2003

Design of optimal equalizers and precoders for MIMO channels

Lijuan Li
Louisiana State University and Agricultural and Mechanical College

DESIGN OF OPTIMAL EQUALIZERS AND PRECODERS FOR MIMO CHANNELS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Electrical and Computer Engineering

by
Lijuan Li
B.S., Liaoning University, P.R.China, 1996
M.S., Fudan University, P.R.China, 1999
M.S.E.E, Louisiana State University, 2001
August, 2003

Acknowledgments

My sincere appreciation and thanks are given to my adviser and committee chair Dr. Guoxiang Gu for his patient academic guidance and constant input throughout the research and preparation of this dissertation. His insight and expertise in this field have deeply influenced me and my work recorded herein. Without his constructive direction and invaluable advice, this work would not have been completed.

My gratitude is also extended to my co-adviser Dr. Kemin Zhou for his intellectual suggestions for the improvement of this dissertation, special thank for his support and encouragement during the entire course of the study under which the research work progressed.

I would like to thank Dr. Subhash C Kak and Dr. Hsiao-Chun Wu for their willingness to serve as my examining committee member and their sincere encouragement. I am grateful to Dr. Peter Wolenski for taking his time and effort to be my minor professor and his valuable feedback on my dissertation.

Thanks also go to the faculty, stuff and students at the LSU Electrical Engineering Department for their assistance. This research was supported by U.S. Air Force Office for Scientific Research, a research grant from NASA, and EDA fellowship.

Table of Contents

List of Figures

Notation and Symbols

A^{-1}:	Inverse of A
$\det(A)$:	Determinant of A
$\text{trace}(A)$:	Trace of A, $\text{trace}(A)=\sum_i A_{ii}$
A':	Transpose of A
A^*:	Complex conjugate transpose of A
$\lambda(A)$:	Eigenvalue of A
$\overline{\sigma}(\cdot)$:	Maximum singular value
I_n:	Identity matrix of size $n \times n$
$0_{n\times m}$:	Zero matrix of size $n \times m$
$:=$	Defined as
$\star$:	Convolution
$\left[\begin{array}{c\|c} A & B \\ \hline C & D \end{array}\right]$:	Shorthand for state space realization $C(zI-A)^{-1}B+D$

List of Acronyms

ARE:	Algebraic Riccati equation
BC:	Broadcast channel
BER:	Bit error rate
CDMA:	Code division multiple access
CSI:	Channel state information
FDMA:	Frequency division multiple access
FIR:	Finite impulse response
IIR:	Infinite impulse response
ICI:	Inter-channel interference
ISI:	Inter-symbol interference
LTI:	Linear time invariant
MAC:	Multiple access channels
MIMO:	Multi-input, multi-output
OFDM:	Orthogonal frequency division multiplexing
PR:	Perfect reconstruction
PSD:	Power spectral density
QoS:	Quality of service
RMS:	Root mean-squared
SNR:	Signal to noise ratio
SVD:	Singular value decomposition
WSS:	Wide-sense stationary
ZF:	Zero-forcing

Abstract

Channel equalization has been extensively studied as a method of combating ISI and ICI for high speed MIMO data communication systems. This dissertation focuses on optimal channel equalization in the presence of non-white observation noises with unknown PSD but bounded power-norm. A worst-case approach to optimal design of channel equalizers leads to an equivalent optimal $\mathcal{H}_\infty$ filtering problem for the MIMO communication systems. An explicit design algorithm is derived which not only achieves the zero-forcing (ZF) condition, but also minimizes the RMS error between the transmitted symbols and the received symbols. The second part of this dissertation investigates the design of optimal precoders which minimize the bit error rate (BER) subject to a fixed transmit-power constraint for the multiple antennas downlink communication channels under the perfect reconstruction (PR) condition. The closed form solutions are derived and an efficient design algorithm is proposed. The performance evaluations indicate that the optimal precoder design for multiple antennas communication systems proposed herein is an attractive/reasonable alternative to the existing precoder design techniques.

Chapter 1

Introduction

Future wireless communication systems will operate at considerably higher data rates with higher reliability than what we have today. Digital communication using multi-input multi-output (MIMO) wireless links is emerging as one of the most promising research areas in wireless communications to increase the spectrum efficiency and the quality of service (QoS). MIMO systems are revolutionizing the way we transmit and receive data and have key applications in the future high-speed high-spectrum efficiency wireless networks (3G and beyond).

During the last decade, advanced signal processing techniques and the rapid evolution of cheap and fast electronics, have attracted an increased interest in the use of multiple antennas at both the base stations and terminals of wireless communication networks. In a conventional communication system employing one antenna on each side of the channel, the achievable capacity is limited by the transmit power and available spectrum bandwidth. To solve this problem, communication systems employing multiple antennas at both the transmitter and receiver sites, forming a

multi-input/multi-output architecture, have been suggested. Moreover, MIMO systems allow multiple users to transmit their signals and to share the same time- and frequency-slots, and hence achieve drastic capacity enhancement. When communicating over a wireless channel, all users use the same resources, and the radio spectrum.

In the transmission of digital data over the MIMO communication systems, the signal propagates from the transmitter to the receiver through different paths, each associated with a time delay. As such each transmitted symbol is received several times, and each received symbol is disturbed by other symbols in the transmitted sequence, causing *inter-symbol interference* (ISI). While ideally any multiple access scheme should provide each transmitter/receiver pair with an independent channel, in practice this is often not the case. Due to inherent properties of the multiple access scheme in use, different users interfere with each other, causing the so-called *inter-channel interference* (ICI), also known as multiple access interference (MAI). The ISI and ICI, due to the multi-path propagation and multi-user impairment, if left uncompensated, will in turn, give rise to high error rate in symbol detection. A solution to this problem is to adopt a technique which can compensate or reduce the ISI and/or ICI in the received signal prior to detection. Such a technique is called *channel equalization*.

The purpose of this chapter is to give some introductions and background materials to the remainder of the dissertation. Some earlier work in the field is reviewed and the

work in the later chapters is motivated. Also, the contributions of the work presented herein are stated, and the notations used throughout the dissertation are introduced.

1.1 MIMO Channels

In wireless communications, transmitted signals arriving at the receiver propagate through multiple paths, due to the reflection, refraction, or diffraction in the channel. Multi-path propagation results in received signals which are superposition of several delayed and scaled copies of the transmitted signals giving rise to *frequency-selective fading*. Frequency-selective fading is caused by destructive interference among multiple propagation paths. Moreover, the environment around the transmitters and the receivers can change over time, particularly in a mobile scattering, leading to variations in the channel response with time. This gives rise to *time-selective fading*.

Consider a narrow band discrete-time MIMO channel with p transmitters and q receivers. Let $s_j(k)$ be the sequence from transmitter j ($j = 1, 2, \cdots, p$), and let $r_i(k)$ be the received signal from receiver i ($i = 1, 2, \cdots, q$). Then the two are related via a time-variant linear system as

$$r_i(n) = \sum_{j=1}^{p} h_{i,j}(n,k) \star s_j(n) + v_i(n), \quad i = 1, 2, \cdots, q, \tag{1.1}$$

where $h_{i,j}(n,k)$ is the channel impulse response at time n to a unit input at time $n-k$, from the jth transmitter ($j = 1, 2, \cdots, p$) to the ith receiver ($i = 1, 2, \cdots, q$). The sequence $\{v_i(n)\}_{i=1}^{q}$ represents the additive noise at the receiver. Equation (1.1) describes a time- and frequency-selective linear channel. For a slowly time-varying

system, one often simplifies system (1.1) to a time-invariant system as

$$r_i(n) = \sum_{j=1}^{p} h_{i,j}(n) \star s_j(n) + v_i(n), \quad i = 1, 2, \cdots, q, \tag{1.2}$$

where $h_{i,j}(n)$ is the time-invariant channel response to a unit input at time 0. Equation (1.2) describes a frequency-selective linear channel with no time-selective fading.

In the frequency ($\mathcal{Z}$-transform) domain, the MIMO linear time-invariant system (1.2) has the following description:

$$\mathbf{r}(z) = H(z)\mathbf{s}(z) + \mathbf{v}(z), \tag{1.3}$$

where $\mathbf{s}(z)$ is the input signal of size $p \times 1$, $\mathbf{v}(z)$[1] is the additive Gaussian noise of size $q \times 1$, and $\mathbf{r}(z)$ is the output signal of size $q \times 1$, all are the blocked version of the original scalar data streams. Consequently $H(z)$ is the linear time-invariant channel with size $q \times p$. Equation (1.3) is a commonly used model for the design of MIMO communication systems. In the following chapters, this general model will be extended to the different scenarios of interest.

In fact the general model (1.3) is capable of characterizing many different wireless transmission systems, including:

(a) Single-user (point to point) communication systems with multiple transmit antennas and receive antennas.

(b) Multi-user channels, including multiple access channel (MAC) and broadcast channel (BC).

[1]There is an abuse of notation here, as $\mathcal{Z}$-transform of random processes does not exist in general.

(c) Virtual MIMO channels for systems with transmitter-induced diversities including fractional sampling [41], orthogonal frequency division multiplexing (OFDM), and code division multiple access (CDMA), etc.

The MIMO channels corresponding to the above different categories will be studied in Section 2 of Chapter 2 along with the realistic communication models. In fact, the MIMO techniques proposed herein are applicable to any channel represented by a matrix, or equivalently a matrix after the system blocking operation. The channels to be covered in this dissertation are listed below:

(1) The optimal equalizer is designed for the filterbank transceiver (*transmitter + receiver*), which belongs to the item (c);

(2) The optimal precoder is designed for the multiple antennas system, which belongs to the item (a).

Throughout this dissertation, unless stated otherwise, we make the following assumptions.

Assumption 1. *The channel can be modeled as being linear time-invariant for the duration of each burst.*

Most wireless communication channels are time-variant, albeit slowly. However, the channel variation over the transmission duration of a burst can be ignored under certain conditions and can be regarded as static, although the time variations between bursts can be significant. For instance, this assumption is valid in a wireless environment under low mobility scenarios.

Assumption 2. *The channel can be modeled as a finite impulse response (FIR) filter.*

This is a standard assumption made in the literature. Although the delay spreads of the channel impulse response depends on the transmission environment, the FIR approximation is typically valid. For the causal linear channel model with a finite length, i.e., FIR, the received sequence is a function of not only the current transmitted symbol, but also components from the earlier transmitted symbols. This phenomenon is known as ISI, as we have introduced earlier. When the FIR filter is of length one, the system is called *flat fading*. Alternatively, the channel with delay spread is termed *memory* channel, and *memoryless* channel otherwise (with no channel delay spread).

Assumption 3. *The channel impulse response is known at the receiver.*

In typical communication systems, the channel response is not known at the receiver. Most burst based communication systems provide a training segment in the data burst. A priori knowledge of the training sequence is used at the receiver to estimate the channel response. The accuracy of the channel estimation is dependent on the length of the training sequence and the noise variance. However, we assume perfect knowledge of the channel at the receiver.

It is also well known that the feature of the channel affects the performance of the communication systems. For the purpose of efficient evaluation and design, the fundamental characteristics about the MIMO channel need be better understood:

A. *Different assumptions about the amount of channel side information available at the transmitter (Based on Assumption 3 of perfect* channel state information (CSI) *at the receiver).*

- CSI is known perfectly at the transmitter. This can be achieved when a closed loop system with feedback channel is present, or when the transmitter and receiver operate in time division duplex (TDD) where the two links (downlink and uplink) share the same frequency band so that the channel impulse response can be estimated from the uplink for downlink processing [43].

- Imperfect CSI (partial CSI) at the transmitter. This happens in the frequency division duplex (FDD) systems where the two links use different frequency bands and therefore, the impulse responses are also different. However, there commonly exists a feedback channel from the receiver to the transmitter that provides the transmitter with some partial CSI.

- No CSI at the transmitter. Without knowing the channels information, the transmitter must maintain a strategy based on knowledge of the channel statistics, which include the full channel distribution, or just its mean and variance.

B. *Different assumptions about the channel correlation.*

The high spectral efficiencies associated with MIMO channels are based on the premise that a rich scattering environment provides independent transmission paths from each transmit antenna to each receive antenna. In a realistic MIMO commu-

nication environment, the different spatial channels between the antennas will be correlated in space. For example when the multiple antennas at a portable device have small spacing, the signals received at different antennas will be correlated, and the performance will degrade. Although correlations sometimes increase the system performance and sometimes decrease, throughout this dissertation we will assume that the channels are independent from each other.

C. *Different assumptions about the channel noise.*

It is a common assumption in the literature that the channel corrupting noise is white, with known power spectral; while there are cases that the non-white disturbance noise with unknown spectral exists, for example, colored noise at the receiver incorporates cross-talk, inter-channel interference, and residual echos [41]. This dissertation will tackle both situations in their respective communication system frameworks for the optimal channel equalization design.

1.2 Equalization and Precoding

As discussed earlier, multi-path propagation in band-limited time dispersive channels causes ISI, which has been recognized as the major obstacle in achieving increased digital transmission rates with the required accuracy. Briefly ISI means that the transmitted pulses are smeared out so that pulses that correspond to different symbols are not separable. Meanwhile, data received from a desired user may be disturbed by other transmitters, due to imperfections in the multiple access scheme, giving rise to ICI. Obviously for a reliable digital transmission system, it is crucial to reduce

the effects of ISI and ICI. This problem can be resolved by channel equalization in which the aim is to construct an equalizer such that the impulse response of the channel/equalizer combination is as close to a pure delay as possible. Figure 1.1 is a block diagram for the baseband equivalent communication system with channel equalization, where the equalizer is designed to reverse the effect of the ISI/ICI channel.

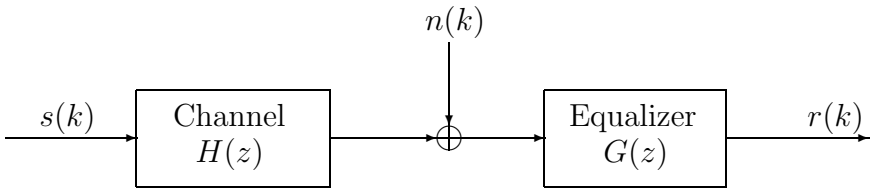

Figure 1.1: Baseband equivalent communication system with equalizer

Channel equalization has received great attention for the past two decades, and it has received renewed interest in the research community due to the need for wireless communications, where the channels are inherently distorted due to multipath and dispersion. See [6, 10, 14, 34, 41, 52, 54, 58], and the references therein. Among the research on channel equalization, filterbank approach as proposed in [14, 41] is of particular interest, which encompasses existing modulations and equalization schemes, as well as general channel identifiability conditions, leading to possible optimal design of transmitter/receiver due to the redundancy, and block transmission introduced by filterbanks. While minimization of BER is most desirable, it is the minimum mean-squared error (MMSE) design for symbol detection among all chan-

nel equalizers that attracts the most attention in the research community due to its mathematically tractable nature. Indeed optimal channel equalization in the sense of MMSE has been studied for its applications in various data networks, such as DSL, FDMA, as well as DS/SS-CDMA systems [5, 31, 32, 40, 41]. Although the results in [41] cover only a limited class of channels having the drawback of high complexity, the filterbank approach shows the promise for its future role in optimization for channel equalization, and minimization of the mean-squared error.

For the MIMO system employing multiple antennas at both the transmitter site and receiver site, channel equalization has also been studied extensively as a method of combating ISI and ICI to improve the received signal quality in digital communications. Optimal designs for the MIMO system developed in the past such as [29, 43, 39, 63, 65, 66] attract attentions because of the interest in joint transmit-receive diversity schemes. Other results have been previously obtained such as [26] where the optimal equalizer/precoder was derived independently by using Bezout theory for MIMO system, hence reduced the hardware implementation compared to the joint design. Recently, there has been investigations regarding block precoding transmission technique to shift the signal processing burden from the receiver to transmitter [40, 63]. In this case, a precoding filter (precoder) is designed at the transmitter side and is applied prior to transmission to equalize the signal at the output of the receive filter. In fact, the concept of precoding is an old idea dating back to the early work on the pre-equalizer at the transmitter [10, 20, 50]. This is

particularly of interest in the downlink mobile radio channel where it is desirable to keep the receiver units as simple as possible. Precoding is used in transmitters to compensate for distortion introduced by the channel response, and to reduce the effects of ISI and ICI, and thus allows for more reliable data transmission at high data rates. Precoder design is expected to allow the receiver to be considerably simplified, which in turn, reduces computational complexity and power consumption. Designing the appropriate precoding strategy has been studied under variety of criteria. See [39] and the references therein. Some results on the precoding technique for CDMA communication systems can be found in [4, 38, 57, 62].

1.3 Dissertation Contributions

This dissertation is a continuation of the existing research in channel equalization. The first part of this dissertation focuses on the optimal equalizer design for the filterbank transceivers which minimizes the detection error variance in the worst-case. The second part focuses on the optimal precoder design for the MIMO system employing multiple transmit antennas and receiver antennas to minimize the BER associated with symbol detection. The contributions of this dissertation are as follows.

- Among all the existing work for the channel equalization under the multirate filterbank transceiver framework, many results have focused on the white noise, and design of FIR channel equalizers. However, the problem of optimal channel equalization remained open for the non-white case, which will be investigated

in this dissertation. No knowledge on the non-white noise is assumed, except that its power-norm or RMS value is bounded. Consequently MMSE design method is not applicable, especially for the case that the non-white noise has a narrow-band feature. A worst-case approach is adopted to optimal channel equalization, leading to $\mathcal{H}_\infty$ optimization [17, 69], which has been well studied in the control area. However the results on discrete-time $\mathcal{H}_\infty$ optimization as in [22] can not be used directly, because of its complex form, which is very different from its continuous-time counterpart. Hence a different approach needs be taken. All channel equalizers, or zero-forcing receiving filterbanks will first be parameterized in an affine form with the free parameter over the set of linear, causal, stable, and time-invariant systems, and $\mathcal{H}_\infty$ optimization will then be carried out over the set of all zero-forcing receiving filterbanks. An explicit design algorithm for suboptimal channel equalizers with the performance index arbitrarily close to the optimal one is derived and a simulation example is provided to illustrate the proposed optimal channel equalization technique.

- For the MIMO system employing multiple transmit antennas and receiver antennas, much research about the channel equalization has been carried out to design optimal precoders and/or equalizers based on minimum mean-squared error (MMSE), maximum information rate, and etc. However, it is the BER in transmission with which users are most concerned. In this dissertation, the problem of designing optimal channel precoders which minimize the BER under

the PR/ZF condition will be addressed assuming the perfect channel knowledge. Analytic solution to the optimal precoder is derived, and an effective design algorithm is provided. The performance evaluations indicate that the optimal precoder design proposed in this dissertation is an attractive/reasonable alternative to the existing precoder design techniques for MIMO systems.

The first part of this dissertation is an extended investigation of our paper [19] accepted by *IEEE transactions on signal processing*. The second part of this dissertation is a more thorough version of a paper [30] submitted to the same journal that is currently under review. It is hoped that the results in this dissertation will contribute to the design technology for channel equalization and precoding in wireless communication.

1.4 Organization of the Dissertation

In the following contents, Chapter 2 presents some necessary preliminaries used in the later chapters along with the realistic communication models; Chapter 3 investigates the worst-case design for optimal channel equalization followed by the optimal precoder design in Chapter 4. Finally Chapter 5 concludes the dissertation by suggesting some possible future research. Throughout this dissertation, symbols for matrices are in capital letters, and vectors are in boldface small letters. Other notational conventions are listed in the earlier page of this dissertation titled as "Notations and Symbols".

Chapter 2

Preliminaries and MIMO Models

This chapter will first introduce some necessary preliminaries to be used in the later chapters. Since most of these results are available in standard textbooks, they are stated without citations and proofs. Typical MIMO models are then presented for different system configurations. More specifically we will study two types of systems. The first describes a MIMO structure based on the use of two banks of filters, called multirate filterbank transceiver. The optimal design for such a system under non-white noise will be discussed in Chapter 3. The second consists of multiple transmit antennas and multiple receive antennas for the single user transmission, forming a MIMO setup which will be used for the optimal design in Chapter 4.

2.1 Signals and Norms

There are two classes of stochastic signals which are of interest in this dissertation for the development of the optimal design problem. The first class is called *bounded power signal* and the second is the well-known *white noise signal*. The system norm can be defined in relation to these two classes of signals, leading to $\mathcal{H}_\infty$ norm and $\mathcal{H}_2$

norm, which will be used as the performance measures for the optimal design in the later chapters.

A norm, denoted by $\|\cdot\|$, is a real-valued function defined on some vector space X (of signals or systems). It satisfies the following properties:

1. $\|x\| \geq 0$
2. $\|x\| = 0$ if and only if $x = 0$
3. $\|\alpha x\| = |\alpha|\|x\|$, for any scalar α
4. $\|x + y\| \leq \|x\| + \|y\|$

for any $x \in X$ and $y \in X$. A real-valued functional $\|\cdot\|$ is called a semi-norm on X if it satisfies properties 1, 3 and 4 but not necessarily 2.

2.1.1 Bounded Power Signals and $\mathcal{H}_\infty$ Norm

Consider a complex-valued discrete-time vector random process $\mathbf{u}(k)$. Assume that it is wide sense stationary (WSS)[2] and has zero mean. Then its autocorrelation function is given as:

$$R_u(\tau) := E[\mathbf{u}(k)\mathbf{u}^*(k+\tau)] \tag{2.1}$$

where $E[\,\cdot\,]$ is the expectation. The power spectral density (PSD) of $\mathbf{u}(k)$ is defined via the Fourier transform pair:

$$S_u(f) = \sum_{\tau=-\infty}^{\infty} R_u(\tau)e^{-j2\pi f\tau}, \quad R_u(\tau) = \int_0^1 S_u(f)e^{j2\pi f\tau}\, df. \tag{2.2}$$

[2]If a random process has constant mean, and its auto-correlation function depends only on the time difference, then it is said to be *wide-sense stationary* (WSS).

We shall be interested in the set of signals $\mathbf{u}(k)$ for which both $R_u(\tau)$ and $S_u(f)$ exist. Often we are interested in the mean-squared value or variance of $\mathbf{u}(k)$, given by

$$E[\mathbf{u}^*(k)\mathbf{u}(k)] = \text{trace}\{E[\mathbf{u}(k)\mathbf{u}^*(k)]\} = \text{trace}\{R_u(0)\} = \int_0^1 \text{trace}\{S_u(f)\}\ df. \quad (2.3)$$

Suppose that the autocorrelation function $R_u(\tau)$ and the PSD $S_u(f)$ of $\mathbf{u}(k)$ exist. Then $\mathbf{u}(k)$ is said to have *bounded power* if

$$E[\mathbf{u}^*(k)\mathbf{u}(k)] < \infty. \quad (2.4)$$

The set of all random process signals having bounded power is denoted by

$$\mathcal{P} := \{\mathbf{u}(k) : \quad E[\mathbf{u}^*(k)\mathbf{u}(k)] < \infty\}. \quad (2.5)$$

A semi-norm, called the *power norm* of $\mathbf{u}(k)$, can be defined on $\mathcal{P}$ by

$$\|\mathbf{u}\|_{\mathcal{P}} := \sqrt{E[\mathbf{u}^*(k)\mathbf{u}(k)]} = \sqrt{\text{trace}\{R_u(0)\}} = \sqrt{\int_0^1 \text{trace}\{S_u(f)\}\ df}, \quad (2.6)$$

which is actually the RMS, or root mean-squared value of $\mathbf{u}(k)$.

Consider a linear, causal and time-invariant discrete-time stable system with transfer function matrix $G(z)$, which is analytic outside of the unit circle, driven by the random process $\mathbf{u}(k)$, with output $\mathbf{y}(k)$, shown in Figure 2.1.

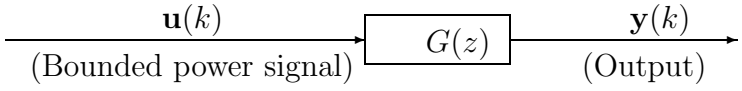

Figure 2.1: A stable system driven by bounded power signal

The input and output are described as

$$\mathbf{y}(k) = g(k) \star \mathbf{u}(k) = \sum_{i=-\infty}^{\infty} g(k-i)\mathbf{u}(i) = \sum_{i=-\infty}^{k} g(k-i)\mathbf{u}(i),$$

by the causality of the system, where $g(k)$ is the impulse response of $G(z)$, and $\star$ denotes the convolution operation. It follows that the output PSD $S_y(f)$ is given by

$$S_y(f) = G(e^{j2\pi f}) S_u(f) \left[G(e^{j2\pi f})\right]^* \tag{2.7}$$

where $S_y(f) = \mathcal{F}[R_y(\tau)] = \mathcal{F}(E[\mathbf{y}(k)\mathbf{y}^*(k+\tau)])$. When $\mathbf{u}(k)$ is a bounded power signal with no knowledge on the PSD, the power norm of the output has a complicated form. Because each input signal $\mathbf{u}(k) \in \mathcal{P}$ may have a different PSD, the ratio $\frac{\|\mathbf{y}\|_{\mathcal{P}}}{\|\mathbf{u}\|_{\mathcal{P}}}$ is not fixed unless its PSD is specified. That is, when the PSD of $\mathbf{u}(k)$ is not known at the receiver, the ratio $\frac{\|\mathbf{y}\|_{\mathcal{P}}}{\|\mathbf{u}\|_{\mathcal{P}}}$ is unknown, which can take any bounded value corresponding to any bounded power-norm input signal $\mathbf{u}(k)$. However, the supreme of this ratio is dependent only on the system, given by [69]

$$\sup_{\|\mathbf{u}\|_{\mathcal{P}} \neq 0} \frac{\|\mathbf{y}\|_{\mathcal{P}}}{\|\mathbf{u}\|_{\mathcal{P}}} = \|G\|_\infty := \sup_{|z|>1} \overline{\sigma}\left(G(z)\right) = \operatorname*{ess\,sup}_{f \in [0,\, 1)} \overline{\sigma}\left(G(e^{j2\pi f})\right), \tag{2.8}$$

where $\overline{\sigma}(\cdot)$ denotes the maximum singular value. Equation (2.8) defines the *induced power norm*, i.e., $\mathcal{H}_\infty$ norm, which is the *worst-case* measure of the power (norm) amplification (or the maximum energy amplification or gain) from the input to the output over all possible power norm bounded input random processes. It should be pointed out that these results are derived for stationary process. Actually it remains valid for quasi-stationary signals also. Such signals are stochastic process with deterministic components.

2.1.2 Gaussian White Noise Signals and $\mathcal{H}_2$ Norm

Many sources of noise signals in engineering are normally modeled as the well-known Gaussian white noise. Mathematically, a Gaussian white noise $\mathbf{u}(k)$ is a stationary random process that satisfies:

(1) $E\{\mathbf{u}(k)\} \equiv 0$,

(2) $R_u(\tau) = \frac{N_0}{2}\delta(\tau) = \begin{cases} \frac{N_0}{2}, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}$,

i.e., $\mathbf{u}(k)$ is a zero mean stationary process with power spectrum $\frac{N_0}{2}$. A more rigorous description of the white noise can be found in standard textbooks for stochastic processes (see, e.g., [61]).

White noise signals have a natural relation with the $\mathcal{H}_2$ norm of a stable system as shown in Figure 2.2:

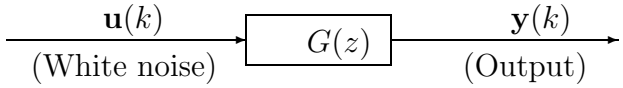

Figure 2.2: A stable system driven by white noise signal

For a system as in Figure 2.2, the power norm of the output signal $\mathbf{y}(k)$ can be computed according to the following formula:

$$\|\mathbf{y}\|_{\mathcal{P}} = \sqrt{\text{trace}\left\{\int_0^1 G(e^{j2\pi f})\,[G(e^{j2\pi f})]^*\,df\right\}}\sqrt{\frac{N_0}{2}}, \tag{2.9}$$

in light of (2.6) and (2.7). In this case, we have the $\mathcal{H}_2$ norm as:

$$\|G\|_2 := \sqrt{\text{trace}\left\{\int_0^1 G(e^{j2\pi f})\,[G(e^{j2\pi f})]^*\,df\right\}}, \tag{2.10}$$

which is a measure of the total output energy given the unit impulse excitation at each input channel one at a time.

2.2 MIMO Models

A MIMO system is typically characterized by the transmission of multiple input signals through a linear, dispersive, noisy channel and results in multiple output signals at the receiver. The received signals are composed of a summation of several transmitted signals corrupted by ISI, ICI, and noise. Examples of MIMO channels include the following:

(a) *Single-user Multiple Transmit and Receive Antennas.*

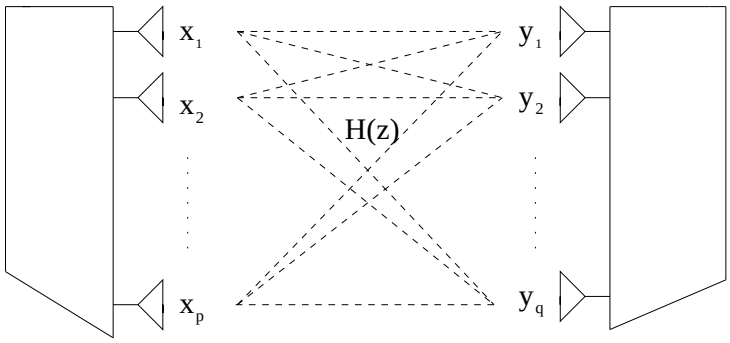

Figure 2.3: Multiple transmit and receive antennas communication system

Multiple transmit and receive antennas arise to achieve the goal of higher data rate for systems that are power, bandwidth, and complexity limited. Pioneering work by Foschini [11, 12] and Telatar [49] ignited much interests in this area by predicting remarkable spectral efficiency for wireless systems with multiple antennas when the

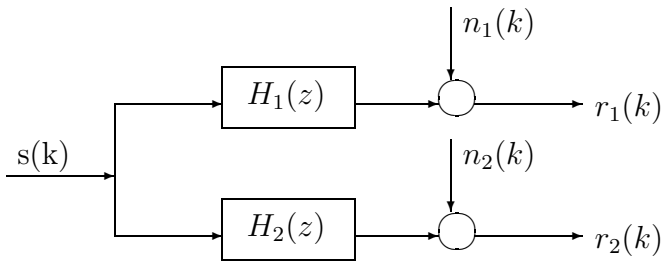

Figure 2.4: Broadcast channel

channel exhibits rich scattering and its variations can be accurately tracked. Figure 2.3 is a typical picture of a multiple transmit and receive antennas communication system.

(b) *Multi-user Channel.*

In a typical multi-user communication scenario with a single base station and many users in different locations, the broadcast channel means that a single transmitter sends independent information to multiple receivers from the downlink (from base station to mobile users) direction; vice versa, the multiple access channel means that multiple transmitters send independent information to a single receiver from the uplink (from mobile users to base station) direction. The simplified system diagrams with two transmitters/receivers for each individual case can be found in Figure 2.4 and Figure 2.5.

(c) *Filterbank Transceivers, OFDM, and CDMA.*

Filterbank transceivers (transmultiplexers) have been studied for the past years in the signal processing community, due to their capability for block transmission, and

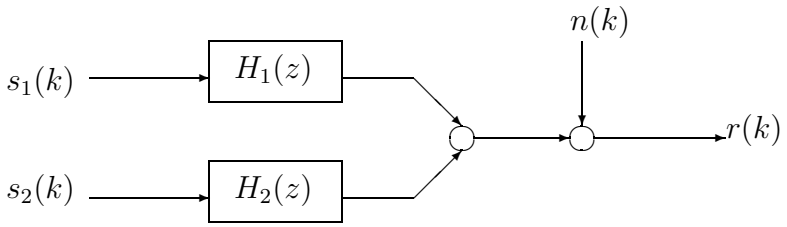

Figure 2.5: Multi-access channel

use of redundancies for precoding, enabling complete elimination of ISI, which is one of the major obstacles for data communications. The filterbank transceiver model was proposed in [41], which introduces transmitter redundancy using filterbank precoders, and generalizes existing modulations in wireless communications. Figure 2.6 shows the discrete-time multirate filterbank model for the baseband communication system.

It is seen from Figure 2.6 that it consists of the transmitter filterbank $\{F_m(z)\}_{m=0}^{M-1}$, and the receiver filterbank $\{G_p(z)\}_{p=0}^{P-1}$, each having down-samplers and up-samplers. The communication channel is represented by the transfer function $H(z)$, which is assumed to be causal and stable. The corrupting noise at the output of the channel is assumed to be stationary, and non-white for our design. It was demonstrated in [41] that the filterbank transceiver provides a useful framework that unifies modulation, precoding, and equalization. In Chapter 3, we will focus on the optimal equalizer design for the filterbanks, that achieve perfect reconstruction (PR) and minimization of $\mathcal{H}_\infty$ norm subject to the non-white noise with unknown PSD.

A mutual influence and interaction have been recently recognized between the fields of multirate signal processing and digital communication. Some of the popular

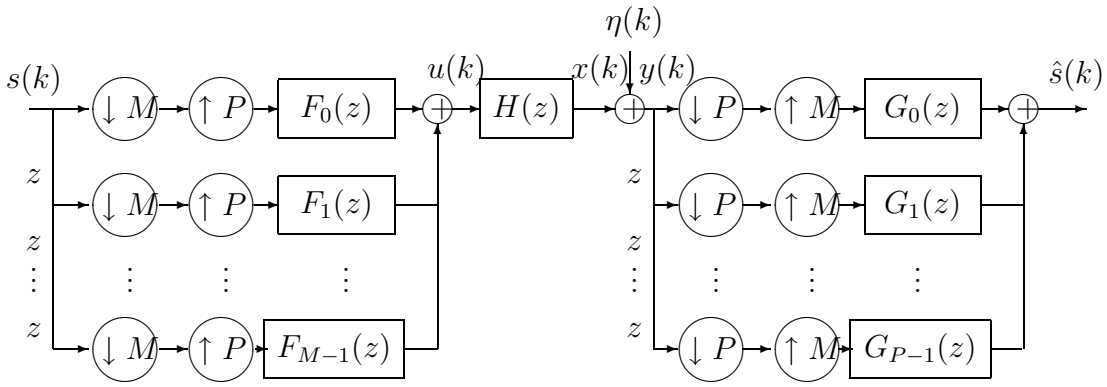

Figure 2.6: Base-band equivalent multirate filterbank model

communication applications can be described in terms of filterbank configurations of subband transforms. Such applications include the following.

- Orthogonal frequency division multiplexing (OFDM).

  OFDM has been widely used in wireless communications. See Figure 2.7 for the illustrations. In OFDM, individual channels are assigned to individual users. Each user is allocated an unique frequency band or channel. These channels are assigned on demand to users who request service.

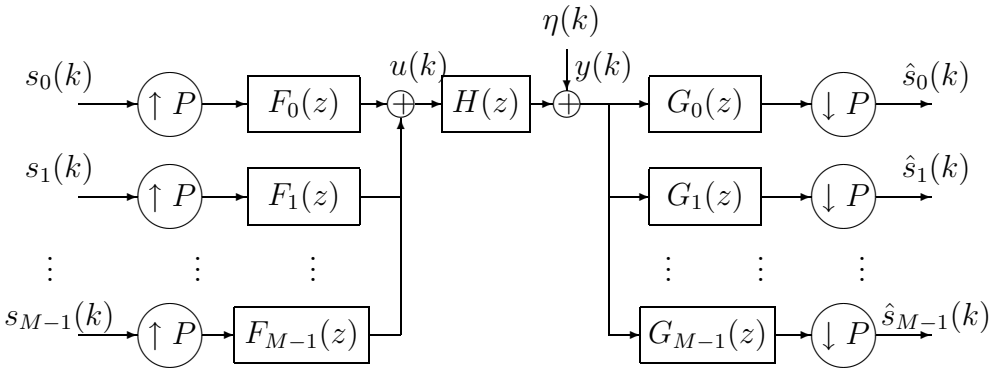

Figure 2.7: Orthogonal frequency division multiplexing (OFDM) system

- Code Division Multiple Access (CDMA)

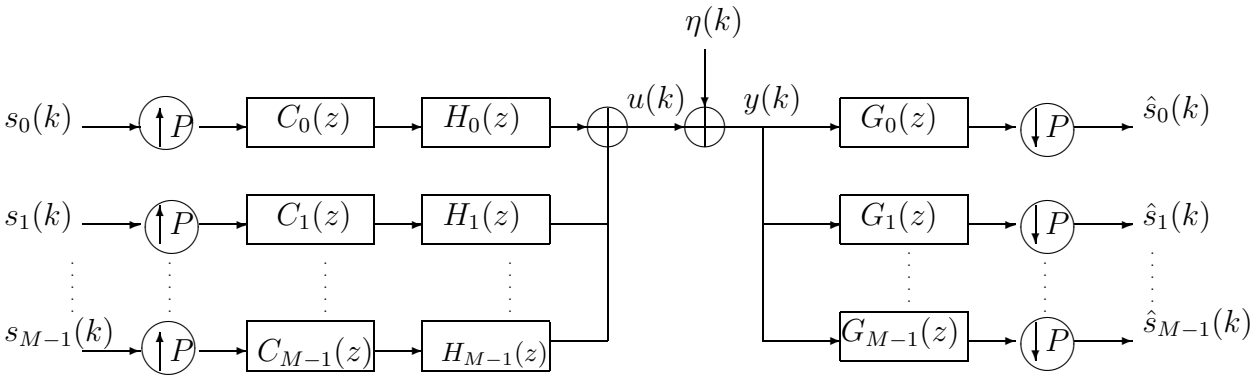

Figure 2.8: The equivalent MIMO model for the CDMA channel

In the CDMA communication system shown in Figure 2.8, the narrowband information signal is multiplied by a very large bandwidth signal called spreading signal. The spreading signal is a pseudo-noise code sequence based on a chip rate which is orders of magnitudes greater than the data rate of the message. All users in a CDMA system use the same carrier frequency, and may transmit simultaneously. Each user has its own pseudo-random codeword, which is approximately orthogonal to all other codewords. The receiver performs a time correlation operation to detect only the specific desired codeword.

In the next chapter, we will show that the MIMO models in Figure 2.6 and Figure 2.7 are equivalent to each other, by mapping the polyphase components of the input signal in Figure 2.6 to different users' signals.

Chapter 3

Worst-case Design for Optimal Channel Equalization

In this chapter, a worst-case approach is adopted to tackle optimal channel equalization under the framework of multirate filterbank transceivers which is widely used in data communication networks. It is assumed that in such applications the observation noise is non-white with bounded power-norm or root-mean-squared (RMS) value. The objective is to design the optimal receiving filterbanks which not only achieve the zero-forcing (ZF) condition or channel equalization, but also minimize the RMS error between the transmitted symbols and the received symbols in the presence of the worst-case non-white noise. All zero-forcing receiving filterbanks are parameterized, and optimal design for channel equalization are converted into an equivalent optimal $\mathcal{H}_\infty$ filtering problem for the augmented receiving filterbanks with RMS error preserved. The main results in this chapter cover computation of the optimal RMS error achievable for the worst-case noise, and an explicit design algorithm for suboptimal channel equalizers with the performance index arbitrarily close to the optimal one.

A simulation example is used to illustrate the proposed optimal channel equalization design technique.

3.1 Background Materials

This section introduces some standard techniques used in this chapter to treat multirate systems.

- Sampling Rate Conversion

  When dealing with signals having different sampling rate, sampling rate conversion is needed. To achieve sampling rate conversion for discrete-time data streams in multirate signal processing, two basic devices are used. One is called down-sampler (decimator, or compressor), and the other is called up-sampler (interpolator, or expander). A down-sampler by M and an up-sampler by P are shown in Figure 3.1 where M and P are integers. Passing an input data stream $s(k)$ through a down-sampler by M produces an output signal $x_d(k)$ with $x_d(k) = s(Mk)$. In the $\mathcal{Z}$-transform domain,

$$X_d(z) = \frac{1}{M} \sum_{n=0}^{M-1} S(z^{1/M} W_M^n), \quad W_M = e^{-j2\pi/M}, \quad j = \sqrt{-1} \tag{3.1}$$

  The down-sampler retains only those samples of $s(k)$ with $k = nM$ and $n = 0, 1, 2, \cdots$. The resulting signal $x_d(k)$ has a sampling rate $1/M$ of that of $s(k)$. A properly designed lowpass filter is usually used before the down-sampling process to eliminate the aliasing. Otherwise it may not be possible to recover $s(k)$ from $x_d(k)$ because of the loss of information due to the aliasing.

Figure 3.1: A down-sampler by M and an up-sampler by P

Passing an input data stream $s(k)$ through an up-sampler by P produces an output signal $x_u(k)$ where

$$x_u(k) = \begin{cases} s(k/P), & k = nP, \quad n = 0, 1, 2, \cdots \\ 0, & k \neq nP, \end{cases} \tag{3.2}$$

which has a sampling rate P times that of the original signal $s(k)$. In $\mathcal{Z}$-transform domain,

$$X_u(z) = S(z^P). \tag{3.3}$$

The input signal $s(k)$ can be completely recovered from $x_u(k)$. In practice, a properly designed lowpass filter is used after the up-sampling process to remove the unwanted images in the spectrum of the up-sampled signal, which was caused by the up-sampling process.

- Blocking Signals

  The standard technique in signal processing for treating periodic/multi-rate systems is called *blocking*. Let ℓ_+ be the space of causal and discrete-time signals defined on the time index set $\{0, 1, 2, \cdots\}$. A signal s in ℓ_+ can be written as

  $$s = \{s(0), s(1), s(2), \cdots\} = \{s(k)\}_{k=0}^{\infty}.$$

  For an integer $M > 0$, define the M-fold blocking operator, L_M, via $\underline{\mathbf{s}} = L_M s$ (where underlining denotes the blocked signal). Blocking a signal $\{s(k)\}_{k=0}^{\infty}$ by

a factor M converts the serial data stream into M parallel substreams $\{s_m(k) := s(kM+m)\}_{k=0}^{\infty}$ for $m = 0, 1, \cdots, M-1$. Thus $s_m(k)$ is the mth symbol in the kth block of symbols, yielding a blocked signal $\{\underline{\mathbf{s}}(k)\}_{k=0}^{\infty}$:

$$\begin{aligned} s &= \{s(0), s(1), s(2), \cdots\} \longmapsto \\ \underline{\mathbf{s}} &= \{\underline{\mathbf{s}}(0), \underline{\mathbf{s}}(1), \underline{\mathbf{s}}(2), \cdots\} \\ &= \left\{ \begin{bmatrix} s(0) \\ s(1) \\ \vdots \\ s(M-1) \end{bmatrix}, \begin{bmatrix} s(M) \\ s(M+1) \\ \vdots \\ s(2M-1) \end{bmatrix}, \cdots \right\} \end{aligned} \tag{3.4}$$

If the underlying sampling period for s is T_s, then the underlying sampling period for $\underline{\mathbf{s}}$ is MT_s. The blocking operation results in no loss of information.

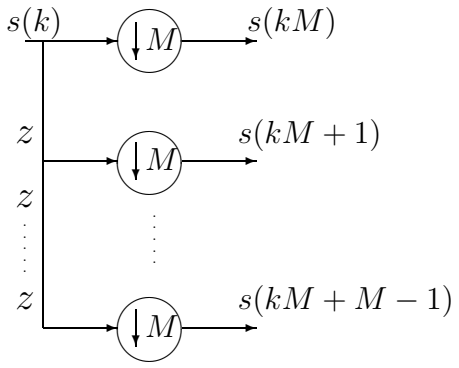

Figure 3.2: Blocking a signal

The blocking operation can be achieved using unit advance elements and down-samplers by the blocking factor M as in Figure 3.2. In reality the unit advance element cannot be implemented but blocking operation can still be achieved by using unit delay elements which is implementable.

- Blocked Signals and Polyphase Decomposition

  Applying the blocking operator L_M to a signal $\{s(k)\}_{k=0}^{\infty}$ yields a blocked signal $\underline{\mathbf{s}} = \{\underline{\mathbf{s}}(k)\}_{k=0}^{\infty}$. In the frequency domain, bring in the M-fold polyphase decomposition for the $\mathcal{Z}$-transform of $s(k)$:

$$S(z) = \mathcal{Z}\{s\} = \sum_{k=0}^{\infty} s(k)z^{-k} = \sum_{m=0}^{M-1} z^{-m} S_m(z^M), \quad S_m(z) = \sum_{k=0}^{\infty} s(kM+m)z^{-k}. \tag{3.5}$$

  The representation is unique for the given integer M. It follows that the mth polyphase component $S_m(z)$ is the $\mathcal{Z}$-transform of $s_m(k) = s(kM+m)$, and there holds

$$\underline{S}(z) = \sum_{k=0}^{\infty} \underline{s}(k)z^{-k} = [S_0(z)\ S_1(z)\ \cdots\ S_{M-1}(z)]^*. \tag{3.6}$$

3.2 Channel Equalization in Face of the Worst-Case Noise

Channel equalization has been effective in coping with ISI in multirate filterbank transmission. A typical filterbank transceiver in data communications is the multirate transmultiplexer in Figure 3.3 (which is the same as Figure 2.7, redrawn here to make this chapter self-contained), widely used in FDMA, DSL and DWMT [7, 25, 27, 31, 32].

On the transmitter side, the M input data sequences $\{s_i(k)\}_{i=0}^{M-1}$ are up-sampled by P, filtered or frequency shaped, and then transmitted through the distorted channel. At the receiver end, the received signal is contaminated by additive noise $\eta(k)$,

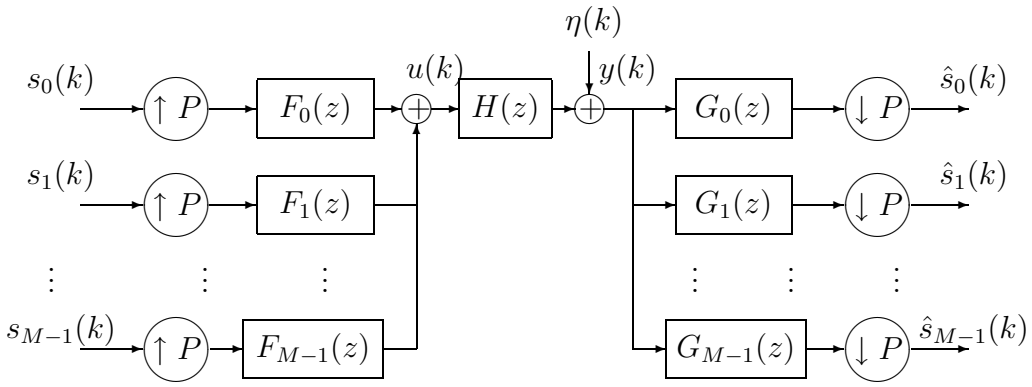

Figure 3.3: Transmultiplexer model with $P > M$

which is filtered, and then down-sampled, in hope of recovering the transmitted data sequences. It is assumed that $P > M$, which introduces the redundancies, enabling perfect reconstruction (PR) or zero-forcing (ZF) in absence of noise. The advantages of using such multirate transmultiplexers lie in the fact that the channel can be divided into M subbands, and M transmitter filters can be designed to confine the modulated signals within their respective subbands, so that the transmission power and bits can be judiciously allocated according to the SNR in each band to improve the BER performance. Even though DS-CDMA network has no relation to multirate transmultiplexers, it is also a multirate system due to the use of spreading sequences at a much higher rate than the symbol rate. In fact, [52] has obtained an equivalent multirate filterbank transceiver for DS-CDMA system, similar to the one in Figure 3.3, despite that the frequency domain interpretation such as subband for transmultiplexer is lost. For this reason it is not assumed in this dissertation that the frequency channel is separated into subbands for each transmitted signal sequence.

For the model in Figure 3.3, denote the input and output vector signals as follows:

$$\underline{\mathbf{s}}(k) = \left[\begin{array}{cccc} s_0^*(k) & s_1^*(k) & \cdots & s_{M-1}^*(k) \end{array} \right]^*, \quad \hat{\underline{\mathbf{s}}}(k) = \left[\begin{array}{cccc} \hat{s}_0^*(k) & \hat{s}_1^*(k) & \cdots & \hat{s}_{M-1}^*(k) \end{array} \right]^*. \tag{3.7}$$

Block the input and output of the channel with size P as follows:

$$\underline{\mathbf{u}}(k) = \left[\begin{array}{c} u(kP) \\ \vdots \\ u(kP+P-1) \end{array} \right], \quad \underline{\eta}(k) = \left[\begin{array}{c} \eta(kP) \\ \vdots \\ \eta(kP+P-1) \end{array} \right], \quad \underline{\mathbf{y}}(k) = \left[\begin{array}{c} y(kP) \\ \vdots \\ y(kP+P-1) \end{array} \right]. \tag{3.8}$$

Then $\underline{\mathbf{u}}(k)$, $\underline{\eta}(k)$, and $\underline{\mathbf{y}}(k)$ have the same sampling rate as the symbol rate for $\{s_i(k)\}_{i=0}^{M-1}$. Let

$$F_k(z) = \sum_{i=0}^{P-1} z^{-i} F_{k,i}(z^P), \qquad G_k(z) = \sum_{i=0}^{P-1} z^{-i} G_{k,i}(z^P), \tag{3.9}$$

be polyphase decomposition of $F_k(z)$ and $G_k(z)$ [55], respectively for $0 \leq k < M$. Let

$$H(z) = \sum_{i=0}^{P-1} z^{-i} H_i(z^P) \tag{3.10}$$

be the polyphase decomposition of $H(z)$. Then there holds

$$\hat{\mathbf{s}}(k) = \underline{\mathcal{G}}(k) \star \underline{\mathcal{H}}(k) \star \underline{\mathcal{F}}(k) \star \underline{\mathbf{s}}(k) + \underline{\mathcal{G}}(k) \star \underline{\eta}(k), \tag{3.11}$$

where $\underline{\mathcal{F}}(k), \underline{\mathcal{G}}(k)$ and $\underline{\mathcal{H}}(k)$ are the impulse responses of the blocked filters given by [55]:

$$\underline{F}(z) \;=\; \mathcal{Z}[\underline{\mathcal{F}}(k)] = \left[\begin{array}{cccc} F_{0,0}(z) & F_{1,0}(z) & \cdots & F_{M-1,0}(z) \\ F_{0,1}(z) & F_{1,1}(z) & \cdots & F_{M-1,1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ F_{0,P-1}(z) & F_{1,P-1}(z) & \cdots & F_{M-1,P-1}(z) \end{array} \right], \tag{3.12}$$

$$\underline{G}(z) = \mathcal{Z}[\underline{\mathcal{G}}(k)] = \begin{bmatrix} G_{0,0}(z) & z^{-1}G_{0,1}(z) & \cdots & z^{-1}G_{0,P-1}(z) \\ G_{1,0}(z) & z^{-1}G_{1,1}(z) & \cdots & z^{-1}G_{1,P-1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ G_{M-1,0}(z) & z^{-1}G_{M-1,1}(z) & \cdots & z^{-1}G_{M-1,P-1}(z) \end{bmatrix} \quad (3.13)$$

$$\underline{H}(z) = \mathcal{Z}[\underline{\mathcal{H}}(k)] = \begin{bmatrix} H_0(z) & z^{-1}H_{P-1}(z) & \cdots & z^{-1}H_1(z) \\ H_1(z) & H_0(z) & \cdots & z^{-1}H_2(z) \\ \vdots & \vdots & \cdots & \vdots \\ H_{P-1}(z) & H_{P-2}(z) & \cdots & H_0(z) \end{bmatrix}. \quad (3.14)$$

Channel equalization requires that the PR condition

$$\underline{G}(z)\underline{H}(z)\underline{F}(z) = \text{diag}\left(z^{-d_0}, z^{-d_1}, \cdots, z^{-d_{M-1}}\right) \quad (3.15)$$

holds true for some nonnegative integers d_i with $0 \leq i < M$. That is, for the noise-free case, the PR condition is equivalent to

$$\hat{s}_i(k) = s_i(k - d_i), \qquad 0 \leq i < M. \quad (3.16)$$

Optimal channel equalization requires not only (3.15) to hold, but also to minimize the noise impact at the receiver side in terms of the mean-squared error.

In this dissertation it is assumed that $\underline{H}(z)$, and $\underline{F}(z)$ are given, because the dynamical behavior of the channel cannot be altered, and the transmitter filters are often designed to shape the transmitted (base-band) signals. On the other hand, the fact $P > M$ provides redundancies which can be used for precoding to reduce the transmission error probability. More importantly, if (3.15) holds for some causal and stable $\underline{G}(z)$, more than one such $\underline{G}(z)$ exists such that (3.15) is true. Our objective is to synthesize an optimal receiving filterbank, among all possible $\underline{G}(z)$ satisfying

(3.15), such that the RMS error

$$J(\underline{G}) = \sqrt{\text{trace}\left\{\int_0^1 \underline{G}(e^{j2\pi f}) S_{\underline{\eta}}(f)\, [\underline{G}(e^{j2\pi f})]^*\ df\right\}}, \tag{3.17}$$

is minimized. If the observation noise $\eta(k)$ is white, then $S_{\underline{\eta}}(f)$ is a constant positive definite (diagonal) matrix. In this case, optimal channel equalization in the sense of minimization of $J(\underline{G})$, subject to the PR condition (3.15) is solved in [18], which is an equivalent Kalman filtering problem. But if the observation noise is not white with an unknown PSD function, then the optimal channel equalization problem becomes more complicated. It is assumed that no information on the PSD of $\underline{\eta}(k)$ is available at the receiver except that

$$S_{\underline{\eta}}(f) \in \mathcal{S}_\Gamma := \left\{\Gamma\Omega(f)\Gamma^* :\ \ \Omega(f) = \Omega^*(f) \geq 0\ \ \forall f,\ \ \text{trace}\left\{\int_0^1 \Omega(f)\ df\right\} \leq 1\right\}, \tag{3.18}$$

where ≥ 0 means positive semi-definite. The subset $\mathcal{S}_\Gamma$ provides the structural information through the full column rank $P \times r$ matrix Γ about $S_{\underline{\eta}}(f)$, while $\Omega(f)$ represents the unknown portion of the PSD function of the blocked observation noise $\underline{\eta}(k)$. Since $R_{\underline{\eta}}(0) = E[\eta(k)\eta^*(k)] = \int_0^1 S_{\underline{\eta}}(f)\ df$, it is most probable that $r = P$. Otherwise there exists a full row rank matrix Θ of size $(P-r) \times P$ such that $\Theta\eta(k) = 0$ for all k, which is not likely to occur. The worst-case approach will be adopted with aim to minimize

$$\sup_{S_{\underline{\eta}}(f)\in\mathcal{S}_\Gamma} J(\underline{G}) = \sup_{S_{\underline{\eta}}(f)\in\mathcal{S}_\Gamma} \sqrt{\text{trace}\left\{\int_0^1 \underline{G}(e^{j2\pi f}) S_{\underline{\eta}}(f) [\underline{G}(e^{j2\pi f})]^*\ df\right\}} = \|\underline{G}\Gamma\|_\infty, \tag{3.19}$$

subject to the PR condition (3.15), in light of (2.8), and (3.18).

3.3 Parameterization of All Channel Equalizers

As discussed in the previous section, our goal is: Given $\underline{H}(z)$ and $\underline{F}(z)$, design a suboptimal channel equalizer such that $\|\underline{G}\Gamma\|_\infty$ is arbitrarily close to an optimal (minimum possible) value, subject to the PR condition (3.15). Such a problem is termed as $\mathcal{H}_\infty$ optimization. Unfortunately the existing results on $\mathcal{H}_\infty$ optimization cannot be used directly. The design problem will be converted into an equivalent $\mathcal{H}_\infty$ filtering problem as in [64], and derive a suboptimal solution. It should be mentioned that the optimal solution to the worst-case channel equalization does not exist, in general. Hence suboptimal channel equalizers will be sought, and the design algorithm will be developed.

It is noted that the PR condition (3.15) can be equivalently written as

$$\underline{G}(z)T(z) \equiv I_M, \qquad T(z) = \underline{H}(z)\underline{F}(z)\text{diag}\left(z^{d_0}, z^{d_1}, \cdots, z^{d_{M-1}}\right), \tag{3.20}$$

with I_s the identity matrix of $s \times s$. We assume temporarily that the nonnegative integers $\{d_i\}_{i=0}^{M-1}$ are chosen such that

$$\text{rank}\left[D = \lim_{z\to\infty} \underline{H}(z)\underline{F}(z)\text{diag}\left(z^{d_0}, z^{d_1}, \cdots, z^{d_{M-1}}\right)\right] = M. \tag{3.21}$$

It follows that the integer $d_i \geq 0$ represents the propagation delay from the ith user to the receiver, which can be determined by sending some training signals (from the ith user) known to the receiver. See also [45] and [53] for the related work on estimation of the propagation delays. The equivalent PR condition in (3.20) requires that $\underline{G}(z)$ be a causal and stable left inverse of $T(z)$, and optimal channel equalization in the

MMSE sense requires that $\|\underline{G}\Gamma\|_\infty$ be minimized over all causal and stable left inverses of $T(z)$. It turns out that the full column rank of the D matrix is critical. In the following we will show that in case that (3.21) is violated, we can find a different $T(z)$ from that in (3.20) such that the full rank condition of $T(\infty)$ holds.

Proposition 3.1 *Suppose that the causal and stable transfer function matrix $\underline{H}(z)\underline{F}(z)$ of size $P \times M$ has normal rank M. Then there exists a square polynomial matrix $V(z)$ of size $M \times M$ which is allpass in the sense that $[V(z)]^*V(z) = V(z)[V(z)]^* = I_M \ \forall |z| = 1$, such that*

$$\text{rank}\left\{D = \lim_{z\to\infty} \underline{H}(z)\underline{F}(z)V(z)\right\} = M. \tag{3.22}$$

Proof: Without loss of generality we begin with

$$T^{(0)}(z) = \underline{H}(z)\underline{F}(z)z^{\beta_0} = \sum_{k=0}^{\infty} H_f^{(0)}(k)z^{-k}, \quad \text{rank}[H_f^{(0)}(0)] = m,$$

where $0 < m < M$ and $\beta_0 \geq 0$. The constant nonzero matrix $H_f^{(0)}(0) = T^{(0)}(\infty)$ is called the direct transmission term of $T^{(0)}(z)$. By the singular value decomposition (SVD), $H_f(0) = U_0\Sigma_0V_0^*$ where U_0 of size $P\times P$ and V_0 of size $M\times M$ are unitary, and Σ_0 of size $P \times M$ has m nonzero singular values on its diagonal. Set $V_0(z) = z^{\beta_0}V_0$, and

$$T^{(1)}(z) = U_0^*T^{(0)}(z)V_0(z) = \Sigma_0 + \sum_{k=1}^{\infty} H_f^{(1)}(k)z^{-k}, \quad \Sigma_0 = \begin{bmatrix} R_0 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.23}$$

where R_0 is diagonal of size $m \times m$, and nonsingular. Then $V_0(z)$ is a polynomial matrix, and allpass. Denote the $(m+1)th$ column of $T^{(1)}(z)$ by $T_{m+1}^{(1)}(z)$. Then

$$T_{m+1}^{(1)}(z) = \sum_{k=\beta_1}^{\infty} H_{f,m+1}^{(1)}(k)z^{-k}, \quad H_{f,m+1}^{(1)}(\beta_1) \neq 0.$$

We call β_1 the relative degree of $T^{(1)}_{m+1}(z)$. The direct transmission term of

$$T^{(1)}_{m+1}(z)\text{diag}(I_m, z^{\beta_1}, I_s) \quad (s = M - m - 1)$$

has the expression with the SVD as follows:

$$\begin{bmatrix} R_0 & H^{(1)}_{f,m+1}(\beta_1) & 0 \\ 0 & & 0 \end{bmatrix} = U_1 \Sigma_1 V_1^*.$$

If Σ_1 has rank $m+1$, then we can proceed for the $(m+2)$th column of

$$T^{(1)}_{m+1}(z)\text{diag}(I_m, z^{\beta_1}, I_{M-m-1}).$$

If not, then we set $V_1(z) = \text{diag}(I_m, z^{\beta_1}, I_{M-m-1})V_1$, and

$$T^{(2)}(z) = U_1^* T^{(1)}(z) V_1(z) = \Sigma_1 + \sum_{k=1}^{\infty} H^{(1)}_f(k) z^{-k}, \quad \Sigma_1 = \begin{bmatrix} R_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where R_1 is diagonal with size $m \times m$, and is nonsingular. Hence we come back to the same form as in (3.23). Now applying the same procedure as the above, which can be repeated until the direct transmission term of $T^{(\ell_1)}(z)$ has rank $m+1$ for some $\ell_1 > 1$. We claim that $\ell_1 > 1$ is finite. Indeed by the above procedure, we have

$$U^{(\ell_1)} T^{(\ell_1)}(z) = \underline{H}(z)\underline{F}(z) V^{(\ell_1)}(z), \quad \det\left(V^{(\ell_1)}(z)\right) = z^{M\beta_0 + \beta_1 + \cdots + \beta_{\ell_1 - 1}},$$

where $U^{(\ell_1)} = U_0 \cdots U_{\ell_1 - 1}$ and $V^{(\ell_1)}(z) = V_0(z) \cdots V_{\ell_1 - 1}(z)$. The hypothesis of full normal rank for $\underline{H}(z)\underline{F}(z)$ implies that all column vectors of $\underline{H}(z)\underline{F}(z)$ are linearly independent. Thus at least M rows of $\underline{H}(z)\underline{F}(z)$ exist, denoted by $H_f(z)$, whose determinant has a finite relative degree. If ℓ_1 is unbounded, then $\det\left(H_f(z)V^{(\ell_1)}(z)\right)$ becomes improper, implying that $U^{(\ell_1)}T^{(\ell_1)}(z)$ is improper, which is a contradiction to

the procedures repeatedly applied for ℓ_1 times. This concludes the fact that we need only finite steps to achieve rank increase to $m+1$ for the direct transmission term of $\underline{H}(z)\underline{F}(z)V^{(\ell_1)}(z)$. Moreover the degree of $\det\left(V^{\ell_1}(z)\right)$ is clearly bounded by the relative degree of $\det\left(H_f(z)\right)$, which is dependent on the propagation delays from the transmitter to the receiver. Note also that $V^{(\ell_1)}(z)$ is a polynomial matrix of z, which admits the allpass property. By induction, a unitary matrix $U = U^{(\ell_{M-m})}\cdots U^{(\ell_1)}$, and an allpass polynomial matrix $V(z) = V^{(\ell_{M-m})}(z)\cdots V^{(\ell_1)}(z)$ exist such that

$$T(z) = UT^{(\ell_1+\cdots+\ell_{M-m})} = \underline{H}(z)\underline{F}(z)V(z) \tag{3.24}$$

has a direct transmission term, which is of full column rank equal to M, thus concluding (3.22). ∎

Remark 3.1 Proposition 3.1 shows that the assumption in (3.21) has no loss of generality. Indeed if (3.21) does not hold we can replace $T(z)$ as in (3.20) by that in (3.24). Since

$$\det(V(z)) = \prod_{i=1}^{M-m} \det\left(V^{(\ell_i)}(z)\right),$$

and each $V^{(\ell_i)}(z)$ is column reduced [24], there exist nonnegative integers $\{d_i\}_{i=0}^{M-1}$ such that

$$\mathrm{diag}(z^{-d_0}, z^{-d_1}, \cdots, z^{-d_{M-1}})V(z)$$

is a polynomial matrix of z^{-1}, with the direct transmission term nonsingular. By (3.24), we can set

$$\underline{G}(z) = \mathrm{diag}(z^{-d_0}, z^{-d_1}, \cdots, z^{-d_{M-1}})V(z)G(z),$$

which is causal and stable, and which implies that

$$\underline{G}(z)\underline{H}(z)\underline{F}(z) = \mathrm{diag}(z^{-d_0}, z^{-d_1}, \cdots, z^{-d_{M-1}}),$$

provided that $G(z)T(z) = I_M$, with $T(z)$ as defined in (3.24) satisfying (3.22). Because $\|\underline{G}\Gamma\|_\infty = \|G\Gamma\|_\infty$, due to allpass property of $V(z)$, our optimal channel equalization problem amounts again to finding $G(z)$, a causal and stable inverse of $T(z)$ as in (3.24) satisfying (3.22), such that $\|G\Gamma\|_\infty$ is minimized. ∎

In the remainder of the section we will investigate causal and stable inverses of $T(z)$ given by either (3.20) satisfying (3.21) or (3.24) satisfying (3.22). We will parameterize all such left inverses in an explicit form. For this purpose, we associate with $T(z)$ of size $P \times M$ with $P > M$ a state-space realization, and write

$$T(z) = D + C(zI_N - A)^{-1}B =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \tag{3.25}$$

with A of size $N \times N$, B of size $N \times M$, C of size $P \times N$, and D of size $P \times M$. Note that D is the same as in (3.21). The strict minimum phase of $T(z)$ is crucial for (3.20) to hold, which is equivalent to

$$\mathrm{rank}\begin{bmatrix} A - zI_N & B \\ C & D \end{bmatrix} = N + M \ \ \forall\ |z| \geq 1. \tag{3.26}$$

If D has full column rank, then there exist D^+ of size $M \times P$, $D_\perp$ of size $P \times (P-M)$, and $D_\perp^+$ of size $(P-M) \times P$ such that

$$\begin{bmatrix} D^+ \\ D_\perp^+ \end{bmatrix}\begin{bmatrix} D & D_\perp \end{bmatrix} = \begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}. \tag{3.27}$$

To find D^+, $D_\perp$, $D_\perp^+$, define

$$D_0^+ := (D^*D)^{-1}D^*, \quad D_{0\perp}D_{0\perp}^* := I_P - D(D^*D)^{-1}D^*, \tag{3.28}$$

where $D_{0\perp}$ has the least rank among all the Cholesky factors above. Then $D_0^+D = I_M$, $D_0^+D_{0\perp} = 0$, and $D_{0\perp}^*D_{0\perp} = I_{P-M}$. In addition, D^+ and $D_\perp$ satisfying (3.27) have the following general form:

$$D^+ = D_0^+ + \chi D_{0\perp}^+, \quad D_\perp = -D\chi + D_{0\perp}, \tag{3.29}$$

where χ is a free matrix of size $M \times (P-M)$, and $D_\perp^+ = D_{0\perp}^+ = D_{0\perp}^*$ is taken.

The assumptions (3.21), (3.26), and the identity (3.27) imply the existence of causal and stable left inverses for $T(z)$. Indeed,

$$T(z) = \left[I_P + C(zI_N - A)^{-1}(BD^+ - KD_\perp^+)\right] D$$

for any constant matrix K of size $N \times (P-M)$. Hence with $B_K := BD^+ - KD_\perp^+$,

$$T^+(z) = D^+\left[I_P + C(zI_N - A)^{-1}B_K\right]^{-1} = D^+\left[I_P - C(zI_N - A_K)^{-1}B_K\right] \tag{3.30}$$

is a left inverse of $T(z)$ with $A_K := A - B_KC = A - BD^+C + KD_\perp^+C$. The existence of stable left inverses is tied to the existence of the constant matrix K such that A_K is stable, i.e., A_K has all its eigenvalues on the open unit disk. Such a K is called *stabilizing*.

Lemma 3.1 *Consider $T(z)$ as in (3.25) with D full column rank. Then the strict minimum phase condition (3.26) is true, if and only if there exists a stabilizing K, i.e., A_K is stable.*

This result was obtained in [36], and can be easily proven by noting that

$$\begin{aligned}\operatorname{rank}\left\{\begin{bmatrix} A-zI_N & B \\ C & D \end{bmatrix}\right\} &= \operatorname{rank}\left\{\begin{bmatrix} I_N & KD_\perp^+ - BD^+ \\ 0 & D^+ \\ 0 & D_\perp^+ \end{bmatrix}\begin{bmatrix} A-zI_N & B \\ C & D \end{bmatrix}\right\} \\ &= \operatorname{rank}\left\{\begin{bmatrix} A_K - zI_N & 0 \\ D^+C & I_M \\ D_\perp^+C & 0 \end{bmatrix}\right\} = M + \operatorname{rank}\left\{\begin{bmatrix} A_K - zI_N \\ D_\perp^+C \end{bmatrix}\right\}.\end{aligned}$$

Remark 3.2 Because $P > M$, generically $T(z)$ admits no transmission zeros. That is, the rows of $T(z)$ are coprime, which is assumed, and explained very recently in [26] for channel equalization. Our assumption on strictly minimum phase is weaker than the coprimeness condition, and in fact the weakest condition on the existence of channel equalizers. If this condition fails, then channel equalizers do not exist. ∎

Consider the augmented square transfer function matrix:

$$T_a(z) = \begin{bmatrix} T(z) & T_\perp(z) \end{bmatrix} := \left[I_P + C(zI_N - A)^{-1}B_K\right]\begin{bmatrix} D & D_\perp \end{bmatrix}. \tag{3.31}$$

Then $T_a(z)$ is strict minimum phase, provided that $T(z)$ is strict minimum phase, and K is stabilizing. It follows that

$$\begin{aligned}T_a^{-1}(z) &= \begin{bmatrix} D & D_\perp \end{bmatrix}^{-1}\left[I_P + C(zI_N - A)^{-1}B_K\right]^{-1} \\ &= \begin{bmatrix} D^+ \\ D_\perp^+ \end{bmatrix}\left[I_P - C(zI_N - A_K)^{-1}B_K\right] =: \begin{bmatrix} T^+(z) \\ T_\perp^+(z) \end{bmatrix},\end{aligned} \tag{3.32}$$

is stable as well. This gives rise to the identity

$$\begin{bmatrix} T^+(z) \\ T_\perp^+(z) \end{bmatrix}\begin{bmatrix} T(z) & T_\perp(z) \end{bmatrix} = \begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}. \tag{3.33}$$

Similar to (3.29), we can also parameterize all the causal and stable left inverses of $T(z)$. For completeness, we include its proof. See also [18].

Proposition 3.2 *Assume that $T(z)$ as in (3.25) satisfies (3.26) with D full column rank. If $T_a^{-1}(z)$ as in (3.32) exists, and is stable, then all causal and stable left inverses of $T(z)$ are given by*

$$T^{\text{inv}}(z) = T^+(z) + Q(z)T_\perp^+(z), \tag{3.34}$$

where $Q(z)$ is an arbitrary causal and stable transfer function matrix.

Proof: It is noted that $T^{\text{inv}}(z)$ is causal and stable, by the causality and stability of $T_a^{-1}(z)$ in (3.32), and of $Q(z)$. By the fact that $T^+(z)T(z) = I_M$, and $T_\perp^+(z)T(z) = 0$, $T^{\text{inv}}(z)$ is indeed a causal and stable left inverse of $T(z)$. Conversely consider any causal and stable left inverse $T^+(z)$ of $T(z)$. Then

$$T^+(z)T_a(z) = \left[\begin{array}{cc} I_M & T^+(z)T_\perp(z) \end{array}\right], \quad T^{\text{inv}}(z)T_a(z) = \left[\begin{array}{cc} I_M & Q(z) \end{array}\right],$$

where $T_a(z)$ is as in (3.31), and $T^{\text{inv}}(z)$ as in (3.34). Hence with $Q(z) = T^+(z)T_\perp(z)$, there holds $T^+(z)T_a(z) = T^{\text{inv}}(z)T_a(z)$, which is equivalent to $T^+(z) = T^{\text{inv}}(z)$, by the fact that $T_a(z)$ as in (3.32) is a $P \times P$ square matrix and nonsingular. The proof is thus complete. ∎

Before proceeding further, we would like to examine further the parameterization of all causal and stable left inverses in Proposition 3.2. Substituting D^+ as (3.29) to equation (3.30), then the left inverse of $T(z)$ has the form

$$T^+(z) = \left[\begin{array}{c|c} A - B(D^*D)^{-1}D^*C + LD_{0\perp}^*C & -B(D^*D)^{-1}D^* + LD_{0\perp}^* \\ \hline (D^*D)^{-1}D^*C + \chi D_{0\perp}^*C & (D^*D)^{-1}D^* + \chi D_{0\perp}^* \end{array}\right], \tag{3.35}$$

where $L = (K - B\chi)$. Furthermore the inverse of the augmented system in (3.32) has the form

$$T_a^{-1}(z) = \begin{bmatrix} T^+(z) \\ T_\perp^+(z) \end{bmatrix} = \left(\begin{bmatrix} D_0^+ \\ D_{0\perp}^* \end{bmatrix} + \begin{bmatrix} \chi \\ 0 \end{bmatrix} D_{0\perp}^* \right) \left[\begin{array}{c|c} A_0 + LD_{0\perp}^* C & -BD_0^+ + LD_{0\perp}^* \\ \hline C & I_P \end{array} \right], \tag{3.36}$$

where $A_0 = A - BD_0^+ C = A - B(D^*D)^{-1}D^*C$. Substituting (3.36) into (3.34) yields a simpler form for parameterization of all causal and stable left inverses of $T(z)$.

Corollary 3.1 *With the same hypothesis as in Proposition 3.2, and $L = (K - B\chi)$, all causal and stable left inverses of $T(z)$ are parameterized by*

$$T^{\text{inv}}(z) = \left(D_0^+ + \tilde{Q}(z) D_{0\perp}^* \right) \left[\begin{array}{c|c} A_0 + LD_{0\perp}^* C & -BD_0^+ + LD_{0\perp}^* \\ \hline C & I_P \end{array} \right]$$

where $\tilde{Q}(z) = \chi + Q(z)$ with χ a free constant matrix as in (3.29), and $Q(z)$ an arbitrary stable transfer function matrix as in (3.34).

Corollary 3.1 shows that all causal and stable left inverses can be parameterized through left inverses of the constant matrix D, with the free constant matrix χ absorbed into the causal and stable transfer matrix $Q(z)$ of the same size. In light of Proposition 3.2 and Corollary 3.1, the optimal channel equalization problem in $\mathcal{H}_\infty$-norm now seeks γ_{opt} defined by

$$\gamma_{\text{opt}} := \inf \left\{ \|(T^+ + QT_\perp^+)\Gamma\|_\infty : \quad Q \in \mathcal{RH}_\infty,\ K \text{ stabilizing} \right\} \tag{3.37}$$

and for each $\gamma > \gamma_{\text{opt}}$, searches for $T^{\text{inv}}(z)$ such that $\|T^{\text{inv}}\Gamma\|_\infty < \gamma$, where $\mathcal{RH}_\infty$ is the collection of all causal and stable rational transfer function matrices.

3.4 Suboptimal Solution to the Worst-Case Channel Equalization

For channel equalization, the PR condition (3.20) requires $\underline{G}(z) = T^{\text{inv}}(z)$, the causal and stable left inverse of $T(z)$ of the form in (3.34). The worst-case optimal channel equalization requires minimization of $\|\underline{G}\Gamma\|_\infty = \|T^{\text{inv}}\Gamma\|_\infty$ over all causal and stable left inverses of $T(z)$, where Γ is given in (3.18). In the following, we will show that the worst-case design of optimal channel equalization in $\mathcal{H}_\infty$-norm can be converted into an equivalent $\mathcal{H}_\infty$ filtering problem. For this purpose, it is noted that for any left inverse $T^{\text{inv}}(z)$ as in Proposition 3.1, $\tilde{T}(z)\Gamma = z^{-1}T^{\text{inv}}(z)\Gamma$ has the same $\mathcal{H}_\infty$-norm. We thus turn our attention to seek a strictly causal and stable $\tilde{T}(z)$ such that $\tilde{T}(z)\Gamma$ has the least (or suboptimal) $\mathcal{H}_\infty$-norm. It is interesting to note that with $Q(z)$ a causal and stable transfer function matrix,

$$
\begin{aligned}
\tilde{T}(z)\Gamma &= z^{-1}T^{\text{inv}}(z)\Gamma \\
&= \left[\begin{array}{c|c} 0 & I_M \\ \hline I_M & 0 \end{array}\right] \left[\begin{array}{c|c} A_0 + LD_{0\perp}^* C & -BD_0^+\Gamma + LD_{0\perp}^*\Gamma \\ \hline D_0^+ C + QD_{0\perp}^* C & D_0^+\Gamma + QD_{0\perp}^*\Gamma \end{array}\right] \\
&= \left[\begin{array}{cc|c} A_0 + LD_{0\perp}^* C & 0 & -BD_0^+\Gamma + LD_{0\perp}^*\Gamma \\ D_0^+ C + QD_{0\perp}^* C & 0 & D_0^+\Gamma + QD_{0\perp}^*\Gamma \\ \hline 0 & I_M & 0 \end{array}\right].
\end{aligned}
$$

Even though Q is a stable transfer function matrix, it is treated as a constant matrix in the above realization, by an abuse of notation. In order to apply the $\mathcal{H}_\infty$ filtering results in [64] to our problem, we assume that

$$
\text{rank}\left\{R^2 := D_{0\perp}^*\Gamma\Gamma^* D_{0\perp}\right\} = P - M. \tag{3.38}
$$

This rank condition implies that r, the rank of Γ, is at least $P-M$. In light of the discussion at the end of Section 3, Γ is most likely to be of rank P, and thus (3.38) is a valid assumption. Consider variable substitution

$$\begin{bmatrix} L \\ Q \end{bmatrix} = \begin{bmatrix} L_1 \\ Q_1 \end{bmatrix} + \begin{bmatrix} L_2 \\ Q_2 \end{bmatrix} R^{-1}, \qquad \begin{bmatrix} L_1 \\ Q_1 \end{bmatrix} = \begin{bmatrix} B \\ -I_M \end{bmatrix} D_0^+ \Gamma\Gamma^* D_{0\perp} R^{-2}. \tag{3.39}$$

Let $A_{L_1} = A_0 + L_1 D_{0\perp}^* C$, and $C_{Q_1} = D_0^+ C + Q_1 D_{0\perp}^* C$. Denote

$$\tilde{A} = \begin{bmatrix} A_{L_1} & 0 \\ C_{Q_1} & 0 \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} -B \\ I_M \end{bmatrix} D_0^+ \Gamma (I_r - \Gamma^* D_{0\perp} R^{-2} D_{0\perp}^* \Gamma), \tag{3.40}$$

$$\tilde{C}_1 = \begin{bmatrix} 0 & I_M \end{bmatrix}, \qquad \tilde{C}_2 = \begin{bmatrix} R^{-1} D_{0\perp}^* C & 0 \end{bmatrix}, \qquad \tilde{D} = R^{-1} D_{0\perp}^* \Gamma. \tag{3.41}$$

Then it can be seen that

$$\tilde{T}(z)\Gamma = z^{-1} T^{\text{inv}}(z)\Gamma = \left[\begin{array}{c|c} \tilde{A} + \tilde{L}\tilde{C}_2 & \tilde{B} + \tilde{L}\tilde{D} \\ \hline \tilde{C}_1 & 0 \end{array}\right], \qquad \tilde{L} = \begin{bmatrix} L_2 \\ Q_2 \end{bmatrix}, \tag{3.42}$$

has the form of a standard state estimator (the error system), where $\tilde{L}$ can be allowed to be dynamical. Hence in light of the $\mathcal{H}_\infty$ filtering results in [64], minimization of $\|\tilde{T}\Gamma\|_\infty$ over all dynamical and stabilizing state estimator gain $\tilde{L}$ is the same as minimization of $\|\tilde{T}\Gamma\|_\infty$ over all static stabilizing state estimator gain $\tilde{L} \in \mathbf{C}^{(N+M)\times(P-M)}$, the set of all complex matrices of size $(N+M)\times(P-M)$. That is,

$$\gamma_{\text{opt}} = \inf_{\tilde{L}\in\mathbf{C}^{(N+M)\times(P-M)}} \|\tilde{T}\Gamma\|_\infty = \inf\left\{\|T^{\text{inv}}\Gamma\|_\infty : \ Q \in \mathcal{RH}_\infty, \ K \text{ stabilizing}\right\}.$$

Consequently we need consider only the case when $\tilde{L}$ is a constant matrix in realization of $\tilde{T}(z)$.

Remark 3.3 It can be verified that with $\tilde{B}$ in (3.40) and $\tilde{D}$ in (3.41),

$$\tilde{D}\begin{bmatrix} \tilde{B}^* & \tilde{D}^* \end{bmatrix} = \begin{bmatrix} 0 & I_{P-M} \end{bmatrix}, \tag{3.43}$$

by nonsingularity of $R^2 := D_{0\perp}^* \Gamma\Gamma^* D_{0\perp}$. Furthermore

$$\begin{aligned}
\operatorname{rank}\left\{\begin{bmatrix} \tilde{A}-zI & \tilde{B} \\ \tilde{C}_2 & \tilde{D} \end{bmatrix}\right\} &= \operatorname{rank}\left\{\begin{bmatrix} I_N & 0 & L_1 R \\ 0 & I_M & Q_1 R \\ 0 & 0 & R \end{bmatrix}\begin{bmatrix} A_{L_1}-zI_N & 0 & \tilde{B}_1\Gamma \\ C_{Q_1} & -zI_M & \tilde{B}_2\Gamma \\ R^{-1}D_{0\perp}^* C & 0 & R^{-1}D_{0\perp}^*\Gamma \end{bmatrix}\right\} \\
&= \operatorname{rank}\left\{\begin{bmatrix} A_0-zI_N & 0 & -BD_0^+ \\ D_0^+ C & -zI_M & D_0^+ \\ D_{0\perp}^* C & 0 & D_{0\perp}^* \end{bmatrix}\right\} \\
&= M + \operatorname{rank}\left\{\begin{bmatrix} A_0-zI_N & -BD_0^+ \\ D_{0\perp}^* C & D_{0\perp}^* \end{bmatrix}\begin{bmatrix} I_N & 0 \\ -C & I_{P-M} \end{bmatrix}\right\} \\
&= M + \operatorname{rank}\left\{\begin{bmatrix} A-zI_N & -BD_0^+ \\ 0 & D_{0\perp}^* \end{bmatrix}\right\} \\
&= P+N \quad (\forall\ |z|\geq 1)
\end{aligned} \tag{3.44}$$

by full rank of Γ, and by stability of A. ∎

We are now ready to prove the main result of this chapter. By applying the results in [64] to our optimal channel equalization problem, the following theorem establishes an equivalent condition for solving the suboptimal worst-case equalizer design problem, and gives the construction of a specific suboptimal left inverse $T^{\text{inv}}(z)$.

Theorem 3.3 *Let γ_{opt} be the optimal value for $\|T^{\text{inv}}\Gamma\|_\infty$ among all causal and stable left inverses of $T(z)$. Then $\gamma > \gamma_{\text{opt}}$, if and only if the discrete-time filtering algebraic Riccati equation (ARE)*

$$Y = A_\pi Y \left(I_N + C^*\Psi C Y\right)^{-1} A_\pi^* + B\Pi(I_M - \gamma^{-2}\Pi)^{-1}B^* \tag{3.45}$$

admits a stabilizing[3] *solution* $Y \geq 0$*, where with* R *as in (3.38),* L_1 *and* Q_1 *as in (3.39),*

$$\begin{aligned} A_\pi &= A_{L_1} - \gamma^{-2} B\Pi(I_M - \gamma^{-2}\Pi)^{-1} C_{Q_1}, \quad A_{L_1} = A_0 + BL_1, \quad C_{Q_1} = D_0^+ C + Q_1 D_{0\perp}^* C, \\ \Psi &= D_{0\perp} R^{-2} D_{0\perp}^* - D_R (D_0^+)^* (\gamma^2 I_M - \Pi)^{-1} D_0^+ D_R^*, \quad D_R = I_P - D_{0\perp} R^{-2} D_{0\perp}^* \Gamma\Gamma^*, \\ \Pi &= D_0^+ \Gamma (I_r - \Gamma^* D_{0\perp} R^{-2} D_{0\perp}^* \Gamma) \Gamma^* (D_0^+)^*. \end{aligned}$$

Let Y *be a stabilizing solution to the ARE (3.45) for any* $\gamma > \gamma_{\text{opt}}$*. Then a specific suboptimal left inverse* $T^{\text{inv}}(z)$ *achieving* $\|T^{\text{inv}}\Gamma\|_\infty < \gamma$ *is given by*

$$T^{\text{inv}}(z) = \left[\begin{array}{c|c} A_{L_1}(I_N + YC^* D_{0\perp} R^{-2} D_{0\perp}^* C)^{-1} & -B_0 \\ \hline C_{Q_1}(I_N + YC^* D_{0\perp} R^{-2} D_{0\perp}^* C)^{-1} & D_0^+ \end{array}\right] (I_P + CYC^* D_{0\perp} R^{-2} D_{0\perp}^*)^{-1} \tag{3.46}$$

with $B_0 = AYC^* D_{0\perp} R^{-2} D_{0\perp}^* + BD_0^+$.

Proof: Since (3.43) and (3.44) hold true, the result in [64] can be applied to the state estimator problem in (3.42). By Theorem 2 of [64], $\|\tilde{T}(z)\Gamma\|_\infty < \gamma$, if and only if the following ARE

$$Z = \tilde{A} Z \tilde{A}^* + \tilde{B}\tilde{B}^* - \tilde{A} Z \tilde{C}_2^* (I_{P-M} + \tilde{C}_2 Z \tilde{C}_2^*)^{-1} \tilde{C}_2 Z \tilde{A}^* + \gamma^{-2} Z \tilde{C}_1^* (I_M + \gamma^{-2} \tilde{C}_1 Z \tilde{C}_1^*)^{-1} \tilde{C}_1 Z \tag{3.47}$$

has a stabilizing solution $Z \geq 0$. Partition Z into a 2×2 block matrix Z_{ij} for $i, j = 1, 2$, compatible with the partition of $\tilde{A}$. If the stabilizing solution $Z \geq 0$ exists, then the suboptimal state estimator gain is given by [64]

$$\tilde{L} = -\tilde{A} Z \tilde{C}_2^* (I_{P-M} + \tilde{C}_2 Z \tilde{C}_2^*)^{-1}$$

[3] Y is stabilizing if $(I_N + C^* \Psi C Y)^{-1} A_\pi^*$ is stable.

$$
\begin{aligned}
&= -\begin{bmatrix} A_{L_1} \\ C_{Q_1} \end{bmatrix} Z_{11} C^* D_{0\perp} R^{-1} (I_{P-M} + R^{-1} D_{0\perp}^* C Z_{11} C^* D_{0\perp} R^{-1})^{-1} \\
&= -\begin{bmatrix} A_{L_1} \\ C_{Q_1} \end{bmatrix} Z_{11} (I_N + C^* D_{0\perp} R^{-2} D_{0\perp}^* C Z_{11})^{-1} C^* D_{0\perp} R^{-1}, \qquad (3.48)
\end{aligned}
$$

and $\tilde{A} + \tilde{L}\tilde{C}_2$ is stable. Substituting the expression of $\tilde{A}$ gives

$$
\tilde{A} Z \tilde{A}^* = \begin{bmatrix} A_{L_1} Z_{11} A_{L_1}^* & A_{L_1} Z_{11} C_{Q_1}^* \\ C_{Q_1} Z_{11} A_{L_1}^* & C_{Q_1} Z_{11} C_{Q_1}^* \end{bmatrix}. \qquad (3.49)
$$

Since $(I_r - \Gamma^* D_{0\perp} R^{-2} D_{0\perp}^* \Gamma)^2 = I_r - \Gamma^* D_{0\perp} R^{-2} D_{0\perp}^* \Gamma$, which is a projection matrix,

$$
\begin{aligned}
\tilde{B}\tilde{B}^* &= \begin{bmatrix} -B \\ I_M \end{bmatrix} D_0^+ \Gamma (I_r - \Gamma^* D_{0\perp} R^{-2} D_{0\perp}^* \Gamma)^2 \Gamma^* (D_0^+)^* \begin{bmatrix} -B \\ I_M \end{bmatrix}^* \\
&= \begin{bmatrix} B\Pi B^* & -B\Pi \\ -\Pi B^* & \Pi \end{bmatrix}. \qquad (3.50)
\end{aligned}
$$

Via direct calculation,

$$
\begin{aligned}
\tilde{A} Z \tilde{C}_2^* &= \begin{bmatrix} A_{L_1} \\ C_{Q_1} \end{bmatrix} Z_{11} C^* D_{0\perp} R^{-1}, \\
(I_{P-M} + \tilde{C}_2 Z \tilde{C}_2^*)^{-1} &= (I_{P-M} + R^{-1} D_{0\perp}^* C Z_{11} C^* D_{0\perp} R^{-1})^{-1}.
\end{aligned}
$$

It follows that

$$
-\tilde{A} Z \tilde{C}_2^* (I_{P-M} + \tilde{C}_2 Z \tilde{C}_2^*)^{-1} \tilde{C}_2 Z \tilde{A}^* = -\begin{bmatrix} A_{L_1} T A_{L_1}^* & A_{L_1} T C_{Q_1}^* \\ C_{Q_1} T A_{L_1}^* & C_{Q_1} T C_{Q_1}^* \end{bmatrix} \qquad (3.51)
$$

where $T = Z_{11} C^* D_{0\perp} R^{-1} (I_{P-M} + R^{-1} D_{0\perp}^* C Z_{11} C^* D_{0\perp} R^{-1})^{-1} R^{-1} D_{0\perp}^* C Z_{11}$. Note that

$$
Z \tilde{C}_1^* = \begin{bmatrix} Z_{12} \\ Z_{22} \end{bmatrix}, \quad (I_M + \gamma^{-2} \tilde{C}_1 Z \tilde{C}_1^*)^{-1} = (I_M + \gamma^{-2} Z_{22})^{-1}.
$$

We now have that

$$\gamma^{-2}Z\tilde{C}_1^*(I_M+\gamma^{-2}\tilde{C}_1Z\tilde{C}_1^*)^{-1}\tilde{C}_1Z = \gamma^{-2}\begin{bmatrix} Z_{12} \\ Z_{22} \end{bmatrix}(I_M+\gamma^{-2}Z_{22})^{-1}\begin{bmatrix} Z_{21} & Z_{22} \end{bmatrix}. \quad (3.52)$$

In light of (3.50), (3.49), (3.51) and (3.52), the ARE (3.47) can now be decomposed into:

$$\begin{aligned} Z_{11} &= A_{L_1}Z_{11}A_{L_1}^* + B\Pi B^* - A_{L_1}TA_{L_1}^* + \gamma^{-2}Z_{12}(I_M+\gamma^{-2}Z_{22})^{-1}Z_{21}, & (3.53)\\ Z_{12} &= A_{L_1}Z_{11}C_{Q_1}^* - B\Pi - A_{L_1}TC_{Q_1}^* + \gamma^{-2}Z_{12}(I_M+\gamma^{-2}Z_{22})^{-1}Z_{22}, & (3.54)\\ Z_{22} &= C_{Q_1}Z_{11}C_{Q_1}^* + \Pi - C_{Q_1}TC_{Q_1}^* + \gamma^{-2}Z_{22}(I_M+\gamma^{-2}Z_{22})^{-1}Z_{22}. & (3.55) \end{aligned}$$

Let $S = Z_{11} - T$. We have

$$\begin{aligned} S &= Z_{11} - Z_{11}C^*D_{0\perp}R^{-1}(I_N + R^{-1}D_{0\perp}^*CZ_{11}C^*D_{0\perp}R^{-1})^{-1}R^{-1}D_{0\perp}^*CZ_{11}\\ &= Z_{11} - (I_N + Z_{11}C^*D_{0\perp}R^{-2}D_{0\perp}^*C)^{-1}Z_{11}C^*D_{0\perp}R^{-2}D_{0\perp}^*CZ_{11}\\ &= (I_N + Z_{11}C^*D_{0\perp}R^{-2}D_{0\perp}^*C)^{-1}Z_{11}\\ &= Z_{11}(I_N + C^*D_{0\perp}R^{-2}D_{0\perp}^*CZ_{11})^{-1}. \end{aligned} \quad (3.56)$$

Then the expressions in (3.53) — (3.55) can be simplified into

$$\begin{aligned} Z_{11} - \gamma^{-2}Z_{12}(I_M+\gamma^{-2}Z_{22})^{-1}Z_{21} &= B\Pi B^* + A_{L_1}SA_{L_1}^*, & (3.57)\\ Z_{12}(I_M+\gamma^{-2}Z_{22})^{-1} &= -B\Pi + A_{L_1}SC_{Q_1}^*, & (3.58)\\ Z_{22}(I_M+\gamma^{-2}Z_{22})^{-1} &= \Pi + C_{Q_1}SC_{Q_1}^*, & (3.59) \end{aligned}$$

respectively. In light of (3.59) we have the following:

$$(I_M+\gamma^{-2}Z_{22})^{-1} = I_M - \gamma^{-2}Z_{22}(I_M+\gamma^{-2}Z_{22})^{-1} = I_M - \gamma^{-2}\left(\Pi + C_{Q_1}SC_{Q_1}^*\right).$$

It follows from (3.58) that

$$Z_{21} = Z_{12}^* = \left[I_M - \gamma^{-2}\left(\Pi + C_{Q_1} S C_{Q_1}^*\right)\right]^{-1}\left(-B\Pi + A_{L_1} S C_{Q_1}^*\right)^*$$

Now equation (3.57) becomes

$$\begin{aligned} Z_{11} &= \gamma^{-2}\left(-B\Pi + A_{L_1} S C_{Q_1}^*\right)\left[I_M - \gamma^{-2}\left(\Pi + C_{Q_1} S C_{Q_1}^*\right)\right]^{-1}\left(-B\Pi + A_{L_1} S C_{Q_1}^*\right)^* \\ &\quad + B\Pi B^* + A_{L_1} S A_{L_1}^*. \end{aligned} \tag{3.60}$$

Write the above equation as $Z_{11} = \Sigma_A + \Sigma_B - \Sigma_{AB} - \Sigma_{AB}^*$ with

$$\Sigma_A = A_{L_1} S A_{L_1}^* + \gamma^{-2} A_{L_1} S C_{Q_1}^*\left[I_M - \gamma^{-2}\left(\Pi + C_{Q_1} S C_{Q_1}^*\right)\right]^{-1}\left(A_{L_1} S C_{Q_1}^*\right)^*, \tag{3.61}$$

$$\Sigma_B = B\Pi B^* + \gamma^{-2} B\Pi\left[I_M - \gamma^{-2}\left(\Pi + C_{Q_1} S C_{Q_1}^*\right)\right]^{-1}(B\Pi)^*, \tag{3.62}$$

$$\Sigma_{AB} = \gamma^{-2} B\Pi\left[I_M - \gamma^{-2}\left(\Pi + C_{Q_1} S C_{Q_1}^*\right)\right]^{-1}\left(A_{L_1} S C_{Q_1}^*\right)^*. \tag{3.63}$$

Through a tedious derivation, equations (3.61) — (3.63) become the following:

$$\Sigma_A = A_{L_1} Z_{11}\left(I_N + C^* \Psi C Z_{11}\right)^{-1} A_{L_1}^* \tag{3.64}$$

$$\begin{aligned} \Sigma_B &= (\gamma^{-2})^2 B\Pi(I_M - \gamma^{-2}\Pi)^{-1} C_{Q_1} Z_{11}\left(I_N + C^*\Psi C Z_{11}\right)^{-1} C_{Q_1}^* (I_M - \gamma^{-2}\Pi)^{-1}\Pi^* B^* \\ &\quad + B\Pi(I_M - \gamma^{-2}\Pi)^{-1} B^* \end{aligned} \tag{3.65}$$

$$\Sigma_{AB} = \gamma^{-2} B\Pi(I_M - \gamma^{-2}\Pi)^{-1} C_{Q_1} Z_{11}\left(I_N + C^*\Psi C Z_{11}\right)^{-1} A_{L_1}^* \tag{3.66}$$

Via direct computation (we skip the detailed steps),

$$Z_{11} = A_\pi Z_{11}\left(I_N + C^*\Psi C Z_{11}\right)^{-1} A_\pi^* + B\Pi(I_M - \gamma^{-2}\Pi)^{-1} B^*. \tag{3.67}$$

Equation (3.67) is identical to (3.45) with $Y = Z_{11}$. Since $\tilde{L}$ in (3.48) depends only on Z_{11}, Z is stabilizing, if and only if Z_{11} is stabilizing. It is thus concluded that

$\gamma > \gamma_{\text{opt}}$, if and only if (3.45) has a stabilizing solution $Y = Z_{11} \geq 0$, in light of [64]. Indeed, if (3.45) has a stabilizing solution $Y \geq 0$, then $Z_{22} > 0$ can be computed from (3.59), and Z_{12} can be obtained from (3.58), based on $Z_{11} = Y$. Since any solution to (3.47) satisfies $Z \geq 0$, and the state estimator gain in (3.48) depends only on Z_{11}, Z is stabilizing as well. We thus conclude that $\|\tilde{T}\Gamma\|_\infty = \|T^{\text{inv}}\Gamma\|_\infty < \gamma$ with $L = L_1 + L_2 R^{-1}$, and $Q = Q_1 + Q_2 R^{-1}$, where L_2 and Q_2 are computed from (3.48), and with $K = B\chi + L$. The construction of suboptimal $T^{\text{inv}}(z)$ can be carried out through routine calculation. ∎

3.5 Computation of γ_{opt} and Summary of the Design Algorithm

We consider computation of γ_{opt}, which is the largest γ such that the ARE (3.45) fails to produce the stabilizing solution. If $T(z)$ as in (3.25) satisfies the power complementary condition [55]

$$[T(e^{j2f\pi})]^* T(e^{j2f\pi}) = I_M \ \ \forall\ f\ \in\ [0,\ 1], \tag{3.68}$$

and $\Gamma = I_M$, then an analytical expression of γ_{opt} can be obtained.

Theorem 3.4 *Suppose that $T(z)$ as in (3.25) is causal, stable, and satisfies the strict minimum phase condition (3.26) with D full column rank. Let X and P be the unique solutions to*

$$X - AXA^* = BB^*, \qquad P - A^*PA = C^*C,$$

respectively. If (3.68) holds, then

$$\gamma_{\text{opt}} = \inf_{Q \in \mathcal{RH}_\infty} \|T^{\text{inv}}\|_\infty = \frac{1}{\sqrt{1 - \lambda_{\max}(XP)}}. \tag{3.69}$$

Proof: Finding minimum induced norm for operators in Hilbert space is termed Corona problem. It has been studied in [13, 15] for the $\mathcal{H}_\infty$-norm case in the continuous-time. It should be clear that the same procedure in [13, 15] applies to the discrete-time systems to yield the analytic expression in (3.69), if the strict minimum phase condition (3.26) with D full column rank, and (3.68) are satisfied. The details are omitted. ∎

In light of the results in [41], joint transceiver optimization requires the power complementary condition (3.68). Thus the assumption on (3.68) makes sense, if joint transceiver optimization is allowed in design of the transmitter filters $\{F_i(z)\}_{i=0}^{M-1}$ in Figure 3.3. However in some applications, transmitter filters may not be free for design, or channel characteristics may vary with respect to time. In this case, the condition (3.68) may not hold, and there is no analytic expression for γ_{opt}. As an alternative, bisection method can be used to iteratively search for γ_{opt}, which is outlined as follows.

Bisection method for computation of γ_{opt}

- Step 1: Choose the tolerance $\epsilon > 0$, and initial upper and lower bounds $(\overline{\gamma}, \underline{\gamma})$ satisfying $\underline{\gamma} < \gamma_{\text{opt}} < \overline{\gamma}$.

- Step 2: Set $\gamma = (\overline{\gamma} + \underline{\gamma})/2$. If $|\overline{\gamma} - \underline{\gamma}| \leq \epsilon$, set $\gamma_{\rm opt} = \gamma$. Stop. Otherwise do the following:
  - Compute the stabilizing solution to ARE (3.45). If the stabilizing solution $Y \geq 0$ exists, set $\overline{\gamma} = \gamma$. Otherwise, set $\underline{\gamma} = \gamma$.
  - Repeat Step 2.

End.

Since $\gamma_{\rm opt}^2 D^* D \geq I_M$, an initial lower bound $\underline{\gamma} = \underline{\gamma}_0 := \sqrt{\lambda_{\max}[(D^*D)^{-1}]}$ can be chosen. As for the initial upper bound, one may first compute a stable left inverse $T^+(z)$ which has the smallest mean-square value as in [18] for the case of white noise, and then set the initial upper bound as $\overline{\gamma} = \overline{\gamma}_0 := \|T^+\|_\infty$. Let n be the total number of iterations in the algorithm. Then the bisection algorithm implies that

$$\frac{\overline{\gamma}_0 - \underline{\gamma}_0}{2^n} \leq \epsilon \iff n \geq \log_2 \left(\frac{\overline{\gamma}_0 - \underline{\gamma}_0}{\epsilon} \right).$$

Hence the computational complexity is dependent on the initial guess on the upper and lower bounds of $\gamma_{\rm opt}$, as well as the required accuracy ϵ for $\gamma_{\rm opt}$.

Finally we summarize our proposed design algorithm as follows.

Design Algorithm for Optimal Channel Equalizers:

- Step 1: Find minimal realizations for the blocked transfer function matrices: $\underline{H}(z)$ and $\underline{F}(z)$.
- Step 2: Find a realization for $T(z)$ in (3.25) satisfying (3.26).

- Step 3: If equation (3.68) is satisfied and $\Gamma = I_M$, γ_{opt} can be computed by (3.69). Solve the discrete time ARE (3.45) by taking $\gamma > \gamma_{\text{opt}}$.

- Step 4: If equation (3.68) is not satisfied, solve ARE (3.45) along with the process of iteratively finding γ_{opt} using the bisection method described above.

- Step 5: Based on the ARE solution from Step 3 or Step 4, compute $\underline{G} = T^{\text{inv}}(z)$ as in (3.46).

- Step 6: Set $G_k(z) = \sum_{i=0}^{P-1} z^{-i} G_{k,i}(z^P)$ for $0 \leq k < M$, where $G_{k,i}(z)$ is the $(i\text{-}k)$th element of $\underline{G}(z)$, obtained in Step 5.

End.

3.6 An Illustrative Example

In this section, we consider the example studied in [41] and [18]. For the purposes of comparison, we choose Γ to be an identity matrix, and the channel to be the same as the one in [41], which is given by

$$H(z) = \sum_{i=0}^{L} h(i) z^{-i} = 1 - 0.3z^{-1} + 0.5z^{-2} - 0.4z^{-3} + 0.1z^{-4} - 0.02z^{-5} + 0.3z^{-6} - 0.1z^{-7}.$$

Thus $L = 7$. We take $P = 4 > M = 3$ and assume that $\underline{F}(z) = \begin{bmatrix} I_M & 0 \end{bmatrix}^T$. It can be verified that [41]

$$\underline{H}(z)\underline{F}(z) = \sum_{i=0}^{2} H_i^f z^{-i}, \quad H_i^f = \begin{bmatrix} h(iP) & \cdots & h(iP - P + 2) \\ h(iP+1) & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ h(iP+P-1) & \cdots & h(iP+1) \end{bmatrix}. \quad (3.70)$$

It follows that a simple realization

$$\underline{H}(z)\underline{F}(z) = \left[\begin{array}{cc|c} 0_{M\times M} & 0_{M\times M} & I_M \\ I_M & 0_{M\times M} & 0_{M\times M} \\ \hline H_1^f & H_2^f & H_0^f \end{array}\right]$$

can be easily obtained. In order to satisfy (3.68) to achieve joint optimization [41], we compute spectral factorization:

$$\underline{F}^*(\bar{z}^{-1})\underline{H}^*(\bar{z}^{-1})\underline{H}(z)\underline{F}(z) = \Omega^*(\bar{z}^{-1})\Omega(z),$$

and set $\underline{F}_{\rm new}(z) = \left[\begin{array}{cc} I_M & 0 \end{array}\right]^T \Omega^{-1}(z)$. The transmitter filters $\{F_k(z)\}_{k=0}^2$ can be easily computed based on their polyphase components in $\underline{F}_{\rm new}(z)$. In this case $T(z) = \underline{H}(z)\underline{F}_{\rm new}(z)$ satisfies (3.68). See also [18]. To design a channel equalizer, which is arbitrarily close to the optimal one in the worst-case, Theorem 3.4 is used to obtain $\gamma_{\rm opt} = 1.104299$. Then $\gamma = 1.0001 * \gamma_{\rm opt} = 1.104409$ is taken to compute the stabilizing solution to the discrete-time ARE (3.45). The corresponding suboptimal channel equalizer is then obtained with $\underline{G}(z) = T^{\rm inv}(z)$ as in (3.46), from which the optimal receiver filters $\{G_k(z)\}_{k=0}^2$ can be obtained via elementary manipulations of $\underline{G}(z)$.

Define the average frequency responses of the transmitter and receiver filters by

$$|\mathcal{F}(f)| := \frac{1}{M}\sum_{m=0}^{M-1} |F_m(e^{j2f\pi})|, \quad |\mathcal{G}(f)| := \frac{1}{M}\sum_{p=0}^{M-1} |G_p(e^{j2f\pi})|.$$

We plot $|\mathcal{F}(f)|$ (solid line) versus the channel frequency response (dashed line) in Figure 3.4, and plot $|\mathcal{G}(f)|$ (solid line) versus the channel frequency response (dashed

line) in Figure 3.5. All the graphs are normalized with respect to the respective maximum.

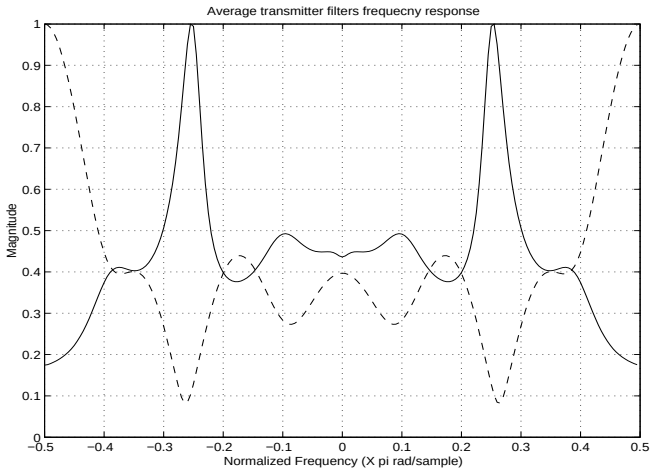

Figure 3.4: Average frequency responses of $|\mathcal{F}(f)|$ versus $|H(e^{j2f\pi})|$

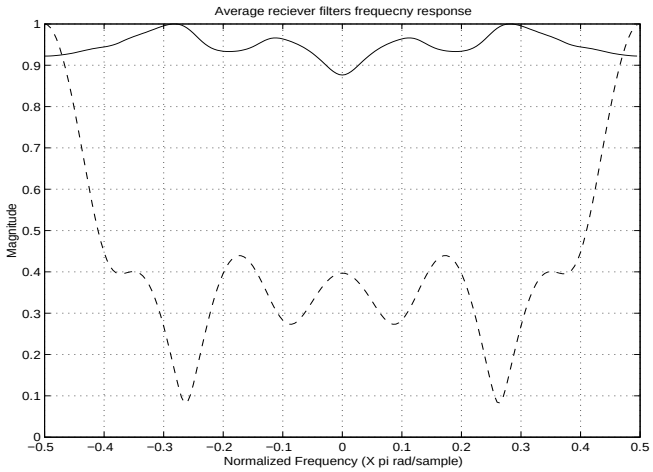

Figure 3.5: Average frequency responses of $|\mathcal{G}(f)|$ versus $|H(e^{j2f\pi})|$

For comparison, we calculated both the $\mathcal{H}_\infty$-norm and $\mathcal{H}_2$-norm for this example and the method in [18], in which the same channel and transmitter are taken, but with

different optimization technique. The computation results are as follows: For $\mathcal{H}_\infty$ criterion in this dissertation, $\|\underline{G}\|_\infty = 1.104380$, $\|\underline{G}\|_2 = 1.794239$; For $\mathcal{H}_2$ criterion in [18], $\|\underline{G}\|_\infty = 1.166611$, $\|\underline{G}\|_2 = 1.785585$. In other words, our worst-case design does not suffer poor performance even if the white noise is present, with only 4.8% deterioration compared with the optimal case.

In the remainder of the section we consider design of the optimal channel equalizer for the same channel $H(z)$ as earlier with again $M = 3$, and $P = 4$, but a different set of transmitter filters $\{F_k(z)\}_{k=0}^2$. We adopt the CDMA model as in [41, 52], and take the coefficient $\mathbf{f}_k$ of $F_k(z)$ as the kth spreading sequence of length 7 for $0 \le k \le 2$, which are given by

$$\mathbf{f}_0 = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & 1 & -1 \end{bmatrix},$$

$$\mathbf{f}_1 = \begin{bmatrix} 1 & 1 & -1 & -1 & 1 & -1 & 1 \end{bmatrix},$$

$$\mathbf{f}_2 = \begin{bmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix}.$$

Thus the equivalent channel model $T(z) = \underline{H}(z)\underline{F}(z)$, and its state-space realization can be easily constructed, which is strictly minimum phase.

The bisection algorithm in Section 3.5 is used to compute γ_{opt}, which yields $\gamma_{\text{opt}} = 2.902226$. An $\mathcal{H}_\infty$ channel equalizer is designed for the case $\gamma = 1.0001\gamma_{\text{opt}}$. Its magnitude frequency responses are plotted in Figure 3.6 in terms of the maximum singular value with solid line (The frequency is normalized with respect to the maximum). The optimal $\mathcal{H}_2$ channel equalizer is designed using the algorithm in [18]. Its maximum singular value at each frequency sample is also computed, and shown

in the dash-dot line in Figure 3.6. As seen, there is a sharp contrast between the two frequency responses. In fact the optimal $\mathcal{H}_2$ channel equalizer has an $\mathcal{H}_\infty$-norm, which is 28.1% greater than that of the $\mathcal{H}_\infty$ channel equalizer. But there is only about 14.3% difference in the RMS values between the two channel equalizers.

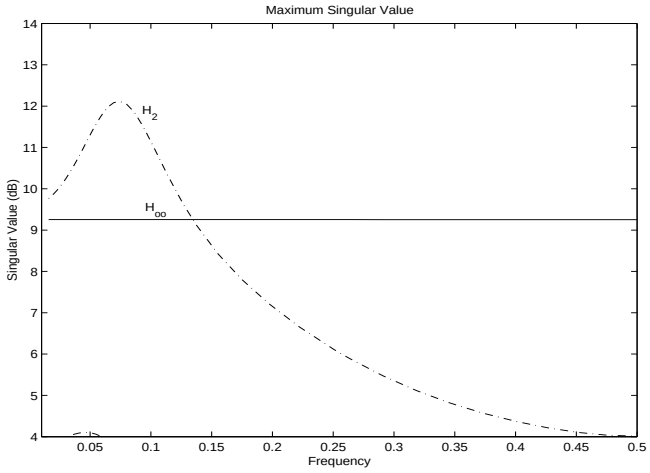

Figure 3.6: Maximum singular value response for $\mathcal{H}_\infty$ and $\mathcal{H}_2$ channel equalizers

To better illustrate the advantages of the $\mathcal{H}_\infty$ channel equalizer, we assume that the colored noise with unknown PSD is generated by passing the white noise through a bandpass filter whose passband is $[\frac{0.4}{2\pi}, \frac{0.5}{2\pi}]$, over which the two channel equalizers differ the most. The averaged BERs for the two different channel equalizers are plotted in Figure 3.7, assuming BPSK using the same scheme as in [41]. It is seen that the $\mathcal{H}_\infty$ channel equalizer outperforms the $\mathcal{H}_2$ channel equalizer. On the other

hand if the noise is white, then the $\mathcal{H}_2$ channel equalizer outperforms the $\mathcal{H}_\infty$ channel equalizer. See Figure 3.8.

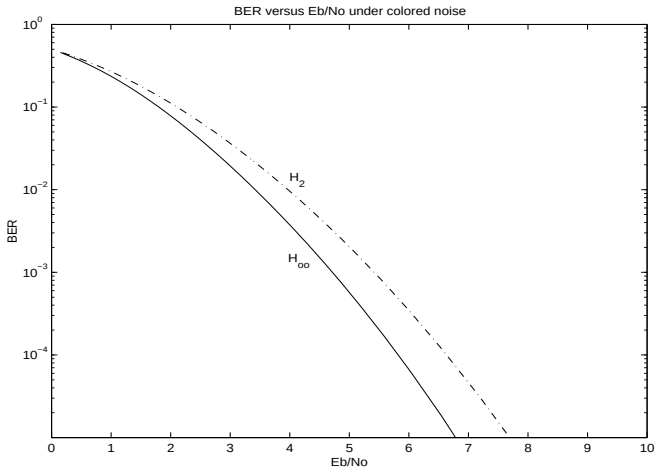

Figure 3.7: BER comparisons between the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ equalizers (colored noise)

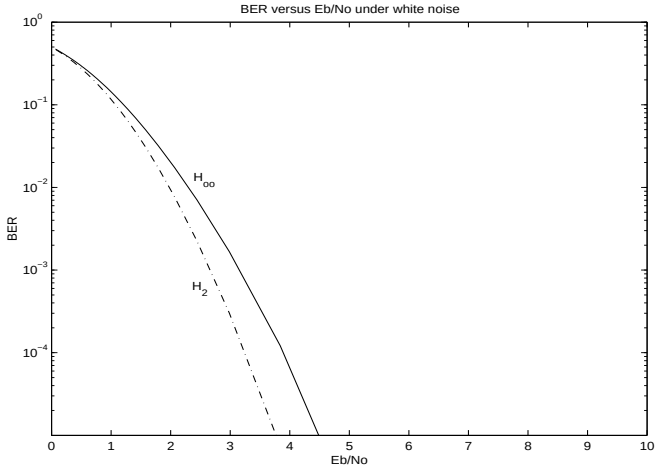

Figure 3.8: BER comparisons between the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ equalizers (white noise)

3.7 Concluding Remark

In this chapter, optimal channel equalization is investigated for multirate filterbank transceivers such as transmultiplexers. The observation noise is assumed to be non-white (colored) with unknown spectral but bounded power norm. Because the colored noise incorporates cross-talk, inter-channel interference, and residual echos [40], it is an important research problem for channel equalization. In the case of known PSD for the observation noise, MMSE design can be extended to achieve optimal channel equalization. However if no knowledge on the PSD of the observation noise, except some crude bound, is available, then the worst-case approach to optimal channel equalization in the presence of non-white noise yields an $\mathcal{H}_\infty$ optimization problem, provided that the equivalent blocked channel satisfies the strict minimum phase condition. The worst-case approach provides the performance guarantee for those colored noises with bounded power norm no matter what the PSD shape is. The design algorithm is obtained for the suboptimal channel equalizer which is arbitrarily close to the optimal one via solving the stabilizing solution of an algebraic Riccati equation associated with $\mathcal{H}_\infty$ filtering. The complexity for the $\mathcal{H}_\infty$ design is the same as the $\mathcal{H}_2$ design, except that the optimal performance bound γ_{opt} needs be computed before hand. Numerical examples are worked out to illustrate the proposed design method.

Chapter 4

Optimal Precoder Design with Minimum BER

The wireless channel often suffers from attenuation due to destructive addition of multipaths in the propagation media, and interferences from other users, which render it difficult for the receiver to reliably detect the transmitted signal. One of the most important techniques to solve this problem is to use diversity, which provides a less-attenuated replica of the transmitted signal to the receiver [46, 48]. Among all the diversity techniques, antenna diversity (spatially separated or differently polarized antenna) is more appealing because of its practical and effective nature. The advantages of deploying multiple antennas at both the transmitter and receiver to achieve spatial diversity is the added reliability due to multiple receive chains. Should one of the receive chains fail, and the other chains are operational, may the signal still be detected. The large capacities available to communication systems employing multiple antennas have been the subject of much research in recent years. The information capacity of wireless communication systems may be increased dramatically by de-

ploying multiple antennas at both the transmitter and the receiver [12, 49, 47], thus achieving very high data rates on narrow-band wireless channels.

In the transmission of digital data over the MIMO high-speed communication systems, channel induced ISI and ICI are major performance limiting factors. To mitigate such effects, channel equalization has been studied extensively as a method of combating ISI and ICI. Optimal designs developed in the past, have received great attention because of the interest in joint transmit-receive diversity schemes [29, 43, 39, 63, 65, 66]. Other results have been previously obtained such as [26] where the optimal equalizer/precoder is derived independently and hence reduced the hardware implementation compared to the joint design. Recently, there has been investigations regarding block precoding transmission technique to shift the signal processing burden from the receiver to transmitter [40, 63]. In this case, a precoding filter (precoder) is designed at the transmitter side and is applied prior to transmission to equalize the signal at the output of the receive filter. This is particularly of interest in the downlink mobile radio channel where it is desired to keep the receiver units as simple as possible. Precoder design is expected to allow the receiver to be considerably simplified, which in turn, reduces computational complexity and power consumption. Much research has been carried out to design optimal precoders based on minimum mean-squared error (MMSE), maximum information rate, and etc. However, it is the bit error rate (BER) in transmission with which users are most concerned. In this chapter, the problem of designing a minimum BER precoder for MIMO systems with

additive noise having known spectra is addressed under the perfect reconstruction (PR) condition.

The remainder of this chapter is organized as follows: A general MIMO communication system model is introduced in Section 4.1 along with the formulation of the design problem. Section 4.2 describes an optimal control problem and its solution, which is then used for the optimal precoder design in Section 4.3, where the optimal precoder is derived in closed form and the corresponding design algorithm is developed. Finally Section 4.4 reports the numerical simulation results, and Section 4.5 concludes this chapter.

4.1 Problem Formulation

System performance can be significantly benefited from precoding, when the channel information is known at the transmitter side. Throughout this chapter, it is assumed that the channel state information (CSI) is available at the transmitter. CSI can be acquired at the transmitter either if a closed loop system with feedback channel is present, or when the transmitter and receiver operate in time division duplex (TDD) where the two links (downlink and uplink) share the same frequency band so that the time-invariant MIMO channel transfer function is the same in both ways, and the channel impulse response can be estimated from the uplink for downlink processing [43]. Although theoretically our design will be valid for time-variant channels as well, we will only consider the precoder design for the time-invariant system in this section. This is due to the fact that the assumption of the knowledge of the CSI

at the transmitter is not realistic in the case of time-varying channels. For practical rapidly varying fading channels, feedback of CSI is not up-to-date, which may result in performance degradation.

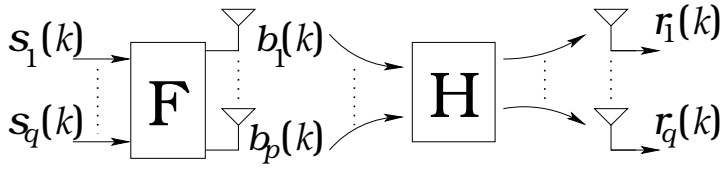

Figure 4.1: MIMO channel with p transmitters and q receivers

Consider a narrow-band MIMO time-invariant channel with p transmitters and q receivers, as shown in Figure 4.1. We deal with the case when $p > q$ for the precoder design which is a commonly used assumption for the downlink model, where the transmitter side can have abundant resource uses. The precoder $\mathbf{F}$ maps q streams of input symbols $\{s_i(k)\}_{i=1}^{q}$ to p streams of symbols $\{b_\ell(k)\}_{\ell=1}^{p}$ at the transmitter site, where the integer k denotes the time index. The precoded p data streams $\{b_\ell(k)\}_{\ell=1}^{p}$ are transmitted through the MIMO channel $\mathbf{H}$. By the assumption, the received signal consists of q data streams $\{r_j(k)\}_{j=1}^{q}$, corrupted by additive observation noise. Let $h_{j,\ell}(k)$ be the channel impulse response from the ℓth transmitter ($\ell = 1, 2, \cdots, p$) to the jth receiver ($j = 1, 2, \cdots, q$). Let $f_{\ell,i}(k)$ be the precoder impulse response from the ith input symbol stream $\{s_i(k)\}$ ($i = 1, 2, \cdots, q$) to the ℓth transmitter. Then the

received distorted sequence at the receiver j is given by

$$r_j(k) = \sum_{\ell=1}^{p} h_{j,\ell}(k) \star b_\ell(k) + v_j(k), \quad j = 1, 2, \cdots, q, \tag{4.1}$$

where the noise $\{v_j(k)\}_{j=1}^{q}$ is assumed to be i.i.d. additive white Gaussian noise (AWGN), and the signal $b_\ell(k)$ at the ℓth transmitter is generated via

$$b_\ell(k) = \sum_{j=1}^{q} f_{\ell,i}(k) \star s_i(k), \quad \ell = 1, 2, \cdots p. \tag{4.2}$$

The q streams of symbols $\{s_i(k)\}_{i=1}^{q}$ are assumed to be uncorrelated and white. If not, a pre-whitening operation can be performed over the symbol blocks prior to precoding [43]. Hence both input symbols, and AWGN are assumed to be zero mean with covariance matrices given by

$$R_s = \sigma_s^2 I, \quad R_v = \sigma_v^2 I, \tag{4.3}$$

respectively. It follows that the signal-to-noise ratio (SNR) equals to σ_s^2/σ_v^2 for each sub-channel, which is a constant. Our goal is to design precoder $\mathbf{F}$ to eliminate the ISI and ICI, and to minimize the BER of a downlink MIMO communication system by preprocessing the transmitted signal.

Since both the MIMO channel $\mathbf{H}$, and the precoder $\mathbf{F}$ are linear, and time-invariant, they admit transfer function matrices $H(z)$, and $F(z)$ respectively. A baseband equivalent MIMO communication system with precoder can be illustrated with block diagram in Figure 4.2 where $\mathbf{s}(k)$ is the input signal of size $q \times 1$, $\mathbf{v}(k)$ is the additive white Gaussian noise of size $q \times 1$, and $\mathbf{r}(k)$ is the output signal of size $q \times 1$,

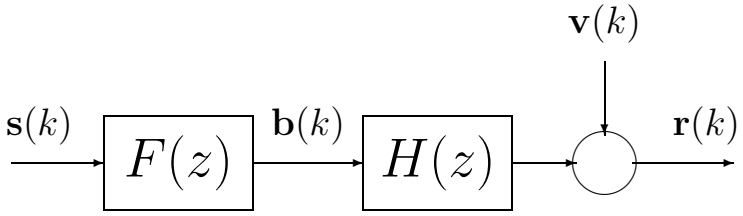

Figure 4.2: A narrow-band communication system with precoder

all are the blocked version of the original scalar data streams. Consequently $F(z)$ represents the precoder with size $p \times q$, and $H(z)$ the linear time-invariant channel with size $q \times p$. Assume that the linear channel $H(z)$ is causal, and stable. We say that the channel $H(z)$ is *perfectly recoverable* (PR), if there exists a linear, causal, and stable $F(z)$ such that

$$H(z)F(z) = \text{diag}\left(z^{-d_1}, z^{-d_2}, \cdots, z^{-d_q}\right), \tag{4.4}$$

for some positive integers $\{d_i\}_{i=1}^q$. Roughly speaking $d_i \geq 0$ is the transmission delay from the input signal $s_i(k)$ to the observed output $r_i(k)$ with $1 \leq i \leq q$. The PR condition implies that the q data streams $\{s_i(k)\}_{i=1}^q$ can be perfectly recovered in the absence of noise, and thus achieve the channel equalization. Without loss of generality, we can assume that

$$\text{diag}\left(z^{d_1}, z^{d_2}, \cdots, z^{d_q}\right) H(z)$$

is a causal transfer function. For this reason we may consider directly

$$H(z)F(z) = I, \tag{4.5}$$

which is termed as the *zero-forcing* condition. The following result holds.

Proposition 4.1 *Let $H(z)$ be a stable transfer function matrix with size $q \times p$, and $q < p$. Then the zero-forcing condition is true for some stable transfer function matrix $F(z)$, if and only if*

$$\text{rank}\{H(z)\} = q \ \ \forall \ |z| \geq 1 \text{ and } z = \infty. \tag{4.6}$$

Proof: Suppose that the zero-forcing condition (4.6) is true. Then the hypothesis $F(z)$ being stable implies that $F(z)$ is bounded for each z on and outside the unit circle. It follows from the zero-forcing condition (4.5) that $H(z)$ is bounded, and has full rank q for all z on and outside the unit circle. Moreover in the limit $H(\infty)F(\infty) = I$ implying that $H(\infty)$ has full rank q as well by the causality of $H(z)$, and $F(z)$. Hence the zero-forcing condition (4.5) implies (4.6). It can also be shown that the condition (4.6) ensures the zero-forcing condition (4.5), which will be made clear in the next section, when it is used to construct the optimal precoder. ∎

It should be clear by $p > q$ that the zero-forcing condition (4.5), once satisfied for some causal and stable $F(z)$, admits infinitely many solutions $F(z)$, which are causal, and stable. In the next section we will parameterize all causal and stable $F(z)$ such that (4.5) is true. Our design objective in this chapter is to optimize certain performance index among all such $F(z)$s. Clearly BER is the most desirable performance measure, which is determined by the related SNR. To analyze the associated BER performance, suppose that the zero-forcing condition (4.5) holds, then the received signal is given by

$$r_i(k) = s_i(k) + v_i(k), \quad i = 1, 2, \cdots, q. \tag{4.7}$$

Or in a blocked (vector) version, we can write (4.7) as

$$\mathbf{r}(k) = \mathbf{s}(k) + \mathbf{v}(k). \tag{4.8}$$

In light of (4.3), and the discussion in the previous section, the SNR for each sub-channel i at the receiver site is given by

$$\text{SNR} = \frac{\sigma_s^2(i)}{\sigma_v^2(i)} = \frac{\sigma_s^2}{\sigma_v^2} = \frac{1}{q}\sum_{i=1}^{q} \frac{\sigma_s^2(i)}{\sigma_v^2(i)} = \text{Average SNR}. \tag{4.9}$$

That is, the SNR for each sub-channel i is the same at the receiver site. It follows that

$$\text{BER} = \text{Q}\left(\sqrt{\text{Average SNR}}\right) = \text{Q}\left(\sqrt{\sigma_s^2/\sigma_v^2}\right), \tag{4.10}$$

by assuming the binary phase shift keying (BPSK) modulation, where

$$\text{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\tau^2/2)\ d\tau$$

is a monotonically decreasing function. Hence minimization of BER is equivalent to maximization of σ_s^2. However the total transmission power at the transmitter site is limited, and given by

$$\begin{aligned}
\sigma_{\mathbf{b}}^2 &= E[\mathbf{b}^*(k)\mathbf{b}(k)] = \text{trace}\left\{E[\mathbf{b}(k)\mathbf{b}^*(k)\right\} \\
&= \text{trace}\left\{\int_0^1 F(e^{j2f\pi})E[\mathbf{s}(k)\mathbf{s}^*(k)]F^*(e^{-j2f\pi})\ df\right\} \\
&= \text{trace}\left\{\int_0^1 F(e^{j2f\pi})F^*(e^{-j2f\pi})\ df\right\}\sigma_s^2 \\
&=: \|F\|_2^2\sigma_s^2 \le pE_b
\end{aligned} \tag{4.11}$$

with E_b the bit energy at each transmitter, and $\|F\|_2$ the $\mathcal{H}_2$-norm for the precoder.

It follows that the SNR in (4.9) at the receiver site is bounded as

$$\text{SNR} = \frac{\sigma_s^2}{\sigma_v^2} \leq \frac{pE_b}{\sigma_v^2 \|F\|_2^2} \tag{4.12}$$

with the right hand side the maximum SNR achievable, yielding the smallest BER. Hence minimization of BER (4.10) is equivalent to maximization of SNR (4.9), or the right hand side of (4.12), which is in turn equivalent to minimization of $\|F\|_2$.

Remark 4.1 The optimal precoder design considered in this chapter aims to synthesize the precoder $F(z)$ such that BER performance is optimized, or equivalently $\|F\|_2$ is minimized, subject to the zero-forcing condition (4.6). It will be shown in the next section that the optimal precoder can be designed using the latest result in optimal control theory. We note that minimization of $\|F\|_2$ as entailed for BER performance is the same in spirit with that in [26] for the precoder design, where minimization of each sub-channel of the precoder is the focal point. We also note that the MIMO channel under investigation can be equivalently decomposed into parallel independent sub-channels [31, 43], and such a rearrangement should not result in the loss of information. More importantly our results in the next two sections will show that the optimal precoder is IIR in general, among all precoders satisfying the zero-forcing condition. Thus optimal FIR precoders as used in the existing literature are suboptimal at best, and they approach to the optimal one obtained in this chapter as their orders increase to infinity.

4.2 Full Information Optimal Control

As discussed in the previous section, optimal precoder with the minimum BER criterion is equivalent to a constrained optimization problem: Design a causal and stable $F(z)$ such that $\|F\|_2$ is minimized, subject to the zero-forcing condition (4.6). We will derive an optimal solution in the closed form based on the latest result from optimal control theory.

We consider the following state-space model

$$\begin{aligned} \mathbf{x}(k+1) &= A\mathbf{x}(k) + B_1\mathbf{d}(k) + B_2\mathbf{u}(k), \quad \mathbf{x}(0) = 0 \\ \mathbf{w}(k) &= C\mathbf{x}(k) + D_1\mathbf{d}(k) + D_2\mathbf{u}(k), \end{aligned} \tag{4.13}$$

where $\mathbf{x}(k)$ of dimension n is the system state vector at the discrete time k. The disturbance input $\mathbf{d}(k)$ of size m_1 is a white random process with mean zero, and covariance identity. The control input $\mathbf{u}(k)$ with size m_2 has the access of both state vector $\mathbf{x}(k)$, and disturbance $\mathbf{d}(k)$ (full information). The design objective is to synthesize a full information control law

$$\mathbf{u}(k) = K[\mathbf{x}(k), \mathbf{d}(k)] \tag{4.14}$$

with $K[\cdot,\cdot]$ continuous such that the mean-squared value of the controlled output

$$J = E\left[\mathbf{w}^*(k)\mathbf{w}(k)\right] = \mathrm{tr}\left\{E\left[\mathbf{w}(k)\mathbf{w}^*(k)\right]\right\} \tag{4.15}$$

is minimized. If the full information control law in (4.14) is stabilizing, then the above J is finite. This problem has a known solution [69].

Theorem 4.2 *[69] Consider the state-space system* (4.13). *Assume that* $\text{rank}\{D_2\} = m_2$, *and*

$$\text{rank}\begin{bmatrix} A - zI & B_2 \end{bmatrix} = n \;\forall\; |z| \geq 1, \quad \text{rank}\begin{bmatrix} A - zI & B_2 \\ C & D_2 \end{bmatrix} = n + m_2 \;\forall\; |z| = 1. \tag{4.16}$$

Then the following algebraic Riccati equation (*ARE*)

$$X = A^*XA - (A^*XB_2 + C^*D_2)(D_2^*D_2 + B_2^*XB_2)^{-1}(A^*XB_2 + C^*D_2)^* + C^*C \tag{4.17}$$

admits an unique solution $X \geq 0$. *In this case the optimal full information control law has the form of linear constant feedback*

$$\mathbf{u}_{\text{opt}}(k) = K_1\mathbf{x}(k) + K_2\mathbf{d}(k) \tag{4.18}$$

which is stabilizing, with K_1, K_2 *given by*

$$K_1 = -(D_2^*D_2 + B_2^*XB_2)^{-1}(B_2^*XA + D_2^*C), \tag{4.19}$$

$$K_2 = -(D_2^*D_2 + B_2^*XB_2)^{-1}(B_2^*XB_1 + D_2^*D_1). \tag{4.20}$$

It is interesting to observe that the optimal control law (4.18) results in the feedback system

$$\begin{aligned} \mathbf{x}(k+1) &= (A + B_2K_1)\mathbf{x}(k) + (B_1 + B_2K_2)\mathbf{d}(k), \quad \mathbf{x}(0) = 0, \\ \mathbf{w}(k) &= (C + D_2K_1)\mathbf{x}(k) + (D_1 + D_2K_2)\mathbf{d}(k). \end{aligned} \tag{4.21}$$

Hence the transfer function matrix for the feedback system (4.21) is given by

$$T_{dw}(z) = (D_1 + D_2K_2) + (C + D_2K_1)(zI - A - B_2K_1)^{-1}(B_1 + B_2K_2). \tag{4.22}$$

We will denote $T_{dw}(z)$ compactly as

$$T_{dw}(z) =: \left[\begin{array}{c|c} A + B_2K_1 & B_1 + B_2K_2 \\ \hline C + D_2K_1 & D_1 + D_2K_2 \end{array}\right], \tag{4.23}$$

which is IIR in general. Because $\mathbf{d}(k)$ is white with covariance identity, we obtain the minimum achievable mean-squared value for the controlled output:

$$J_{\min} = \mathrm{tr}\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}[T_{dw}(e^{j\omega})][T_{dw}(e^{j\omega})]^*\, d\omega\right\} = \|T_{dw}\|_2^2. \tag{4.24}$$

Remark 4.2 It is noted that (4.14) includes all admissible control laws, including nonlinear, and dynamical ones. However the optimal full information control law is given as in (4.18), which has the linear constant feedback form, and which results in IIR closed-loop systems. In light of the well known Parserval's Theorem, the optimal full information control also minimizes the power norm of the output of the feedback system (4.21) in time-domain. Moreover the optimal feedback gains have the closed-form as in (4.19), and (4.20), which can be used to design optimal precoders.

4.3 Optimal Precoders for MIMO Channels

We consider the MIMO channel of p inputs, and q outputs, which admits a state-space realization

$$H(z) = D + C(zI - A)^{-1}B =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]. \tag{4.25}$$

Every channel $H(z)$ admits a state-space realization with A stable, or $\rho(A) < 1$ where $\rho(\cdot)$ denotes the spectral radius, because of its causality and stability. In particular for the FIR channel

$$H(z) = H_0 + H_1z^{-1} + H_2z^{-2} + \cdots + H_lz^{-l}, \quad H_l \neq 0, \tag{4.26}$$

where each H_i is a $q \times p$ matrix for $0 \le i \le l$ with l the length of the FIR channel, it admits a simple realization (A, B, C, D) as

$$A = \begin{bmatrix} 0 & I_q & 0 & \cdots & 0 \\ 0 & 0 & I_q & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & I_q \\ 0 & \cdots & \cdots & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ \vdots \\ H_l \end{bmatrix}, \quad C = \begin{bmatrix} I_q & 0 & \cdots & 0 \end{bmatrix}, \quad D = H_0. \tag{4.27}$$

In the above A is of size $n \times n$ with $n = ql$, B of size $n \times p$, C of size $q \times n$, and D of size $q \times p$. Clearly such an A is stable. Recall that the optimal precoder design seeks an optimal precoder $F(z)$ such that $\|F\|_2$ is minimized, subject to the zero-forcing condition (4.6). We will show that such an optimal design problem is equivalent to the optimal full information control as in the previous subsection.

We assume that D in (4.25) has full row rank. Then there exists $D_\perp$ such that $D_a = \begin{bmatrix} D \\ D_\perp \end{bmatrix}$ is square and nonsingular. A specific $D_\perp$ can be chosen from the minimum rank Cholesky factorization

$$D_\perp^* D_\perp = I - D^*(DD^*)^{-1}D. \tag{4.28}$$

Thus there exist right inverses of D and $D_\perp$, denoted by D^+ and $D_\perp^*$, respectively, and satisfy

$$D_a D_a^{-1} = \begin{bmatrix} D \\ D_\perp \end{bmatrix} \begin{bmatrix} D^+ & D_\perp^* \end{bmatrix} = I. \tag{4.29}$$

A particular right inverse can be chosen as

$$D^+ = D^*(DD^*)^{-1}. \tag{4.30}$$

It follows that $H(z) = D\left[I + (D^+C - D_\perp^* K)(zI - A)^{-1}B\right]$ for any K of size $p \times n$. Taking right inverse of $H(z)$ we have

$$H^+(z) = \left[\begin{array}{c|c} A_0 + BD_\perp^* K & BD^+ \\ \hline -D^+C + D_\perp^* K & D^+ \end{array}\right], \quad A_0 = A - BD^+C.$$

By the definition of $D_\perp$, $H_a(z) = D_a\left[I + (D^+C - D_\perp^* K)(zI - A)^{-1}B\right]$ is square, and strict minimum phase, if $H(z)$ is, and K is stabilizing. In this case,

$$H_a^{-1}(z) = \left[\begin{array}{cc} H^+(z) & H_\perp^+(z) \end{array}\right] = \left[\begin{array}{c|c} A_0 + BD_\perp^* K & B \\ \hline -D^+C + D_\perp^* K & I \end{array}\right] \left[\begin{array}{cc} D^+ & D_\perp^* \end{array}\right]. \tag{4.31}$$

There thus holds

$$H_a(z)H_a^{-1}(z) = \left[\begin{array}{c} H(z) \\ H_\perp(z) \end{array}\right] \left[\begin{array}{cc} H^+(z) & H_\perp^+(z) \end{array}\right] = I. \tag{4.32}$$

Proposition 4.3 *Consider channel transfer function matrix $H(z)$ of size $q \times p$ as in* (4.25) *with $\rho(A) < 1$, and* $\mathrm{rank}\{D\} = q$. *Suppose that $H(z)$ is strict minimum phase. Then there exists stabilizing K such that $A_0 + BD_\perp^+ K$ is stable. In this case, $H^+(z)$ and $H_\perp^+(z)$ as in* (4.31) *are stable, and satisfies* (4.32). *Moreover all stable right inverses of $H(z)$ are given by*

$$H^{(\mathrm{rinv})}(z) = H^+(z) + H_\perp^+(z)Q(z), \tag{4.33}$$

where $Q(z)$ is an arbitrary stable transfer function matrix.

Proof: Suppose that $H(z)$ is strict minimum phase, then

$$\mathrm{rank}\left\{\left[\begin{array}{cc} A - zI & B \\ C & D \end{array}\right]\right\} = n + q \;\forall\; |z| \geq 1.$$

Since elementary operations do not change rank,

$$\begin{aligned} n+q &= \operatorname{rank}\left\{\begin{bmatrix} A-zI & B \\ C & D \end{bmatrix}\begin{bmatrix} I & 0 & 0 \\ D_\perp^* K - D^+ C & D_\perp^* & D^+ \end{bmatrix}\right\} \\ &= \operatorname{rank}\left\{\begin{bmatrix} A_0 + BD_\perp^* K - zI & BD_\perp^* & BD^+ \\ 0 & 0 & I \end{bmatrix}\right\} \\ &= q + \operatorname{rank}\left\{\begin{bmatrix} A_0 + BD_\perp^* K - zI & BD^+ \end{bmatrix}\right\} \end{aligned}$$

which implies the existence of K such that $A_0 + BD_\perp^* K$ is stable. Thus causal and stable right inverses $H^+(z)$ exist.

By $H(z)H^+(z) = I$, and $H_\perp(z)H^+(z) = 0$, we have for $H^{(\mathrm{rinv})}(z)$ as in (4.33),

$$H(z)H^{(\mathrm{rinv})}(z) = I.$$

Thus $H^{(\mathrm{rinv})}(z)$ as in (4.33) is a stable right inverse of $H(z)$. Conversely for any stable right inverse $H^{(\mathrm{rinv})}(z)$, we claim that it has the form (4.33) with $Q(z) = H_\perp(z)H^{(\mathrm{rinv})}(z)$, due to

$$\begin{bmatrix} H(z) \\ H_\perp(z) \end{bmatrix} H^{(\mathrm{rinv})}(z) = \begin{bmatrix} I \\ Q(z) \end{bmatrix}$$

where $Q(z)$ is indeed stable. ∎

It is noted that the parameterized stable right inverse can be written as

$$\begin{aligned} H^{(\mathrm{rinv})}(z) &= \left[\begin{array}{c|c} A_0 + BD_\perp^* K & B \\ \hline -D^+ C + D_\perp^* K & I \end{array}\right](D^+ + D_\perp^* Q) \\ &= \left[\begin{array}{c|c} A_0 + BD_\perp^* K & BD^+ + BD_\perp^* Q \\ \hline -D^+ C + D_\perp^* K & D^+ + D_\perp^* Q \end{array}\right] \end{aligned} \tag{4.34}$$

by an abuse of notation. We thus observe that (4.34) has the same form as (4.23), based on which the following result can be obtained.

Theorem 4.4 *Under the same hypotheses as in Proposition 4.3, there hold* $D_\perp D_\perp^* > 0$, *and*

$$\begin{aligned} &\text{(i)} \quad \text{rank}\begin{bmatrix} A_0 - zI & BD_\perp^* \end{bmatrix} = n \;\forall\; |z| \geq 1, \\ &\text{(ii)} \quad \text{rank}\begin{bmatrix} A_0 - zI & BD_\perp^* \\ -D^+C & D_\perp^* \end{bmatrix} = n + p - q \;\forall\; |z| = 1, \end{aligned} \tag{4.35}$$

with $A_0 = A - BD^+C$. *Then the optimal precoder is the right inverse* $H^{(\text{rinv})}(z)$ *of* $H(z)$ *with the minimum* $\mathcal{H}_2$*-norm, given by*

$$F_{\text{opt}}(z) = H_{\text{opt}}^{(\text{rinv})}(z) = \left[\begin{array}{c|c} A_0 + BD_\perp^* K_{\text{opt}} & BD^+ + BD_\perp^* Q_{\text{opt}} \\ \hline -D^+C + D_\perp^* K_{\text{opt}} & D^+ + D_\perp^* Q_{\text{opt}} \end{array}\right], \tag{4.36}$$

$$K_{\text{opt}} = -(I + D_\perp B^* X B D_\perp^*)^{-1} D_\perp B^* X A_0, \tag{4.37}$$

$$Q_{\text{opt}} = -(I + D_\perp B^* X B D_\perp^*)^{-1} D_\perp B^* X B D^+, \tag{4.38}$$

where $X \geq 0$ *is the unique solution to the ARE*

$$X = A_0^* X A_0 - A_0^* X B D_\perp^* (I + D_\perp B^* X B D_\perp^*)^{-1} D_\perp B^* X A_0 + C^* (D D^*)^{-1} C. \tag{4.39}$$

Proof: The expression for $F(z) = H^{(\text{rinv})}(z)$ as in (4.34) indicates that design of the optimal precoder is equivalent to that of the optimal full information control for the state-space system

$$\mathbf{x}(k+1) = A_0 \mathbf{x}(k) + BD^+ \mathbf{s}(k) + BD_\perp^* \mathbf{u}(k), \quad \mathbf{x}(0) = 0, \tag{4.40}$$

$$\mathbf{w}(k) = -D^+ C \mathbf{x}(k) + D^+ \mathbf{s}(k) + D_\perp^* \mathbf{u}(k) \tag{4.41}$$

where the input to $F(z)$ is $\mathbf{s}(k)$, which is white by the assumption, and $\mathbf{u}(k) = K\mathbf{x}(k) + Q\mathbf{s}(k)$. The strict minimum phase assumption implies (i) of (4.35). For (ii)

of (4.35), the same assumption implies

$$\text{rank}\left\{\begin{bmatrix} A_0 - zI & BD_\perp^* \\ -D^+C & D_\perp^* \end{bmatrix}\right\} = \text{rank}\left\{\begin{bmatrix} A_0 - zI & BD_\perp^* \\ -C & 0 \\ 0 & I \end{bmatrix}\right\} = n + p - q \;\; \forall\, |z| = 1$$

by elementary operation. Hence Theorem 4.2 applies to optimal control problem for the state-space system in (4.40) and (4.41). Since the optimal control input over all stabilizing K, and stable dynamical $Q(z)$ is given in Theorem 4.2, it is straightforward to obtain the results of the theorem. ∎

As discussed earlier, the existing optimal precoders are almost exclusively FIR type, and the design of optimal precoders assumes the coprimeness condition for the channel in order to satisfy the zero-forcing condition. Theorem 4.4 indicates that the true optimal precoder is IIR in general. Thus the optimal precoder results in the existing literature are only suboptimal, and they approach to the optimal one, as the orders of the FIR precoders approach to infinity. It should also be noted that the strict minimum phase condition in this chapter is weaker than the coprimeness condition in the existing literature.

Design Algorithm and Its Complexity

We summarize the design algorithm as follows.

- Step 1: For a given MIMO FIR channel $H(z)$, find its state-space realization (A, B, C, D) according to (4.27).
- Step 2: For the case of full row rank D, compute $D_\perp$, as in (4.28), and D^+ as in (4.30), respectively.

- Step 3: Verify condition (i) and (ii) as in (4.35), and if so, compute the unique solution $X \geq 0$ to the discrete control ARE (4.39).
- Step 4: Compute the optimal gains K, and Q according to (4.37), and (4.38), respectively.
- Step 5: Set the optimal precoder $F(z)$ as in (4.36).

For the design algorithm summarized above, we can see that the computational complexity is mainly in Step 3. In fact conditions (i) and (ii) are the existence condition for $X \geq 0$ to the ARE (4.39) with computational complexity $\mathcal{O}(n^3)$. MATLAB command "dare" can be used to accomplish Step 3.

4.4 Illustrative Examples

In this section, we provide two simulation examples to demonstrate the feasibility and effectiveness of the optimal precoder design technique established in Theorem 4.4.

Example 4.1 Our first example focuses on the performance evaluation of the optimal precoder designed with our proposed technique, and compared with that of the Bezout precoder in [26]. Often the performance of the wireless communication system is measured by the average BER and by the average bit rate (bit/s) that the wireless link can provide. The channel model used for simulation matches the single user downlink communication system using BPSK modulation, as specified in Section 4.1. One hundred ISI channels of order $l = 5$ are randomly generated, each having 4-input/2-output. Then the optimal precoders are designed for each random channel,

and average BER curves associated with each optimal precoder are computed for different SNR values, defined as E_b/N_o, based on the results presented in Section 4.1. The obtained BER curves are then averaged over all the randomly generated channels and plotted in Figure 4.3 with the solid line to provide the BER performance for the optimal precoder design.

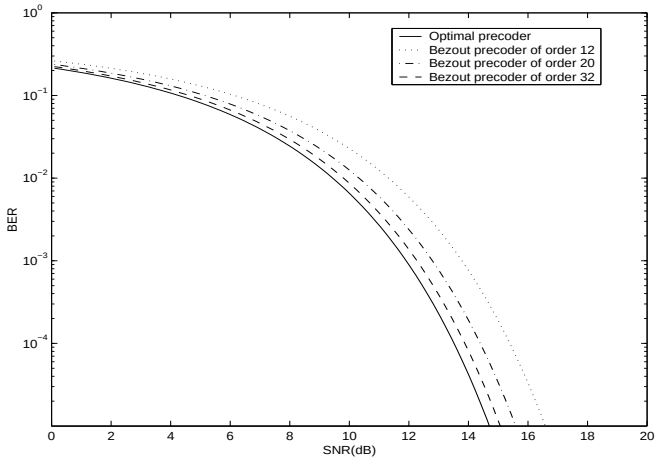

Figure 4.3: BER comparison for the optimal precoder and Bezout precoders

To compare with the Bezout precoder design proposed in [26], BER curves for the Bezout precoder of order 12, 20 and 32 are plotted with the dotted line, dash-dotted line, and dashed line in Figure 4.3, respectively. These curves are also averaged over 100 designs in order to make a fair comparison. The optimal precoder designed with our proposed technique is labeled "Optimal precoder", and the one from [26] is labeled "Bezout precoder". It is seen that when the order of the Bezout precoder becomes larger, the BER performance gets better, which approaches to that of the optimal precoder. Hence the Bezout precoder in [26] is suboptimal in terms of BER, and the

optimality can be reached asymptotically as the order of the Bezout precoders tends to infinity, as we have pointed out at the end of Remark 4.1.

We now examine the system performance in terms of the transmission rate. First we compute the MIMO channel capacity in the Shannon theoretic sense, which is plotted in Figure 4.4 in dashed line, and serves as the upper bound for the transmission rate of the MIMO system (see [3], [21]). The Shannon capacity of a single user time-invariant channel corresponds to the maximum data rate that can be transmitted over the channel with arbitrary small error probability.

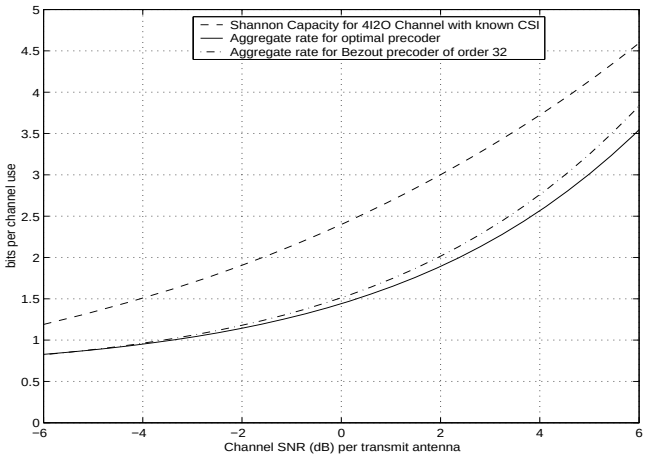

Figure 4.4: Aggregate rate comparison for the optimal precoder and Bezout precoder

Then we define the aggregate rate in terms of SNR as [26]

$$\text{Aggregate Rate} =: \sum_{k=1}^{q} \frac{1}{4\pi} \int_{-\pi}^{\pi} \log_2(1 + \text{SNR}\|F_k(e^{j\theta})\|^{-2}) d\theta,$$

where $F_k(z)$ for $k = 1, \cdots, q$, corresponds to the kth column vector of the precoder $F(z)$. In Figure 4.4, the solid line is the aggregate rate versus SNR for our optimal

precoder, while the dash-dotted line is the aggregate rate for the Bezout precoder of order 32. Both are averaged over the same randomly generated channels as before.

We can see from Figure 4.4 that for the optimal precoder proposed in this chapter, there is an average of about 1 bit difference from the Shannon capacity, indicating an effective design with the achievement of improved BER performance at the cost of a little bit worse rate performance than the Bezout precoder at high SNR.

■

Example 4.2 The multirate filterbank transceivers have been widely studied as an unified MIMO system framework [26, 41, 63]. They are applicable to OFDM (orthogonal frequency division multiplexing), DMT (discrete multitone), TDMA (time-division multiple access), and CDMA/DWMA (code-division/discrete-wavelet multiple access). Our second example focuses on the performance comparison between our optimal design and that in [63], where the SISO channel is chosen as

$$h(z) = \frac{1}{9}(1 + 2z^{-1} + 2.5z^{-2} + 2z^{-3} + z^{-4}).$$

The precoder (transmitter filterbank) in [63] is specified as

$$G = \begin{bmatrix} I_q \\ 0_{(p-q)\times q} \end{bmatrix}, \tag{4.42}$$

where p denotes the number of transmitter filters, and q the receiver filters (In [63] N is used in place of p to denote the N channels and K is used in place of q to denote the decimation K for the nonmaximally decimated multirate filterbank). We

first consider the case for $p = 3$ and $q = 2$. The $p \times p$ pseudo-circulant matrix with polyphase component [55] is computed as

$$\bar{H}(z) = \frac{1}{9}\begin{bmatrix} 1+2z^{-1} & 2.5z^{-1} & 2z^{-1}+z^{-2} \\ 2+z^{-1} & 1+2z^{-1} & 2.5z^{-1} \\ 2.5 & 2+z^{-1} & 1+2z^{-1} \end{bmatrix}.$$

Hence a virtual MIMO system is formulated. The optimal receiver satisfying the zero-forcing condition ($R(z)\bar{H}(z)G(z) = I$) is obtained as [63]:

$$\begin{aligned} R(z) &= [\bar{H}(z)G(z)]^{+} \\ &= 9\begin{bmatrix} -\frac{16}{35} - \frac{56}{35}z^{-1} - \frac{24}{35}z^{-2} & \frac{68}{35} + \frac{130}{35}z^{-1} + \frac{48}{35}z^{-2} & -\frac{34}{35} - \frac{96}{35}z^{-1} - \frac{36}{35}z^{-2} \\ \frac{4}{7} + \frac{12}{7}z^{-1} & -\frac{17}{7} - \frac{24}{7}z^{-1} & \frac{12}{7} + \frac{18}{7}z^{-1} \end{bmatrix}. \end{aligned} \quad (4.43)$$

For fair comparison, our example aims at designing an optimal precoder for the combined system consisting of $\bar{H}(z)$ and $R(z)$, i.e., to seek an optimal precoder $F(z)$ for $H(z) = R(z)\bar{H}(z)$ such that it satisfies $H(z)F(z) = R(z)\bar{H}(z)F(z) = I$ and achieves the minimum $\|F\|_2$ possible. By following the design recipe in Section 4.3, the optimal precoder $F(z)$ of size 3×2 is obtained with $\|F\|_2 = 1.2292$. For the precoder (4.42) with $p = 3$ and $q = 2$, it is obvious that $\|G\|_2 = 1.4142$. The frequency response of the precoder in (4.42) is plotted (solid line) along with the channel magnitude response (dashed line) in Figure 4.5.

As a comparison, the plot in Figure 4.6 depicts the average frequency response of our designed optimal precoder (solid line), defined by

$$|\mathcal{F}(f)| := \frac{1}{q}\sum_{k=1}^{q} |F_k(e^{j2f\pi})|,$$

together with the channel magnitude response (dashed line).

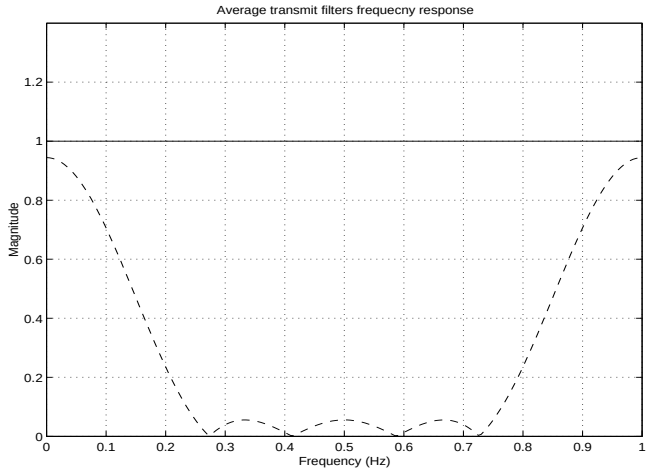

Figure 4.5: (Normalized) Frequency response of precoder versus channel

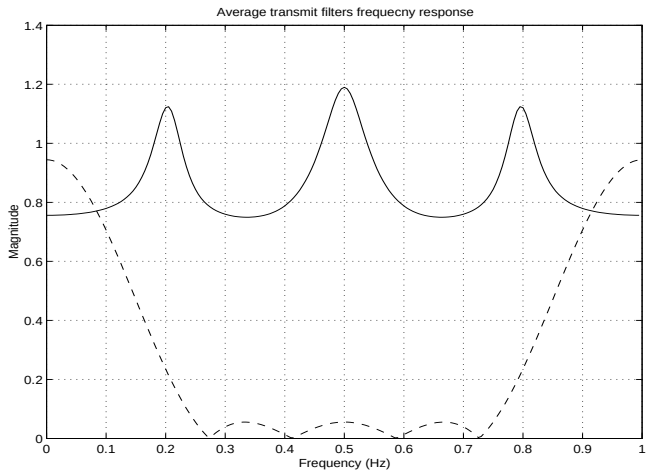

Figure 4.6: (Normalized) Frequency response of precoder versus channel

The BER performance for our optimal design and that in [63] is evaluated, and the BER curves are plotted in Figure 4.7 (solid line for our design and dashed line

for [63]). We can see from Figure 4.7 that our design outperforms the result in [63] regarding the average BER, which is ensured by the norm computation shown earlier as $\|F\|_2 = 1.2292 < 1.4142 = \|G\|_2$. This example shows that the BER performance in [63] can be further improved, in addition to the optimal receiver design in [63], via optimal design of precoders rather than using the simple precoder as in (4.42).

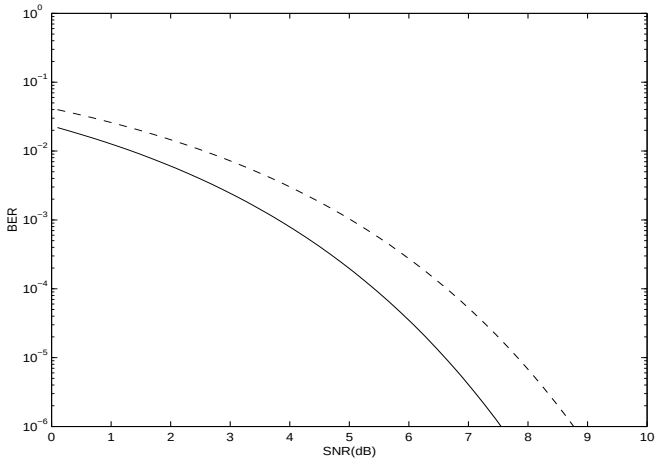

Figure 4.7: BER comparison for $p = 3$ and $q = 2$

If we choose $p = 4$ and $q = 2$, we have that the pseudo-circulant matrix is

$$\bar{H}(z) = \frac{1}{9}\begin{bmatrix} 1+z^{-1} & 2z^{-1} & 2.5z^{-1} & 2z^{-1} \\ 2 & 1+z^{-1} & 2z^{-1} & 2.5z^{-1} \\ 2.5 & 2 & 1+z^{-1} & 2z^{-1} \\ 2 & 2.5 & 2 & 1+z^{-1} \end{bmatrix}. \tag{4.44}$$

A simple receiver can be obtained as a constant matrix:

$$R = [\bar{H}(z)G]^{+} = \begin{bmatrix} 0 & 0 & 10 & -8 \\ 0 & 0 & -8 & 10 \end{bmatrix}. \tag{4.45}$$

The BER curve for such a design is plotted in Figure 4.8 with the dashed line.

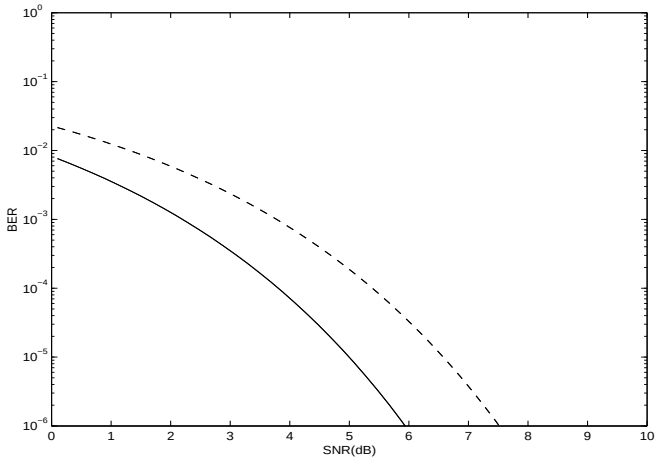

Figure 4.8: BER comparison for $p = 4$ and $q = 2$

Using our proposed design technique, an optimal precoder $F(z)$ of size 4×2 for the combined system $H(z) = R\bar{H}(z)$ is obtained with $\|F\|_2 = 1.1790$ satisfying

$$H(z)F(z) = R\bar{H}(z)F(z) = I.$$

Its BER curve is plotted in Figure 4.8 with solid line. We can see the better BER performance compared with those in Figure 4.7, which indicates the improvement by increasing the number of transmitters, and shows the benefit of adopting transmit diversity. Finally we also would like to point out that our optimal design does not suffer from the spectral-null property of the channel due to the introduction of redundancy, as can be clearly seen from the result in this example. ∎

4.5 Concluding Remark

In this chapter, the design of channel precoders which minimize the BER under PR (or zero-forcing) condition subject to a limited transmit power—a task entailing more than simple mitigation of ISI and ICI, was proposed for a MIMO downlink communication channel assuming the perfect channel knowledge. Analytic solution to the optimal precoder was derived, and an effective design algorithm was provided. The performance evaluations indicate that the optimal precoder design proposed herein is an attractive/reasonable alternative to the existing precoder design techniques when the highly constrained resource at the mobile unit is a priority.

Chapter 5

Conclusion and Future Research

This dissertation investigated two research problems in wireless data communications. The first one is optimal channel equalization for MIMO systems, where the channel is linear and time-invariant, and the corrupting noises are non-white with unknown power spectral densities. Under the multirate filterbank framework and the mean-squared error criterion, a worst-case approach to optimal channel equalization leads to an $\mathcal{H}_\infty$ optimization problem. The second one is optimal precoder design for multiple antennas systems, where the channel is again linear and time-invariant, but the corrupting noises are white and Gaussian. The design objective is to minimize the BER, assuming the perfect channel knowledge. It was shown that the optimal precoder has the form of optimal full information controllers. Complete solutions were derived for these two research problems, and simulation examples were worked out to illustrate the proposed design algorithms.

Much progress has been made in the research area of wireless data communications. This dissertation advances further the research work for MIMO communication

systems. Nevertheless the MIMO research on the equalization and precoding topics has largely yet to mature and there exist a number of unsolved problems. This chapter suggests several topics for possible future research, as a conclusion of this dissertation.

- The results in this dissertation are based on somewhat unrealistic assumptions about the underlying channel model and how well it can be tracked at the receiver as well as at the transmitter. More realistic assumptions can dramatically impact the potential performance of the proposed MIMO design techniques. In practice, signal detection is complicated by the imprecise channel knowledge in a dynamic mobile communication environment. The channel imprecision is mainly caused by channel fading, which can also come from the channel estimation error as a result of imperfect training, or alternatively be attributed to the gradual change in the channel response. Hence robust channel equalization and precoding in the presence of modeling errors need be investigated.

- For most of the cases in real applications, the channel model information is unavailable. When the channel model is not available, it is desirable to be estimated adaptively and blindly, because the use of training signals, or pilot tones can consume significant spectrum resources. For this reason *blind adaptive channel estimation* has been active [28, 59, 68]. If blind and adaptive channel estimation is used in conjunction with optimal channel equalization, *blind channel equalization* can be made possible. However, the overall performance employing the optimal channel equalization in conjunction with the blind and

adaptive channel estimation algorithm may not be the best achievable due to the existence of the modeling error. Even though blind channel equalization has been studied extensively in the literature [6, 8, 9, 23, 44, 42, 51], how to analyze, improve, and achieve the optimal overall performance for blind channel equalization does not have a complete solution, which needs exploration in the future.

- The optimal equalizer design for the filterbank transceiver in Chapter 3 of this dissertation, and that in [18] assume that the precoders are previously given. In most of the real applications, the precoder is a design subject also along with the equalizer, thus giving rise to the problem of joint optimal design, i.e., to seek the optimal precoder and equalizer jointly to achieve the best performance. Currently existing results for the joint design such as [29, 43, 39, 63, 65, 66] are almost exclusively FIR type. The main results obtained in this dissertation indicate that the true optimal equalizers and precoders are IIR in general. Hence the optimal design results in the existing literature are suboptimal at best. A possible approach to optimal joint design is to employ the Lagrange multiplier method used in the $\mathcal{H}_2$ and mixed $\mathcal{H}_2/\mathcal{H}_\infty$ optimal control problems, by setting the MMSE equalizer as a function of the parameterized precoder, and then searching for the optimal precoder which minimizes the BER in data detection. However our initial attempts encountered extremely complicated and lengthy equations which prevented us from obtaining the final optimal solution. Other

methods different from the Lagrange multiplier may have to be employed to tackle the optimal joint design problem.

There are many open problems for wireless data communications, especially for the MIMO wireless systems. The aforementioned research problems are only small portion of the growing research area, which are related to this dissertation work. It is hoped that the results in this dissertation, and in future research will eventually lead to considerably higher data rates with higher reliability for future wireless data communications than what we have today.

Bibliography

[1] S. L. Ariyavisitakul, "Turbo space time processing to improve wireless channel capacity," *IEEE Trans. Commun.*, pp. 1347-1359, Aug. 2000.

[2] R.S. Blum, Y. Li, J.H. Winters, and Q. Yan, " Improved Space Time Coding for MIMO-OFDM Wireless Communications," *IEEE Trans. Commun.*, No. 11, vol. 49, pp. 1873-1878, Nov. 2001.

[3] L.H. Brandenburg and A.D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case", *Bell Syst. Tech. J.*, vol. 53, no. 5, pp. 745-778, May-June 1974.

[4] M. Brandt-Pearce and A. Dharap, "Transmitter-bsaed multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environment," *IEEE Journal on Selected Area of Communications*, pp. 407-417, Mar. 2000.

[5] B.-S. Chen, C.-W. Lin, and Y.-L. Chen, "Optimal signal reconstruction in noisy filter systems: Multirate Kalman synthesis filtering approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 2496-2505, Nov. 1995.

[6] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Blind channel identification and equalization using non-redundant periodic modulation precoders: performance analysis," *IEEE Trans. Signal Processing*, vol. 48, no. 6, pp. 1570-1586, June 2000.

[7] P. S. Chow, J. C. Tu, and J.M. Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and VHDSL services," *IEEE J. Select. Areas Commun.*, vol. 9, no. 6, pp. 909-919, Aug. 1991.

[8] Z. Ding, "Blind Equalization Based on Joint Minimum Mean Square Error Criterion," *IEEE Transactions on Communications*, vol.42, pp. 648-654, Feb/March/April 1994.

[9] Z. Ding and Y. (Geoffrey) Li, *Blind Equalization and Identification*, Marcel Dekker, Inc., New York, 2000.

[10] G.D. Forney, Jr., and M.V. Eyuboğlu, "Combined equalization and coding using precoding," *IEEE Trans. Commun. Mag.*, pp. 25-34, Dec. 1991.

[11] J. Foschini, "Layered Space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol.1, no.2, pp. 41-59, 1996.

[12] J. Foschini and M. Gans, "On the limit of wireless communications in a fading environment when using multiple antennas", *Wireless Personal Communications*, vol.6, no.3, pp. 311-335, 1998.

[13] T.T. Georgiou and M.C. Smith, "Optimal robustness in the gap metric," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 673-685, 1990.

[14] G.B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Letters*, vol. 4, pp. 184-187, June 1997.

[15] K. Glover and D. McFarlane, "Robust stabilization of normalized coprime factor plant description with $\mathcal{H}_\infty$ bounded uncertainty," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 821-830, 1989.

[16] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867-1875, 1980.

[17] M. Green and D.J.N. Limebeer, *Linear Robust Control*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[18] G. Gu and E.F. Badran, "Optimal design for channel equalization via filterbank approach," www.ee.lsu.edu/ggu/new_pubs.html. Submitted to *IEEE Trans. Signal Processing*, July 2001.

[19] G. Gu and L. Li, "Worst-case design for optimal channel equalization in filterbank transceivers," to appear, *IEEE Trans. Signal Processing*, 2003.

[20] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Communication*, pp. 774-780, Aug. 1972.

[21] W. Hirt and J.L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 34, pp. 380-388, May 1998.

[22] P.A. Iglesias and K. Glover, "State-space approach to discrete-time $\mathcal{H}_\infty$ control," *Int. J. Contr.*, vol. 54, pp. 1031-1073, 1991.

[23] C.R. Johnson, Jr., P. Schniter, T.J. Endres, J. Behm, D.R. Brown, and R.A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review," *Proceedings of the IEEE, Special Issue on Blind System Identification and Estimation*, October 1998.

[24] T. Kailath, *Linear Systems*, Englewood Cliffs, NJ: Printice-Hall 1980.

[25] I. Kalet, "The multitone channel," *IEEE Trans. Commun.,* vol. 37, no. 2, pp. 119-224, Feb. 1989.

[26] S.-Y. Kung, Y. Wu, and X. Zhang, "Bezout space-time precoders and equalizers for MIMO channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2499-2514, Oct. 2002.

[27] J. W. Lechleider, "High bit rate digital subscriber lines: A review of HDSL progress," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 769-784, Aug. 1991.

[28] Y. Li, L.J. Cimini, Jr., and N.R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902-915, July 1998.

[29] T. Li and Z. Ding, " Joint transmitter-receiver optimization for partial response channels based on non-maximally decimated filterbank precoding technique, " *IEEE Transactions on Signal Processing*, vol. 47, No. 9, pp. 2407-2414, Sept. 1999.

[30] L. Li, G. Gu and H.-C. Wu, "Design of optimal precoders for MIMO channels," *IEEE 2003 Global Communications Conference*, Dec. 2003, San Francisco, CA.

[31] Y.-P. Lin and S.-M. Phoong, "ISI-free FIR filterbank transceivers for frequency-selective channels," *IEEE Transactions on Signal Processing*, vol. 49, pp. 2648-2658, 2001.

[32] Y.-P. Lin and S.-M. Phoong, "Optimal ISI-free DMT transceivers for distorted channels with colored noise", *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2702-2712, Nov. 2001.

[33] R. Liu and G. Dong, "A fundamental theorem for multiple-channel blind equalization, " *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 472-473, 1997.

[34] H. Luo and R.-W. Liu, "Blind equalization for MIMO FIR channels based only on second order statistics by use of pre-Filters," *Proc. of IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Annapolis, MD, USA, pp. 106-109, 1999.

[35] H.S. Malvar and D.H. Staelin, "Optimal pre-and post filters for multichannel signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 287-289, Feb. 1988.

[36] K.D. Minto, *Design of Reliable Control Systems: Theory and Computations*, Ph.D Dissertation, Dept. Elec. Eng., Univ. Waterloo, Waterloo, Canada, 1985.

[37] A. Paulraj, C. Papadias, "Space-time Processing for Wireless Communications," *IEEE Signal Processing Magazine*, Nov. 1997.

[38] G.S. Rajappan and M.L. Honig, "Signature sequence adaptation for DS-CDMA with multipath," *IEEE Journal On Selected Areas In Communication*, vol. 20, pp. 384-395, Feb. 2002.

[39] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198-2206, Dec. 2001.

[40] A. Scaglione, S. Barbarossa, and G.B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1019-1032, Apr. 1999.

[41] A. Scaglione, G.B. Giannakis and S. Barbarossa, "Redundant filterbank precoders and equalizers, Part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1983-2006, July 1999.

[42] A. Scaglione, G. B. Giannakis, S. Barbarossa, "Redundant Filterbank Precoders and Equalizers Part II: Blind Channel Estimation, Synchronization and Direct Equalization, " *IEEE Trans. Signal Processing*, vo. 47, no. 7, pp. 2007-2022, July 1999.

[43] A. Scaglione, P. Stoica, S. Barbarossa, G.B. Giannakis and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051-1064, May 2002.

[44] J. Shen and Z. Ding, "Direct Blind MMSE Channel Equalization Based on Second Order Statistics", *IEEE Transactions on Signal Processing*, SP-48(4): 1015-1022, April 2000.

[45] E. G. Ström, S. Parkval, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 84-93, 1996.

[46] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. Select. Areas Commun.*, vol. 17, no. 3, pp. 451-460, March 1999.

[47] V. Tarokh, A. Naguib, N. Seshadri, and A.R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1121-1128, May 1999.

[48] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, March 1998.

[49] E. Telatar, "Capacity of multiantenna Gaussian channels," *AT & T-Bell Lab., Internal Tech. Memo.*, June 1995.

[50] M. Tomlinson, "New automatic equaliser employing modulo arithmetric," *Electronics Letter*, pp. 138-139, March 1971.

[51] L. Tong, G. Xu, T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, no.2, pp. 340-349, March 1994.

[52] M.K. Tsatsanis and G.B. Giannakis, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, pp. 3044-3055, 1996.

[53] M.K. Tsatsanis and G.B. Giannakis, "Multirate filter banks for code division multiple access systems," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Detroit MI, vol. II, pp. 1484-1487, May 1995.

[54] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Processing*, vol. 53, pp. 211-229, 1996.

[55] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[56] S. Verdu, " Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319-1343, June 2002.

[57] B. R. Vojoic and W. M. Jang, "Transmitter precoding in synchronous multiuser communcations," *IEEE Transactions on Communications*, pp. 1346-1355, Oct. 1998.

[58] X. Wang and H.V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 40, pp. 91-103, 1998.

[59] X. Wang and K.J. Ray Liu, "Adaptive channel estimation using cyclic prefix in multicarrier modulation system," *IEEE Commun. Lett.*, vol. 3, pp. 291-293, Oct. 1999.

[60] J.H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment, " *IEEE J. Select. Areas Commun.*, pp. 871-878, June 1987.

[61] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*, Springer-Verlag, New York, 1985.

[62] T. F. Wong and T. M. Lok, " Transmitter adaptation in multicode DS-CDMA systems *IEEE Journal On Selected Areas In Communication*, vol.19, pp. 69-82, Jan. 2001.

[63] X.-G. Xia, "New precoding for intersymbol interference cancellation using non-maximally decimated multirate filterbanks with ideal FIR equalizers," *IEEE Trans. on Signal Processing*, vol. 45, pp. 2431-2441, Oct. 1997.

[64] I. Yaesh and U. Shaked, "A transfer function approach to the problem of discrete-time systems: $\mathcal{H}_\infty$-optimal linear control and filtering," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1264-1271, 1991.

[65] J. Yang and S. Roy, "Joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) with decision feedback," *IEEE Trans. Inform. Theory*, vol. 42, pp. 3221-3231, Sept. 1994.

[66] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221-3231, Dec. 1994.

[67] W. Yu, G. Ginis, and J. Cioffi, "An adaptive multiuser pwoer control algorithm for VDSL," *Proceedings of IEEE Globecom*, pp. 394-398, Oct. 2001.

[68] Q. Zhao and L. Tong, "Adaptive Blind Channel Estimation by Least Squares Smoothing," *IEEE Transactions on Signal Processing*, vol. 47, no. 11, pp. 3000-3012, Nov. 1999.

[69] K. Zhou, J. Doyle and K. Glover, *Robust and Optimal Control*, Englewood Cliffs, NJ: Printice-Hall 1996.

Vita

Lijuan Li was born in Liaoning, People's Republic of China. She received her Bachelor of Science degree in mathematics in 1996 from Liaoning University, P.R.China. Then she was awarded entry into graduate program with the admission test waived by Fudan University in 1996. She received the Master of Science degree in applied mathematics in 1999 from Fudan University, Shanghai, P.R.China. In spring 2000, she enrolled in the graduate program in Electrical and Computer Engineering department, Louisiana State University and has received the Master of Science degree in Electrical Engineering in May 2001 from Louisiana State University, Baton Rouge, Louisiana. She is currently a candidate for the degree of Doctor of Philosophy in the field of Electrical Engineering at Louisiana State University.