilde{y}$ in l_+ such that

$$\begin{bmatrix} I & -G_{zu} & 0 \\ 0 & I & -C \\ 0 & -G_{yu} & I \end{bmatrix} \begin{bmatrix} z \\ u \\ \tilde{y} \end{bmatrix} = 0.$$

The solution to (2.8) is therefore not unique for *any* exogeneous inputs w, v_u, and v_y in l_+. If $\Delta \neq 0$ then the solution is given uniquely for every set of inputs by (2.9), but the closed loop system is not causal (for example, $T_{uv_u} = (I - CG_{yu})^{-1} \notin l$). $\square$

There is also a simple test, which follows immediately from Fact 2.2.4, for stability of any well posed pair.

Proposition 2.3.5 *If $\mathcal{G}, \mathcal{C} \in \Sigma_{clti}$ and $(\mathcal{G}, \mathcal{C})$ is well posed then $(\mathcal{G}, \mathcal{C})$ is stable if and only if $T(\mathcal{G}, \mathcal{C}) \in l_1$.*

Stabilizability is considered next. General tests are not available even for $\mathcal{G}, \mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ but, under a mild assumption, necessary and sufficient conditions for stabilizability can be established. Three preliminary lemmas facilitate the statement and proof of these conditions. The first lemma is fairly standard, and is repeated here for completeness; its proof is only sketched. (See Definition A.1.5 for the definition of a coprime factorization.)

Lemma 2.3.6 *If $\mathcal{G} \in \boldsymbol{\Sigma}_{clti}$ and G_{yu} has either a left or a right coprime factorization over l_1 then it has both left and right coprime factorizations $G_{yu} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ over l_1 and there exist additional matrices X, Y, $\tilde{X}$, and $\tilde{Y}$ over l_1 satisfying the Bezout equation*

$$\begin{bmatrix} \tilde{X} & -\tilde{Y} \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & Y \\ N & X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{2.10}$$

Proof: Suppose first that $G_{yu} = \tilde{M}^{-1}\tilde{N}$ is a left coprime factorization over l_1. Then $[\, -\tilde{N} \;\; \tilde{M} \,] \in l_1$ is right-invertible in l_1 (by Definition A.1.5) and, because l_1 is Hermite (Fact A.2.7), it can be complemented (by Definition A.1.1). The right coprime factorization $G_{yu} = NM^{-1}$ is then guaranteed by [17, Lemma 8.1.45] which states that a matrix with a left coprime factorization $G_{yu} = \tilde{M}^{-1}\tilde{N}$ over any commutative domain with identity also has a right coprime factorization provided that $[\, -\tilde{N} \;\; \tilde{M} \,] \in l_1$ can be complemented. The complete identity (2.10) is constructed in the proof of that lemma. The preceding argument is easily modified (and [17, Corollary 8.1.53] used) in case G_{yu} has instead a right coprime factorization. □

The second lemma shows that every $\mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ which stabilizes a given $\mathcal{G} \in \boldsymbol{\Sigma}_{clti}$ has an impulse response which can be expressed uniquely in terms of a parameter matrix belonging to a certain subset of l_1. The parametrization is very similar to the usual YJBK parametrization of stabilizing compensators for the system $\mathcal{G}_{yu}$ alone connected in feedback with $\mathcal{C}$ (cf., e.g., [17, Theorem 8.3.12]), but the range of the parameter matrix is restricted. This restriction results from the restriction of $\mathcal{C}$ to the causal systems, which is not done explicitly in most versions of the YJBK parametrization.

Lemma 2.3.7 *Let $\mathcal{G}$ satisfy the hypotheses of Lemma 2.3.6 and define sets*

$$\boldsymbol{Q}(\mathcal{G}_{yu}) := \left\{ Q \in l_1 : \det(\tilde{X} - Q\tilde{N}), \det(X - NQ) \in \boldsymbol{U}(l) \right\}$$

and $\mathcal{C}_Q(\mathcal{G}_{yu}) :=$

$$\left\{\mathcal{C} : C = (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1},\ Q \in \boldsymbol{Q}(\mathcal{G}_{yu})\right\}$$

where $G_{yu} = \tilde{M}^{-1}\tilde{N} = NM^{-1}$ *are coprime factorizations of* G_{yu} *over* l_1 *and* $\tilde{X}$, $\tilde{Y}$, X, *and* Y *are arbitrary matrices over* l_1 *satisfying the Bezout equation (2.10).*

If $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$ *then there exists a unique* $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ *such that* $C = (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1}$. *Hence* $\boldsymbol{\mathcal{C}}(\mathcal{G}) \subset \boldsymbol{\mathcal{C}}_Q(\mathcal{G}_{yu})$.

Proof: Assume $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$ and note first that

$$\det(\tilde{M} - \tilde{N}C) = \det\tilde{M}\det(I - G_{yu}C) \text{ and } \det(M - CN) = \det(I - CG_{yu})\det M. \tag{2.11}$$

$G_{yu} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ are coprime factorizations over l as well as l_1 since l_1 is a subring of l (Fact A.2.7), and $G_{yu} \in l$ since $\mathcal{G} \in \boldsymbol{\Sigma}_{clti}$. Hence, by Proposition A.1.6, both $\det M, \det\tilde{M} \in \boldsymbol{U}(l)$. Since $(\mathcal{G}, \mathcal{C})$ is well posed, Proposition 2.3.4 shows that $\det(I - G_{yu}C) = \det(I - CG_{yu}) \in \boldsymbol{U}(l)$. (2.11) thus implies that $\det(\tilde{M} - \tilde{N}C), \det(M - CN) \in \boldsymbol{U}(l)$.

In particular, both determinants are non-zero and it is easy to check, using the identity (2.10), that

$$(\tilde{Y} - \tilde{X}C)(\tilde{M} - \tilde{N}C)^{-1} = (M - CN)^{-1}(Y - CX) =: Q. \tag{2.12}$$

It is also easy to check, again using (2.10), that

$$(X - NQ)(\tilde{M} - \tilde{N}C) = I \text{ and } (M - CN)(\tilde{X} - Q\tilde{N}) = I$$

and hence $\det(\tilde{X} - Q\tilde{N})$, $\det(X - NQ) \in \boldsymbol{U}(l)$. Moreover, (2.12) can be solved (in the two obvious ways) to yield

$$C = (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1}. \tag{2.13}$$

To show that $Q \in l_1$, use the representations (2.13) to compute the lower right-hand two by two block corner of $T(\mathcal{G}, \mathcal{C})$ via (2.6):

$$\begin{bmatrix} (I - CG_{yu})^{-1} & C(I - G_{yu}C)^{-1} \\ G_{yu}(I - CG_{yu})^{-1} & (I - G_{yu}C)^{-1} \end{bmatrix} = \begin{bmatrix} M\tilde{X} & Y\tilde{M} \\ N\tilde{X} & X\tilde{M} \end{bmatrix} - \begin{bmatrix} M \\ N \end{bmatrix} Q \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix}.$$

Hence

$$\begin{bmatrix} M \\ N \end{bmatrix} Q \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \in l_1.$$

$Q \in l_1$ since $[\tilde{N}\ \tilde{M}]$ is right invertible in l_1 and $[M^T\ N^T]^T$ is left invertible in l_1.

Finally, to show that the correspondence is unique, suppose that $Q_1, Q_2 \in \boldsymbol{Q}(\mathcal{G}_{yu})$ and

$$C = (\tilde{X} - Q_1\tilde{N})^{-1}(\tilde{Y} - Q_1\tilde{M}) = (Y - MQ_2)(X - NQ_2)$$

and verify using (2.10) that $Q_1 = Q_2$. □

The third lemma shows, in particular, that $\boldsymbol{Q}(\mathcal{G}_{yu}) \neq \emptyset$. It shows in fact much more, and is used in subsequent chapters as well.

Lemma 2.3.8 *If $\mathcal{G}$ satisfies the hypotheses of Lemma 2.3.6 then $\boldsymbol{Q}(\mathcal{G}_{yu})$ defined in Lemma 2.3.7 is open and dense in l_1.*

Proof: To prove the claim it is sufficient to show that both

$$\boldsymbol{Q}_L := \left\{ Q \in l_1 : \det(\tilde{X} - Q\tilde{N}) \in \boldsymbol{U}(l) \right\}$$

and

$$\boldsymbol{Q}_R := \{ Q \in l_1 : \det(X - NQ) \in \boldsymbol{U}(l) \}$$

are open and dense in l_1, since then $\boldsymbol{Q}(\mathcal{G}_{yu}) = \boldsymbol{Q}_L \cap \boldsymbol{Q}_R$ is as well. We will show only the first of these since the proof of the second is entirely similar.

To show that $\boldsymbol{Q}_L$ is open, the proof of [17, Lemma 5.2.11] carries over precisely. To show that $\boldsymbol{Q}_L$ is dense the proof of that lemma requires some modification, although the main idea is the same; given $Q \in l_1 \backslash \boldsymbol{Q}_L$, a sequence $\{Q_i\}_{i \in \mathbf{Z}_+} \subset \boldsymbol{Q}_L$ is constructed which converges to Q in the l_1 norm.

In view of Corollary A.2.5, $\det(\tilde{X} - Q\tilde{N}) \notin \boldsymbol{U}(l)$ implies that the real matrix $\tilde{X}(0) - Q(0)\tilde{N}(0)$ is singular. However, $[\ \tilde{X}^T \quad -\tilde{N}^T\ ]^T$ is left invertible in l because of (2.10). This implies that the real matrix $[\ \tilde{X}^T(0) \quad -\tilde{N}^T(0)\ ]^T$ has full column rank and hence so does the real matrix $M := [\ [\tilde{X}(0) - Q(0)\tilde{N}(0)]^T \quad -\tilde{N}^T(0)\ ]^T$, since

$$M = \begin{bmatrix} I & Q(0) \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{X}(0) \\ -\tilde{N}(0) \end{bmatrix}.$$

Now select a non-zero full-size minor of M containing a minimal number of rows of $-\tilde{N}$. Let $i_1, \ldots, i_k$ denote the indices of the rows of $\tilde{X}(0) - Q(0)\tilde{N}(0)$ omitted and

$j_1, \ldots, j_k$ the indices of the rows of $-\tilde{N}(0)$ replacing them in forming this minor. Define $R^\epsilon \in \mathbf{R}^{n_u \times n_y}$ by

$$\begin{aligned} R^\epsilon_{i_1 j_1} = \cdots = R^\epsilon_{i_k j_k} &= \epsilon \\ R^\epsilon_{ij} &= 0 \text{ for all other } i, j \end{aligned}$$

By following the proof of [17, Lemma 4.4.21], one can verify that if $\epsilon \neq 0$ then

$$\det [\tilde{X}(0) - (Q(0) + R^\epsilon)\tilde{N}(0)] = \pm\epsilon.$$

Now let $\{\epsilon_i\}_{i \in \mathbf{Z}_+}$ be any set of non-zero numbers converging to zero and defined for each i a matrix $Q_i \in l_1$ by

$$Q_i(k) := \begin{cases} Q(0) + R^{\epsilon_i} & k = 0 \\ Q(k) & k > 0 \end{cases}$$

Then $\lim_{i \to \infty} \|Q - Q_i\|_{l_1} = 0$ and $\det(\tilde{X}(0) - Q_i(0)\tilde{N}(0)) \neq 0$ for each i. Using Corollary A.2.5, the second fact implies that $\det(\tilde{X} - Q_i\tilde{N}) \in \boldsymbol{U}(l)$ for each i, and the proof is complete. □

Using the preceding lemmas, the following proposition is easy to establish. It gives necessary and sufficient conditions for stabilizability of $\mathcal{G}$ and parametrizes, for a given stabilizable $\mathcal{G}$, the set of all $\mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ which stabilize $\mathcal{G}$. The assumption that $\mathcal{G}$ satisfy the hypotheses of Lemma 2.3.6, i.e., that G_{yu} have a coprime factorization over l_1 is the standard assumption required for compensator parametrization and has, in the case of $\boldsymbol{H}_\infty$, been shown to be necessary for stabilizability [19].

Proposition 2.3.9 *Let $\mathcal{G}$ satisfy the hypotheses of Lemma 2.3.6, let all symbols be defined as in Lemma 2.3.7, and define matrices*

$$H := G_{zw} + G_{zu}M\tilde{Y}G_{yw}; \quad U := G_{zu}M; \quad V := \tilde{M}G_{yw}\ .$$

$\mathcal{G}$ is stabilizable if and only if $H, U, V \in l_1$. If $\mathcal{G}$ is stabilizable and $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ then $\mathcal{C}$ with $C := (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1}$ is in $\boldsymbol{\mathcal{C}}(\mathcal{G})$. Hence $\boldsymbol{\mathcal{C}}(\mathcal{G}) = \boldsymbol{\mathcal{C}}_Q(\mathcal{G}_{yu})$.

Proof: First note that for any $\mathcal{C} \in \boldsymbol{\mathcal{C}}_Q(\mathcal{G}_{yu})$, $(\mathcal{G}, \mathcal{C})$ is well posed since

$$\det(I - G_{yu}C) = \frac{1}{\det \tilde{M} \det(X - NQ)} \in \boldsymbol{U}(l)$$

and

$$\det(I - CG_{yu}) = \frac{1}{\det M \det(\tilde{X} - Q\tilde{N})} \in \boldsymbol{U}(l).$$

Also, computing $T(\mathcal{G},\mathcal{C})$ via (2.6), we find that

$$T(\mathcal{G},\mathcal{C}) = \begin{bmatrix} H & U\begin{bmatrix} \tilde{X} & \tilde{Y} \end{bmatrix} \\ \begin{bmatrix} Y \\ X \end{bmatrix} V & \begin{bmatrix} M \\ N \end{bmatrix} \tilde{X} \quad \begin{bmatrix} Y \\ X \end{bmatrix} \tilde{M} \end{bmatrix} - \begin{bmatrix} U \\ M \\ N \end{bmatrix} Q \begin{bmatrix} V & \tilde{N} & \tilde{M} \end{bmatrix}. \quad (2.14)$$

For the "if" part of the first claim, Lemma 2.3.8 shows that there exists at least one $\mathcal{C} \in \mathcal{C}_Q(\mathcal{G}_{yu})$; for this $\mathcal{C}$, $T(\mathcal{G},\mathcal{C})$ has the form (2.14). Clearly if $H, U, V \in l_1$ then so is $T(\mathcal{G},\mathcal{C})$, and hence $(\mathcal{G},\mathcal{C})$ is stable. $\mathcal{C} \in \boldsymbol{\Sigma}_{lti}$ by its definition and, because $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$, $\mathcal{C} \in \boldsymbol{\Sigma}_{clti}$ (using Proposition A.1.6).

For the "only if" part, Lemma 2.3.7 shows that $T(\mathcal{G},\mathcal{C})$ has the form (2.14) for *every* $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$. Hence if there exists a $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$, it follows that

$$U\begin{bmatrix} \tilde{X} & \tilde{Y} \end{bmatrix} \in l_1 \text{ and } \begin{bmatrix} Y \\ X \end{bmatrix} V \in l_1.$$

Thus $U \in l_1$ since $[\ \tilde{X} \quad \tilde{Y}\ ]$ is right invertible in l_1 and $V \in l_1$ since $[\ Y^T \quad X^T\ ]^T$ is left invertible in l_1. Finally, these conclusions and (2.14) combine to show that $H = T_{zw} + UQV \in l_1$.

For the second claim, if $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ then $C := (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1} \in l$ since $\det(\tilde{X} - Q\tilde{N}), \det(X - NQ) \in \boldsymbol{U}(l)$. Hence $\mathcal{C} \in \boldsymbol{\Sigma}_{clti}$. $(\mathcal{G},\mathcal{C})$ is also well posed since, using (2.10),

$$\det(I - G_{yu}C) = \frac{\det \tilde{M}}{\det(X - NQ)}$$

and $\det \tilde{M} \in \boldsymbol{U}(l)$. $T(\mathcal{G},\mathcal{C})$ is given by (2.14) where stabilizability of $\mathcal{G}$ implies that H, U, and V are in l_1. Hence $T(\mathcal{G},\mathcal{C}) \in l_1^{n_z \times n_w}$ and $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$.

Finally, Lemma 2.3.7 establishes that $\boldsymbol{\mathcal{C}}(\mathcal{G}) \subset \mathcal{C}_Q(\mathcal{G}_{yu})$ and the reverse inclusion has just been established. □

Thus from every stabilizable $\mathcal{G}$ three matrices in l_1 can be computed. These matrices are shown below to characterize all closed loop impulse responses from w to z achievable using some $\mathcal{C} \in \boldsymbol{\mathcal{C}}$; hence their properties will be important for the solution of all the design problems. It is therefore of interest to know if H, U, and V computed as in Proposition 2.3.9 have any special properties. The next proposition shows that they do not; to every triple of matrices of appropriate dimensions in l_1 there corresponds some stabilizable $\mathcal{G}$ (in the sense of Proposition 2.3.9).

Proposition 2.3.10 *Given positive integers m, n, m_Q, n_Q and matrices $T_1 \in l_1^{m\times n}$, $T_2 \in l_1^{m\times m_Q}$ and $T_3 \in l_1^{n_Q\times n}$, there exists a stabilizable $\mathcal{G}$ satisfying the hypotheses of Lemma 2.3.6 such that $G_{zw} + G_{zu}M\tilde{Y}G_{yw} = T_1$, $G_{zu}M = T_2$, $\tilde{M}G_{yw} = T_3$.*

Proof: Given T_1, T_2, and T_3, the required $\mathcal{G}$ is simple to construct. Define

$$G := \begin{bmatrix} T_1 & T_2 \\ T_3 & 0 \end{bmatrix}$$

Then $G_{yu} = I^{-1}0 = 0I^{-1}$ are left and right coprime factorizations over l_1, and the Bezout equation 2.10 is satisfied by choosing

$$\begin{bmatrix} \tilde{X} & -\tilde{Y} \\ -\tilde{N} & \tilde{M} \end{bmatrix} = \begin{bmatrix} M & Y \\ N & X \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

It is easy to check that

$$G_{zw} + G_{zu}M\tilde{Y}G_{yw} = T_1, \qquad G_{zu}M = T_2, \qquad \tilde{M}G_{yw} = T_3 \ .$$

so that $\mathcal{G}$ is stabilizable. □

Next is a parameterization, for stabilizable $\mathcal{G}$, of the set of closed loop impulse responses achievable using some causal linear time invariant stabilizing compensator.

Proposition 2.3.11 *Let $\mathcal{G}$ satisfy the hypotheses of Lemma 2.3.6, let all symbols be defined as in Proposition 2.3.9, and define the set*

$$\boldsymbol{T}_{zw}(\mathcal{G}) := \{T_{zw}(\mathcal{G},\mathcal{C}) : \mathcal{C} \in \boldsymbol{C}(\mathcal{G})\}\,.$$

If $\mathcal{G}$ is stabilizable then

- *given $\mathcal{C} \in \boldsymbol{C}(\mathcal{G})$, $T_{zw}(\mathcal{G},\mathcal{C}) = H - UQV$, where $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ is the unique Q guaranteed by Lemma 2.3.7.*

- *given $T_0 = H - UQ_0V$ for some $Q_0 \in \boldsymbol{Q}(\mathcal{G}_{yu})$, let $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ be any Q such that $H - UQV = T_0$. Then $\mathcal{C}$ with $C = (\tilde{X} - Q\tilde{N})^{-1}(\tilde{Y} - Q\tilde{M}) = (Y - MQ)(X - NQ)^{-1}$ is in $\boldsymbol{C}(\mathcal{G})$ and $T_{zw}(\mathcal{G},\mathcal{C}) = T_0$.*

Hence $\boldsymbol{T}_{zw}(\mathcal{G}) = \{H - UQV : Q \in \boldsymbol{Q}(\mathcal{G}_{yu})\}$.

Proof: For the first item, if $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$ then $\mathcal{C} \in \boldsymbol{\mathcal{C}}_Q(\mathcal{G}_{yu})$ by Proposition 2.3.7 and $T(\mathcal{G},\mathcal{C})$ has the form (2.14). In particular, $T_{zw}(\mathcal{G},\mathcal{C}) = H - UQV$.

For the second item, let $T_0 = H - UQ_0V$ be given and let $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ be any Q such that $H - UQV = T_0$. Then the given $\mathcal{C}$ is in $\boldsymbol{\mathcal{C}}_Q(\mathcal{G}_{yu})$ and hence, by Proposition 2.3.9, also $\boldsymbol{\mathcal{C}}(\mathcal{G})$. again $T(\mathcal{G},\mathcal{C})$ has the form (2.14) and, in particular, $T_{zw}(\mathcal{G},\mathcal{C}) = H - UQV = T_0$.

The last claim is a consequence of the two items just established. □

Finally, the next proposition shows that for a large class of generalized plants the set $\boldsymbol{Q}(\mathcal{G}_{yu})$ of admissible Qs is all of l_1.

Proposition 2.3.12 *If* $\mathcal{G}$ *satisfies the hypotheses of Lemma 2.3.6 and* $G_{yu}(0) = 0$ *then* $\boldsymbol{Q}(\mathcal{G}_{yu}) = l_1$.

Proof: We will show that, under the hypotheses, $\det(X - NQ) \in \boldsymbol{U}(l)$ for all $Q \in l_1$; that $\det(\tilde{X} - Q\tilde{N}) \in \boldsymbol{U}(l)$ for each Q as well can be established similarly. By Lemma 2.3.6, let $G_{yu} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ and construct the Bezout identity (2.10). If $G_{yu}(0) = 0$ then $\tilde{N}(0) = \tilde{M}(0)G_{yu}(0) = 0$ and $N(0) = G_{yu}(0)M(0) = 0$. $N(0) = 0$ implies that $(X - NQ)(0) = X(0)$ for all Q and $\tilde{N}(0) = 0$ implies that $X(0)$ is non-singular since, from (2.10), $-\tilde{N}(0)Y(0) + \tilde{M}(0)X(0) = I$. Thus, for any $Q \in l_1$, $\det[(X - NQ)(0)] = \det X(0) \neq 0$ and hence, by Corollary A.2.5, $\det(X - NQ) \in \boldsymbol{U}(l)$. □

2.4 Discussion

The main aim of this chapter has been to provide a common basis for the formulation and solution of the three design problems considered in the remainder of the thesis, and not to develop new results per se. However, some aspects of the classes of signals and systems defined, their algebraic properties, and the parametrizations of stabilizing compensators and closed loop impulse responses in the standard problem setting are of interest.

While it is common to consider suddenly applied disturbances, the approach taken here is unique and has some advantages. The customary approach is to model such signals as sequences on $\mathbf{Z}_+$. When this approach is taken, the natural definition of time invariance of systems taken here (i.e. commuting with the delay) is not satisfying; under it, the notions of causality and time invariance are no longer independent (time invariance implies causality). When signals are defined as sequences on **Z** supported only on the right, all suddenly applied excitations can be considered and time invariance and causality of systems defined naturally and independently.

The algebraic properties of sequence spaces given in Appendix A are not difficult to establish and are most likely not new, but their application to the linear time invariant systems defined here is. As a consequence, signals and linear time invariant systems become simply vectors and matrices over the same field, with the causal systems being matrices over a subring with easily characterized units. This makes well posedness issues particularly clear in feedback interconnections, as they have to do with causal invertibility of systems.

The general approach taken to the parametrization of stabilizing compensators and closed loop impulse responses, i.e., the algebraic approach, is not new. It originated in [20][21] and has been studied extensively (e.g., [22] [23][24]), and an excellent book [17] has appeared on the subject from which this chapter borrows heavily.

Nonetheless, some aspects of the parametrization obtained here are unique; in particular the explicit restriction to causal stabilizing compensators and the doubly coprime factorization (with the same free parameter) of each compensator. The proof of the denseness of admissible values of the free parameter in l_1 is also new, although the ideas are similar to those used in [17] for the case of stable rational functions. Stabilizability tests of various types exist, but the one established here is particularly convenient; it is based on quantities that must already be computed to solve the design problem. Finally, the simple example of Proposition 2.3.10 answers conclusively the question of whether any special structure can be assumed on the H, U, and V matrices which determine the parametrization of closed loop impulse responses.

Chapter 3

l_∞ Disturbance Rejection

Recall the standard problem setting of Chapter 1, (repeated in Figure 3.1 for reference), where w and u are the disturbance and control inputs, respectively, and z and y are the error and measured outputs, respectively, of the generalized plant $\mathcal{G}$. The respective dimensions of the signals are n_w, n_u, n_z, and n_y. Given the definitions of Chapter 2, the l_∞ disturbance rejection specification can be stated precisely.

l_∞ **DRS:**

- $\mathcal{C} \in \Sigma_{clti}$, $(\mathcal{G}, \mathcal{C})$ is stable, and
- $w \in l_{\infty+}^{n_w}$ and $\|w\|_{l_\infty} \leq 1$ implies $z \in l_{\infty+}^{n_z}$ and $\|z\|_{l_\infty} \leq 1$.

To satisfy the specification, $\mathcal{C}$ must be causal, linear, and time invariant, it must stabilize $\mathcal{G}$, and all suddenly applied disturbances of peak magnitude less than or equal to 1 must result in errors of peak magnitude less than or equal to 1. (The requirement that $\mathcal{C} \in \Sigma_{clti}$ ensures that all $\mathcal{C}$s satisfying the specification can be implemented.)

If $\mathcal{G} \in \Sigma_{clti}$ then the design problem can be formulated as minimization of a norm of the closed loop system $\mathcal{T}_{zw}(\mathcal{G}, \mathcal{C})$ from w to z over all $\mathcal{C} \in \Sigma_{clti}$ which stabilize $\mathcal{G}$:

$$DR(\mathcal{G}): \qquad \inf\left\{\|T_{zw}(\mathcal{G}, \mathcal{C})\|_{l_1} : \mathcal{C} \in \mathbf{C}(\mathcal{G})\right\} =: \mu_{DR}$$

If $\mathcal{C} \in \mathbf{C}(\mathcal{G})$ then $T_{zw}(\mathcal{G}, \mathcal{C}) \in l_1^{n_z \times n_w}$ and hence the norm above is always defined. Since the norm is non-negative whenever it is defined, μ_{DR} is defined if and only if $\mathcal{G}$

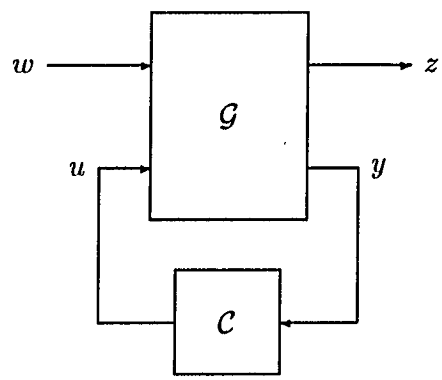

Figure 3.1: Standard problem setting

is stabilizable. Moreover, if $T_{zw}(\mathcal{G},\mathcal{C}) \in l_1^{n_z \times n_w}$ then $z \in l_{\infty+}^{n_z}$ whenever $w \in l_{\infty+}^{n_w}$ and

$$\|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_1} = \|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_\infty\text{-i}} = \sup\left\{\|z\|_{l_\infty} : w \in l_{\infty+}^{n_w} \text{ and } \|w\|_{l_\infty} \leq 1\right\}.$$

$DR(\mathcal{G})$ represents the design problem in the following sense. The feasible solutions for $DR(\mathcal{G})$ are the stabilizing $\mathcal{C}$s for $\mathcal{G}$ in Σ_{clti}; for each $\mathcal{C}$ $T_{zw}(\mathcal{G},\mathcal{C}) \in l_1^{n_z \times n_w}$, and the cost of each, $\|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_1}$, is precisely the worst-case $\|z\|_{l_\infty}$ over all $w \in l_{\infty+}$ with $\|w\|_{l_\infty} \leq 1$ when that $\mathcal{C}$ is used. Hence the first item of the l_∞ DRS can be met if and only if $DR(\mathcal{G})$ is feasible. If $DR(\mathcal{G})$ is feasible then μ_{DR} is the smallest number that can be guaranteed to bound $\|z\|_{l_\infty}$ over all $w \in l_{\infty+}$ with $\|w\|_{l_\infty} \leq 1$. Hence the second specification can be met if and only if $DR(\mathcal{G})$ is feasible and $\mu_{DR} \leq 1$.

The remainder of the chapter is concerned with the solution of $DR(\mathcal{G})$ and is organized as follows. In Section 3.1 $DR(\mathcal{G})$ is reformulated as a minimum distance problem OPT in $l_1^{n_z \times n_w}$ using the parametrizations of stabilizing $\mathcal{C}$s and corresponding $T_{zw}(\mathcal{G},\mathcal{C})$s from Chapter 2. Because it is almost precisely equivalent to $DR(\mathcal{G})$, the rest of the chapter considers OPT. In Section 3.2 it is shown that under some assumptions outlined below a minimizer for OPT always exists and, in some cases, corresponds to a finitely supported $T_{zw}(\mathcal{G},\mathcal{C})$, i.e., an FIR closed loop system. These cases are of interest because they guarantee that if $\hat{G}$ is rational then so is the optimal $\hat{C}$. OPT is in general infinite dimensional, so Sections 3.3 and 3.4 are devoted to its approximate solution. Section 3.3 gives a sequence of finite dimensional infimization problems whose costs approach $\mu_{OPT}(= \mu_{DR})$ from above; its minimizers yield stabilizing $\mathcal{C}$s whose performance approaches optimal as closely as desired. Section 3.4

gives a method for bounding the optimal performance from below by solving finite dimensional infimization problems. Stabilizing $\mathcal{C}$s can also be obtained using this method and, although it is not guaranteed that their performance approaches optimal, it often does in examples. OPT and all the approximating infimization problems formulated in the chapter are equivalent to (infinite or finite) linear programs. Section 3.5 gives detailed formulations of all these linear programs, and Section 3.6 gives a simple example demonstrating some key aspects of both the FIR and lower bound computations. The chapter concludes with a discussion in Section 3.7 of the results, related work, and the main contributions of the chapter.

Notation and Assumptions

In the remainder of this chapter, $\mathcal{G}$ is a given system; the following notational conventions and assumptions are in effect unless otherwise noted.

- $\mathcal{G} \in \Sigma_{clti}$ and satisfies the hypotheses of Lemma 2.3.6 (i.e., G_{yu} has a left or a right coprime factorization over l_1).
- $\mathcal{G}$ is stabilizable ($DR(\mathcal{G})$ is feasible only if this is true).

The first assumption above allows the parametrization of stabilizing $\mathcal{C}$s and corresponding $T_{zw}(\mathcal{G},\mathcal{C})$s in Propositions 2.3.9 and 2.3.11, respectively, to be used. It is a standard assumption, and has been shown to be necessary for stabilizability in the case of $\boldsymbol{H}_\infty$ [19].

H, U, and V denote the matrices over l_1 derived from $\mathcal{G}$ as in Proposition 2.3.9 (using *arbitrary* coprime factorizations of G_{yu}). Although Proposition 2.3.10 shows that H, U, and V have no special structure all results, except where noted, require the following assumption.

Assumption 3.0.1 *U and V have decompositions of the form*

$$U = U_L \Sigma_U U_R \qquad V = V_L \Sigma_V V_R \tag{3.1}$$

where

- $\Sigma_U \in l_1^{r_U \times r_U}$, $\Sigma_V \in l_1^{r_V \times r_V}$ *are diagonal and nonsingular,*

- *$U_L, V_L \in l_1$ are left invertible in l_1, and $U_R, V_R \in l_1$ are right invertible in l_1*

Existence of a decomposition of the above form is equivalent to existence of the Smith form over l_1 [17, Appendix B]. The Smith form over a Bezout domain is guaranteed but l_1 is only Hermite [25], so the assumption poses some restriction. It is not a serious practical restriction, however, because if $\hat{G}$ and hence $\hat{U}$ and $\hat{V}$ are rational then U and V have Smith forms over the finitely supported sequences. Since the finitely supported matrices are in l_1, this satisfies the assumption.

U_L, Σ_U, U_R, V_L, Σ_V, and V_R denote the factors in *arbitrary* decompositions of U and V of the form specified in Assumption 3.0.1. Because l_1 is Hermite, there exist additional matrices over l_1 satisfying the Bezout equations

$$\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} \begin{bmatrix} U_L & U_L^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} V_R \\ V_R^c \end{bmatrix} \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{3.2}$$

$$\begin{bmatrix} U_R \\ U_R^c \end{bmatrix} \begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} V_L^{-L} \\ V_L^{\perp} \end{bmatrix} \begin{bmatrix} V_L & V_L^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{3.3}$$

In the remainder of the chapter, the additional symbols above will denote *arbitrary* choices satisfying the equations, given U_L, U_R, V_L, and V_R.

For each $(i,j) \in \{1,\ldots,r_U\} \times \{1,\ldots,r_V\}$ the set of zeros of either $(\hat{\Sigma}_U)_{ii}$ or $(\hat{\Sigma}_V)_{jj}$ in the closed unit disk is denoted by $\boldsymbol{Z}_{ij}$, and for each $z_0 \in \boldsymbol{Z}_{ij}$ the sum of its multiplicities as a zero of $(\hat{\Sigma}_U)_{ii}$ and $(\hat{\Sigma}_V)_{jj}$ is denoted by $m_{ij}(z_0)$. That is,

$$\begin{aligned} \boldsymbol{Z}_{ij} &:= \left\{ z \in \overline{\mathbf{D}} : [(\hat{\Sigma}_U)_{ii}(z)][(\hat{\Sigma}_V)_{jj}(z)] = 0 \right\} \\ m_{ij}(z_0) &:= \text{multiplicity of the zero } z_0 \text{ in } (\hat{\Sigma}_U)_{ii}(\hat{\Sigma}_V)_{jj}. \end{aligned}$$

The set of all zeros of any element of either $\hat{\Sigma}_U$ or $\hat{\Sigma}_V$ in the closed unit disk is denoted by $\boldsymbol{Z}$, and the total of the multiplicities of all zeros of any element of $\hat{\Sigma}_U$ or $\hat{\Sigma}_V$ is denoted by m_t. That is,

$$\begin{aligned} \boldsymbol{Z} &:= \bigcup_{i=1}^{r_U} \bigcup_{j=1}^{r_V} \boldsymbol{Z}_{ij} \\ m_t &:= \sum_{i=1}^{r_U} \sum_{j=1}^{r_V} \sum_{z_0 \in \boldsymbol{Z}_{ij}} m_{ij}(z_0). \end{aligned}$$

The index set

$$\boldsymbol{S} := \{\{1,\ldots,n_z\} \times \{1,\ldots,n_w\}\} \setminus \{\{1,\ldots,r_U\} \times \{1,\ldots,r_V\}\}.$$

will be convenient notationally; it contains the indices of all elements of an $n_z \times n_w$ matrix *except* those of the upper left hand corner block of dimension $r_U \times r_V$.

Finally, the following assumption is crucial; it says that both U and V have only finitely many zeros in the open unit disk and that neither U nor V has any zeros on the unit circle.

Assumption 3.0.2 *$Z \subset \mathbf{D}$ and $m_t < \infty$.*

3.1 Formulation as a Minimum Distance Problem

The parametrization of stabilizing $\mathcal{C}$s and corresponding $T_{zw}(\mathcal{G},\mathcal{C})$s in Propositions 2.3.9 and 2.3.11 suggest that $DR(\mathcal{G})$ is closely related to the problem of finding points in a certain subspace of l_1 which minimize the distance to a given point. In particular, since H, U, and V are all in l_1, an optimization problem

$$OPT(H,U,V): \qquad \inf\left\{\|H-K\|_{l_1} \ : \ K \in \boldsymbol{K}(U,V)\right\} =: \mu_{OPT}$$

can be defined, where

$$\boldsymbol{K}(U,V) := \left\{K \in l_1^{n_z \times n_w} : \exists Q \in l_1^{n_u \times n_y} \text{ satisfying } K = UQV\right\} \tag{3.4}$$

is easily verified to be a subspace of $l_1^{n_z \times n_w}$.

$OPT(H,U,V)$ is essentially equivalent to $DR(\mathcal{G})$, as the next theorem shows, and the remainder of the chapter is concerned with its solution.

Theorem 3.1.1 *$DR(\mathcal{G})$ and $OPT(H,U,V)$ are equivalent in the following sense:*

1. *If $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$ and $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ is constructed from it as in Proposition 2.3.11 then $K := UQV \in \boldsymbol{K}(U,V)$, $H - K = T_{zw}(\mathcal{G},\mathcal{C})$, and hence $\|H-K\|_{l_1} = \|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_1}$.*

2. *If $K \in \boldsymbol{K}(U,V)$ and $Q \in l_1$ solves $K = UQV$ then*

   (a) if $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ and $\mathcal{C}$ is constructed from it as in Proposition 2.3.11 then $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$, $T_{zw}(\mathcal{G},\mathcal{C}) = H - K$, and hence $\|T_{zw}(\mathcal{G},\mathcal{C})\|_{l_1} = \|H-K\|_{l_1}$.

   (b) if $Q \notin \boldsymbol{Q}(\mathcal{G}_{yu})$ then, for any $\epsilon > 0$, there exists $Q_\epsilon \in \boldsymbol{Q}(\mathcal{G}_{yu})$ such that $\|U\|_{l_1}\|Q-Q_\epsilon\|_{l_1}\|V\|_{l_1} \le \epsilon$. If $\mathcal{C}_\epsilon$ is constructed from it as in Proposition 2.3.11 then $\mathcal{C}_\epsilon \in \boldsymbol{\mathcal{C}}(\mathcal{G})$, $T_{zw}(\mathcal{G},\mathcal{C}) = H - UQ_\epsilon V$, and $\|T_{zw}(\mathcal{G},\mathcal{C}_\epsilon)\|_{l_1} \le \|H-K\|_{l_1} + \epsilon$.

3. $\mu_{DR} = \mu_{OPT}$.

Proof: For item 1, the first conclusion follows from the definition of $\boldsymbol{K}(U,V)$ and the second from Proposition 2.3.11.

For items 2a and 2b, the existence of a $Q \in l_1^{n_u \times n_y}$ is guaranteed by the definition of $\boldsymbol{K}(U,V)$. For item 2a, both conclusions follow from Proposition 2.3.11. For item 2b, the first statement follows from Lemma 2.3.8. For the second statement, the first two conclusions follow from Proposition 2.3.11 and the third from Fact 2.1.8 and the fact that

$$\begin{aligned}
\|H - UQ_\epsilon V\|_{l_1} &= \|H - UQV + U(Q - Q_\epsilon)V\|_{l_1} \leq \|H - K\|_{l_1} + \|U(Q - Q_\epsilon)V\|_{l_1} \\
&\leq \|H - K\|_{l_1} + \|U\|_{l_1} \|Q - Q_\epsilon\|_{l_1} \|V\|_{l_1} \\
&\leq \|H - K\|_{l_1} + \epsilon
\end{aligned}$$

Item 3 is immediate from the preceding items. □

Among the consequences of Theorem 3.1.1 are the following:

- To each feasible solution for $DR(\mathcal{G})$ there corresponds a feasible solution for $OPT(H,U,V)$ of the same cost. From every feasible solution $K \in \boldsymbol{K}(U,V)$ for $OPT(H,U,V)$, one or more corresponding feasible solutions $\mathcal{C} \in \boldsymbol{\mathcal{C}}(\mathcal{G})$ for $DR(\mathcal{G})$ can be constructed of cost arbitrarily close to the same. However, given a feasible solution $K \in \boldsymbol{K}(U,V)$, a corresponding feasible solution for $DR(\mathcal{G})$ of the exactly the same cost can be constructed if and only if there exists $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ such that $K = UQV$.

- To each minimizer for $DR(\mathcal{G})$ there corresponds a minimizer for $OPT(H,U,V)$; hence *all* minimizers for $DR(\mathcal{G})$ can be found from minimizers for $OPT(H,U,V)$ via the constructions of items 2a and 2b. However, given a minimizer K_0 for $OPT(H,U,V)$, a corresponding minimizer for $DR(\mathcal{G})$ can be constructed if and only if there exists $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ such that $K_0 = UQV$.

Although item 2b of Theorem 3.1.1 does not provide a construction for Q_ϵ in the event that $Q \in l_1^{n_u \times n_y} \setminus \boldsymbol{Q}(\mathcal{G}_{yu})$ (recall that the construction in the proof of Lemma 2.3.8 can guarantee only $\det(\tilde{X} - Q\tilde{N}) \in \boldsymbol{U}(l)$ and not necessarily $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$), the fact that $\boldsymbol{Q}(\mathcal{G}_{yu})$ is open and dense in $l_1^{n_u \times n_y}$ guarantees that in this event, given $\epsilon > 0$, $Q_\epsilon := Q + \Delta_\epsilon$ where Δ_ϵ is chosen at random from set $\boldsymbol{\Delta} := \left\{\Delta \in l_1^{n_u \times n_y} : \|U\|_{l_1} \|\Delta\|_{l_1} \|V\|_{l_1} \leq \epsilon\right\}$ will be in $\boldsymbol{Q}(\mathcal{G}_{yu})$ (with probability one under any probability distribution on $\boldsymbol{\Delta}$). Hence a construction is not necessary to produce, given a feasible solution for $OPT(H,U,V)$, feasible solutions for $DR(\mathcal{G})$ of arbitrarily close to the same cost.

Proposition 2.3.12 gives a condition under which the equivalence between $DR(\mathcal{G})$ and $OPT(H,U,V)$ is stronger. It shows that if $\mathcal{G}$ satisfies the hypotheses of Theorem 3.1.1 above and $G_{yu}(0) = 0$ then $\boldsymbol{Q}(\mathcal{G}_{yu}) = l_1^{n_u \times n_y}$ and hence item 2b can be deleted from the theorem statement. In this case, a feasible solution for $DR(\mathcal{G})$ of equal cost can be constructed from *every* feasible solution for $OPT(H,U,V)$; in particular, every minimizer for $OPT(H,U,V)$ yields a minimizer for $DR(\mathcal{G})$.

The next proposition shows how to compute stabilizing compensators for $\mathcal{G}$ given feasible solutions for $OPT(H,U,V)$. According to Theorem 3.1.1, the set of all stabilizing compensators corresponding to a given a feasible solution $K \in \boldsymbol{K}(U,V)$ for $OPT(H,U,V)$ is parameterized by the set of $Q \in l_1^{n_u \times n_y}$ satisfying $K = UQV$. Because of the assumed decompositions of U and V, this set can in turn be parametrized.

Proposition 3.1.2 *If $K \in \boldsymbol{K}(U,V)$ then the set*

$$\boldsymbol{Q}_{UV}(K) := \left\{Q \in l_1^{n_u \times n_y} : K = UQV\right\}$$

is given by

$$\left\{\begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} K V_R^{-R} \Sigma_V^{-1} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} V_L^{-L} \\ V_L^{\perp} \end{bmatrix} : Q_{12}, Q_{21}, Q_{22} \in l_1 \right\}$$

Proof: If $K \in \boldsymbol{K}(U,V)$ then $K = UQ_0V$ for some $Q_0 \in l_1^{n_u \times n_y}$. If Q has the claimed form then $Q \in l_1^{n_u \times n_y}$ since

$$Q = U_R^{-R} U_R Q_0 V_L V_L^{-L} + U_R^{\perp} Q_{12} V_L^{-L} + U_R^{-R} Q_{21} V_L^{\perp} + U_R^{\perp} Q_{22} V_L^{\perp}$$

and all matrices on the right are in l_1. Moreover,

$$\begin{aligned} &UQV = \\ &\quad U_L \Sigma_U U_R (U_R^{-R} U_R Q_0 V_L V_L^{-L} + U_R^{\perp} Q_{12} V_L^{-L} + U_R^{-R} Q_{21} V_L^{\perp} + U_R^{\perp} Q_{22} V_L^{\perp}) V_L \Sigma_V V_R \\ &\quad = U_L \Sigma_U U_R Q_0 V_L \Sigma_V V_R = UQ_0V = K \end{aligned}$$

using the Bezout equations (3.3). Hence $Q \in \boldsymbol{Q}_{UV}(K)$.

Conversely, if $Q \in \boldsymbol{Q}_{UV}(K)$ then $Q \in l_1^{n_u \times n_y}$ and $K = UQV$. Now

$$\begin{aligned}
&K = UQV \\
&\Rightarrow\ K \ = \ U_L \Sigma_U U_R Q V_L \Sigma_V V_R \\
&\Rightarrow\ \Sigma_U^{-1} U_L^{-L} K V_R^{-R} \Sigma_V^{-1} \ = \ U_R Q V_L \\
&\Rightarrow\ \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} K V_R^{-R} \Sigma_V^{-1} & U_R Q V_L^c \\ U_R^c Q V_L & U_R^c Q V_L^c \end{bmatrix} \ = \ \begin{bmatrix} U_R \\ U_R^c \end{bmatrix} Q \begin{bmatrix} V_L & V_L^c \end{bmatrix} \\
&\Rightarrow\ Q = \begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} K V_R^{-R} \Sigma_V^{-1} & U_R Q V_L^c \\ U_R^c Q V_L & U_R^c Q V_L^c \end{bmatrix} \begin{bmatrix} V_L^{-R} \\ V_L^{\perp} \end{bmatrix} \qquad (3.5)
\end{aligned}$$

where the first implication follows from (3.1) and the last by reversing (3.3). That Q has the claimed form follows since Q, U_R, U_R^c, V_L, and V_L^c are matrices over l_1. □

3.2 Existence of a Minimizer

In this section OPT is shown to have a minimizer under only the assumptions outlined at the beginning of the chapter. The first step is an alternate characterization of the feasible subspace $\boldsymbol{K}(U, V)$ (this characterization is also useful in the formulation of OPT as a linear program in Section 3.5). If U and V have full row and column rank, respectively (this corresponds to having more independent measurements than disturbances and more independent controls than errors) then a finitely supported minimizer is shown to exist.

Much use is made in this section of basic results from functional analysis applied to the present setting of matrices over sequence spaces. Appendix B contains a list of notation for standard objects and some basic results on duality and minimum distance problems. Appendix C contains detailed results on normed linear spaces of sequences and their duals and on several classes of operators and their adjoints.

Theorem 3.2.1 $K \in \boldsymbol{K}(U, V)$ *if and only if*

1. $K \in l_1^{n_z \times n_w}$,
2. $\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}$

   where $*$ denotes an irrelevant block, and

3. $[(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j}]^{(n)}(z_0) = 0$ *for each* $(i,j) \in \{1,\ldots,r_U\} \times \{1,\ldots,r_V\}$, *each* $z_0 \in \boldsymbol{Z}_{ij}$, *and each* $n \in \{0,\ldots,m_{ij}(z_0)-1\}$.

Proof: If $K \in \boldsymbol{K}(U,V)$ then by definition condition 1 holds and $K = UQV$ for some $Q \in l_1^{n_u \times n_y}$. Hence

$$\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix} \Sigma_U U_R Q V_L \Sigma_V \begin{bmatrix} I & 0 \end{bmatrix} = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}$$

using (3.1) and (3.2). This establishes condition 2. Also, condition 3 is satisfied since $(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j} = (\hat{\Sigma}_U)_{ii}(\hat{U}_R\hat{Q}\hat{V}_L)_{ij}(\hat{\Sigma}_V)_{jj}$ and U_RQV_L is an l_1 matrix; in particular its entries are all analytic in $\mathbf{D}$.

Conversely, if conditions 1, 2, and 3 hold then $K \in l_1^{n_z \times n_w}$ and, defining $\bar{Q} := \Sigma_U^{-1}U_L^{-L}KV_R^{-R}\Sigma_V^{-1}$, we can compute by condition 3 since

$$\hat{\bar{Q}}_{ij} = \frac{(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j}}{(\hat{\Sigma}_U)_{ii}(\hat{\Sigma}_V)_{jj}}, \qquad (i,j) \in \{1,\ldots,r_U\} \times \{1,\ldots,r_V\}$$

Because $\boldsymbol{Z}_{ij} \subset \mathbf{D}$ for each i and j, the numerator above can be factored

$$(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j} = \prod_{z_0 \in Z_{ij}} (z - z_0)^{m_{ij}(z_0)}\hat{\tilde{Q}}$$

where $\tilde{Q} \in l_1$. Hence $\bar{Q} \in l_1^{r_U \times r_V}$, $Q := U_R^{-R}\bar{Q}V_L^{-L} \in l_1^{n_u \times n_y}$ and

$$UQV = U_LU_L^{-L}KV_R^{-R}V_R = (I - U_L^cU_L^{\perp})K(I - V_R^{\perp}V_R^c) = K$$

where the first equality follows using (3.1) and (3.3), the second by reversing (3.2), and the last using condition 2. Hence $K \in \boldsymbol{K}(U,V)$. □

Two preliminary lemmas will facilitate the proof of the existence theorem; the first characterizes all matrices in $l_1^{n_u \times n_y}$ satisfying condition 2 of Theorem 3.2.1 as the null space of a bounded linear operator from $l_1^{n_z \times n_w}$ to a subspace of $l_1^{n_z \times n_w}$. Moreover, this operator is the adjoint of a bounded linear operator on a subspace of $c_0^{n_z \times n_w}$.

Lemma 3.2.2 *There exists* $\tilde{T}_C \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$ *such that*

$$\boldsymbol{K}_C := \left\{K \in l_1^{n_z \times n_w} : K \text{ satisfies condition 2 of Theorem 3.2.1}\right\} = \mathcal{N}\left(\tilde{T}_C^*\right).$$

Proof: Begin by defining $\tilde{\mathcal{T}}_C$ on c_0^S by

$$\tilde{\mathcal{T}}_C := \tilde{\mathcal{T}}_{U,V}\tilde{\mathcal{E}}_S \tag{3.6}$$

where $\tilde{\mathcal{E}}_S$ is defined, given $G \in c_0^S$,

$$(\tilde{\mathcal{E}}_S G)_{ij} := \begin{cases} G_{ij} & (i,j) \in \boldsymbol{S} \\ 0 & \text{otherwise} \end{cases},$$

and $\tilde{\mathcal{T}}_{U,V}$ is defined, given $G \in c_0^{n_z \times n_w}$,

$$\tilde{\mathcal{T}}_{U,V}K := \begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix}^T \triangleleft K \triangleright \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}^T.$$

The operations $\triangleleft$ and $\triangleright$ of left and right matrix correlation are defined in C.1.5. $\tilde{\mathcal{E}}_S$ is an embedding operator of the type defined in Proposition C.1.6 and $\tilde{\mathcal{T}}_{U,V}$ is a correlation operator of the type defined in Proposition C.1.3. Hence $\tilde{\mathcal{E}}_S \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$ and $\tilde{\mathcal{T}}_{U,V} \in \mathcal{B}(c_0^{n_z \times n_w})$, and, as a result, $\tilde{\mathcal{T}}_C \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$. Moreover, using these two propositions,

$$\tilde{\mathcal{T}}_C^* = \tilde{\mathcal{E}}_S^*\tilde{\mathcal{T}}_{U,V}^* = \Pi_S \mathcal{T}_{UV} = \mathcal{T}_C \tag{3.7}$$

where $\mathcal{T}_{UV}$ is defined, given $K \in l_1^{n_z \times n_w}$,

$$\mathcal{T}_{UV}K := \begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}$$

and Π_S is defined, given $K \in l_1^{n_z \times n_w}$,

$$(\Pi_S K)_{ij} := K_{ij}, \quad (i,j) \in \boldsymbol{S}.$$

It is clear that $\boldsymbol{K}_C = \mathcal{N}(\mathcal{T}_C) = \mathcal{N}\left(\tilde{\mathcal{T}}_C^*\right)$. □

The second lemma characterizes all matrices in $l_1^{n_u \times n_y}$ satisfying condition 3 of Theorem 3.2.1 as the null space of a bounded linear operator from $l_1^{n_z \times n_w}$ to $\mathbf{R}^{m_t}$. Moreover, this operator is the adjoint of bounded linear operator on $\mathbf{R}^{m_t}$.

Lemma 3.2.3 *There exists $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z \times n_w})$ such that*

$$\boldsymbol{K}_I := \left\{K \in l_1^{n_z \times n_w} : K \textit{ satisfies condition 3 of Theorem 3.2.1}\right\} = \mathcal{N}\left(\tilde{\mathcal{T}}_I^*\right).$$

Proof: Begin by defining $\tilde{\mathcal{T}}_I$, given $\alpha \in \mathbf{R}^{m_t}$,

$$\tilde{\mathcal{T}}_I\alpha := \tilde{\mathcal{T}}_{U,V}\sum_{i=1}^{r_U}\sum_{j=1}^{r_V}\left[\sum_{z_0\in Z_{ij}}\sum_{n=0}^{m_{ij}(z_0)-1}\tilde{\mathcal{E}}_{ij}\tilde{\mathcal{D}}^{\Re}_{n,z_0}\alpha_{i,j,n,z_0} + \sum_{z_0\in Z^+_{ij}}\sum_{n=0}^{m_{ij}(z_0)-1}\tilde{\mathcal{E}}_{ij}\tilde{\mathcal{D}}^{\Im}_{n,z_0}\alpha_{i,j,n,z_0}\right] \tag{3.8}$$

where $\tilde{\mathcal{E}}_{ij} : c_0 \mapsto c_0^{n_z\times n_w}$ is defined, given $G \in c_0$,

$$(\tilde{\mathcal{E}}_{ij}G)_{mn} := \begin{cases} G_{ij} & m = i \text{ and } n = j \\ 0 & \text{otherwise} \end{cases}$$

and $\tilde{\mathcal{D}}^{\Re}_{n,z_0}$ and $\tilde{\mathcal{D}}^{\Im}_{n,z_0}$ are defined, given $\alpha \in \mathbf{R}$,

$$(\tilde{\mathcal{D}}^{\Re}_{n,z_0}\alpha)(k) := \begin{cases} 0 & k < n \\ \frac{\alpha k!}{(k-n)!}\Re(z_0^{k-n}) & k \geq n \end{cases}$$

and

$$(\tilde{\mathcal{D}}^{\Im}_{n,z_0}\alpha)(k) := \begin{cases} 0 & k < n \\ \frac{\alpha k!}{(k-n)!}\Im(z_0^{k-n}) & k \geq n \end{cases}.$$

Note that for each i and j $\tilde{\mathcal{E}}_{ij}$ is an embedding operator of the type defined in Proposition C.1.6 and that for each z_0 and n $\tilde{\mathcal{D}}^{\Re}_{n,z_0}$ and $\tilde{\mathcal{D}}^{\Im}_{n,z_0}$ are operators of the type defined on $\mathbf{R}$ in Proposition C.1.9. Hence all the $\tilde{\mathcal{E}}_{ij} \in \mathcal{B}(c_0, c_0^{n_z\times n_w})$ and all the $\tilde{\mathcal{D}}^{\Re}_{n,z_0}, \tilde{\mathcal{D}}^{\Im}_{n,z_0} \in \mathcal{B}(\mathbf{R}, c_0)$. The first item of Proposition B.1.2 shows that an operator defined as a a finite sum of component operators each of which is bounded (as $\tilde{\mathcal{T}}_I$ is in (3.8)) is itself bounded. Hence $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z\times n_w})$. Moreover, by the second item of Proposition B.1.2, the adjoint of an operator defined in this way can be written as the cartesian product of the adjoints of the component operators. Using this fact and using Propositions C.1.6 and C.1.9 to find the adjoints of the components, it is straightforward to check that

$$\tilde{\mathcal{T}}_I^* = \mathcal{T}_I \tag{3.9}$$

where $\mathcal{T}_I$ is defined, given $K \in l_1^{n_z\times n_w}$,

$$\mathcal{T}_I K := \prod_{i=1}^{r_U}\prod_{j=1}^{r_V}\left\{\prod_{z_0\in Z_{ij}}\prod_{n=0}^{m_{ij}(z_0)-1}\left\{\mathcal{D}^{\Re}_{n,z_0}\Pi_{ij}\mathcal{T}_{UV}K\right\}\right\}\times\left\{\prod_{z_0\in Z^+_{ij}}\prod_{n=0}^{m_{ij}(z_0)-1}\left\{\mathcal{D}^{\Im}_{n,z_0}\Pi_{ij}\mathcal{T}_{UV}K\right\}\right\}$$

where $Z^+_{ij} := \{z_0 \in Z_{ij} : \Im z_0 > 0\}$, $\mathcal{T}_{UV}$ is defined in the proof of Lemma 3.2.2, Π_{ij} is defined, given $K \in l_1^{n_z\times n_w}$,

$$\Pi_{ij}K := K_{ij}$$

and $\mathcal{D}^{\Re}_{n,z_0}$ and $\mathcal{D}^{\Im}_{n,z_0}$ are defined, given $K \in l_1$,

$$\mathcal{D}^{\Re}_{n,z_0} := \sum_{k=n}^{\infty} K(k)\frac{k!}{(k-n)!}\Re(z_0^{k-n}) \text{ and } \mathcal{D}^{\Im}_{n,z_0} := \sum_{k=n}^{\infty} K(k)\frac{k!}{(k-n)!}\Im(z_0^{k-n}).$$

To see that $\boldsymbol{K}_I = \mathcal{N}(\mathcal{T}_I) = \mathcal{N}\left(\tilde{\mathcal{T}}_I^*\right)$ note that for each i, j, z_0, and n,

$$\mathcal{D}^{\Re}_{n,z_0}\Pi_{ij}\mathcal{T}_{UV}K = \Re\left\{[(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j}]^{(n)}(z_0)\right\}$$

and

$$\mathcal{D}^{\Im}_{n,z_0}\Pi_{ij}\mathcal{T}_{UV}K = \Im\left\{[(\hat{U}_L^{-L})_{i\cdot}\hat{K}(\hat{V}_R^{-R})_{\cdot j}]^{(n)}(z_0)\right\}$$

and use the fact that, for any $f \in l_1$, $n \in \mathbf{Z}_+$, and $z_0 \in \mathbf{D}$, $\overline{\hat{f}^{(n)}(z_0)} = \hat{f}^{(n)}(\overline{z}_0)$, where $\overline{z}$ denotes the complex conjugate of $z \in \mathbf{C}$ (this is why we can consider only the zero with a positive imaginary part in each complex conjugate pair in $\boldsymbol{Z}_{ij}$). □

Theorem 3.2.4 *$OPT(H,U,V)$ has a minimizer.*

Proof: If a bounded linear operator $\tilde{\mathcal{T}}_K$ can be constructed such that $\boldsymbol{K}(U,V) = \mathcal{N}\left(\tilde{\mathcal{T}}_K^*\right)$ then $\boldsymbol{K}(U,V)$ is *weak**-closed since it is the null space of the adjoint of a bounded linear operator [26, Theorem 4.12]. Corollary B.2.4 states that every minimum distance problem whose feasible subspace is *weak**-closed has a minimizer and hence $OPT(H,U,V)$ has a minimizer.

The construction is simple given Lemmas 3.2.2 and 3.2.3; define $\tilde{\mathcal{T}}_K$, given $(G,\alpha) \in c_0^S \times \mathbf{R}^{m_t}$,

$$\tilde{\mathcal{T}}_K\left[(G,\alpha)\right] := \tilde{\mathcal{T}}_C G + \tilde{\mathcal{T}}_I \alpha \tag{3.10}$$

where $\tilde{\mathcal{T}}_C$ and $\tilde{\mathcal{T}}_I$ are the operators guaranteed by Lemma 3.2.2 and Lemma 3.2.3, respectively. By Fact B.1.1, the space $c_0^S \times \mathbf{R}^{m_t}$ is a normed linear space under the norm $\|(G,\alpha)\| := \left\|(\|G\|_{l_\infty}, \|\alpha\|_p)\right\|_p$ where $1 \le p \le \infty$. The choice of p is not important to this proof; what is of interest is that for any p, again by Fact B.1.1, $\left(c_0^S \times \mathbf{R}^{m_t}\right)^* = l_1^S \times \mathbf{R}^{m_t}$ with the norm $\left\|(\|G\|_{l_\infty}, \|\alpha\|_q)\right\|_q$, where the q-norm is the conjugate of the p-norm.

Recall that $\tilde{\mathcal{T}}_C \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$ and $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z \times n_w})$. Hence, by the first item of Proposition B.1.2, $\tilde{\mathcal{T}}_K \in \mathcal{B}(c_0^S \times \mathbf{R}^{m_t}, c_0^{n_z \times n_w})$. Using the second item of Proposition B.1.2 we find that $\tilde{\mathcal{T}}_K^* \in \mathcal{B}(l_1^{n_z \times n_w}, l_1^S \times \mathbf{R}^{m_t})$ can be written, given $K \in l_1^{n_z \times n_w}$,

$$\tilde{\mathcal{T}}_K^* K = (\tilde{\mathcal{T}}_C^* K, \tilde{\mathcal{T}}_I^* K). \tag{3.11}$$

Clearly $\boldsymbol{K}(U,V) = \boldsymbol{K}_C \cap \boldsymbol{K}_I = \mathcal{N}\left(\tilde{T}_C^*\right) \cap \mathcal{N}\left(\tilde{T}_I^*\right) = \mathcal{N}\left(\tilde{T}_K^*\right)$, and the proof is complete. □

As a consequence of Theorem 3.2.4, $DR(\mathcal{G})$ has a minimizer provided that there exists $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ satisfying $K_0 = UQV$ for some minimizer K_0 for $OPT(H,U,V)$. Proposition 2.3.12 leads to the following corollary.

Corollary 3.2.5 *If $G_{yu}(0) = 0$ then $DR(\mathcal{G})$ has a minimizer.*

Theorem 3.2.7 will show that if U and V have full row rank and full column rank, respectively, then there exists a minimizer K_{fs} for $OPT(H,U,V)$ which is such that $H - K_{\text{fs}}$ is finitely supported; its proof is facilitated by the following lemma.

Lemma 3.2.6 *If $X \in l_1^{m\times n}$, $\boldsymbol{M}$ is a subspace of $l_1^{m\times n}$, $\boldsymbol{M}^\perp \subset c_0^{m\times n}$, and*

$$\|X - M_0\|_{l_1} = \inf\left\{\|X - M\|_{l_1} : M \in \boldsymbol{M}\right\}$$

then at least one row of $X - M_0$ is finitely supported.

Proof: If M_0 is such that $\|X - M_0\|_{l_1} = \inf\left\{\|X - M\|_{l_1} : M \in \boldsymbol{M}\right\}$ then, by the duality theorem B.2.2, there exists $G_0 \in \boldsymbol{M}^\perp$ with $\|G_0\|_{l_\infty} = 1$ such that M_0 and G_0 are aligned (Definition B.2.1). If $\boldsymbol{M}^\perp \subset c_0^{m\times n}$, then $G_0 \in c_0^{m\times n}$. Moreover, since $G_0 \neq 0$ there is a row, say the i-th, such that $\gamma_i := \max\left\{\|(G_0)_{ij}\|_\infty : j \in \{1,\ldots,n\}\right\} > 0$. Since $G_0 \in c_0^{m\times n}$, there exists N such that $|(G_0)_{ij}(k)| < \gamma_i$ for each $j \in \{1,\ldots,n\}$ and all $k > N$. Alignment of G_0 and $X - M_0$ then implies that $(X - M_0)_{ij}(k) = 0$ for each $j \in \{1,\ldots,n\}$ and all $k > N$ [5, Fact ?], i.e., that the i-th row of $X - M_0$ consists entirely of finitely supported sequences. □

Theorem 3.2.7 *If U and V have full row rank and full column rank, respectively, then every minimizer K_0 for $OPT(H,U,V)$ is such that at least one row of $H - K_0$ is finitely supported. Moreover, there exists a minimizer K_{fs} such that all entries of $H - K_{\text{fs}}$ are finitely supported.*

Proof: For the first statement, note that if U and V satisfy the hypotheses of Theorem 3.2.1 and the stated rank hypotheses, condition 2 in Theorem 3.2.1 is satisfied vacuously and hence $\boldsymbol{K}(U,V) = \boldsymbol{K}_I$, defined in Lemma 3.2.3. By that lemma, $\boldsymbol{K}(U,V) = \mathcal{N}\left(\tilde{T}_I^*\right)$ and hence

$$\boldsymbol{K}(U,V)^\perp = \mathcal{N}\left(\tilde{T}_I^*\right)^\perp = \left[{}^\perp\mathcal{R}\left((\tilde{T}_I^*)^*\right)\right]^\perp = \mathcal{R}\left((\tilde{T}_I^*)^*\right) = \mathcal{R}\left(\tilde{T}_I\right) \subset c_0^{n_z\times n_w} \quad (3.12)$$

where $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z \times n_w})$ is defined in the proof of Lemma 3.2.3. The second equality holds because $\mathcal{N}(\mathcal{T}) =^{\perp} \mathcal{R}(\mathcal{T}^*)$ for every bounded linear operator [26, Theorem 4.12]. $\mathcal{R}\left(\tilde{\mathcal{T}}_I^*\right)$ is closed because it is a subspace of the finite dimensional space $\mathbf{R}^{m_t}$ and hence it is also *weak**-closed [26, Theorem 4.14]. The third equality then follows because the *weak**-closure of any subspace is the right annihilator of its left annihilator [26, Theorem 4.7]. The last equality holds by applying Propositions C.1.9, C.1.6, C.1.5 and B.1.2 to compute $(\tilde{\mathcal{T}}_I^*)^*$; it is equal to $\tilde{\mathcal{T}}_I$. (Although the adjoint of a map in $\mathcal{B}(l_1^{n_z \times n_w}, \mathbf{R}^{m_t})$ in general maps $\mathbf{R}^{m_t}$ into $l_1^{n_z \times n_w}$, $(\tilde{\mathcal{T}}_I^*)^*$ maps $\mathbf{R}^{m_t}$ into $c_0^{n_z \times n_w}$ because $\boldsymbol{Z} \subset \mathbf{D}$.) Finally, by Lemma 3.2.6, every minimizer K^0 for $OPT(H, U, V)$ is such that $H - K^0$ has at least one finitely supported row.

For the second statement, note that Theorem 3.2.4 guarantees the existence of a minimizer K^1 for $OPT(H, U, V)$. By the argument above, at least one row of $H - K^1$ is finitely supported. Next we show that, given $p \in \{1, \ldots, n_z\}$ and any minimizer K^p such that at least p rows of $H - K^p$ are finitely supported, there exists a minimizer K^{p+1} such that at least $p+1$ rows of $H - K^{p+1}$ are finitely supported. Hence, there is at least one minimizer K_{fs} such that $H - K_{\text{fs}}$ is finitely supported.

Suppose, then, that $p \in \{1, \ldots, n_z\}$ and K^p is any minimizer for $OPT(H, U, V)$ such that at least p rows (assumed, without loss of generality, to be the first p) of $H - K^p$ are finitely supported. Partition after the p-th row,

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}, \qquad K^p = \begin{bmatrix} K_1^p \\ K_2^p \end{bmatrix},$$

and consider the problem

$$OPT': \qquad \inf\left\{\|(H_2 - K_2^p) - K'\|_{l_1} : K' \in \boldsymbol{K}'(K^p, U, V)\right\}$$

where

$$\boldsymbol{K}'(K^p, U, V) := \left\{K' \in l_1^{(n_z - p) \times n_w} : \begin{bmatrix} K_1^p \\ K_2^p + K' \end{bmatrix} \in \boldsymbol{K}(U, V)\right\}.$$

It is easy to check that any minimizer for OPT' is also a minimizer for $OPT(H, U, V)$.

Moreover, $K^p \in \boldsymbol{K}(U, V)$ implies, together with Theorem 3.2.1 and the rank assumptions on U and V, that $K' \in \boldsymbol{K}'(K^p, U, V)$ if and only if $K' \in l_1^{(n_z - p) \times n_w}$ and K' satisfies

$$[(\hat{U}_L^{-L})_{i\cdot}^2 \hat{K}'(\hat{V}_R^{-R})_{\cdot j}]^{(k)}(z_0) = 0, \qquad k \in \{0, \ldots, m_{ij}(z_0) - 1\}$$

for each $(i,j) \in \{1,\ldots,r_U\} \times \{1,\ldots,r_V\}$ and each $z_0 \in \boldsymbol{Z}_{ij}$, where $(U_L^{-L})^2$ denotes the last $n_z - p$ columns of U_L^{-L}.

Thus, arguing as in Lemma 3.2.3, there exists $\mathcal{T} \in \mathcal{B}(l_1^{(n_z-p)\times n_w}, \mathbf{R}^{m_t})$ such that $\boldsymbol{K}'(K^p, U, V) = \mathcal{N}(\mathcal{T})$. Moreover, arguing as in the proof of Theorem 3.2.4, there exists $\tilde{\mathcal{T}} \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{(n_z-p)\times n_w})$ such that $\tilde{\mathcal{T}}^* = \mathcal{T}$. Hence $\boldsymbol{K}'(K^p, U, V)$ is *weak**-closed and CorollaryB.2.4 guarantees the existence of a minimizer $(K')^1$ for OPT'. Arguing as for the proof of the first statement above, $\boldsymbol{K}'(K^p, U, V)^\perp \subset c_0^{(n_z-p)\times n_w}$ so that, by Lemma 3.2.6, $(K')^1$ is such that at least one row of $H - K_2^p - (K')^1$ is finitely supported. Finally, then,

$$K^{p+1} = \begin{bmatrix} K_1^{p+1} \\ K_2^{p+1} \end{bmatrix} := \begin{bmatrix} K_1^p \\ K_2^p + (K')^1 \end{bmatrix}$$

is a minimizer for $OPT(H, U, V)$ and at least $p+1$ rows of $H - K^{p+1}$ are finitely supported. □

3.3 Suboptimal Solutions via FIR Approximation

$OPT(H, U, V)$ is in general an infinite dimensional problem (it is shown in Section 3.5 to be equivalent to a linear program with infinitely many variables and constraints). As a result, only approximate solutions can be computed. In this section, a sequence of finite dimensional minimization problems is formulated which can be used under certain conditions to compute feasible solutions for $OPT(H, U, V)$ which are arbitrarily close to optimal.

The idea is to consider only feasible solutions $K \in \boldsymbol{K}(U, V)$ for which $H - K$ is finitely supported; this corresponds to considering only $\mathcal{C}$s for which $T_{zw}(\mathcal{G}, \mathcal{C})$ is finitely supported, i.e., for which the closed loop system is FIR. Each problem corresponds to allowing closed loop impulse responses of a fixed length; increasing the length approximates the optimal solution.

Of course it is not always possible to choose a stabilizing controller which gives an FIR closed loop system. A test for this is given in Proposition 3.3.2. On the other hand, Theorem 3.2.7 showed that when U and V have full row rank and full column rank, respectively, the minimizer for OPT corresponds to an FIR closed loop system. Hence, the method of this section can be used to find such a minimizer.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\overline{OPT}_n(H,U,V) \qquad\qquad \inf\left\{\|H-K\|_{l_1} : K \in \overline{\boldsymbol{K}}_n(H,U,V)\right\} =: \overline{\mu}_n$$

where

$$\overline{\boldsymbol{K}}_n(H,U,V) := \{K \in \boldsymbol{K}(U,V) : \text{supp } (H-K) \subset \{0,\ldots,n\}\}.$$

Note that $\overline{\boldsymbol{K}}_n(H,U,V)$ is a subset and not a subspace of $l_1^{n_z \times n_w}$.

Proposition 3.3.1 *If there exists N such that $\overline{\boldsymbol{K}}_N(H,U,V) \neq \emptyset$ and the finitely supported matrices are dense in $\boldsymbol{K}(U,V)$ then*

$$\{\overline{\mu}_n\}_{n=N}^{\infty} \searrow \mu_{OPT} \ \textit{as } n \to \infty.$$

Proof: Note first that $n_1 > n_2 \geq N$ implies $\overline{\boldsymbol{K}}_{n_1}(H,U,V) \supset \overline{\boldsymbol{K}}_{n_2}(H,U,V) \supset \overline{\boldsymbol{K}}_N(H,U,V)$. Hence, if $\overline{\boldsymbol{K}}_N(H,U,V) \neq \emptyset$ then $\overline{\mu}_n$ is well defined for every $n \geq N$ and $\{\overline{\mu}_n\}_{n=N}^{\infty}$ is clearly monotonically non-decreasing.

To show convergence, note first that if $K_N \in \overline{\boldsymbol{K}}_N(H,U,V)$ then, by Lemma B.2.5,

$$\begin{aligned} \mu_{OPT} &:= \inf\left\{\|H-K\|_{l_1} : K \in \boldsymbol{K}(U,V)\right\} \\ &= \inf\left\{\|H_N-K\|_{l_1} : K \in \boldsymbol{K}(U,V)\right\} \end{aligned} \tag{3.13}$$

where $H_N := H - K_N$ is finitely supported. Now let $\epsilon > 0$ be given. In view of (3.13), we can choose $K^\epsilon \in \boldsymbol{K}(U,V)$ such that $\|H_N - K^\epsilon\|_{l_1} \leq \mu_{OPT} + \frac{\epsilon}{2}$. By the density hypothesis, there exist $\bar{N}$ and $K^\epsilon_{\bar{N}} \in \boldsymbol{K}(U,V)$ such that supp $(K^\epsilon_{\bar{N}}) \subset \left\{0,\ldots,\bar{N}\right\}$ and $\left\|K^\epsilon - K^\epsilon_{\bar{N}}\right\|_{l_1} \leq \frac{\epsilon}{2}$. Finally, defining $N^* := \max\left\{N,\bar{N}\right\}$, note that $K^\epsilon_{\bar{N}} \in \overline{\boldsymbol{K}}_{N^*}(H,U,V)$ so that

$$\begin{aligned} \overline{\mu}_{N^*} &\leq \|H_N - K^\epsilon_{\bar{N}}\|_{l_1} = \|(H_N - K^\epsilon) + (K^\epsilon - K^\epsilon_{\bar{N}})\|_{l_1} \\ &\leq \|H_N - K^\epsilon\|_{l_1} + \|K^\epsilon - K^\epsilon_{\bar{N}}\|_{l_1} \\ &\leq \mu_{OPT} + \epsilon \end{aligned}$$

□

If $\hat{U}$ and $\hat{V}$ are rational then the finitely supported matrices are dense in $\boldsymbol{K}(U,V)$; whether the standing assumptions $\boldsymbol{Z} \subset \mathbf{D}$ and $m_t < \infty$ are sufficient to guarantee this is unclear.

The next proposition gives necessary and sufficient conditions for the existence of a feasible solution K such that $H-K$ is finitely supported when decompositions of U and V as in Assumption 3.0.1 exist such that Bezout equations (3.2) can be constructed entirely with finitely supported matrices. This is possible whenever $\hat{U}$ and $\hat{V}$ are rational; the Smith forms of U and V over the finitely supported matrices suffice.

Proposition 3.3.2 *Suppose there exists a decomposition of U and V as in Assumption 3.0.1 such that all matrices in the Bezout equations (3.2) can be chosen to have finite support. There exists n such that $\overline{\boldsymbol{K}}_n(H,U,V) \neq \emptyset$ if and only if*

$$\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} H \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} \right)_{ij}$$

is finitely supported for all $(i,j) \in \boldsymbol{S}$, where the matrices satisfying (3.2) have been chosen to be finitely supported.

Proof: For the "only if" part of the second statement, let $K_n \in \overline{\boldsymbol{K}}_n(H,U,V)$ for some n, and let $(i,j) \in \boldsymbol{S}$. Then $(H-K_n)_{ij}$ is finitely supported. Also

$$\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} H \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} \right)_{ij} = \left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} (H-K_n) \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} \right)_{ij}.$$

The latter is finitely supported because all matrices in the expression are, and the convolution of finitely supported sequences is finitely supported.

For the "if" part, begin by choosing a finitely supported matrix T_{fs} such that for each $i \in \{1,\ldots,r_U\}$ and $j \in \{1,\ldots,r_V\}$ $(\hat{T}_{\text{fs}} - \hat{U}_L^{-L}\hat{H}\hat{V}_R^{-R})$ has all the zeros of $(\hat{\Sigma}_U)_{ii}$ and $(\hat{\Sigma}_V)_{jj}$ in $\mathbf{D}$ (including multiplicities). This is equivalent to finding a matrix of polynomials satisfying a finite set of interpolation conditions and is hence always possible. With this choice,

$$Q_{\text{fs}} := U_R^{-R}\Sigma_U^{-1}(T_{\text{fs}} - U_L^{-L}HV_R^{-R})\Sigma_V^{-1}V_L^{-L} \in l_1$$

and if $K_{\text{fs}} := UQ_{\text{fs}}V$ then $K_{\text{fs}} \in \boldsymbol{K}(U,V)$ and

$$\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} (H-K_{\text{fs}}) \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} = \begin{bmatrix} T_{\text{fs}} & U_L^{-L}HV_R^{\perp} \\ U_L^{\perp}HV_R^{-R} & U_L^{\perp}HV_R^{\perp} \end{bmatrix} =: T'_{\text{fs}}$$

is finitely supported because the matrices satisfying (3.2) are and because of the hypothesis on H. Moreover

$$H - K_{\mathrm{fs}} = \begin{bmatrix} U_L & U_L^c \end{bmatrix} T'_{\mathrm{fs}} \begin{bmatrix} V_R \\ V_R^c \end{bmatrix}$$

is then finitely supported as well. □

3.4 A Converging Lower Bound on μ_{DR}

The sequence of FIR approximations defined in Section 3.3 approximates OPT sub-optimally. As a result, its infimal costs constitute a sequence of upper bounds converging to μ_{DR}; its minimizers yield stabilizing $\mathcal{C}$s whose performance approaches optimal. There is no information on the rate of this convergence, so a corresponding scheme for super-optimal approximation is needed to determine μ_{DR} with certainty.

In this section, a sequence of super-optimal approximating problems is formulated and their costs shown to converge to μ_{DR} from below. They are not finite dimensional problems, however it is shown that each has a finitely supported minimizer. Hence each problem in the sequence can in turn be approximated from above by a *computable* sequence much like that of Section 3.3, each problem corresponding to a fixed "length" of the support. It is crucial for the success of this scheme that these finitely supported minimizers exist; otherwise an individual problem's cost cannot be guaranteed to bound μ_{DR} either above or below. The results allow a double iteration consisting of finite dimensional problems to be carried out to bound μ_{DR} below to any desired degree of accuracy; the process is described at the end of the section.

An important feature of these problems is that they are always feasible, unlike the FIR approximations, and, while their feasible solutions are not feasible for OPT, feasible solutions for OPT can be computed from them. This yields stabilizing compensators for $\mathcal{G}$; the implications of this are shown in Section 3.6 and discussed in Section 3.7.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\underline{OPT}_n(H,U,V) \qquad \inf\left\{\|H-K\|_{l_1} : K \in \underline{\mathbf{K}}_n(U,V)\right\} =: \underline{\mu}_n$$

where

$$K \in \underline{\boldsymbol{K}}_n(U,V) \iff K \in l_1^{n_z \times n_w}, \quad \text{K satisfies condition 3 of Theorem 3.2.1, and} \quad \left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} \right)(k) = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} \quad \text{for all } k \in \{0,\ldots,n\} \tag{3.14}$$

Note that $\underline{\boldsymbol{K}}_n(U,V)$ is a subspace of $\boldsymbol{K}(U,V)$.

The following theorem shows that $\underline{OPT}_n(H,U,V)$ is feasible for every n and that its infimal costs form a non-decreasing sequence which converges to μ_{DR} from below, and that each problem has a minimizer.

Theorem 3.4.1 *$\{\underline{\mu}_n\}_{n=0}^{\infty} \nearrow \mu_{DR}$ as $n \to \infty$. Moreover, for each n, $\underline{OPT}_n(H,U,V)$ has a minimizer.*

Proof: $\{\underline{\mu}_n\}_{n=0}^{\infty}$ is well defined because, for any n, $0 \in \underline{\boldsymbol{K}}_n(U,V)$ and, for any $K \in \underline{\boldsymbol{K}}_n(U,V)$, $\|H-K\|_{l_1} \geq 0$. Also, $\underline{\mu}_n \leq \mu_{DR}$ for each n and the sequence is non-decreasing since $n_1 > n_2$ implies $\underline{\boldsymbol{K}}_{n_2}(U,V) \supset \underline{\boldsymbol{K}}_{n_1}(U,V) \supset \boldsymbol{K}(U,V)$.

To show convergence, begin by defining for each n a map $\underline{\tilde{\mathcal{T}}}_n$, given $(G,\alpha) \in c_0^S \times \mathbf{R}^{m_t}$,

$$\underline{\tilde{\mathcal{T}}}_n\left[(G,\alpha)\right] := \tilde{\mathcal{T}}_C \tilde{\mathcal{P}}_n G + \tilde{\mathcal{T}}_I \alpha \tag{3.15}$$

where $\tilde{\mathcal{T}}_C \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$ is defined in (3.6) and $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z \times n_w})$ in (3.8), in the proofs of Lemmas 3.2.2 and 3.2.3, respectively. $\tilde{\mathcal{P}}_n$ is the n-th truncation operator restricted to c_0^S and is bounded by Proposition C.1.7. Proposition B.1.2 shows that an operator defined as a finite sum of component operators each of which is bounded (as $\underline{\tilde{\mathcal{T}}}_n$ is in (3.15) above) is itself bounded. Hence $\underline{\tilde{\mathcal{T}}}_n \in \mathcal{B}(c_0^S \times \mathbf{R}^{m_t}, c_0^{n_z \times n_w})$. Moreover, Proposition B.1.2 also shows that the adjoint of such an operator can be written as the cartesian product of the adjoints of the component operators. Hence $\underline{\tilde{\mathcal{T}}}_n^*$ can be written, given $K \in l_1^{n_z \times n_w}$,

$$\underline{\tilde{\mathcal{T}}}_n^* K = (\tilde{\mathcal{P}}_n^* \tilde{\mathcal{T}}_C^* K, \tilde{\mathcal{T}}_I^* K).$$

Because $\{\underline{\mu}_n\}_{n=0}^{\infty}$ is monotonic and $\underline{\mu}_n \leq \mu_{DR}$ for every n, the sequence convergences and

$$\lim_{n\to\infty} \underline{\mu}_n = \sup_n \underline{\mu}_n \leq \mu_{DR}.$$

To establish that the supremum is equal to μ_{DR} use the fact [26, Theorem 4.12] that $\mathcal{R}\left(\tilde{\underline{\mathcal{I}}}_n\right)^{\perp} = \mathcal{N}\left(\tilde{\underline{\mathcal{I}}}_n^*\right) = \underline{\boldsymbol{K}}_n(U,V)$ and apply the duality theorem B.2.3 to obtain

$$\begin{aligned}
\underline{\mu}_n &= \sup\left\{\langle G,H\rangle : \boldsymbol{B}\mathcal{R}\left(\tilde{\underline{\mathcal{I}}}_n\right)\right\} \\
&= \sup\left\{\left\langle \tilde{\underline{\mathcal{I}}}_n\tilde{G},H\right\rangle : \tilde{G}\in c_0^S\times\mathbf{R}^{m_t},\ \left\|\tilde{\underline{\mathcal{I}}}_n\tilde{G}\right\|_{l_\infty}\le 1\right\} \\
&= \sup\left\{\left\langle \tilde{\mathcal{T}}_K\tilde{G},H\right\rangle : \tilde{G}\in\mathcal{R}\left(\tilde{\mathcal{P}}_n\right)\times\mathbf{R}^{m_t},\ \left\|\tilde{\mathcal{T}}_K\tilde{G}\right\|_{l_\infty}\le 1\right\} \qquad (3.16)
\end{aligned}$$

where $\tilde{\mathcal{T}}_K$ is defined in equation (3.10). Hence

$$\begin{aligned}
\sup_n \underline{\mu}_n &= \sup_n\sup\left\{\left\langle \tilde{\mathcal{T}}_K\tilde{G},H\right\rangle : \tilde{G}\in\mathcal{R}\left(\tilde{\mathcal{P}}_n\right)\times\mathbf{R}^{m_t},\ \left\|\tilde{\mathcal{T}}_K\tilde{G}\right\|_{l_\infty}\le 1\right\} \\
&= \sup\left\{\left\langle \tilde{\mathcal{T}}_K\tilde{G},H\right\rangle : \tilde{G}\in (c_0)_{\text{fs}}^S\times\mathbf{R}^{m_t},\ \left\|\tilde{\mathcal{T}}_K\tilde{G}\right\|_{l_\infty}\le 1\right\} \\
&= \sup\left\{\left\langle \tilde{\mathcal{T}}_K\tilde{G},H\right\rangle : \tilde{G}\in c_0^S\times\mathbf{R}^{m_t},\ \left\|\tilde{\mathcal{T}}_K\tilde{G}\right\|_{l_\infty}\le 1\right\} \\
&= \sup\left\{\langle G,H\rangle : \boldsymbol{B}\mathcal{R}\left(\tilde{\mathcal{T}}_K\right)\right\} = \inf\left\{\|H-K\|_{l_1} : K\in\mathcal{N}\left(\tilde{\mathcal{T}}_K^*\right)\right\} \\
&= \mu_{DR}
\end{aligned}$$

where $(c_0)_{\text{fs}}^S$ denotes the finitely supported matrices in c_0^S. The second equality follows because $(c_0)_{\text{fs}}^S\times\mathbf{R}^{m_t} = \cup_{n\in\mathbf{Z}_+}\mathcal{R}\left(\tilde{\mathcal{P}}_n\right)\times\mathbf{R}^{m_t}$. The third follows from the facts that $(c_0)_{\text{fs}}^S\times\mathbf{R}^{m_t}$ is dense in $c_0^S\times\mathbf{R}^{m_t}$ and $\tilde{\mathcal{T}}_K\in\mathcal{B}(c_0^S\times\mathbf{R}^{m_t}, c_0^{n_z\times n_w})$ (the image under a bounded operator of a dense subspace is dense). The last equality follows because $\boldsymbol{K}(U,V) = \mathcal{N}\left(\tilde{\mathcal{T}}_K^*\right)$.

For the second statement, it is clear that $\underline{\boldsymbol{K}}_n(U,V) = \mathcal{N}\left(\tilde{\underline{\mathcal{I}}}_n^*\right)$; hence it is *weak**-closed and $\underline{OPT}_n(H,U,V)$ has a minimizer. □

As a consequence of this theorem, the solution of a sequence of super-optimal problems $\underline{OPT}_n(H,U,V)$ for increasing n provides a converging lower bound on μ_{DR}. The next theorem shows that the set of minimizers for $\underline{OPT}_n(H,U,V)$ has properties similar to that for $OPT(H,U,V)$ when U and V have full row rank and full column rank, respectively (see Theorem 3.2.7).

Theorem 3.4.2 *For every n every minimizer K_0 for $\underline{OPT}_n(H,U,V)$ is such that at least one row of $H-K_0$ is finitely supported. Moreover, there exists a minimizer K_{fs} such that all entries of $H-K_{\text{fs}}$ are finitely supported.*

Proof: The proof is very similar to that of Theorem 3.2.7. For the first statement, recall that $\underline{\boldsymbol{K}}_n(U,V) = \mathcal{N}\left(\underline{\tilde{\mathcal{I}}}_n^*\right)$ where $\underline{\tilde{\mathcal{I}}}_n$ is defined in (3.15). The results of Appendix C imply that $(\underline{\tilde{\mathcal{I}}}_n^*)^*$ can be written, given $(G,\alpha) \in l_\infty^S \times \mathbf{R}^{m_t}$,

$$(\underline{\tilde{\mathcal{I}}}_n^*)^* [(G,\alpha)] = \tilde{\mathcal{T}}_C \bar{\mathcal{P}}_n G + \tilde{\mathcal{T}}_I \alpha$$

where $\tilde{\mathcal{T}}_C \in \mathcal{B}(c_0^S, c_0^{n_z \times n_w})$ is defined in (3.6) and $\tilde{\mathcal{T}}_I \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_z \times n_w})$ in (3.8), in the proofs of Lemmas 3.2.2 and 3.2.3, respectively. $\bar{\mathcal{P}}_n$ is the n-th truncation operator restricted to l_∞^S and is bounded by Proposition C.1.7. The expression for $(\underline{\tilde{\mathcal{I}}}_n^*)^*$ makes sense since clearly $\mathcal{R}\left(\bar{\mathcal{P}}_n\right) \subset c_0^S$. Now,

$$\underline{\boldsymbol{K}}_n(U,V)^\perp = \mathcal{N}\left(\underline{\tilde{\mathcal{I}}}_n^*\right)^\perp = \left[{}^\perp\mathcal{R}\left((\underline{\tilde{\mathcal{I}}}_n^*)^*\right)\right]^\perp = \mathcal{R}\left((\underline{\tilde{\mathcal{I}}}_n^*)^*\right) = \mathcal{R}\left(\underline{\tilde{\mathcal{I}}}_n\right) \subset c_0^{n_z \times n_w}$$

The reasoning here is identical to that for (3.12). Finally, by Lemma 3.2.6, every minimizer K_0 for $\underline{OPT}_n(H,U,V)$ is such that $H - K_0$ has at least one finitely supported row.

For the second statement, note that Theorem 3.4.1 guarantees the existence of a minimizer K_1 for $\underline{OPT}_n(H,U,V)$. By the argument above, at least one row of $H - K_1$ is finitely supported. Next it is shown that, given $p \in \{1,\ldots,n_z\}$ and any minimizer K_p such that at least p rows of $H - K_p$ are finitely supported, there exists a minimizer K_{p+1} such that at least $p+1$ rows of $H - K_{p+1}$ are finitely supported. Hence, there is at least one minimizer K_{fs} such that $H - K_{\text{fs}}$ is finitely supported.

Suppose, then, that $p \in \{1,\ldots,n_z\}$ and K_p is any minimizer for $\underline{OPT}_n(H,U,V)$ such that at least p rows (assumed, without loss of generality, to be the first p) of $H - K_p$ are finitely supported. Partition after the p-th row,

$$H = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix}, \quad K_p = \begin{bmatrix} (K_p)_1 \\ (K_p)_2 \end{bmatrix},$$

and consider the problem

$$\underline{OPT}' : \qquad \inf\left\{\|(H_2 - (K_p)_2) - K'\|_{l_1} : K' \in \underline{\boldsymbol{K}}_n'(K_p,U,V)\right\}$$

where

$$\underline{\boldsymbol{K}}_n'(K_p,U,V) := \left\{K' \in l_1^{(n_z - p)\times n_w} : \begin{bmatrix} (K_p)_1 \\ (K_p)_2 + K' \end{bmatrix} \in \underline{\boldsymbol{K}}_n(U,V)\right\}.$$

It is easy to check that any minimizer for $\underline{OPT}'$ is also a minimizer for $\underline{OPT}_n(H,U,V)$.

Moreover, $K_p \in \boldsymbol{K}(U,V)$ implies, using Theorem 3.2.1, that $K' \in \underline{\boldsymbol{K}}'(K_p,U,V)$ if and only if $K' \in l_1^{(n_z-p)\times n_w}$, K' satisfies

$$[(\hat{U}_L^{-L})^2_{i\cdot}\hat{K}'(\hat{V}_R^{-R})_{\cdot j}]^{(k)}(z_0) = 0, \qquad k \in \{0,\dots,m_{ij}(z_0)-1\}$$

for each $(i,j) \in \{1,\dots,r_U\} \times \{1,\dots,r_V\}$ and each $z_0 \in \boldsymbol{Z}_{ij}$, where $(U_L^{-L})^2$ denotes the last $n_z - p$ columns of U_L^{-L}, and K' satisfies

$$\left(\begin{bmatrix} (U_L^{-L})^2 \\ (U_L^{\perp})^2 \end{bmatrix} K' \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)(k) = \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{for all } k \in \{0,\dots,n\}$$

where $(U_L^{\perp})^2$ denotes the last $n_z - p$ columns of $U_L^{\perp}$.

Thus, arguing as in Theorem 3.4.1, there exists $\mathcal{T} \in \mathcal{B}(l_1^{(n_z-p)\times n_w}, c_0^{S'} \times \mathbf{R}^{m_t})$ (for an appropriate choice of index set S') such that $\underline{\boldsymbol{K}}'_n(K_p,U,V) = \mathcal{N}(\mathcal{T})$. Moreover, $\exists \tilde{\mathcal{T}} \in \mathcal{B}(c_0^{S'} \times \mathbf{R}^{m_t}, c_0^{(n_z-p)\times n_w})$ such that $\tilde{\mathcal{T}}^* = \mathcal{T}$ and hence $\underline{\boldsymbol{K}}'_n(K_p,U,V)$ is *weak**-closed. This guarantees the existence of a minimizer $(K')^1$ for $\underline{OPT}'$ by CorollaryB.2.4. Arguing as for the proof of the first statement, $\underline{\boldsymbol{K}}'_n(K_p,U,V)^{\perp} \subset c_0^{(n_z-p)\times n_w}$ so that, by Lemma 3.2.6, $(K')^1$ is such that at least one row of $H - (K_p)_2 - (K')^1$ is finitely supported. Finally, then,

$$K_{p+1} = \begin{bmatrix} (K_{p+1})_1 \\ (K_{p+1})_2 \end{bmatrix} := \begin{bmatrix} (K_p)_1 \\ (K_p)_2 + (K')^1 \end{bmatrix}$$

is a minimizer for $\underline{OPT}_n(H,U,V)$ and at least $p+1$ rows of $H - K_{p+1}$ are finitely supported. □

Theorem 3.4.2 has an important practical consequence which is exploited in the remainder of this section. A doubly indexed family of optimization problems is defined each member of which is equivalent to a finite dimensional optimization problem, and which provides a computable double iteration method for bounding μ_{DR} from below arbitrarily closely.

Given a problem $\underline{OPT}_n(H,U,V)$ for some n define for each $m \in \mathbf{Z}_+$ an optimization problem

$$\underline{OPT}_{n,m}(H,U,V) \qquad\qquad \inf\left\{\|H-K\|_{l_1} : K \in \underline{\boldsymbol{K}}_n(U,V)\right\} =: \underline{\mu}_{n,m}$$

where

$$\underline{\boldsymbol{K}}_{n,m}(H,U,V) := \{K \in \underline{\boldsymbol{K}}_n(U,V) : \text{supp } (H-K) \subset \{0,\ldots,m\}\} \tag{3.17}$$

and $\underline{\boldsymbol{K}}_n(U,V)$ is defined in (3.14).

Hence for each problem $\underline{OPT}_n(H,U,V)$ there is a sequence of problems

$$\left\{\underline{OPT}_{n,m}(H,U,V)\right\}_{m=1}^{\infty}.$$

The following proposition shows that as m is increased each such sequence not only becomes feasible for all greater m but, for large enough m, a minimizer for $\underline{OPT}_{n,m}(H,U,V)$ yields a minimizer for $\underline{OPT}_n(H,U,V)$. If $\boldsymbol{Z} = \emptyset$ the situation is particularly simple.

Proposition 3.4.3 *If $m \geq n$ and $m \geq m_t$ then $\underline{OPT}_{n,m}(H,U,V)$ is feasible and $\underline{\mu}_{n,m} \geq \underline{\mu}_n$. Moreover*

1. *For each n there exists $M_n \geq n$ such that $\underline{\mu}_{n,m} = \underline{\mu}_n$.*

2. *If $\boldsymbol{Z} = \emptyset$ then, for each n, $\underline{OPT}_{n,n}(H,U,V)$ is feasible and $\underline{\mu}_{n,n} = \underline{\mu}_n$.*

Proof: For the first sentence, assume $m \geq n$ and $m \geq m_t$ and define $K := K_H - K_m$ where $K_H := (\mathcal{I} - \mathcal{P}_m)H$, K_m satisfies

$$\text{supp } K_m \subset \{0,\ldots,m\} \text{ and } \mathcal{T}_I K_m = \mathcal{T}_I K_H,$$

and $\mathcal{T}_I$ is defined in Lemma 3.2.3. Finding K_m to satisfy these requirements is equivalent to constructing a polynomial matrix of degree at most m which satisfies m_t interpolation requirements; this is possible since $m \geq m_t$. Moreover, it is straightforward to check that K satisfies all conditions of Theorem 3.2.1 by its construction since $m \geq n$. Hence $\underline{\boldsymbol{K}}_{n,m}(H,U,V) \neq \emptyset$ and $\underline{\mu}_{n,m}$ is well defined. $\underline{\mu}_{n,m} \geq \underline{\mu}_n$ since $\underline{\boldsymbol{K}}_{n,m}(H,U,V) \subset \underline{\boldsymbol{K}}_n(U,V)$.

Item 1: Theorem 3.4.2 shows that there is a finitely supported minimizer K_{fs} for $\underline{OPT}_n(H,U,V)$; just take $M_n := \max\{n, \text{supp } K_{\text{fs}}\}$.

Item 2: Under the hypotheses, $(\mathcal{I} - \mathcal{P}_n)H \in \underline{\boldsymbol{K}}_n(U,V)$ and $\mathcal{P}_n K \in \underline{\boldsymbol{K}}_n(U,V)$ for all $K \in \underline{\boldsymbol{K}}_n(U,V)$. (This is because condition 3 of Theorem 3.2.1 is vacuous in this case.) Therefore if $K \in \underline{\boldsymbol{K}}_n(U,V)$ then $\tilde{K} := \mathcal{P}_n K + (\mathcal{I} - \mathcal{P}_n)H \in \underline{\boldsymbol{K}}_n(U,V)$. Moreover it is easy to check that supp $(H - \tilde{K}) \subset \{0, \ldots, n\}$ so that $\tilde{K} \in \underline{\boldsymbol{K}}_{n,n}(H, U, V)$. Also, $(H-\tilde{K})(k) = (H-K)(k)$ for $k \le n$ so that $\left\|H - \tilde{K}\right\|_{l_1} \le \|H - K\|_{l_1}$. This shows that $\underline{\mu}_{n,n} \le \underline{\mu}_n$; the reverse inequality follows from the first sentence of the proposition. □

Finally, the following proposition gives a method for constructing feasible solutions for $OPT(H, U, V)$ and hence sub-optimal stabilizing compensators for $\mathcal{G}$ from feasible solutions for any $\underline{OPT}_n(H, U, V)$. Since feasible solutions for $\underline{OPT}_{n,m}(H, U, V)$ (for any $m \ge n$) are also feasible for $\underline{OPT}_n(H, U, V)$, sub-optimal compensators can be obtained from these computable problems.

Proposition 3.4.4 *If* $\underline{K} \in \underline{\boldsymbol{K}}_n(U, V)$ *for any* n *then* $K := U_L U_L^{-L} \underline{K} V_R^{-R} V_R \in \boldsymbol{K}(U, V)$.

Proof: It is easy to verify using the definition of $\underline{\boldsymbol{K}}_n(U, V)$ in (3.14) that if $K \in \underline{\boldsymbol{K}}_n(U, V)$ then K satisfies the three conditions of Theorem 3.2.1 and hence $K \in \boldsymbol{K}(U, V)$. First, $\underline{K} \in \underline{\boldsymbol{K}}_n(U, V) \Rightarrow K \in l_1^{n_z \times n_w}$ so that condition 1 is satisfied. Also,

$$\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} = \begin{bmatrix} U_L^{-L} \underline{K} V_R^{-R} & 0 \\ 0 & 0 \end{bmatrix}$$

so that condition 2 is satisfied. In particular,

$$U_L^{-L} K V_R^{-R} = U_L^{-L} \underline{K} V_R^{-R}$$

so that condition 3 is also satisfied. □

Using the results of this section, a double iteration can be carried out to bound μ_{DR} from below for any problem. The procedure is choose a desired initial n. Next solve $\underline{OPT}_{n,m}(H, U, V)$ for increasing $m \ge n$ until no further change is observed in $\underline{\mu}_{n,m}$; hence $\underline{\mu}_n$ has been found. (Although no results have been shown bounding the required m, an abrupt convergence is generally obvious in examples. Moreover, if $\boldsymbol{Z} = \emptyset$ then no iteration on m is required.) n can then be increased until $\underline{\mu}_n$ appears to converge. This convergence can be confirmed when a feasible sequence of FIR approximations $\overline{OPT}_n(H, U, V)$ exists by solving them in parallel to upper bound μ_{DR}.

3.5 Linear Programming Formulations

In this section linear programs are formulated corresponding to all the optimization problems defined thus far. The linear programs are all equivalent to the original problems in the following sense.

Definition 3.5.1 *Let MIN and $\widetilde{MIN}$ be two optimization problems defined*

$$MIN: \qquad \inf \{J(x) : x \in \boldsymbol{X}\} =: \gamma_0$$

$$\widetilde{MIN}: \qquad \inf \left\{\tilde{J}(\tilde{x}) : \tilde{x} \in \tilde{\boldsymbol{X}}\right\} =: \tilde{\gamma}_0$$

MIN is equivalent to $\widetilde{MIN}$ under the map ψ *if* $\psi : \boldsymbol{X} \mapsto \tilde{\boldsymbol{X}}$ *and there exists a map* $\tilde{\psi} : \tilde{\boldsymbol{X}} \mapsto \boldsymbol{X}$ *such that* $\psi\tilde{\psi} = \mathcal{I}$, $\tilde{J}(\psi x) \leq J(x)$ *for all* $x \in \boldsymbol{X}$, *and* $J(\tilde{\psi}\tilde{x}) \leq \tilde{J}(\tilde{x})$ *for all* $\tilde{x} \in \tilde{\boldsymbol{X}}$.

If MIN is equivalent to $\widetilde{MIN}$ under the map ψ then

- all feasible solutions for $\widetilde{MIN}$ can be recovered from feasible solutions for MIN via the map ψ; more precisely,

  $$\tilde{\boldsymbol{X}} = \psi \boldsymbol{X}$$

  In particular, $\widetilde{MIN}$ is feasible ($\tilde{\boldsymbol{X}} \neq \emptyset$) if and only if MIN is feasible ($\boldsymbol{X} \neq \emptyset$).

- all feasible solutions for $\widetilde{MIN}$ achieving at most a given cost can be recovered from feasible solutions for MIN achieving at most the same cost via the map ψ; more precisely, if $\gamma \in \mathbf{R}$, $\tilde{\boldsymbol{X}}_\gamma := \left\{\tilde{x} \in \tilde{\boldsymbol{X}} : \tilde{J}(x) \leq \gamma\right\}$, and $\boldsymbol{X}_\gamma := \{x \in \boldsymbol{X} : J(x) \leq \gamma\}$ then

  $$\tilde{\boldsymbol{X}}_\gamma = \psi \boldsymbol{X}_\gamma$$

  In particular, the cost for $\widetilde{MIN}$ is bounded below if and only if the cost for MIN is bounded below.

- all optimal solutions for $\widetilde{MIN}$ can be recovered from optimal solutions for MIN via the map ψ; more precisely, if the costs are bounded below, $\tilde{\boldsymbol{X}}_0 := \left\{\tilde{x} \in \tilde{\boldsymbol{X}} : \tilde{J}(x) = \gamma_0\right\}$, and $\boldsymbol{X}_0 := \{x \in \boldsymbol{X} : J(x) = \gamma_0\}$ then

  $$\tilde{\boldsymbol{X}}_0 = \psi \boldsymbol{X}_0$$

  In particular, there exists a minimizer for $\widetilde{MIN}$ if and only if there exists a minimizer for MIN.

- if either γ_0 or $\tilde{\gamma}_0$ is well defined then both are and $\gamma_0 = \tilde{\gamma}_0$.

For each linear program, a map is given under which it is equivalent to the corresponding optimization problem. Proving equivalence is done by showing that the map satisfies the conditions of Definition 3.5.1.

In order to simplify the statement of the linear programs the notation l_1^+ will be adopted for the positive cone in l_1, i.e.,

$$l_1^+ := \{x \in l_1 : x(k) \geq 0 \ \forall k \in \mathbf{Z}_+\}$$

The maps used to prove equivalence will be defined partially in terms of the projection operator $\Pi_+ : l_1^{n_z \times n_w} \mapsto (l_1^+)^{n_z \times n_w}$ defined, given $x \in l_1^{n_z \times n_w}$,

$$(\Pi_+ x)_{ij}(k) := \begin{cases} x_{ij}(k) & \text{if } x_{ij}(k) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This operator simply sets the negative elements of each sequence in the matrix to zero. It is also convenient to define

$$\tilde{H} := \begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} H \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}$$

The projection operator Π_+ has some useful properties which will be exploited in the proofs in this section.

Remark 3.5.2 *If $X \in l_1^{n_z \times n_w}$ then*

1. $X = \Pi_+ X - \Pi_+(-X)$.

2. $|X_{ij}(k)| = [\Pi_+ X + \Pi_+(-X)]_{ij}(k)$ *for each i, j, and k and hence*

$$\|X\|_{l_1} = \|\Pi_+ X\|_{l_1} + \|\Pi_+(-X)\|_{l_1}.$$

The first linear program corresponds to $OPT(H, U, V)$ itself, defined in Section 3.1. The linear program is

$LP(H, U, V):$ $\qquad\qquad\qquad \inf \mu$

subject to:

$$
\begin{aligned}
\left[\sum_{j=1}^{n_w}\sum_{k=0}^{\infty}(T_{zw}^{+}+T_{zw}^{-})_{ij}(k)\right]-\mu &\leq 0 \qquad i=1,\ldots,n_z && (3.18)\\
\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix}(T_{zw}^{+}-T_{zw}^{-})\begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)_{ij}(k) &= \tilde{H}_{ij}(k) && (3.19)\\
& \qquad (i,j)\in \boldsymbol{S}, \quad k\in \mathbf{Z}_+ &&\\
\mathcal{T}_I(T_{zw}^{+}-T_{zw}^{-}) &= \mathcal{T}_I H && (3.20)\\
(\mu, T_{zw}^{+}, T_{zw}^{-}) &\in \mathbf{R}_+\times (l_1^{+})^{n_z\times n_w} && (3.21)\\
&\qquad \times (l_1^{+})^{n_z\times n_w} &&
\end{aligned}
$$

where $\mathcal{T}_I$ is defined in Lemma 3.2.3.

The variables in $LP(H,U,V)$ are $\mu \in \mathbf{R}$ and

$$\left\{(T_{zw}^{+})_{ij}(k),\ (T_{zw}^{-})_{ij}(k) : i\in\{1,\ldots,n_z\},\ j\in\{1,\ldots,n_w\},\ k\in\mathbf{Z}_+\right\}.$$

Note that they are infinite in number. Equation (3.18) defines n_z inequality constraints called *cost constraints.* The equality constraints in (3.19), called *convolution constraints*, are infinite in number. The equality constraints in (3.20), called *interpolation constraints*, are finite in number since $\mathcal{T}_I : l_1^{n_z\times n_w} \mapsto \mathbf{R}^{m_t}$. Equation (3.21) defines an infinite number of inequality constraints called *positivity constraints.*

Proposition 3.5.3 *$LP(H,U,V)$ is equivalent to $OPT(H,U,V)$ under the map ψ defined, given a feasible solution $(\mu, T_{zw}^{+}, T_{zw}^{-})$ for $LP(H,U,V)$,*

$$\psi(\mu, T_{zw}^{+}, T_{zw}^{-}) := H - (T_{zw}^{+} - T_{zw}^{-}).$$

Proof: Let $(\mu, T_{zw}^{+}, T_{zw}^{-})$ be a feasible solution for $LP(H,U,V)$. We will show that $\psi(\mu, T_{zw}^{+}, T_{zw}^{-}) \in \boldsymbol{K}(U,V)$ by checking the conditions of Theorem 3.2.1. $H, T_{zw}^{+}, T_{zw}^{-} \in l_1^{n_z\times n_w}$ implies $\psi(\mu, T_{zw}^{+}, T_{zw}^{-}) = H - (T_{zw}^{+} - T_{zw}^{-}) \in l_1^{n_z\times n_w}$ and hence satisfies condition 1. The convolution constraints (3.19) ensure that condition 2 is satisfied, and the interpolation constraints (3.20) ensure that condition 3 is satisfied. Hence $\psi(\mu, T_{zw}^{+}, T_{zw}^{-}) \in \boldsymbol{K}(U,V)$. Moreover the cost of $\psi(\mu, T_{zw}^{+}, T_{zw}^{-})$ is

$$\left\|H - \psi(\mu, T_{zw}^{+}, T_{zw}^{-})\right\|_{l_1} = \left\|T_{zw}^{+} - T_{zw}^{-}\right\|_{l_1} \leq \left\|T_{zw}^{+}\right\|_{l_1} + \left\|T_{zw}^{-}\right\|_{l_1} \leq \mu,$$

which is the cost of $(\mu, T^+_{zw}, T^-_{zw})$. The last inequality is ensured by the cost and positivity constraints (3.18) and (3.21).

Next define a map $\tilde{\psi}$, given $K \in \boldsymbol{K}(U,V)$,

$$\tilde{\psi}K := \left(\|H-K\|_{l_1}, \; \Pi_+(H-K), \; \Pi_+(K-H)\right).$$

Item 2 of Remark 3.5.2 ensures that $\tilde{\psi}K$ satisfies the cost constraints (3.18) (with equality). Using item 1 of Remark 3.5.2 it is easy to check that conditions 2 and 3 ensure, respectively, that the convolution and interpolation constraints (3.19) and (3.20) are satisfied. The positivity constraints (3.21) are satisfied since $\Pi_+ : l_1^{n_z \times n_w} \mapsto (l_1^+)^{n_z \times n_w}$. Moreover the cost of $\tilde{\psi}K$ is $\|H-K\|_{l_1}$ which equals the cost of K. Finally, again using item 1 of Remark 3.5.2,

$$\begin{aligned} \psi\tilde{\psi}K &= \psi\left(\|H-K\|_{l_1}, \; \Pi_+(H-K), \; \Pi_+(K-H)\right) \\ &= H - [\Pi_+(H-K) - \Pi_+(K-H)] \\ &= K \end{aligned}$$

so that $\psi\tilde{\psi} = \mathcal{I}$. □

The next linear program corresponds to an FIR sub-optimal approximation problem $\overline{OPT}_n(H,U,V)$, defined in Section 3.3. The linear program is

$\overline{LP}_n(H,U,V)$: $\quad\quad$ $\inf \mu$

subject to:

$$\left[\sum_{j=1}^{n_w}\sum_{k=0}^{n}(T^+_{zw}+T^-_{zw})_{ij}(k)\right] - \mu \;\leq\; 0 \qquad i=1,\ldots,n_z \tag{3.22}$$

$$\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} \mathcal{E}_n(T^+_{zw} - T^-_{zw}) \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)_{ij}(k) \;=\; \tilde{H}_{ij}(k) \tag{3.23}$$

$$(i,j) \in \boldsymbol{S}, \quad k \in \{0,\ldots,\bar{n}\}$$

$$\mathcal{T}_I \mathcal{E}_n(T^+_{zw} - T^-_{zw}) \;=\; \mathcal{T}_I H \tag{3.24}$$

$$(\mu, T^+_{zw}, T^-_{zw}) \;\in\; \mathbf{R}_+ \times \mathbf{R}_+^{n_z \times n_w \times (n+1)} \times \mathbf{R}_+^{n_z \times n_w \times (n+1)} \tag{3.25}$$

where the matrices in (3.23) have been chosen to be finitely supported. $\bar{n} := n + n_U + n_V$ where n_U and n_V are integers such that

$$\operatorname{supp}\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix}\right) \subset \{0,\ldots,n_U\} \quad \text{and} \quad \operatorname{supp}\left(\begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right) \subset \{0,\ldots,n_V\}.$$

$\mathcal{E}_n : \mathbf{R}_+^{n_z \times n_w \times (n+1)} \mapsto l_1^{n_z \times n_w}$ is the natural embedding operator which pads with zeros. S and $\mathcal{T}_I$ are as defined in the formulation of LP above.

There are $2n_z n_w(n+1)+1$ variables in $\overline{LP}_n(H,U,V)$: μ and

$$\left\{(T_{zw}^+)_{ij}(k),\ (T_{zw}^-)_{ij}(k) : i \in \{1,\ldots,n_z\},\ j \in \{1,\ldots,n_w\},\ k \in \{0,\ldots,n\}\right\}.$$

Proposition 3.5.4 *If there exists $\bar{n}$ such that $\overline{\boldsymbol{K}}_{\bar{n}}(H,U,V) \neq \emptyset$ then there exists n_H such that*

$$\operatorname{supp}\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} H \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)_{ij} \subset \{0,\ldots,n_H\} \quad \forall (i,j) \in \boldsymbol{S}$$

and if $n \geq n_H - n_U - n_V$ then $\overline{LP}_n(H,U,V)$ is equivalent to $\overline{OPT}_n(H,U,V)$ under the map $\overline{\psi}_n$ defined, given a feasible solution $(\mu, T_{zw}^+, T_{zw}^-)$ for $\overline{LP}_n(H,U,V)$,

$$\overline{\psi}_n(\mu, T_{zw}^+, T_{zw}^-) := H - \mathcal{E}_n(T_{zw}^+ - T_{zw}^-).$$

Proof: Proceeding as in the proof of Proposition 3.5.3, let $(\mu, T_{zw}^+, T_{zw}^-)$ be a feasible solution for $\overline{LP}_n(H,U,V)$. $H, T_{zw}^+, T_{zw}^- \in l_1^{n_z \times n_w}$ implies $\overline{\psi}_n(\mu, T_{zw}^+, T_{zw}^-) = H - \mathcal{E}_n(T_{zw}^+ - T_{zw}^-) \in l_1^{n_z \times n_w}$ and hence satisfies condition 1 of Theorem 3.2.1. The convolution constraints (3.23) ensure that, for each $(i,j) \in \boldsymbol{S}$ and for $k \in \{0.\ldots,\bar{n}\}$,

$$\begin{aligned}
&\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)_{ij}(k) = \\
&\qquad \tilde{H}_{ij}(k) - \left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} \mathcal{E}_n(T_{zw}^+ - T_{zw}^-) \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}\right)_{ij}(k).
\end{aligned}$$

$n \geq n_H - n_U - n_V$ implies that $\bar{n} \geq n_H$ so that for each $(i,j) \in \boldsymbol{S}$ and all $k > \bar{n}$ each term on the right above is zero. Hence condition 2 is satisfied. The interpolation constraints (3.24) ensure that condition 3 is satisfied, and it is obvious by the construction that supp $[H - \mathcal{E}_n(T^+_{zw} - T^-_{zw})] \subset \{0, \ldots, n\}$. Hence $\overline{\psi}_n(\mu, T^+_{zw}, T^-_{zw}) \in \overline{\boldsymbol{K}}_n(H, U, V)$. Moreover the cost of $\overline{\psi}_n(\mu, T^+_{zw}, T^-_{zw})$ is

$$\left\|H - \overline{\psi}_n(\mu, T^+_{zw}, T^-_{zw})\right\|_{l_1} = \left\|\mathcal{E}_n(T^+_{zw} - T^-_{zw})\right\|_{l_1} \leq \left\|\mathcal{E}_n T^+_{zw}\right\|_{l_1} + \left\|\mathcal{E}_n T^-_{zw}\right\|_{l_1} \leq \mu,$$

which is the cost of $(\mu, T^+_{zw}, T^-_{zw})$. The last inequality is ensured by the cost and positivity constraints (3.22) and (3.25).

Define a map $\tilde{\overline{\psi}}_n$, given $K \in \overline{\boldsymbol{K}}_n(H, U, V)$,

$$\tilde{\overline{\psi}}_n K := \left(\|H - K\|_{l_1},\ \Pi_n \Pi_+(H - K),\ \Pi_n \Pi_+(K - H)\right).$$

where $\Pi_n : l_1^{n_z \times n_w} \mapsto \mathbf{R}_+^{n_z \times n_w \times (n+1)}$ is the obvious projection operator. Item 2 of Remark 3.5.2 and the fact that supp $(H - K) \subset \{0, \ldots, n\}$ ensure that $\tilde{\overline{\psi}}_n K$ satisfies the cost constraints (3.22) (with equality). Using item 1 of Remark 3.5.2 and the fact that $n \geq n_H - n_U - n_V$ it is easy to check that condition 2 ensures that the convolution constraints (3.23) are satisfied. Also using item 1 of Remark 3.5.2 is easy to check that condition 3 ensures that the interpolation constraints (3.23) are satisfied. The positivity constraints (3.21) are satisfied since $\Pi_n \Pi_+ : l_1^{n_z \times n_w} \mapsto \mathbf{R}_+^{n_z \times n_w \times (n+1)}$. Moreover the cost of $\tilde{\overline{\psi}}_n K$ is $\|H - K\|_{l_1}$ which equals the cost of K. Finally, using item 1 of Remark 3.5.2 and the fact that $\Pi_n(H - K) = H - K$,

$$\begin{aligned}
\overline{\psi}_n \tilde{\overline{\psi}}_n K &= \overline{\psi}_n \left(\|H - K\|_{l_1},\ \Pi_n \Pi_+(H - K),\ \Pi_n \Pi_+(K - H)\right) \\
&= H - [\Pi_n \Pi_+(H - K) - \Pi_n \Pi_+(K - H)] \\
&= K
\end{aligned}$$

so that $\overline{\psi}_n \tilde{\overline{\psi}}_n = \mathcal{I}$. □

The next linear program corresponds to an infinite dimensional super-optimal approximation problem $\underline{OPT}_n(H, U, V)$, defined in Section 3.4. The linear program

$\underline{LP}_n(H, U, V)$: $\inf \mu$

subject to:

$$\left[\sum_{j=1}^{n_w} \sum_{k=0}^{\infty} (T^+_{zw} + T^-_{zw})_{ij}(k)\right] - \mu \ \leq\ 0 \qquad i = 1, \ldots, n_z$$

$$
\begin{aligned}
\left(\begin{bmatrix} U_L^{-L} \\ U_L^{\perp} \end{bmatrix} (T_{zw}^+ - T_{zw}^-) \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix} \right)_{ij} (k) &= \tilde{H}_{ij}(k) \\
& \qquad (i,j) \in \boldsymbol{S}, \quad k \in \{0, \ldots, n\} \\
\mathcal{T}_I (T_{zw}^+ - T_{zw}^-) &= \mathcal{T}_I H \\
(\mu, T_{zw}^+, T_{zw}^-) &\in \mathbf{R}_+ \times (l_1^+)^{n_z \times n_w} \times (l_1^+)^{n_z \times n_w}
\end{aligned}
$$

where all symbols are as defined in the formulation of LP above.

Note that $\underline{LP}_n(H,U,V)$ is very similar to $LP(H,U,V)$. The variables are $\mu \in \mathbf{R}$ and

$$
\left\{ (T_{zw}^+)_{ij}(k),\ (T_{zw}^-)_{ij}(k) : i \in \{1, \ldots, n_z\},\ j \in \{1, \ldots, n_w\},\ k \in \mathbf{Z}_+ \right\}.
$$

as in $LP(H,U,V)$ and are infinite in number. The constraints are identical as well, except that only a finite number of the convolution constraints (3.19) are enforced. Hence, although there are still infinitely many positivity constraints, there are only a finite number of equality constraints.

Proposition 3.5.5 *For every n, $\underline{LP}_n(H,U,V)$ is equivalent to $\underline{OPT}_n(H,U,V)$ under the map $\underline{\psi}_n$ defined, given a feasible solution $(\mu, T_{zw}^+, T_{zw}^-)$ for $\underline{LP}_n(H,U,V)$,*

$$
\underline{\psi}_n(\mu, T_{zw}^+, T_{zw}^-) := H - (T_{zw}^+ - T_{zw}^-).
$$

Proof: For every n $\underline{\psi}_n$ is defined identically to ψ of Proposition 3.5.3 and if we define, given $K \in \underline{\boldsymbol{K}}_n(U,V)$,

$$
\underline{\tilde{\psi}}_n K := \left(\|H-K\|_{l_1},\ \Pi_+(H-K),\ \Pi_+(K-H) \right)
$$

identically to $\tilde{\psi}$ then the proof of that proposition is easily adapted to serve here. We need only note that the only difference between $LP(H,U,V)$ and $\underline{LP}_n(H,U,V)$ are identical except that only a finite number of convolution constraints have been enforced in the latter. This this corresponds precisely to the difference between the characterization of $\boldsymbol{K}(U,V)$ in Theorem 3.2.1 and the definition of $\underline{\boldsymbol{K}}_n(U,V)$ in (3.14).
□

The last linear program corresponds to a finite dimensional super-optimal approximation problem $\underline{OPT}_{n,m}(H,U,V)$, defined in Section 3.4.

$\underline{LP}_{n,m}(H,U,V)$: $\qquad\qquad \inf \mu$

subject to:

$$
\begin{aligned}
\left[\sum_{j=1}^{n_w}\sum_{k=0}^{m}(T_{zw}^{+}+T_{zw}^{-})_{ij}(k)\right]-\mu &\le 0 && i=1,\ldots,n_z\\
\left(\begin{bmatrix} U_L^{-L}\\ U_L^{\perp}\end{bmatrix}\mathcal{E}_m(T_{zw}^{+}-T_{zw}^{-})\begin{bmatrix} V_R^{-R} & V_R^{\perp}\end{bmatrix}\right)_{ij}(k) &= \tilde{H}_{ij}(k)\\
& && (i,j)\in \boldsymbol{S},\quad k\in\{0,\ldots,n\}\\
\mathcal{T}_I\mathcal{E}_m(T_{zw}^{+}-T_{zw}^{-}) &= \mathcal{T}_I H\\
(\mu,T_{zw}^{+},T_{zw}^{-}) &\in \mathbf{R}_{+}\times\mathbf{R}_{+}^{n_z\times n_w\times(m+1)}\\
&\qquad\times\mathbf{R}_{+}^{n_z\times n_w\times(m+1)}
\end{aligned}
$$

Where $\mathcal{E}_m:\mathbf{R}_{+}^{n_z\times n_w\times(m+1)}\mapsto l_1^{n_z\times n_w}$ is the same embedding operator as in the formulation of $\overline{LP}_n$ above.

There are $2n_zn_w(m+1)+1$ variables in $\underline{LP}_{n,m}(H,U,V)$: μ and

$$\left\{(T_{zw}^{+})_{ij}(k),\ (T_{zw}^{-})_{ij}(k): i\in\{1,\ldots,n_z\},\ j\in\{1,\ldots,n_w\},\ k\in\{0,\ldots,m\}\right\}.$$

Proposition 3.5.6 *For every n,m $\underline{LP}_{n,m}(H,U,V)$ is equivalent to $\underline{OPT}_{n,m}(H,U,V)$ under the map $\underline{\psi}_{n,m}$ defined, given a feasible solution $(\mu,T_{zw}^{+},T_{zw}^{-})$ for $\underline{LP}_n(H,U,V)$,*

$$\underline{\psi}_{n,m}(\mu,T_{zw}^{+},T_{zw}^{-}) := H-\mathcal{E}_m(T_{zw}^{+}-T_{zw}^{-}).$$

Proof: The proof of Proposition 3.5.5 goes through identically if we define the map $\tilde{\underline{\psi}}_{n,m}$, given $K\in\underline{\boldsymbol{K}}_{n,m}(H,U,V)$,

$$\tilde{\underline{\psi}}_{n,m}K := \left(\|H-K\|_{l_1},\ \Pi_m\Pi_{+}(H-K),\ \Pi_n\Pi_{+}(K-H)\right).$$

□

3.6 Example

In this section an l_∞ disturbance rejection example first considered in [27] and later in [28] is first described and then solved using the results of Sections 3.3 and 3.4 to obtain FIR approximations to the optimal closed loop impulse response and a converging lower bound on the optimal performance, respectively.

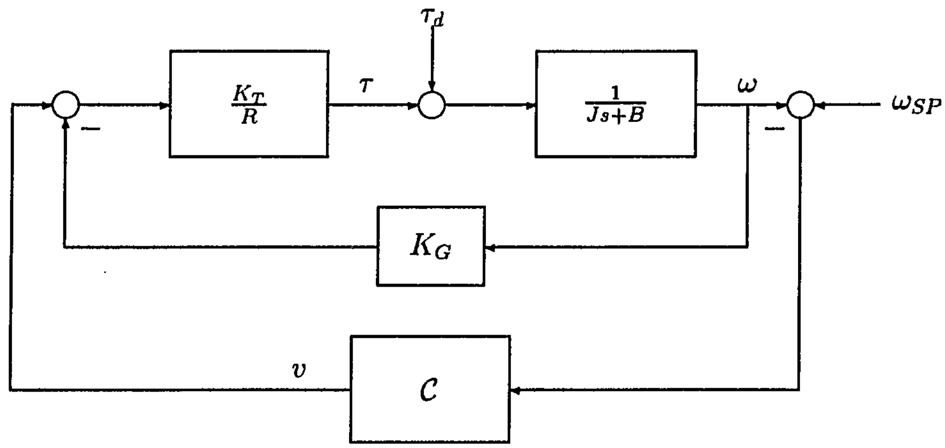

Figure 3.2: Block diagram of DC motor control problem

The objective is to limit the peak magnitude of the control effort required in controlling the speed of a DC motor near a given setpoint. Both required torque and compensator output voltage are to be bounded. The situation is pictured in Figure 3.2, where ω, τ, and v denote the motor speed, the control torque, and the compensator output voltage, respectively. τ_d denotes an unknown but bounded disturbance torque. If all constants are taken to be one, the system is linearized about the setpoint $\omega_{SP} = 1$ rad/sec, $\tau_{SP} = 1$ N-m, $v_{SP} = 2$ V, the resulting linear model in the deviation variables discretized at a sampling period of $T = .1$ sec, and

$$z := \begin{bmatrix} \delta\tau \\ \delta v \end{bmatrix}, \quad w := \tau_d, \quad u := \delta v, \quad y := \delta\omega$$

then the generalized plant transfer function matrix $\hat{G}$ is

$$\hat{G} = \left[\begin{array}{c|c} \frac{-(1-\alpha)z}{1-\beta z} & \frac{1-\alpha z}{1-\beta z} \\ 0 & 1 \\ \hline \frac{-(1-\alpha)z}{1-\beta z} & \frac{-(1-\alpha)z}{1-\beta z} \end{array}\right]$$

where $\alpha \approx .9094$ and $\beta \approx .8187$. Note that $G_{yu}(0) = \hat{G}_{yu}(0) = 0$ so that $\boldsymbol{Q}(\mathcal{G}_{yu}) = l_1$. $\hat{G}_{yu}$ has a polynomial coprime factorization (i.e., G_{yu} has a coprime factorization over the finitely supported sequences, and hence l_1) so that a Bezout equation as in Lemma 2.3.6 can be found. In terms of z-transforms it is

$$\begin{bmatrix} 1 & \frac{-\beta}{1-\alpha} \\ (1-\alpha)z & 1-\beta z \end{bmatrix} \begin{bmatrix} 1-\beta z & \frac{\beta}{1-\alpha} \\ -(1-\alpha)z & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

n	$\overline{\mu}_n$	$\underline{\mu}_{n,n}$
10	.4398	.2524
20	.3488	.3058
40	.3336	.3295
60	.3333	.3328
80	.3333	.3332

Table 3.1: Convergence of upper and lower bounds on μ_{OPT} in DC motor example.

The z-transforms of H, U, and V of Proposition 2.3.9 are then found to be

$$\hat{H} = \begin{bmatrix} -\alpha \\ -\beta \end{bmatrix}, \quad \hat{U} = \begin{bmatrix} 1-\alpha z \\ 1-\beta z \end{bmatrix}, \quad \hat{V} = \alpha - 1.$$

Hence H, U, and V are all finitely supported, therefore in l_1, and therefore $\mathcal{G}$ is stabilizable.

Simple decompositions of U and V of the form in Assumption 3.0.1 are given by

$$U_L = U, \; \hat{\Sigma}_U = \hat{U}_R = 1 \text{ and } V_L = V, \; \hat{\Sigma}_V = \hat{V}_R = 1$$

Note that $\boldsymbol{Z} = \emptyset$ and only one of the Bezout equations (3.2), (3.3) is non-trivial; it is given in terms of z-transforms by

$$\begin{bmatrix} \frac{-\beta}{\alpha-\beta} & \frac{\alpha}{\alpha-\beta} \\ 1-\beta z & -1+\alpha z \end{bmatrix} \begin{bmatrix} 1-\alpha z & \frac{\alpha}{\alpha-\beta} \\ 1-\beta z & \frac{\beta}{\alpha-\beta} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This is all the data required to solve the problem. Because all factorizations and Bezout equations have been in terms of finitely supported sequences, Proposition 3.3.2 can be used to show that $\overline{OPT}_n(H,U,V)$ will be feasible for some n (this is trivial in this case since H is finitely supported, and $K = 0$ is always a feasible solution). Moreover, because $\boldsymbol{Z} = \emptyset$, $\underline{\mu}_{n,n}$ is a lower bound on μ_{OPT} for every n. Table 3.1 shows the results of solving $\overline{OPT}_n(H,U,V)$ and $\underline{OPT}_{n,n}(H,U,V)$ for several values of n. Evidently $\mu_{OPT} \approx \frac{1}{3}$ so that the l_∞ disturbance rejection specification can be satisfied with a safety margin of roughly a factor of 3.

3.7 Discussion

In this chapter, a solution to the l_∞ disturbance rejection problem, or l_1 problem, has been obtained. The approach is to formulate the problem first in the standard setting and then as a minimum distance problem in l_1 using the parametrization of stabilizing controllers in Chapter 2. In this formulation it is easiest to establish existence of the minimizer, and insight into the properties of the FIR sub-optimal and the super-optimal approximation problems is clearest. For actual computation, the problems are recast as linear programs.

From a practical viewpoint, the results of this chapter allow approximate or exact solutions to the l_∞ disturbance rejection problem to be found in many cases. For example:

- If U and V have full row and full column rank, respectively, a sequence of feasible FIR approximation problems $\overline{OPT}_n(H,U,V)$ is guaranteed and can be solve for increasing n to yield an optimal solution to OPT and hence an optimal compensator. Although no bound is given on the n which might be required to obtain the optimal solution (and it has been shown in [29] that this can be arbitrarily large), bounds can be computed [5] if desired. Otherwise, experience shows that simply increasing n until no further improvement is obtained is satisfactory. Convergence can be checked by solving super-optimal problems $\underline{OPT}_n(H,U,V)$; they are always feasible and if the infimal costs agree, then an optimal compensator has been found.

- If U and V do not satisfy the above rank assumptions, approximate solutions and the value of μ_{DR} can be obtained to arbitrary accuracy using sub- and super-optimal approximations. If there is no feasible sequence of FIR approximations, one can be obtained by approximating $\mathcal{G}$ by an FIR system and a sequence of sub-optimal compensators obtained. A sequence of super-optimal problems can be solved as well to bound μ_{DR} from below.

- Stabilizing compensators can be obtained from feasible or optimal solutions of $\underline{OPT}_n(H,U,V)$, and this is another principal attraction of solving them. It has been noted in examples [30] that optimal compensators can sometimes be discovered in this way which have significantly lower order than those produced by FIR approximation. Moreover, even if optimal compensators are not found, the suboptimal ones which can be computed from the solutions of $\underline{OPT}_n(H,U,V)$ often approach optimal, and are often of lower order than those obtained via

FIR approximation, for the same performance level.

As mentioned in Chapter 1, the key elements of the solution to the l_1 problem are due to Dahleh and Pearson in [4][5][6]. In these three papers, the results on existence (Theorem 3.2.4), and the finite support property of the minimizer in case U and V have full row rank and full column rank, respectively, were obtained and the FIR approximation methods were described. The problem setting in each of the three papers was successively more general; only [6] considered the case in which the just mentioned rank assumptions on U and V were not satisfied, and it did so for a special case ($n_z = n_w = 2$; $n_y = n_u = 1$).

An earlier version of Theorem 3.2.1 characterizing the feasible subspace of OPT in the general case was first given by the author in [15], along with earlier versions of the existence theorems 3.2.4 and 3.2.7. An earlier version of Proposition 3.3.2 providing a test for the applicability of the FIR approximation method also appeared for the first time in [15].

The super-optimal approximation method was first suggested in [30] in which an example was solved. The idea of obtaining stabilizing compensators from the solutions to super-optimal problems was also suggested there, and was used in that example to discover an optimal compensator of much lower order than the sub-optimal ones obtained via FIR approximation. The promise of possibly obtaining low order optimal compensators makes the study of these problems particularly interesting. The results of [30] have been gradually generalized in [31] and [32].

Although the results reported here overlap those of [30][31][32], the formulations are sufficiently different that it is unclear whether the results are equivalent. Moreover, the lower bound theorem 3.4.1 of this chapter combined with the details of the relationship between correlation and convolution operators on c_0 outlined in Appendix C make the key results (the norm computations) of chapters 4 and 5 possible.

An alternate method for the formulation and solution of converging super-optimal problems by a "delay augmentation" method has recently been developed [33][34]. Its properties are similar in many respects to those described here, although the approach taken is seemingly very different.

Chapter 4

Incremental Weighted l_∞ Disturbance Rejection

The problem setting is again the general feedback system of Chapters 2 and 3 and the problem to be considered is the incremental weighted l_∞ design problem described in Chapter 1. Given the definitions of Chapter 2, the associated specification can be stated precisely.

Incremental Weighted l_∞ DRS:

- $\mathcal{C} \in \Sigma_{clti}$, $(\mathcal{G}, \mathcal{C})$ is stable, and
- $\|\mathcal{P}_N \mathcal{W}_w w\|_{l_\infty} \leq 1$ implies $\|\mathcal{P}_N \mathcal{W}_z z\|_{l_\infty} \leq 1$.

where the weights $\mathcal{W}_z$ and $\mathcal{W}_w$ are systems in $\Sigma^{n_z \mapsto n_{\tilde{z}}}$ and $\Sigma^{n_w \mapsto n_{\tilde{w}}}$, respectively. Note that the above norms are always defined since every truncation of every signal in l_+ is in $l_{\infty+}$.

To satisfy the specification $\mathcal{C}$ must be causal, linear, and time invariant and it must stabilize $\mathcal{G}$. Moreover, for any time and for any suddenly applied disturbance whose weighted peak magnitude *up until that time* is less than or equal to one, the resulting error's weighted peak magnitude must not have exceeded one *up until that time.*

In Chapter 3, the design problem could be formulated immediately as minimization of the l_1 norm of the closed loop system $\mathcal{T}_{zw}(\mathcal{G}, \mathcal{C})$. For this problem, an appropriate norm must first be defined and, to solve the problem, a means of computing the norm is needed.

Hence Section 4.1 will define, under some assumptions on the weights, an appropriate notion of stability and corresponding norm on the causal linear time invariant stable systems. This is followed by the formulation of the design problem in Section 4.2 under the same assumptions on $\mathcal{G}$ as were in effect in Chapter 3, and $\mathcal{W}_z$ and $\mathcal{W}_w$ are assumed to be in $\boldsymbol{\Sigma}_{clti}$, stable, and to have stable left inverses in $\boldsymbol{\Sigma}_{clti}$.

Sections 4.3 through 4.7 mirror Sections 3.1 through 3.5. In Section 4.3 the norm minimization problem is reformulated as a minimum distance problem in $l_1^{n_{\bar{z}} \times n_{\bar{w}}}$. In Section 4.4 it is shown that a minimizer exists. The minimum distance problem is again infinite dimensional as in Chapter 3 and approximate solution methods are given in Sections 4.5 and 4.6. Section 4.5 gives a sequence of finite dimensional optimization problems whose solutions converge to the optimal performance from above, and Section 4.6 gives a method for bounding the optimal performance from below via finite dimensional optimization problems. Stabilizing compensators can be obtained using this method as in Chapter 3. Section 4.7 gives linear programming formulations of all the problems posed in the chapter, and Section 4.8 contains an example of norm computation for a fixed system (i.e., not a design problem). The chapter concludes in Section 4.9 with a discussion of the results, related work, and the main contributions of the chapter.

4.1 Incremental Weighted Stability and Gain

In this section, the symbol $\mathcal{G}$ is used to denote some given system in $\boldsymbol{\Sigma}$, and should not be confused with the generalized plant $\mathcal{G}$. The following notion of stability is appropriate to the incremental weight design problem.

Definition 4.1.1 *A system* $\mathcal{G} \in \boldsymbol{\Sigma}^{n \mapsto m}$ *is* incrementally stable w.r.t. $\mathcal{W}_o \in \boldsymbol{\Sigma}^{m \mapsto \bar{m}}$, $\mathcal{W}_i \in \boldsymbol{\Sigma}^{n \mapsto \bar{n}}$ *if there exists* $c < \infty$ *such that*

$$\|\mathcal{P}_N \mathcal{W}_o \mathcal{G} x\|_{l_\infty} \leq c \|\mathcal{P}_N \mathcal{W}_i x\|_{l_\infty} \ \forall x \in l_+ \text{ and } N \in \mathbf{Z} \tag{4.1}$$

If $\mathcal{G}$ *is incrementally stable w.r.t.* $\mathcal{W}_o$, $\mathcal{W}_i$, *the smallest* c *satisfying (4.1) is the* incremental gain of $\mathcal{G}$ w.r.t. $\mathcal{W}_o, \mathcal{W}_i$ *and is denoted by* $\rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i)$.

Under this definition, a system which is incrementally stable w.r.t. given weights is guaranteed at every time to produce weighted errors not exceeding *some* finite bound, provided that the weighted disturbances are bounded.

The next proposition provides a test for incremental stability w.r.t. a large class of weights. There is no restriction on the output weights, and the requirement on the input weight is quite mild.

Proposition 4.1.2 *Let* $\mathcal{G} \in \Sigma^{n \mapsto m}$, $\mathcal{W}_\mathrm{o} \in \Sigma^{m \mapsto \tilde{m}}$, *and* $\mathcal{W}_\mathrm{i} \in \Sigma^{n \mapsto \tilde{n}}$, *and let* $\mathcal{W}_\mathrm{i}^{-L} \in \Sigma^{\tilde{n} \mapsto n}$ *be such that* $\mathcal{W}_\mathrm{i}^{-L}\mathcal{W}_\mathrm{i} = \mathcal{I}$ *and* $\mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}$ *is stable and causal.* $\mathcal{G}$ *is incrementally stable w.r.t.* $\mathcal{W}_\mathrm{o}$, $\mathcal{W}_\mathrm{i}$ *if and only if* $\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}$ *is stable and causal.*

Proof: (if) Suppose $\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}$ is stable and causal and let $x \in l_+^n$ and $N \in \mathbf{Z}$ be arbitrary. Then

$$\begin{aligned}\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}x\|_{l_\infty} &= \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\mathcal{W}_\mathrm{i}x\right\|_{l_\infty} = \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\mathcal{P}_N\mathcal{W}_\mathrm{i}x\right\|_{l_\infty} \\ &\leq \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\right\|_{l_\infty\text{-i}} \|\mathcal{P}_N\mathcal{W}_\mathrm{i}x\|_{l_\infty} \\ &\leq \left\|\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\right\|_{l_\infty\text{-i}} \|\mathcal{P}_N\mathcal{W}_\mathrm{i}x\|_{l_\infty}\end{aligned}$$

and (4.1) is satisfied by $c := \left\|\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\right\|_{l_\infty\text{-i}}$. Hence $\mathcal{G}$ is incrementally stable w.r.t. $\mathcal{W}_\mathrm{o}$, $\mathcal{W}_\mathrm{i}$.

(only if) Suppose $\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}$ is not causal. Then there exist $\tilde{x} \in l_+$ and $N \in \mathbf{Z}$ such that $\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\tilde{x} \neq \mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\mathcal{P}_N\tilde{x}$ and $x := \mathcal{W}_\mathrm{i}^{-L}(\mathcal{I} - \mathcal{P}_N)\tilde{x}$ is such that

$$\begin{aligned}\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}x\|_{l_\infty} &= \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}(\mathcal{I} - \mathcal{P}_N)\tilde{x}\right\|_{l_\infty} \\ &= \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\tilde{x} - \mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\mathcal{P}_N\tilde{x}\right\|_{l_\infty} \\ &> 0.\end{aligned}$$

On the other hand,

$$\mathcal{P}_N\mathcal{W}_\mathrm{i}x = \mathcal{P}_N\mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}(\mathcal{I} - \mathcal{P}_N)\tilde{x} = \mathcal{P}_N\mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}\mathcal{P}_N(\mathcal{I} - \mathcal{P}_N)\tilde{x} = 0$$

since $\mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}$ is causal and $\mathcal{P}_N$ is a projection. Hence there does not exist $c < \infty$ satisfying (4.1) and $\mathcal{G}$ is not incrementally stable w.r.t. $\mathcal{W}_\mathrm{o}$, $\mathcal{W}_\mathrm{i}$.

Suppose $\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}$ is unstable. If there exists $\tilde{x} \in l_{\infty+}^{\tilde{n}}$ such that $\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\tilde{x} \notin l_{\infty+}^{\tilde{m}}$, define $x := \mathcal{W}_\mathrm{i}^{-L}\tilde{x}$. Then, since $\mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}$ is stable, $\mathcal{W}_\mathrm{i}x = \mathcal{W}_\mathrm{i}\mathcal{W}_\mathrm{i}^{-L}\tilde{x} \in l_{\infty+}^{\tilde{n}}$. Moreover, given any $c < \infty$, there exists $N \in \mathbf{Z}$ such that

$$\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}x\|_{l_\infty} = \left\|\mathcal{P}_N\mathcal{W}_\mathrm{o}\mathcal{G}\mathcal{W}_\mathrm{i}^{-L}\tilde{x}\right\|_{l_\infty} > c\,\|\mathcal{W}_\mathrm{i}x\|_{l_\infty} \geq c\,\|\mathcal{P}_N\mathcal{W}_\mathrm{i}x\|_{l_\infty}.$$

Hence there does not exist $c < \infty$ satisfying (4.1) and $\mathcal{G}$ is not incrementally stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$.

If such an $\tilde{x}$ does not exist then, given any $c < \infty$, there must exist $\tilde{x} \in l_{\infty+}^{\tilde{n}}$ such that

$$\left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x}\right\|_{l_\infty} > c\left\|\mathcal{W}_i\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\tilde{x}\|_{l_\infty}.$$

If $x := \mathcal{W}_i^{-L}\tilde{x}$ then

$$\begin{aligned}\sup\left\{\|\mathcal{P}_N\mathcal{W}_o\mathcal{G}x\|_{l_\infty} : N \in \mathbf{Z}\right\} &= \|\mathcal{W}_o\mathcal{G}x\|_{l_\infty} = \left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x}\right\|_{l_\infty} \\ &> c\left\|\mathcal{W}_i\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\tilde{x}\|_{l_\infty}\end{aligned}$$

Hence there exists $N \in \mathbf{Z}$ such that

$$\|\mathcal{P}_N\mathcal{W}_o\mathcal{G}x\|_{l_\infty} > c\left\|\mathcal{W}_i\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\tilde{x}\|_{l_\infty} \geq c\,\|\mathcal{W}_ix\|_{l_\infty} \geq c\,\|\mathcal{P}_N\mathcal{W}_ix\|_{l_\infty}$$

and therefore c does not satisfy (4.1). Since c was arbitrary, the proof is complete. □

Assumption 4.1.3 *In the remainder of this section $\mathcal{G}$ denotes a system in $\Sigma_{clti}^{n\mapsto m}$. $\mathcal{W}_o \in \Sigma_{clti}^{n\mapsto\tilde{n}}$ and $\mathcal{W}_i \in \Sigma_{clti}^{m\mapsto\tilde{m}}$ denote stable systems with stable left inverses in Σ_{clti}. Hence they have impulse response matrices $W_o \in l_1^{\tilde{n}\times n}$ and $W_i \in l_1^{\tilde{m}\times m}$ which are left invertible in l_1. Because l_1 is Hermite, there exist additional matrices over l_1 satisfying the Bezout equations*

$$\begin{bmatrix} W_o^{-L} \\ W_o^{\perp} \end{bmatrix}\begin{bmatrix} W_o & W_o^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} W_i^{-L} \\ W_i^{\perp} \end{bmatrix}\begin{bmatrix} W_i & W_i^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.2}$$

In the remainder of this section, the additional symbols above will denote arbitrary *choices satisfying the equations, given W_o and W_i.*

The next proposition shows that the set of systems in Σ_{clti} which are incrementally stable w.r.t. weights as in Assumption 4.1.3 is precisely the set of stable systems. Moreover, the incremental weighted gain for any such choice of weights is a norm on the space of stable systems in Σ_{clti}.

Proposition 4.1.4 *$\mathcal{G}$ is incrementally stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ if and only if $G \in l_1$. Moreover, $\|\cdot\|_{\mathcal{W}_o,\mathcal{W}_i}^{i} := \rho_i(\cdot;\mathcal{W}_o,\mathcal{W}_i)$ is a norm on the space of stable systems in $\Sigma_{clti}^{n\mapsto m}$.*

Proof: Because $\mathcal{W}_i$ satisfies the hypotheses of Proposition 4.1.2 that proposition can be applied to prove the first statement; it says that $\mathcal{G}$ is incrementally stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ if and only if $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is stable and causal, where $\mathcal{W}_i^{-L}$ is any stable left inverse of $\mathcal{W}_i$ in Σ_{clti}. Now if $G \in l_1$ then $W_oGW_i^{-L} \in l_1$ and $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is stable and causal. Conversely, if $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is stable and causal then $W_oGW_i^{-L} \in l_1$; hence $G = W_o^{-L}(W_oGW_i^{-L})W_i \in l_1$.

For the second statement, $\|\cdot\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ is non-negative by definition and is well defined for every stable system. It is easy to check that for linear $\mathcal{W}_o$ and $\mathcal{W}_i$, $\|\cdot\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ scales and is subadditive. Moreover, $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} = 0$ implies that $\mathcal{P}_N\mathcal{G}x = 0$ for all $x \in l_+$ and all $N \in \mathbf{Z}$ which implies that $\mathcal{G}$ is the zero system. Hence $\|\cdot\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ is a norm on the stable systems. □

Theorem 4.1.6 will show that $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ can be computed for any $\mathcal{G}$ which is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ by solving a minimum distance problem $OPT(H,U,V)$ as defined in Section 3.1 for appropriate choices of H, U, and V. Specifically, since $W_oGW_i^{-L}$ and $W_i^{\perp}$ are in l_1 the infimization

$$\inf\left\{\left\|W_oGW_i^{-L} - K\right\|_{l_1} : K \in \boldsymbol{K}(I, W_i^{\perp})\right\} =: \gamma. \tag{4.3}$$

is precisely $OPT(W_oGW_i^{-L}, I, W_i^{\perp})$. It is easy to see that I and $W_i^{\perp}$ have decompositions as in Assumption 3.0.1; the obvious ones are:

$$U_L = \Sigma_U = U_R = I \quad \text{and} \quad V_L = \Sigma_V = I,\ V_R = W_i^{\perp}. \tag{4.4}$$

Clearly $\boldsymbol{Z} = \emptyset$ and m_ι is finite $(= 0)$. Hence all results of Chapter 3 can be applied to the problem (4.3).

The Bezout equations (3.2) and (3.3) associated with the decompositions (4.4) have a particularly simple form. None are defined for U_R or V_L since $U_R = I$ and $V_L = I$, and none is defined for U_L when $U_L = I$. The equations for V when $V = W_i^{\perp}$ have the form

$$\begin{bmatrix} W_i^{\perp} \\ W_i^{-L} \end{bmatrix} \begin{bmatrix} W_i^{c} & W_i \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.5}$$

The special form of (4.3), in particular the fact that $U = I$, is used in the next lemma, which will facilitate the proof of Theorem 4.1.6.

Lemma 4.1.5 *If μ_i is defined for each $i \in \{1, \ldots, \tilde{m}\}$*

$$\mu_i := \inf \left\{ \left\| (W_o G W_i^{-L} - K)_{i\cdot} \right\|_{l_1} : K \in \boldsymbol{K}(I, W_i^{\perp}) \right\}$$

then $\gamma = \max\{\mu_i : i \in \{1, \ldots, \tilde{m}\}\}$.

Proof: Clearly $\gamma \geq \max\{\mu_i : i \in \{1, \ldots, \tilde{m}\}\}$ since

$$\left\| W_o G W_i^{-L} - K \right\|_{l_1} \geq \left\| (W_o G W_i^{-L} - K)_{i\cdot} \right\|_{l_1}$$

for all i and all $K \in \boldsymbol{K}(I, W_i^{\perp})$. To show the reverse inequality note that for each i, $K \in \boldsymbol{K}(I, W_i^{\perp})$ if and only if $K_{i\cdot} = QW_i^{\perp}$ for some $Q \in l_1^{1\times(\tilde{n}-n)}$. Hence, given $\epsilon > 0$, there exists for each i $Q_i \in l_1^{1\times(\tilde{n}-n)}$ such that $\left\| (W_o G W_i^{-L})_{i\cdot} - Q_i W_i^{\perp} \right\|_{l_1} \leq \mu_i + \epsilon$. Defining

$$Q_0 := \begin{bmatrix} Q_1 \\ \vdots \\ Q_{\tilde{m}} \end{bmatrix} \in l_1^{\tilde{m}\times(\tilde{n}-n)}$$

we have that $K_0 := Q_0 W_i^{\perp} \in \boldsymbol{K}(I, W_i^{\perp})$ and

$$\begin{aligned}
\left\| W_o G W_i^{-L} - K_0 \right\|_{l_1} &= \max\left\{ \left\| (W_o G W_i^{-L} - K_0)_{i\cdot} \right\|_{l_1} : i \in \{1, \ldots, \tilde{m}\} \right\} \\
&= \max\left\{ \left\| (W_o G W_i^{-L})_{i\cdot} - Q_i W_i^{\perp} \right\|_{l_1} : i \in \{1, \ldots, \tilde{m}\} \right\} \\
&\leq \max\{\mu_i : i \in \{1, \ldots, \tilde{m}\}\} + \epsilon
\end{aligned}$$

Since ϵ was arbitrary, the proof is complete. □

Theorem 4.1.6 $\|\mathcal{G}\|^{i}_{\mathcal{W}_o, \mathcal{W}_i} = \gamma$.

Proof: $\|\mathcal{G}\|^{i}_{\mathcal{W}_o, \mathcal{W}_i} \leq \gamma$: By Theorem 3.2.4, $OPT(W_o G W_i^{-L}, I, W_i^{\perp})$ has a minimizer and there exists $Q_0 \in l_1^{\tilde{m}\times(\tilde{n}-n)}$ such that $\left\| W_o G W_i^{-L} - Q_0 W_i^{\perp} \right\|_{l_1} = \gamma$. Moreover, for all $N \in \mathbf{Z}$ and $w \in l^n_{\infty+}$,

$$\begin{aligned}
\|\mathcal{P}_N \mathcal{W}_o \mathcal{G} w\|_{l_\infty} &= \left\| \mathcal{P}_N (\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L} - \mathcal{Q}_0 \mathcal{W}_i^{\perp}) \mathcal{W}_i w \right\|_{l_\infty} \\
&= \left\| \mathcal{P}_N (\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L} - \mathcal{Q}_0 \mathcal{W}_i^{\perp}) \mathcal{P}_N \mathcal{W}_i w \right\|_{l_\infty} \\
&\leq \left\| \mathcal{P}_N (\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L} - \mathcal{Q}_0 \mathcal{W}_i^{\perp}) \right\|_{l_\infty\text{-i}} \|\mathcal{P}_N \mathcal{W}_i w\|_{l_\infty} \\
&\leq \left\| W_o G W_i^{-L} - Q_0 W_i^{\perp} \right\|_{l_1} \|\mathcal{P}_N \mathcal{W}_i w\|_{l_\infty} \\
&= \gamma \|\mathcal{P}_n \mathcal{W}_i w\|_{l_\infty}.
\end{aligned}$$

Hence (4.1) is satisfied by $c := \gamma$ and $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} \leq \gamma$.

$\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} \geq \gamma$: Given any $\epsilon > 0$ we will find $w \in l^n_{\infty+}$ and $N \in \mathbf{Z}$ such that $\|\mathcal{P}_N\mathcal{W}_i w\|_{l_\infty} \leq 1$ but $\|\mathcal{P}_N\mathcal{W}_o\mathcal{G}w\|_{l_\infty} \geq \gamma - \epsilon$. Hence, using Definition 4.1.1, $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} \geq \gamma - \epsilon$. Since ϵ was arbitrary, $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} \geq \gamma$.

Accordingly, let $\epsilon > 0$ be given. Lemma 4.1.5 implies that

$$\inf\left\{\left\|(W_oGW_i^{-L})_{i\cdot} - QW_i^{\perp}\right\|_{l_1} : Q \in l_1^{1\times(\tilde{n}-n)}\right\} \geq \gamma$$

for some $i \in \{1,\ldots,\tilde{m}\}$. Let this value of i be fixed for the remainder of the proof. By the lower bound theorem 3.4.1 there exists $N \in \mathbf{Z}_+$ such that $\underline{\mu}_N > \gamma - \frac{\epsilon}{2}$. Moreover, using equation (3.16) from the proof of that theorem, we find that there exists $x \in \mathcal{P}_N c_0^{1\times n}$ such that

$$\left\langle \tilde{\mathcal{T}}_K x, (W_oGW_i^{-L})_{i\cdot}\right\rangle > \gamma - \epsilon \tag{4.6}$$

and

$$\left\|\tilde{\mathcal{T}}_K x\right\|_{l_\infty} \leq 1 \tag{4.7}$$

Now define $w \in l^n_{\infty+}$ by

$$w(k) := \begin{cases} x(-k) & k \leq 0 \\ 0 & k > 0 \end{cases}$$

This is the offending w.

According to the definition of $\tilde{\mathcal{T}}_K$ in equation (3.10) $\tilde{\mathcal{T}}_K x = x \triangleright W_i^T$ where $\triangleright$ denotes right correlation (see Definitions C.1.2 and C.1.4) and we have used the Bezout equation (4.5). Moreover, for each $i \in \{1,\ldots,\tilde{n}\}$ and $k \in \mathbf{Z}_+$,

$$\begin{aligned}
(\tilde{\mathcal{T}}_K x)_i(k) &= (x \triangleright W_i^T)_i(k) = \sum_{j=1}^{n}\left[x_j \triangleright (W_i)_{ij}\right](k) \\
&= \sum_{j=1}^{n}\sum_{n=0}^{\infty}(W_i)_{ij}(n)x_j(n+k) = \sum_{j=1}^{n}\sum_{n=0}^{\infty}(W_i)_{ij}(n)w_j(-n-k) \\
&= \sum_{j=1}^{n}\left[(W_i)_{ij} * w_j\right](-k)
\end{aligned}$$

and hence, for all $k \in \mathbf{Z}_+$,

$$(\tilde{\mathcal{T}}_K x)(k) = (W_i * w)^T(-k).$$

Using (4.7) and the definition of $\|\cdot\|_{l_\infty}$, we conclude that $\|\mathcal{P}_0\mathcal{W}_\mathrm{i}w\|_{l_\infty} \leq 1$.

On the other hand,

$$\begin{aligned}\left\langle \tilde{\mathcal{T}}_K x, (W_\mathrm{o}GW_\mathrm{i}^{-L})_{i\cdot}\right\rangle &= \left\langle x \triangleright W_\mathrm{i}^T, (W_\mathrm{o}GW_\mathrm{i}^{-L})_{i\cdot}\right\rangle = \langle x, (W_\mathrm{o}G)_{i\cdot}\rangle \\ &= \sum_{j=1}^{n}\sum_{n=0}^{\infty}(W_\mathrm{o}G)_{ij}(n)x_j(n) = \sum_{j=1}^{n}\sum_{n=0}^{\infty}(W_\mathrm{o}G)_{ij}(n)w_j(-n) \\ &= (W_\mathrm{o}G * x)_i(0)\end{aligned}$$

where the first line follows using the fact (Proposition C.1.5) that the adjoint of a correlation operator is a convolution operator and the second using the definitions of functional evaluation (see Fact C.1.1) and w. Finally, using (4.6), we conclude that $\|\mathcal{P}_0\mathcal{W}_\mathrm{o}\mathcal{G}w\|_{l_\infty} > \gamma - \epsilon$, and the proof is complete. □

As a consequence of this theorem, computation of the incremental gain of a system w.r.t. given weights or, equivalently, determining if the system satisfies the second item of the corresponding incremental weighted specification requires the solution of $OPT(W_\mathrm{o}GW_\mathrm{i}^{-L}, I, W_\mathrm{i}^{\perp})$. As mentioned above, Assumption 3.0.1 is satisfied so we can apply the results of Chapter 3 as follows:

- Theorem 3.2.4 implies that a minimizer exists for (4.3).
- Proposition 3.3.1 implies that if there exists $K_\mathrm{fs} \in \boldsymbol{K}(I, W_\mathrm{i}^{\perp})$ such that

  $$W_\mathrm{o}GW_\mathrm{i}^{-L} - K_\mathrm{fs}$$

  is finitely supported then a sequence of problems $\overline{OPT}_n(H, U, V)$ can be solved to bound $\|\mathcal{G}\|^i_{\mathcal{W}_\mathrm{o},\mathcal{W}_\mathrm{i}}$ from above.
- If W_i is finitely supported and has a finitely supported left inverse then Proposition 3.3.2 implies that $K_\mathrm{fs} \in \boldsymbol{K}(I, W_\mathrm{i}^{\perp})$ exists if and only if $(W_\mathrm{o}GW_\mathrm{i}^{-L})W_\mathrm{i} = W_\mathrm{o}G$ is finitely supported.
- The results of Section 3.4 give a method for computing a converging lower bound on $\|\mathcal{G}\|^i_{\mathcal{W}_\mathrm{o},\mathcal{W}_\mathrm{i}}$. In particular, since $\boldsymbol{Z} = \emptyset$, Proposition 3.4.3 implies that it suffices to solve a sequence of problems $\underline{OPT}_{n,n}(H, U, V)$ for increasing n.
- $\boldsymbol{Z} = \emptyset$ also implies that there are no interpolation constraints in the linear programming formulation of (4.3).

4.2 Problem Statement

Recall the specification:

Incremental Weighted l_∞ DRS:

- $\mathcal{C} \in \Sigma_{clti}$, $(\mathcal{G},\mathcal{C})$ is stable, and
- $\|\mathcal{P}_N \mathcal{W}_w w\|_{l_\infty} \leq 1$ implies $\|\mathcal{P}_N \mathcal{W}_z z\|_{l_\infty} \leq 1$.

where the weights $\mathcal{W}_z$ and $\mathcal{W}_w$ are systems in $\Sigma^{n_z \mapsto n_{\bar{z}}}$ and $\Sigma^{n_w \mapsto n_{\bar{w}}}$, respectively.

If $\mathcal{G} \in \Sigma_{clti}$ and $\mathcal{W}_z, \mathcal{W}_w \in \Sigma_{clti}$ are stable and have stable left inverses then the results of Section 4.1 can be used to formulate the design problem as follows.

$$DR^{iw}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w): \qquad \inf\left\{\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|^{i}_{\mathcal{W}_z,\mathcal{W}_w} : \mathcal{C} \in \mathcal{C}(\mathcal{G})\right\} =: \mu_{IWDR}$$

If $\mathcal{C} \in \mathcal{C}(\mathcal{G})$ then $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ is stable and hence the norm above is always defined; μ_{IWDR} is defined if and only if $\mathcal{G}$ is stabilizable. Moreover, if $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ is stable then

$$\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|^{i}_{\mathcal{W}_z,\mathcal{W}_w} = \sup\left\{\|\mathcal{P}_N \mathcal{W}_z z\|_{l_\infty} : N \in \mathbf{Z} \text{ and } \|\mathcal{P}_N \mathcal{W}_w w\|_{l_\infty} \leq 1\right\}.$$

$DR^{iw}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ represents the design problem in the same sense that $DR(\mathcal{G})$ represents the unweighted l_∞ disturbance rejection design problem. In particular, the feasible solutions for $DR^{iw}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ are the stabilizing compensators for $\mathcal{G}$ and the cost of each is the worst-case $\|\mathcal{P}_N \mathcal{W}_z z\|_{l_\infty}$ over all $N \in \mathbf{Z}$ and all $\|\mathcal{P}_N \mathcal{W}_w w\|_{l_\infty} \leq 1$ when that $\mathcal{C}$ is used. Hence $DR^{iw}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ is feasible if and only if the first item of the specification is met, and $\mu_{IWDR} \leq 1$ if and only if the second item of the specification can be met.

Notation and Assumptions

In the remainder of this chapter all assumptions on $\mathcal{G}$ made in Chapter 3 are in effect; it is in Σ_{clti}, G_{yu} has a coprime factorization over l_1, it is stabilizable, and the associated H, U and V matrices have decompositions as in Assumption 3.0.1. All notation in Chapter 3 has the same meaning here.

The weights $\mathcal{W}_z$ and $\mathcal{W}_w$ are assumed to be in $\Sigma_{clti}^{n_z \mapsto n_{\tilde{z}}}$ and $\Sigma_{clti}^{n_w \mapsto n_{\tilde{w}}}$, respectively. Both are assumed to be stable, and to have stable left inverses in Σ_{clti}. Hence W_z and W_w are in l_1 and have left inverses in l_1 and there exist Bezout equations

$$\begin{bmatrix} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} \begin{bmatrix} W_z & W_z^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} W_w^{-L} \\ W_w^{\perp} \end{bmatrix} \begin{bmatrix} W_w & W_w^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.8}$$

In the remainder of the chapter the additional symbols on the left hand sides of equations (4.8) will denote *arbitrary* matrices over l_1 satisfying the equations.

The index set which it is convenient to define here (it plays a role similar to that of $\boldsymbol{S}$ in Chapter 3) is

$$\boldsymbol{S}_W := \{\{1,\ldots,n_{\tilde{z}}\} \times \{1,\ldots,n_{\tilde{w}}\}\} \setminus \{\{1,\ldots,r_U\} \times \{1,\ldots,r_V\}\},$$

containing the indices of all elements of an $n_{\tilde{z}} \times n_{\tilde{w}}$ matrix except those of the upper left hand corner block of dimension $r_U \times r_V$.

4.3 Formulation as a Minimum Distance Problem

The parametrization of stabilizing compensators allowed $DR(\mathcal{G})$ to be formulated as a minimum distance problem in $l_1^{n_z \times n_w}$ in Chapter 3. Because computation of $\|\mathcal{G}\|^i_{\mathcal{W}_z,\mathcal{W}_w}$ is a minimum distance problem in $l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$, $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ can also be formulated as one:

$$OPT^{iw}: \qquad \inf\left\{ \left\| W_z H W_w^{-L} - K \right\|_{l_1} : K \in \boldsymbol{K}^{iw} \right\} =: \mu^{iw}$$

where

$$\boldsymbol{K}^{iw} := \tag{4.9}$$
$$\left\{ K \in l_1 : \exists Q_C, Q_W \in l_1 \text{ satisfying } K = \begin{bmatrix} W_z U & I \end{bmatrix} \begin{bmatrix} Q_C & 0 \\ 0 & Q_W \end{bmatrix} \begin{bmatrix} V W_w^{-L} \\ W_w^{\perp} \end{bmatrix} \right\}$$

is easily verified to be a subspace of $l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$. Note that OPT^{iw} depends on H, U, V, W_z, and W_w and $\boldsymbol{K}^{iw}$ depends on all of these except H. These dependences are suppressed for notational convenience.

The following theorem shows that OPT^{iw} is equivalent to DR^{iw}. Its statement is simplified somewhat as noted after the theorem.

Theorem 4.3.1 *$DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ and OPT^{iw} are equivalent in the following sense:*

1. *If $\mathcal{C} \in \mathcal{C}(\mathcal{G})$ and $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ is constructed from it as in Proposition 2.3.11 then $K := W_z UQVW_w^{-L} + K_\rho \in \boldsymbol{K}^{iw}$ where K_ρ is a minimizer for $OPT(W_z(H - UQV)W_w^{-L}, I, W_w^{\perp})$. Moreover $\left\|W_z H W_w^{-L} - K\right\|_{l_1} = \|\mathcal{T}_{zw}(\mathcal{G}, \mathcal{C})\|^i_{\mathcal{W}_z, \mathcal{W}_w}$.*

2. *If $K \in \boldsymbol{K}^{iw}$, $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ solves $K = W_z UQVW_w^{-L} + Q_W W_w^{\perp}$ for some $Q_W \in l_1$, and $\mathcal{C}$ is constructed from it as in Proposition 2.3.11 then $\mathcal{C} \in \mathcal{C}(\mathcal{G})$ and $\|\mathcal{T}_{zw}(\mathcal{G}, \mathcal{C})\|^i_{\mathcal{W}_z, \mathcal{W}_w} \leq \left\|W_z H W_w^{-L} - K\right\|_{l_1}$.*

3. *$\mu^{iw} = \mu_{IWDR}$.*

Proof: For item 1, such a choice of K is possible because a minimizer always exists for $OPT(W_z(H - UQV)W_w^{-L}, I, W_w^{\perp})$. That $K \in \boldsymbol{K}^{iw}$ follows because every feasible solution $\bar{K}$ for $OPT(W_z(H - UQV)W_w^{-L}, I, W_w^{\perp})$ can be written $\bar{K} = Q_W W_w^{\perp}$ for some $Q_W \in l_1$; hence so can $\boldsymbol{K}_\rho$. The equality of the norms is immediate using Theorem 4.1.6 and the fact that, by Proposition 2.3.11 and the definition of K, $W_z H W_w^{-L} + K = W_z(H - UQV)W_w^{-L} + K_\rho = W_z T_{zw}(\mathcal{G}, \mathcal{C})W_w^{-L} + K_\rho$.

For item 2, the existence of a $Q \in l_1^{n_u \times n_y}$ is guaranteed by the definition of $\boldsymbol{K}^{iw}$. By Proposition 2.3.11 and the choice of Q

$$W_z T_{zw}(\mathcal{G}, \mathcal{C})W_w^{-L} - Q_W W_w^{\perp} = W_z(H - UQV)W_w^{-L} + Q_W W_w^{\perp} = W_z H W_w^{-L} - K.$$

The norm inequality follows from Theorem 4.1.6 which shows that $\|\mathcal{T}_{zw}(\mathcal{G}, \mathcal{C})\|^i_{\mathcal{W}_z, \mathcal{W}_w}$ is the infimum over $Q_W \in l_1$ of the l_1 norm of the left hand side above.

Item 3 follows from the preceding items. □

Among the consequences of Theorem 4.3.1 are the following:

- To each feasible solution for $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ there corresponds a feasible solution for OPT^{iw} of the same cost. From every feasible solution $K \in \boldsymbol{K}^{iw}$ for OPT^{iw}, one or more corresponding feasible solutions for $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ can be constructed of no greater cost.

- To each minimizer for $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ there corresponds a minimizer for OPT^{iw}; hence *all* minimizers for $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ can be found from minimizers for OPT^{iw} via the construction of item 2.

Note that item 2 of the theorem and the above discussion are somewhat simplified. Item 2 should contain a provision as in the corresponding theorem 3.1.1 of Chapter 3 which addresses the case in which $Q \notin \boldsymbol{Q}(\mathcal{G}_{yu})$. An argument similar to that in Chapter 3 shows that a random perturbation on Q can produce a stabilizing compensator $\mathcal{C}_\epsilon$ such that $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C}_\epsilon)\|^i_{W_z,W_w} \leq \left\|W_z H W_w^{-L} - K\right\|_{l_1} + \epsilon$ for any $\epsilon > 0$. Of course if $G_{yu}(0) = 0$ then Proposition 2.3.12 guarantees that this case does not arise.

The next proposition is the analog of Proposition 3.1.2; it shows how to compute, given a feasible solution K for OPT^{iw}, stabilizing compensators for $\mathcal{G}$ whose performance does not exceed $\left\|W_z H W_w^{-L} - K\right\|_{l_1}$.

Proposition 4.3.2 *If $K \in \boldsymbol{K}^{iw}$ then the set*

$$\boldsymbol{Q}^{iw}(K) := \left\{Q \in l_1^{n_u \times n_y} : K = W_z U Q V W_w^{-L} + Q_W W_w^{\perp} \text{ for some } Q_W \in l_1\right\}$$

is given by

$$\left\{\begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} W_z^{-L} K W_w V_R^{-R} \Sigma_V^{-1} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} V_L^{-L} \\ V_L^{\perp} \end{bmatrix} : Q_{12}, Q_{21}, Q_{22} \in l_1 \right\}.$$

Proof: If $K \in \boldsymbol{K}^{iw}$ then $K = W_z U \bar{Q} V W_w^{-L} + \bar{Q}_W W_w^{\perp}$ for some $\bar{Q}, \bar{Q}_W \in l_1^{n_u \times n_y}$. If Q has the claimed form then $Q \in l_1^{n_u \times n_y}$ since

$$Q = U_R^{-R} U_R \bar{Q} V_L V_L^{-L} + U_R^{\perp} Q_{12} V_L^{-L} + U_R^{-R} Q_{21} V_L^{\perp} + U_R^{\perp} Q_{22} V_L^{\perp}$$

and all matrices on the right are in l_1. Moreover,

$$W_z U Q V W_w^{-L} + \bar{Q}_W W_w^{\perp} = W_z U \bar{Q} V W_w^{-L} + \bar{Q}_W W_w^{\perp} = K$$

using the Bezout equations (3.3) and (4.8). Hence $Q \in \boldsymbol{Q}^{iw}(K)$.

Conversely, if $Q \in \boldsymbol{Q}^{iw}(K)$ then $Q \in l_1^{n_u \times n_y}$ and $K = W_z U Q V W_w^{-L} + Q_W W_w^{\perp}$ for some $Q_W \in l_1$. Now

$$\begin{aligned}
& K = W_z U Q V W_w^{-L} + Q_W W_w^{\perp} \\
& \Rightarrow \quad K = W_z U_L \Sigma_U U_R Q V_L \Sigma_V V_R W_w^{-L} + Q_W W_w^{\perp} \\
& \Rightarrow \quad \Sigma_U^{-1} U_L^{-L} W_z^{-L} K W_w V_R^{-R} \Sigma_V^{-1} = U_R Q V_L \\
& \Rightarrow \quad \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} W_z^{-L} K W_w V_R^{-R} \Sigma_V^{-1} & U_R Q V_L^c \\ U_R^c Q V_L & U_R^c Q V_L^c \end{bmatrix} = \begin{bmatrix} U_R \\ U_R^c \end{bmatrix} Q \begin{bmatrix} V_L & V_L^c \end{bmatrix} \\
& \Rightarrow \quad Q = \begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} W_z^{-L} K W_w V_R^{-R} \Sigma_V^{-1} & U_R Q V_L^c \\ U_R^c Q V_L & U_R^c Q V_L^c \end{bmatrix} \begin{bmatrix} V_L^{-R} \\ V_L^{\perp} \end{bmatrix}
\end{aligned}$$

Hence Q has the claimed form. □

4.4 Existence of a Minimizer

The following lemma shows that the feasible subspace $\boldsymbol{K}^{iw}$ of OPT^{iw} is closely related to the feasible subspace $\boldsymbol{K}(U,V)$ of $OPT(H,U,V)$. This relationship makes it easy to obtain results for OPT^{iw} similar to those found for OPT in Chapter 3.

Lemma 4.4.1 $\boldsymbol{K}^{iw} = \left\{K \in l_1^{n_{\hat{z}} \times n_{\hat{w}}} : KW_w \in \boldsymbol{K}(W_zU, V)\right\}$.

Proof: If $K \in \boldsymbol{K}^{iw}$ then $K \in l_1^{n_{\hat{z}} \times n_{\hat{w}}}$ and, performing the matrix multiplication in the definition (4.9), $K = W_zUQ_CVW_w^{-L} + Q_WW_w^{\perp}$. Hence, using the Bezout equation (4.8), $KW_w = W_zUQ_CV$. Since $Q_C \in l_1$, $KW_w \in \boldsymbol{K}(W_zU, V)$. Conversely, if $KW_w \in \boldsymbol{K}(W_zU,V)$ then $K \in l_1^{n_{\hat{z}} \times n_{\hat{w}}}$ and $KW_w = W_zUQ_CV$ for some $Q_C \in l_1$. Using the reverse of (4.8),

$$\begin{aligned} K &= K(W_wW_w^{-L} + W_w^cW_w^{\perp}) = W_zUQ_CVW_w^{-L} + KW_w^cW_w^{\perp} \\ &= \begin{bmatrix} W_zU & I \end{bmatrix} \begin{bmatrix} Q_C & 0 \\ 0 & Q_W \end{bmatrix} \begin{bmatrix} VW_w^{-L} \\ W_w^{\perp} \end{bmatrix} \end{aligned}$$

where $Q_W := KW_w^c \in l_1$ since both K and W_w^c are. Hence $K \in \boldsymbol{K}^{iw}$. □

Theorem 4.4.2 $K \in \boldsymbol{K}^{iw}$ *if and only if*

1. $K \in l_1^{n_{\hat{z}} \times n_{\hat{w}}}$,
2. $$\begin{bmatrix} U_L^{-L}W_z^{-L} \\ U_L^{\perp}W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_wV_R^{-R} & W_wV_R^{\perp} \end{bmatrix} = \begin{bmatrix} * & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$
   where $*$ denotes an irrelevant block, and
3. $[(\hat{U}_L^{-L}\hat{W}_z^{-L})_{i\cdot}\hat{K}(\hat{W}_w\hat{V}_R^{-R})_{\cdot j}]^{(n)}(z_0) = 0$ *for each* $(i,j) \in \{1,\ldots,r_U\} \times \{1,\ldots,r_V\}$, *each* $z_0 \in \boldsymbol{Z}_{ij}$, *and each* $n \in \{0,\ldots,m_{ij}(z_0)-1\}$.

Proof: U has a decomposition $U = U_L\Sigma_UU_R$ of the form specified in Assumption 3.0.1. Since W_z is left invertible, $W_zU = (W_zU_L)\Sigma_UU_R$ is a decomposition of W_zU of the same form. The Bezout equation for U_R in (3.3) are unchanged, and the Bezout equation for W_zU_L is

$$\begin{bmatrix} U_L^{-L}W_z^{-L} \\ U_L^{\perp}W_z^{-L} \\ W_z^{\perp} \end{bmatrix} \begin{bmatrix} W_zU_L & W_zU_L^c & W_z^c \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

The conclusion is now immediate using the fact (Lemma 4.4.1) that $K \in \boldsymbol{K}^{iw}$ if and only if $KW_w \in \boldsymbol{K}(W_zU,V)$ and applying Theorem 3.2.1 to $\boldsymbol{K}(W_zU,V)$. $\square$

The proof of the upcoming existence theorem 4.4.5 is similar to that of its counterpart Theorem 3.2.4 for *OPT*; two lemmas analogous to Lemmas 3.2.2 and 3.2.3 precede its statement and proof.

Lemma 4.4.3 *There exists* $\tilde{\mathcal{T}}_C^{iw} \in \mathcal{B}(c_0^{S_W}, c_0^{n_{\tilde{z}}\times n_{\tilde{w}}})$ *such that*

$$\boldsymbol{K}_C^{iw} := \left\{K \in l_1^{n_{\tilde{z}}\times n_{\tilde{w}}} : K \textit{ satisfies condition 2 of Theorem 4.4.2}\right\} = \mathcal{N}\left((\tilde{\mathcal{T}}_C^{iw})^*\right).$$

Proof: Begin by defining $\tilde{\mathcal{T}}_C^{iw}$ on $c_0^{S_W}$ by

$$\tilde{\mathcal{T}}_C^{iw} := \tilde{\mathcal{T}}_{W_w}\tilde{\mathcal{T}}_{W_zU,V}\tilde{\mathcal{E}}_{S_W} \tag{4.10}$$

where $\tilde{\mathcal{E}}_{S_W}$ is defined, given $G \in c_0^{S_W}$,

$$(\tilde{\mathcal{E}}_{S_W}G)_{ij} := \begin{cases} G_{ij} & (i,j) \in \boldsymbol{S}_W \\ 0 & \text{otherwise} \end{cases},$$

$\tilde{\mathcal{T}}_{W_zU,V}$ is defined, given $G \in c_0^{n_{\tilde{z}}\times n_w}$,

$$\tilde{\mathcal{T}}_{W_zU,V}G := \begin{bmatrix} U_L^{-L}W_z^{-L} \\ U_L^{\perp}W_z^{-L} \\ W_z^{\perp} \end{bmatrix}^T \triangleleft G \triangleright \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}^T,$$

and $\tilde{\mathcal{T}}_{W_w}$ is defined, given $G \in c_0^{n_{\tilde{z}}\times n_w}$, $\tilde{\mathcal{T}}_{W_w}G := G \triangleright W_w^T$.

Arguing as in the proof of Lemma 3.2.2, $\tilde{\mathcal{T}}_C^{iw} \in \mathcal{B}(c_0^{S_W}, c_0^{n_{\tilde{z}}\times n_{\tilde{w}}})$ and

$$(\tilde{\mathcal{T}}_C^{iw})^* = \tilde{\mathcal{E}}_{S_W}^*\tilde{\mathcal{T}}_{W_zU,V}^*\tilde{\mathcal{T}}_{W_w}^* = \Pi_{S_W}\mathcal{T}_{W_zU,V}\mathcal{T}_{W_w} =: \mathcal{T}_C^{iw} \tag{4.11}$$

where $\mathcal{T}_{W_w}$ is defined, given $K \in l_1^{n_{\tilde{z}}\times n_{\tilde{w}}}$, $\mathcal{T}_{W_w}K := KW_w$, $\mathcal{T}_{W_zU,V}$ is defined, given $K \in l_1^{n_{\tilde{z}}\times n_w}$,

$$\mathcal{T}_{W_zU,V}K := \begin{bmatrix} U_L^{-L}W_z^{-L} \\ U_L^{\perp}W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} V_R^{-R} & V_R^{\perp} \end{bmatrix}$$

and Π_{S_W} is defined, given $K \in l_1^{n_{\tilde{z}}\times n_w}$, $(\Pi_{S_W}K)_{ij} := K_{ij}, \quad (i,j) \in \boldsymbol{S}_W$.

By the construction of $\tilde{\mathcal{T}}_C^{iw}$, $\boldsymbol{K}_C^{iw} = \mathcal{N}\left(\mathcal{T}_C^{iw}\right) = \mathcal{N}\left((\tilde{\mathcal{T}}_C^{iw})^*\right)$. $\square$

Lemma 4.4.4 *There exists* $\tilde{\mathcal{T}}_I^{iw} \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_{\hat{z}} \times n_{\hat{w}}})$ *such that*

$$\boldsymbol{K}_I^{iw} := \left\{K \in l_1^{n_{\hat{z}} \times n_{\hat{w}}} : K \textit{ satisfies condition 3 of Theorem 4.4.2}\right\} = \mathcal{N}\left((\tilde{\mathcal{T}}_I^{iw})^*\right).$$

Proof: Begin by defining $\tilde{\mathcal{T}}_I^{iw} := \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{T}}_{W_z U, V} \tilde{\mathcal{T}}_{IZ}^{iw}$ where $\tilde{\mathcal{T}}_{IZ}^{iw}$ is defined, given $\alpha \in \mathbf{R}^{m_t}$,

$$\tilde{\mathcal{T}}_{IZ}^{iw}\alpha := \sum_{i=1}^{r_U}\sum_{j=1}^{r_V}\left[\sum_{z_0 \in Z_{ij}} \sum_{n=0}^{m_{ij}(z_0)-1} \tilde{\mathcal{E}}_{ij}\tilde{\mathcal{D}}_{n,z_0}^{\Re}\alpha_{i,j,n,z_0} + \sum_{z_0 \in Z_{ij}^{+}} \sum_{n=0}^{m_{ij}(z_0)-1} \tilde{\mathcal{E}}_{ij}\tilde{\mathcal{D}}_{n,z_0}^{\Im}\alpha_{i,j,n,z_0}\right] \tag{4.12}$$

where $\tilde{\mathcal{T}}_{W_w}$ is defined in the proof of Lemma 4.4.3, $\tilde{\mathcal{E}}_{ij} : c_0 \mapsto c_0^{n_{\hat{z}} \times n_w}$ is defined

$$(\tilde{\mathcal{E}}_{ij}G)_{mn} := \begin{cases} G_{ij} & m = i \text{ and } n = j \\ 0 & \text{otherwise} \end{cases}$$

and $\tilde{\mathcal{D}}_{n,z_0}^{\Re}$ and $\tilde{\mathcal{D}}_{n,z_0}^{\Im}$ are defined, given $\alpha \in \mathbf{R}$,

$$(\tilde{\mathcal{D}}_{n,z_0}^{\Re}\alpha)(k) := \begin{cases} 0 & k < n \\ \frac{\alpha k!}{(k-n)!}\Re(z_0^{k-n}) & k \geq n \end{cases}$$

and

$$(\tilde{\mathcal{D}}_{n,z_0}^{\Im}\alpha)(k) := \begin{cases} 0 & k < n \\ \frac{\alpha k!}{(k-n)!}\Im(z_0^{k-n}) & k \geq n \end{cases}$$

Arguing as in the proof of Lemma 3.2.3, $\tilde{\mathcal{T}}_I^{iw} \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_{\hat{z}} \times n_{\hat{w}}})$. $(\tilde{\mathcal{T}}_I^{iw})^* =: \mathcal{T}_I^{iw}$ can be computed straightforwardly (if tediously) to verify that, because of the construction of $\tilde{\mathcal{T}}_I^{iw}$, $\boldsymbol{K}_I^{iw} = \mathcal{N}\left(\mathcal{T}_I^{iw}\right) = \mathcal{N}\left((\tilde{\mathcal{T}}_I^{iw})^*\right)$. □

Theorem 4.4.5 OPT^{iw} *has a minimizer.*

Proof: The argument is the same as in Theorem 3.2.4; $\boldsymbol{K}^{iw} = \mathcal{N}\left((\tilde{\mathcal{T}}_K^{iw})^*\right)$ where $\tilde{\mathcal{T}}_K^{iw}$ is defined, given $(G, \alpha) \in c_0^{S_W} \times \mathbf{R}^{m_t}$,

$$\tilde{\mathcal{T}}_K^{iw}[(G,\alpha)] := \tilde{\mathcal{T}}_C^{iw}G + \tilde{\mathcal{T}}_I^{iw}\alpha \tag{4.13}$$

and $\tilde{\mathcal{T}}_C^{iw}$ and $\tilde{\mathcal{T}}_I^{iw}$ are defined in equations (4.10) and (4.12), respectively. Hence $\boldsymbol{K}^{iw}$ is *weak**-closed and existence is guaranteed by Corollary B.2.4 □

Of course a corollary analogous to Corollary 3.2.5 holds in this case as well.

Corollary 4.4.6 *If* $G_{yu}(0) = 0$ *then* $DR^{iw}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ *has a minimizer.*

4.5 Suboptimal Solutions via FIR Approximation

OPT^{iw} is in general an infinite dimensional problem as OPT is. In this section, a sequence of finite dimensional optimization problems is formulated which can be used under certain conditions to compute feasible solutions for OPT^{iw} which are arbitrarily close to optimal.

The idea is similar to that of Section 3.3; in that section, the sequence corresponded to allowing finitely supported closed loop impulse responses of a given length. In this section, OPT^{iw} is approximated by problems which correspond to allowing finitely supported $W_zHW_w^{-L} - K$ instead. Similar difficulties arise with respect to feasibility of such a sequence, but it will be shown that for finitely supported weights with finitely supported left inverses (i.e., with polynomial transfer functions having polynomial left inverses) the sequence for OPT^{iw} is feasible if and only if the sequence for OPT is.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\overline{OPT}_n^{iw}: \qquad \inf\left\{\left\|W_zHW_w^{-L} - K\right\|_{l_1} : K \in \overline{\boldsymbol{K}}^{iw}\right\} =: \overline{\mu}_n^{iw}$$

where

$$\overline{\boldsymbol{K}}_n^{iw} := \left\{K \in \boldsymbol{K}^{iw} : \text{supp } (W_zHW_w^{-L} - K) \subset \{0, \ldots, n\}\right\}$$

Note that $\overline{\boldsymbol{K}}_n^{iw}$, like $\overline{\boldsymbol{K}}_n(H, U, V)$, is a subset and not a subspace of $l_1^{n_{\bar{z}} \times n_{\bar{w}}}$. Both $\overline{OPT}_n^{iw}$ and $\overline{\boldsymbol{K}}_n^{iw}$ depend on H, U, V, W_z, and W_w; these dependences are suppressed for notational convenience.

The following proposition shows that the sequence of problems, when feasible, approximates OPT^{iw} from above.

Proposition 4.5.1 *If there exists $N \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}_N^{iw} \neq \emptyset$ and the finitely supported matrices are dense in $\boldsymbol{K}^{iw}$ then*

$$\{\overline{\mu}_n^{iw}\}_{n=N}^{\infty} \searrow \mu_{OPT} \text{ as } n \to \infty.$$

Proof: The proof is virtually identical to that of Proposition 3.3.1. $n_1 > n_2 \geq N$ implies $\overline{\boldsymbol{K}}_{n_1}^{iw} \supset \overline{\boldsymbol{K}}_{n_2}^{iw} \supset \overline{\boldsymbol{K}}_N^{iw}$. Hence if $\overline{\boldsymbol{K}}_N^{iw} \neq \emptyset$ then $\overline{\mu}_n^{iw}$ is well defined for every $n \geq N$ and $\{\mu_n^{iw}\}_{n=N}^{\infty}$ is monotonically non-decreasing. To show convergence, again use Lemma B.2.5 to obtain

$$\mu^{iw} = \inf\left\{\left\|W_zHW_w^{-L} - K\right\|_{l_1} : K \in \boldsymbol{K}^{iw}\right\} = \inf\left\{\|H_N - K\|_{l_1} : K \in \boldsymbol{K}^{iw}\right\}$$

where $K_N \in \boldsymbol{K}^{iw}$ and $H_N := W_z H W_w^{-L} - K_N$ is finitely supported. Now, given $\epsilon > 0$, choose $K^\epsilon \in \boldsymbol{K}^{iw}$ such that $\|H_N - K^\epsilon\|_{l_1} \leq \mu^{iw} + \frac{\epsilon}{2}$ and use the density hypothesis to find $K^\epsilon_{\text{fs}} \in \boldsymbol{K}^{iw}$ such that $\left\|K^\epsilon - K^\epsilon_{\hat{N}}\right\|_{l_1} \leq \frac{\epsilon}{2}$. Then $\overline{\mu}^{iw}_n \leq \epsilon$ for n such that supp $(H_N - K^\epsilon_{\text{fs}}) \subset \{0, \ldots, n\}$. □

Note that the hypotheses of the following proposition are equivalent to the rows of $\hat{W}_z$ and $\hat{W}_w$ being polynomial and having polynomial left inverses. Hence weights which bound magnitude and any number of nth order differences satisfy the hypotheses.

Proposition 4.5.2 *If W_z and W_w are finitely supported and have finitely supported left inverses then there exists $n \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}^{iw} \neq \emptyset$ if and only if there exists $n \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}_n(H, U, V) \neq \emptyset$.*

Proof: If W_z and W_w are finitely supported with finitely supported left inverses then there exist l_1 matrices Q_z and Q_w such that $W_z^{-L} + Q_z W_z^\perp$ and $W_w^{-L} + Q_w W_w^\perp$ are finitely supported (since all left inverses in l_1 are parameterized in this form). Suppose first that there exists $K \in \boldsymbol{K}(U, V)$ such that $H - K$ is finitely supported. Then it is easy to check that $\tilde{K} := W_z K W_w^{-L} + W_z(H - K) Q_w W_w^\perp \in \boldsymbol{K}^{iw}$. Moreover $W_z H W^{-L} - \tilde{K} = W_z(H - K)(W_w^{-L} + Q_w W_w^\perp)$ and is hence finitely supported, since it is a product of finitely supported matrices.

Conversely, suppose there exists $K \in \boldsymbol{K}^{iw}$ such that $W_z H W_w^{-L} - K$ is finitely supported. It is easy to check that $(W_z^{-L} + Q_z W_z^\perp)(W_z H W_w^{-L} - K) W_w = H - \tilde{K}$ where $\tilde{K} \in \boldsymbol{K}(U, V)$. Moreover, $H - \tilde{K}$ is finitely supported since it is a product of finitely supported matrices. □

4.6 A Converging Lower Bound on μ_{IWDR}

In this section, which parallels Section 3.4, a sequence of super-optimal approximating problems is formulated whose infimal costs converge to μ^{iw} from below. The same key features as found for the corresponding problems $\underline{OPT}_n(H, U, V)$ carry over here; they are always feasible, they are infinite dimensional, they have finitely supported minimizers, and stabilizing Cs can be obtained from their feasible solutions. The proofs of the results here are generally similar to those of the corresponding results in Section 3.4; they are presented somewhat more briefly and sometimes only sketched.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\underline{OPT}_n^{iw} \qquad \inf\left\{\left\|W_z H W_w^{-L} - K\right\|_{l_1} : K \in \underline{\boldsymbol{K}}_n^{iw}\right\} =: \underline{\mu}_n^{iw}$$

where

$$K \in \underline{\boldsymbol{K}}_n^{iw} \iff K \in l_1^{n_{\bar{z}} \times n_{\bar{w}}}, \tag{4.14}$$

K satisfies Condition 3 of Theorem 4.4.2, and

$$\left(\begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix}\right)(k) = \begin{bmatrix} * & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

for $k \in \{0, \ldots, n\}$

Note that $\underline{\boldsymbol{K}}_n^{iw}$ is a subspace of $\boldsymbol{K}^{iw}$.

Theorem 4.6.1 $\{\underline{\mu}_n^{iw}\}_{n=0}^{\infty} \nearrow \mu^{iw}$ *as* $n \to \infty$*. Moreover, for each* $n \in \mathbf{Z}_+$*,* $\underline{OPT}_n^{iw}$ *has a minimizer.*

Proof: The proof is very similar to that of Theorem 3.4.1. $\{\underline{\mu}_n^{iw}\}_{n=0}^{\infty}$ is well defined because, for any n, $0 \in \underline{\boldsymbol{K}}_n^{iw}$ and, for any $K \in \underline{\boldsymbol{K}}_n^{iw}$, $\left\|W_z H W_w^{-L} - K\right\|_{l_1} \geq 0$. Also, $\underline{\mu}_n^{iw} \leq \mu^{iw}$ for each n and the sequence is non-decreasing since $n_1 > n_2$ implies $\underline{\boldsymbol{K}}_{n_2} \supset \underline{\boldsymbol{K}}_{n_1} \supset \boldsymbol{K}^{iw}$.

To show convergence, begin by defining for each n a map $\underline{\tilde{\mathcal{I}}}_n^{iw}$, given $(G, \alpha) \in c_0^{S_W} \times \mathbf{R}^{m_t}$,

$$\underline{\tilde{\mathcal{I}}}_n^{iw}[(G,\alpha)] := \tilde{\mathcal{T}}_C^{iw} \tilde{\mathcal{P}}_n G + \tilde{\mathcal{T}}_I^{iw} \alpha \tag{4.15}$$

where $\tilde{\mathcal{T}}_C^{iw} \in \mathcal{B}(c_0^{S_W}, c_0^{n_{\bar{z}} \times n_{\bar{w}}})$ is defined in (4.10) and $\tilde{\mathcal{T}}_I^{iw} \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_{\bar{z}} \times n_{\bar{w}}})$ in (4.12), in the proofs of Lemmas 4.4.3 and 4.4.4, respectively. $\tilde{\mathcal{P}}_n$ is the n-th truncation operator restricted to $c_0^{S_W}$ and is bounded by Proposition C.1.7. Arguing as in the proof of Theorem 3.4.1, $(\underline{\tilde{\mathcal{I}}}_n^{iw})^* \in \mathcal{B}(c_0^{n_{\bar{z}} \times n_{\bar{w}}}, c_0^{S_W} \times \mathbf{R}^{m_t})$ can be written, given $K \in l_1^{n_{\bar{z}} \times n_{\bar{w}}}$,

$$(\underline{\tilde{\mathcal{I}}}_n^{iw})^* K = (\tilde{\mathcal{P}}_n^* (\tilde{\mathcal{T}}_C^{iw})^* K, (\tilde{\mathcal{T}}_I^{iw})^* K).$$

Continuing with the argument, $\mathcal{R}\left(\underline{\tilde{\mathcal{I}}}_n^{iw}\right)^{\perp} = \mathcal{N}\left((\underline{\tilde{\mathcal{I}}}_n^{iw})^*\right) = \underline{\boldsymbol{K}}_n^{iw}$ so the duality theorem B.2.3 gives

$$\begin{aligned}
\underline{\mu}_n^{iw} &= \sup\left\{\left\langle \tilde{G}, W_z H W_w^{-L}\right\rangle : \boldsymbol{B}\mathcal{R}\left(\underline{\tilde{\mathcal{I}}}_n^{iw}\right)\right\} \\
&= \sup\left\{\left\langle \tilde{\mathcal{T}}_K^{iw} \tilde{G}, W_z H W_w^{-L}\right\rangle : \tilde{G} \in \mathcal{R}\left(\tilde{\mathcal{P}}_n\right) \times \mathbf{R}^{m_t},\ \left\|\tilde{\mathcal{T}}_K^{iw} \tilde{G}\right\|_{l_\infty} \leq 1\right\}
\end{aligned}$$

where $\tilde{T}_K^{iw}$ is defined in equation (4.13). Hence

$$\begin{aligned}
\sup_n \underline{\mu}_n^{iw} &= \sup\left\{\left\langle \tilde{T}_K^{iw}\tilde{G}, W_z H W_w^{-L}\right\rangle : \tilde{G} \in c_0^{Sw} \times \mathbf{R}^{m_t},\ \left\|\tilde{T}_K^{iw}\tilde{G}\right\|_{l_\infty} \leq 1\right\} \\
&= \sup\left\{\left\langle \tilde{G}, W_z H W_w^{-L}\right\rangle : \boldsymbol{B}\mathcal{R}\left(\tilde{T}_K^{iw}\right)\right\} \\
&= \inf\left\{\left\|W_z H W_w^{-L} - K\right\|_{l_1} : K \in \mathcal{N}\left((\tilde{T}_K^{iw})^*\right)\right\} \\
&= \mu^{iw}
\end{aligned}$$

For the second statement, $\underline{\boldsymbol{K}}_n^{iw} = \mathcal{N}\left((\tilde{\underline{\mathcal{T}}}_n^{iw})^*\right)$; hence it is *weak**-closed and $\underline{OPT}_n^{iw}$ has a minimizer. □

Theorem 4.6.2 *For every n, every minimizer K_0 for $\underline{OPT}_n^{iw}$ is such that at least one row of $W_z H W_w^{-L} - K_0$ is finitely supported. Moreover, there exists a minimizer K_{fs} such that all entries of $W_z H W_w^{-L} - K_{\text{fs}}$ are finitely supported.*

Proof: For the first statement, the arguments which prove the first statement of Theorem 3.4.2 go through with $\tilde{\underline{\mathcal{T}}}_n$, $\tilde{T}_C$, and $\tilde{T}_I$ replaced by $\tilde{\underline{\mathcal{T}}}_n^{iw}$, $\tilde{T}_C^{iw}$, and $\tilde{T}_I^{iw}$, respectively. The somewhat long and tedious proof of the second statement of that theorem can be modified to serve here. □

Next a doubly indexed family of finite dimensional problems analogous to the family of problems $\underline{OPT}_{n,m}(H, U, V)$ of Section 3.4 is defined as follows.

Given a problem $\underline{OPT}_n^{iw}$ for some n define for each $m \in \mathbf{Z}_+$ an optimization problem

$$\underline{OPT}_{n,m}^{iw} \qquad \inf\left\{\|H - K\|_{l_1} : K \in \underline{\boldsymbol{K}}_n^{iw}\right\} =: \underline{\mu}_{n,m}^{iw}$$

where

$$\underline{\boldsymbol{K}}_{n,m}^{iw} := \left\{K \in \underline{\boldsymbol{K}}_n^{iw} : \text{supp}\ (W_z H W_w^{-L} - K) \subset \{0, \ldots, m\}\right\} \tag{4.16}$$

and $\underline{\boldsymbol{K}}_n^{iw}$ is defined in (4.14).

Hence for each problem $\underline{OPT}_n^{iw}$ there is a sequence of problems $\left\{\underline{OPT}_{n,m}^{iw}\right\}_{m=0}^{\infty}$. The following proposition corresponds to Proposition 3.4.3 and shows that a finitely supported minimizer for $\underline{OPT}_n^{iw}$ can be found by solving $\underline{OPT}_{n,m}^{iw}$ for some finite m.

Proposition 4.6.3 *If $m \geq n$ and $m \geq m_t$ then $\underline{OPT}_{n,m}^{iw}$ is feasible and $\underline{\mu}_{n,m}^{iw} \geq \underline{\mu}_n^{iw}$. Moreover*

1. *For each n there exists $M_n \geq n$ such that $\underline{\mu}_{n,m}^{iw} = \underline{\mu}_n^{iw}$.*

2. If $\boldsymbol{Z} = \emptyset$ then, for each n, $\underline{OPT}^{iw}_{n,n}$ is feasible and $\underline{\mu}^{iw}_{n,n} = \underline{\mu}^{iw}_{n}$.

Proof: As in the proof of Proposition 4.6.3, if $m \geq n$ and $m \geq m_t$ then there exists $K := K_H - K_m$ where $K_H := (\mathcal{I} - \mathcal{P}_m)W_z H W_w^{-L}$, K_m satisfies

$$\text{supp } K_m \subset \{0, \ldots, m\} \text{ and } \mathcal{T}^{iw}_I K_m = \mathcal{T}^{iw}_I K_H,$$

and $\mathcal{T}^{iw}_I$ is defined in Lemma 4.4.4. Hence $\underline{\boldsymbol{K}}^{iw}_{n,m} \neq \emptyset$ so $\underline{\mu}^{iw}_{n,m}$ is well defined and $\underline{\mu}^{iw}_{n,m} \geq \underline{\mu}^{iw}_{n}$ since $\underline{\boldsymbol{K}}^{iw}_{n,m} \subset \underline{\boldsymbol{K}}^{iw}_{n}$.

Item 1: Theorem 4.6.2 shows that there is a finitely supported minimizer K_{fs} for $\underline{OPT}^{iw}_{n}$; just take $M_n := \max\{n, \text{supp } K_{\text{fs}}\}$.

Item 2: $(\mathcal{I} - \mathcal{P}_n)W_z H W_w^{-L} \in \underline{\boldsymbol{K}}^{iw}_{n}$ and $\mathcal{P}_n K \in \underline{\boldsymbol{K}}^{iw}_{n}$ for all $K \in \underline{\boldsymbol{K}}^{iw}_{n}$. Thus $K \in \underline{\boldsymbol{K}}^{iw}_{n}$ implies $\tilde{K} := \mathcal{P}_n K + (\mathcal{I} - \mathcal{P}_n)W_z H W_w^{-L} \in \underline{\boldsymbol{K}}^{iw}_{n}$, and it is easy to check that $\text{supp } (W_z H W_w^{-L} - \tilde{K}) \subset \{0, \ldots, n\}$ so that $\tilde{K} \in \underline{\boldsymbol{K}}^{iw}_{n,n}$. Also, $(W_z H W_w^{-L} - \tilde{K})(k) = (W_z H W_w^{-L} - K)(k)$ for $k \leq n$ so that $\left\|W_z H W_w^{-L} - \tilde{K}\right\|_{l_1} \leq \left\|W_z H W_w^{-L} - K\right\|_{l_1}$. Conclude that $\underline{\mu}^{iw}_{n,n} \leq \underline{\mu}^{iw}_{n}$, and the reverse inequality follows from the first sentence of the proposition. □ Stabilizing compensators can be computed from feasible solutions to $\underline{OPT}^{iw}_{n,m}$ in a fashion very similar to the unweighted case.

Proposition 4.6.4 *If $\underline{K} \in \underline{\boldsymbol{K}}^{iw}_{n,m}$ for any n, m then*

$$K := W_z U_L U_L^{-L} W_z^{-L} \underline{K} W_w V_R^{-R} V_R W_w^{-L} \in \boldsymbol{K}^{iw}.$$

Proof: If $\underline{K} \in \underline{\boldsymbol{K}}^{iw}_{n,m}$ then $\underline{K} \in \underline{\boldsymbol{K}}^{iw}_{n}$ by definition. Next check using the definition of $\underline{\boldsymbol{K}}^{iw}_{n}$ in (4.14) that K satisfies the three conditions of Theorem 4.4.2 and hence $K \in \boldsymbol{K}(U, V)$. First, $\underline{K} \in \underline{\boldsymbol{K}}^{iw}_{n} \Rightarrow K \in l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$ so that condition 1 is satisfied. Also,

$$\begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix} = \begin{bmatrix} U_L^{-L} W_z^{-L} \underline{K} W_w V_R^{-R} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

so that conditions 2 and 3 are satisfied. □

4.7 Linear Programming Formulations

In this section linear programs are formulated corresponding to all the optimization problems defined thus far. The linear programs are all equivalent to the original problems in the sense described in Section 3.5. The notation of that section is also in effect here: l_1^+ for the positive cone in l_1 and a projection operator $\Pi_+ : l_1^{n_{\tilde{z}} \times n_{\tilde{w}}} \mapsto (l_1^+)^{n_{\tilde{z}} \times n_{\tilde{w}}}$ which sets the negative elements of each sequence in a matrix to zero. For compactness of notation,

$$\tilde{U} := \begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix}, \quad \tilde{V} := \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix},$$

and

$$\tilde{H} := \tilde{U} H \tilde{V}$$

throughout the section. Also recall from Remark 3.5.2 that if $X \in l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$ then

- $X = \Pi_+ X - \Pi_+(-X)$.
- $|X_{ij}(k)| = [\Pi_+ X + \Pi_+(-X)]_{ij}(k)$ for each i, j, and k and hence
$$\|X\|_{l_1} = \|\Pi_+ X\|_{l_1} + \|\Pi_+(-X)\|_{l_1}.$$

The first linear program corresponds to OPT^{iw} defined in Section 4.3. The linear program is

LP^{iw} : $\qquad \inf \mu$

subject to:

$$\begin{array}{rcll} \left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=0}^{\infty} (T^+ + T^-)_{ij}(k) \right] - \mu & \leq & 0 & i = 1, \ldots, n_{\tilde{z}} \\ \left[\tilde{U}(T^+ - T^-)\tilde{V} \right]_{ij}(k) & = & \tilde{H}_{ij}(k) & (i,j) \in \boldsymbol{S}_W, \quad k \in \mathbf{Z}_+ \\ \mathcal{T}_I^{iw}(T^+ - T^-) & = & \mathcal{T}_I^{iw} H & \\ (\mu, T^+, T^-) & \in & \mathbf{R}_+ \times (l_1^+)^{n_{\tilde{z}} \times n_{\tilde{w}}} \times (l_1^+)^{n_{\tilde{z}} \times n_{\tilde{w}}} & \end{array}$$

where $\mathcal{T}_I^{iw}$ is defined in Lemma 4.4.4.

The variables in LP^{iw} are $\mu \in \mathbf{R}$ and

$$\left\{ T_{ij}^+(k),\ T_{ij}^-(k) : i \in \{1, \ldots, n_{\tilde{z}}\},\ j \in \{1, \ldots, n_{\tilde{w}}\},\ k \in \mathbf{Z}_+ \right\}.$$

Note that they are infinite in number. There are $n_{\tilde{z}}$ (inequality) cost constraints, an infinite number of (equality) convolution constraints, a finite number of (equality) interpolation constraints, and an infinite number of (inequality) positivity constraints.

Proposition 4.7.1 *LP^{iw} is equivalent to OPT^{iw} under the map ψ^{iw} defined, given a feasible solution (μ, T^+, T^-) for LP^{iw},*

$$\psi^{iw}(\mu, T^+, T^-) := W_z H W_w^{-L} - (T^+ - T^-).$$

Proof: Use $\tilde{\psi}^{iw}$ defined, given $K \in \boldsymbol{K}^{iw}$,

$$\tilde{\psi}^{iw} K := \left(\left\| W_z H W_w^{-L} - K \right\|_{l_1}, \ \Pi_+(W_z H W_w^{-L} - K), \ \Pi_+(K - W_z H W_w^{-L}) \right).$$

□

The next linear program corresponds to an FIR approximation problem $\overline{OPT}_n^{iw}$, defined in Section 4.5. The linear program is

$\overline{LP}_n^{iw}$: $\inf \mu$

subject to:

$$\begin{aligned}
\left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=0}^{n} (T^+ + T^-)_{ij}(k) \right] - \mu &\leq 0 & i &= 1, \ldots, n_{\tilde{z}} \\
\left[\mathcal{E}_n (T^+ - T^-) \right]_{ij}(k) &= \tilde{H}_{ij}(k) & (i,j) &\in \boldsymbol{S}_W, \quad k \in \{0, \ldots, \bar{n}\} \\
\mathcal{T}_I^{iw} \mathcal{E}_n (T^+ - T^-) &= \mathcal{T}_I^{iw} H & & \\
(\mu, T^+, T^-) &\in \mathbf{R}_+ \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (n+1)} \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (n+1)} & &
\end{aligned}$$

where the matrices in the convolution constraints have been chosen to be finitely supported. $\bar{n} := n + n_U + n_V$ where n_U and n_V are integers such that

$$\operatorname{supp} \tilde{U} \subset \{0, \ldots, n_U\} \text{ and } \operatorname{supp} \tilde{V} \subset \{0, \ldots, n_V\}.$$

$\mathcal{E}_n : \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (n+1)} \mapsto l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$ is the natural embedding operator which pads with zeros.

There are $2 n_{\tilde{z}} n_{\tilde{w}} (n+1) + 1$ variables in $\overline{LP}_n^{iw}$: μ and

$$\left\{ T_{ij}^+(k), \ T_{ij}^-(k) : i \in \{1, \ldots, n_{\tilde{z}}\}, \ j \in \{1, \ldots, n_{\tilde{w}}\}, \ k \in \{0, \ldots, n\} \right\}.$$

Proposition 4.7.2 *If there exists $\bar{n}$ such that $\overline{\boldsymbol{K}}_{\bar{n}}^{iw} \neq \emptyset$ then there exists n_H such that*

$$\operatorname{supp} \tilde{H}_{ij} \subset \{0, \ldots, n_H\} \quad \forall (i,j) \in \boldsymbol{S}_W$$

and if $n \geq n_H - n_U - n_V$ then $\overline{LP}_n^{iw}$ is equivalent to $\overline{OPT}_n^{iw}$ under the map $\overline{\psi}_n^{iw}$ defined, given a feasible solution (μ, T^+, T^-) for $\overline{LP}_n^{iw}$,

$$\overline{\psi}_n^{iw}(\mu, T^+, T^-) := W_z H W_w^{-L} - \mathcal{E}_n (T^+ - T^-).$$

Proof: The first statement follows from Proposition 4.5.2 and the equivalence is shown using $\overline{\tilde{\psi}}_n^{iw}$ defined, given $K \in \overline{\boldsymbol{K}}_n^{iw}$,

$$\overline{\tilde{\psi}}_n^{iw} K := \left(\left\| W_z H W_w^{-L} - K \right\|_{l_1}, \ \Pi_n \Pi_+ (W_z H W_w^{-L} - K), \ \Pi_n \Pi_+ (K - W_z H W_w^{-L}) \right).$$

where $\Pi_n : l_1^{n_{\bar{z}} \times n_{\bar{w}}} \mapsto \mathbf{R}_+^{n_{\bar{z}} \times n_{\bar{w}} \times (n+1)}$ is the obvious projection operator. □

The next linear program corresponds to an infinite dimensional super-optimal approximation problem $\underline{OPT}_n^{iw}$, defined in Section 4.6. The linear program

$$\underline{LP}_n^{iw}: \qquad \inf \mu$$

subject to:

$$\begin{aligned}
\left[\sum_{j=1}^{n_{\bar{w}}} \sum_{k=0}^{\infty} (T^+ + T^-)_{ij}(k) \right] - \mu \ &\leq \ 0 && i = 1, \ldots, n_{\bar{z}} \\
\left[\tilde{U}(T^+ - T^-) \tilde{V} \right]_{ij}(k) \ &= \ \tilde{H}_{ij}(k) && (i,j) \in \boldsymbol{S}_W, \quad k \in \{0, \ldots, n\} \\
\mathcal{T}_I^{iw}(T^+ - T^-) \ &= \ \mathcal{T}_I^{iw} H && \\
(\mu, T^+, T^-) \ &\in \ \mathbf{R}_+ \times (l_1^+)^{n_{\bar{z}} \times n_{\bar{w}}} \times (l_1^+)^{n_{\bar{z}} \times n_{\bar{w}}} &&
\end{aligned}$$

Note that $\underline{LP}_n^{iw}$ is very similar to LP^{iw}. The variables are $\mu \in \mathbf{R}$ and

$$\left\{ T_{ij}^+(k), \ T_{ij}^-(k) : i \in \{1, \ldots, n_{\bar{z}}\}, \ j \in \{1, \ldots, n_{\bar{w}}\}, \ k \in \mathbf{Z}_+ \right\}.$$

as in LP^{iw} and are infinite in number. The constraints are identical as well, except that only a finite number of the convolution constraints are enforced. Hence, although there are still infinitely many positivity constraints, there are only a finite number of equality constraints.

Proposition 4.7.3 *$\underline{LP}_n^{iw}$ is equivalent to $\underline{OPT}_n^{iw}$ under the map $\underline{\psi}_n^{iw}$ defined, given a feasible solution (μ, T^+, T^-) for $\underline{LP}_n^{iw}$,*

$$\underline{\psi}_n^{iw}(\mu, T^+, T^-) := W_z H W_w^{-L} - (T^+ - T^-).$$

Proof: $\underline{\psi}_n^{iw}$ is defined identically to ψ^{iw} of Proposition 4.7.1, so use $\underline{\tilde{\psi}}_n^{iw}$ defined, given $K \in \underline{\boldsymbol{K}}_n^{iw}$,

$$\underline{\tilde{\psi}}_n^{iw} K := \left(\left\| W_z H W_w^{-L} - K \right\|_{l_1}, \ \Pi_+ (W_z H W_w^{-L} - K), \ \Pi_+ (K - W_z H W_w^{-L}) \right)$$

which is defined identically to $\tilde{\psi}^{iw}$. □

The last linear program corresponds to a finite dimensional super-optimal approximation problem $\underline{OPT}^{iw}_{n,m}$, defined in Section 4.6.

$\underline{LP}^{iw}_{n,m}$: $\inf \mu$

subject to:

$$\begin{array}{rcll} \left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=0}^{m} (T^+ + T^-)_{ij}(k)\right] - \mu & \leq & 0 & i = 1, \ldots, n_{\tilde{z}} \\ \left[\tilde{U}\mathcal{E}_m(T^+ - T^-)\tilde{V}\right]_{ij}(k) & = & \tilde{H}_{ij}(k) & (i,j) \in \boldsymbol{S}_W, \quad k \in \{0, \ldots, n\} \\ \mathcal{T}_I^{iw}\mathcal{E}_m(T^+ - T^-) & = & \mathcal{T}_I^{iw}H & \\ (\mu, T^+, T^-) & \in & \mathbf{R}_+ \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (m+1)} \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (m+1)} & \end{array}$$

Where $\mathcal{E}_m : \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (m+1)} \mapsto l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$ is the same embedding operator as in the formulation of $\overline{LP}^{iw}_n$ above.

There are $2n_{\tilde{z}}n_{\tilde{w}}(m+1)+1$ variables in $\underline{LP}^{iw}_{n,m}$: μ and

$$\left\{(T^+)_{ij}(k),\ (T^-)_{ij}(k) : i \in \{1, \ldots, n_{\tilde{z}}\},\ j \in \{1, \ldots, n_{\tilde{w}}\},\ k \in \{0, \ldots, m\}\right\}.$$

Proposition 4.7.4 *For every n and m, $\underline{LP}^{iw}_{n,m}$ is equivalent to $\underline{OPT}^{iw}_{n,m}$ under the map $\underline{\psi}^{iw}_{n,m}$ defined, given a feasible solution (μ, T^+, T^-) for $\underline{LP}^{iw}_{n}$,*

$$\underline{\psi}^{iw}_{n,m}(\mu, T^+, T^-) := W_z H W_w^{-L} - \mathcal{E}_m(T^+ - T^-).$$

Proof: Can be proved identically to Proposition 4.7.3 if $\underline{\tilde{\psi}}^{iw}_{n,m}$ is defined, given $K \in \underline{\boldsymbol{K}}^{iw}_{n,m}$,

$$\underline{\tilde{\psi}}_{n,m}K := \left(\left\|W_z H W_w^{-L} - K\right\|_{l_1},\ \Pi_m\Pi_+(W_z H W_w^{-L} - K),\ \Pi_n\Pi_+(K - W_z H W_w^{-L})\right).$$

□

4.8 Example

In this section the incremental gain of a simple FIR high pass filter is computed w.r.t. several input and output weights. The system is given in terms of its z-transform by

$$\hat{G} = .1 \sum_{i=0}^{9} (-1)^i z^i.$$

	$a_i = 1$	$a_i = 5$	$a_i = 10$
$a_o = 1$	1	.2	.1
$a_o = 5$	5	1	.5
$a_o = 10$	10	2	1

Table 4.1: Incremental weighted gain of high pass w.r.t. various weights

Note that $G \in l_1$ so that it is incrementally stable w.r.t. any weights satisfying the assumptions of Section 4.1.

The input and output weights to be considered are given in terms of their z-transforms by

$$\hat{W}_o = \begin{bmatrix} 1 \\ a_o(1-z) \end{bmatrix}, \quad \hat{W}_i = \begin{bmatrix} 1 \\ a_i(1-z) \end{bmatrix}$$

and the associated Bezout equations (4.2) in terms of z transforms are

$$\begin{bmatrix} 1 & 0 \\ a_o(1-z) & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_o(1-z) & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 \\ a_i(1-z) & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a_i(1-z) & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The parameters a_o and a_i can be increased to reflect decreasing bounds on the input and output rates, or decreased to reflect increasing bounds on the rates.

$\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ was computed for several values of a_o and a_i as an l_1 problem per Theorem 4.1.6. The FIR approximation scheme of Section 3.3 was used to solve the problem, and the optimality of the results were confirmed by solving super-optimal approximation problems as in Section 3.4. The results are given in Table 4.1.

By way of constrast, $\|\mathcal{G}_{lp}\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ was computed for the same weights and values of

As might be expected, increasing a_o for fixed a_i causes the gain to increase. This is because the specification corresponding to the choice of $\mathcal{W}_o$ is becoming increasingly stringent; a smaller and smaller output rate is being required. On the other hand, increasing a_i for fixed a_o causes the gain to decrease. This is because the corresponding specification is becoming less stringent; a smaller and smaller possible input rate is being assumed.

	$a_i = 1$	$a_i = 5$	$a_i = 10$
$a_o = 1$	1	1	1
$a_o = 5$	1	1	1
$a_o = 10$	2	2	2

Table 4.2: Incremental weighted gain of low pass w.r.t. various weights

By way of constrast, $\|\mathcal{G}_{\mathrm{lp}}\|^{i}_{\mathcal{W}_o,\mathcal{W}_i}$ was computed for the same weights and values of a_o and a_i, where $\mathcal{G}_{\mathrm{lp}}$ is an FIR low pass filter given in terms of its z-transform by

$$\hat{G}_{\mathrm{lp}} = .1 \sum_{i=0}^{9} z^i$$

and the results shown in Table 4.2. In this case increasing a_o for fixed a_i causes the gain to increase, due to the more stringent requirement on the output rate. However for fixed a_o, the gain is constant. This reflects the fact that although a smaller bound is guaranteed on the input rate, this is of little use since relatively slowly varying inputs are the ones which cause large outputs from a low pass system.

4.9 Discussion

In this chapter, the incremental weighted l_∞ disturbance rejection problem has been formulated and a solution given. The first step was to define a notion of system gain appropriate to the weighted l_∞ specification, and show how it can be computed by solving a standard l_1 optimization problem. Thereafter, results analogous to those of Chapter 3 for the unweighted problem were obtained.

The approach, following that of Chapter 3, is to formulate the problem first in the standard problem setting as a norm (i.e., system gain) minimization and then as a minimum distance problem in l_1. This problem is distinct from the standard l_1 problem in that the minimization is done with respect to *two* free parameter matrices in l_1; one representing the choice of stabilizing compensator and one the computation of the system gain when that compensator is used. Although it is distinct from the standard l_1 problem, there is enough similarity to allow analogous methods for FIR sub-optimal and for super-optimal approximation to be established. For actual computation, the problems are finally recast as linear programs.

The results are not as complete as those of Chapter 3; it has not been established, for example, when optimal compensators can be computed exactly via finite dimensional optimizations (as is the case for the l_1 problem when U and V have appropriate rank). However, the definition and computation of the incremental gain w.r.t. given weights together with the approximate solution methods provides a basis for trial design and experimentation.

The most important aspect of the chapter is the following: the results show that a wide range of time domain specifications can be reflected by appropriate choices of weights, and that the corresponding design problems are as tractable theoretically and computationally as the standard l_1 problem. This helps to close a serious gap between the l_1 and $\boldsymbol{H}_\infty$ theories from this standpoint. Designers have a great deal of intuition about the frequency domain and the practical meaning of frequency response based specifications, and such specifications are readily addressed by the $\boldsymbol{H}_\infty$ theory through the use of cascade weights as described in Chapter 1. It seems possible that specifications of the type addressed here are a natural basis for the development of time domain based design intuition which is compatible with a tractable theory.

As mentioned in Chapter 1, problems similar to this have been addressed in [13][14]. In those papers similar weighted specifications were formulated, but the exact meaning of the problem solved remains unclear. In particular, the substantial example which was the main subject of [14] was solved under the assumption that disturbances were periodic because the theory could guarantee performance only over a finite time span. The notion of incremental weighted gain, its relation to the standard l_1 problem, its computation, and the synthesis of compensators which are optimal in the sense of this gain all appear here for the first time.

Chapter 5

Weighted l_∞ Disturbance Rejection

The problem setting is again the general feedback system of Chapters 2, 3, and 4 and the problem to be considered is the weighted l_∞ design problem described in Chapter 1. The precise statement of the specification is

Weighted l_∞ DRS:

- $\mathcal{C} \in \Sigma_{clti}$, $(\mathcal{G}, \mathcal{C})$ is stable, and
- $\mathcal{W}_w w \in l_{\infty+}^{n_{\tilde{w}}}$ and $\|\mathcal{W}_w w\|_{l_\infty} \leq 1$ implies $\mathcal{W}_z z \in l_{\infty+}^{n_{\tilde{z}}}$ and $\|\mathcal{W}_z z\|_{l_\infty} \leq 1$.

where the weights $\mathcal{W}_z$ and $\mathcal{W}_w$ are systems in $\Sigma_{clti}^{n_z \mapsto n_{\tilde{z}}}$ and $\Sigma_{clti}^{n_z \mapsto n_{\tilde{z}}}$, respectively.

As in the previous two problems, a $\mathcal{C}$ which satisfies the specification must be causal, linear, and time invariant, and it must stabilize $\mathcal{G}$. The second item of the specification differs, however, from the incremental version of the problem in that all whose weighted peak magnitude is less than or equal to one *for all time* must result in a weighted error with peak magnitude less than or equal to one *for all time.*

Similar results are obtained as for the incremental weighted problem, and the chapter is organized identically to Chapter 4. In Section 5.1 a notion of stability and a norm on the causal linear time invariant systems appropriate to the weighted specification are defined. It is shown here that, unlike the norm for the incremental problem, this one is in fact an induced norm between weighted versions of l_∞. The design problem is formulated in terms of this norm in Section 5.2 and assumptions are stated there. Essentially they are the same as those in effect for Chapter 4; the problem is solved for $\mathcal{G}$, $\mathcal{W}_z$, and $\mathcal{W}_w$ in $\boldsymbol{\Sigma}_{clti}$ with $\mathcal{W}_z$ and $\mathcal{W}_w$ stable and having stable left inverses in $\boldsymbol{\Sigma}_{clti}$.

In Section 5.3 the norm minimization problem is reformulated as a minimum distance problem, not in $l_1^{n_{\bar{z}} \times n_{\bar{w}}}$, but in $l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$. (Even though the system under consideration and the weights are causal and hence have impulse responses in l_1, the feasible subspace must be allowed to include two-sided sequences; Section C.2 contains the pertinent results analogous to those for the one-sided case.) These technical differences reduce the similarity to the unweighted problem, hence this problem is treated last, and fewer results are obtained. Existence of a minimizer is established in Section 5.4, sub-optimal and super-optimal approximations are treated in Sections 5.5 and 5.6, respectively, and linear program formulations are found in Section 5.7. In the case of super-optimal approximation, it is not readily apparent that a double iteration method involving finite dimensional optimizations can work, although this is a reasonable conjecture. The chapter concludes in Section 5.8 with a discussion of the results, related work, and the main contributions of the chapter.

5.1 Weighted Stability and Gain

Stability and gain are defined in this section analogously to the way they were in Section 4.1. For systems in $\boldsymbol{\Sigma}_{clti}$ and for stable weights in $\boldsymbol{\Sigma}_{clti}$ with stable left inverses in $\boldsymbol{\Sigma}_{clti}$, both notions of stability w.r.t. given weights are equivalent (i.e., both are equivalent to just stability). However in this case the norm computation is not equivalent to a standard l_1 problem, but to one in which the free parameter is allowed to range over $l_1(\mathbf{Z})$. As a consequence the norm defined here is in general smaller than that defined in Section 4.1.

Note: In this section the operations $\triangleleft$ and $\triangleright$ denote left and right correlation for sequences and matrices on $\mathbf{Z}$ (as opposed to $\mathbf{Z}_+$). See Section C.2 for the definition and for properties of correlation operators and their adjoints. Impulse responses of systems in $\boldsymbol{\Sigma}_{clti}$ are to be considered as elements of l_+ supported on $\mathbf{Z}_+$ instead of as elements of l, as has been the case.

In this section $\mathcal{G}$ again denotes a given system in $\boldsymbol{\Sigma}$, and not the generalized plant.

Definition 5.1.1 *A system $\mathcal{G} \in \Sigma^{n \mapsto m}$ is* stable w.r.t. $\mathcal{W}_o \in \Sigma^{m \mapsto \bar{m}}$, $\mathcal{W}_i \in \Sigma^{n \mapsto \bar{n}}$ *if*

1. $\mathcal{W}_o \mathcal{G} x \in l_{\infty+}^{\bar{m}}$ *for all* $x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\bar{n}}\right)$ *and*

2. *there exists* $c < \infty$ *such that*

$$\|\mathcal{W}_o \mathcal{G} x\|_{l_\infty} \leq c \|\mathcal{W}_i x\|_{l_\infty} \quad \forall x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{n}\right). \tag{5.1}$$

If $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$, the smallest c satisfying (5.1) is the gain of $\mathcal{G}$ w.r.t. $\mathcal{W}_o, \mathcal{W}_i$ *and is denoted by* $\rho(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i)$.

The motivation for this definition is the same as that of Definition 4.1.1; it guarantees *some* bound on weighted error if the weighted disturbance is bounded. There is a similar stability test provided by the following proposition.

Proposition 5.1.2 *Let $\mathcal{G} \in \Sigma^{n \mapsto m}$, $\mathcal{W}_o \in \Sigma^{m \mapsto \bar{m}}$, and $\mathcal{W}_i \in \Sigma^{n \mapsto \bar{n}}$, and let $\mathcal{W}_i^{-L} \in \Sigma^{\bar{n} \mapsto n}$ be such that $\mathcal{W}_i^{-L}\mathcal{W}_i = \mathcal{I}$ and $\mathcal{W}_i\mathcal{W}_i^{-L}$ is stable. $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ if and only if $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is stable.*

Proof: (if) Suppose $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is stable and consider any $x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\bar{n}}\right)$. Then $\mathcal{W}_o\mathcal{G}x = \mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\mathcal{W}_i x \in l_{\infty+}^{\bar{m}}$ so that condition 1 is satisfied. Also, $\|\mathcal{W}_o\mathcal{G}x\|_{l_\infty} \leq \left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\mathcal{W}_i x\|_{l_\infty}$ so that condition 2 is satisfied by $c := \left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}}$ and $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$.

(only if) Suppose $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}$ is unstable. If there exists $\tilde{x} \in l_{\infty+}^{\bar{n}}$ such that $\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x} \notin l_{\infty+}^{\bar{m}}$, then $x := \mathcal{W}_i^{-L}\tilde{x} \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\bar{n}}\right)$ since $\mathcal{W}_i x = \mathcal{W}_i\mathcal{W}_i^{-L}\tilde{x}$ and $\mathcal{W}_i\mathcal{W}_i^{-L}$ is stable, but $\mathcal{W}_o\mathcal{G}x = \mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x} \notin l_{\infty+}^{\bar{m}}$. Hence condition 1 is violated and $\mathcal{G}$ is not stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$. If such an $\tilde{x}$ does not exist then, given any $c < \infty$, there must exist $\tilde{x} \in l_{\infty+}^{\bar{n}}$ such that

$$\left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x}\right\|_{l_\infty} > c\left\|\mathcal{W}_i\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\tilde{x}\|_{l_\infty}.$$

If $x := \mathcal{W}_i^{-L}\tilde{x}$ then, as above, $x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\bar{n}}\right)$ but

$$\|\mathcal{W}_o\mathcal{G}x\|_{l_\infty} = \left\|\mathcal{W}_o\mathcal{G}\mathcal{W}_i^{-L}\tilde{x}\right\|_{l_\infty} > c\left\|\mathcal{W}_i\mathcal{W}_i^{-L}\right\|_{l_\infty\text{-i}} \|\tilde{x}\|_{l_\infty} \geq c\|\mathcal{W}_i x\|_{l_\infty}.$$

Hence condition 2 is violated and $\mathcal{G}$ is not stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$. □

The next proposition shows that in general incremental stability w.r.t. given weights is a stronger requirement than stability w.r.t. the same weights.

Proposition 5.1.3 *Let $\mathcal{G} \in \Sigma^{n \mapsto m}$, $\mathcal{W}_o \in \Sigma^{m \mapsto \tilde{m}}$, and $\mathcal{W}_i \in \Sigma^{n \mapsto \tilde{n}}$. If $\mathcal{G}$ is incrementally stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ then $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ and $\rho(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i) \leq \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i)$.*

Proof: Consider any $x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\tilde{n}}\right)$. Since $\mathcal{G}$ is incrementally stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$,

$$\begin{aligned} \|\mathcal{P}_N \mathcal{W}_o \mathcal{G} x\|_{l_\infty} &\leq \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i) \, \|\mathcal{P}_N \mathcal{W}_i x\|_{l_\infty} \\ &\leq \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i) \, \|\mathcal{W}_i x\|_{l_\infty} \end{aligned}$$

for all $N \in \mathbf{Z}$. Hence $\mathcal{W}_o \mathcal{G} x \in l_{\infty+}^{\tilde{m}}$ and $\|\mathcal{W}_o \mathcal{G} x\|_{l_\infty} \leq \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i) \, \|\mathcal{W}_i x\|_{l_\infty}$. Since $x \in \mathcal{W}_i^{-1}\left(l_{\infty+}^{\tilde{n}}\right)$ was arbitrary, (5.1) is satisfied by $c := \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i)$, which implies that $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o, \mathcal{W}_i$ and $\rho(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i) \leq \rho_i(\mathcal{G}; \mathcal{W}_o, \mathcal{W}_i)$. □

Assumption 5.1.4 *In the remainder of this section $\mathcal{G}$ denotes a system in $\Sigma_{clti}^{n \mapsto m}$. $\mathcal{W}_o \in \Sigma_{clti}^{n \mapsto \tilde{n}}$ and $\mathcal{W}_i \in \Sigma_{clti}^{m \mapsto \tilde{m}}$ denote stable systems with stable left inverses in Σ_{clti}. Hence they have impulse response matrices $W_o \in l_1^{\tilde{n} \times n}$ and $W_i \in l_1^{\tilde{m} \times m}$ which are left invertible in l_1. Because l_1 is Hermite, the Bezout equations*

$$\begin{bmatrix} W_o^{-L} \\ W_o^{\perp} \end{bmatrix} \begin{bmatrix} W_o & W_o^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} W_i^{-L} \\ W_i^{\perp} \end{bmatrix} \begin{bmatrix} W_i & W_i^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5.2}$$

can be constructed where all additional matrices on the left hand sides of (5.2) are in l_1. In the remainder of this section, they denote arbitrary *choices satisfying the equations, given W_o and W_i.*

Assumption 5.1.4 is identical to Assumption 4.1.3. The next proposition shows that stability w.r.t. any weights satisfying this assumption is equivalent to stability when $\mathcal{G} \in \Sigma_{clti}$, and hence to incremental stability (because of Proposition 4.1.4). Moreover the gain w.r.t. given weights, like the incremental gain, is a norm on the stable systems.

Proposition 5.1.5 *$\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ if and only if $G \in l_1$. Moreover, $\|\cdot\|_{\mathcal{W}_o, \mathcal{W}_i} := \rho(\cdot; \mathcal{W}_o, \mathcal{W}_i)$ is a norm on the space of stable systems in $\Sigma_{clti}^{n \mapsto m}$.*

Proof: Because $\mathcal{W}_i$ satisfies the hypotheses of Proposition 5.1.2 that proposition can be applied to prove the first statement; it says that $\mathcal{G}$ is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ if and only if $\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L}$ is stable, where $\mathcal{W}_i^{-L}$ is any stable left inverse of $\mathcal{W}_i$ in Σ_{clti}. Now if $G \in l_1$ then $W_o G W_i^{-L} \in l_1$ and $\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L}$ is stable and causal. Conversely, if $\mathcal{W}_o \mathcal{G} \mathcal{W}_i^{-L}$ is stable and causal then $W_o G W_i^{-L} \in l_1$; hence $G = W_o^{-L}(W_o G W_i^{-L}) W_i \in l_1$.

For the second statement, $\|\cdot\|_{\mathcal{W}_o,\mathcal{W}_i}$ is non-negative by definition and is well defined for every stable system. It is easy to check that for linear $\mathcal{W}_o$ and $\mathcal{W}_i$, $\|\cdot\|_{\mathcal{W}_o,\mathcal{W}_i}$ scales and is subadditive. Moreover, $\|\mathcal{G}\|_{\mathcal{W}_o,\mathcal{W}_i} = 0$ implies that $\mathcal{G}x = 0$ for all $x \in \mathcal{W}_i^{-1}(l_{\infty+})$, which implies that $\mathcal{G}$ is the zero system. Hence $\|\cdot\|_{\mathcal{W}_o,\mathcal{W}_i}$ is a norm on the stable systems. $\square$

Theorem 5.1.7 will show that $\|\mathcal{G}\|_{\mathcal{W}_o,\mathcal{W}_i}$ can be computed for any $\mathcal{G}$ which is stable w.r.t. $\mathcal{W}_o$, $\mathcal{W}_i$ by solving a minimum distance problem, as in the case of $\|\cdot\|^i_{\mathcal{W}_z,\mathcal{W}_w}$. However in this case the computation is not precisely a version of OPT because the free parameter must be allowed to range over $l_1(\mathbf{Z})$ instead of just l_1. This is the only difference, however; $W_oGW_i^{-L}$ and $W_i^{\perp}$ are still in l_1. Hence the infimization

$$\inf\left\{\left\|W_oGW_i^{-L} - K\right\|_{l_1} : K \in \boldsymbol{K}_Z(I, W_i^{\perp})\right\} =: \gamma_Z. \tag{5.3}$$

where

$$\boldsymbol{K}_Z(I, W_i^{\perp}) := \left\{K \in l_1(\mathbf{Z})^{m\times n} : \exists Q \in l_1(\mathbf{Z})^{m\times p} \text{ satisfying } K = QW_i^{\perp}\right\} \tag{5.4}$$

can be posed and Bezout equations

$$\begin{bmatrix} W_i^{\perp} \\ W_i^{-L} \end{bmatrix} \begin{bmatrix} W_i^c & W_i \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5.5}$$

exist for $W_i^{\perp}$.

Before stating and proving Theorem 5.1.7, a lemma is needed to establish that the key features of OPT needed for the proof of Theorem 4.1.6 for the incremental weighted case have appropriate analogs for the infimization 5.3.

Lemma 5.1.6 *The infimization (5.3) has the following properties:*

- *there exists $K_0 \in \boldsymbol{K}_Z(I, W_i^{\perp})$ such that $\left\|W_oGW_i^{-L} - K_0\right\|_{l_1} = \gamma_Z$, and*
- *given $\epsilon > 0$, there exists $N \in \mathbf{Z}$ and $\tilde{G} \in c_0(\mathbf{Z})^{\tilde{m}\times n}$ such that*

$$\begin{aligned} \left\langle \tilde{G} \triangleright W_i^T, W_oHW_i^{-L}\right\rangle &> \gamma_Z - \epsilon \\ \left\|\tilde{G} \triangleright W_i^T\right\|_{l_\infty} &\leq 1 \\ \operatorname{supp} \tilde{G} &\subset \{\ldots, N-1, N\} \end{aligned}$$

Proof: (First item): If $K \in \boldsymbol{K}_Z(V)$ then (5.5) implies $KW_\text{i} = 0$. Conversely, if $K \in l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}$ and $KW_\text{i} = 0$ then, using the reverse of (5.5),

$$\begin{aligned} K &= K \begin{bmatrix} W_\text{i}^c & W_\text{i} \end{bmatrix} \begin{bmatrix} W_\text{i}^\perp \\ W_\text{i}^{-L} \end{bmatrix} = KW_\text{i}^c W_\text{i}^\perp \\ &= QW_\text{i}^\perp \end{aligned}$$

where $Q := KW_\text{i}^c \in l_1(\mathbf{Z})^{\bar{m}\times(\bar{n}-n)}$. Hence $K \in \boldsymbol{K}_Z(I, W_\text{i}^\perp)$. Thus $K \in \boldsymbol{K}_Z(I, W_\text{i}^\perp)$ if and only if $K \in l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}$ and $KW_\text{i} = 0$.

Define a convolution operator $\mathcal{T}_{W_\text{i}}$ on $l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}$, given $K \in l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}$,

$$\mathcal{T}_{W_\text{i}} K := KW_\text{i}.$$

$\mathcal{T}_{W_\text{i}} \in \mathcal{B}(l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}, l_1(\mathbf{Z})^{\bar{m}\times n})$ because $W_\text{i} \in l_1$, and it is clear that $\boldsymbol{K}_Z(I, W_\text{i}^\perp) = \mathcal{N}\left(\mathcal{T}_{W_\text{i}}\right)$.

Next define a correlation operator $\tilde{\mathcal{T}}_{W_\text{i}}$ on $c_0(\mathbf{Z})^{\bar{m}\times n}$, given $\tilde{G} \in c_0(\mathbf{Z})^{\bar{m}\times n}$,

$$\tilde{\mathcal{T}}_{W_\text{i}} := \tilde{G} \triangleright W_\text{i}^T.$$

$\tilde{\mathcal{T}}_{W_\text{i}} \in \mathcal{B}(c_0(\mathbf{Z})^{\bar{m}\times n}, c_0(\mathbf{Z})^{\bar{m}\times\bar{n}})$ by Proposition C.2.5. Moreover, by that proposition, $\tilde{\mathcal{T}}_{W_\text{i}}^* = \mathcal{T}_{W_\text{i}}$. Hence $\boldsymbol{K}_Z(I, W_\text{i}^\perp) = \mathcal{N}\left(\tilde{\mathcal{T}}_{W_\text{i}}^*\right)$ and is *weak**-closed by [26, Theorem 4.12]. Since $\boldsymbol{K}_Z(I, W_\text{i}^\perp)$ is a subspace and $l_1(\mathbf{Z})^{\bar{m}\times\bar{n}}$ is the dual of $c_0(\mathbf{Z})^{\bar{m}\times\bar{n}}$, Corollary B.2.4 applies to show that there exists a minimizer for the infimization (5.3).

(Second item): Consider the following:

$$\begin{aligned} \gamma_Z &= \inf\left\{ \left\| W_\text{o} H W_\text{i}^{-L} - K \right\|_{l_1} : K \in \boldsymbol{K}_Z(I, W_\text{i}^\perp) \right\} \\ &= \inf\left\{ \left\| W_\text{o} H W_\text{i}^{-L} - K \right\|_{l_1} : K \in \mathcal{N}\left(\mathcal{T}_{W_\text{i}}\right) \right\} \\ &= \sup\left\{ \left\langle \tilde{G}, W_\text{o} H W_\text{i}^{-L} \right\rangle : \boldsymbol{B}\mathcal{R}\left(\tilde{\mathcal{T}}_{W_\text{i}}\right) \right\} \\ &= \sup\left\{ \left\langle \tilde{G} \triangleright W_\text{i}^T, W_\text{o} H W_\text{i}^{-L} \right\rangle : \tilde{G} \in c_0(\mathbf{Z})^{\bar{m}\times n},\ \left\| \tilde{G} \triangleright W_\text{i}^T \right\|_{l_\infty} \leq 1 \right\} \\ &= \sup\left\{ \left\langle G \triangleright W_\text{i}^T, W_\text{o} H W_\text{i}^{-L} \right\rangle : \tilde{G} \in c_0(\mathbf{Z})_-^{\bar{m}\times n},\ \left\| \tilde{G} \triangleright W_\text{i}^T \right\|_{l_\infty} \leq 1 \right\} \end{aligned}$$

where $c_0(\mathbf{Z})_-^{\bar{m}\times n}$ denotes the matrices in $c_0(\mathbf{Z})^{\bar{n}\times n}$ each of whose entries is left supported. The second line follows using the duality theorem B.2.3 and the fact that $\mathcal{R}(\mathcal{T})^\perp = \mathcal{N}(\mathcal{T}^*)$ for any bounded linear operator $\mathcal{T}$. The third line uses the definition of $\tilde{\mathcal{T}}_{W_\text{i}}$ and the fourth follows because the left supported sequences are dense in $c_0(\mathbf{Z})$ (this is easily shown).

Hence, given $\epsilon > 0$, there exists $\tilde{G} \in c_0(\mathbf{Z})_-^{\tilde{m}\times n}$ such that $\left\langle \tilde{G} \triangleright W_{\mathrm{i}}^T, W_{\mathrm{o}}HW_{\mathrm{i}}^{-L} \right\rangle > \gamma_Z - \epsilon$. Take any such $\tilde{G}$; because it is left supported, it is possible to choose N large enough so that $\operatorname{supp} \tilde{G} \subset \{\ldots, N-1, N\}$. □

The proof is very similar to that of Theorem 4.1.6 but the fact that the infimization (5.3) is set in $l_1(\mathbf{Z})$ rather than l_1 and the consequent fact that $\boldsymbol{K}_Z(I, W_{\mathrm{i}}^{\perp})$ is the annihilator of the range of a correlation operator defined on *two-sided* sequences allows construction of an offending disturbance whose weighted l_∞-norm is bounded by 1 *for all time.* In the incremental weighted case the fact that one-sided correlation operators arose prevented this; the weighted l_∞ norm bound on the offending disturbance could be guaranteed only up until some given time.

Theorem 5.1.7 $\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} = \gamma_Z$.

Proof:

$\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} \leq \gamma_Z$: By Lemma 5.1.6 there exists $Q_0 \in l_1(\mathbf{Z})^{\tilde{m}\times(\tilde{n}-n)}$ such that

$$\left\| W_{\mathrm{o}}GW_{\mathrm{i}}^{-L} - Q_0 W_{\mathrm{i}}^{\perp} \right\|_{l_1} = \gamma_Z.$$

Moreover, for all $w \in l_{\infty+}^n$,

$$\begin{aligned} \|\mathcal{W}_{\mathrm{o}}\mathcal{G}w\|_{l_\infty} &= \left\| (\mathcal{W}_{\mathrm{o}}\mathcal{G}\mathcal{W}_{\mathrm{i}}^{-L} - \mathcal{Q}_0\mathcal{W}_{\mathrm{i}}^{\perp})\mathcal{W}_{\mathrm{i}}w \right\|_{l_\infty} \\ &\leq \left\| (\mathcal{W}_{\mathrm{o}}\mathcal{G}\mathcal{W}_{\mathrm{i}}^{-L} - \mathcal{Q}_0\mathcal{W}_{\mathrm{i}}^{\perp}) \right\|_{l_\infty\text{-i}} \|\mathcal{W}_{\mathrm{i}}w\|_{l_\infty} \\ &= \left\| W_{\mathrm{o}}GW_{\mathrm{i}}^{-L} - Q_0W_{\mathrm{i}}^{\perp} \right\|_{l_1} \|\mathcal{W}_{\mathrm{i}}w\|_{l_\infty} \\ &= \gamma_Z \|\mathcal{W}_{\mathrm{i}}w\|_{l_\infty}. \end{aligned}$$

Hence (5.1) is satisfied by $c := \gamma_Z$ and $\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} \leq \gamma_Z$.

$\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} \geq \gamma_Z$: Given any $\epsilon > 0$ we will find $w \in l_{\infty+}^n$ such that $\|\mathcal{W}_{\mathrm{i}}w\|_{l_\infty} \leq 1$ but $\|\mathcal{W}_{\mathrm{o}}\mathcal{G}w\|_{l_\infty} \geq \gamma_Z - \epsilon$. Hence, using Definition 5.1.1, $\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} \geq \gamma_Z - \epsilon$. Since ϵ was arbitrary, $\|\mathcal{G}\|_{\mathcal{W}_{\mathrm{o}},\mathcal{W}_{\mathrm{i}}} \geq \gamma_Z$.

Accordingly, let $\epsilon > 0$ be given. Lemma 4.1.5 can easily be modified for $l_1(\mathbf{Z})$ case to show that

$$\inf\left\{ \left\| (W_{\mathrm{o}}GW_{\mathrm{i}}^{-L})_{i\cdot} - QW_{\mathrm{i}}^{\perp} \right\|_{l_1} : Q \in l_1(\mathbf{Z})^{1\times(\tilde{n}-n)} \right\} \geq \gamma_Z$$

for some $i \in \{1, \ldots, \tilde{m}\}$. Let this value of i be fixed for the remainder of the proof. By Lemma 5.1.6, there exists $N \in \mathbf{Z}$ and $x \in \mathcal{P}_N c_0(\mathbf{Z})^{1\times n}$ such that

$$\left\langle x \triangleright W_{\mathrm{i}}^T, (W_{\mathrm{o}}GW_{\mathrm{i}}^{-L})_{i\cdot} \right\rangle > \gamma_Z - \epsilon \tag{5.6}$$

and

$$\left\| x \triangleright W_{\mathrm{i}}^{T} \right\|_{l_\infty} \leq 1 \tag{5.7}$$

Now define $w \in l_{\infty+}^{n}$ by $w := \{x(-k)\}_{k \in \mathbf{Z}}$; This is the offending w since, for each $i \in \{1, \ldots, \tilde{n}\}$ and $k \in \mathbf{Z}$,

$$\begin{aligned}
(x \triangleright W_{\mathrm{i}}^{T})_i(k) &= \sum_{j=1}^{n} [x_j \triangleright (W_{\mathrm{i}})_{ij}](k) \\
&= \sum_{j=1}^{n} \sum_{n=-\infty}^{\infty} (W_{\mathrm{i}})_{ij}(n) x_j(n+k) = \sum_{j=1}^{n} \sum_{n=-\infty}^{\infty} (W_{\mathrm{i}})_{ij}(n) w_j(-n-k) \\
&= \sum_{j=1}^{n} [(W_{\mathrm{i}})_{ij} * w_j](-k).
\end{aligned}$$

Hence

$$x \triangleright W_{\mathrm{i}}^{T} = \left\{ (W_{\mathrm{i}} * w)^{T}(-k) \right\}_{k \in \mathbf{Z}}.$$

Using (5.7) and the definition of $\|\cdot\|_{l_\infty}$, we conclude that $\|\mathcal{W}_{\mathrm{i}} w\|_{l_\infty} \leq 1$.

On the other hand,

$$\begin{aligned}
\left\langle x \triangleright W_{\mathrm{i}}^{T}, (W_{\mathrm{o}} G W_{\mathrm{i}}^{-L})_{i\cdot} \right\rangle &= \langle x, (W_{\mathrm{o}} G)_{i\cdot} \rangle = \sum_{j=1}^{n} \sum_{n=-\infty}^{\infty} (W_{\mathrm{o}} G)_{ij}(n) x_j(n) \\
&= \sum_{j=1}^{n} \sum_{n=-\infty}^{\infty} (W_{\mathrm{o}} G)_{ij}(n) w_j(-n) \\
&= (W_{\mathrm{o}} G * x)_i(0)
\end{aligned}$$

where the first line follows using the fact (Proposition C.2.5) that the adjoint of a correlation operator is a convolution operator and the second using the definitions of functional evaluation (see Proposition C.2.1) and w. Finally, using (5.6), we conclude that $\|\mathcal{W}_{\mathrm{o}} \mathcal{G} w\|_{l_\infty} > \gamma_Z - \epsilon$, and the proof is complete. □

Proposition 5.1.3 shows that $\|\cdot\|_{\mathcal{W}_{\mathrm{o}}, \mathcal{W}_{\mathrm{i}}} \leq \|\cdot\|^{i}_{\mathcal{W}_{\mathrm{o}}, \mathcal{W}_{\mathrm{i}}}$ in general; of course, the two norms could nonetheless be identical. The following simple example shows that this is not the case. (For ease of notation, the space $\boldsymbol{A}$ of all z-transforms of sequences in l_1 with norm $\left\|\hat{G}\right\|_{A} := \|G\|_{l_1}$ is introduced; this prevents us from needing a notation for sequences. Recall that z is taken as the delay in the definition of the z-transform.)

Example: Let $\hat{H} = \hat{W}_{\mathrm{o}} = 1$,

$$\hat{W}_{\mathrm{i}} = \begin{bmatrix} 1 \\ -1 + 3z \end{bmatrix}$$

and choose $\hat{W}_i^{-L} = [\,1 \;\; 0\,]$ and $\hat{W}_i^{\perp} = [\,1-3z \;\; 1\,]$ to satisfy the Bezout equations (5.2). $\|\mathcal{H}\|^i_{\mathcal{W}_o,\mathcal{W}_i}$ and $\|\mathcal{H}\|_{\mathcal{W}_o,\mathcal{W}_i}$ are computed by solving $\inf \|[\,1 \;\; 0\,] + \hat{q}[\,1-3z \;\; 1\,]\|_A =: \gamma$ where q ranges over l_1 and $l_1(\mathbf{Z})$, respectively. It is not hard to check that $\gamma \geq 1$ when q ranges over l_1 since

$$\|[\,1 \;\; 0\,] + \hat{q}[\,1-3z \;\; 1\,]\|_A = \|1+\hat{q}(1-3z)\|_A + \|\hat{q}\|_A \geq (1-|q_0|) + |q_0| = 1$$

where q_0 is the first element of q. On the other hand, if we take $\hat{q} = \frac{1}{3}z^{-1}$ then $q \in l_1(\mathbf{Z})$ and

$$\|[\,1 \;\; 0\,] + \hat{q}[\,1-3z \;\; 1\,]\|_A = \|1+\hat{q}(1-3z)\|_A + \|\hat{q}\|_A = \frac{2}{3}.$$

Hence $\gamma \leq \frac{2}{3}$ if q is allowed to range over $l_1(\mathbf{Z})$.

In terms of signals it is not hard to see the difference in the norms. It is not hard to check that if $|w(k)| > \frac{2}{3}$ for any k then $\|(\mathcal{W}_i w)(k+1)\|_\infty > 1$. Hence $\|\mathcal{W}_i w\|_{l_\infty} \leq 1 \Rightarrow \|w\|_{l_\infty} \leq \frac{2}{3}$ so that $\|\mathcal{G}\|_{\mathcal{W}_o,\mathcal{W}_i} \leq \frac{2}{3}$. On the other hand, if $w = \delta$ then $\|\mathcal{P}_0\mathcal{W}_i w\|_{l_\infty} = 1$ and $|(\mathcal{W}_o\mathcal{G}w)(0)| = 1$. Hence $\|\mathcal{G}\|^i_{\mathcal{W}_o,\mathcal{W}_i} \geq 1$.

Next it is shown that the gain of a system w.r.t. weights satisfying Assumption 5.1.4 is an induced norm between weighted versions of l_∞. In particular, given any linear weight a weighted norm can be defined on signals which are mapped into l_∞ by the weight.

Definition 5.1.8 *If $x \in l_+$, $\mathcal{W} \in \Sigma_l$, and $\mathcal{W}x \in l_\infty$ then $\rho_{\mathcal{W}}(x) := \|\mathcal{W}x\|_\infty$ is the* $\mathcal{W}$-weighted l_∞-norm *of x.*

With no further assumptions on $\mathcal{W}$, $\rho_{\mathcal{W}}(\cdot)$ is actually only a semi-norm, as it can have a null space. Moreover, it need not be defined on all of l_∞ (e.g., if $W \notin l_1$) and can be defined for signals not in l_∞ (e.g., if W not left invertible in l_1). Under Assumption 4.1.3, however, it is defined precisely on l_∞ and is a norm.

Proposition 5.1.9 *If $W \in l_1$ and has a left inverse in l_1 then $\mathcal{W}x \in l_\infty$ if and only if $x \in l_\infty$, and $\|\cdot\|_{\mathcal{W}} := \rho_{\mathcal{W}}(\cdot)$ is a norm on l_∞.*

Proof: Let W^{-L} denote any left inverse of W in l_1. For the first claim, if $x \in l_\infty$ then $\mathcal{W}x \in l_\infty$ since $W \in l_1$. Conversely, if $\mathcal{W}x \in l_\infty$ then $x = \mathcal{W}^{-L}\mathcal{W}x \in l_\infty$ since $W^{-L} \in l_1$. For the second claim, the properties of a semi-norm follow from the linearity of $\mathcal{W}$ and the corresponding properties of $\|\cdot\|_\infty$. Moreover, $\|x\|_{\mathcal{W}} = 0 \Rightarrow \mathcal{W}x = 0 \Rightarrow x = \mathcal{W}^{-L}\mathcal{W}x = 0$. $\square$

l_∞ under $\|\cdot\|_{\mathcal{W}}$ can be called *$\mathcal{W}$-weighted l_∞*. It is clear that $\|\mathcal{H}\|_{\mathcal{W}_o,\mathcal{W}_i}$ is the induced norm of $\mathcal{H}$ viewed as a map from $\mathcal{W}_i$-weighted l_∞ to $\mathcal{W}_o$-weighted l_∞ since

$$\begin{aligned}\|\mathcal{H}\|_{\mathcal{W}_o,\mathcal{W}_i} &= \sup\{\|\mathcal{W}_o\mathcal{H}x\|_\infty : \mathcal{W}_i x \in l_\infty,\ \|\mathcal{W}_i x\|_\infty \le 1\} \\ &= \sup\left\{\|\mathcal{H}x\|_{\mathcal{W}_o} : x \in l_\infty,\ \|x\|_{\mathcal{W}_i} \le 1\right\}.\end{aligned}$$

using the definition of $\|\mathcal{H}\|_{\mathcal{W}_o,\mathcal{W}_i}$ for the first equality and Definition 5.1.8 and Proposition 5.1.9 for the second.

5.2 Problem Statement

Recall the specification:

Weighted l_∞ DRS:

- $\mathcal{C} \in \Sigma_{clti}$, $(\mathcal{G}, \mathcal{C})$ is stable, and
- $\mathcal{W}_w w \in l_{\infty+}^{n_{\tilde{w}}}$ and $\|\mathcal{W}_w w\|_{l_\infty} \le 1$ implies $\mathcal{W}_z z \in l_{\infty+}^{n_{\tilde{z}}}$ and $\|\mathcal{W}_z z\|_{l_\infty} \le 1$.

where the weights $\mathcal{W}_z$ and $\mathcal{W}_w$ are systems in $\Sigma_{clti}^{n_z \mapsto n_{\tilde{z}}}$ and $\Sigma_{clti}^{n_z \mapsto n_{\tilde{z}}}$, respectively.

If $\mathcal{G} \in \Sigma_{clti}$ and $\mathcal{W}_z, \mathcal{W}_w \in \Sigma_{clti}$ are stable and have stable left inverses then the results of Section 5.1 can be used to formulate the design problem as follows.

$$DR^w(\mathcal{G}): \qquad \inf\left\{\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|_{\mathcal{W}_z,\mathcal{W}_w} : \mathcal{C} \in \mathcal{C}(\mathcal{G})\right\} =: \mu_{WDR}$$

If $\mathcal{C} \in \mathcal{C}(\mathcal{G})$ then $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ is stable and hence the norm above is always defined; μ_{WDR} is defined if and only if $\mathcal{G}$ is stabilizable. Moreover, if $\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})$ is stable then

$$\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|_{\mathcal{W}_z,\mathcal{W}_w} = \sup\left\{\|\mathcal{W}_z z\|_{l_\infty} : w \in l_{\infty+}^{n_w},\ \|\mathcal{W}_w w\|_{l_\infty} \le 1\right\}.$$

$DR^w(\mathcal{G})$ represents the design problem in the usual sense; the feasible solutions for $DR^w(\mathcal{G})$ are the stabilizing compensators for $\mathcal{G}$ and the cost of each is the worst-case $\|\mathcal{W}_z z\|_{l_\infty}$ over all $\|\mathcal{W}_w w\|_{l_\infty} \le 1$ when that $\mathcal{C}$ is used. $DR^w(\mathcal{G})$ is feasible if and only if the first item of the specification is met, and $\mu_{WDR} \le 1$ if and only if the second item of the specification can be met.

Notation and Assumptions

In the remainder of this chapter all assumptions on $\mathcal{G}$, $\mathcal{W}_z$, and $\mathcal{W}_w$ made in Chapter 4 are in effect. $\mathcal{G} \in \Sigma_{clti}$, G_{yu} has a coprime factorization over l_1, $\mathcal{G}$ stabilizable, and associated H, U and V matrices have decompositions as in Assumption 3.0.1. $\mathcal{W}_z$ and $\mathcal{W}_w$ are in $\Sigma_{clti}^{n_z \mapsto n_{\bar{z}}}$ and $\Sigma_{clti}^{n_w \mapsto n_{\bar{w}}}$, respectively, are stable, and have stable left inverses in Σ_{clti}. Recall the associated Bezout equations

$$\begin{bmatrix} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} \begin{bmatrix} W_z & W_z^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} W_w^{-L} \\ W_w^{\perp} \end{bmatrix} \begin{bmatrix} W_w & W_w^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad (5.8)$$

where the additional symbols on the left hand sides denote *arbitrary* matrices over l_1 satisfying the equations.

All other notation in Chapter 4 has the same meaning here.

5.3 Formulation as a Minimum Distance Problem

Because computation of $\|\mathcal{G}\|_{\mathcal{W}_z,\mathcal{W}_w}$ is a minimum distance problem in $l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$, $DR^{w}(\mathcal{G})$ can also be formulated as one:

$$OPT^{w}: \qquad \inf\left\{ \left\| W_z H W_w^{-L} - K \right\|_{l_1} : K \in \boldsymbol{K}^{w} \right\} =: \mu^{w}$$

where

$$\boldsymbol{K}^{w} := \left\{ K : \exists Q_C \in l_1,\ Q_W \in l_1(\mathbf{Z}) \text{ s. t. } K = \begin{bmatrix} W_z U & I \end{bmatrix} \begin{bmatrix} Q_C & 0 \\ 0 & Q_W \end{bmatrix} \begin{bmatrix} V W_w^{-L} \\ W_w^{\perp} \end{bmatrix} \right\}$$

is a subspace of $l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$. OPT^{w} depends on H, U, V, W_z, and W_w and $\boldsymbol{K}^{w}$ depends on all of these except H; these dependences are suppressed.

The following analog of Theorem 4.3.1 (whose statement is simplified somewhat as well) shows that OPT^{w} is equivalent to $DR^{w}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$.

Theorem 5.3.1 *$DR^{w}(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ and OPT^{w} are equivalent in the following sense:*

1. *If $C \in \mathcal{C}(\mathcal{G})$ and $Q \in \mathbf{Q}(\mathcal{G}_{yu})$ is constructed from it as in Proposition 2.3.11 then $K := W_z U Q V W_w^{-L} + K_\rho \in \boldsymbol{K}^{w}$ where K_ρ is a minimizer for (5.3). Moreover $\left\| W_z H W_w^{-L} - K \right\|_{l_1} = \|\mathcal{T}_{zw}(\mathcal{G}, C)\|_{\mathcal{W}_z,\mathcal{W}_w}$.*

2. *If $K \in \boldsymbol{K}^{\mathrm{w}}$, $Q \in \boldsymbol{Q}(\mathcal{G}_{yu})$ solves $K = W_z U Q V W_w^{-L} + Q_W W_w^{\perp}$ for some $Q_W \in l_1(\mathbf{Z})$, and $\mathcal{C}$ is constructed from it as in Proposition 2.3.11 then $\mathcal{C} \in \mathcal{C}(\mathcal{G})$ and $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|_{\mathcal{W}_z,\mathcal{W}_w} \leq \left\|W_z H W_w^{-L} - K\right\|_{l_1}$.*

3. *$\mu^{\mathrm{w}} = \mu_{WDR}$.*

Proof: For item 1, such a choice of K is possible because, by Lemma 5.1.6, a minimizer always exists for (5.3). That $K \in \boldsymbol{K}^{\mathrm{w}}$ follows because every feasible solution $\bar{K}$ for the norm computation (5.3) can be written $\bar{K} = Q_W W_w^{\perp}$ for some $Q_W \in l_1(\mathbf{Z})$; hence so can $\boldsymbol{K}_\rho$. The equality of the norms is immediate using Theorem 5.1.7 and the fact that, by Proposition 2.3.11 and the definition of K, $W_z H W_w^{-L} + K = W_z(H - UQV)W_w^{-L} + K_\rho = W_z T_{zw}(\mathcal{G},\mathcal{C})W_w^{-L} + K_\rho$.

For item 2, the existence of a $Q \in l_1^{n_u \times n_y}$ is guaranteed by the definition of $\boldsymbol{K}^{\mathrm{w}}$. By Proposition 2.3.11 and the choice of Q

$$W_z T_{zw}(\mathcal{G},\mathcal{C})W_w^{-L} - Q_W W_w^{\perp} = W_z(H - UQV)W_w^{-L} + Q_W W_w^{\perp} = W_z H W_w^{-L} - K.$$

The norm inequality follows from Theorem 5.1.7 which shows that $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C})\|_{\mathcal{W}_z,\mathcal{W}_w}$ is the infimum over $Q_W \in l_1(\mathbf{Z})$ of the l_1 norm of the left hand side above.

Item 3 follows from the preceding items. □

The consequences of Theorem 5.3.1 are the same as those of Theorem 4.3.1:

- To each feasible solution for $DR^{\mathrm{w}}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ there corresponds a feasible solution for OPT^{w} of the same cost. From every feasible solution $K \in \boldsymbol{K}^{\mathrm{w}}$ for OPT^{w}, one or more corresponding feasible solutions for $DR^{\mathrm{w}}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ can be constructed of no greater cost.

- To each minimizer for $DR^{\mathrm{w}}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ there corresponds a minimizer for OPT^{w}; hence *all* minimizers for $DR^{\mathrm{w}}(\mathcal{G},\mathcal{W}_z,\mathcal{W}_w)$ can be found from minimizers for OPT^{w} via the construction of item 2.

Item 2 of the theorem and the above discussion are simplified as for the incremental case: the provision in item 2 for the case in which the constructed Q is not in $\boldsymbol{Q}(\mathcal{G}_{yu})$ is omitted. A random perturbation on Q can also be used here to produce a stabilizing $\mathcal{C}_\epsilon$ such that $\|\mathcal{T}_{zw}(\mathcal{G},\mathcal{C}_\epsilon)\|_{\mathcal{W}_z,\mathcal{W}_w} \leq \left\|W_z H W_w^{-L} - K\right\|_{l_1} + \epsilon$ for any $\epsilon > 0$. If $G_{yu}(0) = 0$ then this case does not arise.

To compute stabilizing $\mathcal{C}$s whose performance does not exceed $\left\|W_z H W_w^{-L} - K\right\|_{l_1}$ given a feasible solution K for OPT^{w}, the same parametrization as given in Proposition 4.3.2 applies, i.e., if $K \in \boldsymbol{K}^{w}$ then the set

$$\boldsymbol{Q}^{w}(K) := \left\{Q \in l_1^{n_u \times n_y} : K = W_z U Q V W_w^{-L} + Q_W W_w^{\perp} \text{ for some } Q_W \in l_1(\mathbf{Z})\right\}$$

is given by

$$\left\{\begin{bmatrix} U_R^{-R} & U_R^{\perp} \end{bmatrix} \begin{bmatrix} \Sigma_U^{-1} U_L^{-L} W_z^{-L} K W_w V_R^{-R} \Sigma_V^{-1} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} V_L^{-L} \\ V_L^{\perp} \end{bmatrix} : Q_{12}, Q_{21}, Q_{22} \in l_1\right\}$$

The proof of Proposition 4.3.2 goes through to establish this; the fact that Q_W is allowed to be in $l_1(\mathbf{Z})$ is immaterial to the parametrization.

5.4 Existence of a Minimizer

The following lemma shows that relationship between the feasible subspace $\boldsymbol{K}^{w}$ of OPT^{w} and $\boldsymbol{K}(W_z U, V)$ is analogous to that between $\boldsymbol{K}^{iw}$ and $\boldsymbol{K}(U_z, V)$. The proof is essentially identical to that of Lemma 4.4.1 but is given to reinforce the similarity of the situations.

Lemma 5.4.1 $\boldsymbol{K}^{w} = \{K \in l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}} : K W_w \in \boldsymbol{K}(W_z U, V)\}$.

Proof: If $K \in \boldsymbol{K}^{w}$ then $K \in l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$ and $K = W_z U Q_C V W_w^{-L} + Q_W W_w^{\perp}$ and hence, using (5.8), $K W_w = W_z U Q_C V$. Since $Q_C \in l_1$, $K W_w \in \boldsymbol{K}(W_z U, V)$. Conversely, if $K \in l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$ and $K W_w \in \boldsymbol{K}(W_z U, V)$ then $K W_w = W_z U Q_C V$ for some $Q_C \in l_1$. Using the reverse of (5.8),

$$\begin{aligned} K &= K(W_w W_w^{-L} + W_w^{c} W_w^{\perp}) = W_z U Q_C V W_w^{-L} + K W_w^{c} W_w^{\perp} \\ &= \begin{bmatrix} W_z U & I \end{bmatrix} \begin{bmatrix} Q_C & 0 \\ 0 & Q_W \end{bmatrix} \begin{bmatrix} V W_w^{-L} \\ W_w^{\perp} \end{bmatrix} \end{aligned}$$

where $Q_W := K W_w^{c} \in l_1(\mathbf{Z})$ since both K and W_w^{c} are. Hence $K \in \boldsymbol{K}^{w}$. □

As a consequence, a characterization of the same form as that of $\boldsymbol{K}^{iw}$ is possible. The difference is that the convolution of item 2 is *two-sided* because $K \in l_1(\mathbf{Z})^{n_{\bar{z}} \times n_{\bar{w}}}$ rather than $l_1^{n_{\bar{z}} \times n_{\bar{w}}}$, and an additional requirement (item 4) is needed.

Theorem 5.4.2 *$K \in \boldsymbol{K}^{w}$ if and only if*

1. $K \in l_1(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}}$,

2. $$\begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix} = \begin{bmatrix} * & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$
   where $*$ denotes an irrelevant block, and

3. $[(\hat{U}_L^{-L} \hat{W}_z^{-L})_{i\cdot} \hat{K} (\hat{W}_w \hat{V}_R^{-R})_{\cdot j}]^{(n)}(z_0) = 0$ *for each* $(i,j) \in \{1, \ldots, r_U\} \times \{1, \ldots, r_V\}$, *each* $z_0 \in \boldsymbol{Z}_{ij}$, *and each* $n \in \{0, \ldots, m_{ij}(z_0) - 1\}$, *and*

4. $KW_w \in l_1^{n_{\tilde{z}} \times n_w}$.

Proof: The proof is similar to that of Theorem 4.4.2, but enforcing the requirement of Lemma 5.4.1 that $KW_w \in l_1^{n_{\tilde{z}} \times n_w}$ and not just $KW_w \in l_1(\mathbf{Z})^{n_{\tilde{z}} \times n_w}$ requires the extra item 4 (the point is that the convolution must be zero for negative time so that the free parameter corresponding to the compensator is ensured to be causal). □

The proof of the existence theorem in this case is similar to previous ones but complicated by the need for a mix of two-sided and one-sided spaces. It is presented as a whole as the requirements of Theorem 5.4.2 do not partition neatly. In spite of the added complication, there is no essential difficulty in establishing the result.

Theorem 5.4.3 *OPT^{w} has a minimizer.*

Proof: An operator $\tilde{\mathcal{T}}_K^{w} \in \mathcal{B}(c_0^{S_W} \times \mathbf{R}^{m_t} \times c_0^{n_{\tilde{z}} \times n_{\tilde{w}}}, c_0(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}})$ will be constructed from various previously defined operators such that $\boldsymbol{K}^{w} = \mathcal{N}\left((\tilde{\mathcal{T}}_K^{w})^*\right)$, establishing existence.

Define $\tilde{\mathcal{T}}_K^{w}$, given $(G_C, \alpha, G_-) \in c_0^{S_W} \times \mathbf{R}^{m_t} \times c_0^{n_{\tilde{z}} \times n_{\tilde{w}}}$,

$$\tilde{\mathcal{T}}_K^{w}[(G_C, \alpha, G_-)] := \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_+ \tilde{\mathcal{T}}_{W_z U,V} \tilde{\mathcal{E}}_{S_W} G_C + \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_+ \tilde{\mathcal{T}}_{W_z U,V} \tilde{\mathcal{T}}_{IZ}^{iw} \alpha + \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_- G_- \quad (5.9)$$

where $\tilde{\mathcal{T}}_{IZ}^{iw} \in \mathcal{B}(\mathbf{R}^{m_t}, c_0^{n_{\tilde{z}} \times n_w})$ is defined in the proof of Lemma 4.4.4 and $\tilde{\mathcal{T}}_{W_z U,V} \in \mathcal{B}(c_0^{n_{\tilde{z}} \times n_w})$ and $\tilde{\mathcal{E}}_{S_W} \in \mathcal{B}(c_0^{S_W}, c_0^{n_{\tilde{z}} \times n_{\tilde{w}}})$ in the proof of Lemma 4.4.3.

$\tilde{\mathcal{T}}_{W_w} \in \mathcal{B}(c_0(\mathbf{Z})^{n_{\tilde{z}} \times n_w}, c_0(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}})$ is defined

$$\tilde{\mathcal{T}}_{W_w} G := G \triangleright W_w^T,$$

$\tilde{\mathcal{E}}_+ \in \mathcal{B}(c_0^{n_{\tilde{z}}\times n_w}, c_0(\mathbf{Z})^{n_{\tilde{z}}\times n_w})$ is defined

$$(\tilde{\mathcal{E}}_+ G)_{ij}(k) := \begin{cases} G_{ij}(k) & k \in \mathbf{Z}_+ \\ 0 & k < 0 \end{cases}$$

and $\tilde{\mathcal{E}}_- \in \mathcal{B}(c_0^{n_{\tilde{z}}\times n_{\tilde{w}}}, c_0(\mathbf{Z})^{n_{\tilde{z}}\times n_{\tilde{w}}})$ is defined

$$(\tilde{\mathcal{E}}_- G)_{ij}(k) := \begin{cases} G_{ij}(-k-1) & k < 0 \\ 0 & k \in \mathbf{Z}_+ \end{cases}$$

Computing the adjoint using previous computations for $\tilde{\mathcal{T}}_{IZ}^{iw}$, $\tilde{\mathcal{T}}_{W_zU,V}$, and $\tilde{\mathcal{E}}_{S_W}$, Proposition C.2.5 for $\tilde{\mathcal{T}}_{W_w}$, and the facts that $\mathcal{E}_+^* = \Pi_+ \in \mathcal{B}(l_1(\mathbf{Z})^{n_{\tilde{z}}\times n_w}, l_1^{n_{\tilde{z}}\times n_w})$ defined, given $K \in l_1^{n_{\tilde{z}}\times n_w}$, $\Pi_+ K := \{K(k)\}_{k\in\mathbf{Z}_+}$, and that $\mathcal{E}_-^* = \Pi_- \in \mathcal{B}(l_1(\mathbf{Z})^{n_{\tilde{z}}\times n_{\tilde{w}}}, l_1^{n_{\tilde{z}}\times n_{\tilde{w}}})$ defined, given $K \in l_1^{n_{\tilde{z}}\times n_{\tilde{w}}}$, $\Pi_+ K := \{K(-k-1)\}_{k\in\mathbf{Z}_+}$ we find that $(\tilde{\mathcal{T}}_K^w)^*$ can be written, given $K \in l_1^{n_{\tilde{z}}\times n_{\tilde{w}}}$,

$$(\tilde{\mathcal{T}}_K^w)^* K = (\Pi_{S_W} \mathcal{T}_{W_zU,V} \Pi_+ \mathcal{T}_{W_w} K,\ \mathcal{T}_I^{iw} \mathcal{T}_{W_zU,V} \Pi_+ \mathcal{T}_{W_w} K,\ \Pi_- \mathcal{T}_{W_w} K)$$

In spite of the complication, it is not hard to check that $\boldsymbol{K}^w = \mathcal{N}\left((\tilde{\mathcal{T}}_K^w)^*\right)$; the last component ensures item 4 of Theorem 5.4.2, i.e., that $KW_w \in l_1^{n_{\tilde{z}}\times n_{\tilde{w}}}$. The first two components combined with the third ensure, respectively, items 2 and 3. □

Of course a corollary analogous to Corollary 3.2.5 holds in this case as well.

Corollary 5.4.4 *If $G_{yu}(0) = 0$ then $DR^w(\mathcal{G}, \mathcal{W}_z, \mathcal{W}_w)$ has a minimizer.*

5.5 Suboptimal Solutions via FIR Approximation

OPT^w is again infinite dimensional problem as OPT^{iw} is. In this section, an FIR upper bound approximation method for OPT^w is formulated which closely mirrors that for OPT^{iw}. It is also applicable in precisely the cases in which its analog is for OPT^{iw}.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\overline{OPT}_n^w: \qquad \inf\left\{\left\|W_z H W_w^{-L} - K\right\|_{l_1} : K \in \overline{\boldsymbol{K}}^w\right\} =: \overline{\mu}_n^w$$

where

$$\overline{\boldsymbol{K}}_n^w := \left\{K \in \boldsymbol{K}^w : \text{supp } (W_z H W_w^{-L} - K) \subset \{-n, \dots, n\}\right\}$$

Note that $\overline{\boldsymbol{K}}_n^{\mathrm{w}}$, is a subset but not a subspace of $l_1(\mathbf{Z})^{n_{\hat{z}}\times n_{\hat{w}}}$. Both $\overline{OPT}_n^{\mathrm{w}}$ and $\overline{\boldsymbol{K}}_n^{\mathrm{w}}$ depend on H, U, V, W_z, and W_w; these dependences are suppressed.

The following proposition shows that the sequence of problems, when feasible, approximates OPT^{w} from above under the same assumptions as in the case of OPT^{iw}.

Proposition 5.5.1 *If there exists $N \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}_N^{\mathrm{w}} \neq \emptyset$ and the finitely supported matrices are dense in $\boldsymbol{K}^{\mathrm{w}}$ then*

$$\{\overline{\mu}_n^{\mathrm{w}}\}_{n=N}^{\infty} \searrow \mu_{OPT} \text{ as } n \to \infty.$$

Proof: The proof is virtually identical to that of Proposition 4.5.1 with a minor adjustment for the two-sided truncation; the argument is not repeated. □

For finitely supported weights (i.e., $\hat{W}_z$ and $\hat{W}_w$ polynomial with polynomial left inverses) the FIR approximation scheme for OPT^{w} can still be used precisely when it can for OPT.

Proposition 5.5.2 *If W_z and W_w are finitely supported and have finitely supported left inverses in l_1 then there exists $n \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}_n^{\mathrm{w}} \neq \emptyset$ if and only if there exists $n \in \mathbf{Z}_+$ such that $\overline{\boldsymbol{K}}_n(H, U, V) \neq \emptyset$.*

Proof: If there exists $K \in \boldsymbol{K}(U, V)$ such that $H - K$ is finitely supported then by Proposition 4.5.2 there is a feasible solution $\tilde{K}$ for OPT^{iw} such that $W_z H W_w^{-L} - \tilde{K}$ is finitely supported; embedding this in $l_1(\mathbf{Z})^{n_{\hat{z}}\times n_{\hat{w}}}$ gives a feasible solution for OPT^{w} such that $W_z H W_w^{-L} - K$ is finitely supported. The converse is established identically to the proof of Proposition 4.5.2. □

5.6 A Converging Lower Bound on μ_{WDR}

The super-optimal approximation scheme, although only the infinite dimensional version, is developed in this section. The additional complication of two-sided sequence spaces makes it unclear whether or not a finitely supported minimizer exists for each such problem as it does in the previous problems, although it is a reasonable conjecture. In any event, a double iteration scheme can be formulated as in Section 4.6 and computational experience may help determine whether this is indeed the case.

Define for each $n \in \mathbf{Z}_+$ an optimization problem

$$\underline{OPT}_n^w \qquad \inf\left\{\left\|W_z H W_w^{-L} - K\right\|_{l_1} : K \in \underline{\boldsymbol{K}}_n^w\right\} =: \underline{\mu}_n^w$$

where

$$K \in \underline{\boldsymbol{K}}_n^w \iff K \in l_1^{n_{\bar{z}} \times n_{\bar{w}}}, \tag{5.10}$$

K satisfies Conditions 3 and 4 of Theorem 5.4.2, and

$$\left(\begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix}\right)(k) = \begin{bmatrix} * & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

for all $k \in \{0, \ldots, n\}$

$\underline{\boldsymbol{K}}_n^w$ is a subspace of $\boldsymbol{K}^w$.

Theorem 5.6.1 *$\{\underline{\mu}_n^w\}_{n=0}^{\infty} \nearrow \mu^w$ as $n \to \infty$. Moreover, for each $n \in \mathbf{Z}_+$, $\underline{OPT}_n^w$ has a minimizer.*

Proof: The proof is very similar to that of Theorem 3.4.1. $\{\underline{\mu}_n^w\}_{n=0}^{\infty}$ is well defined because, for any n, $0 \in \underline{\boldsymbol{K}}_n^w$ and, for any $K \in \underline{\boldsymbol{K}}_n^w$, $\left\|W_z H W_w^{-L} - K\right\|_{l_1} \geq 0$. Also, $\underline{\mu}_n^w \leq \mu^w$ for each n and the sequence is non-decreasing since $n_1 > n_2$ implies $\underline{\boldsymbol{K}}_{n_2} \supset \underline{\boldsymbol{K}}_{n_1} \supset \boldsymbol{K}^w$.

To show convergence define for each n a map $\tilde{\underline{\mathcal{I}}}_n^{iw}$, given $(G_C, \alpha, G_-) \in c_0^{S_W} \times \mathbf{R}^{m_t} \times c_0^{n_{\bar{z}} \times n_{\bar{w}}}$,

$$\tilde{\underline{\mathcal{I}}}_n^w[(G_C, \alpha, G_-] := \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_+ \tilde{\mathcal{T}}_{W_z U, V} \tilde{\mathcal{E}}_{S_W} \tilde{\mathcal{P}}_n G_C + \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_+ \tilde{\mathcal{T}}_{W_z U, V} \tilde{\mathcal{T}}_{IZ}^{iw} \alpha + \tilde{\mathcal{T}}_{W_w} \tilde{\mathcal{E}}_- G_- \tag{5.11}$$

where all operators are defined as in the proof of Theorem 5.4.3 except $\tilde{\mathcal{P}}_n$, which is the n-th truncation operator restricted to $c_0^{S_W}$ all operators are bounded, as has been established previously. Moreover, using previous computations and the fact that $\tilde{\mathcal{P}}_n^* = \mathcal{P}_n$,

$$(\tilde{\underline{\mathcal{I}}}_n^w)^* K = (\mathcal{P}_n \Pi_{S_W} \mathcal{T}_{W_z U, V} \Pi_+ \mathcal{T}_{W_w} K,\ \mathcal{T}_I^{iw} \mathcal{T}_{W_z U, V} \Pi_+ \mathcal{T}_{W_w} K,\ \Pi_- \mathcal{T}_{W_w} K).$$

Continuing with the argument, $\mathcal{R}\left(\tilde{\underline{\mathcal{I}}}_n^w\right)^{\perp} = \mathcal{N}\left((\tilde{\underline{\mathcal{I}}}_n^w)^*\right) = \underline{\boldsymbol{K}}_n^w$ so the duality theorem B.2.3 gives

$$\begin{aligned} \underline{\mu}_n^w &= \sup\left\{\left\langle \tilde{G}, W_z H W_w^{-L}\right\rangle : \boldsymbol{B}\mathcal{R}\left(\tilde{\underline{\mathcal{I}}}_n^w\right)\right\} \\ &= \sup\left\{\left\langle \tilde{\mathcal{T}}_K^w \tilde{G}, W_z H W_w^{-L}\right\rangle : \tilde{G} \in \mathcal{R}\left(\tilde{\mathcal{P}}_n\right) \times \mathbf{R}^{m_t} \times c_0^{n_{\bar{z}} \times n_{\bar{w}}},\ \left\|\tilde{\mathcal{T}}_K^w \tilde{G}\right\|_{l_\infty} \leq 1\right\} \end{aligned}$$

where $\tilde{\mathcal{T}}_K^w$ is defined in equation (5.9). Hence

$$\begin{aligned}
\sup_n \underline{\mu}_n^w &= \sup\left\{\left\langle \tilde{\mathcal{T}}_K^w \tilde{G}, W_z H W_w^{-L}\right\rangle : \tilde{G} \in c_0^{Sw} \times \mathbf{R}^{m_t} \times c_0^{n_{\tilde{z}} \times n_{\tilde{w}}},\ \left\|\tilde{\mathcal{T}}_K^w \tilde{G}\right\|_{l_\infty} \le 1\right\} \\
&= \sup\left\{\left\langle G, W_z H W_w^{-L}\right\rangle : \boldsymbol{B\mathcal{R}}\left(\tilde{\mathcal{T}}_K^w\right)\right\} \\
&= \inf\left\{\left\|W_z H W_w^{-L} - K\right\|_{l_1} : K \in \mathcal{N}\left((\tilde{\mathcal{T}}_K^w)^*\right)\right\} \\
&= \mu^w
\end{aligned}$$

For the second statement, $\underline{\boldsymbol{K}}_n^w = \mathcal{N}\left((\tilde{\underline{\mathcal{T}}}_n^w)^*\right)$; hence it is *weak**-closed and $\underline{OPT}_n^w$ has a minimizer. □

Although finitely supported minimizers are not guaranteed, if they should exist they can be used to compute stabilizing compensators as in the previous problems.

Proposition 5.6.2 *If* $\underline{K} \in \underline{\boldsymbol{K}}_{n,m}^w$ *for any* n, m *then*

$$K := W_z U_L U_L^{-L} W_z^{-L} \underline{K} W_w V_R^{-R} V_R W_w^{-L} \in \boldsymbol{K}^w.$$

Proof: If $\underline{K} \in \underline{\boldsymbol{K}}_n^w$ it is easy to check using the definition in (5.10) that K satisfies the three conditions of Theorem 5.4.2 and hence $K \in \boldsymbol{K}(U, V)$. First, $\underline{K} \in \underline{\boldsymbol{K}}_n^w \Rightarrow K \in l_1^{n_{\tilde{z}} \times n_{\tilde{w}}}$ so that condition 1 is satisfied. Also,

$$\begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix} K \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix} = \begin{bmatrix} U_L^{-L} W_z^{-L} \underline{K} W_w V_R^{-R} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

so that conditions 2 and 3 are satisfied. □

5.7 Linear Programming Formulations

In this section linear programs are formulated corresponding to all the optimization problems defined thus far. The linear programs are all equivalent to the original problems in the sense described in Section 3.5. The notation of that section is also in effect here: $l_1^+(\mathbf{Z})$ for the positive cone in $l_1(\mathbf{Z})$ and a projection operator $\Pi_+ : l_1(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}} \mapsto l_1^+(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}}$ which sets the negative elements of each sequence in a matrix to zero. For compactness of notation,

$$\tilde{U} := \begin{bmatrix} U_L^{-L} W_z^{-L} \\ U_L^{\perp} W_z^{-L} \\ W_z^{\perp} \end{bmatrix}, \quad \tilde{V} := \begin{bmatrix} W_w V_R^{-R} & W_w V_R^{\perp} \end{bmatrix},$$

and

$$\tilde{H} := \tilde{U} H \tilde{V}$$

throughout the section. In addition,

$$\mathcal{T}_I^{w} := \mathcal{T}_I^{iw} \mathcal{T}_{W_z U,V} \Pi_+ \mathcal{T}_{W_w}$$

where $\mathcal{T}_I^{iw}$ is defined in the proof of Lemma 4.4.3, $\mathcal{T}_{W_z U,V}$ in the proof of Lemma 4.4.3, and Π_+ and $\mathcal{T}_{W_w}$ in the proof of Theorem 5.4.3. (Note that Π_+ here is a different projection; there should be no confusion as it does not arise again separately for the rest of this section.)

The first linear program corresponds to OPT^{w} defined in Section 5.3. The linear program is

$$LP^{w}: \qquad \inf \mu$$

subject to:

$$\begin{array}{rcll}
\left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=-\infty}^{\infty} (T^+ + T^-)_{ij}(k)\right] - \mu & \leq & 0 & i = 1, \ldots, n_{\tilde{z}} \\
\left[\tilde{U}(T^+ - T^-)\tilde{V}\right]_{ij}(k) & = & \tilde{H}_{ij}(k) & (i,j) \in \boldsymbol{S}_W, \quad k \in \mathbf{Z}_+ \\
\mathcal{T}_I^{w}(T^+ - T^-) & = & \mathcal{T}_I^{w} H & \\
\left[(T^+ - T^-)W_w\right]_{ij}(k) & = & 0 & k \in \{\ldots, -2, -1\} \qquad (5.12) \\
(\mu, T^+, T^-) & \in & \mathbf{R}_+ \times l_1^+(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}} \times l_1^+(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}} &
\end{array}$$

The variables in LP^{w} are $\mu \in \mathbf{R}$ and

$$\left\{T_{ij}^+(k),\ T_{ij}^-(k) : i \in \{1, \ldots, n_{\tilde{z}}\},\ j \in \{1, \ldots, n_{\tilde{w}}\},\ k \in \mathbf{Z}\right\}.$$

Note that they are infinite in number. There are $n_{\tilde{z}}$ (inequality) cost constraints, an infinite number of (equality) convolution constraints, a finite number of (equality) interpolation constraints, and an infinite number of (inequality) positivity constraints as in the corresponding linear program in the incremental case. The additional infinite set of constraints (5.12) arise from Condition 4 of Theorem 5.4.2 and are called *causality constraints.*

Proposition 5.7.1 *LP^w is equivalent to OPT^w under the map ψ^w defined, given a feasible solution (μ, T^+, T^-) for LP^w,*

$$\psi^w(\mu, T^+, T^-) := W_z H W_w^{-L} - (T^+ - T^-).$$

Proof: Use $\tilde{\psi}^w$ defined, given $K \in \boldsymbol{K}^w$,

$$\tilde{\psi}^w K := \left(\left\| W_z H W_w^{-L} - K \right\|_{l_1}, \ \Pi_+(W_z H W_w^{-L} - K), \ \Pi_+(K - W_z H W_w^{-L}) \right).$$

$\square$

The next linear program corresponds to an FIR approximation problem $\overline{OPT}_n^w$, defined in Section 5.5. The linear program is

$\overline{LP}_n^w$: $\inf \mu$

subject to:

$$\begin{array}{rcll}
\left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=-n}^{n} (T^+ + T^-)_{ij}(k) \right] - \mu & \leq & 0 & i = 1, \ldots, n_{\tilde{z}} \\
\left[\mathcal{E}_n(T^+ - T^-) \right]_{ij}(k) & = & \tilde{H}_{ij}(k) & (i,j) \in \boldsymbol{S}_W, \quad k \in \{-\bar{n}, \ldots, \bar{n}\} \\
\mathcal{T}_I^w \mathcal{E}_n(T^+ - T^-) & = & \mathcal{T}_I^w H & \\
\left[(T^+ - T^-) W_w \right]_{ij}(k) & = & 0 & k \in \{\ldots, -2, -1\} \\
(\mu, T^+, T^-) & \in & \mathbf{R}_+ \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (2n+1)} \times \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (2n+1)} &
\end{array}$$

where the matrices in the convolution constraints have been chosen to be finitely supported. $\bar{n} := n + n_U + n_V$ where n_U and n_V are integers such that

$$\text{supp } \tilde{U} \subset \{0, \ldots, n_U\} \text{ and supp } \tilde{V} \subset \{0, \ldots, n_V\}.$$

$\mathcal{E}_n : \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (2n+1)} \mapsto l_1(\mathbf{Z})^{n_{\tilde{z}} \times n_{\tilde{w}}}$ is the natural embedding operator which centers and pads with zeros on both sides.

There are $2 n_{\tilde{z}} n_{\tilde{w}} (2n+1) + 1$ variables in $\overline{LP}_n^w$: μ and

$$\left\{ T_{ij}^+(k), \ T_{ij}^-(k) : i \in \{1, \ldots, n_{\tilde{z}}\}, \ j \in \{1, \ldots, n_{\tilde{w}}\}, \ k \in \{-n, \ldots, n\} \right\}.$$

Proposition 5.7.2 *If there exists $\tilde{n}$ such that $\overline{\boldsymbol{K}}^{w}_{\tilde{n}} \neq \emptyset$ then there exists n_H such that*

$$\operatorname{supp} \tilde{H}_{ij} \subset \{-n_H, \ldots, n_H\} \quad \forall (i,j) \in \boldsymbol{S}_W$$

and if $n \geq n_H - n_U - n_V$ then $\overline{LP}^{w}_{n}$ is equivalent to $\overline{OPT}^{w}_{n}$ under the map $\overline{\psi}^{w}_{n}$ defined, given a feasible solution (μ, T^+, T^-) for $\overline{LP}^{w}_{n}$,

$$\overline{\psi}^{w}_{n}(\mu, T^+, T^-) := W_z H W_w^{-L} - \mathcal{E}_n(T^+ - T^-).$$

Proof: The first statement follows from Proposition 5.5.2 and the equivalence is shown using $\tilde{\overline{\psi}}^{w}_{n}$ defined, given $K \in \overline{\boldsymbol{K}}^{w}_{n}$,

$$\tilde{\overline{\psi}}^{w}_{n} K := \left(\left\| W_z H W_w^{-L} - K \right\|_{l_1}, \ \Pi_n \Pi_+ (W_z H W_w^{-L} - K), \ \Pi_n \Pi_+ (K - W_z H W_w^{-L}) \right).$$

where $\Pi_n : l_1^{n_{\tilde{z}} \times n_{\tilde{w}}} \mapsto \mathbf{R}_+^{n_{\tilde{z}} \times n_{\tilde{w}} \times (2n+1)}$ is the obvious projection operator. □

The next linear program corresponds to an infinite dimensional super-optimal approximation problem $\underline{OPT}^{w}_{n}$, defined in Section 5.6. The linear program

$\underline{LP}^{w}_{n}$: $\inf \mu$

subject to:

$$\begin{array}{rcll}
\left[\sum_{j=1}^{n_{\tilde{w}}} \sum_{k=-\infty}^{\infty} (T^+ + T^-)_{ij}(k) \right] - \mu & \leq & 0 & i = 1, \ldots, n_{\tilde{z}} \\
\left[\tilde{U}(T^+ - T^-)\tilde{V} \right]_{ij}(k) & = & \tilde{H}_{ij}(k) & (i,j) \in \boldsymbol{S}_W, \ k \in \{0, \ldots, n\} \\
\mathcal{T}_I^{w}(T^+ - T^-) & = & \mathcal{T}_I^{w} H & \\
\left[(T^+ - T^-) W_w \right]_{ij}(k) & = & 0 & k \in \{\ldots, -2, -1\} \\
(\mu, T^+, T^-) & \in & \mathbf{R}_+ \times (l_1^+)^{n_{\tilde{z}} \times n_{\tilde{w}}} \times (l_1^+)^{n_z \times n_w} &
\end{array}$$

Note that $\underline{LP}^{w}_{n}$ is very similar to LP^{w}. The variables are $\mu \in \mathbf{R}$ and

$$\left\{ T_{ij}^+(k), \ T_{ij}^-(k) : i \in \{1, \ldots, n_{\tilde{z}}\}, \ j \in \{1, \ldots, n_{\tilde{w}}\}, \ k \in \mathbf{Z} \right\}.$$

as in LP^{w} and are infinite in number. The constraints are identical as well, except that only a finite number of the convolution constraints are enforced. There are still infinitely many positivity constraints and an infinite set of causality constraints.

Proposition 5.7.3 *$\underline{LP}_n^w$ is equivalent to $\underline{OPT}_n^w$ under the map $\underline{\psi}_n^w$ defined, given a feasible solution (μ, T^+, T^-) for $\underline{LP}_n^w$,*

$$\underline{\psi}_n^w(\mu, T^+, T^-) := W_z H W_w^{-L} - (T^+ - T^-).$$

Proof: $\underline{\psi}_n^w$ is defined identically to ψ^w of Proposition 5.7.1, so use $\tilde{\underline{\psi}}_n^w$ defined, given $K \in \underline{K}_n^w$,

$$\tilde{\underline{\psi}}_n^w K := \left(\left\|W_z H W_w^{-L} - K\right\|_{l_1},\ \Pi_+(W_z H W_w^{-L} - K),\ \Pi_+(K - W_z H W_w^{-L})\right)$$

which is defined identically to $\tilde{\psi}^w$. □

5.8 Discussion

In this chapter the weighted l_∞ disturbance rejection problem has been formulated and a partial solution given. A system gain appropriate to the specification was first defined and it was shown that it can be computed via an optimization similar to a standard l_1 optimization problem, but with the feasible subspace allowed to range over all of $l_1(\mathbf{Z})$. As a consequence, it is not clear that the norm can be computed via finite dimensional optimization except in certain cases. In particular, if $\mathcal{G}$ is approximated by an FIR system and finitely supported weights with finitely supported left inverses are chosen then the norm computation can be approximated finite dimensionally.

The remainder of the results are a subset of those obtained in Chapter 4 for the incremental problem. It is shown that an FIR sub-optimal approximation scheme can be applied for precisely the same generalized plants and weights as in the incremental case. (FIR approximation was the only solution method available in l_1 optimization when it was first introduced as well.) The super-optimal approximation scheme is given in its infinite dimensional form. While it is not shown that such problems have finitely supported optimal solutions as required for a finite dimensional double iteration to produce a true lower bound, such a scheme could nonetheless be implemented. Based on computational experience gained in the corresponding method for the standard l_1 problem, it seems likely that this would succeed.

As in Chapter 4, the most important aspect of the results obtained here is that they represent a substantial step in the direction of a tractable theory appropriate to a very practically appealing design specification. The same references to related work as in Chapter 4 ([13][14]) are the only ones available here. As noted in Chapter 4 the problem addressed there is not well defined; in particular no distinction between weighted and incremental weighted specifications is made.

Chapter 6

Conclusion

A brief summary of the thesis is given in Section 6.1 followed by a discussion of possible future related research in Section 6.2.

6.1 Summary

Three discrete time disturbance rejection design problems for linear time invariant systems aimed at satisfying l_∞, or peak magnitude, specifications have been considered. All were posed in the very general problem setting which has become standard in control design. This setting was carefully defined in Chapter 2 such that *all* suddenly applied signals can be considered as potential disturbances, and such that all compensators obtained by the given design methods are implementable (i.e., causal) or can be approximated arbitrarily closely in terms of performance by implementable compensators.

Each design problem was solved through reduction, by turns, to the minimization of an appropriate norm of the closed loop system over stabilizing compensators, the solution of a minimum distance problem in l_1, and the solution of an infinite linear program. For each design problem, it was shown, in the minimum distance problem setting, how to obtain converging upper and lower bounds on the optimal performance of any stabilizing compensator by solving sequences of optimization problems, and *finite* linear programs equivalent to these were formulated (the sole exception is for the lower bound problems corresponding to the weighted specification of Chapter 5; these are left in infinite dimensional form). The upper bound optimizations are not always feasible but when they are they yield a sequence of stabilizing compensators whose performances approach the optimum. The lower bound optimizations are always feasible and a sequence of stabilizing compensators can be obtained from them as well. Although the performance of this sequence is not guaranteed to converge to the optimum, it often does in examples and often yields compensators of much lower order than the upper bound method for the same performance level.

The unweighted l_∞ disturbance rejection problem, or l_1 problem, considered in Chapter 3 has been the subject of considerable previous research and the key aspects of the solution discovered. The results here encompass all major results previously obtained, presenting them under minimal assumptions. Care has also been taken to provide complete details of all proofs, problem formulations, and equivalences of optimization problems. It is hoped that these will both help to give insight into the solution of unresolved theoretical problems (see Section 6.2 for a few of these) and facilitate the development of software tools for the problem.

The value of some of these details is evidenced by the solution in Chapters 4 and 5 of the incremental weighted and the weighted l_∞ disturbance rejection problems. In each of these problems it was necessary first to define an appropriate norm of the closed loop system to be minimized. In each case the norm computation is done via an l_1 (or l_1-like) optimization problem; the key elements in establishing these computations depended in each case on the detailed structure of the feasible subspace for each and its close relationship to that of the l_1 problem.

The incremental weighted and weighted problems are important because they provide the designer the ability to reflect his actual requirements more precisely in their specifications, which can potentially help to alleviate the problem of conservatism inherent in norm-based design for disturbance rejection. In particular, the ability to bound rates and accelerations is practically appealing.

Although the results obtained for these problems were not quite as extensive as those for the unweighted problem, they were nearly so and provide all the information necessary to establish complete design methods. Moreover, they share with the unweighted problem the fact that the ultimate computations required are finite linear programs.

6.2 Suggestions for Further Research

This concluding section enumerates a few possible directions for further research suggested by the results seen so far. These will be grouped according to the design problem they concern. Because of the similarity of the motivation, theory and computational methods of each problem, most questions raised with respect to one problem apply equally to the subsequent problems. With respect to each of the problems solved there are both theoretical and more application-oriented investigations which could be pursued; both will be addressed.

Theory

l_∞ Disturbance Rejection

This is the oldest of the three problems and the most is known about it. Nonetheless, many gaps remain both in the theory and its practical applicability. The theory seems to stand at a crossroads, and a fresh approach may be needed in order to advance it significantly; as it is presently viewed, no general mathematical tools are available. As a consequence, results are difficult to come by, the proofs tedious, and relatively little insight gained with respect to broader questions. One potential direction to pursue, both from the point of view of theory and of computation, is to exploit any available results concerning semi-infinite linear programs, as the super-optimal approximation problems are all equivalent to these. In particular, general results concerning finitely supported minimizers would be invaluable in simplifying the present theory.

More specific near term issues to be resolved include

- When does the performance of a sequence of sub-optimal compensators obtained from super-optimal approximation problems converge to the optimum?
- When (and precisely how) can exactly optimal compensators be constructed from super-optimal approximation problems?
- When are optimal compensators unique and, when they are not, how can they be parametrized (perhaps using results from the theory of linear programming)?

Incremental Weighted l_∞ Disturbance Rejection

The results here have all the same deficiencies as those for the unweighted problem; however there are few additional questions concerning the design methodology and computation due to its similarity to the unweighted case. Of greater interest is to investigate the properties of weights and to gain experience with the design method, both from a computational and an applications viewpoint. In particular

- Not all choices of weights lead to problems distinct from standard l_1 problems, i.e., cascade weights as described in Chapter 1 can sometimes be found which produce the same class of possible disturbances. (A trivial example is $\hat{W}_w = [\begin{array}{cc} 1 & 0 \end{array}]^T$; it is obviously equivalent to no weight at all.) Since standard l_1 problems are more easily solved, it is of interest to obtain a characterization of weights which can be replaced by cascade equivalents.

- There is a dual specification in which both weights are treated as the disturbance weight is in the cascade approach, i.e., both possible disturbances and acceptable errors are treated as having resulted from signals in the unit ball of l_∞ and then passed through a weight. No practical interpretation has been suggested for such a specification but, if one exists, the results of Chapter 4 are readily dualized to obtain a solution.

- Some weights which are unstable and/or not stably invertible can be treated in this framework as well; they lead to tracking-type problems with a norm criterion in addition. The details of the theory and design methods require further investigation, as does the appropriate interpretation of such problems.

- Computational experience with the design methods is necessary both to determine whether there are any unforeseen difficulties with the computations and to improve theoretical insight.

- Experience in applications is necessary to determine whether weighted specifications can in fact capture the objectives of designers. In particular, the incremental nature of the specification is a unique feature which may make it more or less practically useful; in the final analysis designed compensators must be implemented and tested to determine this.

Weighted l_∞ Disturbance Rejection

Additional questions exist concerning the weight problem, in particular:

- Some theoretical issues remain unresolved with respect to super-optimal approximations. While it is likely that they can be computed in a fashion similar to the incremental case, this remains to be worked out.

- Just as some incremental weighted specifications are equivalent to cascade specifications, so are some weighted specifications. Moreover, there is a question of equivalence between weighted and incremental weighted specifications; they have shown to be different for certain weights and they are identical for others (such as the trivial example cited above). If a useful class of weights can be found for which they are equivalent then the simpler incremental design methods can be used.

Robustness

One of the natural advantages shared by theories based on minimization of a system norm is that certain robustness problems can be easily solved. In particular, if the nominal generalized plant is poorly known, it is modelled as containing "perturbation" subsystems which are unknown, but of bounded norm.

For both the $\boldsymbol{H}_\infty$ and the l_1 norms, necessary and sufficient conditions for robust stability with respect to a single such perturbation are found in terms of the respective norms. Moreover, robust performance problems in both settings can be reduced to robust stability problems for multiple perturbations. Because the methods here are norm-based and in particular because of their similarity to standard l_1 problems it is reasonable to conjecture that similar results can be achieved.

Presently the class of possible perturbations is modelled in terms of a cascade norm, which has no natural interpretation in the l_1 setting. A natural interpretation in terms of weighted or incremental weighted norm is not readily apparent either but, at a minimum, it would increase the latitude of the designer to specify perturbation classes.

The same questions arise in this setting as to which weighted specification is more natural. The answer from a practical standpoint is no more clear, but there may be a natural choice from a theoretical standpoint. One or the other norm (or both) may lead to a nice robustness theory. A possible advantage goes to the weighted norm because it is an induced norm, a property it shares with the $\boldsymbol{H}_\infty$ and l_1 norms.

Continuous Time/Sampled Data

Discrete time systems are necessarily approximations to reality, so it is natural to be interested in continuous time versions of the problems considered, i.e., L_∞ specifications and, in the unweighted case, the L_1 problem. While continuous and discrete time is more or less equally tractable in $\boldsymbol{H}_\infty$, this is far from the case in L_1. Optimal compensators have been shown to be necessarily infinite dimensional [8], and only recently have finite dimensional approximation schemes begun to become available [35]. It seems unlikely, therefore, that much progress can be made in terms of weighted continuous time specifications.

The most realistic control system design problem given current technology is probably the sampled data problem, in which the plant is modelled as continuous time and the compensator as discrete time (i.e., a digital system). Complete design methods have recently been obtained for the sampled data L_1 problem [11], and it is natural to investigate how weighted specifications can be incorporated in this setting. If this is possible, it will allow the designer to bound actual velocities and accelerations and not just discrete time approximations to them.

Appendix A

Rings and Fields of Sequences

This appendix contains supporting definitions, facts, and results (primarily) for Chapter 2. Algebraic terms not defined are meant in their most standard sense. The principal source of definitions and facts is [17].

Section A.1 defines some terms and some classes of rings which are significant for algebraic system theory. Coprimeness and coprime factorizations are defined, and there is a simple proposition concerning them.

Section A.2 establishes that the sequence spaces l_+ and l are related rings under convolution defined on two-sided and one-sided sequences, respectively.

Throughout the section, $\boldsymbol{R}$ is a ring and $\boldsymbol{F}$ is a field. If $\boldsymbol{R}$ is commutative and $x, y \in \boldsymbol{R}$, $x|y$ if there exists $q \in \boldsymbol{R}$ such that $y = qx$. If $\boldsymbol{R}$ is a ring with identity, an element $x \in \boldsymbol{R}$ is a *unit* of $\boldsymbol{R}$ if there is another element $x^{-1} \in \boldsymbol{R}$ such that $xx^{-1} = x^{-1}x = 1$. The set of all units in $\boldsymbol{R}$ is denoted by $\boldsymbol{U}(\boldsymbol{R})$. If $\boldsymbol{R}$ is a commutative domain with identity, $\boldsymbol{F}_R$ denotes the field of fractions corresponding to $\boldsymbol{R}$.

A.1 Algebraic Preliminaries

Definition A.1.1 *A commutative domain with identity $\boldsymbol{R}$ is said to be*

- Euclidean *if there exists a degree function* $\deg : \boldsymbol{R} \setminus \{0\} \mapsto \mathbf{Z}_+$ *satisfying, for all $x, y \in \boldsymbol{R}$:*

  1. $x \neq 0 \Rightarrow \exists q \in \boldsymbol{R}$ *such that either* $r := y - qx = 0$ *or* $\deg(r) < \deg(x)$

2. $x|y \Rightarrow \deg(x) \leq \deg(y)$

and proper Euclidean *if, in addition, $\boldsymbol{R}$ is not a field and* $\deg(xy) = \deg x + \deg y$ *for all $x, y \in \boldsymbol{R}$.*

- principal ideal *if every ideal in $\boldsymbol{R}$ is principal.*
- Bezout *if every finitely generated ideal in $\boldsymbol{R}$ is principal.*
- Hermite *if every left-invertible matrix over $\boldsymbol{R}$ can be complemented (i.e., $M \in \boldsymbol{R}^{m\times n}$ left-invertible in $\boldsymbol{R} \Rightarrow \exists M^c \in \boldsymbol{R}^{m\times(m-n)}$ such that* $\det \begin{bmatrix} M & M^c \end{bmatrix} \in U(\boldsymbol{R})$*).*

Fact A.1.2 *If M and $\tilde{M}$ are matrices over a Hermite domain $\boldsymbol{R}$ with identity and are left- and right-invertible, respectively, then* Bezout equations

$$\begin{bmatrix} M^{-L} \\ M^{\perp} \end{bmatrix} \begin{bmatrix} M & M^c \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \qquad \begin{bmatrix} \tilde{M} \\ \tilde{M}^c \end{bmatrix} \begin{bmatrix} \tilde{M}^{-R} & \tilde{M}^{\perp} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

can be constructed, where all additional symbols denote matrices over $\boldsymbol{R}$.

Fact A.1.3 *The following inclusions hold:*

- *Every Euclidean domain is principal ideal.*
- *Every principal ideal domain is Bezout.*
- *Every Bezout domain is Hermite.*

Definition A.1.4 *Let $\boldsymbol{R}$ be a domain with identity.*

- *Two matrices N and M are* right coprime over $\boldsymbol{R}$ *if they have the same number of columns and there exist matrices $\tilde{X}$ and $\tilde{Y}$ over $\boldsymbol{R}$ such that $\tilde{X}M + \tilde{Y}N = I$.*
- *Two matrices $\tilde{N}$ and $\tilde{M}$ are* left coprime over $\boldsymbol{R}$ *if they have the same number of rows and there exist matrices X and Y over $\boldsymbol{R}$ such that $\tilde{M}X + \tilde{N}Y = I$.*

Definition A.1.5 *Let $\boldsymbol{R}$ be a commutative domain with identity and let $G \in \boldsymbol{G}$, where $\boldsymbol{G}$ is a ring containing $\boldsymbol{R}$.*

- *If* $N \in \boldsymbol{R}^{m\times n}$ *and* $M \in \boldsymbol{R}^{n\times n}$ *are right coprime over* $\boldsymbol{R}$, $\det M \neq 0$, *and* $G = NM^{-1}$, *then* $G = NM^{-1}$ *is a* right coprime factorization of G over $\boldsymbol{R}$.
- *If* $\tilde{N} \in \boldsymbol{R}^{m\times n}$ *and* $\tilde{M} \in \boldsymbol{R}^{m\times m}$ *are left coprime over* $\boldsymbol{R}$, $\det \tilde{M} \neq 0$, *and* $G = \tilde{M}^{-1}\tilde{N}$, *then* $G = \tilde{M}^{-1}\tilde{N}$ *is a* left coprime factorization of G over $\boldsymbol{R}$.

An arbitrary matrix over a ring $\boldsymbol{G}$ need not have either a left or a right coprime factorization. However, every matrix over a Bezout ring has both, and every matrix over a Hermite domain has both if it has one.

Proposition A.1.6 *Let* $\boldsymbol{R}$ *be a commutative domain with identity and let* $G \in \boldsymbol{G}$, *where* $\boldsymbol{G}$ *is a ring containing* $\boldsymbol{R}$.

- *If* $G = NM^{-1}$ *is a right coprime factorization of* G *over* $\boldsymbol{R}$ *then*

$$G \in \boldsymbol{R} \iff M^{-1} \in \boldsymbol{R}$$

- *If* $G = \tilde{M}^{-1}\tilde{N}$ *is a left coprime factorization of* G *over* $\boldsymbol{R}$ *then*

$$G \in \boldsymbol{R} \iff \tilde{M}^{-1} \in \boldsymbol{R}$$

Proof: Only the first item is proven; the proof of the second is entirely similar. If $M^{-1} \in \boldsymbol{R}$ then clearly $G = NM^{-1} \in \boldsymbol{R}$. For the converse, note that if $G = NM^{-1}$ is a right coprime factorization over $\boldsymbol{R}$ then by Definition A.1.5 N and M are right coprime and by Definition A.1.4 there exist matrices $\tilde{X}$ and $\tilde{Y}$ over $\boldsymbol{R}$ such that $\tilde{X}M + \tilde{Y}N = I$. M is non-singular and hence, multiplying on the right by its inverse, $\tilde{X} + \tilde{Y}NM^{-1} = \tilde{X} + \tilde{Y}G = M^{-1}$. Thus if $G \in \boldsymbol{R}$ then so is M^{-1}. □

A.2 Rings and Fields of Sequences

Several sequence spaces of interest in Chapter 2 have useful algebraic structures under pointwise addition and convolution.

Proposition A.2.1 l_+ *is a field when addition and multiplication are defined, given* $x_1, x_2 \in l_+$,

$$\begin{aligned} x_1 + x_2 &:= \{x_1(k) + x_2(k)\}_{k\in\mathbf{Z}_+} \\ x_1x_2 &:= \left\{\sum_{n=-\infty}^{\infty} x_1(n)x_2(k-n)\right\}_{k\in\mathbf{Z}} \end{aligned}$$

The additive identity is $\{0\}_{k\in\mathbf{Z}_+}$ *and the multiplicative identity is* δ *(defined in Chapter 2). Given* $x \in l_+$, $x \neq 0$, *the multiplicative inverse of* x, *denoted* x^{-1}, *is defined recursively*

$$x^{-1}(k) = \begin{cases} 0 & k < -k_x \\ \frac{1}{x(k_x)} & k = -k_x \\ -\frac{1}{x(k_x)} \sum_{n=-k_x}^{-k_x+k-1} x^{-1}(n)x(k-n) & k > -k_x \end{cases}$$

where k_x *is defined*

$$k_x := \max\{k \in \mathbf{Z} : \text{supp } x \subset \{k, k+1, \ldots\}\}$$

Proof: It is easy to check that l_+ is a commutative ring under the defined operations, that the additive and multiplicative identities are as claimed, and that the inverse of every non-zero element is as claimed. There are certainly more than two elements in l_+ so that it is a field. □

Fact A.2.2 *l is a commutative domain with identity when addition and multiplication are defined, given* $x_1, x_2 \in l$:

$$\begin{aligned} x_1 + x_2 &:= \{x_1(k) + x_2(k)\}_{k\in\mathbf{Z}_+} \\ x_1 x_2 &:= \left\{\sum_{n=0}^{k} x_1(n)x_2(k-n)\right\}_{k\in\mathbf{Z}_+} \end{aligned}$$

The additive identity is $\{0\}_{k\in\mathbf{Z}_+}$ *and the multiplicative identity is* δ *(defined on* $\mathbf{Z}_+$*).*

Proposition A.2.3 *The ring l is isomorphic to a subring of the field l_+ containing the identity under the map* $\phi : l \mapsto l_+$ *defined, given* $x \in l$,

$$(\phi x)(k) := \begin{cases} 0 & k < 0 \\ x(k) & k \geq 0 \end{cases}$$

Proof: ϕ is clearly well defined and maps l into l_{1+}. It is also one-to-one; in fact, its inverse is given by

$$\phi^{-1}y := \{y(k)\}_{k\in\mathbf{Z}_+} \tag{A.1}$$

for all $y = \phi x$, $x \in l$. Moreover, it is easy to check that ϕ is a homomorphism. Hence, ϕ is an isomorphism mapping l to the subring ϕl of l_+. l contains the identity since $\phi^{-1}\delta = \delta$ (defined on $\mathbf{Z}_+$). □

Proposition A.2.4 *The commutative domain with identity l is proper Euclidean with degree function Δ_l defined, given $x \in l$:*

$$\Delta_l(x) := \max \{k \in \mathbf{Z}_+ : \operatorname{supp} x \subset \{k, k+1, \ldots\}\}$$

Proof: Fact A.2.2 shows that l is a commutative domain. We first establish that the degree function meets the conditions required by Definition A.1.1.

Given $x, y \in l$, $x \neq 0$, define $q \in l_+$ as follows:

$$q := \begin{cases} \phi^{-1}\left[(\phi y)(\phi x)^{-1}\right] & \Delta_l(x) \leq \Delta_l(y) \\ 0 & \Delta_l(x) > \Delta_l(y) \end{cases}$$

where ϕ is defined in Proposition A.2.3, ϕ^{-1} in (A.1), and $(\phi x)^{-1}$ is the multiplicative inverse of ϕx in l_+, defined in Proposition A.2.1. It is easy to check that $(\phi y)(\phi x)^{-1}$ is always in ϕl if $\Delta_l(x) \leq \Delta_l(y)$ so that the above definition of q makes sense. Moreover, if a remainder r is defined $r := y - qx$, $\Delta_l(x) \leq \Delta_l(y) \Rightarrow r = y - \phi^{-1}[(\phi y)(\phi x)^{-1}]x = 0$ (i.e., x divides y), and $\Delta_l(x) > \Delta_l(y) \Rightarrow r = y \Rightarrow \Delta_l(r) = \Delta_l(y) < \Delta_l(x)$. Moreover, $\Delta_l(xy) = \Delta_l(x) + \Delta_l(y)$ $\forall x, y \in l$ for which both sides are defined (this follows easily from the definition of convolution).

To complete the proof, note that l is not a field (for this it is sufficient that there exist elements, e.g. any $x \neq 0$ with $x(0) = 0$, with non-zero degree). □

Corollary A.2.5 *The set $U(l)$ of units in l is given by*

$$U(l) = \{x \in l : x(0) \neq 0\}.$$

Proof: It is easy to see that in a Euclidean domain the units are precisely the elements of zero degree. The conclusion then follows since l is Euclidean by Proposition A.2.4 with degree function Δ_l and if $x \in l$ then $\Delta_l(x) = 0 \iff x(0) \neq 0$. □

The proof of the following proposition is straightforward, and is omitted.

Proposition A.2.6 *The field of fractions F_l corresponding to l is isomorphic to l_+ under the map $\phi_f : F_l \mapsto l_+$ defined, given $n, d \in l$ with $d \neq \{0\}_{k \in \mathbf{Z}_+}$,*

$$\phi_f\left(\frac{n}{d}\right) := (\phi n)(\phi d)^{-1}$$

where ϕ is defined in Proposition A.2.3 and $(\phi d)^{-1}$ is the multiplicative inverse of ϕd in l_+, defined in Proposition A.2.1.

Fact A.2.7 *l_1 is a subdomain of l containing the identity and is Hermite.*

Appendix B

Duality and Minimum Distance Problems

This appendix contains supporting notation, facts, and results for Chapters 3 (in particular), 4, and 5. The facts and results are mainly concerned with duality and the solution of minimum distance problems in general normed linear spaces; the principal reference is [36].

Section B.1 concerns duality and adjoint operators in product spaces, and Section B.2 contains the basic facts on duality and its relation to minimum distance problems. First, some standard notation is defined.

Notation

In the following definitions $\boldsymbol{X}$ and $\boldsymbol{Y}$ are real normed linear spaces.

$\boldsymbol{BX}$ the unit ball in $\boldsymbol{X}$.

$\boldsymbol{X}^*$ the dual space of $\boldsymbol{X}$.

$\langle x, x^* \rangle$ the linear functional $x^* \in \boldsymbol{X}^*$ evaluated on $x \in \boldsymbol{X}$.

$\boldsymbol{M}^\perp$ the right annihilator of the subspace $\boldsymbol{M}$ of $\boldsymbol{X}$ (a subspace of $\boldsymbol{X}^*$).

${}^\perp\boldsymbol{M}$ the left annihilator of the subspace $\boldsymbol{M}$ of $\boldsymbol{X}^*$ (a subspace of $\boldsymbol{X}$).

$\mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ the space of bounded linear operators from $\boldsymbol{X}$ into $\boldsymbol{Y}$.

$\mathcal{B}(\boldsymbol{X})$ the space of bounded linear operators from $\boldsymbol{X}$ into $\boldsymbol{X}$.

$\mathcal{T}^*$ the adjoint of $\mathcal{T} \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Y})$ ($\mathcal{T}^* \in \mathcal{B}(\boldsymbol{Y}^*, \boldsymbol{X}^*)$).

B.1 Duality and Adjoints in Product Spaces

Fact B.1.1 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be real normed linear spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and let $\|\cdot\|_p$, $1 \leq p \leq \infty$, denote the usual p-norm on $\mathbf{R}^2$.*

- *With addition and scalar multiplication defined componentwise, $\boldsymbol{X} \times \boldsymbol{Y}$ is a real normed linear space under the norm $\|\cdot\|_{X \times Y}$ defined, given $(x, y) \in \boldsymbol{X} \times \boldsymbol{Y}$,*
$$\|(x,y)\|_{X\times Y} := \|(\|x\|_X, \|y\|_Y)\|_p$$
- *$(\boldsymbol{X} \times \boldsymbol{Y})^* = \boldsymbol{X}^* \times \boldsymbol{Y}^*$ with addition and scalar multiplication defined componentwise, under the norm $\|\cdot\|_{X^* \times Y^*}$ defined, given $(x^*, y^*) \in \boldsymbol{X}^* \times \boldsymbol{Y}^*$,*
$$\|(x^*,y^*)\|_{X^*\times Y^*} := \|(\|x^*\|_{X^*}, \|y^*\|_{Y^*})\|_q$$
where $\frac{1}{p} + \frac{1}{q} = 1$ if $1 < p < \infty$, $q = \infty$ if $p = 1$, and $q = 1$ if $p = \infty$, and with linear functional evaluation defined, given $(x, y) \in \boldsymbol{X} \times \boldsymbol{Y}$ and $(x^*, y^*) \in \boldsymbol{X}^* \times \boldsymbol{Y}^*$,
$$\langle (x,y), (x^*,y^*) \rangle := \langle x, y \rangle + \langle x^*, y^* \rangle .$$

Proposition B.1.2 *Let $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ be real normed linear spaces and define the real normed linear space $\boldsymbol{X} \times \boldsymbol{Y}$ as in Fact B.1.1. Define $\mathcal{T} : \boldsymbol{X} \times \boldsymbol{Y} \mapsto \boldsymbol{Z}$, given $(x, y) \in \boldsymbol{X} \times \boldsymbol{Y}$, $\mathcal{T}[(x,y)] := \mathcal{T}_X x + \mathcal{T}_Y y$, where $\mathcal{T}_X \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Z})$ and $\mathcal{T}_Y \in \mathcal{B}(\boldsymbol{Y}, \boldsymbol{Z})$.*

- *$\mathcal{T} \in \mathcal{B}(\boldsymbol{X} \times \boldsymbol{Y}, \boldsymbol{Z})$.*
- *$\mathcal{T}^*$ can be written, given $z^* \in \boldsymbol{Z}^*$, $\mathcal{T}^* z^* = (\mathcal{T}_X^* z^*, \mathcal{T}_Y^* z^*)$.*

Proof: For the first item, recall that all norms, in particular the p-norms, on $\mathbf{R}^2$ are equivalent. Thus, $\exists \alpha_p > 0$ such that $\|x\|_p \geq \alpha_p \|x\|_1$ for all $x \in \mathbf{R}^2$. For this choice of α_p and for any $(x, y) \in \boldsymbol{X} \times \boldsymbol{Y}$,
$$\begin{aligned}
\frac{\|\mathcal{T}[(x,y)]\|_Z}{\|(x,y)\|_{X\times Y}} &= \frac{\|\mathcal{T}_X x + \mathcal{T}_Y y\|_Z}{\|(\|x\|_X, \|y\|_Y)\|_p} \\
&\leq \frac{\|\mathcal{T}_X\| \, \|x\|_X + \|\mathcal{T}_Y\| \, \|y\|_Y}{\alpha_p (\|x\|_X + \|y\|_Y)} \leq \frac{\max\{\|\mathcal{T}_X\|, \|\mathcal{T}_Y\|\}}{\alpha_p} \\
&< \infty.
\end{aligned}$$

For the second item, note that for all $(x,y) \in \boldsymbol{X} \times \boldsymbol{Y}$ and $z^* \in \boldsymbol{Z}^*$,

$$\begin{aligned}
\langle (x,y), \mathcal{T}^* z^* \rangle &= \langle \mathcal{T}\left[(x,y)\right], z^* \rangle \\
&= \langle \mathcal{T}_X x + \mathcal{T}_Y y, z^* \rangle = \langle \mathcal{T}_X x, z^* \rangle + \langle \mathcal{T}_Y y, z^* \rangle = \langle x, \mathcal{T}_X^* z^* \rangle + \langle y, \mathcal{T}_Y^* z^* \rangle \\
&= \langle (x,y), (\mathcal{T}_X^* z^*, \mathcal{T}_Y^* z^*) \rangle .
\end{aligned}$$

where the last equality follows from the definition of linear functional evaluation in Fact B.1.1. Thus it must be that $\mathcal{T}^* z^* = (\mathcal{T}_X^* z^*, \mathcal{T}_Y^* z^*)$ for every $z^* \in \boldsymbol{Z}^*$. □

Proposition B.1.3 *Let $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ be real normed linear spaces and define the real normed linear space $\boldsymbol{X} \times \boldsymbol{Y}$ as in Fact B.1.1 above. Define $\mathcal{T} : \boldsymbol{Z} \mapsto \boldsymbol{X} \times \boldsymbol{Y}$, given $z \in \boldsymbol{Z}$, $\mathcal{T} z := (\mathcal{T}_X x, \mathcal{T}_Y y)$, where $\mathcal{T}_X \in \mathcal{B}(\boldsymbol{X}, \boldsymbol{Z})$ and $\mathcal{T}_Y \in \mathcal{B}(\boldsymbol{Y}, \boldsymbol{Z})$.*

- $\mathcal{T} \in \mathcal{B}(\boldsymbol{Z}, \boldsymbol{X} \times \boldsymbol{Y})$.
- *$\mathcal{T}^*$ can be written, given $(x^*, y^*) \in \boldsymbol{X}^* \times \boldsymbol{Y}^*$, $\mathcal{T}^*\left[(x^*, y^*)\right] = \mathcal{T}_X^* x^* + \mathcal{T}_Y^* y^*)$.*

Proof: For the first item, recall that all norms, in particular the p-norms, on $\mathbf{R}^2$ are equivalent. Thus, $\exists \beta_p > 0$ such that, for all $x \in \mathbf{R}^2$, $\|x\|_p \le \beta_p \|x\|_1$. For this choice of β_p and for any $z \in \boldsymbol{Z}$,

$$\begin{aligned}
\frac{\|\mathcal{T} z\|_{X \times Y}}{\|z\|_Z} &= \frac{\left\|\left(\|\mathcal{T}_X z\|_X, \|\mathcal{T}_Y z\|_Y\right)\right\|_p}{\|z\|_Z} \\
&\le \frac{\beta_p \left(\|\mathcal{T}_X z\|_X + \|\mathcal{T}_Y z\|_Y\right)}{\|z\|_Z} \le \beta_p \left(\|\mathcal{T}_X\| + \|\mathcal{T}_Y\|\right) \\
&< \infty.
\end{aligned}$$

For the second item, note that for all $z \in \boldsymbol{Z}$ and $(x^*, y^*) \in \boldsymbol{X}^* \times \boldsymbol{Y}^*$,

$$\begin{aligned}
\langle z, \mathcal{T}^*\left[(x^*, y^*)\right] \rangle &= \langle \mathcal{T} z, (x^*, y^*) \rangle = \langle (\mathcal{T}_X z, \mathcal{T}_Y z), (x^*, y^*) \rangle \\
&= \langle \mathcal{T}_X z, x^* \rangle + \langle \mathcal{T}_Y z, y^* \rangle = \langle z, \mathcal{T}_X^* x^* \rangle + \langle z, \mathcal{T}_Y^* y^* \rangle \\
&= \langle z, \mathcal{T}_X^* x^* + \mathcal{T}_Y^* y^* \rangle .
\end{aligned}$$

where the third equality follows from the second item of Fact B.1.1. Thus it must be that $\mathcal{T}^*\left[(x^*, y^*)\right] = \mathcal{T}_X^* x^* + \mathcal{T}_Y^* y^*$ for every $(x^*, y^*) \in \boldsymbol{X}^* \times \boldsymbol{Y}^*$. □

B.2 Minimum Distance Problems

Definition B.2.1, Theorem B.2.2, and Theorem B.2.3 can be found, respectively, in [36, pp. 116, 119, 121].

Definition B.2.1 *If $\boldsymbol{X}$ is a real normed linear space then a pair of elements $x \in \boldsymbol{X}$ and $x^* \in \boldsymbol{X}^*$ are* aligned *if $\langle x, x^* \rangle = \|x\| \, \|x^*\|$.*

Theorem B.2.2 *If $\boldsymbol{X}$ is a real normed linear space, $\boldsymbol{M}$ is a subspace of $\boldsymbol{X}$, and $x \in \boldsymbol{X}$ then*

$$\inf \{\|x - m\| : m \in \boldsymbol{M}\} = \sup \left\{\langle x, x^* \rangle : x^* \in \boldsymbol{B}\boldsymbol{M}^{\perp}\right\}$$

where the supremum on the right is achieved for some $x_0^* \in \boldsymbol{B}\boldsymbol{M}^{\perp}$ with $\|x_0^*\| = 1$. If the infimum on the left is achieved for some $m_0 \in \boldsymbol{M}$ then $x - m_0$ and x_0^* are aligned.

Theorem B.2.3 *If $\boldsymbol{X}$ is a real normed linear space, $\boldsymbol{M}$ is a subspace of $\boldsymbol{X}$, and $x^* \in \boldsymbol{X}^*$ then*

$$\inf \left\{\|x^* - m^*\| : m^* \in \boldsymbol{M}^{\perp}\right\} = \sup \{\langle x, x^* \rangle : x \in \boldsymbol{B}\boldsymbol{M}\}$$

where the infimum on the left is achieved for some $m_o^* \in \boldsymbol{M}^{\perp}$. If the supremum on the right is achieved for some $x_0 \in \boldsymbol{M}$ then $x^* - m_0^*$ and x_0 are aligned.

Corollary B.2.4 *If $\boldsymbol{X}$ is a real normed linear space, $\boldsymbol{M}^*$ is weak*-closed a subspace of $\boldsymbol{X}^*$, and $x^* \in \boldsymbol{X}^*$ then there exists an element $m_0^* \in \boldsymbol{M}^*$ such that*

$$\|x^* - m_0^*\| = \inf \{\|x^* - m^*\| : m^* \in \boldsymbol{M}^*\}.$$

Proof: If $\boldsymbol{M}^*$ is a *weak**-closed, then $\boldsymbol{M}^* = [{}^{\perp}\boldsymbol{M}^*]^{\perp}$ [26, Theorem 4.7]. Thus Theorem B.2.3 applies to show existence of m_0^*. □

Corollary B.2.4 provides only a sufficient condition for existence of an element in $\boldsymbol{M}^*$ of infimal distance from x_0^*, but it is in fact precisely as powerful as Theorem B.2.3. This is because, for every subspace $\boldsymbol{M}$ of a normed linear space $\boldsymbol{X}$, $\boldsymbol{M}^{\perp}$ is a *weak**-closed subspace of $\boldsymbol{X}^*$[26, p. 91].

Lemma B.2.5 *If $\boldsymbol{X}$ is a real normed linear space, $\boldsymbol{M}$ is a subspace of $\boldsymbol{X}$, $x \in \boldsymbol{X}$, and $m_0 \in \boldsymbol{M}$ then*

$$\inf \{\|x - m\| : m \in \boldsymbol{M}\} = \inf \{\|(x - m_0) - m\| : m \in \boldsymbol{M}\}.$$

Proof: Let $d := \inf\{\|x-m\| : m \in M\}$ and $d_0 := \inf\{\|(x-m_0)-m\| : m \in M\}$. If $m \in M$ then $m - m_0 \in M$ since M is a subspace, and

$$\|(x-m_0)-(m-m_0)\| = \|x-m\|.$$

Hence $d_0 \leq d$. Conversely, if $m \in M$ then $m + m_0 \in M$ and

$$\|x-(m+m_0)\| = \|(x-m_0)-m\|.$$

Hence $d \leq d_0$. □

Appendix C

Duality and Adjoints in Sequence Spaces

This appendix contains supporting notation, facts, and results for Chapters 3, 4, and 5. The facts and results are mainly concerned with duality in spaces of matrices over sequences, operators on those spaces, and their adjoints.

Section C.1 considers matrices over sequences on $\mathbf{Z}_+$, Section C.2 considers sequences on $\mathbf{Z}$, as required for the results of Chapter 5. First, some convenient notation is introduced:

If m and n are positive integers, $\boldsymbol{S}$ is a subset of $\{1, \ldots, m\} \times \{1, \ldots, n\}$, and $\boldsymbol{X}$ is any set then $\boldsymbol{X}^S$ denotes the set of matrices X whose elements $X_{ij} \in \boldsymbol{X}$ for $(i, j) \in S$ and whose other elements are undefined. $\boldsymbol{S}_I$ denotes the set of $i \in \{1, \ldots, m\}$ such that $(i, j) \in \boldsymbol{S}$ for some $j \in \{1, \ldots, n\}$.

C.1 Spaces of Sequences on $\mathbf{Z}_+$

l_1, l_∞, and c_0 denote the classical sequence spaces with their usual norms; recall that c_0 is a subspace of l_∞ defined

$$c_0 := \{x \in l_\infty : \lim_{k \to \infty} x(k) = 0\}.$$

The following is a simple generalization of the usual matrix versions of these spaces and duality relations among them, which can be found in [5], for example.

Fact C.1.1 *With addition and scalar multiplication defined componentwise,*

1. *the following all define real normed linear spaces:*

   (a) l_1^S *with the norm* $\|X\|_{l_1} := \max_{i \in S_I} \sum_{(i,j) \in S} \|X_{ij}\|_{l_1}$

   (b) l_∞^S *with the norm* $\|X\|_{l_\infty} := \sum_{i \in S_I} \max_{(i,j) \in S} \|X_{ij}\|_{l_\infty}$

   (c) *the subspace* c_0^S *of* l_∞^S

2. *the duality relations*

   (a) $(c_0^S)^* = l_1^S$,

   (b) $(l_1^S)^* = l_\infty^S$,

   hold when linear functional evaluation is defined, given X *in a primal space and* X^* *in its dual,*

$$\langle X, X^* \rangle := \sum_{(i,j) \in S} \left\langle X_{ij}, X_{ij}^* \right\rangle$$

Definition C.1.2 *The binary operations* $\triangleleft$ *of* left correlation *and* $\triangleright$ *of* right correlation *are defined, given* $f, g \in l$,

$$f \triangleleft g := \left\{ \sum_{n=0}^{\infty} f(n) g(n+k) \right\}_{k \in \mathbf{Z}_+} =: g \triangleright f.$$

Proposition C.1.3 *Given* $f \in l_1$, *define the* correlation operator $\bar{\mathcal{F}}$ *on* l_∞, *given* $x \in l_\infty$,

$$\bar{\mathcal{F}} x := f \triangleleft x = x \triangleright f$$

and let $\tilde{\mathcal{F}}$ *denote its restriction to* c_0. *Define also the* convolution operator $\mathcal{F} \in \mathcal{B}(l_1)$, *given* $x \in l_1$,

$$\mathcal{F} x := f * x$$

With these definitions, $\bar{\mathcal{F}} \in \mathcal{B}(l_\infty)$, $\tilde{\mathcal{F}} \in \mathcal{B}(c_0)$, *and*

1. $\mathcal{F}^* = \bar{\mathcal{F}}$

2. $\tilde{\mathcal{F}}^* = \mathcal{F}$

with linear functional evaluation defined as in Fact C.1.1.

Proof: It is easy to see that if $\bar{\mathcal{F}}$ is well defined on l_∞ then it is also linear. The following, where $x \in l_\infty$ is arbitrary, shows that $\bar{\mathcal{F}}$ is both well defined and bounded on l_∞:

$$\begin{aligned}
\left\|\bar{\mathcal{F}}x\right\|_{l_\infty} &= \sup_k \left|\sum_{n=0}^{\infty} f(n)x(n+k)\right| \\
&\leq \sup_k \sum_{n=0}^{\infty} |f(n)||x(n+k)| \leq \|x\|_{l_\infty} \sum_{n=0}^{\infty} |f(n)| \\
&= \|x\|_{l_\infty} \|f\|_{l_1}
\end{aligned}$$

$\tilde{\mathcal{F}}$ is then immediately well defined, linear, and bounded as a map from c_0 into l_∞. To show that $\tilde{\mathcal{F}} \in \mathcal{B}(c_0)$ we need only show in addition that if $x \in c_0$ then so is $\tilde{\mathcal{F}}x$. Thus, assume $x \in c_0$ and let $\epsilon > 0$ be given. There exists k_ϵ such that $\forall k > k_\epsilon$, $|x(k)| < \frac{\epsilon}{\|f\|_{l_1}}$. Hence, $\forall k > k_\epsilon$,

$$\begin{aligned}
|(\tilde{\mathcal{F}}x)(k)| &= \left|\sum_{n=0}^{\infty} f(n)x(n+k)\right| \\
&\leq \sum_{n=0}^{\infty} |f(n)||x(n+k)| \leq \frac{\epsilon}{\|f\|_{l_1}} \sum_{n=0}^{\infty} |f(n)| \\
&= \epsilon
\end{aligned}$$

and $\tilde{\mathcal{F}}x \in c_0$.

To show that $\mathcal{F}^* = \bar{\mathcal{F}}$, let $x \in l_1$ and $x^* \in l_\infty$ be given. Then

$$\begin{aligned}
\langle \mathcal{F}x, x^* \rangle &= \langle f * x, x^* \rangle = \sum_{k=0}^{\infty} x^*(k) \sum_{n=0}^{k} f(k-n)x(n) \\
&= \sum_{n=0}^{\infty} x(n) \sum_{k=n}^{\infty} f(k-n)x^*(k) = \sum_{n=0}^{\infty} x(n) \sum_{k=0}^{\infty} f(k)x^*(k+n) = \langle x, f \triangleleft x^* \rangle \\
&= \left\langle x, \bar{\mathcal{F}}x^* \right\rangle
\end{aligned}$$

so that $\mathcal{F}^* = \bar{\mathcal{F}}$.

To show that $\tilde{\mathcal{F}}^* = \mathcal{F}$ note that, given $x \in c_0$ and $x^* \in l_1$,

$$\left\langle \tilde{\mathcal{F}}x, x^* \right\rangle = \left\langle x^*, \tilde{\mathcal{F}}x \right\rangle = \langle \mathcal{F}x^*, x \rangle = \langle x, \mathcal{F}x^* \rangle$$

using the fact that $l_1 \subset c_0 \subset l_\infty$, the symmetry of the definition of functional evaluation and the fact that $\mathcal{F}^* = \bar{\mathcal{F}}$. $\square$

Definition C.1.4 *Given positive integers m, n, and p and matrices $L \in l^{m\times p}$ and $R \in l^{p\times n}$, define the operation $\triangleleft$ of* matrix left correlation *by*

$$(L \triangleleft R)_{ij} := \sum_{k=1}^{p} L_{ik} \triangleleft R_{kj}, \qquad i \in \{1,\ldots,m\}\,,\ j \in \{1,\ldots,n\}$$

and the operation $\triangleright$ of matrix right correlation *by*

$$(L \triangleright R)_{ij} := \sum_{k=1}^{p} L_{ik} \triangleright R_{kj}, \qquad i \in \{1,\ldots,m\}\,,\ j \in \{1,\ldots,n\}$$

Proposition C.1.5 *Given positive integers m, n, p and q and matrices $L \in l_1^{m\times p}$ and $R \in l_1^{q\times n}$, define the* correlation operator *$\bar{\mathcal{F}}$ on $l_\infty^{m\times n}$, given $X \in l_\infty^{m\times n}$,*

$$\bar{\mathcal{F}} := L^T \triangleleft X \triangleright R^T$$

and let $\tilde{\mathcal{F}}$ denote its restriction to $c_0^{m\times n}$. Define also the convolution operator *$\mathcal{F} \in \mathcal{B}(l_1^{p\times q}, l_1^{m\times n})$, give $X \in l_1^{p\times q}$,*

$$\mathcal{F}X := L * X * R$$

With these definitions, $\bar{\mathcal{F}} \in \mathcal{B}(l_\infty^{m\times n}, l_\infty^{p\times q})$, $\tilde{\mathcal{F}} \in \mathcal{B}(c_0^{m\times n}, c_0^{p\times q})$, and

1. *$\mathcal{F}^* = \bar{\mathcal{F}}$*

2. *$\tilde{\mathcal{F}}^* = \mathcal{F}$*

with linear functional evaluation defined as in Fact C.1.1.

Proof: It is easy to see that if $\bar{\mathcal{F}}$ is well defined on $l_\infty^{m\times n}$ then it is also linear. The following, where $X \in l_\infty^{m\times n}$ is arbitrary, shows that $\bar{\mathcal{F}}$ is both well defined and bounded on $l_\infty^{m\times n}$:

$$\begin{aligned}
\left\|\bar{\mathcal{F}}X\right\|_{l_\infty} &= \sum_{i=1}^{p} \max_{j\in\{1,\ldots,q\}} \left\| \sum_{r=1}^{m}\sum_{s=1}^{n} L_{ir}^T \triangleleft X_{rs} \triangleright R_{sj}^T \right\|_{l_\infty} \\
&\le \sum_{i=1}^{p} \max_{j\in\{1,\ldots,q\}} \sum_{r=1}^{m}\sum_{s=1}^{n} \left\| L_{ir}^T \triangleleft X_{rs} \triangleright R_{sj}^T \right\|_{l_\infty} \\
&\le pqmn \max_{i,j,r,s} \left\| L_{ir}^T \triangleleft X_{rs} \triangleright R_{sj}^T \right\|_{l_\infty} \\
&\le pqmn \left\| L_{i_0 r_0}^T \right\|_{l_1} \left\| R_{s_0 j_0}^T \right\|_{l_1} \left\| X_{r_0 s_0} \right\|_{l_\infty}
\end{aligned}$$

Where i_0, j_0, r_0, s_0 denote the indices of the maximal scalar norm. $\tilde{\mathcal{F}}$ is then immediately well defined, linear, and bounded as a map from $c_0^{m\times n}$ into $l_\infty^{p\times q}$. To show that

$\tilde{\mathcal{F}} \in \mathcal{B}(c_0^{m\times n}, c_0^{p\times q})$, we need only show in addition that if $X \in c_0^{m\times n}$ then $\tilde{\mathcal{F}}X \in c_0^{p\times q}$. But this is immediate since for each i and j,

$$(\bar{\mathcal{F}}X)_{ij} = \sum_{r=1}^{m}\sum_{s=1}^{n} L_{ir}^T \triangleleft X_{rs} \triangleright R_{sj}^T$$

where the summands are all in c_0, and any finite sum of c_0 sequences is again a c_0 sequence.

To show that $\mathcal{F}^* = \bar{\mathcal{F}}$, let $X \in l_1^{p\times q}$ and $X^* \in l_\infty^{m\times n}$ be given. Then

$$\begin{aligned}
\langle \mathcal{F}X, X^*\rangle &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left\langle \sum_{r=1}^{p}\sum_{s=1}^{q} L_{ir} * X_{rs} * R_{sj}, X_{ij}^*\right\rangle \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{r=1}^{p}\sum_{s=1}^{q}\left\langle L_{ir} * X_{rs} * R_{sj}, X_{ij}^*\right\rangle \\
&= \sum_{r=1}^{p}\sum_{s=1}^{q}\sum_{i=1}^{m}\sum_{j=1}^{n}\left\langle X_{rs}, L_{ir} \triangleleft X_{ij}^* \triangleright R_{sj}\right\rangle \\
&= \sum_{r=1}^{p}\sum_{s=1}^{q}\left\langle X_{rs}, \sum_{i=1}^{m}\sum_{j=1}^{n} L_{ir} \triangleleft X_{ij}^* \triangleright R_{sj}\right\rangle \\
&= \left\langle X, L^T \triangleleft X^* \triangleright R^T\right\rangle = \left\langle X, \bar{\mathcal{F}}X^*\right\rangle
\end{aligned}$$

To show that $\tilde{\mathcal{F}}^* = \mathcal{F}$ note that, given $X \in c_0^{m\times n}$ and $X^* \in l_1^{p\times q}$,

$$\left\langle \tilde{\mathcal{F}}X, X^*\right\rangle = \left\langle X^*, \tilde{\mathcal{F}}X\right\rangle = \langle \mathcal{F}X^*, X\rangle = \langle X, \mathcal{F}X^*\rangle$$

using the fact that $l_1 \subset c_0 \subset l_\infty$, the symmetry of the definition of functional evaluation and the fact that $\mathcal{F}^* = \bar{\mathcal{F}}$. $\square$

Proposition C.1.6 *Given positive integers m and n and a subset*

$$\boldsymbol{S} \subset \{1, \ldots, m\} \times \{1, \ldots, n\}$$

define the spatial embedding operator $\bar{\mathcal{E}}_S$ on l_∞^S, given $X \in l_\infty^S$,

$$(\bar{\mathcal{E}}_S X)_{ij} := \begin{cases} X_{ij} & (i,j) \in \boldsymbol{S} \\ 0 & (i,j) \in \{1, \ldots, m\} \times \{1, \ldots, n\} \setminus \boldsymbol{S} \end{cases}$$

and let $\tilde{\mathcal{E}}_S$ denote its restriction to c_0^S. Define also the spatial projection operator Π_S on $l_1^{m\times n}$, given $X \in l_1^{m\times n}$,

$$(\Pi_S X)_{ij} := X_{ij} \quad \text{for all } (i,j) \in S$$

With these definitions, $\bar{\mathcal{E}}_S \in \mathcal{B}(l_\infty^S, l_\infty^{m\times n})$, $\tilde{\mathcal{E}}_S \in \mathcal{B}(c_0^S, c_0^{m\times n})$, $\Pi_S \in \mathcal{B}(l_1^{m\times n}, l_1^S)$, and

1. $\Pi_S^* = \bar{\mathcal{E}}_S$,

2. $\tilde{\mathcal{E}}_S^* = \Pi_S$,

with linear functional evaluation defined as in Fact C.1.1.

Proof: It is easy to see that $\bar{\mathcal{E}}_S$, $\tilde{\mathcal{E}}_S$, and Π_S are all bounded linear operators as claimed. To show that $\Pi_S^* = \bar{\mathcal{E}}_S$, let $X \in l_1^{m\times n}$ and $X^* \in l_\infty^S$ be given. Then

$$\begin{aligned}\langle \Pi_S X, X^* \rangle &= \sum_{(i,j)\in S} \left\langle (\Pi_S X)_{ij}, X_{ij}^* \right\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \left\langle X_{ij}, (\bar{\mathcal{E}}_S X^*)_{ij} \right\rangle \\ &= \left\langle X, \bar{\mathcal{E}}_S X^* \right\rangle\end{aligned}$$

To show that $\tilde{\mathcal{E}}_S^* = \Pi_S$ note that, given $X \in c_0^S$ and $X^* \in l_1^{m\times n}$,

$$\left\langle \tilde{\mathcal{E}}_S X, X^* \right\rangle = \left\langle X^*, \tilde{\mathcal{E}}_S X \right\rangle = \langle \Pi_S X^*, X \rangle = \langle X, \Pi_S X^* \rangle$$

using the fact that $l_1 \subset c_0 \subset l_\infty$, the symmetry of the definition of functional evaluation and the fact that $\Pi_S^* = \bar{\mathcal{E}}_S$. □

Proposition C.1.7 *Given positive integers m and n, a subset*

$$\boldsymbol{S} \subset \{1,\ldots,m\} \times \{1,\ldots,n\}$$

and $N \in \mathbf{Z}_+$, define the truncation operator $\bar{\mathcal{P}}_N$ on l_∞^S, given $X \in l_\infty^S$,

$$(\bar{\mathcal{P}}_N X)_{ij}(k) := \begin{cases} X_{ij} & k \le N \\ 0 & k > N \end{cases} \qquad \text{for all } (i,j) \in \boldsymbol{S}$$

and let $\tilde{\mathcal{P}}_N$, $\mathcal{P}_N$ denote its restrictions to c_0^S, l_1^S respectively. With these definitions, $\bar{\mathcal{P}}_N \in \mathcal{B}(l_\infty^S)$, $\tilde{\mathcal{P}}_N \in \mathcal{B}(c_0^S)$, $\mathcal{P}_N \in \mathcal{B}(l_1^S)$, and

1. $\mathcal{P}_N^* = \bar{\mathcal{P}}_N$,

2. $\tilde{\mathcal{P}}_N^* = \mathcal{P}_N$,

with linear functional evaluation defined as in Fact C.1.1.

Proof: It is easy to see that $\bar{\mathcal{P}}_N$, $\tilde{\mathcal{P}}_N$, and $\mathcal{P}_N$ are all bounded linear operators as claimed. To show that $\mathcal{P}_N^* = \bar{\mathcal{P}}_N$, let $X \in l_1^S$ and $X^* \in l_\infty^S$ be given. Then

$$\begin{aligned}\langle \mathcal{P}_N X, X^* \rangle &= \sum_{(i,j)\in S} \left\langle X_{ij}, X_{ij}^* \right\rangle = \sum_{(i,j)\in S} \sum_{k=0}^{N} X_{ij}(k) X_{ij}^*(k) = \sum_{(i,j)\in S} \left\langle X_{ij}, (\bar{\mathcal{P}}_N X^*)_{ij} \right\rangle \\ &= \left\langle X, \bar{\mathcal{P}}_N X^* \right\rangle\end{aligned}$$

To show that $\tilde{\mathcal{P}}_N^* = \mathcal{P}_N$ note that, given $X \in c_0^S$ and $X^* \in l_1^S$,

$$\left\langle \tilde{\mathcal{P}}_N X, X^* \right\rangle = \left\langle X^*, \tilde{\mathcal{P}}_N X \right\rangle = \langle \mathcal{P}_N X^*, X \rangle = \langle X, \mathcal{P}_N X^* \rangle$$

using the fact that $l_1 \subset c_0 \subset l_\infty$, the symmetry of the definition of functional evaluation and the fact that $\mathcal{P}_N^* = \bar{\mathcal{P}}_S$. □

Lemma C.1.8 *Given a positive integer n and a complex number $z_0 \in \mathbf{D}$, the two sequences $d_\Re$ and $d_\Im$ defined*

$$d_\Re(k) = \begin{cases} 0 & k < n \\ \frac{k!}{(k-n)!}\Re(z_0^{k-n}) & k \geq n \end{cases}$$

and

$$d_\Im(k) = \begin{cases} 0 & k < n \\ \frac{k!}{(k-n)!}\Im(z_0^{k-n}) & k \geq n \end{cases}$$

are in l_1.

Proof: We will prove only $d_\Re \in l_1$; the proof for $d_\Re$ is analogous.

$$\begin{aligned}\|d_\Re\|_{l_1} &= \sum_{k=0}^{\infty} |d_\Re(k)| = \sum_{k=n}^{\infty} \left| \frac{k!}{(k-n)!}\Re(z_0^{k-n}) \right| = \sum_{k=0}^{\infty} \left| \frac{(k+n)!}{k!}\Re(z_0^k) \right| \\ &\leq \sum_{k=0}^{\infty} \left| \frac{(k+n)!}{k!} \right| \left| z_0^k \right| \leq \sum_{k=0}^{\infty} |(k+n)^n| \left| z_0^k \right| \\ &= \sum_{k=0}^{\infty} \left| \sum_{r=0}^{n} \binom{n}{r} k^r n^{n-r} \right| \left| z_0^k \right| \leq (n+1)n!n^n \sum_{k=0}^{\infty} k^n \left| z_0^k \right| \\ &< \infty\end{aligned}$$

where the last line follows since $z_0 \in \mathbf{D}$. □

Proposition C.1.9 *Given a positive integer n and a complex number $z_0 \in \mathbf{D}$, define two operators $\mathcal{D}_\Re$ and $\mathcal{D}_\Im$ on l_1, given $x \in l_1$,*

$$\mathcal{D}_\Re x := \sum_{k=0}^{\infty} x(k) d_\Re(k) \quad \textit{and} \quad \mathcal{D}_\Im x := \sum_{k=0}^{\infty} x(k) d_\Im(k)$$

where $d_{\Re}$ and $d_{\Im}$ are defined as in Lemma C.1.8. Define also two operators $\bar{\mathcal{D}}_{\Re}$ and $\bar{\mathcal{D}}_{\Im}$ on $\mathbf{R}$, given $x \in \mathbf{R}$,

$$\bar{\mathcal{D}}_{\Re}x := xd_{\Re} \quad \textit{and} \quad \bar{\mathcal{D}}_{\Im}x := xd_{\Im}$$

With these definitions, $\bar{\mathcal{D}}_{\Re}, \bar{\mathcal{D}}_{\Im} \in \mathcal{B}(\mathbf{R}, l_\infty)$, $\bar{\mathcal{D}}_{\Re}, \bar{\mathcal{D}}_{\Im} \in \mathcal{B}(\mathbf{R}, c_0)$, $\mathcal{D}_{\Re}, \mathcal{D}_{\Im} \in \mathcal{B}(l_1, \mathbf{R})$,

1. *$\mathcal{D}_{\Re}^* = \bar{\mathcal{D}}_{\Re}$, and $\mathcal{D}_{\Im}^* = \bar{\mathcal{D}}_{\Im}$,*

2. *$\bar{\mathcal{D}}_{\Re}^* = \mathcal{D}_{\Re}$, and $\bar{\mathcal{D}}_{\Im}^* = \mathcal{D}_{\Im}$,*

with linear functional evaluation defined as in Fact C.1.1.

Proof: We will prove all facts concerning $\mathcal{D}_{\Re}$ and $\bar{\mathcal{D}}_{\Re}$; the corresponding proofs for $\mathcal{D}_{\Im}$ and $\bar{\mathcal{D}}_{\Re}$ are analogous. The fact that $\bar{\mathcal{D}}_{\Re} \in \mathcal{B}(\mathbf{R}, l_\infty)$ follows since $d_{\Re} \in l_1 \subset l_\infty$ and, given $x \in \mathbf{R}$, $\left\|\bar{\mathcal{D}}_{\Re}x\right\|_{l_\infty} = \|xd_{\Re}\|_{l_\infty} = |x|\,\|d_{\Re}\|_{l_\infty}$. $\bar{\mathcal{D}}_{\Re} \in \mathcal{B}(\mathbf{R}, c_0)$ because also $d_{\Re} \in l_1 \subset c_0 \Rightarrow \mathcal{R}\left(\bar{\mathcal{D}}_{\Re}\right) \subset c_0$. The fact that $\mathcal{D}_{\Re} \in \mathcal{B}(l_1, \mathbf{R})$ follows since, given $x \in l_1$,

$$\mathcal{D}_{\Re}x = \sum_{k=0}^{\infty} x(k)d_{\Re}(k) = \langle x, d_{\Re}\rangle$$

where we have used the facts that $d_{\Re} \in l_1$ (Lemma C.1.8), $l_1 \subset l_\infty$, and $l_1^* = l_\infty$ (Fact C.1.1).

To show that $\mathcal{D}_{\Re}^* = \bar{\mathcal{D}}_{\Re}$, let $x \in l_1$ and $x^* \in \mathbf{R}$ be given. Then

$$\begin{aligned} \langle \mathcal{D}_{\Re}x, x^*\rangle &= x^* \sum_{k=0}^{\infty} x(k)\frac{k!}{(k-n)!}\Re(z_0^{k-n}) = \sum_{k=n}^{\infty} x(k)\left[x^*\frac{k!}{(k-n)!}\Re(z_0^{k-n})\right] \\ &= \sum_{k=0}^{\infty} x(k)\left[\bar{\mathcal{D}}_{\Re}x^*\right] = \left\langle x, \bar{\mathcal{D}}_{\Re}x^*\right\rangle \end{aligned}$$

To show that $\bar{\mathcal{D}}_{\Re}^* = \mathcal{D}_{\Re}$ note that, given $x \in \mathbf{R}$ and $x^* \in l_1$,

$$\left\langle \bar{\mathcal{D}}_{\Re}x, x^*\right\rangle = \left\langle x^*, \bar{\mathcal{D}}_{\Re}x\right\rangle = \langle \mathcal{D}_{\Re}x^*, x\rangle = \langle x, \mathcal{D}_{\Re}x^*\rangle$$

using the fact that $l_1 \subset c_0 \subset l_\infty$, the symmetry of the definition of functional evaluation and the fact that $\mathcal{D}_{\Re}^* = \bar{\mathcal{D}}_{\Re}$. $\square$

C.2 Spaces of Sequences on Z

$l_1(\mathbf{Z})$, $l_\infty(\mathbf{Z})$, and $c_0(\mathbf{Z})$ denote the counterparts of the classical sequence spaces with norms defined analogously; $c_0(\mathbf{Z})$ is a subspace of $l_\infty(\mathbf{Z})$ defined

$$c_0(\mathbf{Z}) := \{x \in l_\infty(\mathbf{Z}) : \lim_{k\to\pm\infty} x(k) = 0\}.$$

Matrix versions and duality relations analogous to those of the preceding section also hold for these spaces.

Fact C.2.1 *With addition and scalar multiplication defined componentwise,*

1. *the following all define real normed linear spaces:*

   (a) $l_1(\mathbf{Z})^S$ *with the norm* $\|X\|_{l_1} := \max_{i\in S_I} \sum_{(i,j)\in S} \|X_{ij}\|_{l_1}$

   (b) $l_\infty(\mathbf{Z})^S$ *with the norm* $\|X\|_{l_\infty} := \sum_{i\in S_I} \max_{(i,j)\in S} \|X_{ij}\|_{l_\infty}$

   (c) *the subspace* $c_0(\mathbf{Z})^S$ *of* $l_\infty(\mathbf{Z})^S$

2. *the duality relations*

   (a) $(c_0(\mathbf{Z})^S)^* = l_1(\mathbf{Z})^S$,

   (b) $(l_1(\mathbf{Z})^S)^* = l_\infty(\mathbf{Z})^S$,

   hold when linear functional evaluation is defined, given X *in a primal space and* X^* *in its dual,*

$$\langle X, X^*\rangle := \sum_{(i,j)\in S} \left\langle X_{ij}, X_{ij}^*\right\rangle$$

Definition C.2.2 *The binary operations* $\triangleleft$ *of* left correlation *and* $\triangleright$ *of* right correlation *are defined, given* $f, g \in l(\mathbf{Z})$,

$$f \triangleleft g := \left\{ \sum_{n=-\infty}^{\infty} f(n)g(n+k) \right\}_{k\in\mathbf{Z}} =: g \triangleright f$$

Proposition C.2.3 *Given* $f \in l_1(\mathbf{Z})$, *define the* correlation operator $\bar{\mathcal{F}}$ *on* $l_\infty(\mathbf{Z})$, *given* $x \in l_\infty(\mathbf{Z})$,

$$\bar{\mathcal{F}}x := f \triangleleft x = x \triangleright f$$

and let $\tilde{\mathcal{F}}$ denote its restriction to $c_0(\mathbf{Z})$. Define also the convolution operator $\mathcal{F} \in \mathcal{B}(l_1(\mathbf{Z}))$*, given* $x \in l_1(\mathbf{Z})$,

$$\mathcal{F}x := f * x$$

With these definitions, $\bar{\mathcal{F}} \in \mathcal{B}(l_\infty(\mathbf{Z}))$, $\tilde{\mathcal{F}} \in \mathcal{B}(c_0(\mathbf{Z}))$, *and*

1. $\mathcal{F}^* = \bar{\mathcal{F}}$

2. $\tilde{\mathcal{F}}^* = \mathcal{F}$

with linear functional evaluation defined as in Fact C.1.1.

Proof: It is easy to see that if $\bar{\mathcal{F}}$ is well defined on $l_\infty(\mathbf{Z})$ then it is also linear. The following, where $x \in l_\infty(\mathbf{Z})$ is arbitrary, shows that $\bar{\mathcal{F}}$ is both well defined and bounded on $l_\infty(\mathbf{Z})$:

$$\begin{aligned}
\left\|\bar{\mathcal{F}}x\right\|_{l_\infty} &= \sup_k \left| \sum_{n=-\infty}^{\infty} f(n)x(n+k) \right| \\
&\leq \sup_k \sum_{n=-\infty}^{\infty} |f(n)|\,|x(n+k)| \leq \|x\|_{l_\infty} \sum_{n=-\infty}^{\infty} |f(n)| \\
&= \|x\|_{l_\infty} \|f\|_{l_1}
\end{aligned}$$

$\tilde{\mathcal{F}}$ is then immediately well defined, linear, and bounded as a map from $c_0(\mathbf{Z})$ into $l_\infty(\mathbf{Z})$. To show that $\tilde{\mathcal{F}} \in \mathcal{B}(c_0(\mathbf{Z}))$ we need only show in addition that if $x \in c_0(\mathbf{Z})$ then so is $\tilde{\mathcal{F}}x$. Given $x \in c_0(\mathbf{Z})$ we will show that $\lim_{k\to\infty}(\tilde{\mathcal{F}}x)(k) = 0$; to show that $\lim_{k\to\infty}(\tilde{\mathcal{F}}x)(k) = 0$ is entirely similar and the conclusion follows. Accordingly, let $\epsilon > 0$ be given. Because $f \in l_1(\mathbf{Z})$, there exists $k_f \in \mathbf{Z}$ such that $\sum_{-\infty}^{-k_f} |f(k)| < \frac{\epsilon}{2\|x\|_{l_\infty}}$ and, because $x \in c_0(\mathbf{Z})$, there exists $k_x \in \mathbf{Z}$ such that $\sup\{|x(k)| : k \geq k_x\} < \frac{\epsilon}{2\|f\|_{l_1}}$. Hence, for $k \geq k_f + k_x$,

$$\begin{aligned}
\left|(\tilde{\mathcal{F}}x)(k)\right| &= \left| \sum_{n=-\infty}^{\infty} f(n)x(n+k) \right| \\
&\leq \sum_{n=-\infty}^{-k_f-1} |f(n)|\,|x(n+k)| + \sum_{n=-k_f}^{\infty} |f(n)|\,|x(n+k)| \\
&\leq \|x\|_{l_\infty} \sum_{n=-\infty}^{-k_f-1} |f(n)| + \|f\|_{l_1} \sup\{|x(k)| : k \geq k_x\} \\
&< \epsilon.
\end{aligned}$$

To show that $\mathcal{F}^* = \bar{\mathcal{F}}$, let $x \in l_1(\mathbf{Z})$ and $x^* \in l_\infty(\mathbf{Z})$ be given. Then

$$\begin{aligned}
\langle \mathcal{F}x, x^* \rangle &= \langle f * x, x^* \rangle = \sum_{k=-\infty}^{\infty} x^*(k) \sum_{n=-\infty}^{\infty} f(k-n)x(n) \\
&= \sum_{n=-\infty}^{\infty} x(n) \sum_{k=-\infty}^{\infty} f(k-n)x^*(k) = \sum_{n=-\infty}^{\infty} x(n) \sum_{k=-\infty}^{\infty} f(k)x^*(k+n) \\
&= \langle x, f \triangleleft x^* \rangle = \left\langle x, \bar{\mathcal{F}}x^* \right\rangle
\end{aligned}$$

so that $\mathcal{F}^* = \bar{\mathcal{F}}$.

To show that $\tilde{\mathcal{F}}^* = \mathcal{F}$ note that, given $x \in c_0(\mathbf{Z})$ and $x^* \in l_1(\mathbf{Z})$,

$$\left\langle \tilde{\mathcal{F}}x, x^* \right\rangle = \left\langle x^*, \tilde{\mathcal{F}}x \right\rangle = \langle \mathcal{F}x^*, x \rangle = \langle x, \mathcal{F}x^* \rangle$$

using the fact that $l_1(\mathbf{Z}) \subset c_0(\mathbf{Z}) \subset l_\infty(\mathbf{Z})$, the symmetry of the definition of functional evaluation and the fact that $\mathcal{F}^* = \bar{\mathcal{F}}$. □

Definition C.2.4 *Given positive integers m, n, and p and matrices $L \in l^{m\times p}$ and $R \in l^{p\times n}$, define the operation $\triangleleft$ of* matrix left correlation *by*

$$(L \triangleleft R)_{ij} := \sum_{k=1}^{p} L_{ik} \triangleleft R_{kj}, \qquad i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}$$

and the operation $\triangleright$ of matrix right correlation *by*

$$(L \triangleright R)_{ij} := \sum_{k=1}^{p} L_{ik} \triangleright R_{kj}, \qquad i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}$$

Proposition C.2.5 *Given positive integers m, n, p and q and matrices $L \in l_1(\mathbf{Z})^{m\times p}$ and $R \in l_1(\mathbf{Z})^{q\times n}$, define the* correlation operator *$\bar{\mathcal{F}}$ on $l_\infty(\mathbf{Z})^{m\times n}$, given $X \in l_\infty(\mathbf{Z})^{m\times n}$,*

$$\bar{\mathcal{F}} := L^T \triangleleft X \triangleright R^T$$

and let $\tilde{\mathcal{F}}$ denote its restriction to $c_0(\mathbf{Z})^{m\times n}$. Define also the convolution operator *$\mathcal{F} \in \mathcal{B}(l_1(\mathbf{Z})^{p\times q}, l_1(\mathbf{Z})^{m\times n})$, give $X \in l_1(\mathbf{Z})^{p\times q}$,*

$$\mathcal{F}X := L * X * R$$

With these definitions, $\bar{\mathcal{F}} \in \mathcal{B}(l_\infty(\mathbf{Z})^{m\times n}, l_\infty(\mathbf{Z})^{p\times q})$, $\tilde{\mathcal{F}} \in \mathcal{B}(c_0(\mathbf{Z})^{m\times n}, c_0(\mathbf{Z})^{p\times q})$, and

1. $\mathcal{F}^* = \bar{\mathcal{F}}$
2. $\tilde{\mathcal{F}}^* = \mathcal{F}$

with linear functional evaluation defined as in Fact C.2.1.

Proof: The proof of Proposition C.1.5 carries over exactly to the present case. □

Bibliography

[1] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses," *IEEE Trans. Automat. Control*, vol. AC-26, pp. 301–320, Apr. 1981.

[2] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard $\mathcal{H}_{\in}$ and $\mathcal{H}_{\infty}$ control problems," *IEEE Trans. Automat. Control*, vol. AC-34, pp. 831–847, Aug. 1989.

[3] M. Vidyasagar, "Optimal rejection of persistent bounded disturbances," *IEEE Trans. Automat. Control*, vol. AC-31, pp. 527–534, 1986.

[4] M. A. Dahleh and J. B. Pearson, "ℓ^1-optimal feedback controllers for discrete-time systems," in *Proceedings of the 1986 American Control Conference*, pp. 1964–1968, 1986.

[5] M. A. Dahleh and J. B. Pearson, "ℓ^1-optimal feedback control for MIMO discrete-time systems," *IEEE Trans. Automat. Control*, vol. AC-32, pp. 314–323, 1987.

[6] M. A. Dahleh and J. B. Pearson, "Optimal rejection of persistent bounded disturbances, robust stability, and mixed sensitivity minimization," *IEEE Trans. Automat. Control*, vol. AC-33, pp. 722–731, 1988.

[7] P. Iglesias and K. Glover, "State space approach to discrete time $\mathcal{H}_{\infty}$ control," Technical Report, Engineering Department, Cambridge University, 1990.

[8] M. A. Dahleh and J. B. Pearson, "l^1-optimal compensators for continuous-time systems," *IEEE Trans. Automat. Control*, vol. AC-32, pp. 889–895, 1987.

[9] M. A. Dahleh, P. G. Voulgaris, and L. S. Valavani, "Optimal and robust controllers for periodic and multirate systems," *IEEE Trans. Automat. Control*, vol. 37, pp. 90–99, Jan. 1992.

[10] J. S. Shamma and M. A. Dahleh, "Time-varying versus time-invariant compensation for rejection of persistent bounded disturbances," *IEEE Trans. Automat. Control*, vol. AC-36, pp. 838–847, July 1991.

[11] B. A. Bamieh, *Analysis and Robust Control of Hybrid Continuous/Discrete Time Systems.* PhD thesis, Rice University, Houston, TX, July 1991.

[12] G. E. Dullerud and B. A. Francis, "$\mathcal{L}_\infty$ analysis and design of sampled-data systems," *IEEE Trans. Automat. Control*, vol. AC-37, pp. 436–446, Apr. 1992.

[13] M. R. Keenan and J. C. Kantor, "An l_∞ optimal approach to robust feedback control," in *Proceedings of the 1989 American Control Conference*, 1989.

[14] M. R. Keenan and J. C. Kantor, "An l_∞ optimal approach to linear feedback control," in *Second Shell Workshop on Process Control*, 1988.

[15] J. S. McDonald and J. B. Pearson, "ℓ^1-optimal control of multivariable system with output norm constraints," *Automatica*, vol. 27, no. 3, 1991.

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

[17] M. Vidyasagar, *Control System Synthesis: A Factorization Approach.* Cambridge, Massachusetts: MIT Press, 1985.

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

[19] M. C. Smith, "On stabilization and the existence of coprime factorizations," *IEEE Trans. Automat. Control*, vol. AC-34, pp. 1005–1007, Sept. 1989.

[20] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, Jr., "Modern Wiener-Hopf design of optimal controllers, part I," *IEEE Trans. Automat. Control*, vol. AC-21, pp. 3–15, Jan. 1976.

[21] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, Jr., "Modern Wiener-Hopf design of optimal controllers, part II: The multivariable case," *IEEE Trans. Automat. Control*, vol. AC-21, pp. 319–338, June 1976.

[22] C. A. Desoer, R.-W. Liu, J. Murray, and R. Saeks, "Feedback system design: The fractional representation approach to analysis and synthesis," *IEEE Trans. Automat. Control*, vol. AC-25, pp. 399–412, June 1980.

[23] R. Saeks and J. Murray, "Fractional representation, algebraic geometry, and the simultaneous stabilization problem," *IEEE Trans. Automat. Control*, vol. AC-27, pp. 895–903, Aug. 1982.

[24] C. A. Desoer and C. L. Gustafson, "Algebraic theory of linear multivariable feedback systems," *IEEE Trans. Automat. Control*, vol. AC-29, pp. 909–916, Oct. 1984.

[25] M. Vidyasagar, H. Schneider, and B. A. Francis, "Algebraic and topological aspects of feedback stabilization," *IEEE Trans. Automat. Control*, vol. AC-27, pp. 880–894, Aug. 1982.

[26] W. Rudin, *Functional Analysis.* McGraw-Hill, 1973.

[27] A. G. Parlos, A. F. Henry, F. C. Schweppe, L. A. Gould, and D. D. Lanning, "Nonlinear multivariable control of nuclear power plants based on the unknown-but-bounded disturbance model," *IEEE Trans. Automat. Control*, vol. AC-33, pp. 130–137, Feb. 1988.

[28] J. B. Pearson and B. Bamieh, "On minimizing maximum errors," *IEEE Trans. Automat. Control*, vol. AC-35, pp. 598–601, May 1990.

[29] D. G. Meyer and S. P. Boyd, "A note on the order of l_1 optimal compensators," in *Proceedings of the 25th IEEE Conference on Decision and Control*, (Athens, Greece), pp. 14–15, 1986.

[30] O. J. Staffans, "The mixed sensitivity minimization problem has a rational l^1-optimal solution," Research Report A274, Helsinki Institute of Mathematics, Espoo, Finland, 1990.

[31] O. J. Staffans, "MIMO l^1-optimization with a scalar control," Research Report A280, Helsinki Institute of Mathematics, Espoo, Finland, June 1990.

[32] O. J. Staffans, "On the four-block model matching problem in l^1," Research Report A289, Helsinki Institute of Mathematics, Espoo, Finland, December 1990.

[33] I. J. Diaz-Bobillo and M. A. Dahleh, "Minimization of the maximum peak-to-peak gain: The general multiblock problem." submitted for publication, Mar. 1992.

[34] I. J. Diaz-Bobillo, *The General ℓ_1 Multiblock Problem: Exact and Approximate Solutions.* PhD thesis, Massachusetts Institute of Technology, Feb. 1992.

[35] Y. Ohta, H. Maeda, and S. Kodama, "Rational approximation of l_1-optimal controller for SISO systems," in *Proceedings of the 30th IEEE Conference on Decision and Control*, 1991.

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