ON THE PERFORMANCE OF LINEAR MIMO TRANSCEIVERS

by

Ahmed Mohammed Hesham Mehana

APPROVED BY SUPERVISORY COMMITTEE:

_____

Dr. Aria Nosratinia, Chair

_____

Dr. Naofal Al-Dhahir

_____

Dr. Hlaing Minn

_____

Dr. Mohammad Saquib

Dedicated to my family.

ON THE PERFORMANCE OF LINEAR MIMO TRANSCEIVERS

by

AHMED MOHAMMED HESHAM MEHANA, B.S., M.S.

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

October 2012

ACKNOWLEDGMENTS

I give thanks and my appreciation:

To God who helped me during my life and during my Ph.D. period

To my parents, my sister and my wife who gave me (and are still giving) their support, understanding and help

To all my relatives who cared for me and supported me

To my adviser who guided me through many carefully-chosen research avenues

To my colleagues who were part of a friendly environment.

October 2012

PREFACE

This dissertation was produced in accordance with guidelines which permit the inclusion as part of the dissertation the text of an original paper or papers submitted for publication. The dissertation must still conform to all other requirements explained in the "Guide for the Preparation of Master's Theses and Doctoral Dissertations at The University of Texas at Dallas." It must include a comprehensive abstract, a full introduction and literature review, and a final overall conclusion. Additional material (procedural and design data as well as descriptions of equipment) must be provided in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported.

It is acceptable for this dissertation to include as chapters authentic copies of papers already published, provided these meet type size, margin, and legibility requirements. In such cases, connecting texts which provide logical bridges between different manuscripts are mandatory. Where the student is not the sole author of a manuscript, the student is required to make an explicit statement in the introductory material to that manuscript describing the student's contribution to the work and acknowledging the contribution of the other author(s). The signatures of the Supervising Committee which precede all other material in the dissertation attest to the accuracy of this statement.

ON THE PERFORMANCE OF LINEAR MIMO TRANSCEIVERS

Publication No. _____

Ahmed Mohammed Hesham Mehana, Ph.D.
The University of Texas at Dallas, 2012

Supervising Professor: Dr. Aria Nosratinia

In this dissertation, the reliability and throughput of equalization and precoding techniques are analyzed in both several multiple-input multiple-output (MIMO) as well as single-input single-output (SISO) scenarios, resulting in several new discoveries as well as explanation certain phenomena that were known but not fully understood.

The first part of this dissertation establishes the diversity of the minimum mean square error (MMSE) MIMO receiver for all fixed rates (spectral efficiencies) in the quasi-static flat-fading MIMO channel. It is shown that full spatial diversity is achieved for all antenna configurations if and only if the rate is below a threshold which itself is a function of the number of antennas. The diversity of the MIMO multiple access channel (MAC) is also obtained. Linear receivers for the quasi-static frequency selective MIMO channel are also analyzed.

The second part of the dissertation is dedicated to linear MIMO precoders, including Wiener filtering, regularized zero-forcing filtering and matched filtering. It is shown that regularized zero-forcing or matched filter suffer from error floors for all positive multiplexing gains. In the fixed-rate regime, these precoders achieve full diversity up to a certain spectral efficiency

and zero diversity at rates above it. The diversity in the presence of both linear precoding and linear equalization is also analyzed.

The third part of the dissertation investigates the performance of common transmit diversity techniques, such as Alamouti and cyclic-delay diversity schemes, when used with linear receivers. The effect of block length on the system performance is fully characterized.

The fourth and final part of the dissertation investigates decision feedback equalizers in SISO ISI channels. As part of the developments of this part, the notion of the *spectral representation* of random processes is used for a rigorous analytical framework of decision feedback equalizers.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Background

Wireless communications has seen a rise in popularity that has driven new discoveries. Many factors have contributed to this trend, including the increasing demand for connectivity, the progress of integrated circuit technology as well as the successful deployment of standards that have rapidly expanded the availability of wireless connectivity, thus creating more demand for new products and services, and the fundamental discoveries that make them possible. Despite great advances, many open problems in this area persist. This dissertation is dedicated to the solution of several long-standing open problems in the analysis of popular wireless transceivers.

Two phenomena in wireless communication produce challenges that are unique to the medium: fading and interference phenomena. The former describes the variation of the channel strength due to small-sale and large-scale effects of multipath fading; the latter describes the interactions between different transmitted signals in the same medium. Interference can naturally occur in a multi-user scenario, but it can also occur in a single-user scenario when multiple signal components of the same user interfere with each other, which can be thought of as self-interference. Examples of self-interference include signals emitting from multi-antenna transmitters (interference in space) or inter-symbol interference (ISI) which is interference across time. This dissertation analyzes various transceivers that are designed to *efficiently* address the question of fading and interference in wireless systems. With very few exceptions, most of the transceivers analyzed in this dissertation are linear and operate on multiple-input multiple-output (MIMO) channels.

1

Several metrics have been widely adopted to characterize the performance of wireless systems, among them the average probability of error and the outage probability. In the fading channel, these quantities are partially characterized by the notion of diversity [1, 2], i.e., the slope of error probability as a function of SNR in the log-log scale. While in non-fading channels the probability of error decreases exponentially with the signal-to-noise ratio (SNR), in fading channels the probability of error averaged over the channel distribution is proportional to SNR^{-d}, where d is the channel diversity.

At high SNR the channel throughput increases proportionally with $\log SNR$. It has been shown that a fundamental tradeoff exists between this throughput and the channel reliability as expressed by the diversity. This tradeoff was characterized for the MIMO channel under maximum likelihood decoding by the seminal paper of Zheng and Tse [2], leading to the widespread use of the Diversity-Multiplexing Tradeoff (DMT) as a metric of quality for a variety of wireless systems and algorithms.

1.2   Motivation and Objectives

The primary objective of this dissertation is an in-depth study of the performance limits of multiple input multiple output (MIMO) systems in the presence of linear receivers and/or linear precoders. Although several DMT results are also obtained, the main thrust of this work is in the fixed-rate regime, i.e., when the spectral efficiency R is independent of SNR.

The DMT is a powerful framework but it is not able to describe the diversity in the fixed rate regime, because the DMT cannot distinguish between different spectral efficiencies R that correspond to the same multiplexing gain r. In several practical systems, various spectral efficiencies R, all corresponding to the same multiplexing gain, indeed give rise to different diversities [1, 2]. The difference between calculating the diversity in the fixed rate regime versus in the DMT is further clarified in Chapter 2.

Fixed rate analysis of diversity requires new tools compared with DMT analysis, since certain terms and mathematical expressions in DMT analysis are asymptotically negligible

and can be ignored, while the counterparts in fixed-rate analysis are not negligible and must be handled delicately.

1.3   Outline and Contributions

Chapter 2 briefly discusses the basics and background material on which the remainder of the dissertation rests.

Chapter 3 analyzes the MIMO MMSE receiver under flat fading channel assumption and the diversity is explicitly characterized as a function of spectral efficiency R and the number of antennas. The MIMO multiple access channel (MAC) is similarly studied and the diversity of reception is computed.

Chapter 4 extends the results of the previous chapter and studies the performance of the MIMO MMSE receiver in the frequency selective channel under two common transmission schemes: the zero-padding and the cyclic-prefix tranmission.

Chapter 5 studies the MIMO precoding systems when channel state information is available at the transmitter. Several precoding filters are analyzed, including the zero-forcing, the regularized zero-forcing and the matched filter. The analysis reveals that the matched filter and the regularized zero-forcing are not always interference-limited. Their performance depends on the spectral efficiency and in some cases these two filters achieve the maximum spatial diversity in contrast to the zero-forcing filtering. Several other results pertaining to MIMO precoders were also obtained.

Chapter 6 extends the analysis of the previous chapters to MIMO systems with linear transmit *and* receive filters under the flat fading channel.

Chapter 7 analyzes two common transmit diversity techniques, Alamouti signaling and cyclic delay diversity, in the presence of the MMSE receiver. The effect of the block length, the antenna configurations, and the spectral efficiency on the system performance is fully characterized.

In order to improve the performance of the MIMO linear receivers, lattice-reduction aided equalization has been proposed in the literature for MIMO flat fading and SISO frequency selective channels. Chapter 8 analyzes the LR-aided equalization for MIMO frequency selective channel.

Chapter 9 analyzes the decision-feedback equalizer for the SISO ISI channel. As part of the developments of this chapter, the spectral representation of random processes is used to put the analysis of decision feedback equalizers on a more solid mathematical foundation.

PRELIMINARIES

This chapter provides a review some information-theoretic concepts and performance measures used throughout this dissertation.

2.1 Information Rate and Decoding Reliability

The framework to study performance limits in communication systems is *information theory* established by Claude Shannon in 1948. Shannon characterizes the limits of reliable communication by proving the surprising result that one can communicate at strictly positive rate with as small error probability as required. [1] Consider the complex baseband representation of the channel

$$y = hx + n \tag{2.1}$$

where x and y are the input and the output of the channel respectively, h is the channel (possibly fading) and n is the thermal noise at the receiver modeled as AWGN $\sim \mathcal{N}(0, \sigma_n^2)$ where σ_n^2 is the noise variance. The maximum rate of reliable communication supported by this channel depends on the process $\{h\}$ as well as how much information about $\{h\}$ is available at the transmitter and the receiver.

Conditional on the realization of h and assuming that the transmitter does not know this realization but the receiver does (this information denoted by the channel state information at the receiver CSI-Rx), the maximum rate of reliable communication supported by this channel is

$$\log(1 + |h|^2\rho) \quad \text{bits/s/Hz} \tag{2.2}$$

---

[1]Before Shannon it was widely believed that the only way to achieve reliable communication over a noisy channel was to reduce the data rate [3].

which is the capacity of AWGN channel with a fixed gain h and ρ is the signal-to-noise ratio (SNR). The quantity in (2.2) is indeed the Shannon capacity for AWGN under the above assumptions.

2.2 Shannon Capacity vs. Fading Capacity

If the channel $\{h\}$ is a fading process, as it is usually the case for wireless channel, then the quantity in (2.2) is random. If the channel h remains fixed over the transmission period of the the codeword then the model is *slow fading* model. It is quite possible that for some realizations of $\{h\}$ the channel cannot support the transmission rate. That is, if the transmitter encodes the data at a rate (spectral efficiency) of R bits/s/Hz then there is a non-zero probability that $\log(1 + |h|^2\rho) < R$ in which case the error probability is strictly positive. The Shannon capacity in this case is strictly zero. Instead it is possible to define the *outage* probability

$$P_{out}(R) \triangleq \mathbb{P}(\log(1 + |h|^2\rho) < R) \tag{2.3}$$

and correspondingly define the ϵ-*outage capacity* [4], which is the largest rate for which $P_{out(R)} \leqslant \epsilon$,

$$C_\epsilon \triangleq \sup \left\{ R : P_{out} \leqslant \epsilon \right\}. \tag{2.4}$$

These two measures, P_{out} and C_ϵ, are used to characterize the performance limits of the slow fading model where the codeword only spans one coherence period (i.e. the channel is fixed over this period). When the codeword spans more than one coherent periods the model is *fast fading* model. The capacity of the fast fading channel can be shown to be

$$C = \mathbb{E}[(\log(1 + |h|^2\rho)]. \tag{2.5}$$

which is achieved by large number of coherence time intervals.

An extreme case for the fast fading model is when the channel is varying so fast such that *each* codeword symbol experiences a different channel gain. An intermediate case between the *so-fast fading* model and the *slow fading* model is when h remains fixed over a period of

codeword symbols and is i.i.d. across different coherent periods. This model is called *block fading* model and is widely adopted in practical applications.

2.3 Diversity: Fixed-Rate Regime vs. Variable-Rate Regime

Reliable communication over fading channel depends on the strength of the signal path, i.e. the channel gain, which can sometime be in deep fade and therefore results in very poor communication. A natural solution to improve the communication performance is to ensure that the information symbol experience multiple independent channel gains through *diversity* techniques. Diversity techniques can be performed over time or frequency (e.g. coding and interleaving), space (e.g. multiple antennas) or networks (e.g. cellular networks) [3]. These techniques result in a probability of error P_e that is proportional ρ^{-d} at high SNR, where d is the diversity order of the system.

Another important parameter, especially in MIMO systems, is the *multiplexing gain* or *degree of freedom*. Consider an $M \times N$ MIMO channel model (where M and N are the number of transmit and receive antennas respectively)

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.6}$$

where $\mathbf{H}$ is a stationary ergodic random fading process. For the fast fading model and CSI-Rx and no CSI-Tx, a celebrated capacity result is given by

$$C_{\text{MIMO}} = \mathbb{E}\Big[\Big(\log \det \Big(1 + \frac{\rho}{M}\mathbf{H}\mathbf{H}^H\Big)\Big].$$ \tag{2.7}

with the assumption of equal power allocations for all transmitted streams. At high SNR, the MIMO capacity is approximated by

$$C_{\text{MIMO}} = n_{\min} \log \Big(\frac{\rho}{M}\Big) \approx n_{\min} C_{\text{AWGN}}. \tag{2.8}$$

where $n_{\min} \triangleq \min(M, N)$ and C_{AWGN} is the high-SNR capacity of AWGN single antenna channel with transmit equivalent SNR $\frac{\rho}{M}$.

For a square MIMO system where $M = N = n$ we also have the following interesting result. For a fixed SNR and large value of n [3]

$$C_{\text{MIMO}} \approx n\, c^*(\rho). \tag{2.9}$$

where

$$c^*(\rho) = 2\log\left(0.25(\sqrt{4\rho+1}+1)\right) - \frac{\log e}{4\rho}(\sqrt{4\rho+1}-1)^2$$

Equation (2.9) implies that the capacity increases linearly with n at any SNR.

The pre-log factor in (2.8) and (2.9) is called the multiplexing gain or degree of freedom, and is usually denoted by r .

The dual benefits of MIMO communication at high SNR (represented by the multiplexing gain r and the diversity gain d) is well captured by the diversity-multiplexing tradeoff (DMT) framework formulated in the seminal work of Zheng and Tse [2].

Definition 2.3.1 *Let the rate R scales with $\log \rho$ as [2, 3]*

$$R = r\log \rho.$$

An outage diversity gain $d_{out}(r)$ is achieved at multiplexing gain r if the outage probability satisfies

$$\lim_{\rho \to \infty} \frac{\log P_{out}(r\log \rho)}{\log \rho} = -d_{out}(r)$$

This definition is used to characterize the slow fading performance of the *channel* [2, 3]. We can correspondingly define a DMT framework for a *communication scheme* by replacing P_{out} by P_e.

Definition 2.3.2 *A space-time scheme is a family of codes, indexed by the signal-to-noise ratio ρ, achieves a multiplexing gain r and a diversity gain d is the data rate scales as [2, 3]*

$$R = r\log \rho$$

and the error-probability satisfies

$$\lim_{\rho \to \infty} \frac{\log P_e}{\log \rho} = -d$$

Usually outage analysis in many MIMO systems is more tractable than error analysis, therefore DMT is often obtained via outage analysis and then rigorous proof is usually carried out to bound, if possible, the error via outage and hence characterize the DMT.

The diversity in the fixed rate regime (i.e. R is not function of ρ) cannot be in general deduced from the DMT result by setting $r = 0$. The reason is that there is no solid relationship between $d(r)$ and $d(0)$. The standard DMT arguments in [2] are based on developments that depend critically on the positivity of r in two specific instances. One instance is the proof of [2, Lemma 5], which depends critically on r being strictly positive.[2] More importantly, the asymptotic outage calculations in [2, p. 1079] implicitly use $r > 0$ and result in the outage region:

$$\mathcal{A} = \{\alpha : \sum_i (1 - \alpha_i)^+ < r\}$$

where α_i are the exponential order of the channel eigenvalues, i.e., $\lambda_i = \rho^{-\alpha_i}$. If we set $r = 0$ this expression implies that the outage region is always empty, which is clearly not true.

Thus, the DMT as calculated by the standard methods of [2] cannot be assumed to extend to $r = 0$. It is true that sometimes the function $d(r)$ is indeed continuous at zero, including all the examples in [2]. But it should not be assumed that this continuity always holds. In fact, there are systems where $d(0)$, the diversity at multiplexing gain zero, is not even uniquely defined. Rather, it takes multiple values as a function of rate R. This fact has been observed and analyzed, e.g., in [5–7]. The work in the present dissertation also produces several examples of this phenomenon.

---

[2]This is not the most critical point, because a slightly reformulated version of this lemma can be developed that holds at fixed rates [5].

CHAPTER 3
LINEAR RECEIVERS IN MIMO FLAT FADING

3.1 Introduction

Linear receivers are widely used for their low complexity compared to maximum likelihood (ML) receivers. In the context of MIMO systems, linear receivers such as the minimum mean square error (MMSE) receiver are adopted in some of the emerging standards, e.g. IEEE 802.11n and 802.16e. Therefore the analysis of MMSE receivers is strongly motivated by both theoretical and practical considerations.

A significant amount of research has focused on linear receivers, however, their performance is not fully understood in the MIMO channel. For instance, the distribution of the output signal-to-interference-plus-noise ratio (SINR) of the linear MIMO receiver is still unknown except in asymptotic regimes (large number of antennas, and high/low SNR) [8–11]. The outage and diversity of MMSE receiver have also been a subject of interest. It has been observed [6, 7, 12] that while the MMSE receiver can extract the full spatial diversity of the MIMO quasi-static channel at low rates, it does not enjoy this feature at high rates.

Figure 3.1 shows the outage probabilities (for various spectral efficiencies R bps/Hz) of MMSE and ML receivers respectively. Clearly, one of the main differences between the two characteristics is the slope of the error curves, i.e., the *diversity*. Figure 3.1 shows that in a 2×2 MIMO system the ML receiver achieves diversity 4 at all rates. However, the MMSE receiver diversity varies with the operating spectral efficiency. From a system design perspective, obtaining the MMSE diversity is important in order to understand the broad tradeoffs involved in the determination of the operating point of the system and predicting its performance.
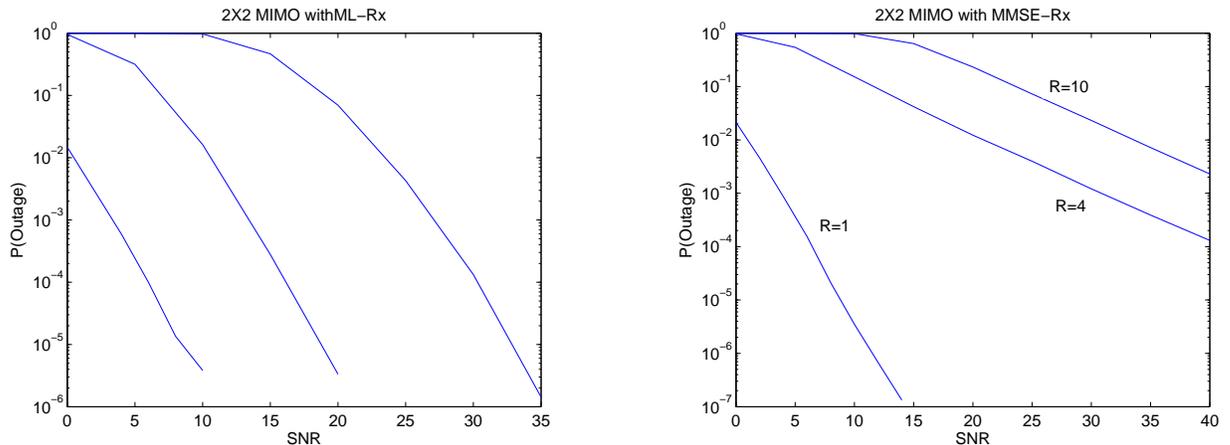
Figure 3.1. Outage probability of ML receiver (left) and MMSE (right) with $M = N = 2$ antennas and for rates R= 1, 4, and 10 bps/Hz

In this work we seek answers for the following questions: when can the MMSE receiver exploit the full diversity in MIMO channel? More generally, how does the diversity of the MMSE receiver vary with the system parameters such as spectral efficiency R, the number of antennas, and in case of inter-symbol interference channel (ISI), the channel memory?

The well-known and powerful framework of diversity-multiplexing tradeoff (DMT) is not sufficient to answer the above questions, because the DMT framework cannot distinguish between different spectral efficiencies that correspond to the same multiplexing gain. In the MIMO MMSE receiver, rates that correspond to the same multiplexing gain can produce different diversities.

We approach the problem of MMSE reception in MIMO flat fading channels through a rate-dependent approximation of the outage probability and then proceed with bounding the pairwise error probability (PEP) from both sides using the outage. This leads to a closed-form expression for the diversity-rate tradeoff which reveals the relationship between diversity, spectral efficiency, and number of transmit and receive antennas. The approximation of outage and PEP as functions of rate requires more delicate handling compared with the DMT analysis, as certain ratios and terms that simply vanish in the DMT analysis are in our case relevant and must be carefully handled.

We then analyze the *frequency-selective*, quasi-static MIMO channel. Specifically we consider single carrier (SC) MMSE equalization under zero-padding (ZP) and cyclic-prefix (CP) transmission. SC-MMSE provides an attractive alternative to orthogonal frequency division multiplexing (OFDM) due to its low complexity and natural avoidance of the peak-to-average power ratio problem. The use of cyclic prefix and zero padding has been investigated in the literature, but the explicit tradeoff between the spectral efficiency and diversity of MIMO SC-MMSE under these two schemes has been unknown and is the subject of our work. We show that the diversity is a function of number of antennas, channel memory and spectral efficiency, and obtain the explicit tradeoff in the special case of SIMO under CP transmission.

The results of this chapter fully characterize the MIMO MMSE diversity in the fixed rate flat quasi-static regime. We analyze both the cases $N \geq M$ and $N < M$, showing that in either case it is possible for the system to be limited to a diversity strictly less than MN. More specifically, the central result of the chapter is as follows: with M transmit and N receive antennas (for any N and M) the diversity is $d = \lceil (M2^{-\frac{R}{M}} - (M-N)^+)^+ \rceil^2 + |N - M| \lceil (M2^{-\frac{R}{M}} - (M-N)^+)^+ \rceil$, where $(\cdot)^+ = \max(0, \cdot)$ and $\lceil \cdot \rceil$ denotes rounding up to the next higher integer. Our results confirm and refine the earlier approximate results on the diversity of MMSE MIMO receivers that were obtained for very high and very low rates [6,7,12]. The MIMO MAC channel is also studied.

Some of the related literature is as follows. The performance of MMSE receiver in terms of reliability goes back to [13] where outage analysis was performed for MMSE SIMO diversity combiner in a Rayleigh fading channel with multiple interferers. In the context of point-to-point MIMO systems, Gore et al. [14] compared the performance of MMSE D-BLAST with the ordered successive cancellation V-BLAST. They show that the former has better throughput at low- and moderate SNR. Onggosanusi et al. [12] studied MMSE and zero-forcing (ZF) MIMO receivers and noticed their distinct outage performance at high-SNR, specifically for large number of transmit antennas and low spectral efficiencies R, but provided no analysis.

Hedayat and Nosratinia [6] considered the outage probability as a function of fixed rates R under joint and separate spatial encoding, but for MMSE they obtained results only in the extremes of very high and very low rates. Kumar et al. [7] provided a DMT analysis for the system of [6] and observed that the DMT analysis does not predict the diversity of MMSE receivers at lower rates. We note that all existing analyses are limited to the case where the number of receive antennas (N) is greater than or equal the number of the transmit antennas (M).

This chapter is organized as follows. Section 3.2 describes the system model. Section 3.3 finds the exponential order of outage. Section 3.4 bounds the codeword error probabilities using the outage values, and derives the final result. Section 3.5 extends the result to the MAC channel. Section 3.6 provides simulations that illuminate our results.

3.2   Linear Receivers

The input-output system model for flat fading MIMO channel with M transmit and N receive antennas is given by

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} \tag{3.1}$$

where $\mathbf{H} \in \mathbb{C}^{N \times M}$ is the channel matrix whose entries are independent and identically distributed complex Gaussian, $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is the transmitted vector, $\mathbf{n} \in \mathbb{C}^{N \times 1}$ is the Gaussian noise vector. The vectors $\mathbf{x}$ and $\mathbf{n}$ are assumed independent. We assume a quasi-static flat fading channel and perfect channel state information (CSI) at the receiver (CSIR) and no CSI at the transmitter (CSIT), therefore transmit antennas operate with equal power.

We aim to characterize the diversity gain, $d(R, M, N)$, as a function of the spectral efficiency R (bits/sec/Hz) and the number of transmit and receive antennas. This requires a pairwise error probability (PEP) analysis which is not directly tractable. Instead, we find the exponential order of outage probability and then demonstrate that outage and PEP exhibit identical exponential orders.
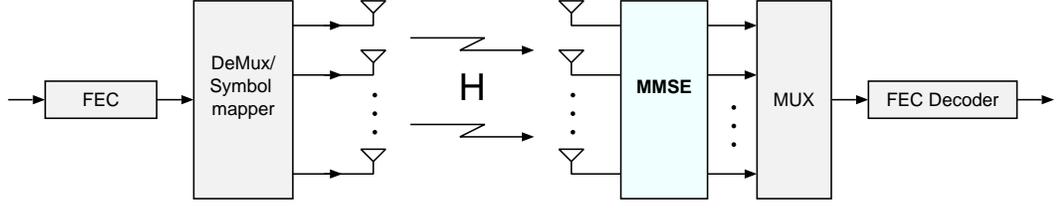
Figure 3.2. MIMO system with linear MMSE receiver

Following the notation of [15], we define the outage-type quantities

$$P_{out}(R, N, M) \triangleq \mathbb{P}(I(\mathbf{x}; \mathbf{y}) < R) \tag{3.2}$$

$$d_{out}(R, N, M) \triangleq -\lim_{\rho \to \infty} \frac{\log P_{out}(R, M, N)}{\log \rho} \tag{3.3}$$

where ρ is the per-stream signal-to-noise ratio (SNR).

We say that the two functions $f(\rho)$ and $g(\rho)$ are *exponentially equal*, denoted by $f(p) \doteq g(p)$ when

$$\lim_{\rho \to \infty} \frac{\log f(\rho)}{\log(\rho)} = \lim_{\rho \to \infty} \frac{\log g(\rho)}{\log(\rho)}$$

The ordering operators $\dot{\leq}$ and $\dot{\geq}$ are also defined accordingly. If $f(\rho) \doteq \rho^d$, we say that d is the *exponential order* of $f(p)$.

3.2.1   MMSE Equalizer

The equalizer, denoted by $\mathbf{W}$, decouples the M transmitted data streams at the receiver (Figure 3.2). The MMSE equalizer is obtained by minimizing the mean square error (MSE) defined as $\mathbb{E}[||\mathbf{x} - \mathbf{W}^H \mathbf{y}||^2]$. It is usually assumed [6,7] that the number of transmit antennas M is no more than that of receive antennas N. In the following, we start with $N \geq M$ but later generalize it to $N < M$ as well.

For $N \geq M$, using the orthogonality principle, the MMSE equalizer is given by [12, 16]

$$\mathbf{W} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \rho^{-1}I)^{-1}$$
$$= (\mathbf{H}^H\mathbf{H} + \rho^{-1}I)^{-1}\mathbf{H}^H \tag{3.4}$$

The corresponding signal-to-interference and noise ratio (SINR) of the output stream k of the MMSE detector is

$$\gamma_k = \frac{1}{(\mathbf{I} + \rho\mathbf{H}^H\mathbf{H})_{kk}^{-1}} - 1, \quad 1 \leqslant k \leqslant M \tag{3.5}$$

where $(\cdot)^H$ denotes matrix Hermitian, $(\cdot)_{kk}^{-1}$ denotes the diagonal element k of the matrix inverse.

For the case $N < M$, it can be shown using a technique[1] very similar to [13, Appendix A] that the SINR expression (3.5) is again valid.

The square matrix $\mathcal{W} = \mathbf{H}^H\mathbf{H}$ is random, non-negative definite, and obeys the *Wishart Distribution* [17,18]. In this work, the joint distribution of the eigenvalues of this equivalent channel matrix opens the door to the development of our analysis, as is also the case in many other MIMO results.

The equalizer output is

$$\mathbf{y} = \mathbf{W}\mathbf{H}\mathbf{x} + \mathbf{W}\mathbf{n}. \tag{3.6}$$

The signal streams of the transmit antennas may be either separately or jointly encoded. Separate encoding is simpler and has been fully analyzed [6], but we mention the central result for completeness.

Theorem 3.2.1 ([6, 7]) *In a MIMO system consisting of M transmit and N receive antennas ($N \geqslant M$), under separate spatial encoding, the MMSE receiver achieves the diversity*

$$d_{out}(R, N, M) = N - M + 1 \tag{3.7}$$

under either uniform or non-uniform rate assignment.

Furthermore, it has been established [6, 7] that the zero forcing equalizer achieves diversity $N - M + 1$ under both joint or separate spatial encoding.

According to Theorem 3.2.1, a MMSE receiver operating under separate spatial encoding (e.g. horizontal encoding V-BLAST) will have no more diversity gain than ZF receiver.

---

[1] In [13] an MMSE diversity combiner is used at the receiver in the presence of one transmit antenna and M interferers.

3.3   Outage Analysis

We now consider the MMSE diversity where the data stream is first encoded then multiplexed into M sub-streams, each transmitted by one antenna. This approach is known to improve the performance compared with separate coding of the streams [3]. Outage occurs if the channel fails to support the target rate [17]. After channel equalization, the M sub-streams x_k are decoupled and thus the mutual information between the transmitted vector $\mathbf{x}$ and the received vector $\mathbf{y}$ given CSIR is [12]

$$I(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{M} I(x_k, y_k) \tag{3.8}$$

Thus from (3.2) and (3.8), P_{out} is given by

$$P_{out} = \mathbb{P}\left(\sum_{k=1}^{M} \log(1 + \gamma_k) < R \right) \tag{3.9}$$

Substituting MMSE SINR (γ_k) from (3.5) in (3.9) we get

$$P_{out} = \mathbb{P}\left(\sum_{k=1}^{M} \log(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1} > -R \right) \tag{3.10}$$

The dependence on the diagonal elements of the random matrix $(\mathbf{I}+\rho\mathcal{W})_{kk}^{-1}$ makes further analysis intractable. We instead proceed to provide lower and upper bounds on the outage probability. In Section 3.4 we will show that outage probability (P_{out}) and pairwise error probability (PEP) exhibit identical exponential error.

3.3.1   Outage Upper Bound

Lemma 3.3.1 *For an MMSE MIMO system consisting of M transmit and N receive antennas, under quasi-static Rayleigh fading, we have*

$$P_{out}(R, M, N) \dot{\leqslant} \rho^{-d_{out}(R,M,N)}$$

where

$$d_{out}(R, M, N) = \left[\left(M2^{-\frac{R}{M}} - (M-N)^+ \right)^+ \right]^2 +$$

$$|N - M| \left[\left(M2^{-\frac{R}{M}} - (M-N)^+ \right)^+ \right]. \tag{3.11}$$

where $()^+$ *denotes the* $\max(0, \cdot)$.

Proof We begin by bounding the sum in (3.10) via Jensen's inequality

$$\sum_{k=1}^{M} \log \left(\mathbf{I} + \rho \mathcal{W} \right)_{kk}^{-1} \leq M \log \left(\sum_{k=1}^{M} \frac{1}{M} (\mathbf{I} + \rho \mathcal{W})_{kk}^{-1} \right)$$

$$= M \log \left(\frac{1}{M} \text{tr} \left((\mathbf{I} + \rho \mathcal{W})^{-1} \right) \right)$$

$$= M \log \left(\frac{1}{M} \sum_{k=1}^{M} \frac{1}{1 + \rho \lambda_k} \right) \tag{3.12}$$

where (3.12) is true because trace is equal to the sum of eigenvalues.

Notice that for $N < M$ only N eigenvalues are non-zero. hence (3.12) can be written as

$$M \log \left(\frac{1}{M} \sum_{k=1}^{L} \frac{1}{1 + \rho \lambda_k} + (M - N)^+ \right) \tag{3.13}$$

where $L = \min(M, N)$.

Substituting (3.13) in (3.10), we have

$$P_{out} \leq \mathbb{P} \left(\sum_{k=1}^{L} \frac{1}{1 + \rho \lambda_k} \geqslant M2^{-\frac{R}{M}} - (M - N)^+ \right) \tag{3.14}$$

Define:

$$\alpha_k \triangleq -\frac{\log \lambda_k}{\log \rho}, \quad \text{for} \quad k = 1, ..., n \quad , \tag{3.15}$$

based on which we can write the exponential equality

$$\frac{1}{1 + \rho \lambda_k} \doteq \begin{cases} \rho^{\alpha_k - 1} & \alpha_k < 1 \\ 1 & \alpha_k > 1 \end{cases} \tag{3.16}$$

Define $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_n]$ and a new random variable

$$M(\boldsymbol{\alpha}) \triangleq \sum_{\alpha_k > 1} 1 \tag{3.17}$$

This definition is based on the observation that the term $\frac{1}{1+\rho\lambda_k}$ defined in (3.16) is either zero or one at high SNR, therefore to characterize $\sum_k \frac{1}{1+\rho\lambda_k}$ at high SNR we count the ones. Thus

$$\sum_{k=1}^{n} \frac{1}{1 + \rho\lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \tag{3.18}$$

$$\doteq M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1} \tag{3.19}$$

$M(\boldsymbol{\alpha})$ inherits its randomness from $\lambda_1, \ldots, \lambda_n$. The bound in (3.14) is evaluated by computing the probability of $\{\boldsymbol{\alpha} \in \mathcal{A}\}$, where $\mathcal{A} = \{\alpha : M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1} > M2^{-\frac{R}{M}} - (M - N)^+\}$ denotes the outage event based on the approximation in (3.14). In order to evaluate the probability of this event we need the joint distribution of the eigenvalues, or equivalently the distribution of $\boldsymbol{\alpha}$. The distribution follows Wishart distribution and was initially discovered by [18] . The distribution of $\boldsymbol{\alpha}$ can be easily evaluated as follows [2].

Let $\mathbf{R}$ be an $m \times n$ $(m \geqslant n)$ random matrix whose entries are $\mathcal{CN}(0, 1)$. The joint PDF of the ordered random variables $\boldsymbol{\alpha}$ (defined in (3.15) for the eigenvalues of $\mathbf{R}^H\mathbf{R}$) is given by

$$\mathbb{P}(\boldsymbol{\alpha}) = K_{m,n}^{-1}(\log \rho)^n \prod_{i=1}^{n} \rho^{-(m-n+1)\alpha_i} \prod_{i<j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \exp\left[-\sum_{i=1}^{n} \rho^{-\alpha_i}\right] \tag{3.20}$$

where $K_{m,n}^{-1}$ is a normalizing factor.

Using the distribution of $\boldsymbol{\alpha}$ for the defined matrix $\mathbf{R}$, the asymptotic outage bound is

$$P_{out} \dot{\leqslant} \int_{\mathcal{A}} \mathbb{P}(\boldsymbol{\alpha}) d\boldsymbol{\alpha}$$

$$= K_{m,n}^{-1}(\log \rho)^n \int_{\mathcal{A}} \prod_{i=1}^{n} \rho^{-(m-n+1)\alpha_i} \prod_{i<j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \exp\left[-\sum_{i=1}^{n} \rho^{-\alpha_i}\right] d\boldsymbol{\alpha} \tag{3.21}$$

The simplification of the integral follows from [2]. The term outside the integral has no effect on the exponent. The term $|\rho^{-\alpha_i} - \rho^{-\alpha_j}|$ is dominated by $\rho^{-\alpha_i}$ at high SNR.

We now divide the integration range into $\mathcal{A}' = \mathcal{A} \cap \mathbb{R}_+^n$ and its complement. If $\boldsymbol{\alpha} \notin \mathcal{A}'$, the exponential term will dominate the other terms and will drive the integral to zero. If $\boldsymbol{\alpha} \in \mathcal{A}'$, the exponential term is approximately 1 at high SNR and will disappear. Therefore

$$
\begin{aligned}
P_{out} &\dot{\leqslant} \int_{\mathcal{A}'} \prod_{i=1}^n \rho^{-(m-n+1)\alpha_i} \prod_{i<j} |\rho^{-\alpha_i} - \rho^{-\alpha_j}|^2 \, d\boldsymbol{\alpha} \\
&\dot{=} \int_{\mathcal{A}'} \prod_{i=1}^n \rho^{-(2i-1+m-n)\alpha_i} \, d\boldsymbol{\alpha}
\end{aligned}
\tag{3.22}
$$

where

$$
\begin{aligned}
\mathcal{A}' &= \{M(\boldsymbol{\alpha}) > M2^{-\frac{R}{M}} - (M-N)^+\} \\
&= \{\alpha_1 > 1, ..., \alpha_S > 1, \alpha_{S+1} > 0, ...\alpha_L > 0\}
\end{aligned}
\tag{3.23}
$$

and $S = \left\lceil \left(M2^{-\frac{R}{M}} - (M-N)^+\right)^+ \right\rceil$. The integration region $\mathcal{A}'$ has boundaries that are parallel to nonnegative orthant $\mathbb{R}_+^n$, therefore the integration over multiple variables in (3.22) can be separated:

$$
\begin{aligned}
P_{out} \quad &\dot{\leqslant} \quad \prod_{i=1}^n \int_{\mathcal{A}'} \rho^{-(2i-1+m-n)\alpha_i} \, d\boldsymbol{\alpha} \tag{3.24} \\
&= \quad \rho^{-\sum_{i=1}^S (2i-1+m-n)} \\
&= \quad \rho^{-(S^2+(m-n)S)}, \quad \text{for } m \geqslant n \tag{3.25} \\
&= \quad \rho^{-(S^2+|m-n|S)}, \quad \text{for general } m, n \tag{3.26} \\
&= \quad \rho^{-d_{out}}
\end{aligned}
$$

which establishes the proof of Lemma 3.3.1.

3.3.2 Outage Lower Bound

Lemma 3.3.2 *For an MMSE MIMO system consisting of M transmit and N receive antennas (and $L = \min\{M, N\}$), operating under quasi-static Rayleigh fading, we have*

$$
P_{out}(R, M, N) \dot{\geqslant} \rho^{-d_{out}(R,M,N)}
$$

where

$$d_{out}(R, M, N) = \left\lceil \left(M2^{-\frac{R}{M}} - (M-N)^+ \right)^+ \right\rceil^2 + |N - M| \left\lceil \left(M2^{-\frac{R}{M}} - (M-N)^+ \right)^+ \right\rceil.$$

Proof The lower bound is also based on Jensen's inequality. Recall

$$P_{out} = \mathbb{P}\left(\sum_{k=1}^{M} \log(1 + \gamma_k) < R \right)$$

$$= \mathbb{P}\left(\sum_{k=1}^{M} \log \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} < R \right)$$

$$\geqslant \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^{M} \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} < R \right) \tag{3.27}$$

Let the eigen decomposition of $\mathbf{H}^H\mathbf{H}$ be given by $\mathbf{H}^H\mathbf{H} = \mathbf{U}^H\Lambda\mathbf{U}$ where $\mathbf{U}$ is unitary and Λ is a diagonal matrix that has the eigenvalues of the Wishart matrix $\mathcal{W}$ on its diagonal. Let the vector $\mathbf{u}_k$ be the column k of the matrix $\mathbf{U}$ and $u_{\ell k}$ be the element ℓ of this column, we have

$$(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1} = \mathbf{u}_k^H(\mathbf{I} + \rho\Lambda)^{-1}\mathbf{u}_k$$

$$= \sum_{\ell=1}^{M} \frac{|u_{\ell k}|^2}{1 + \rho\lambda_\ell}$$

$$\triangleq S_k. \tag{3.28}$$

Let $\bar{k} = \arg\min_k S_k$. Using (3.28), we can bound the sum in (3.27)

$$\frac{1}{M} \sum_{k=1}^{M} \frac{1}{(\mathbf{I} + \rho\mathcal{W})_{kk}^{-1}} = \frac{1}{M} \sum_{k=1}^{M} \frac{1}{S_k}$$

$$\leqslant \frac{1}{\min_k S_k} \tag{3.29}$$

$$= \frac{1}{S_{\bar{k}}} \tag{3.30}$$

thus the outage bound in (3.27) can be further bounded using (3.29)

$$P_{out} \geqslant \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^{M} \frac{1}{(\mathbf{I} + \rho \mathcal{W})_{kk}^{-1}} < R \right)$$

$$\geqslant \mathbb{P}\left(M \log \frac{1}{S_{\bar{k}}} < R \right)$$

$$= \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \right) \tag{3.31}$$

We now bound (3.31) by conditioning on the event $\mathcal{B} \triangleq \{ |u_{\ell \bar{k}}|^2 \geqslant \frac{a}{M} \}$ where a is a positive real number that is slightly smaller than one, i.e. $a = 1 - \epsilon$, and ϵ is a small positive number. We then have

$$\mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \right) \geqslant \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \Big| \mathcal{B} \right) \mathbb{P}(\mathcal{B})$$

$$= \mathbb{P}\left(\sum_{\ell=1}^{M} \frac{|u_{\ell \bar{k}}|^2}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}} \Big| \mathcal{B} \right) \mathbb{P}(\mathcal{B})$$

$$\geqslant \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^{M} \frac{a}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}} \right) \mathbb{P}(\mathcal{B})$$

$$\doteq \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^{M} \frac{a}{1 + \rho \lambda_\ell} > 2^{-\frac{R}{M}} \right) \tag{3.32}$$

$$= \mathbb{P}\left(\sum_{\ell=1}^{M} \frac{1}{1 + \rho \lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}} \right)$$

$$= \mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^{L} \frac{1}{1 + \rho \lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}} - (M - N)^+ \right) \tag{3.33}$$

where (3.32) follows because $\mathbb{P}(\mathcal{B})$ is finite and independent of ρ; this can be proved similarly to [7, Appendix A]. To make the upcoming expressions compact, we introduce a new variabe $\kappa \triangleq \frac{M}{a} 2^{-\frac{R}{M}} - (M - N)^+$

$$\mathbb{P}\left(\frac{1}{M} \sum_{\ell=1}^{L} \frac{1}{1 + \rho \lambda_\ell} > \kappa \right) \tag{3.34}$$

Whenever $M 2^{-\frac{R}{M}}$ is non-integer, the constant a can be chosen such that $\left\lceil \left(M 2^{-\frac{R}{M}} - (M - N)^+ \right)^+ \right\rceil = \left\lceil \left(\frac{M}{a} 2^{-\frac{R}{M}} - (M - N)^+ \right)^+ \right\rceil$. We note this is satisfied for all rates, with the

exception of an isolated set of points. As long as $M2^{\frac{R}{M}} \notin \mathbb{N}$ we have:

$$
\begin{aligned}
P_{out} &\geqslant \mathbb{P}\left(\sum_{\ell=1}^{L} \frac{1}{1+\rho\lambda_\ell} > \kappa \right) \\
&\doteq \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1+\rho\lambda_\ell} > \lceil\kappa\rceil \right)
\end{aligned}
\tag{3.35}
$$

The remaining steps follow similarly to the proof of Lemma 3.3.1. Thus $P_{out} \dot{\geqslant} \rho^{-d_{out}}$ with d_{out} is given by Lemma 3.3.2.

On the set of isolated points $M2^{-\frac{R}{M}} \in \mathbb{N}$, the right hand side of Eq. (3.35) obeys a slightly weaker upper bound by replacing κ with $\kappa + 1$. We can combine the cases where $M2^{-\frac{R}{M}}$ is integer and non-integer to write the upper bound compactly as follows:

$$
d_{out}(R, M, N) \leq \left\lfloor \left(M2^{-\frac{R}{M}} + 1 - (M-N)^+\right)^+ \right\rfloor^2 + |N - M| \left\lfloor \left(M2^{-\frac{R}{M}} + 1 - (M-N)^+\right)^+ \right\rfloor.
$$

Inspection shows that this bound is tight against the lower bound everywhere except its discontinuity points. In other words, the upper bound is left-continuous while the lower bound was right-continuous at the discontinuity points.

3.4   PEP Analysis

Recalling that the diversity is roughly defined as the slope of PEP at high SNR, we now proceed to bound the PEP tightly from both sides using the outage results already obtained.

3.4.1   PEP Upper Bound

We start by a lower bound that is inspired by [2, Lemma 5] but requires a more careful treatment since we are analyzing rate, not the DMT (see the Introduction).

Lemma 3.4.1 *For a quasi-static fading MIMO channel with MMSE receiver we have*

$$
d_{out}(R, M, N) \geqslant d(R, M, N).
$$

Proof Denote E for an error event, and let $x \in \mathcal{C}$ be the transmitted codeword from a codebook $\mathcal{C}$ of size 2^{Rl} where R and l are code rate and code length respectively. Define $\boldsymbol{f} = \mathbf{WHx}$ that accounts for the combined effect of channel and equalizer. The transmit messages are assumed equi-probable so the entropy $\mathcal{H} = \log|\mathcal{C}| = Rl$. Applying the Fano inequality [19]

$$\mathbb{P}(E|\boldsymbol{f} = f) \geqslant \frac{Rl - I(\mathbf{x}; \mathbf{y}|\boldsymbol{f} = f)}{Rl} - \frac{\mathcal{H}(\mathbb{P}(E)|\boldsymbol{f} = f)}{Rl} \tag{3.36}$$

By defining $\mathcal{D}_\delta$ for any $\delta > 0$ as $\mathcal{D}_\delta \triangleq \{f : I(\mathbf{x}; \mathbf{y}|\boldsymbol{f} = f) < l(R - \delta)\}$, and noting that $\mathcal{H}(\mathbb{P}(E)|f \in \mathcal{D}_\delta) \leqslant \mathcal{H}(\mathbb{P}(E))$ from (3.36), we get

$$\begin{aligned}
\mathbb{P}(E|\boldsymbol{f} \in \mathcal{D}_\delta) &\geqslant \frac{Rl - I(\mathbf{x}; \mathbf{y}|\boldsymbol{f} \in \mathcal{D}_\delta)}{Rl} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl} \\
&\geqslant \frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl}.
\end{aligned} \tag{3.37}$$

Also by using the definition of P_{out} we have

$$\mathbb{P}(\boldsymbol{f} \in \mathcal{D}_\delta) = \mathbb{P}\big(I(\mathbf{x}; \mathbf{y}) < l(R - \delta)\big) \doteq \rho^{-d_{out}(R-\delta, M, N)} \tag{3.38}$$

For small enough values of $\delta > 0$, we have $d_{out}(R, M, N) = d_{out}(R-\delta, M, N)$ since $d_{out}(R, M, N)$ is left-continuous with respect to R. Hence, by invoking (3.37) and (3.38), the error probability is given by

$$\begin{aligned}
\mathbb{P}_{\mathrm{err}}(R, M, N) &= \mathbb{P}(E|f \in \mathcal{D}_\delta)\mathbb{P}(f \in \mathcal{D}_\delta) + \mathbb{P}(E|f \notin \mathcal{D}_\delta)\mathbb{P}(f \notin \mathcal{D}_\delta) \\
&\geqslant \mathbb{P}(E|f \in \mathcal{D}_\delta)\mathbb{P}(f \in \mathcal{D}_\delta) \\
&\dot{\geqslant} \left(\frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl}\right)\rho^{-d_{out}} \\
&\doteq \rho^{-d_{out}}
\end{aligned} \tag{3.39}$$

where we have used $\left(\frac{\delta}{R} - \frac{\mathcal{H}(\mathbb{P}(E))}{Rl}\right) \doteq 1$, which was derived in [15]. This establishes the proof of the PEP upper bound.

3.4.2 PEP Lower Bound

We begin by writing the error probability in terms of error event E and outage event O

$$\mathbb{P}_{\mathrm{err}}(R, M, N) = \mathbb{P}(E|O) \cdot P_{out} + \mathbb{P}(E, \bar{O})$$

In Section 3.3.1 we have shown that, based on the event $\left\{ \sum_{k=1}^{L} \frac{1}{1+\rho\lambda_k} \geqslant M2^{-\frac{R}{M}} - (M - N)^+ \right\}$, the outage probability is upper bounded by $P_{out} \dot{\leqslant} \rho^{-d_{out}}$. Hence, the error probability can be bounded as

$$\mathbb{P}_{\mathrm{err}}(R, M, N) \dot{\leqslant} \mathbb{P}(E|O)\, \rho^{-d_{out}} + \mathbb{P}(E, \bar{O})$$

$$\leqslant \rho^{-d_{out}} + \mathbb{P}(E, \bar{O}) \tag{3.40}$$

We intend to show that $\rho^{-d_{out}} \dot{\geqslant} \mathbb{P}(E, \bar{O})$, and thus $\mathbb{P}_{\mathrm{err}}(R, M, N) \dot{\leqslant} \rho^{-d_{out}}$ which produces the following lemma.

Lemma 3.4.2 *For a quasi-static fading MIMO channel with MMSE receiver we have*

$$d_{out}(R, M, N) \leqslant d(R, M, N).$$

Proof We begin by giving a sketch of the proof then we proceed with the details. The first part of the proof consists of developing a bound on PEP conditioned on H, namely $P[s_k \rightarrow s_j | \mathbf{H} = H]$. To do this we obtain an upper bound of the variance of the SINR which is expressed in terms of the eigenvalues of the Wishart matrix $\mathcal{W}$, resulting in $\mathbb{P}[E|\mathbf{H} = H] \leqslant 4\exp\left(-\left(\sum_{k=1}^{L} \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}\right)$. The PEP is used to derive a conditional union bound on error. We then divide the channel events into two sets based on the exponential order of the eigenvalues: the set where $M(\boldsymbol{\alpha}) = 0$ and otherwise. We apply Bayes theorem on the union bound using these two sets. The calculation of the terms of the Bayesian gives $\mathbb{P}(E, \bar{O}) \dot{\leqslant} \rho^{-MN} \leqslant \rho^{-d_{out}}$ as desired.

We now proceed in detail. We want to compute the probability that the transmitted symbol $x(k) = s_l$ is erroneously detected as $x(k) = s_j$.

Recalling the equalizer output given by (3.6), define the noise-plus-interference signal

$$\tilde{\mathbf{n}} = \mathbf{y} - \sqrt{\rho}\mathbf{x} = \sqrt{\rho}(\mathbf{W}\mathbf{H} - \mathbf{I})\mathbf{x} + \mathbf{W}\mathbf{n} \tag{3.41}$$

Using the eigen-decomposition of H and noting that $E(\mathbf{n}) = 0$ and $E(\mathbf{n}\mathbf{n}^H) = I$, we have

$$\mu_{\tilde{\mathbf{n}}} \triangleq E(\tilde{\mathbf{n}}) = \sqrt{\rho}(\mathbf{W}\mathbf{H} - \mathbf{I}) = -\rho^{\frac{1}{2}}(\mathcal{W} + \rho^{-1}\mathbf{I})^{-1}\mathbf{x} \tag{3.42}$$

$$\boldsymbol{R}_{\tilde{\mathbf{n}}} \triangleq E(\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H) = (\mathcal{W} + \rho^{-1}\mathbf{I})^{-1} \tag{3.43}$$

Thus the variance of the noise sample $\tilde{n}(k)$ is given by

$$\sigma_{\tilde{\mathbf{n}}}^2(k) = \boldsymbol{R}_{\tilde{\mathbf{n}}}(k,k) - |\mu_{\tilde{\mathbf{n}}}(k)|^2$$

$$= (\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-1} - \rho^{-1}(\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-2} \tag{3.44}$$

where $|\mu_{\tilde{\mathbf{n}}}(k)|^2$ is the kth diagonal of the matrix $E(\tilde{\mathbf{n}})E(\tilde{\mathbf{n}}^H)$ and k counts from 1 to M.

By defining $e_{jl} \triangleq \frac{s_j - s_l}{|s_j - s_l|}$, the probability of erroneous detection for channel realization is given by

$$\mathbb{P}[s_l \to s_j | \mathbf{H} = H]$$

$$= \mathbb{P}\left[\frac{\rho}{4}|s_j - s_l|^2 \leqslant |e_{jl}^*(y(k) - \sqrt{\rho}s_l)|^2 \,\middle|\, \mathbf{H} = H\right] \tag{3.45}$$

$$\leqslant \mathbb{P}\left[\frac{\rho}{4}|s_j - s_l|^2 \leqslant |\tilde{n}_k|^2 \,\middle|\, \mathbf{H} = H\right]$$

where the inequality holds since $|e_{jl}^*(y(k) - \sqrt{\rho}s_l)| \leqslant |e_{jl}^*||(y(k) - \sqrt{\rho}s_l)| = |(y(k) - \sqrt{\rho}s_j)| = |\tilde{n}(k)|$.

Denoting the real and imaginary parts of $\tilde{n}(k)$ by $\tilde{n}_r(k) \sim \mathcal{N}(\mu_r(k), \sigma_r^2(k))$ and $\tilde{n}_i(k) \sim \mathcal{N}(\mu_i(k), \sigma_i^2(k))$ respectively, we then have

$$\left\{\frac{\rho}{4}|s_j - s_l|^2 \leqslant |\tilde{n}(k)|^2\right\}$$

$$\subset \left\{\frac{\rho}{16}|s_j - s_l|^2 \leqslant |\tilde{n}_r(k)|^2\right\} \cup \left\{\frac{\rho}{16}|s_j - s_l|^2 \leqslant |\tilde{n}_i(k)|^2\right\} \tag{3.46}$$

Applying the property of the Gaussian tail function $Q(x) \leqslant e^{(-x^2/2)}$ for the pairwise error probability, we obtain

$$\mathbb{P}[s_k \to s_j | \mathbf{H} = H] \leqslant e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| - \mu_r(k))^2}{\sigma_r^2(k)}\right)} + e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| + \mu_r(k))^2}{\sigma_r^2(k)}\right)} + e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| - \mu_i(k))^2}{\sigma_i^2(k)}\right)}$$

$$+ e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| + \mu_i(k))^2}{\sigma_i^2(k)}\right)}$$

$$\leqslant e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| - \mu_r(k))^2}{\sigma_{\tilde{\mathbf{n}}}^2(k)}\right)} + e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| + \mu_r(k))^2}{\sigma_{\tilde{\mathbf{n}}}^2(k)}\right)} + e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| - \mu_i(k))^2}{\sigma_{\tilde{\mathbf{n}}}^2(k)}\right)}$$

$$+ e^{\left(-\frac{(\frac{\sqrt{\rho}}{4}|s_j - s_l| + \mu_i(k))^2}{\sigma_{\tilde{\mathbf{n}}}^2(k)}\right)} \tag{3.47}$$

where the last step holds as $\sigma_{\mathbf{n}}^2(k) = \sigma_r^2(k) + \sigma_i^2(k) \geqslant \sigma_r^2(k), \sigma_i^2(k)$.

Now we proceed by showing that $\mu_i(k) \dot{\leqslant} \rho^{\frac{1}{2}}$. Consider the eigen decomposition of

$$
\begin{aligned}
[\mathcal{W} + \rho^{-1}\mathbf{I}]^{-1} &= \mathbf{U}^H[\Lambda + \rho^{-1}\mathbf{I}]^{-1}\mathbf{U} \\
&= \mathbf{U}^H\left[\mathrm{d}iag\{\frac{1}{\lambda_k + \rho^{-1}}\}\right]\mathbf{U}
\end{aligned}
\tag{3.48}
$$

where $\mathbf{U}$ is unitary matrix, and Λ is the eigen decomposition of $\mathcal{W}$. Note that $\lambda_k + \rho^{-1} \dot{\geqslant} \rho^{-1}$ or $\frac{1}{\lambda_k + \rho^{-1}} \dot{\leqslant} \rho$. Therefore, all elements of the matrix $\pm\mathbf{U}^H[\Lambda + \rho^{-1}\mathbf{I}]^{-1}\mathbf{U}$, being linear combination of $\{\frac{1}{\lambda_k + \rho^{-1}}\}$, cannot grow faster than $O(\rho)$, and thus the elements of $\pm\rho^{\frac{1}{2}}[\mathcal{W} + \rho^{-1}\mathbf{I}]^{-1}$ cannot grow faster than $O(\rho^{\frac{1}{2}})$, i.e. $\pm\dot{\mu}_{\tilde{\mathbf{n}}}(k) \dot{\leqslant} \rho^{\frac{1}{2}}$ and therefore $\rho^{\frac{1}{2}} \pm \mu_{\tilde{\mathbf{n}}}(k) \doteq \rho^{\frac{1}{2}}$. The same result holds for $\mu_r(k)$ and $\mu_i(k)$.

As a result, for any s_j and s_l, $\frac{\sqrt{\rho}}{4}|s_j - s_l| \pm \mu_r(k) \doteq \rho^{\frac{1}{2}} \pm \mu_r(k) \doteq \rho^{\frac{1}{2}}$ and similarly $\frac{\sqrt{\rho}}{4}|s_j - s_l| \pm \mu_i(k) \doteq \rho^{\frac{1}{2}}$. Thus from (3.47), we have

$$
\mathbb{P}[s_k \to s_j|\mathbf{H} = H] \quad \dot{\leqslant} \quad 4e^{-\frac{\rho}{\sigma_{\tilde{\mathbf{n}}}^2(k)}}
\tag{3.49}
$$

Now we bound the variance in (3.44) and apply it in (3.49)

$$
\begin{aligned}
\sigma_{\tilde{\mathbf{n}}}^2(k) &\leqslant \sum_{k=1}^{L}\left[(\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-1} - \rho^{-1}(\mathcal{W} + \rho^{-1}\mathbf{I})_{kk}^{-2}\right] \\
&= \sum_{k=1}^{L}\left[\frac{\rho}{1 + \rho\lambda_k} - \frac{\rho}{(1 + \rho\lambda_k)^2}\right] = \sum_{k=1}^{L}\frac{\rho^2\lambda_k}{(1 + \rho\lambda_k)^2}
\end{aligned}
\tag{3.50}
$$

Denoting the error event E and using (3.50), the probability of erroneous detection in (3.49) is bounded as

$$
\mathbb{P}[E|\mathbf{H} = H] \leqslant 4e^{-\left(\sum_{k=1}^{L}\frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}}
\tag{3.51}
$$

Applying the union bound, we get

$$
\mathbb{P}(E|\mathbf{H} = H) \dot{\leqslant} 2^{Rl}e^{-\left(\sum_{k=1}^{L}\frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)^{-1}}
\tag{3.52}
$$

Based on (3.52), we can evaluate $P(E, \bar{O})$ in (3.40) as follows. Recalling the exponential inequality

$$\sum_{k=1}^{n} \frac{1}{1 + \rho \lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \tag{3.53}$$

$$\doteq M(\boldsymbol{\alpha}) + \max_{\{\alpha_k : \alpha_k < 1\}} \rho^{\alpha_k - 1}$$

$$\doteq M(\boldsymbol{\alpha}) \tag{3.54}$$

Consider the two regions: $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) = 0\}$ and $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \geqslant 1\}$. At high SNR the event $\bar{O}$ is equivalent to $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \dot{\leqslant} \lceil M 2^{-\frac{R}{M}} - (M - N)^+ \rceil\}$.

In the first region $\{M(\boldsymbol{\alpha}) = 0\}$, at any rate $R \geqslant 0$ we have $\{\boldsymbol{\alpha} : \lceil M 2^{-\frac{R}{M}} - (M - N)^+ \rceil \dot{\geqslant} M(\boldsymbol{\alpha}) = 0$ so there is no outage.

In the second region $\{M(\boldsymbol{\alpha}) \geqslant 1\}$ the exponent order of the outage probability depends on the rate. We investigate these two regions separately.

In the region $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) = 0\}$, we have $\max_k \alpha_k < 1$ since all $\alpha'_k s < 1$. From (3.52) and (3.54) we conclude that

$$\mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) = 0) \dot{\leqslant} 2^{Rl} e^{-\rho \left(\max_k \alpha_k - 1 \right)^{-1}}$$

$$= 2^{Rl} e^{-\rho \left(1 - \max_k \alpha_k \right)} \tag{3.55}$$

Since exponential function dominates all polynomials and $1 - \max_k \alpha_k > 0$, we get

$$\lim_{\rho \to \infty} \frac{e^{-\rho \left(1 - \max_k \alpha_k \right)}}{\rho^{-MN}} = 0$$

which in turn yields

$$\mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) = 0) \dot{\leqslant} 2^{Rl} e^{-\rho \left(1 - \max_k \alpha_k \right)}$$

$$\dot{\leqslant} \rho^{-MN} \tag{3.56}$$

We next show that the same result holds for the other region $\{\boldsymbol{\alpha} : M(\boldsymbol{\alpha}) \geqslant 1\}$.

Following the same line of argument as we did for (3.56) but for $M(\boldsymbol{\alpha}) \geqslant 1$, we have

$$P(E, \bar{O} | M(\boldsymbol{\alpha}) \geqslant 1) \dot{\leqslant} 2^{Rl} e^{-\left(\sum_{k=1}^{L} \frac{\rho \lambda_k}{(1+\rho \lambda_k)^2}\right)^{-1}}$$

$$\leqslant e^{2Rl} e^{-\left(\sum_k \frac{1}{1+\rho\lambda} - \sum_k \frac{\rho\lambda_k}{(1:\rho\lambda_k)^2}\right)^{-1}}$$

$$= e^{2Rl} \underbrace{e^{-\left(\sum_k \frac{1}{1+\rho\lambda_k}\right)^{-1}}}_{\leqslant 1 \text{ since } M(\boldsymbol{\alpha}) \geqslant 1} e^{\left[-\frac{\sum_k \frac{1}{(\rho\lambda_k+1)^2}}{\left(\sum_k \frac{1}{1+\rho\lambda_k}\right)\left(\sum_k \frac{\rho\lambda_k}{(1+\rho\lambda_k)^2}\right)}\right]}$$

$$\dot{\leqslant} e^{2Rl} e^{\left[-\frac{LM(\boldsymbol{\alpha})}{LM(\boldsymbol{\alpha})\rho^{-\min_k |1-\alpha_k|}}\right]} \tag{3.57}$$

$$\doteq e^{-\rho^{\min_k |1-\alpha_k|}}$$

$$\dot{\leqslant} e^{\rho^{1-\max_k \alpha_k}} \dot{\leqslant} \rho^{-MN} \tag{3.58}$$

where (3.57) is direct application of (3.54) for $M(\boldsymbol{\alpha}) \geqslant 1$, and (3.58) follows from the fact that $|1 - \alpha_k| \geqslant 1$. Note that (3.58) is true for any code length l. Invoking the results of (3.56) and (3.58), we can now evaluate $\mathbb{P}(E, \bar{\mathcal{C}})$ as follows

$$\mathbb{P}(E, \bar{O}) = \int_{M(\boldsymbol{\alpha})=0} \mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) = 0)\mathbb{P}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} + \int_{M(\boldsymbol{\alpha})\geqslant 1} \mathbb{P}(E, \bar{O} | M(\boldsymbol{\alpha}) \geqslant 1)\mathbb{P}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}$$

$$\dot{\leqslant} \rho^{-MN} \int_{M(\boldsymbol{\alpha})=0} \mathbb{P}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} + \rho^{-MN} \int_{M(\boldsymbol{\alpha})\geqslant 1} \mathbb{P}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} \tag{3.59}$$

$$\doteq \rho^{-MN} \tag{3.60}$$

Therefore, $\mathbb{P}(E, \bar{O}) \dot{\leqslant} \rho^{-MN}$ for all regions of $\boldsymbol{\alpha}$. Finally, (3.40) becomes

$$\mathbb{P}_{\text{err}}(R, M, N) \dot{\leqslant} \mathbb{P}(E|O)\, \rho^{-d_{out}} + \mathbb{P}(E, \bar{O})$$

$$\leqslant \rho^{-d_{out}} + \mathbb{P}(E, \bar{O})$$

$$\doteq \rho^{-d_{out}} + \rho^{-MN}$$

$$\doteq \rho^{-d_{out}}$$

$$= P_{out}(R, M, N) \tag{3.61}$$

which establishes the lemma.

From Lemma 3.4.1 and Lemma 3.4.2, we thus get

Theorem 3.4.3 *For MMSE MIMO Receiver under quasi-static channel and joint spatial encoding, the pairwise error probability (PEP) and the outage probability P_{out} are exponentially equal and the diversity gain is $d(R, M, N) = d_{out}(R, M, N)$, where $d_{out}(R, M, N)$ is given in* (3.11).

3.5   Multiple-Access Channel (MAC)

We now extend the result to the MAC channel. Consider a MIMO MAC channel with K users, M transmit antennas per user, N receive antennas (there is no condition on M, N and k). Assume flat fading MIMO channel, the system model is given by

$$\mathbf{y} = \sum_{i=1}^{K} \mathbf{H}_i \mathbf{x}_i + \mathbf{n} = \mathbf{H_e X} + \mathbf{n} \tag{3.62}$$

where $\mathbf{H}_i \in \mathbb{C}^{N \times M}$ is the user i channel matrix whose entries are independent and identically distributed complex Gaussian, $\mathbf{H_e} = [\mathbf{H}_1 \mathbf{H}_2 \ldots \mathbf{H}_K]$ is the overall equivalent channel matrix, $\mathbf{x}_i \in \mathbb{C}^{M \times 1}$ is the transmitted vector of user i, $\mathbf{X} = [\mathbf{x}_1^T \mathbf{x}_2^T \ldots \mathbf{x}_K^T]^T$ is the overall transmitted vector, and $\mathbf{n} \in \mathbb{C}^{N \times 1}$ is the Gaussian noise vector. The vectors $\mathbf{X}$ and $\mathbf{n}$ are assumed independent. We keep the same assumptions about the channel. That is we assume a quasi-static flat fading channel and perfect CSIR and no CSIT. We have the following theorem

Theorem 3.5.1 *In a MIMO MAC system with MMSE receiver consisting of K users, M transmit antennas per user and N receive antennas, the lower and upper bounds on the per user diversity are respectively given by $d_L^{MAC}(R)$ and $d_U^{MAC}(R)$,*

$$d_L^{MAC}(R) = \left\lceil \left(M2^{-R/M} - (M - N)^+ \right)^+ \right\rceil^2 + |N - KM| \left\lceil \left(M2^{-R/M} - (M - N)^+ \right)^+ \right\rceil$$

$$\tag{3.63}$$

$$d_U^{MAC}(R) = \left\lceil \left(KM2^{-R/KM} - (M - N)^+ \right)^+ \right\rceil^2 + |N - KM| \left\lceil \left(KM2^{-R/KM} - (M - N)^+ \right)^+ \right\rceil.$$

$$\tag{3.64}$$

From (3.63) it is straightforward to verify the single user case. The machinery of the proof is mostly similar to the single user case. However, the outage upper and lower bounds are obtained in a different manner that is pointed out in the following analysis for $N \geqslant M$. The case $N < M$ can be similarly obtained.

3.5.1 MAC Outage Upper Bound

The user i outage probability can be written as

$$P_{out}^i = \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(1 + \gamma_k^i) < R \right). \tag{3.65}$$

where γ_k^i is the SINR of the stream k of user i. Specializing this to MMSE receiver we get

$$P_{out}^i = \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right). \tag{3.66}$$

Using Jensen's Inequality the outage probability can be bounded as

$$P_{out}^i \leqslant \mathbb{P}\left(\log \left(\sum_{k=(i-1)M+1}^{iM} \frac{1}{M}(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} \right) > \frac{-R}{M} \right)$$

$$\leqslant \mathbb{P}\left(\log \left(\sum_{k=1}^{KM} \frac{1}{M}(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} \right) > \frac{-R}{M} \right) \tag{3.67}$$

$$= \mathbb{P}\left(\sum_{k=1}^{KM} \frac{1}{1 + \rho \lambda_k} > M 2^{-\frac{R}{M}} \right) \tag{3.68}$$

where (3.67) is true since the summation in the left-hand side of the inequality adds more positive terms (recall that $(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})$ is a positive definite matrix [17]). Following similar steps that were used to obtain (3.26) we can easily show that $P_{out}^i \leqslant \rho^{-d_L^{\text{MAC}}}$, where d_L^{MAC} is given by (3.63).

3.5.2 MAC Outage Lower Bound

The outage probability can be lower bounded as follows

$$P_{out}^i = \mathbb{P}\left(\sum_{k=(i-1)M+1}^{iM} \log(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right)$$

$$\geqslant \mathbb{P}\left(\sum_{k=1}^{KM} \log(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right) \tag{3.69}$$

$$\dot{\geqslant} \mathbb{P}\left(\sum_{k=1}^{KM} (\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > \frac{KM}{a} 2^{\frac{-R}{M}} \right) \tag{3.70}$$

where (3.69) is a trivial bound based on dedicating all KM antennas to *one* user, and (3.70) uses the same technique as in Section 3.3.2, and a is a positive number slightly less than one. Following similar steps that were used to obtain (3.26) we can easily show that $P_{out}^i \dot{\geqslant} \rho^{-d_U^{\text{MAC}}}$, where d_U^{MAC} is given by (3.64).

3.6 Simulation Results

Simulations generate Monte Carlo random channel realizations and calculate outage probability by checking the appropriate linear MIMO receiver mutual information for the quasi-static flat fading model. Figure 3.3 shows the case $M = N = 3$. According to Theorem 3.4.3, $d_{out} = 1$ for $R \geqslant 4.755$, $d_{out} = 4$ for $4.755 > R \geqslant 1.755$, and $d_{out} = 9$ for $R < 1.7549$. Figure 3.3 shows the diversity step between $R = 4.5$ and 4.8bps/Hz. The slope of diversity 9 is difficult to measure precisely with simulations, but it is approximately observed. Figure 3.4 shows the outage probability for $R = 1, 4$ and 10 with the Jensen bound, with a diversity transition at $R = 2$. Figure 3.5 shows the case of $M = 2$, and $N = 3$ again with transition at $R = 2$. In Figure 3.6, simulations results for $N = 2$ and $M = 3$ are given and compared with $N = 3$ and $M = 2$. Theorem 3.4.3 gives the diversity for both systems. It is observed that when $N > M$ the break point of the slopes occurs before its counterparts in $M > N$ case. Lower rates were difficult to simulate precisely.
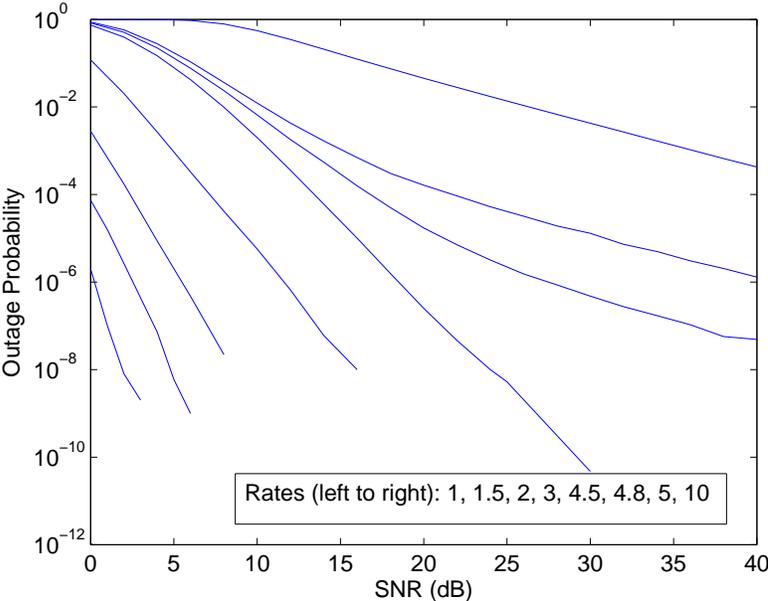
Figure 3.3. Outage probability of MMSE Receiver, $M = N = 3$ for R=1, 1.5, 2, 3, 4.5, 4.8, 5, 10 bps/Hz
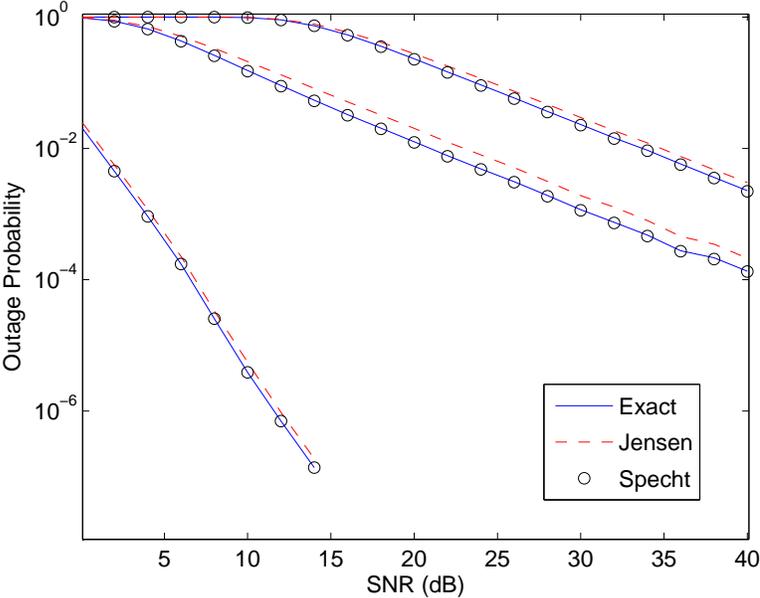
Figure 3.4. Outage probability of MMSE Receiver, $M = N = 2$ for R (left to right)= 1, 4, 10 bps/Hz
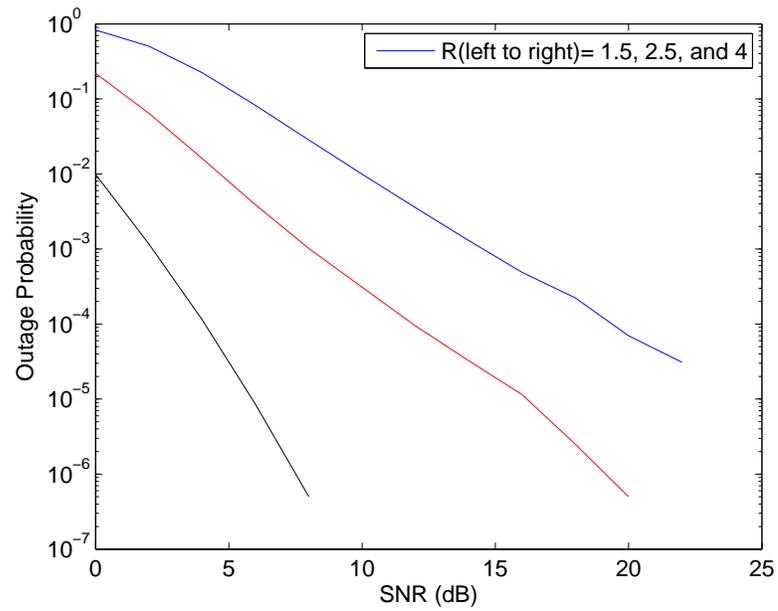
Figure 3.5. Outage probability of MMSE Receiver, $M = 2$, and $N = 3$ for R (left to right)= 1.5, 2.5, 4 bps/Hz
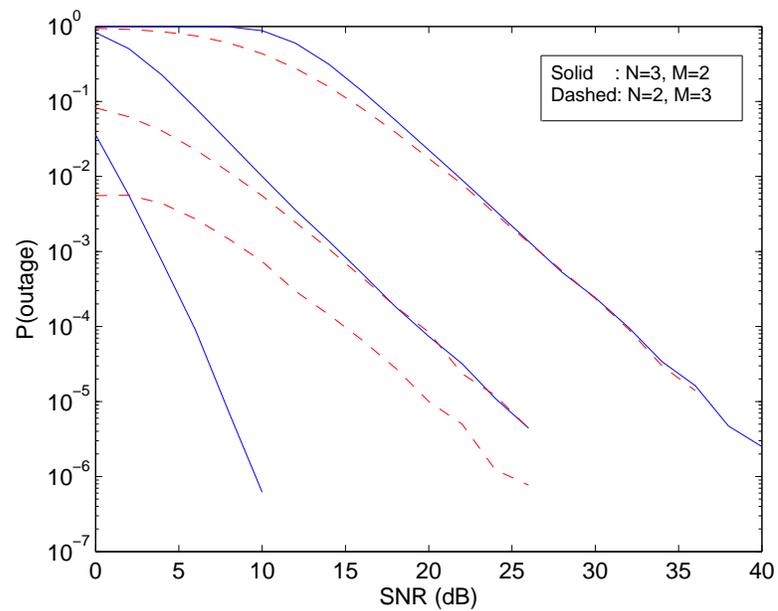
Figure 3.6. Outage probability of MMSE Receiver for both cases $N > M$ (solid) and $M > N$ (dashed). The spectral efficiency R (left to right)= 1.8, 4, and 10 bps/Hz

3.7 Conclusion

This chapter settles the long standing problem of the diversity of the MMSE MIMO receivers under all fixed rates for any number of transmit (M) and receive (N) antennas, giving the result as $d = \lceil M2^{-\frac{R}{M}} - \kappa\rceil^2 + |N - M|\lceil M2^{-\frac{R}{M}} - \kappa\rceil$, where $\kappa = max(0, M - N)$. The analysis confirms the earlier approximate results $[6, 7]$ showing that the system diversity can be as high as MN for low spectral efficiency and as low as $N - M + 1$ for high spectral efficiency. The result is easily extended to the multiple access channel (MAC).

CHAPTER 4

LINEAR RECEIVERS IN MIMO FREQUENCY-SELECTIVE CHANNEL

4.1   Frequency-Selective Channel

4.2   Introduction

Broadband wireless systems usually operate in frequency-selective channels where, in addition to the spatial diversity obtained in MIMO broadband systems, frequency diversity can be achieved. Broadband systems usually employ orthogonal frequency division multiplexing (OFDM) or single carrier (SC) transmission [20]. Specifically, SC was shown to be attractive for broadband wireless channels due to its lower complexity, lower peak-to-average power ratio and reduced sensitivity to carrier frequency errors compared to OFDM [20, 21].

In this section, we investigate the diversity achieved by SC-MMSE receivers for two block transmission schemes, namely cyclic prefix (CP) and zero-padding (ZP) schemes. The CP and ZP are commonly used for guard intervals in block quasi-static channels. Although CP was initially proposed for both single carrier and multi-carrier systems, ZP was lately shown to be an attractive alternative for both systems [22, 23].

4.2.1   System Model

We consider a general MIMO system in a rich scattering quasi-static environment. The equivalent baseband channel is given by multipath model with ν paths referred to as the ISI channel in the sequel. The $(\nu + 1)$-tap channel impulse response between the transmit antenna m and receive antenna n is denoted by the vector $\mathbf{h}_{mn} = [h_{mn,0}, h_{mn,1}, \ldots, h_{mn,\nu}]$. We assume a block-fading model where $\mathbf{h}_{mn}$ remains unchanged during a transmission block. Assuming M transmit and N receive antennas, the received vector $\mathbf{y}_k$ at time instant k is

given by [15, 24]

$$\mathbf{y}_k = \sum_{i=0}^{\nu} \mathbf{H}_i \mathbf{x}_{k-i} + \mathbf{n}_k \tag{4.1}$$

where $\mathbf{H}_i$ is the $M \times N$ channel matrix that has $h_{mn,i}$ as its (m, n) element, $\mathbf{x}_{k-i}$ is $M \times 1$ transmitted vector at time index $k - i$, $\mathbf{y}_k$ is the $N \times 1$ received vector and $\mathbf{n}_k$ is the $N \times 1$ Gaussian noise vector at time index k.

Consider a transmission of $L_d + L_e$ spatial vectors each of size $M \times 1$, where L_d is an integer representing the number of transmissions over the quasi-static channel and L_e is the length of data extension to avoid inter-block interference, in the form of either zero-padding or cyclic prefix. The receiver discards the L_e vectors in the case of cyclic-prefix transmission [24]. Stacking the transmitted vector in an $M(L_d + L_e) \times 1$ vector, we can write the stacked $M(L_d + L_e) \times 1$ transmitted as follows

$$\bar{\mathbf{x}}_k = [\mathbf{x}_{k(L_d+L_e)}^T, \ldots, \mathbf{x}_{k(L_d+L_e)+L_d+L_e-1}^T]$$

We can then rewrite (4.1) as

$$\bar{\mathbf{y}}_{cp} = \bar{\mathbf{H}}\,\bar{\mathbf{x}} + \bar{\mathbf{n}} \tag{4.2}$$

where $\bar{\mathbf{y}}_{cp}$ is the $NL_d \times 1$ received vector, $\bar{\mathbf{x}}$ is the $M(L_d + L_e) \times 1$ transmitted vector, $\bar{\mathbf{n}}$ is the white Gaussian noise vector $\in \mathbb{C}^{NL_d \times 1}$ and $\bar{\mathbf{H}}$ is the channel matrix given by

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & 0 & \cdots & 0 \\ 0 & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & \cdots & \mathbf{H}_0 & \mathbf{H}_1 & \cdots & \mathbf{H}_\nu \end{bmatrix}. \tag{4.3}$$

The linear data extension operation maps the data vector $\hat{\mathbf{x}}$ to the transmitted vector $\bar{\mathbf{x}}$ and is shown by

$$\bar{\mathbf{x}} = \mathbf{U}_{\mathbf{cp}}\hat{\mathbf{x}} \tag{4.4}$$

where $\mathbf{U}_{\mathbf{cp}}$ is given by

$$\mathbf{U}_{\mathbf{cp}} = \begin{bmatrix} & \mathbf{I}_{ML_d} & \\ \mathbf{I}_{ML_e} & & \mathbf{0}_{ML_e \times (L_d - L_e)M} \end{bmatrix} \tag{4.5}$$

The system model in (4.2) can now be written in terms of the unpadded data vector $\hat{\mathbf{x}}$ and an equivalent channel matrix $\mathbf{H_e}$ as follows

$$\bar{\mathbf{y}}_{cp} = \mathbf{H_e}\,\hat{\mathbf{x}} + \bar{\mathbf{n}} \tag{4.6}$$

where in a CP system, $\mathbf{H_e} = \bar{\mathbf{H}}\mathbf{U_{cp}}$ is a $NL_d \times ML_d$ block circulant matrix constructed by block circulations of the matrix $[\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_\nu, 0, \ldots, 0]^T$.

For the zero-padding transmission, we can rewrite (4.1) as

$$\bar{\mathbf{y}}_{zp} = \mathbf{H_e}\,\hat{\mathbf{x}} + \bar{\mathbf{n}} \tag{4.7}$$

where $\bar{\mathbf{y}}_{zp}$ is the $N(L_d + L_e) \times 1$ received vector, $\bar{\mathbf{x}}$ is the $ML_d \times 1$ transmitted vector, $\bar{\mathbf{n}}$ is the white Gaussian noise vector $\in \mathbb{C}^{N(L_d+L_e)\times 1}$ and $\bar{\mathbf{H}}$ is the channel matrix given by

$$\mathbf{H_e} = \begin{bmatrix} \mathbf{H}_0 & 0 & \cdots & 0 \\ \vdots & \mathbf{H}_1 & \ddots & \vdots \\ \mathbf{H}_\nu & \vdots & \ddots & \mathbf{H}_0 \\ 0 & \mathbf{H}_\nu & \ddots & \vdots \\ \vdots & \vdots & \vdots & \mathbf{H}_\nu \end{bmatrix}. \tag{4.8}$$

Assuming perfect channel state information at the receiver (CSIR) and that the channel remains unchanged during the transmission of $L_d + L_e$ vectors, the MMSE equalizer $\mathbf{W}$ is applied to decouple the received streams (after removing the L_e extension vectors in case of cyclic-prefix transmission). The MMSE equalizer is given by

$$\mathbf{W} = (\rho^{-1}\mathbf{I} + \mathbf{H_e}^H\mathbf{H_e})^{-1}\mathbf{H_e}^H \tag{4.9}$$

and the unbiased decision-point SINRs of the equalizers output for detecting the kth transmitted stream are

$$\gamma_k = \frac{1}{(\mathbf{I} + \rho\mathbf{H_e}^H\mathbf{H_e})^{-1}_{kk}} - 1 \qquad k = 1, \ldots, ML_d. \tag{4.10}$$

In the following sections we analyze the outage diversity for the ZP and CP systems. The PEP analysis follows in a direct manner as in the flat fading case so we omit it.

4.2.2 The Zero Padding MMSE Receiver

It is known that in a point-to-point *single-antenna* ISI channel, linear receivers can achieve full multipath diversity under zero-padding transmission [23, 25, 26]. In this section we investigate the similar question for MIMO systems whose receivers use linear MMSE operations in both the spatial and temporal dimensions. We provide lower and upper bounds on diversity. The bounds are not always tight, but the diversity is fully characterized for SIMO systems.

We begin by analyzing the tradeoff between the spectral efficiency R and the diversity of MMSE receiver in the *single-antenna* ISI channel d_{MMSE}^{ISI} under ZP transmission. Tajer *et al* [15] shows that d_{MMSE}^{ISI} varies with R under CP transmission and MMSE equalization, in particular, for a quasi-static single-antenna ISI channel with $\nu + 1$ taps, the diversity of the SC-MMSE receiver under CP transmission is $d_{MMSE}^{CP} = 1 + \min(\nu, \lfloor 2^{-R} L_d \rfloor)$, where L_d is the transmission data block length. We show that the same is not true for ZP transmission.

Lemma 4.2.1 *For a quasi-static single-antenna ISI channel with $\nu + 1$ taps, the diversity of the SC-MMSE receiver under ZP transmission is $d_{MMSE}^{ZP} = \nu + 1$ irrespective of R.*

Proof See Appendix 4.3.1.

We proceed with lower and upper bounds on diversity for MIMO ISI channel.

Diversity Upper Bound

Applying the MMSE equalizer given by (4.9) to the received vector in (4.6), the effective mutual information between $\hat{\mathbf{x}}$ and $\mathbf{W}\bar{\mathbf{y}}$ is equal to the sum of mutual information of their components [12]

$$I(\hat{\mathbf{x}}, \mathbf{W}\bar{\mathbf{y}}) = \frac{1}{L_d} \sum_{k=1}^{ML_d} I(x_k, y_k).$$

Thus the outage probability is given by

$$P_{out} = \mathbb{P}\left(\frac{1}{L_d}\sum_{k=1}^{ML_d}\log(1+\gamma_k) < R\right) \tag{4.11}$$

$$= \mathbb{P}\left(\frac{1}{L_d}\sum_{k=1}^{ML_d}\log\frac{1}{(\mathbf{I}+\rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}} < R\right) \tag{4.12}$$

$$\geqslant \mathbb{P}\left(M\log\frac{1}{ML_d}\sum_{k=1}^{ML_d}\frac{1}{(\mathbf{I}+\rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}} < R\right) \tag{4.13}$$

where we have used Jensen's inequality as in Section 3.3.2. Let the eigen decomposition of $\mathbf{H_e}^H\mathbf{H_e}$ be given by $\mathbf{H_e}^H\mathbf{H_e} = \mathbf{U}^H\Lambda\mathbf{U}$ where $\mathbf{U}$ is unitary and Λ is a diagonal matrix that has the eigenvalues of the matrix $\mathbf{H_e}^H\mathbf{H_e}$ on its diagonal. Let the eigenvalues of $\mathbf{H_e}^H\mathbf{H_e}$ be given by $\{\lambda_\ell\}$ with $\lambda_1 \geqslant \lambda_2\cdots \geqslant \lambda_{ML_d}$. Let the vector $\mathbf{u}_k$ be the column k of the matrix $\mathbf{U}$, we have

$$(\mathbf{I}+\rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1} = \mathbf{u}_k^H(\mathbf{I}+\rho\Lambda)^{-1}\mathbf{u}_k$$

$$= \sum_{\ell=1}^{ML_d}\frac{|u_{\ell k}|^2}{1+\rho\lambda_\ell}$$

$$\triangleq S_k.$$

Let $\bar{k} = \arg\min_k S_k$. we can bound the sum in (4.13)

$$\frac{1}{ML_d}\sum_{k=1}^{ML_d}\frac{1}{(\mathbf{I}+\rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}} = \frac{1}{ML_d}\sum_{k=1}^{ML_d}\frac{1}{S_k}$$

$$\leqslant \frac{1}{\min_k S_k}$$

$$= \frac{1}{S_{\bar{k}}} \tag{4.14}$$

thus the outage bound in (4.13) can be further bounded

$$P_{out} \geqslant \mathbb{P}\left(M\log\frac{1}{ML_d}\sum_{k=1}^{M}\frac{1}{(\mathbf{I}+\rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}} < R\right)$$

$$\geqslant \mathbb{P}\left(M\log\frac{1}{S_{\bar{k}}} < R\right)$$

$$= \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}}\right) \tag{4.15}$$

We now bound (4.15) by conditioning on the event

$$\mathcal{B} \triangleq \left\{ |u_{\ell\bar{k}}|^2 \geqslant \frac{a}{M}, \ \ell = ML_d - M + 1, \cdots, ML_d \right\} \tag{4.16}$$

where a is a positive real number that is slightly smaller than one $a = 1 - \epsilon_1$, and ϵ_1 is a small positive number. We then have

$$P_{out} = \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \right)$$

$$\geqslant \mathbb{P}\left(S_{\bar{k}} > 2^{-\frac{R}{M}} \Big| \mathcal{B} \right) \mathbb{P}(\mathcal{B})$$

$$= \mathbb{P}\left(\sum_{\ell=1}^{ML_d} \frac{|u_{\ell\bar{k}}|^2}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}} \Big| \mathcal{B} \right) \mathbb{P}(\mathcal{B})$$

$$\geqslant \mathbb{P}\left(\sum_{\ell=ML_d-M+1}^{ML_d} \frac{|u_{\ell\bar{k}}|^2}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}} \Big| \mathcal{B} \right) \mathbb{P}(\mathcal{B}) \tag{4.17}$$

$$\geqslant \mathbb{P}\left(\frac{1}{M} \sum_{\ell=ML_d-M+1}^{ML_d} \frac{a}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}} \right) \mathbb{P}(\mathcal{B})$$

$$\doteq \mathbb{P}\left(\frac{1}{M} \sum_{\ell=ML_d-M+1}^{ML_d} \frac{a}{1 + \rho\lambda_\ell} > 2^{-\frac{R}{M}} \right) \tag{4.18}$$

$$= \mathbb{P}\left(\sum_{\ell=ML_d-M+1}^{ML_d} \frac{1}{1 + \rho\lambda_\ell} > \frac{M}{a} 2^{-\frac{R}{M}} \right) \tag{4.19}$$

where (4.17) follows by removing some of the elements of the sum corresponding to the largest eigenvalues. The steps used to obtain Eq. (4.18) are similar to the steps used in Section 3.3.2.

Note that $\mathbf{H_e}^H \mathbf{H_e}$ is not a Wishart matrix, hence the analysis of Section 3.2 does not directly apply here. The block diagonal elements of $\mathbf{H_e}^H \mathbf{H_e}$ are similar and are given by

$$\mathbf{D} = \sum_{i=0}^{\nu} \mathbf{H}_i^H \mathbf{H}_i. \tag{4.20}$$

The matrix $\mathbf{H_e}^H \mathbf{H_e}$ is Toeplitz and Hermitian. Moreover, the matrix $\mathbf{D}$ given by (4.20) is a Wishart matrix[1].

---

[1] Let $\mathcal{W}(n, \sum)$ denote a Wishart distribution with degree of freedom n and covariance (also called scale) matrix $\sum$. Any of the diagonal block matrices $\mathbf{D}_j$ given by (4.20) follows

Observe that the probability in (4.19) depends on the M smallest eigenvalues. We now bound these eigenvalues with the eigenvalues of the matrix $\mathbf{D}$ via the Sturmian separation theorem [27, P.1077].

Theorem 4.2.2 (Sturmian Separation Theorem) *Let $\{\mathbf{A}_r, r = 1, 2, \ldots\}$ be a sequence of symmetric $r \times r$ matrices such that each $\mathbf{A}_r$ is a submatrix of $\mathbf{A}_{r+1}$. Then if $\{\lambda_k(\mathbf{A}_r)$, $k = 1, \ldots, r\}$ denote the ordered eigenvalues of each matrix $\mathbf{A}_r$ in descending order, we have*

$$\lambda_{k+1}(\mathbf{A}_{i+1}) \leq \lambda_k(\mathbf{A}_i) \leq \lambda_k(\mathbf{A}_{i+1}).$$

For our purposes, we consider a special case of the Sturmian Theorem by constructing a set of matrices $\mathbf{A}_M, \mathbf{A}_{M+1}, \ldots, \mathbf{A}_{L_dM}$ starting by the largest one $\mathbf{A}_{L_dM} \triangleq \mathbf{H_e}^H\mathbf{H_e}$ and making all other matrices $\mathbf{A}_i$ to be (successively embedded) $i \times i$ principal submatrices of $\mathbf{H_e}^H\mathbf{H_e}$, such that the smallest matrix is $\mathbf{A}_M = \mathbf{D}_{L_d}$. Then we repeatedly apply the first inequality in the Sturmian to get:

$$\lambda_{ML_d}(\mathbf{A}_{ML_d}) \leq \lambda_{ML_d-1}(\mathbf{A}_{ML_d-1}) \leq \cdots \leq \lambda_M(\mathbf{A}_M)$$

$$\lambda_{ML_d-1}(\mathbf{A}_{ML_d}) \leq \lambda_{ML_d-2}(\mathbf{A}_{ML_d-1}) \leq \cdots \leq \lambda_{M-1}(\mathbf{A}_M)$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\lambda_{ML_d-M+1}(\mathbf{A}_{ML_d}) \leq \lambda_{ML_d-M}(\mathbf{A}_{ML_d-1}) \leq \cdots \leq \lambda_1(\mathbf{A}_M)$$

This implies that the smallest M eigenvalues of $\mathbf{H_e}^H\mathbf{H_e}$ are bounded above by the M eigenvalues of $\mathbf{D}$, respectively. Hence:

$$P_{out} \dot{\geqslant} \mathbb{P}\Big(\sum_{k=1}^{M} \frac{1}{1 + \rho\lambda_k(\mathbf{D})} > \frac{M}{a} 2^{-\frac{R}{M}}\Big). \tag{4.21}$$

$\mathbf{D}$ is a sum of $(\nu + 1)$ central Wishart matrices each with N degrees of freedom and with identity covariance matrix, i.e. $\mathbf{D} \in \mathcal{W}((\nu + 1)N, I)$. Therefore the analysis of Section 3.2 applies here and we have the following lemma.

---

a Wishart distribution since if $\mathbf{B}_1 \in \mathcal{W}(n_1, \sum)$ and $\mathbf{B}_2 \in \mathcal{W}(n_2, \sum)$ then $\mathbf{B}_1 + \mathbf{B}_2 \in \mathcal{W}(n_1 + n_2, \sum)$.

Lemma 4.2.3 *In a MIMO quasi-static frequency-selective system (with channel memory ν) consisting of M transmit and N receive antennas, the MMSE receiver diversity under joint spatial encoding and zero-padding transmission is upper bounded as*

$$d^{ZP} \leqslant \left\lfloor \left(M2^{-\frac{R}{M}} + 1 - (M - \bar{N})^+ \right)^+ \right\rfloor^2 + |\bar{N} - M| \left\lfloor \left(M2^{-\frac{R}{M}} + 1 - (M - \bar{N})^+ \right)^+ \right\rfloor \quad (4.22)$$

where $\bar{N} = (\nu + 1)N$.

Diversity Lower Bound

We can upper bound the outage probability as follows.

$$
\begin{aligned}
P_{out} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log(1 + \gamma_k) < R \right) \\
&= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right) \\
&\leqslant \mathbb{P}\left(M \log \frac{1}{ML_d} \sum_{k=1}^{ML_d} (\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right) \quad (4.23) \\
&\leqslant \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^{ML_d} (\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right) \\
&= \mathbb{P}\left(\sum_{k=1}^{ML_d} \frac{1}{1 + \rho \lambda_k(\mathbf{H_e}^H \mathbf{H_e})} > M2^{-\frac{R}{M}} \right) \\
&\leqslant \mathbb{P}\left(\sum_{k=1}^{M} \frac{1}{1 + \rho \lambda_k(\mathbf{H_e}^H \mathbf{H_e})} + L_d M - M > M2^{-\frac{R}{M}} \right) \quad (4.24) \\
&= \mathbb{P}\left(\sum_{k=1}^{M} \frac{1}{1 + \rho \lambda_k(\mathbf{H_e}^H \mathbf{H_e})} > M2^{-\frac{R}{M}} - (ML_d - M) \right) \quad (4.25)
\end{aligned}
$$

where (4.23) follows from Jensen's inequality and (4.24) follows from setting the smallest $L_d M - M$ eigenvalues to zero.

Now we repeatedly use the second inequality in the Sturmian theorem to get

$$\lambda_M(\mathbf{A}_M) \leq \cdots \quad \leq \lambda_M(\mathbf{A}_{ML_d-1}) \leq \lambda_M(\mathbf{A}_{ML_d})$$

$$\lambda_{M-1}(\mathbf{A}_M) \leq \cdots \leq \lambda_{M-1}(\mathbf{A}_{ML_d-1}) \leq \lambda_{M-1}(\mathbf{A}_{ML_d})$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\lambda_1(\mathbf{A}_M) \leq \cdots \quad \leq \lambda_1(\mathbf{A}_{ML_d-1}) \leq \lambda_1(\mathbf{A}_{ML_d})$$

with $\mathbf{A}_{ML_d} \triangleq \mathbf{H_e}^H \mathbf{H_e}$ and $\mathbf{A}_M \triangleq \mathbf{D}$, similar to the earlier case. Therefore the largest M eigenvalues of $\mathbf{H_e}^H \mathbf{H_e}$ are bounded below by the M eigenvalues of $\mathbf{D}$, respectively. Therefore

$$P_{out} \dot{\leqslant} \mathbb{P}\left(M \log \frac{1}{M} \sum_{k=1}^{M} \frac{1}{1 + \rho \lambda_k(\mathbf{D})} > Q\right). \tag{4.26}$$

where $Q = \max\left(0, M 2^{-\frac{R}{M}} - (ML_d - M)\right)$. Recall that $\mathbf{D}$ is a Wishart matrix, therefore the analysis of Section 3.2 follows and we obtain the following lemma.

Lemma 4.2.4 *In a MIMO quasi-static frequency-selective system (with channel memory ν) consisting of M transmit and N receive antennas, the MMSE receiver diversity is lower bounded as*

$$d^{ZP} \geqslant \lceil Q \rceil^2 + |(\nu + 1)N - M|\lceil Q \rceil \tag{4.27}$$

under joint spatial encoding and zero-padding transmission. $Q = \max\left(0, M 2^{-\frac{R}{M}} - (ML_d - M)\right)$.

Remark 4.2.1 *Notice that both lower and upper bounds differ only in the second term of Q, i.e. $(ML_D - M)$. The diversity lower bound for $L_d = 1$ is tight against the upper bound, but for $L_d > 1$ the lower bound (4.27) is trivial.*

4.2.3 The Cyclic Prefix MMSE Receiver

For the *single-antenna* ISI channel under CP transmission, the explicit tradeoff between spectral efficiency and diversity was found [15] to be $d^{CP}_{MMSE} = 1 + \min(\nu, \lfloor 2^{-R}L_d \rfloor)$. In this
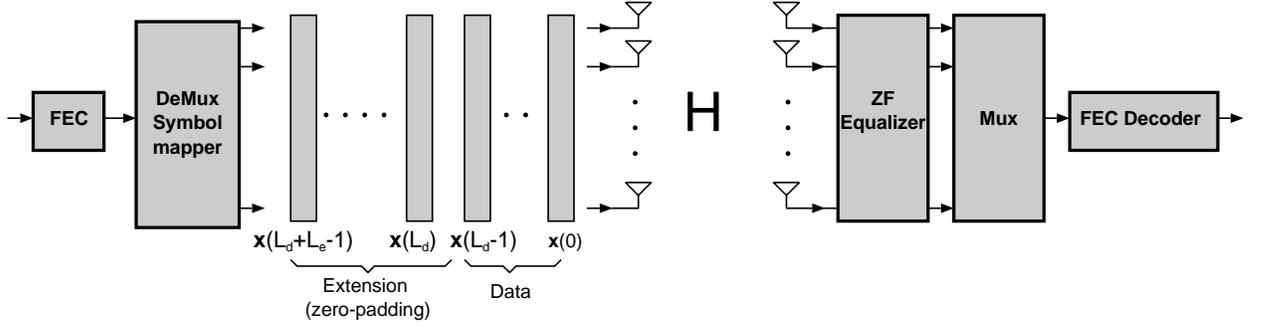
Figure 4.1. Single-carrier block transmission in a frequency-selective channel. In the case of CP, the extension is removed at the receiver prior to equalization.

section, we extend the analysis to the MIMO case. The system model is shown in Figure 4.1. We start with the general $M \times N$ MIMO system.

The system model is again given by (4.6) where $\mathbf{H_e} = \bar{\mathbf{H}}\mathbf{U_{cp}}$ and $\hat{\mathbf{x}}$ is generated by taking the IDFT of the information vector $\mathbf{x}$ [28], i.e.

$$\hat{\mathbf{x}} = \mathbf{Q}_{Tx}^H \mathbf{x} \tag{4.28}$$

where $\mathbf{Q}_{Tx}$ is the augmented DFT matrix given by $\mathbf{Q}_{Tx} = \mathbf{Q} \otimes \mathbf{I}_M$, where $\mathbf{I}_M$ is the identity matrix, $\mathbf{Q}$ is the normalized DFT matrix, and $\otimes$ is the Kroenecker product.

The $NL_d \times ML_d$ block-circulant matrix $\mathbf{H_e}$ has eigen decomposition $\mathbf{H_e} = \mathbf{Q}_{Rx}^H \mathbf{\Lambda} \mathbf{Q}_{Tx}$, where $\mathbf{Q}_{Rx} = \mathbf{Q} \otimes \mathbf{I}_N$. Both $\mathbf{Q}_{Tx}$ and $\mathbf{Q}_{Rx}$ are unitary matrices. The block diagonal matrix $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{B}_1 & & & 0 \\ & \mathbf{B}_2 & & \\ & & \ddots & \\ 0 & & & \mathbf{B}_{L_d} \end{pmatrix} \tag{4.29}$$

where the matrix $\mathbf{B}_k$ is given by [29]

$$\mathbf{B}_k = \sum_{i=0}^{\nu} \mathbf{H}_i e^{-j\frac{2\pi i(k-1)}{L_d}} \quad \text{for } k = 1, \ldots, L_d \tag{4.30}$$

and $\mathbf{H}_i$ is the instantaneous MIMO channel (cf. Section 4.2.1).

Analogous to the proof of [15], we first consider the case where the transmission data-block length is equal to the number of channel taps, i.e. $L_d = \nu + 1$. In this case, the entries of $\mathbf{B}_k's$ are i.i.d. normal complex Gaussian.

Outage upper bound

The outage probability of the MMSE receiver is given by

$$P_{out} = \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log\left(\frac{1}{(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1}} \right) < R \right) \tag{4.31}$$

$$= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log\left((\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} \right) > -R \right)$$

$$\leqslant \mathbb{P}\left(M \log \sum_{k=1}^{ML_d} \frac{1}{ML_d} (\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1} > -R \right) \tag{4.32}$$

$$= \mathbb{P}\left(M \log \sum_{k=1}^{ML_d} \frac{1}{ML_d} (\mathbf{I} + \rho \mathbf{\Lambda}^H \mathbf{\Lambda})_{kk}^{-1} > -R \right) \tag{4.33}$$

$$= \mathbb{P}\left(\sum_{k=1}^{ML_d} (\mathbf{I} + \rho \mathbf{\Lambda}^H \mathbf{\Lambda})_{kk}^{-1} > ML_d 2^{-\frac{R}{M}} \right)$$

$$= \mathbb{P}\left(\sum_{i=1}^{L_d} \mathrm{tr}(\mathbf{I} + \rho \mathbf{B}_i^H \mathbf{B}_i)^{-1} > ML_d 2^{-\frac{R}{M}} \right)$$

$$= \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^{M} \frac{1}{(1 + \rho \lambda_{k,i})} > ML_d 2^{-\frac{R}{M}} \right) \tag{4.34}$$

Where (4.32) follows from Jensen's inequality, (4.33) follows from the eigen decomposition of $\mathbf{H_e}$, and $\lambda_{k,i}$ is k-th eigenvalue of the i-th Wishart matrix $\mathbf{B}_i^H \mathbf{B}_i$.

Recall from Section 3.3 that the eigenvalues of a Wishart matrix have the asymptotic property

$$\sum_{k=1}^{M} \frac{1}{1 + \rho \lambda_k} \doteq \sum_{\alpha_k > 1} 1 + \sum_{\alpha_k < 1} \rho^{\alpha_k - 1} \tag{4.35}$$

based on which we established in Lemmas 3.3.1 and 3.3.2 the following

$$\mathbb{P}\Big(\sum_{k=1}^{M}\frac{1}{1+\rho\lambda_k}\geqslant s\Big)\doteq\rho^{-(s^2+|N-M|s)} \tag{4.36}$$

where α_k is defined in (3.15) and s, M, and N are arbitrary integers. Define

$$\theta_i\triangleq\sum_{\alpha_{k,i}>1}1$$

θ_i are i.i.d. discrete random variables with the following asymptotic distribution (cf. Section 3.3, Equations (3.22)-(3.26))

$$\mathbb{P}\big(\theta_i=n_i\big)\doteq\rho^{-(n_i^2+|N-M|n_i)}\quad\text{for }n_i=1,\ldots,M \tag{4.37}$$

Using (4.36), the outage probability in (4.34) can be evaluated as

$$P_{out}\dot{\leqslant}\mathbb{P}\Big(\sum_{i=1}^{L_d}\sum_{k=1}^{M}\frac{1}{(1+\rho\lambda_{k,i})}>ML_d2^{-\frac{R}{M}}\Big)$$

$$\doteq\mathbb{P}\Big(\sum_{i=1}^{L_d}\theta_i\geqslant\Omega\Big) \tag{4.38}$$

where $\Omega=\lceil ML_d2^{-\frac{R}{M}}\rceil$. Evaluating the probability in (4.38) in a combinatorial manner, we get

$$\mathbb{P}\Big(\sum_{i=1}^{L_d}\theta_i\geqslant\Omega\Big)\doteq\mathbb{P}\Big(\sum_{i=1}^{L_d}\theta_i=\Omega\Big)$$

$$\doteq\sum_{n_1,n_2,\ldots,n_{L_d}}\rho^{-(n_1^2+|N-M|n_1)}\ldots\rho^{-(n_p^2+|N-M|n_{L_d})} \tag{4.39}$$

$$\doteq\max_{n_1,n_2,\ldots,n_{L_d}}\rho^{-(n_1^2+|N-M|n_1)}\ldots\rho^{-(n_{L_d}^2+|N-M|n_{L_d})} \tag{4.40}$$

where $n_i\in[0,M]$ for $(i=1,2,\ldots,L_d)$ is the value of the i-th discrete random variable θ_i, and (4.40) is true since the summation in (4.39) is dominated by the maximum element.

Let the set $\{n_k^*,k=1,\ldots,L_d\}$ be the set of indices of the optimal solution of (4.40). The set $\{n_k^*\}$ is obtained by solving the following optimization problem

$$\min_{n_1,n_2,\ldots,n_{L_d}}\sum_{k=1}^{L_d}(n_k^2+|N-M|n_k)$$

$$\text{subject to} \quad \sum_{k=1}^{L_d} n_k = \Omega$$

$$0 \leqslant n_k \leqslant M$$

or equivalently,

$$\min_{n_1, n_2, \dots, n_{L_d}} \quad \sum_{k=1}^{L_d} n_k^2 \tag{4.41}$$

$$\text{subject to} \quad \sum_{k=1}^{L_d} n_k = \Omega$$

$$n_k \geq 0$$

The problem in (4.41) is a quadratic integer-programming (QIP) problem (see e.g. [30]). Integer programming problems are in general NP-hard. However, due to the simple structure of the objective function in (4.41), we can efficiently solve it, thus obtain a closed form expression for $\{n_k^*\}$ and hence (4.40).

Lemma 4.2.5 *For the QIP given by (4.41), the optimum solution is given by:*

$$n_i^* = u \quad \text{for } 1 \leqslant i \leqslant t$$

$$n_j^* = u + 1 \quad \text{for } t + 1 \leqslant j \leqslant L_d$$

where $u = \lfloor \frac{\Omega}{L_d} \rfloor$ and $t = L_d(u+1) - \Omega$.

Proof *See Appendix 4.3.2*

Using Lemma 4.2.5, we can now evaluate the outage upper bound given by (4.40) as

$$P_{out} \dot{\leqslant} \rho^{-d_{cp}} \tag{4.42}$$

where $d_{cp} = \Omega(2u+1) - uL_d(u+1) + |N - M|\Omega$ and $u = \lfloor \frac{\Omega}{L_d} \rfloor$.

Outage lower bound

The bound is obtained using the same steps to obtain the lower bound in Section 4.2.2. It can be shown that

$$P_{out} = \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{ML_d} \log((\mathbf{I} + \rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}) > -R \right) \tag{4.43}$$

$$\dot{\geqslant} \mathbb{P}\left(\sum_{i=1}^{L_d} \sum_{k=1}^{M} \frac{1}{(1 + \rho\lambda_{k,i})} > ML_d 2^{-\frac{R}{M}} \right) \tag{4.44}$$

The bound in (4.44) is the same as the upper bound in (4.34), thus the bound is tight and the diversity is given by (4.42). The PEP analysis follows in a manner similar to Section 3.4.

Recall that so far we have considered data block length $L_d = \nu + 1$. It can be shown that the diversity for any $L_d > \nu + 1$ is upper bounded by the computed diversity for the case $L_d = \nu + 1$. This bounding is derived from (4.33) via FFT arguments similar to those used in [15], which we omit for brevity. A tight diversity *lower* bound for data block lengths $L_d > \nu + 1$ remains an open problem, except for the SIMO system as discussed in the next section.

Diversity of CP Transmission in the SIMO Channel

Theorem 4.2.6 *In a SIMO quasi-static frequency-selective channel with memory ν, N receive antennas and data-block length L_d, the MMSE receiver diversity is $d_{MMSE}^{CP} = N\min(\nu + 1, \lfloor 2^{-R}L_d \rfloor + 1)$ under joint spatial encoding and cyclic prefix transmission.*

In order to prove Theorem 4.2.6, we first analyze the case of $L_d = \nu + 1$ and then generalize the result for $L_d > \nu + 1$. The system model is given by (4.6) where the $NL_d \times L_d$ equivalent channel matrix is given by

$$\mathbf{H_e} = \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_\nu & 0 & \cdots & 0 \\ 0 & \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_\nu & 0 & \cdots & \mathbf{h}_0 \end{bmatrix}. \tag{4.45}$$

where $\mathbf{h}_i$ (for $i = 0, 1, \ldots, \nu$) is $N \times 1$ SIMO channel. Note that the diagonal elements of $(\mathbf{H_e}^H \mathbf{H_e})$ are identical and equal to $\sum_{i=0}^{\nu} \mathbf{h}_i^H \mathbf{h}_i$. Thus the MMSE SINR for each output information stream is

$$\gamma_k = \frac{1}{(\mathbf{I} + \mathbf{H_e}^H \mathbf{H_e})_{kk}} - 1 = \frac{1}{\frac{1}{L_d} \mathrm{tr}(\mathbf{I} + \mathbf{H_e}^H \mathbf{H_e})_{kk}} - 1 \tag{4.46}$$

Evaluating the outage probability as in (4.31)

$$\begin{aligned} P_{out} &= \mathbb{P}\left(\frac{1}{L_d} \sum_{k=1}^{L_d} \log\left(\frac{1}{(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1}} \right) < R \right) \\ &= \mathbb{P}\left(\log \frac{1}{L_d} \sum_{k=1}^{L_d} \frac{1}{(\mathbf{I} + \rho \mathbf{H_e}^H \mathbf{H_e})_{kk}^{-1}} < R \right) \tag{4.47} \\ &= \mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{(1 + \rho \lambda_k)} > L_d 2^{-R} \right) \tag{4.48} \end{aligned}$$

where (4.47) follows from (4.46) and (4.48) follows similarly to (4.34).

In a manner similar to (4.34) we have $\lambda_k = \mathbf{B}_k^H \mathbf{B}_k$ because now $\mathbf{B}$ is simply a $N \times 1$ vector. For the case $L_d = \nu + 1$, the eigenvalues $\{\lambda_k\}$ are distributed according to Gamma distribution with shape parameter N and scale parameter 1, i.e. $\lambda_k \sim \Gamma(N, 1)$. For $L_d > \nu + 1$ the Gaussian variables in $\mathbf{B}_k$ are no longer independent and thus analyzing this case requires the unknown distribution $\{\lambda_k\}$. Instead, we indirectly show that the diversity of $L_d = \nu + 1$ also holds for $L_d > \nu + 1$.

Lemma 4.2.7 *In a SIMO quasi-static frequency-selective channel with memory ν, N receive antennas and data-block length $L_d = \nu + 1$, the MMSE receiver diversity is $d_{MMSE}^{CP} = N(\lfloor L_d 2^{-R} \rfloor + 1)$ under joint spatial encoding and cyclic prefix transmission.*

Proof The outage probability can be written as

$$\begin{aligned} P_{out} &= \mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{(1 + \rho \lambda_k)} > L_d 2^{-R} \right) \\ &\doteq \mathbb{P}\left(M(\alpha) > L_d 2^{-R} \right) \tag{4.49} \end{aligned}$$

where we use $M(\alpha) = \sum_{\alpha_k > 1} 1$ from (4.35). We thus need to evaluate $\mathbb{P}(\alpha > 1)$. The probability density function of λ_k is

$$f_{\lambda_k}(x) = \frac{1}{\Gamma(N)} x^{N-1} e^{-x} \tag{4.50}$$

The distribution of α_k is thus given by

$$f_{\alpha_k}(x) = \frac{1}{\Gamma(N)} \rho^{-Nx} e^{-x} \ln \frac{1}{\rho} \tag{4.51}$$

The cumulative distribution function of α_k is

$$F_{\alpha_k}(x) = \int_{-\infty}^{x} f_{\alpha_k}(y) \, dy$$

$$= \frac{1}{\Gamma(N)} \int_{\rho^{-x}}^{\infty} r^{N-1} e^{-r} dr \tag{4.52}$$

$$= \frac{1}{\Gamma(N)} \left(\int_{0}^{\infty} r^{N-1} e^{-r} dr - \int_{0}^{\rho^{-x}} r^{N-1} e^{-r} dr \right) \tag{4.53}$$

$$= e^{-\rho^{-\alpha}} \sum_{k=0}^{N-1} \frac{\rho^{-xk}}{k!} \tag{4.54}$$

where we have made a change of variables $r = \rho^{-x}$ in (4.52), and evaluate the integral according to [27, P.334 and P.336]. Thus we have

$$P(\alpha_k > 1) = 1 - e^{-\rho} \sum_{k=0}^{N-1} \frac{\rho^{-k}}{k!}$$

$$\doteq 1 - \left(1 - \frac{1}{N!} \rho^{-N} \right) \tag{4.55}$$

$$\doteq \rho^{-N} \tag{4.56}$$

where (4.55) follows from the Taylor expansion for (4.54).

From the independence of $\{\lambda_k\}$, and subsequently the independence of $\{\alpha_k\}$, we conclude that $M(\alpha)$ in (4.49) is binomially distributed with parameter ρ^{-N}. Hence, similar to [15],

we have

$$\mathbb{P}\left(\sum_{k=1}^{L_d} \frac{1}{1 + \rho\lambda_k} > L_d 2^{-R}\right) \doteq \mathbb{P}(M(\alpha) > L_d 2^{-R})$$

$$= \sum_{i=\lfloor L_d 2^{-R}\rfloor+1}^{L_d} \mathbb{P}(M(\alpha) = i)$$

$$\doteq \sum_{i=\lfloor L_d 2^{-R}\rfloor+1}^{L_d} \binom{L_d}{i} \rho^{-Ni} \underbrace{(1 - \rho^{-N})^{n-i}}_{\doteq 1}$$

$$\doteq \rho^{-N(\lfloor L_d 2^{-R}\rfloor+1)}.$$

which concludes the proof for $L_d = \nu + 1$

For $L_d > \nu + 1$ we follow steps similar to [15].

Lemma 4.2.8 *Consider two SIMO systems both operating under quasi-static frequency-selective channels with memory ν. One system has data block length $L_{d_1} > \nu + 1$ and the other $L_{d_2} \geq L_{d_1}$, we have the following property*

$$\mathbb{P}\left(\sum_{k=1}^{L_{d_1}} \frac{1}{(1 + \rho\lambda_k)} > m\right) \doteq \mathbb{P}\left(\sum_{k=1}^{L_{d_2}} \frac{1}{(1 + \rho\lambda_k)} > m\right)$$

for any $m \in \mathbb{R}$.

Proof The proof has similarities with the SISO case developed in [15, Lemma 2], but is not a trivial extension (see Appendix 4.3.3).

Using Lemma 4.2.8 and the results in [15, Theorem 2], Theorem 4.2.6 is established.

4.3 Appendix

4.3.1 Proof of Lemma 4.2.1

Consider a single-antenna ISI channel $\mathbf{h} = [h_0, \ldots, h_\nu]$, where ν is channel memory. The transmitter sends a block of $L_d + \nu$ symbols (i.e. the extension $L_e = \nu$), the last ν symbols

of which are zeros to remove the inter-block interference. The system model is given by

$$\mathbf{y} = \mathbf{H_e}\mathbf{x} + \mathbf{n} \tag{4.57}$$

where $\mathbf{x}$ is the transmitted length-$(L_d + \nu)$ vector. We consider the case where the padding length is equal to the memory of the channel. The results are also valid for $L_e > \nu$ as a direct result of [15, Theorem 2].

The outage probability of MMSE receiver under ZP transmission is given by [15]

$$P_{out} = \mathbb{P}\Big(\frac{1}{L_d}\sum_{k=1}^{L_d}\log\big(\frac{1}{(\mathbf{I} + \rho\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}}\big) < R\Big)$$

$$\leqslant \mathbb{P}\Big(\frac{1}{L_d}\sum_{k=1}^{L_d}\log(1 + \frac{\rho}{(\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1}}) < R\Big) \tag{4.58}$$

$$\leqslant \mathbb{P}\Big(\log\frac{1}{L_d}\sum_{k=1}^{L_d}\frac{1}{\rho}(\mathbf{H_e}^H\mathbf{H_e})_{kk}^{-1} > -R\Big) \tag{4.59}$$

$$= \mathbb{P}\Big(\frac{L_d\,2^{-R}}{\mathrm{tr}(\mathbf{H_e}^H\mathbf{H_e})^{-1}} < \rho^{-1}\Big) \tag{4.60}$$

where (4.58) represents the outage probability of zero-forcing equalizer which upper bounds that of the MMSE. The bound in (4.59) follows from Jensen's inequality.

We want to show that $\mathrm{tr}(\mathbf{H_e}^H\mathbf{H_e})^{-1}$ in (4.60) is proportional to $||\mathbf{h}||^{-2}$. Thus it is straightforward to obtain full-diversity at any R since [2]

$$\mathbb{P}\big(c\,||\mathbf{h}||^2 < \rho^{-\alpha}\big)\dot{=}\rho^{-L\alpha} \tag{4.61}$$

where c is a constant that is independent of $\mathbf{h}$.

To show that this is indeed the case, we use the result of Tepedelenlioglu [25, 31] which provides a family of linear zero-forcing equalizers that is capable of achieving full multipath diversity in zero-padded systems under certain constraints. We paraphrase the result for convenience.

Lemma 4.3.1 ([25, 31]) *Under zero-padded transmission, there exists a family of left-inverses of $\mathbf{H_e}$, denoted by $\mathbf{G}$, such that $||\mathbf{G}||^{-2} \geqslant C||\mathbf{h}||^2$ for some constant C independent*

of the channel vector $\mathbf{h}$. *Moreover, we have* $||\mathbf{W}_{ZF}|| \leqslant ||\mathbf{G}||$, *for any* $\mathbf{G}$ *satisfying* $\mathbf{GH_e} = \mathbf{I}$, *and* $\mathbf{W}_{ZF}$ *is given by*

$$\mathbf{W}_{ZF} = (\mathbf{H_e}^H \mathbf{H_e})^{-1} \mathbf{H_e}^H. \tag{4.62}$$

Applying the ZF equalizer $\mathbf{W}_{ZF}$ on the channel output given by (4.57) we get the equalized signal $\tilde{\mathbf{y}} = \mathbf{x} + \mathbf{z}$, where $\mathbf{z} = \mathbf{W}_{ZF}\mathbf{n}$. The filtered noise power P_z can be evaluated as

$$\begin{aligned}
P_z &= \mathbb{E}\operatorname{tr}[\mathbf{z}\mathbf{z}^H] \\
&= \operatorname{tr}\left[\mathbb{E}((\mathbf{H_e}^H\mathbf{H_e})^{-1}\mathbf{H_e}^H\mathbf{n}\mathbf{n}^H\mathbf{H_e}(\mathbf{H_e}^H\mathbf{H_e})^{-1})\right] \\
&= \operatorname{tr}[(\mathbf{H_e}^H\mathbf{H_e})^{-1}] \tag{4.63}
\end{aligned}$$

where we assume the noise is uncorrelated and has variance equal to one.

Using the properties of the Frobenius norm, P_z can be bounded as

$$\begin{aligned}
P_z &= \mathbb{E}(||\mathbf{W}_{zf}\mathbf{n}||^2) \\
&\leqslant \mathbb{E}(||\mathbf{W}_{zf}||^2||\mathbf{n}||^2) = L_d||\mathbf{W}_{zf}||^2. \tag{4.64}
\end{aligned}$$

Using (4.63), (4.64) and Lemma 4.3.1, the trace in (4.60) can be bounded by

$$\operatorname{tr}[(\mathbf{H_e}^H\mathbf{H_e})^{-1}] \leqslant L_d||\mathbf{W}_{zf}||^2 \leqslant \frac{L_d}{C\,||\mathbf{h}||^2}. \tag{4.65}$$

Thus from (4.60) we have

$$\begin{aligned}
P_{out} &\dot{\leqslant} \mathbb{P}\left(C_2||\mathbf{h}||^2 < \rho^{-1}\right) \\
&\doteq \rho^{-(\nu+1)}. \tag{4.66}
\end{aligned}$$

where $C_2 = C\,2^{-R}$ is a constant independent of $\mathbf{h}$ and ρ.

Note that the constraints and construction methods in [25, 31] for the zero-forcing equalizers to achieve full multipath diversity in ZP systems do not apply in CP systems. That is, Lemma 4.3.1 is not true for CP transmission. This is because the equivalent channel in CP systems does not have the same properties that were used in [25, 31].

4.3.2 Proof of Lemma 4.2.5:(QIP Problem)

Consider the following Quadratic Integer Programming (QIP) problem

$$\min_{n_1,n_2,\ldots,n_\ell} \quad \sum_{k=1}^{\ell} n_k^2 \tag{4.67}$$

$$\text{subject to} \quad \sum_{k=1}^{\ell} n_k = \Omega$$

$$n_k \geq 0.$$

where Ω and ℓ are integers.

Consider a candidate solution vector $[n_1, \ldots, n_k, \ldots, n_\ell]$. We partition the variables in this vector according to their values into $\Omega + 1$ sets $\mathcal{N}_j = \{n_k : n_k = j\}$ for $0 \leq j \leq \Omega$; clearly some of these sets may be empty. Denote the membership of each set $S_j = |\mathcal{N}_j|$. Furthermore, let $\Omega = m\ell + K$ where m is the divisor and K is the remainder of the division of Ω by ℓ. From the constraint in (4.67) we have

$$\sum_{k=1}^{\ell} n_k = \sum_{j=0}^{\Omega} j S_j = m\ell + \sum_{j=0}^{\Omega} (j-m) S_j = m\ell + K. \tag{4.68}$$

Evaluating the objective function:

$$\sum_{k=1}^{\ell} n_k^2 = \sum_{j=0}^{\Omega} (m + j - m)^2 S_j$$

$$= \ell m^2 + 2m \sum_{j=0}^{\Omega} (j-m) S_j + \sum_{j=0}^{\Omega} (j-m)^2 S_j$$

$$= \ell m^2 + 2mK + \sum_{j=0}^{\Omega} (j-m)^2 S_j \tag{4.69}$$

$$\geqslant \ell m^2 + 2mK + \sum_{j=0}^{\Omega} (j-m) S_j \tag{4.70}$$

$$= \ell m^2 + 2mK + K \tag{4.71}$$

where (4.69) and (4.71) use $\sum_{j=0}^{\Omega} (j-m) S_j = K$, which follows from (4.68).

We now propose that one may achieve optimality when all variables take values either m or $m + 1$. In that case,

$$\sum_k n_k = mS_m + (m+1)(\ell - S_m) = m\ell + (\ell - S_m)$$

$$\sum_k n_k^2 = m^2 S_m + (m+1)^2(\ell - S_m) = \ell m^2 + 2mK + K.$$

where we substituted the value of $\ell - S_m$ from the first equation into the second equation above. This shows that the variables taking values m or $m + 1$ achieves the lower bound in (4.71). At optimality $S_m = (m+1)\ell - \Omega$.

4.3.3 Proof of Lemma 4.2.8

We begin by showing that for any integer multiplier of $L_{d_1} = \nu + 1$ denoted by $L_{d_2} = TL_{d_1}$ ($T \in \mathbb{N}$) and any real-valued $m \in (0, L_{d_1})$, we have

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_1}} \frac{1}{(1 + \rho\lambda_q)} > m \right) \doteq \mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1 + \rho\lambda_q)} > m \right) \tag{4.72}$$

Note that for SIMO-CP system, $\lambda_q = \mathbf{b}_q^H \mathbf{b}_q$, where $\mathbf{b}_q$ is the $N \times 1$ vector given by

$$\mathbf{b}_q^{(i)} = \sum_{n=0}^{\nu} \mathbf{h}_n \, e^{-j\frac{2\pi(q-1)}{L_{d_i}}} \quad \text{for } q = 1, \ldots, L_{d_i} \tag{4.73}$$

where $\mathbf{h}_n$ is the channel gain as a function of the tap index n, and the superscript $i = 1, 2$ is used to distinguish the variables in two systems with data block lengths L_{d_1} and L_{d_2}.

Recall that we can take a L_{d_1}-point signal and apply a L_{d_2}-point DFT on it (after zero-padding), which will result in a resampling in the Fourier domain at L_{d_2} points. Following [15] we can write the explicit relationship between entries of $\mathbf{b}^{(1)}$ and $\mathbf{b}^{(2)}$ as

$$b_{q,l}^{(1)} = \sum_{i=1}^{L_{d_1}} b_{i,l}^{(2)} \psi_i \quad q = 1, 2 \ldots, L_{d_2} \text{ and } l = 1, 2, \ldots, N. \tag{4.74}$$

where

$$\psi_i = \frac{1}{L_{d_1}} \frac{1 - e^{-j\frac{(q-1)2\pi L_{d_1}}{L_{d_2}}}}{1 - e^{-j\left(\frac{2\pi(q-1)}{L_{d_2}} - \frac{2\pi(i-1)}{L_{d_1}}\right)}}.$$

Define $\alpha_{q,l}^{(i)} = \frac{\log |b_{q,l}^{(i)}|^2}{\log \rho}$. Note that $b_{T(q-1),l}^{(1)} = b_{q,l}^{(2)}$ and $\alpha_{T(q-1),l}^{(2)} = \alpha_{q,l}^{(1)}$ for $q = 1, 2 \ldots, L_{d_1}$ since $L_{d_2} = TL_{d_1}$. From (4.74), we have

$$|b_{q,l}^{(1)}|^2 = \sum_{i=1}^{L_{d_1}} |\psi_i|^2 |b_{i,l}^{(2)}|^2 + \underbrace{\sum_{i=1}^{L_{d_1}} \sum_{s=1}^{L_{d_1}} \psi_i \psi_s b_{i,l}^{(2)} b_{s,l}^{*(2)}}_{\triangleq \eta}. \tag{4.75}$$

We now analyze each part of the sum in (4.75). For the set of indices $\mathcal{A} \triangleq \{i : i = T(k-1)+1, k = 1, \ldots, L_{d_1}\}$, the coefficients $\{\psi_i\}$ are non-zero constants, then $|\psi_i|^2 |b_{i,l}^{(2)}|^2 \doteq |b_{i,l}^{(2)}|^2 \quad \forall l$. Noting that η must be real-valued, and defining $\alpha_\eta \triangleq -\frac{\log |\eta|}{\log \rho}$, Eq. (4.75) can be written as

$$\rho^{-\alpha_{q,l}^{(2)}} \doteq \sum_{i=1}^{L_{d_1}} \rho^{-\alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta}$$

$$\doteq \rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta}. \tag{4.76}$$

Note that if $\eta < 0$ the second term in (4.76) should be smaller than the first term since otherwise the right-hand side of (4.76) will be negative while the left-hand side is positive. Thus for $\eta < 0$ we have $\alpha_\eta \geqslant \min_i \alpha_{i,l}^{(1)}$. Also, for $a \geqslant 0$ we have $\rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta} \dot{\geqslant} \rho^{-\min_i \alpha_{i,l}^{(1)}}$. Thus we always have $\rho^{-\min_i \alpha_{i,l}^{(1)}} + \frac{\eta}{|\eta|} \rho^{-\alpha_\eta} \dot{\geqslant} \rho^{-\min_i \alpha_{i,l}^{(1)}}$, leading to the following lemma.

Lemma 4.3.2 *For $\alpha_{q,l}^{(1)}$ and $\alpha_{q,l}^{(2)}$ defined above we have: $\rho^{-\alpha_{q,l}^{(2)}} \dot{\geqslant} \rho^{-\min_i \alpha_{i,l}^{(1)}} \Rightarrow \alpha_{q,l}^{(2)} \leqslant \min_i \alpha_{i,l}^{(1)}$ for $q \in \mathcal{A}$.*

We now partition the DFT points into two sets $\mathcal{A} = \{T(i-1)+1, i = 1, \ldots, L_{d_1}\}$ and $\mathcal{B} = \{1, \ldots, L_{d_2}\} \backslash \{T(i-1)+1, i = 1, \ldots, L_{d_1}\}$ We now define the event:

$$\mathcal{D} \triangleq \{\min_i \alpha_{i,1}^{(1)} < 1 \,, \, \min_i \alpha_{i,2}^{(1)} < 1 \,, \, \ldots \,, \, \min_i \alpha_{i,N}^{(1)} < 1\}$$

and proceed to evaluate the probability

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) = \mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{1+\rho\sum_{l=1}^{N}|b_{q,l}^{(1)}|^2} > m\right) \tag{4.77}$$

$$= \mathbb{P}\left(\sum_{q\in\mathcal{A}} \frac{1}{1+\rho\sum_{l=1}^{N}|b_{q,l}^{(1)}|^2} + \sum_{q\in\mathcal{B}} \frac{1}{1+\rho\sum_{l=1}^{N}|b_{q,l}^{(1)}|^2} > m\right)$$

$$\doteq \mathbb{P}\left(S_1 + S_2 > m\right) \tag{4.78}$$

where (4.77) follows since $\lambda_q = \mathbf{b}_q^H \mathbf{b}_q$ and S_1 and S_2 are given by

$$S_1 \triangleq \sum_{q=1}^{L_{d_1}} \frac{1}{1+\sum_{l=1}^{N}\rho^{1-\alpha_{q,l}^{(1)}}}$$

$$S_2 \triangleq \sum_{q\in\mathcal{B}} \frac{1}{1+\sum_{l=1}^{N}\rho^{1-\alpha_{q,l}^{(2)}}}$$

We now evaluate (4.78)

$$\mathbb{P}\left(S_1 + S_2 > m\right) = \mathbb{P}\left(S_1 + S_2 > m \,\middle|\, \mathcal{D}\right) \times \mathbb{P}(\mathcal{D}) + \mathbb{P}\left(S_1 + S_2 > m \,\middle|\, \bar{\mathcal{D}}\right) \times \mathbb{P}(\bar{\mathcal{D}}) \tag{4.79}$$

Note that subject to the event $\mathcal{D}$, we have

$$S_2 = \sum_{q\in\mathcal{B}} \frac{1}{1+\sum_{l=1}^{N}\rho^{1-\alpha_{q,l}^{(2)}}} \doteq 0$$

Therefore this term can be asymptotically ignored. Also subject to $\bar{\mathcal{D}}$, we have

$$S_1 = \sum_{q=1}^{L_{d_1}} \frac{1}{1+\sum_{l=1}^{N}\rho^{1-\alpha_{q,l}^{(1)}}} \doteq L_{d_1}$$

and since with probability one, $L_{d_1} \geq m$, the other (non-negative) term can be asymptotically ignored. Thus, both the terms involving the set $\mathcal{B}$ can be altogether ignored and we have:

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) \doteq \mathbb{P}\left(S_1 > m|\mathcal{D}\right)\mathbb{P}(\mathcal{D}) + \mathbb{P}\left(S_1 > m|\bar{\mathcal{D}}\right)\mathbb{P}(\bar{\mathcal{D}})$$

$$\doteq \mathbb{P}\left(\sum_{q=1}^{L_{d_1}} \frac{1}{(1+\rho\lambda_q)} > m\right)$$

We have thus established (4.72) when $L_{d_1}|L_{d_2}$. We must now show that the same result holds for any T' when $L_{d_1} \nmid T'$. To do so, let $L_{d_2} = T'L_{d_1}$, then we have

$$\mathbb{P}\left(\sum_{q=1}^{L_{d_2}} \frac{1}{(1+\rho\lambda_q)} > m\right) \doteq \mathbb{P}\left(\sum_{q=1}^{T'} \frac{1}{(1+\rho\lambda_q)} > m\right). \tag{4.80}$$

Using (4.72) when $L_{d_1}|L_{d_2}$ and (4.80) when $T'|L_{d_2}$ together establishes (4.72) for any two positive integers.

CHAPTER 5
MIMO LINEAR PRECODING

5.1 Introduction

Precoding is a preprocessing technique that exploits channel-state information at the transmitter (CSIT) to match the transmission to the instantaneous channel conditions [4, 24, 32]. Linear and non-linear precoding designs are available in the literature [33]. Linear precoding in particular provides a simple and efficient method to utilize CSIT. Linear precoding has been shown to be optimal in certain situations involving partial CSIT [34, 35], however, in many instances the main motivation of linear precoders is to simplify the MIMO receiver.

Linear precoders include zero-forcing (ZF), matched filtering (MF), Wiener filtering, and regularized zero-forcing (RZF). The ZF precoding schemes were extensively studied in multiuser systems as the ZF decouples the multiuser channel into independent single-user channels and has been shown to achieve a large portion of dirty paper coding capacity [36]. ZF precoding often involves *channel inversion*, using the pseudo-inverse of the channel or other generalized inverses [33]. Matched filter (MF) precoding [37], similarly to the MF receiver, is interference limited at high SNR but it outperforms the ZF precoder at low SNR [33]. The regularized ZF precoder, as the name implies, introduces a regularization parameter in channel inversion. If the regularization parameter is inversely proportional to SNR, the RZF of [38] is identical to the Wiener filter precoding [33]. Peel et al. [38] introduce a vector perturbation technique to reduce the transmit power of the RZF method, showing that in this way RZF can operate near channel capacity.

This chapter analyzes the diversity of MIMO linear precoding, with or without linear receivers, under flat fading. We show that in a $M \times N$ MIMO channel with $M \geq N$, the ZF precoder has diversity $M - N + 1$. We show that Wiener precoders produce a diversity that
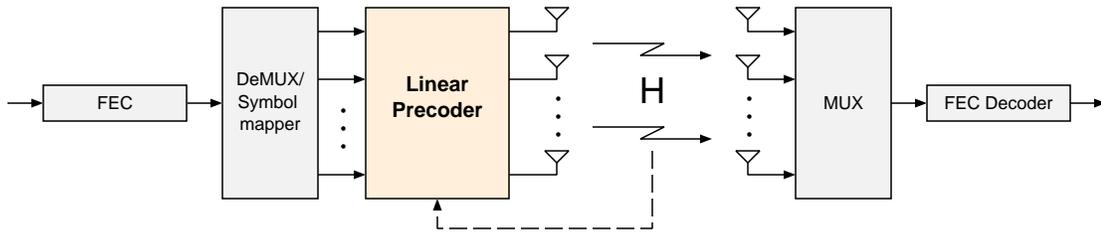
Figure 5.1. MIMO with linear precoder

is a function of spectral efficiency as well as the number of transmit and receive antennas. At very low rates, the Wiener precoder enjoys diversity MN, while at very high rates it achieves diversity $M - N + 1$. These results are reminiscent of MIMO linear equalizers [5], even though in general the behavior of equalizers (receive side) can be very different from precoders (transmit side) and the analysis does not carry from one to the other. We also show that MIMO systems with RZF and MF precoders (together with optimal receivers) exhibit a new kind of rate-dependent diversity that has not to date been observed or reported, i.e., they either have full diversity or zero diversity (error floor) depending on the operating spectral efficiency R.

We also provide DMT analysis for all precoders mentioned above. The fact that DMT and the diversity under fixed-rate regime require separate analyses has been established for MIMO linear equalizers [5, 6] and is explained in Chapter 2 and Chapter 3. Essentially, the reason is that various fixed rates (spectral efficiencies) for MIMO precoding result in distinctly different diversities, whereas DMT analysis assigns only a single value of diversity to all fixed rates (all fixed rates correspond to multiplexing gain zero).

This chapter is organized as follows. Section 5.2 describes the system model. Section 5.3 provides outage analysis of many precoded MIMO systems. Section 5.4 provides the DMT analysis. Section 3.6 provides simulations that illuminate our results.

5.2   System Model

A MIMO system with linear precoding is depicted in Fig. 5.1. This system uses the linear precoder to manage the interference between the streams in a MIMO system to avoid a requirement of optimal joint decoding in the receiver, which is costly. We consider a flat fading channel $\mathbf{H} \in \mathbb{C}^{N \times M}$, where M and N are the number of transmit and receive antennas, respectively. While $M \geqslant N$ when using linear precoding alone, we have $N \geqslant M$ or $M \geqslant N$ when using precoding together with receive-side linear equalization depending on whether the precoder is designed for the equalized channel or the equalizer is designed for the precoded channel (see Figure 6.1). The input-output system model for flat fading MIMO precoded channel with M transmit and N receive antennas is given by

$$\mathbf{y} = \mathbf{HTx} + \mathbf{n}$$

where $\mathbf{T} \in \mathbb{C}^{M \times B}$ is the precoder matrix. In Chapter 6, we will consider the joint effect of precoding and equalization, where the system model will be

$$\mathbf{y} = \mathbf{WHTx} + \mathbf{Wn} \tag{5.1}$$

where $\mathbf{W} \in \mathbb{C}^{B \times N}$ is the receiver side equalizer. The number of information symbols is $B \leqslant \min(M, N)$, the transmitted vector is $\mathbf{x} \in \mathbb{C}^{B \times 1}$, and $\mathbf{n} \in \mathbb{C}^{N \times 1}$ is the Gaussian noise vector. The vectors $\mathbf{x}$ and $\mathbf{n}$ are assumed independent.

We aim to characterize the diversity gain, $d(R, M, N)$, as a function of the spectral efficiency R (bits/sec/Hz) and the number of transmit and receive antennas. This requires a Pairwise Error Probability (PEP) analysis which is not directly tractable. Instead, we find the exponential order of outage probability and then demonstrate that outage and PEP exhibit identical exponential orders.

The objective of linear precoding/equalization is to transform the MIMO channel into $\min(M, N)$ parallel channels that can be described by

$$y_k = \sqrt{\gamma_k} x_k + \tilde{n}_k, \quad k = 1, \dots, B \tag{5.2}$$

where γ_k is the SINR at the k-th receiver output and $B = min(M, N)$, and $\tilde{n}_k$ are the decision point noise coefficients.

The outage probabilities of MIMO systems under joint spatial encoding is respectively given by [6, 7]

$$P_{\text{out}} \triangleq \mathbb{P}\left(\sum_{k=1}^{B} \log(1 + \gamma_k) \leqslant R \right) \tag{5.3}$$

We shall perform outage analysis for different precoders/equalizers as the first step towards deriving the diversity function. We then provide lower and upper bounds on error probability via outage probabilities. This two-step approach was first proposed in [2] due to the intractability of the direct PEP analysis for many MIMO architectures.

We denote the *exponential equality* of two functions $f(\rho)$ and $g(\rho)$ as $f(\rho) \doteq g(\rho)$ when

$$\lim_{\rho \to \infty} \frac{\log f(\rho)}{\log(\rho)} = \lim_{\rho \to \infty} \frac{\log g(\rho)}{\log(\rho)}$$

The exponential inequalities $\dot{\geqslant}$ and $\dot{\leqslant}$ are defined in a similar manner. In the following, we shall need to specify various upper and lower bounds or approximations of the SINR γ, which will give rise to a number of variables such as $\hat{\gamma}$, $\check{\gamma}$, and $\bar{\gamma}$.

5.3 Precoding Diversity

In this section we analyze a linearly precoded MIMO system where $M \geq N$ and the number of data streams B is equal to N. For the purposes of the developments in this section, there is no receive-side equalization.

5.3.1 Zero-Forcing Precoding

The ZF precoder completely eliminates the interference at the receiver. ZF precoding is well studied in the literature via performance measures such as throughput and fairness under a total (or per antenna) power constraint [39, and references therein].

Design Method I

One approach to design the ZF precoder is to solve the following problem [33]

$$\mathbf{T} = \arg \min_{\mathbf{T}} \mathbb{E}\big[||\mathbf{Tx}||_2^2\big] \tag{5.4}$$

$$\text{subject to } \mathbf{HT} = \mathbf{I}$$

The resulting ZF transmit filter is given by

$$\mathbf{T} = \beta\, \mathbf{H}^H(\mathbf{HH}^H)^{-1} \in \mathbb{C}^{M \times N} \tag{5.5}$$

where β is a scaling factor to satisfy the transmit power constraint, that is [33]

$$\beta^2 \mathrm{tr}\big(\mathbf{TT}^H\big) \leqslant \rho \tag{5.6}$$

where we assume that the noise power is one and that the information streams are independent. From (5.6), the received SINR per stream is thus given by

$$\gamma_k^{ZFP} = \frac{\rho}{\mathrm{tr}((\mathbf{HH}^H)^{-1})}. \tag{5.7}$$

Using (5.3), the outage probability is given by

$$P_{\text{out}} = \mathbb{P}\left(N \log \left(1 + \frac{\rho}{\mathrm{tr}((\mathbf{HH}^H)^{-1})}\right) \leqslant R \right) \tag{5.8}$$

A direct evaluation of (5.8) is intractable since the diagonal elements of $(\mathbf{HH}^H)^{-1}$ are distributed according to the inverse-chi-square distribution [6, 40]. We instead bound (5.8) from below and above and show that the two bounds match asymptotically.

Let $\{\lambda_k\}$ be the eigenvalues of $\mathbf{HH}^H$. Equation (5.8) can be written as

$$P_{\text{out}} = \mathbb{P}\left(N \log \left(1 + \frac{\rho}{\sum_{k=1}^N \frac{1}{\lambda_k}}\right) \leqslant R \right)$$

which can be bounded as

$$P_{\text{out}} \leqslant \mathbb{P}\left(N \log(1 + \frac{\rho}{N}\lambda_{\min}) \leqslant R \right) \tag{5.9}$$

$$= \mathbb{P}\left(\lambda_{\min} \leqslant N(2^{\frac{R}{N}} - 1)R\rho^{-1} \right)$$

$$\doteq \mathbb{P}\big(\lambda_{\min} \leqslant \rho^{-1}\big). \tag{5.10}$$

The marginal probability density of $\lambda_{\min}$ is approximately proportional to: $f_1(\lambda) \propto \lambda^{(M-N)}$ whenever $\lambda \ll 1$ [7,17]. Therefore the bound in (5.10) can be evaluated to yield:

$$P_{\text{out}} \;\dot{\leqslant}\; \rho^{-(M-N+1)}. \tag{5.11}$$

We now proceed with a lower bound on outage. The outage probability in (5.8) can be bounded:

$$
\begin{aligned}
P_{\text{out}} &= \mathbb{P}\left(N \log(1 + \frac{\rho}{\text{tr}(\mathbf{HH}^H)^{-1}}) \leqslant R \right) \\
&\geqslant \mathbb{P}\left(N \log(1 + \frac{\rho}{(\mathbf{HH}^H)_{kk}^{-1}}) \leqslant R \right) \\
&\doteq \mathbb{P}\left(z \leqslant \rho^{-1} \right)
\end{aligned}
\tag{5.12}
$$

where we have made a change of variable $z = \frac{1}{(\mathbf{HH}^H)_{kk}^{-1}}$.

The random variable z in (5.12) is distributed according to the chi-square distribution with $2(M - N + 1)$ degree of freedom, i.e. $z \sim \chi^2_{2(M-N+1)}$ [40]. Thus the bound in (5.12) can be evaluated [6] yielding:

$$P_{\text{out}} \;\dot{\geqslant}\; \rho^{-(M-N+1)}. \tag{5.13}$$

From (5.11) and (5.13), we conclude that the diversity of MIMO system using the ZF precoder given by (5.4) and joint spatial encoding is

$$d^{ZFP} = M - N + 1. \tag{5.14}$$

5.3.2 Zero-Forcing Precoding: Design Method II

Notice that the ZF precoder design in (5.4) minimizes the transmitted power. Another approach for ZF precoding design allocates unequal power levels across the transmit antennas to optimize some performance measure. For instance, consider the optimization problem [39]

$$
\begin{aligned}
\max_{p_k, \mathbf{T}} \quad & \sum_k \log(1 + \gamma_k^{ZFP}) \\
\text{subject to} \quad & \mathbf{HT} = \text{diag}\{\sqrt{p_1}, \ldots, \sqrt{p_N}\} \\
& \mathbb{E}||\mathbf{Tx}||^2 \leqslant \rho
\end{aligned}
\tag{5.15}
$$

where p_k is the transmit power for stream k. The optimal solution for (5.15) (assuming independent transmit signaling) has the following form [39, Theorem 1]:

$$\mathbf{T} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}\text{diag}\{\sqrt{p_1}, \ldots, \sqrt{p_N}\} \tag{5.16}$$

where p_k are the solution to:

$$\max_{p_k} \sum_k \log(1 + \gamma_k^{ZFP})$$

$$\text{subject to} \quad \sum_k p_k\big[(\mathbf{H}\mathbf{H}^H)^{-1}\big]_{kk} \leqslant \rho \tag{5.17}$$

Due to the logarithmic form of the cost function, the solution has the familiar form of waterfilling. It is well-known that water-filling may drive some p_k to zero. Depending on the value of ρ and realization of $\mathbf{H}\mathbf{H}^H$, it may also happen that all optimal p_k are positive. The set of realizations of $\mathbf{H}\mathbf{H}^H$ that satisfy this condition are collected into an event that we denote $\mathcal{P}$. Conditioned on $\mathcal{P}$ it is easy to verify that the optimal solution is given by:

$$p_k = \frac{\rho + \sum_{k=1}^{N}(\mathbf{H}\mathbf{H}^H)_{kk}^{-1}}{N(\mathbf{H}\mathbf{H}^H)_{kk}^{-1}} - 1 \qquad k = 1, \ldots, N \tag{5.18}$$

The outage probability can then be evaluated as follows

$$P_{\text{out}} = \mathbb{P}\left(\log \sum_{k=1}^{N}(p_k + 1) < R\right)$$

$$= \mathbb{P}\left(\log \sum_{k=1}^{N}(p_k + 1) < R\Big|\mathcal{P}\right)\mathbb{P}(\mathcal{P}) + \mathbb{P}\left(\log \sum_{k=1}^{N}(p_k + 1) < R\Big|\bar{\mathcal{P}}\right)\mathbb{P}(\bar{\mathcal{P}}) \tag{5.19}$$

We will now calculate $\mathbb{P}(\bar{\mathcal{P}})$. Using (5.18), we have

$$\mathbb{P}(\bar{\mathcal{P}}) = \mathbb{P}\left(N(\mathbf{H}\mathbf{H}^H)_{kk}^{-1} - \sum_{k=1}^{N}(\mathbf{H}\mathbf{H}^H)_{kk}^{-1} > \rho\right)$$

$$\leqslant \mathbb{P}\left(N(\mathbf{H}\mathbf{H}^H)_{kk}^{-1} > \rho\right)$$

$$\leqslant \mathbb{P}\left(N\lambda_{\text{max}}(\mathbf{H}\mathbf{H}^H)^{-1} > \rho\right)$$

$$= \mathbb{P}\left(\lambda_{\text{min}}(\mathbf{H}\mathbf{H}^H) < N\rho^{-1}\right)$$

$$\doteq \rho^{-(M-N+1)}. \tag{5.20}$$

where (5.20) is again due to the marginal distribution of $\lambda_{\min}$ via the method of [7].

We now bound other terms of (5.19).

$$\mathbb{P}\left(\sum_{k=1}^{N}\log(\frac{\rho + \sum_{k=1}^{N}(\mathbf{HH}^H)_{kk}^{-1}}{M(\mathbf{HH}^H)_{kk}^{-1}}) \leqslant R\right) \leqslant \mathbb{P}\left(\sum_{k=1}^{N}\log(\frac{\rho}{M(\mathbf{HH}^H)_{kk}^{-1}}) \leqslant R\right) \tag{5.21}$$

$$= \mathbb{P}\left(\sum_{k=1}^{N}\log(\frac{M(\mathbf{HH}^H)_{kk}^{-1}}{\rho}) \geqslant -R\right) \tag{5.22}$$

$$\leqslant \mathbb{P}\left(N\log\sum_{k=1}^{N}(\frac{(\mathbf{HH}^H)_{kk}^{-1}}{\rho}) \geqslant -R\right) \tag{5.23}$$

$$= \mathbb{P}\left(\sum_{k=1}^{N}\frac{1}{\rho\lambda_k} \geqslant 2^{-\frac{R}{N}}\right)$$

$$\leqslant \mathbb{P}\left(\frac{1}{\rho\lambda_{\min}} \geqslant \frac{1}{N}2^{-\frac{R}{N}}\right)$$

$$\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) \tag{5.24}$$

$$\doteq \rho^{-(M-N+1)}, \tag{5.25}$$

where (5.21) holds by discarding the positive element $\sum_{k=1}^{N}(\mathbf{HH}^H)_{kk}^{-1}$. Equation (5.23) follows from Jensen's inequality, and the transition from (5.24) to (5.25) is again due to the marginal distribution of $\lambda_{\min}$ via the method of [7]. Thus

$$P_{\text{out}} = \mathbb{P}\left(\log\sum_{k=1}^{N}(p_k + 1) < R\Big|\bar{\mathcal{P}}\right)\mathbb{P}(\bar{\mathcal{P}}) + \mathbb{P}\left(\log\sum_{k=1}^{N}(p_k + 1) < R\Big|\mathcal{P}\right)\mathbb{P}(\mathcal{P})$$

$$\leqslant \mathbb{P}\left(\log\sum_{k=1}^{N}(p_k + 1) < R\Big|\bar{\mathcal{P}}\right) + \mathbb{P}(\mathcal{P})$$

$$\doteq \rho^{-(M-N+1)}. \tag{5.26}$$

where we have used (5.20) and (5.25) to obtain (5.26).

A lower bound on the outage probability can be given as follows.

$$P_{\text{out}} = \mathbb{P}\bigg(\log \sum_{k=1}^{N}(p_k + 1) < R \Big| \bar{\mathcal{P}} \bigg)\mathbb{P}(\bar{\mathcal{P}}) + \mathbb{P}\bigg(\log \sum_{k=1}^{N}(p_k + 1) < R \Big| \mathcal{P} \bigg)\mathbb{P}(\mathcal{P})$$

$$\geqslant \mathbb{P}\bigg(\log \sum_{k=1}^{N}(p_k + 1) < R \Big| \bar{\mathcal{P}} \bigg)\mathbb{P}(\bar{\mathcal{P}})$$

$$\doteq \mathbb{P}\bigg(\log \sum_{k=1}^{N}(p_k + 1) < R \Big| \bar{\mathcal{P}} \bigg). \tag{5.27}$$

where (5.27) follows since $\mathbb{P}(\bar{\mathcal{P}}) = 1 - \mathbb{P}(\mathcal{P}) \doteq 1$. Thus the outage probability can be bounded as follows

$$P_{\text{out}} \mathrel{\dot{\geqslant}} \mathbb{P}\bigg(\sum_{k=1}^{N} \log(\frac{\rho + \sum_{k=1}^{N}(\mathbf{HH}^H)_{kk}^{-1}}{M(\mathbf{HH}^H)_{kk}^{-1}}) \leqslant R \bigg)$$

$$\geqslant \mathbb{P}\bigg(\sum_{k=1}^{N} \log(\frac{\rho + \frac{N}{\lambda_{\min}}}{M(\mathbf{HH}^H)_{kk}^{-1}}) \leqslant R \bigg)$$

$$\geqslant \mathbb{P}\bigg(N \log \frac{1}{MN} \sum_{k=1}^{N} \frac{\rho + \frac{1}{\lambda_{\min}}}{(\mathbf{HH}^H)_{kk}^{-1}} \leqslant R \bigg). \tag{5.28}$$

The singular value decomposition of $\mathbf{H}$ and the corresponding eigen decomposition of $\mathbf{HH}^H$ are given by

$$\mathbf{H} = \mathbf{U}\,\Gamma\,\mathbf{V}^H$$

$$\mathbf{HH}^H = \mathbf{U}\Lambda\,\mathbf{U}^H$$

where $\mathbf{U} \in \mathbb{C}^{N \times N}$ and $\mathbf{V} \in \mathbb{C}^{M \times M}$ are unitary matrices, $\Gamma \in \mathbb{R}^{N \times M}$ is a rectangular matrix with non-negative real diagonal elements and zero off-diagonal elements, and $\Lambda = \Gamma\Gamma^T \in \mathbb{R}^{N \times N}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{HH}^H$. Let $\mathbf{u}_k$ be the k-th column of $\mathbf{U}^H$. We have

$$(\mathbf{HH}^H)_{kk}^{-1} = \mathbf{u}_k^H \Lambda^{-1} \mathbf{u}_k = \sum_{l=1}^{N} \frac{|u_{kl}|^2}{\lambda_l} \tag{5.29}$$

where u_{kl} is the (k, l) entry of the matrix $\mathbf{U}$.

The bound in (5.28) can be rewritten

$$P_{\text{out}} \dot{\geqslant} \mathbb{P}\left(N \log \frac{1}{MN} \sum_{k=1}^{N} \frac{1}{\sum_{l=1}^{N} \frac{|u_{kl}|^2}{(\rho+\frac{1}{\lambda_{\min}})\lambda_l}} \leqslant R \right)$$

$$\geqslant \mathbb{P}\left(N \log \frac{1}{MN} \sum_{k=1}^{N} \frac{1}{\sum_{l=1}^{N} \frac{|u_{kl}|^2}{1+(\rho+\frac{1}{\lambda_{\min}})\lambda_l}} \leqslant R \right) \tag{5.30}$$

We can bound the sum in the left hand side of (5.30) similarly to the bound in [7, Eq.(18)]

$$\sum_{k=1}^{N} \frac{1}{\sum_{l=1}^{N} \frac{|u_{kl}|^2}{1+(\rho+\frac{1}{\lambda_{\min}})\lambda_l}} \leqslant \left(1 + (\rho + \frac{N}{\lambda_{\min}})\lambda_{\min} \right) \sum_{k=1}^{N} \frac{1}{\sum_{l=1}^{N} |u_{1\ell}|^2}$$

$$= \left(N + \rho\lambda_{\min} \right) \sum_{k=1}^{N} \frac{1}{\sum_{l=1}^{N} |u_{1\ell}|^2}$$

is similar to [7, Eq.(18)], thus the analysis of [7] applies and we obtain

$$P_{\text{out}} \dot{\geqslant} \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1} \right) = \rho^{-(M-N+1)}. \tag{5.31}$$

Thus, the MIMO ZF precoding with unequal power allocation (5.17) achieves diversity order $M - N + 1$.

Recall that the diversity is defined based on the error probability. In Appendix 5.7.1 we provide the pairwise error probability (PEP) analysis for the zero-forcing and regularized zero-forcing precoded systems and show that the outage and error probabilities exhibit same diversity.

5.3.3 Regularized Zero-Forcing Precoding

In general, direct channel inversion performs poorly due to the singular value spread of the channel matrix [38]. One technique often used is to regularize the channel inversion:

$$\mathbf{T} = \beta\, \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + c\, \mathbf{I})^{-1} \tag{5.32}$$

where β is a normalization factor and c is a fixed constant.

Recall

$$\mathbf{y} = \mathbf{HTx} + \mathbf{n} = \beta \mathbf{U}\Lambda(\Lambda + c\,\mathbf{I})^{-1}\mathbf{U}^H\mathbf{x} + \mathbf{n} \tag{5.33}$$

allowing us to decompose the received waveform at each antenna into signal, interference, and noise terms:

$$y_k = \beta\left(\sum_{l=1}^{N}\frac{\lambda_l}{\lambda_l + c}|u_{kl}|^2\right)x_k \ + \ \beta\sum_{i=1,i\neq k}^{N}\left(\sum_{l=1}^{N}\frac{\lambda_l}{\lambda_l + c}u_{kl}u_{il}^*\right)x_i + n_k \tag{5.34}$$

where the scaling factor β is given by $\beta = \frac{1}{\sqrt{\eta}}$ and

$$
\begin{aligned}
\eta &= \mathrm{tr}\big[(\mathbf{HH}^H + c\,\mathbf{I})^{-1}\mathbf{HH}^H(\mathbf{HH}^H + c\,\mathbf{I})^{-1}\big]\\
&= \mathrm{tr}\big[(\mathbf{U}\Lambda\mathbf{U}^H + c\,\mathbf{I})^{-1}\mathbf{U}\Lambda\mathbf{U}^H(\mathbf{U}\Lambda\mathbf{U}^H + c\,\mathbf{I})^{-1}\big]\\
&= \mathrm{tr}\big[\mathbf{U}(\Lambda + c\,\mathbf{I})^{-1}\Lambda(\Lambda + c\,\mathbf{I})^{-1}\mathbf{U}^H\big]\\
&= \mathrm{tr}\big[\Lambda(\Lambda + c\,\mathbf{I})^{-2}\big] = \sum_{l=1}^{N}\frac{\lambda_l}{(\lambda_l + c\,)^2}.
\end{aligned}
\tag{5.35}
$$

The received signal power is given by

$$
\begin{aligned}
P_T &= \mathbb{E}||\mathbf{HTx}||^2\\
&= \mathbb{E}\left[\beta^2\mathrm{tr}\left(\mathbf{U}\Lambda(\Lambda + c\,\mathbf{I})^{-1}\mathbf{U}^H\mathbf{xx}^H\mathbf{U}(\Lambda + c\,\mathbf{I})^{-1}\Lambda\mathbf{U}^H\right)\right]\\
&= \mathbb{E}\left[\beta^2\mathrm{tr}\left(\Lambda(\Lambda + c\,\mathbf{I})^{-1}\mathbf{U}^H\mathbf{xx}^H\mathbf{U}(\Lambda + c\,\mathbf{I})^{-1}\Lambda\mathbf{U}^H\mathbf{U}\right)\right]\\
&= \beta^2\mathrm{tr}\left(\Lambda(\Lambda + c\,\mathbf{I})^{-1}\mathbf{U}^H E(\mathbf{xx}^H)\mathbf{U}(\Lambda + c\,\mathbf{I})^{-1}\Lambda\right)\\
&= \frac{\beta^2\rho}{N}\mathrm{tr}\big[(\Lambda + c\,\mathbf{I})^{-2}\Lambda^2\big] = \frac{\beta^2\rho}{N}\sum_{l=1}^{N}\frac{\lambda_l^2}{(\lambda_l + c\,)^2}.
\end{aligned}
\tag{5.36}
$$

where we have used $\mathbb{E}(\mathbf{xx}^H) = \frac{\rho}{N}\mathbf{I}$.

The SINR is evaluated by computing the signal and interference powers from (5.34). For a given channel $\mathbf{H}$, the power of desired and interference signals at the k-th receive antenna

are respectively given by

$$P_D^{(k)} = \frac{\beta^2 \rho}{N} \left(\sum_{l=1}^{N} \frac{\lambda_l}{\lambda_l + c} |u_{kl}|^2 \right)^2 \tag{5.37}$$

$$P_I^{(k)} = \frac{\beta^2 \rho}{N} \sum_{i=1, i \neq k}^{N} \left| \sum_{l=1}^{N} \frac{\lambda_l}{\lambda_l + c} u_{kl} u_{il}^* \right|^2. \tag{5.38}$$

Thus the SINR for the k-th signal stream is given by

$$\begin{aligned}
\gamma_k &= \frac{P_D^{(k)}}{P_I^{(k)} + 1} \\
&= \frac{\frac{\beta^2 \rho}{N} \left(\sum_{l=1}^{N} \frac{\lambda_l}{\lambda_l + c} |u_{kl}|^2 \right)^2}{\frac{\beta^2 \rho}{N} \sum_{i=1, i \neq k}^{N} \left| \sum_{l=1}^{N} \frac{\lambda_l}{\lambda_l + c} u_{kl} u_{il}^* \right|^2 + 1}
\end{aligned} \tag{5.39}$$

Defining the exponential order of eigenvalues $\lambda_l = \rho^{-\alpha_l}$ in a manner similar to [2], and using the definition of $\eta = \beta^{-2}$,

$$\begin{aligned}
\gamma_k &= \frac{\left(\sum_l \frac{\rho^{-\alpha_l}}{\rho^{-\alpha_l} + c} |u_{kl}|^2 \right)^2}{\sum_{i \neq k} \left| \sum_{l=1}^{N} \frac{\rho^{-\alpha_l}}{\rho^{-\alpha_l} + c} u_{kl} u_{il}^* \right|^2 + N \rho^{-1} \eta} \\
&\doteq \frac{\left(\sum_l \rho^{-\alpha_l} |u_{kl}|^2 \right)^2}{\sum_{i \neq k} \left| \sum_{l=1}^{N} u_{kl} u_{il}^* \rho^{-\alpha_l} \right|^2 + N \rho^{-1} \sum_{l=1}^{N} \rho^{-\alpha_l}}
\end{aligned} \tag{5.40}$$

where we have substituted for η using (5.35), and the asymptotic equality follows because constant c dominates $\rho^{-\alpha_l}$, a fact that also implies $\eta \doteq \sum_l \rho^{-\alpha_l}$.

Multiplying the numerator and denominator of (5.40) by ρ^2, we have

$$\gamma_k \doteq \frac{\left(\sum_l \rho^{1-\alpha_l} |u_{kl}|^2 \right)^2}{\sum_{i \neq k} \left| \sum_{l=1}^{N} u_{kl} u_{il}^* \rho^{1-\alpha_l} \right|^2 + N \sum_{l=1}^{N} \rho^{1-\alpha_l}}. \tag{5.41}$$

The sum in the numerator of (5.41) is, in the SNR exponent, equivalent to:

$$\sum_l \rho^{1-\alpha_l}|u_{kl}|^2 \doteq \rho^{1-\alpha_{\min}}\sum_l |u_{kl}|^2$$

$$= \rho^{1-\alpha_{\min}} \tag{5.42}$$

where we use the fact that $\sum_l |u_{kl}|^2 = 1$. Similarly, for the first term in the denominator of (5.41)

$$\sum_{i\neq k}\left|\sum_{l=1}^{N} u_{kl}u_{il}^*\rho^{1-\alpha_l}\right|^2 \doteq \rho^{2-2\alpha_{\min}}\sum_{i\neq k}\left|\sum_{l=1}^{N} u_{kl}u_{il}^*\right|^2$$

$$= \rho^{2-2\alpha_{\min}}\sum_{i\neq k} w_{ki} \tag{5.43}$$

where we define $w_{ki} \triangleq \left|\sum_{l=1}^{N} u_{kl}u_{il}^*\right|^2$. Notice that $w_{ki} \leq 1$.

Using (5.42) and (5.43), the SINR in (5.41) is given by

$$\gamma_k \doteq \frac{\left(\rho^{1-\alpha_{\min}}\right)^2}{\rho^{2-2\alpha_{\min}}\sum_{i\neq k} w_{ki} + N\sum_{l=1}^{N}\rho^{1-\alpha_l}}. \tag{5.44}$$

If all $\alpha_\ell > 1$ then the exponents of ρ are negative and the denominator is dominated by its second term, which also dominates the numerator. If at least one of the $\alpha_\ell \leq 1$, then the maximum exponent which corresponds to $\alpha_{\min}$ dominates each summation. Thus we have:

$$\gamma_k \doteq \begin{cases} \rho^{1-\alpha_{\min}} & \alpha_1 > 1,\dots,\alpha_N > 1 \\ \dfrac{\left(\rho^{1-\alpha_{\min}}\right)^2}{\rho^{2-2\alpha_{\min}}\sum_{\substack{i=1\\i\neq k}}^{N} w_{ki} + N\rho^{1-\alpha_{\min}}} & \text{otherwise} \end{cases} \tag{5.45}$$

We now concentrate on the case where there exists at least one $\alpha_\ell \leq 1$. We define

$$\mu_{\min} \triangleq \min_{\substack{k,i \\ k\neq i}} w_{ki} \tag{5.46}$$

therefore in this special case we have:

$$\gamma_k \dot{\leqslant} \frac{\left(\rho^{1-\alpha_{\min}}\right)^2}{(N-1)\left(\rho^{1-\alpha_{\min}}\right)^2 \mu_{\min} + N\,\rho^{1-\alpha_{\min}}} \tag{5.47}$$

$$\doteq \frac{1}{(N-1)\mu_{\min}} \tag{5.48}$$

$$\triangleq \bar{\gamma}$$

Thus in general

$$\gamma_k \dot{\leqslant} \frac{\nu}{(N-1)\mu_{\min}} \tag{5.49}$$

$$\triangleq \bar{\gamma}$$

where ν is a new random variable defined as:

$$\nu = \begin{cases} \kappa_\alpha & \text{if } \alpha_k > 1 \ \forall k \\ \\ 1 & \text{otherwise} \end{cases} \tag{5.50}$$

where $\kappa_\alpha \triangleq \rho^{1-\alpha_{\min}}$.

We can now bound the outage probability as follows

$$\begin{aligned} P_{\text{out}} &= \mathbb{P}\left(\sum_{k=1}^{N} \log(1+\gamma_k) \leqslant R \right) \\ &\dot{\geqslant} \mathbb{P}\left(\sum_{k=1}^{N} \log(1+\bar{\gamma}) \leqslant R \right) \\ &= \mathbb{P}\left(\frac{\nu}{(N-1)\mu_{\min}} \leqslant 2^{R/N} - 1 \right) \\ &= \mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta \right) \end{aligned} \tag{5.51}$$

where $\Theta \triangleq (2^{R/N} - 1)(N-1)$.

The bound in (5.51) can be evaluated as follows

$$\mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta \right) = \mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta \big| \nu = \kappa_\alpha \right) \mathbb{P}(\nu = \kappa_\alpha) + \mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta \big| \nu = 1 \right) \mathbb{P}(\nu = 1)$$

$$= \mathbb{P}\left(\kappa_\alpha \leqslant \Theta\,\mu_{\min} \right) \mathbb{P}(\nu = \kappa_\alpha) + \mathbb{P}\left(\frac{1}{\mu_{\min}} \leqslant \Theta \right) \mathbb{P}(\nu = 1). \tag{5.52}$$

Notice that $\mathbb{P}\big(\kappa_\alpha \leqslant \Theta\, \mu_{\min}\big) \doteq 1$ since κ_α is vanishing at high SNR and Θ and $\mu_{\min}$ are positives. We now need to compute $\mathbb{P}\big(\nu = \kappa_\alpha\big)$ and $\mathbb{P}\big(\nu = 1\big)$, or equivalently $\mathbb{P}\big(\{\alpha_k > 1\ \forall k\}\big)$ and its complement. We use one of the results of [5].

Lemma 5.3.1 *Let* $\{\lambda_n\}$ *denotes the eigenvalues of a Wishart matrix* $\mathbf{HH}^H$, *where* $\mathbf{H}$ *is an* $N \times M$ *matrix with i.i.d Gaussian entries, and let* $\alpha_n = -\frac{\log(\lambda_n)}{\log(\rho)}$. *If* $\mathbf{1}_{\alpha_n}$ *denotes the number of* α_n *that are greater than one, then for any integer* $s \leqslant N$ *we have [5, Section III-A]* [1]

$$\mathbb{P}\big(\mathbf{1}_{\alpha_n} = s\big) \doteq \rho^{-(s^2 + (M-N)s)}. \tag{5.53}$$

Thus setting $s = N$ (i.e. all $\alpha_n > 1$) in (5.53) yields

$$\mathbb{P}\big(\nu = \kappa_\alpha\big) = \mathbb{P}\big(\mathbf{1}_{\alpha_n} = N\big) \doteq \rho^{-MN} \tag{5.54}$$

$$\mathbb{P}\big(\nu = 1\big) \doteq O(1) \tag{5.55}$$

where $O(1)$ is a non-zero constant with respect to ρ.

Evaluating (5.52) depends on the values of Θ which is always real and positive. If $\Theta < 1$ then we have

$$\mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta\right) \doteq \rho^{-MN} \tag{5.56}$$

because $\mathbb{P}\big(\frac{1}{\mu_{\min}} \leqslant \Theta\big) = 0$ as $1/\mu_{\min} > 1$. On the other hand if $\Theta > 1$ then

$$\mathbb{P}\left(\frac{\nu}{\mu_{\min}} \leqslant \Theta\right) \doteq \rho^{-MN} + \mathbb{P}\big(\frac{1}{\mu_{\min}} \leqslant \Theta\big)O(1) \tag{5.57}$$

$$\doteq O(1) \tag{5.58}$$

since $\mathbb{P}\big(\frac{1}{\mu} \leqslant \Theta\big)$ is not a function of ρ because μ is independent ρ. For the set of rates where $\Theta > 1$, equation (5.58) implies that the outage probability in (5.82) is not function of ρ and thus the diversity is zero, i.e. the system will have error floor. The set of rates for which $\Theta > 1$ are

$$R > N \log\left(\frac{N}{N-1}\right) \triangleq R_{th}. \tag{5.59}$$

---

[1]Note that [5] analyzes linear MIMO receiver where it is assumed $N \geqslant M$. It can be easily shown that the above Lemma 5.3.1 applies for the case considered here where $M \geqslant N$.

This concludes the calculation of a lower bound on the outage probability. A similar approach will yield a corresponding upper bound, as follows. Let

$$\mu_{\max} \triangleq \max_{k \neq i} |u_{kl'} u_{il'}^*|^2 \tag{5.60}$$

A lower bound on the SINR is given as

$$\gamma_k \dot{\geqslant} \frac{\nu}{(N-1)\mu_{\max}} \tag{5.61}$$

$$\triangleq \hat{\gamma}.$$

The outage probability is bounded as

$$P_{\text{out}} \dot{\leqslant} \mathbb{P}\left(\sum_{k=1}^{N} \log(1+\hat{\gamma}) \leqslant R \right)$$

$$= \mathbb{P}\left(\frac{\nu}{\mu_{\max}} \leqslant \Theta \right). \tag{5.62}$$

We can evaluate (5.62) in a similar way as (5.52), establishing that the outage diversity $d_{out}^{RZF} = MN$ if the operating spectral efficiency R is less than $R_{th} = N \log\left(\frac{N}{N-1}\right)$, and $d_{out}^{RZF} = 0$ if $R > R_{th}$. This shows that the performance of RZF precoder can be much better than that of the conventional ZF precoder MIMO system whose diversity is $M - N + 1$ independent of rate.

Recall that diversity is the SNR exponent of the probability of codeword error. In Appendix 5.7.1, we show that the outage exponent tightly bounds the SNR exponent of the error probability. Thus we have the following theorem.

Theorem 5.3.2 *For an $M \times N$ MIMO system that utilizes joint spatial encoding and regularized ZF precoder given by (5.32), the outage diversity is $d^{RZF} = MN$ if the operating spectral efficiency R is less than $R_{th} = N \log\left(\frac{N}{N-1}\right)$, and $d^{RZF} = 0$ if $R > R_{th}$.*

Remark 5.3.1 *R_{th} is a monotonically decreasing function of N with the asymptotic value $\lim_{N \to \infty} R_{th} = \frac{1}{\ln 2} \approx 1.44$. Overall we have $1.44 \leq R_{th} \leq 2$, leading to an easily remembered rule of thumb that applies to all antenna configurations. Regularized ZF precoders always exhibit an error floor at spectral efficiencies above 2 b/s/Hz, and enjoy full diversity at spectral efficiencies below 1.44 b/s/Hz.*

5.3.4 Matched Filter Precoding

The transmit matched filter (TxMF) is introduced in [33, 37]. The TxMF maximizes the signal-to-interference ratio (SIR) at the receiver and is optimum for high signal-to-noise-ratio scenarios [33]. The TxMF is also proposed for non-cooperative cellular wireless network [41]. The TxMF is derived by maximizing the ratio between the power of the desired signal portion in the received signal and the signal power under the transmit power constraint, that is [33]

$$\mathbf{T} = \arg\max_{\mathbf{T}} \frac{\mathbb{E}\big(||\mathbf{x}^H \tilde{\mathbf{y}}||^2\big)}{\mathbb{E}\big(||\mathbf{n}||^2\big)} \tag{5.63}$$

$$\text{subject to: } \mathbb{E}||\mathbf{T}\mathbf{x}||^2 \leqslant \rho$$

where $\tilde{\mathbf{y}}$ is the noiseless received signal $\tilde{\mathbf{y}} = \mathbf{T}\mathbf{x}$.

The solution to (5.63) is given by

$$\mathbf{T} = \beta \mathbf{H}^H \tag{5.64}$$

with

$$\beta = \sqrt{\frac{1}{\text{tr}(\mathbf{H}^H \mathbf{H})}}. \tag{5.65}$$

We now analyze the diversity for the MIMO system under TxMF. The received signal is given by

$$\mathbf{y} = \mathbf{H}\mathbf{H}^H \mathbf{x} + \mathbf{n} = \beta \mathbf{U} \Lambda \mathbf{U}^H \mathbf{x} + \mathbf{n}.$$

The received signal at the k-th antenna

$$y_k = \beta \bigg(\sum_{l=1}^{N} \lambda_l |u_{kl}|^2 \bigg) x_k + \beta \sum_{i=1, i \neq k}^{N} \bigg(\sum_{l=1}^{N} \lambda_l u_{kl} u_{il}^* \bigg) x_i + n_k \tag{5.66}$$

The SINR at k-th receive antenna is

$$\gamma_k = \frac{\beta^2 \frac{\rho}{N} \bigg(\sum_{l=1}^{N} \lambda_l |u_{kl}|^2 \bigg)^2}{\beta^2 \frac{\rho}{N} \sum_{i=1, i \neq k}^{N} \bigg| \sum_{l=1}^{N} \lambda_l u_{kl} u_{il}^* \bigg|^2 + 1}$$

Substitute with the value of β and $\lambda_l = \rho^{-\alpha_l}$

$$\gamma_k = \frac{\left(\sum_{l=1}^{N} \rho^{-\alpha_l} |u_{kl}|^2 \right)^2}{\sum_{i=1, i \neq k}^{N} \left| \sum_{l=1}^{N} \rho^{-\alpha_l} u_{kl} u_{il}^* \right|^2 + N \rho^{-1} \sum_{l=1}^{N} \rho^{-\alpha_l}} \tag{5.67}$$

Observe that (5.67) is the same as the SINR of the RZF precoded system given by (5.40). Hence the analysis in the present case follows closely that of the outage lower bound of the RZF precoder, with the following result: the system can achieve full diversity as long as the operating rate is less than R_{th} given in (5.59). The pairwise error probability analysis is also similar to that of the RZF precoding system (given in Appendix 5.7.1) which we omit for brevity. Thus we conclude that Theorem 5.3.2 applies for the TxMF precoder.

5.3.5 Wiener Filter Precoding

The transmit Wiener filter TxWF minimizes the weighted MSE function.

$$\{\mathbf{T}, \beta\} = \operatorname{argmin}_{\mathbf{T}, \beta} \mathbb{E}\left(||\mathbf{x} - \beta^{-1} \tilde{\mathbf{y}}||^2 \right)$$

$$\text{subject to } \mathbb{E}\left(||\mathbf{T}\mathbf{x}||^2 \right) \leqslant \rho. \tag{5.68}$$

Solving (5.68) yields

$$\mathbf{T} = \beta \bar{F}^{-1} \mathbf{H}^H \tag{5.69}$$

with

$$\bar{F} = \left(\mathbf{H}^H \mathbf{H} + \frac{N}{\rho} \mathbf{I} \right)$$

$$\beta = \sqrt{\frac{1}{\operatorname{tr}(\bar{F}^{-2} \mathbf{H}^H \mathbf{H})}} \tag{5.70}$$

where β can be interpreted as the optimum gain for the combined precoder and channel [33].

Notice that the TxWF precoding function is similar to that of the MMSE equalizer [16]. Indeed the SINR of both systems are equivalent. To see this, we first compute the SINR for

the precoded $\mathbf{H} \in \mathbb{C}^{M \times N}$ (with $M \geqslant N$) MIMO channel

$$\gamma_k = \frac{\frac{\rho \beta}{N} |(\mathbf{T} \, \mathbf{H})_{kk}|^2}{\frac{\rho \beta}{N} \sum_{i \neq k}^{N} |(\mathbf{T} \, \mathbf{H})_{ki}|^2 + 1} \tag{5.71}$$

$$= \frac{\frac{\rho}{N} |(\mathbf{T} \, \mathbf{H})_{kk}|^2}{\frac{\rho}{N} \sum_{i \neq k}^{N} |(\mathbf{T} \, \mathbf{H})_{ki}|^2 + \mathrm{tr}(\bar{F}^{-2} \mathbf{H}^H \mathbf{H})} \tag{5.72}$$

where we have used the independence of the transmitted signal to compute (5.71).

Now consider a MIMO channel $\mathbf{H}_2 = \mathbf{H}^T \in \mathbb{C}^{N \times M}$. The MMSE equalizer for this channel is given by

$$\mathbf{W}_e = (\mathbf{H}_2^H \mathbf{H}_2 + \frac{N}{\rho} \mathbf{I})^{-1} \mathbf{H}_2^H. \tag{5.73}$$

The received SINR for that system is given by

$$\gamma_k^{MMSE} = \frac{\frac{\rho}{N} |(\mathbf{W}_e \, \mathbf{H}_2)_{kk}|^2}{\frac{\rho}{N} \sum_{i \neq k}^{N} |(\mathbf{W}_e \, \mathbf{H}_2)_{ki}|^2 + \mathrm{tr}(\mathbf{W}_e \mathbf{W}_e)}. \tag{5.74}$$

Since $\mathbf{W}_e \, \mathbf{H}_2 = \mathbf{T}_{WFP} \mathbf{H}$ and $\mathrm{tr}(\mathbf{W}_e \mathbf{W}_e) = \mathrm{tr}(\bar{F}^{-2} \mathbf{H}^H \mathbf{H})$, we conclude that $\gamma_k^{MMSE} = \gamma_k^{WFP}$. Hence the diversity analysis of [5, 7] for the MIMO MMSE receiver applies for the MIMO Wiener precoding system. It is shown in [5] that this diversity is a function of rate R and number of transmit and receive antennas. We thus conclude the following.

Lemma 5.3.3 *Consider a channel* $\mathbf{H} \in \mathbb{C}^{M \times N}$ *the diversity of the MIMO system under Wiener filter precoding is given by*

$$d^{WFP} = \lceil N 2^{-\frac{R}{N}} \rceil^2 + (M - N) \lceil N 2^{-\frac{R}{N}} \rceil \tag{5.75}$$

where $(\cdot)^+ = \max(\cdot, 0)$ *and* $\lceil \cdot \rceil$.

Remark 5.3.2 *It is commonly stated that MMSE and ZF operators "converge" at high SNR. The developments in this chapter as well as [6] serve to show that although not false, this comment is essentially fruitless because the performance of MMSE and ZF at high SNR are very different. This apparent incongruity is explained in the broadest sense as follows: Even though the MMSE coefficients converge to ZF coefficients as* $\rho \to \infty$*, the high sensitivity*

of logarithm of errors (especially at low error probabilities) to coefficients is such that the convergence of MMSE to ZF coefficients is not fast enough for the logarithm of respective errors to converge.

5.4   Diversity-Multiplexing Tradeoff in Precoding

For increasing sequence of SNRs, consider a corresponding sequence of codebooks $\mathcal{C}(\rho)$, designed at increasing rates $R(\rho)$ and yielding average error probabilities $P_e(\rho)$. Then define

$$r = \lim_{\rho \to \infty} \frac{R(\rho)}{\log \rho}$$
$$d = - \lim_{\rho \to \infty} \frac{\log P_e(\rho)}{\log \rho}.$$

For each r the corresponding diversity $d(r)$ is defined (with a slight abuse of notation) as the supremum of the diversities over all possible codebook sequences $\mathcal{C}(\rho)$.

From the viewpoint of definitions, the traditional notion of diversity can be considered a special case of the DMT by setting $r = 0$. However, from the viewpoint of analysis, the approximations needed in DMT calculation make use of $R(\rho)$ being a *strictly* increasing function, while for diversity analysis R is constant (not strictly increasing function of ρ). Thus, although sometimes DMT analysis may produce results that are luckily consistent with diversity analysis[2] ($r = 0$), in other cases the DMT analysis may produce results that are inconsistent with diversity analysis. Certain equalizers and precoders fall into the latter category. In the following, we calculate the DMT of the various precoders considered up to this point.

---

[2]E.g. the point-to-point MIMO channel with ML decoding.

ZF Precoding

Recall that two ZF precoding designs have been considered. For the ZF precoder minimizing power, given by (5.5), the outage upper bound in (5.9) can be written as

$$P_{out} \leqslant \mathbb{P}\big(\lambda_{\min} \leqslant \rho^{(\frac{r}{N}-1)}\big) \tag{5.76}$$

$$\doteq \rho^{-(M-N+1)(1-\frac{r}{N})} \tag{5.77}$$

where we substitute $R = r \log \rho$ to obtain (5.76), and equation (5.77) follows in a manner identical to the procedure that led to (5.11).

Similarly the outage lower bound (5.12) can be written as

$$P_{out} \geqslant \mathbb{P}\big(z \leqslant \rho^{(\frac{r}{N}-1)}\big)$$

$$\doteq \rho^{-(M-N+1)(1-\frac{r}{N})}. \tag{5.78}$$

From (5.77) and (5.78) we conclude

$$d^{ZFP}(r) = (M - N + 1)\big(1 - \frac{r}{N}\big)^+. \tag{5.79}$$

The DMT of the ZF precoder maximizing the throughput, given by (5.16), is obtained in an essentially similar manner to the above, therefore the discussion is omitted in the interest of brevity.

Regularized ZF Precoding

We begin by producing an outage lower bound. To do so, we start by the bound on the SINR of each stream k obtained in (5.45), and further bound it by discarding some positive

terms in the denominator.

$$\bar{\gamma}_k = \frac{\left(\rho^{1-\alpha_{\min}}\right)^2}{\sum_{i \neq k} \left|u_{kl} u_{il}^* \rho^{1-\alpha_{\min}}\right|^2 + N\rho^{1-\alpha_{\min}}}$$

$$\leqslant \begin{cases} \dfrac{\left(\rho^{1-\alpha_{\min}}\right)^2}{\rho^{2(1-\alpha_{\min})}\left|u_{kl'} u_{2l'}^*\right|^2 + N\rho^{1-\alpha_{\min}}} & k = 1 \\[4ex] \dfrac{\left(\rho^{1-\alpha_{\min}}\right)^2}{\rho^{2(1-\alpha_{\min})}\left|u_{kl'} u_{1l'}^*\right|^2 + N\rho^{1-\alpha_{\min}}} & k > 1 \end{cases}$$

$$\doteq \begin{cases} \dfrac{1}{|u_{kl'} u_{2l'}^*|^2} & k = 1 \\[3ex] \dfrac{1}{|u_{kl'} u_{1l'}^*|^2} & k > 1 \end{cases}$$

We can now bound the outage probability

$$P_{\text{out}} = \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R\right)$$

$$\dot{\geqslant} \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \bar{\gamma}_k) \leqslant R\right)$$

$$\geqslant \mathbb{P}\left(N \log \sum_{k=1}^{N} \frac{1}{N}(1 + \bar{\gamma}_k) \leqslant R\right) \tag{5.80}$$

$$\doteq \mathbb{P}\left(\sum_{k=1}^{N} \frac{1}{N}(1 + \bar{\gamma}_k) \leqslant \rho^{\frac{r}{N}}\right) \tag{5.81}$$

$$\doteq \mathbb{P}\left(\sum_{k=1}^{N} \bar{\gamma}_k \leqslant \rho^{\frac{r}{N}}\right)$$

$$\dot{\geqslant} \mathbb{P}\left(\frac{\nu}{|u_{kl'} u_{2l'}^*|^2} + \sum_{k=2}^{N} \frac{\nu}{|u_{kl'} u_{1l'}^*|^2} \leqslant \rho^{\frac{r}{N}}\right). \tag{5.82}$$

where we have used Jensen's inequality in (5.80).

For notational convenience define

$$\psi \triangleq \frac{1}{|u_{kl'} u_{2l'}^*|^2} + \sum_{k=2}^{N} \frac{1}{|u_{kl'} u_{1l'}^*|^2} \ .$$

Then the bound in (5.82) can be evaluated as follows:

$$\mathbb{P}\big(\nu\psi \leqslant \rho^{\frac{r}{N}}\big)$$

$$= \mathbb{P}\big(\nu\psi \leqslant \rho^{\frac{r}{N}}\big|\nu = 0\big)\mathbb{P}\big(\nu = 0\big) + \mathbb{P}\big(\nu\psi \leqslant \rho^{\frac{r}{N}}\big|\nu = 1\big)\mathbb{P}\big(\nu = 1\big)$$

$$= \mathbb{P}\big(0 \leqslant \rho^{\frac{r}{N}}\big)\mathbb{P}\big(\nu = 0\big) + \mathbb{P}\big(\psi \leqslant \rho^{\frac{r}{N}}\big)\mathbb{P}\big(\nu = 1\big)$$

$$\doteq \rho^{-MN} + \mathbb{P}\big(\psi \leqslant \rho^{\frac{r}{N}}\big)\, O(1). \tag{5.83}$$

$$\geqslant \rho^{-MN} + O(1) \tag{5.84}$$

$$= O(1) \tag{5.85}$$

where (5.83) follows from Lemma 5.3.1, and (5.84) is true as long as $\mathbb{P}\big(\psi \leqslant \rho^{\frac{r}{N}}\big) = O(1)$, the proof of which is relegated to Appendix 5.7.2.

Since the outage lower bound (5.84) is not a function of ρ, the system will always have an error floor. In other words the DMT is given by

$$d^{RZFP}(r) = 0 \qquad 0 < r \leq B \tag{5.86}$$

We saw earlier that in the fixed-rate regime RZF precoding enjoys full diversity for spectral efficiencies below a certain threshold, but it now appears that DMT shows only zero diversity. DMT is not capable of predicting the complex behavior at $r = 0$ because the DMT framework only assigns a single value diversity to all distinct spectral efficiencies at $r = 0$. A similar behavior was observed and analyzed for the MMSE MIMO receiver [5–7].

Matched Filter Precoding

The DMT of the MIMO system with TxMF is the same as the DMT given by (5.86) due to the similarity in the outage analysis (see Section 5.3.4). We omit the details for brevity.

Wiener Filter Precoding

Since the the received SINR of the MIMO system using TxWF precoding is the same as that of MIMO MMSE receiver, we conclude from [7] that the DMT for the TxWF precoding
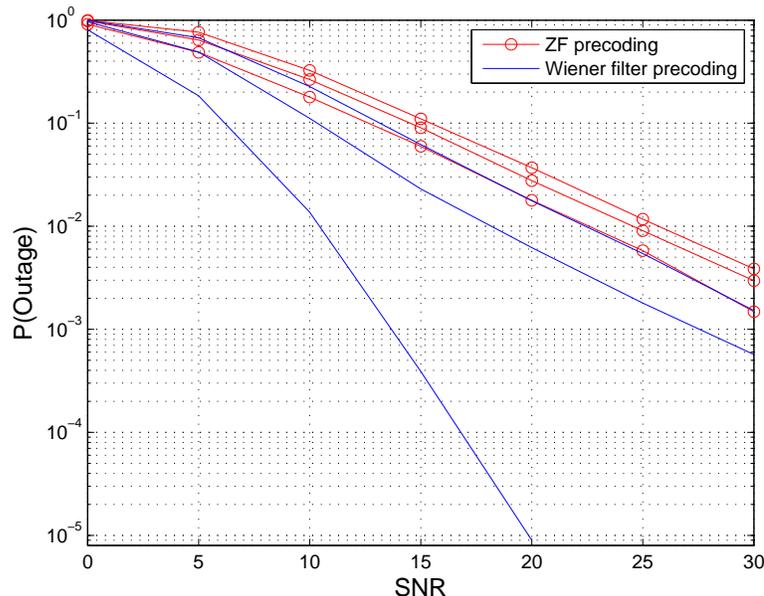
Figure 5.2. Outage probability of the ZF and Wiener filtering precoded MIMO 2×2 system for rates (left to right): $R = 1.9, 2.5,$ and 3 b/s/Hz.

system is

$$d^{WFP}(r) = (M - N + 1)\left(1 - \frac{r}{N}\right)^{+}. \tag{5.87}$$

Similarly to the MIMO MMSE receiver [5,7], we observe that DMT for the MIMO system with TxWF does not always predict the diversity in the fixed rate regime given by (5.75).

5.5 Simulation Results

This section produces numerical results for the outage probabilities of ZF, regularized ZF (RZF), matched filter (MF) and Wiener precoding systems. Figure 5.2 shows the outage probabilities of the ZF and Wiener-filter precoded 2×2 MIMO systems. The diversity in the case of the ZF case is the same as the one predicted by the DMT. In the case of Wiener precoding, the diversity is the same as the one predicted by the DMT for high rate (R) values and it departs from the DMT for low rate values. A complete diversity characterization is given by (5.75) which is similar to that of the MMSE MIMO equalizer [5]. Figure 5.3 shows
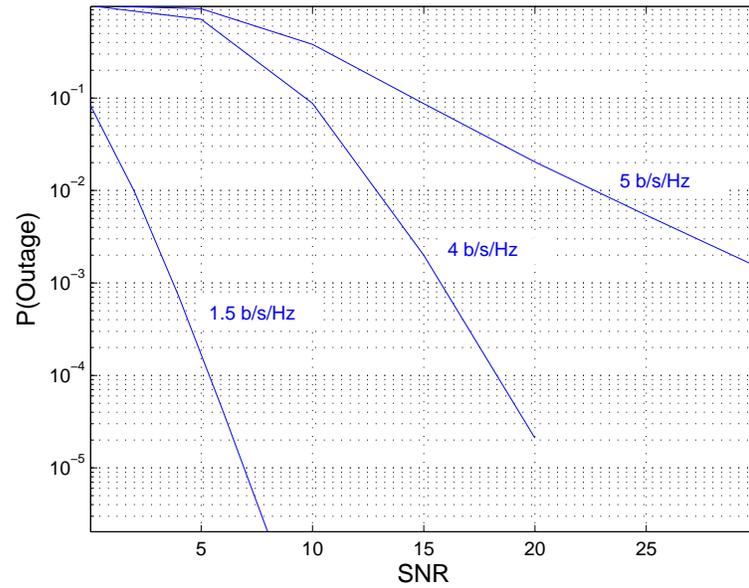
Figure 5.3. Wiener precoded 3×3 MIMO system. The diversities are $d = 9, 4$ and 1 for $R = 1.5, 4$ and 5 b/s/Hz respectively.
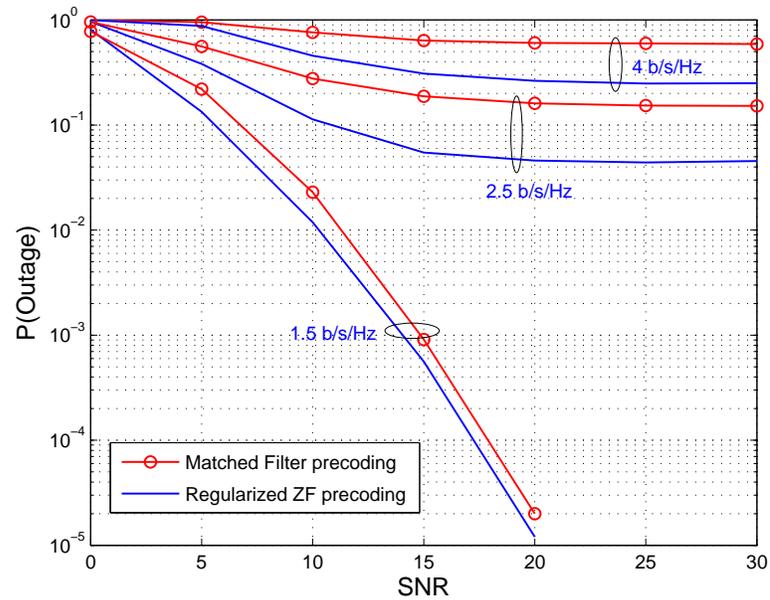
Figure 5.4. MF and regularized ZF precoded 2×2 MIMO system for rates (left to right): $R = 1.9, 2.5$, and 4 b/s/Hz. The diversity is $d = 4$ for $R = 1.5$ b/s/Hz and $d = 0$ otherwise.

outage probabilities for a 3×3 MIMO system with Wiener precoding. The diversity for the rates $R = 1.5, 4$, and 5 b/s/Hz is $9, 4$ and 1 respectively. Figure 5.4 shows an error floor for the regularized ZF and matched filtering precoded 2×2 system at high rates. However we observe that the maximum diversity is achieved for any rate $R < 2$ (c.f. Equation (5.59)).

5.6 Conclusion

Linear precoders provide a simple and efficient processing, and have been shown to be optimal in some scenarios [34–36]. This chapter studies the high-SNR performance of linear precoders. It is shown that the zero-forcing precoder under two common design approaches, maximizing the throughput and minimizing the transmit power, achieves the same DMT as that of MIMO systems with ZF equalizer. When a regularized ZF (RZF) precoder (for a fixed regularization term that is independent of the signal-to-noise ratio) or matched filter (MF) precoder is used, we have $d(r) = 0$ for all r, implying an error floor under all conditions. It is also shown that in the fixed rate regime, RZF and MF precoding achieve full diversity up to a certain spectral efficiency, while at higher spectral efficiencies they produce an error floor. If the regularization parameter in the RZF is optimized in the MMSE sense, the RZF precoded MIMO system exhibits a complex rate-dependent behavior. In particular, the diversity of this system (also known as Wiener filter precoding) is characterized by $d(R) = \lceil N2^{-\frac{R}{N}} \rceil^2 + (M - N)\lceil N2^{-\frac{R}{N}} \rceil$ where M and N are the number of transmit and receive antennas. This is the same behavior observed in linear MMSE MIMO receivers [5].

5.7 Appendix

5.7.1 Pairwise error Probability (PEP) Analysis

In this section we perform PEP analysis for the the zero-forcing (ZF) and the regularized ZF (RZF) precoding systems. The presented analysis can be easily extended to all other precoding systems. The basic strategy is to show the SNR exponent of outage probability

bounds the SNR exponent of PEP from both sides The PEP analysis follows from $[5, 15]$, with careful attention to the system model given by Equation (5.1).

The lower bound immediately follows from [15, Lemma 3] by recognizing that although it was developed for SISO block equalization, nowhere in its development does it depend on the number of receive antennas, therefore we can directly use it for our purposes:

$$P_{err} \,\dot{\geqslant}\, P_{out}. \tag{5.88}$$

The upper bound on PEP for the ZF/RZF precoding systems receiver is developed using the union bound. Denote the channel outage event by O and the error event by E. The PEP is given by

$$P_{\text{err}} = P(E|O) \, P_{out} + P(E, \bar{O})$$
$$\leqslant P_{out} + P(E, \bar{O}). \tag{5.89}$$

In order to show that P_{out} dominates the right hand side of (5.89), it is shown in [5] that the probability $P(E, \bar{O})$ can be bounded as follows using the union bound

$$\mathbb{P}(E, \bar{O}) \quad \dot{\leqslant} \quad 2^{Rl} e^{-\frac{\rho/N}{\sigma_{\tilde{\mathbf{n}}}^2(k)}} \,\dot{\leqslant}\, \rho^{-MN} \tag{5.90}$$

where l is the codeword length and $\sigma_{\tilde{\mathbf{n}}}^2(k)$ is the variance of the interference plus noise signal $\tilde{\mathbf{n}}$ in the k-th receive stream [3]. The proof of [15] does not depend on the codeword length for both upper and lower PEP bounds. The bound are tight and were confirmed by simulations for outage and error probabilities.

We now show that a similar proof holds for regularized zero-forcing (RZFP). Recall that the outage probability of the RZFP can be upper bounded by (5.62)

$$P_{out} \leqslant \mathbb{P}\big(\frac{\nu}{\mu_{\max}} \leqslant \Theta\big) \triangleq P_{out}^b \tag{5.91}$$

We will use P_{out}^b to further bound (5.89). Moreover $P(E, \bar{O})$ can be upper bounded by bounding the noise variance $\sigma_{\tilde{\mathbf{n}}}^2(k)$ in (5.90)

$$\sigma_{\tilde{\mathbf{n}}}^2(k) = P_I + P_n < P_T + 1 \tag{5.92}$$

---

[3] [15] analyzes linear receivers so $\tilde{\mathbf{n}}$ is the k-th output filtered interference plus noise signals. By symmetry assumption all the equalizer outputs have equal noise variance.

where we have used the noise power $P_n = 1$, and bound the interference power by the total received power P_T. We will first consider the case of RZF precoding since the case of ZF precoding can be easily deduced from RZF by substituting setting the regularization parameter $c = 0$. For the RZF precoding system we use the P_T given by (5.36) which can be simplified in a way similar to earlier sections

$$
\begin{aligned}
P_T &= \frac{\beta^2 \rho}{N} \sum_{l=1}^{N} \frac{\lambda_l^2}{(\lambda_l + c)^2} \\
&= \frac{1}{\sum_{l=1}^{N} \frac{\lambda_l}{(\lambda_l+c)^2}} \frac{\rho}{N} \sum_{l=1}^{N} \frac{\lambda_l^2}{(\lambda_l + c)^2} \\
&= \frac{1}{\sum_{l=1}^{N} \frac{\rho^{-\alpha_l}}{(\rho^{-\alpha_l}+c)^2}} \frac{\rho}{N} \sum_{l=1}^{N} \frac{\rho^{-2\alpha_l}}{(\rho^{-\alpha_l} + c)^2} \\
&\doteq \frac{1}{\rho^{-\alpha_{\min}}} \frac{\rho}{N} \rho^{-2\alpha_{\min}} \\
&= \frac{1}{N} \rho^{1-\alpha_{\min}}.
\end{aligned}
\tag{5.93}
$$

Using the union bound (5.90),

$$
P(E, \bar{O}) \dot{\leqslant}
\begin{cases}
2^{Rl} e^{-\rho^{\alpha_{\min}}} & \alpha_{\min} < 1 \\
2^{Rl} e^{-\frac{\rho}{N}} & \alpha_{\min} > 1
\end{cases}
\tag{5.94}
$$

Since the exponential function dominates polynomials we have

$$
\lim_{\rho \to \infty} \frac{e^{-\rho^{\alpha_{min}}}}{\rho^{-MN}} = 0
$$

and

$$
\lim_{\rho \to \infty} \frac{e^{-\rho}}{\rho^{-MN}} = 0
$$

which in turns gives

$$
P(E, \bar{O}) \dot{\leqslant} \rho^{-MN} .
\tag{5.95}
$$

Using (5.91) and (5.95), the PEP given by (5.89) is bounded as

$$
\begin{aligned}
P_{\text{err}} &\mathrel{\dot{\leqslant}} P_{out} + P(E, \bar{O}) \\
&\mathrel{\dot{\leqslant}} P_{out}^b + \rho^{-MN} \\
&\mathrel{\dot{=}} P_{out}^b \\
&= \rho^{-d_{out}}.
\end{aligned}
\tag{5.96}
$$

therefore $d \geqslant d_{out}$ which concludes the proof for the RZF system.

For the ZF precoding system, it can be directly shown that a similar proof holds for both ZF precoding designs.

5.7.2 Proof of Eq. (5.84)

Recall that

$$
\psi \triangleq \frac{1}{|u_{1l'}u_{2l'}^*|^2} + \sum_{k=2}^{N} \frac{1}{|u_{kl'}u_{1l'}^*|^2}.
$$

All terms of ψ the common factor $\frac{1}{|u_{1l'}|^2}$. Thus we have

$$
\begin{aligned}
\psi &= \psi_a \psi_b \\
\psi_a &= \frac{1}{|u_{1l'}|^2} \\
\psi_b &= \left(\frac{1}{|u_{2l'}^*|^2} + \frac{1}{|u_{2l'}|^2} + \frac{1}{|u_{3l'}|^2} + \frac{1}{|u_{4l'}|^2} + \cdots + \frac{1}{|u_{Nl'}|^2} \right).
\end{aligned}
\tag{5.97}
$$

Observe that all the terms of ψ_b are distinct except for the first two.

We now bound the probability $\mathbb{P}\left(\psi \leqslant \rho^{\frac{r}{N}}\right)$.

$$
\begin{aligned}
\mathbb{P}\left(\psi \leqslant \rho^{\frac{r}{N}}\right) &\geqslant \mathbb{P}\left(\psi \leqslant \rho^{\frac{r}{N}} \mid \psi < c\right) \mathbb{P}(\psi < c) \\
&\geqslant \mathbb{P}\left(c \leqslant \rho^{\frac{r}{N}}\right) \mathbb{P}(\psi < c) \\
&\mathrel{\dot{=}} \quad \mathbb{P}(\psi < c)
\end{aligned}
\tag{5.98}
$$

Using $\psi = \psi_a \psi_b$ we can further bound (5.98)

$$\mathbb{P}(\psi < c) = \mathbb{P}(\psi_a \psi_b < c)$$
$$\geqslant \mathbb{P}\big(\psi_a \psi_b \leqslant c \,\big|\, \psi_a < c_2\big)\, \mathbb{P}(\psi_a < c_2)$$
$$\geqslant \mathbb{P}\big(c_2 \psi_b \leqslant c\big)\, \mathbb{P}(\psi_a < c_2).$$

We thus have

$$\mathbb{P}\big(\psi \leqslant \rho^{\frac{r}{N}}\big) \geqslant \mathbb{P}\big(\psi_b \leqslant c'\big)\, \mathbb{P}(\psi_a < c_2) \tag{5.99}$$

and $c' = c/c_2$.

We now evaluate the two probabilities in the right hand side of (5.99). The first probability $\mathbb{P}\big(\psi_b \leqslant c'\big) = O(1)$. The proof easily follows from [7, Appendix A] with the observation that this proof holds even when the two first elements of ψ_b are the same. The second probability $\mathbb{P}(\psi_a < c_2)$ is evaluated as follows. Let $q = |u_{1l'}|^2$. We use the following distributions from [38, Appendix A]

$$f(q) = (N-1)(1-q)^{N-2}, \quad 0 \leqslant q \leqslant 1$$

then

$$\mathbb{P}(\psi_b < c_2) = \mathbb{P}(q > \frac{1}{c_2})$$
$$= \int_{\frac{1}{c_2}}^{1} f(q)\, dq$$
$$= (1 - \frac{1}{c_2})^{N-2} \tag{5.100}$$

Observing that (5.100) is not a function of ρ concludes the proof.

CHAPTER 6

EQUALIZATION FOR LINEARLY PRECODED TRANSMISSION

6.1 Introduction

The objective of a precoded transmitter is to separate the data streams at the receiver. In other words, linear precoding is a method of interference management at the transmitter. In general, precoded systems do not require interference management at the receiver, however, once a transmitter is designed and standardized (as precoders have been), some standards-compliant receivers may opt to further equalize the precoded channel (see Figure 6.1). This section analyzes the equalization of precoded transmissions.

When the transmit and receive filters can be designed jointly and from scratch, singular value decomposition becomes an attractive option whose diversity has been analyzed in [42]. The distinction of the systems analyzed in this section is that the precoders can be used with or without the receive filters, while with the SVD solution neither the transmit nor the receive filters can operate without each other.

A snapshot of some of the results of this section is as follows. It is shown that equalization at the receiver can alleviate the error floor that was observed in matched filter precoding as well as regularized ZF precoding. It is shown that MMSE equalization does not affect the diversity of Wiener filter precoding, but ZF equalization does indeed affect the diversity of Wiener filter precoding in a negative way.
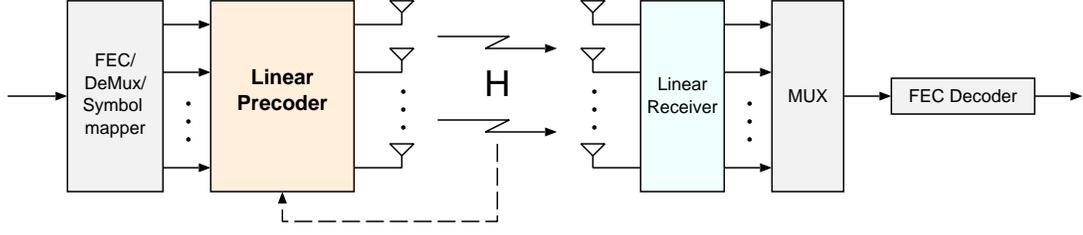
Figure 6.1. MIMO with linear precoder with receive-side equalization

6.2 System Model

The input-output system model for an $M \times N$ flat fading MIMO channel with a precoder matrix $\mathbf{T} \in \mathbb{C}^{M \times B}$ and a receiver equalizer $\mathbf{W} \in \mathbb{C}^{B \times N}$ is

$$\mathbf{y} = \mathbf{WHTx} + \mathbf{Wn}. \tag{6.1}$$

where B is the number of data streams, with $B \leq \min(M, N)$. In most wireless systems, the equalizer at the receiver is designed to equalize the compound channel ($\mathbf{HT}$) composed of the precoder and the channel (rather than designing the precoder for the equalized channel ($\mathbf{WH}$) although it is possible). In such case we have $M \geqslant N$ and we set $B = N$.

6.2.1 ZF Equalizer

The ZF equalizer is analyzed when operating together with various precoders, as follows.

Wiener Filter Precoding

The TxWF precoder is given by

$$\begin{aligned}
\mathbf{T} &= \beta \left(\mathbf{H}^H \mathbf{H} + \frac{N}{\rho} \mathbf{I} \right)^{-1} \mathbf{H}^H \\
&= \beta \mathbf{H}^H \left(\mathbf{H} \mathbf{H}^H + \frac{N}{\rho} \mathbf{I}_N \right)^{-1}
\end{aligned} \tag{6.2}$$

where (6.2) follows from [43, Fact 2.16.16] [1]. The scalar coefficient β is given in (5.70) and, similar to (5.35), it can be written as $\beta = 1/\sqrt{\eta}$

$$\eta = \text{tr}\left[\Lambda(\Lambda + N\rho^{-1}\mathbf{I})^{-2}\right] = \sum_{l=1}^{N} \frac{\lambda_l}{(\lambda_l + N\rho^{-1})^2}$$

The ZF equalizer for the precoder and the channel is given by

$$\mathbf{W}_{ZF} = (\mathbb{H}^H\mathbb{H})^{-1}\mathbb{H}^H \tag{6.3}$$

The composite channel $\mathbb{H}$ is given by

$$\mathbb{H} = \mathbf{HT}.$$

The received signal is given by

$$\mathbf{y} = \mathbf{W}_{ZF}\mathbf{HTx} + \mathbf{W}_{ZF}\mathbf{n}. \tag{6.4}$$

The filtered noise $\tilde{\mathbf{n}} = \mathbf{W}_{ZF}\mathbf{n}$ is is a complex Gaussian vector with zero-mean and covariance matrix $R_{\tilde{n}}$ given by

$$\begin{aligned}
R_{\tilde{n}} &= [\mathbb{H}^H\mathbb{H}]^{-1} \\
&= \left[(\mathbf{HH}^H + N\rho^{-1}\mathbf{I})^{-1}(\mathbf{HH}^H)^2(\mathbf{HH}^H + N\rho^{-1}\mathbf{I})^{-1}\right]^{-1} \\
&= \left[\mathbf{U}\Lambda(\Lambda + N\rho^{-1}\mathbf{I})^{-1}\mathbf{U}^H\mathbf{U}\Lambda(\Lambda + N\rho^{-1}\mathbf{I})^{-1}\mathbf{U}^H\right]^{-1} \\
&= \left[\mathbf{U}\Lambda^2(\Lambda + N\rho^{-1}\mathbf{I})^{-2}\mathbf{U}^H\right]^{-1}
\end{aligned}$$

where we have used the eigen decomposition $\mathbf{HH}^H = \mathbf{U}\Lambda\mathbf{U}^H$. The noise variance of the output stream k is therefore

$$R_{\tilde{n}}(k,k) = \sum_{l=1}^{N} \left(\frac{\lambda_l + N\rho^{-1}}{\lambda_l}\right)^2 |u_{kl}|^2 \tag{6.5}$$

---

[1]Let $\mathbf{A} \in \mathbb{C}^{n\times m}$ and $\mathbf{B} \in \mathbb{C}^{m\times n}$ then $(\mathbf{I}_n + \mathbf{AB})^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I}_m + \mathbf{BA})^{-1}$. This fact can be proved via Matrix Inversion Lemma.

where (6.5) follows in a similar manner as (5.29). We can compute the signal-to-noise ratio of the ZF filter output:

$$
\begin{aligned}
\gamma_k &= \frac{\rho \, \beta^2}{N \, R_{\tilde{n}}(k, k)} \\
&= \frac{\rho/N}{\sum_{j=1}^{N} \frac{\lambda_j}{(\lambda_j + N\rho^{-1})^2} \sum_{l=1}^{N} \left(\frac{\lambda_l + N\rho^{-1}}{\lambda_l}\right)^2 |u_{kl}|^2}.
\end{aligned}
\tag{6.6}
$$

Due to the complexity of (6.6) we proceed to bound the outage from above and below. The upper bound on outage is calculated as follows. Since $|u_{kl}| \leqslant 1$,

$$
\gamma_k \geqslant \frac{\rho/N}{\sum_{j=1}^{N} \frac{\lambda_j}{(\lambda_j + N\rho^{-1})^2} \sum_{l=1}^{N} \left(\frac{\lambda_l + N\rho^{-1}}{\lambda_l}\right)^2}
\tag{6.7}
$$

$$
= \frac{1/N}{\sum_{j=1}^{N} \frac{\rho^{1-\alpha_j}}{(\rho^{1-\alpha_j} + N)^2} \sum_{l=1}^{N} \left(\frac{\rho^{1-\alpha_l} + N}{\rho^{1-\alpha_l}}\right)^2}
\tag{6.8}
$$

$$
\triangleq \hat{\gamma}.
\tag{6.9}
$$

where we have substituted $\lambda_l = \rho^{-\alpha_l}$ in (6.8). Thus the outage probability is bounded as

$$
\begin{aligned}
P_{\text{out}} &= \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R \right) \\
&\leqslant \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \hat{\gamma}) \leqslant R \right) \\
&= \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \right)
\end{aligned}
\tag{6.10}
$$

Similarly to the analyses of earlier cases, we examine the SINR bound $\hat{\gamma}$ for different values of α_l. Define the set $\mathcal{B} = \{l \mid \alpha_l > 1\}$ and the event

$$
\mathcal{L} = \{|\mathcal{B}| = N\}
\tag{6.11}
$$

we have

$$
\begin{aligned}
P_{\text{out}} &\leqslant \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \right) \\
&= \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \,\Big|\, \mathcal{L} \right) \mathbb{P}(\mathcal{L}) + \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \,\Big|\, \bar{\mathcal{L}} \right) \mathbb{P}(\bar{\mathcal{L}}) \\
&\leqslant \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \mathcal{L} \right) + \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}} \right).
\end{aligned}
\tag{6.12}
\tag{6.13}
$$

To calculate the first term in (6.13), we evaluate $\hat{\gamma}$ when $\alpha_l \geqslant 1 \ \forall l$

$$\hat{\gamma} \doteq \frac{1/N}{\sum_{j=1}^{N} \rho^{1-\alpha_j} \sum_{l=1}^{N} \frac{1}{\rho^{2(1-\alpha_l)}}} \tag{6.14}$$

$$\dot{\geqslant} \frac{1/N}{\sum_{l=1}^{N} \frac{1}{\rho^{2(1-\alpha_l)}}} \tag{6.15}$$

$$\doteq \frac{1}{N} \rho^{2(1-\alpha_{\max})} = \frac{1}{N} \rho^2 \lambda_{\min}^2 \tag{6.16}$$

where (6.14) follows because $\rho^{1-\alpha_l} + N \doteq N$, (6.15) follows because $\sum_{j=1}^{N} \rho^{1-\alpha_j} \dot{\leqslant} 1$, and (6.16) follows because the sum in (6.15) is asymptotically dominated by the largest component.

We further bound the first term in (6.13)

$$\mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \mathcal{L}\right) \dot{\leqslant} \mathbb{P}\left(\frac{1}{N} \rho^2 \lambda_{\min}^2 \leqslant 2^{\frac{R}{N}}\right)$$

$$\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) \tag{6.17}$$

$$\doteq \rho^{-(M-N+1)} \tag{6.18}$$

where (6.17) is the same as (5.10), hence (6.18) follows.

To calculate the second term in (6.13), we evaluate $\hat{\gamma}$ when one or more $\alpha_l \leqslant 1$. Consider the the two summations in the denominator of (6.8). The first one can be asymptotically evaluated as

$$\sum_{j=1}^{N} \frac{\rho^{1-\alpha_j}}{(\rho^{1-\alpha_j} + N)^2} \doteq \sum_{\alpha_j < 1} \frac{1}{\rho^{1-\alpha_j}} + \sum_{\alpha_j > 1} \rho^{1-\alpha_j}$$

$$\doteq \begin{cases} \rho^{-(1-\alpha_{\max})} & |\bar{\mathcal{L}}| = N \\ \max(\rho^{-1+\alpha'}, \rho^{1-\alpha''}) \dot{\leqslant} \rho^{-(1-\alpha_{\max})} & 1 \leqslant |\bar{\mathcal{L}}| < N \end{cases} \tag{6.19}$$

where $\alpha' = \max_{\alpha_j < 1} \alpha_j$ and $\alpha'' = \min_{\alpha_j > 1} \alpha_j$ and Eq. (6.19) follows because

$$\min(\rho^{-1+\alpha'}, \rho^{1-\alpha''}) \dot{\leqslant} \rho^{-(1-\alpha_{\max})}.$$

The second summation in the denominator of (6.8) can be evaluated as follows

$$\sum_{l=1}^{N}\left(\frac{\rho^{1-\alpha_l}+N}{\rho^{1-\alpha_l}}\right)^2 \doteq \sum_{\alpha_l<1} 1 + \sum_{\alpha_l>1}\frac{1}{\rho^{2(1-\alpha_l)}}$$

$$\doteq \begin{cases} 1 & |\bar{\mathcal{L}}| = N \\ \rho^{-2(1-\alpha_{\max})} & 1 \leqslant |\bar{\mathcal{L}}| < N \end{cases} \tag{6.20}$$

We now use (6.19) and (6.20) to bound $\hat{\gamma}$

$$\hat{\gamma} \dot{\geqslant} \begin{cases} \rho^{1-\alpha_{\max}} = \rho\lambda_{\min} & |\bar{\mathcal{L}}| = N \\ \rho^{2-2\alpha_{\max}} = \rho^3\lambda_{\min}^3 & 1 \leqslant |\bar{\mathcal{L}}| < N \end{cases}$$

$$\triangleq \bar{\gamma} \tag{6.21}$$

We thus have

$$P_{out} \leqslant \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \middle| \bar{\mathcal{L}}\right)$$

$$\leqslant \mathbb{P}\left(\bar{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \middle| \bar{\mathcal{L}}\right)$$

$$< \mathbb{P}\left(\bar{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \middle| |\mathcal{B}| = 0\right) + \mathbb{P}\left(\bar{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \middle| 0 < |\mathcal{B}| < N\right)$$

$$\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) + \mathbb{P}\left(\lambda_{\min}^3 \leqslant \rho^{-3}\right)$$

$$\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right)$$

$$\doteq \rho^{-(M-N+1)}. \tag{6.22}$$

This concludes the calculation of outage upper bound. We now proceed with the outage lower bound.

Define the event $\mathcal{Q} = \{|a_{kl}| \geqslant \epsilon \quad \forall k, l\}$ where a_{kl} is the (k, l) entry of the unitary matrix U (c.f. equation (5.29)). Define

$$\breve{\gamma} = \frac{1/N}{\sum_{j=1}^{N}\frac{\rho^{1-\alpha_j}}{(\rho^{1-\alpha_j}+N)^2}\sum_{l=1}^{N}\left(\frac{\rho^{1-\alpha_l}+N}{\rho^{1-\alpha_l}}\right)^2\epsilon} \tag{6.23}$$

Notice that $\breve{\gamma} > \gamma$ because $|a_{kl}| \geqslant \epsilon \quad \forall k, l$.

The outage probability is bounded as

$$P_{\text{out}} = \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R \right)$$

$$\geqslant \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R \Big| \mathcal{Q} \right) \mathbb{P}(\mathcal{Q})$$

$$\geqslant \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \breve{\gamma}) \leqslant R \right) \mathbb{P}(\mathcal{Q}) \tag{6.24}$$

$$= \mathbb{P}\left(\breve{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \right) \mathbb{P}(\mathcal{Q}) \tag{6.25}$$

The probability $\mathbb{P}(\mathcal{Q}) = O(1)$, i.e. non-zero constant with respect to ρ. The proof is similar to the one in [7, Appendix A] and omitted here for brevity. We thus have

$$P_{\text{out}} \dot{\geqslant} \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \right)$$

$$= \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \mathcal{L} \right) \mathbb{P}(\mathcal{L}) \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}} \right) \mathbb{P}(\bar{\mathcal{L}})$$

$$\geqslant \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}} \right) \mathbb{P}(\bar{\mathcal{L}})$$

$$\dot{=} \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}} \right) \tag{6.26}$$

where (6.26) holds since $\mathbb{P}(\bar{\mathcal{L}}) \dot{=} O(1)$ as given by (5.55).

We further bound the outage probability by bounding $\hat{\gamma}$ as follows. Once again consider the two summations in the denominator of (6.23). For the first summation of (6.23), we have

$$\sum_{j=1}^{N} \frac{\rho^{1-\alpha_j}}{(\rho^{1-\alpha_j} + N)^2} \dot{=} \sum_{\alpha_j < 1} \frac{1}{\rho^{1-\alpha_j}} + \sum_{\alpha_j > 1} \rho^{1-\alpha_j}$$

$$\dot{=} \begin{cases} \rho^{-(1-\alpha_{\max})} & |\bar{\mathcal{L}}| = N \\ \max(\rho^{-1+\alpha'}, \rho^{1-\alpha''}) \dot{\geqslant} \rho^{1-\alpha_{\max}} & 1 \leqslant |\bar{\mathcal{L}}| < N \end{cases} \tag{6.27}$$

where the bound in the second line (6.27) is true because

$$\sum_{\alpha_j < 1} \frac{1}{\rho^{1-\alpha_j}} + \sum_{\alpha_j > 1} \rho^{1-\alpha_j} \geqslant \sum_{\alpha_j > 1} \rho^{1-\alpha_j} \dot{=} \rho^{1-\alpha_{\max}}$$

Using (6.19) and (6.27) to bound $\hat{\gamma}$ Substituting back in (6.23) gives:

$$\breve{\gamma} \dot{\lessgtr} \begin{cases} \rho^{1-\alpha_{\max}} = \rho\lambda_{\min} & |\bar{\mathcal{L}}| = N \\ \rho^{1-\alpha_{\max}} = \rho\lambda_{\min} & 1 \leqslant |\bar{\mathcal{L}}| < N \end{cases}$$
$$\triangleq \breve{\gamma} \tag{6.28}$$

Thus the outage bound in (6.26) can be then evaluated as we did for the upper bound

$$
\begin{aligned}
P_{out} &\leqslant \mathbb{P}\left(\hat{\gamma} \geqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}}\right) \\
&\leqslant \mathbb{P}\left(\breve{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}}\right) \\
&< \mathbb{P}\left(\breve{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| |\mathcal{B}| = 0\right)\mathbb{P}(|\mathcal{B}| = 0) + \mathbb{P}\left(\breve{\gamma} \leqslant 2^{\frac{R}{N}} - 1 \Big| \bar{\mathcal{L}}, 0 < |\bar{\mathcal{B}}| < N\right)\mathbb{P}(|\bar{\mathcal{L}}| < N) \\
&\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) O(1) \quad + \quad \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) O(1) \\
&\doteq \mathbb{P}\left(\lambda_{\min} \leqslant \rho^{-1}\right) \\
&\doteq \rho^{-(M-N+1)}.
\end{aligned}
$$

(6.29)

(6.30)

where (6.29) follows as a direct result of Lemma 5.3.1. From (6.22) and (6.30), we conclude that the diversity of MIMO system using TxWF precoder and ZF equalizer is

$$d^{WFP-ZF} = M - N + 1.$$

Regularized Zero Forcing Precoding

The ZF equalizer is given by (6.3) where the composite channel $\mathbb{H} = \mathbf{HT}$. The received signal to noise ratio of the k-th output symbol of the ZF filter as

$$
\begin{aligned}
\gamma_k &= \frac{\rho\,\beta^2}{N\,R_{\tilde{n}}(k,k)} \\
&= \frac{\rho/N}{\sum_{j=1}^{N}\frac{\lambda_j}{(\lambda_j+N)^2}\sum_{l=1}^{N}\left(\frac{\lambda_l+N}{\lambda_l}\right)^2|u_{kl}|^2}.
\end{aligned}
\tag{6.31}
$$

The process of obtaining lower and upper bound has many similarities with the developments of Section 6.2.1, therefore we omit many of the steps in the interest of brevity by referring to the previous developments.

We begin with the outage upper bound, which is developed in a manner similar to (6.10).

$$P_{\text{out}} = \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R\right)$$

$$\leqslant \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \hat{\gamma}) \leqslant R\right)$$

$$= \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1\right) \tag{6.32}$$

where

$$\hat{\gamma} = \frac{\rho/N}{\sum_{j=1}^{N} \frac{\lambda_j}{(\lambda_j + N)^2} \sum_{l=1}^{N} \left(\frac{\lambda_l + N}{\lambda_l}\right)^2}$$

$$= \frac{\rho/N}{\sum_{j=1}^{N} \frac{\rho^{-\alpha_j}}{(\rho^{-\alpha_j} + N)^2} \sum_{l=1}^{N} \left(\frac{\rho^{-\alpha_l} + N}{\rho^{-\alpha_l}}\right)^2}$$

$$\doteq \frac{\rho/N}{\sum_{j=1}^{N} \rho^{-\alpha_j} \sum_{l=1}^{N} \rho^{2\alpha_l}} \tag{6.33}$$

$$\dot{\geqslant} \frac{\rho/N}{\sum_{l=1}^{N} \rho^{2\alpha_l}}$$

$$\doteq \frac{\rho/N}{\rho^{2\alpha_{max}}}. \tag{6.34}$$

Thus the outage in (6.32) can be bounded as

$$P_{\text{out}} \leqslant \mathbb{P}\left(\hat{\gamma} \leqslant 2^{\frac{R}{N}} - 1\right)$$

$$\dot{\leqslant} \mathbb{P}\left(\frac{\rho/N}{\rho^{2\alpha_{max}}} \leqslant 2^{\frac{R}{N}} - 1\right)$$

$$\doteq \mathbb{P}(\lambda_{\min} \leqslant \rho^{-0.5})$$

$$\doteq \rho^{-\frac{1}{2}(M-N+1)}. \tag{6.35}$$

We now turn to the lower bound, which is obtained in the same manner as (6.26):

$$P_{\text{out}} = \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \gamma_k) \leqslant R\right)$$

$$\dot{\geqslant} \mathbb{P}\left(\sum_{k=1}^{N} \log(1 + \check{\gamma}) \leqslant R\right)$$

$$= \mathbb{P}\left(\check{\gamma} \leqslant 2^{\frac{R}{N}} - 1\right) \tag{6.36}$$

where

$$\check{\gamma} = \frac{\rho/N}{\sum_{j=1}^{N} \frac{\lambda_j}{(\lambda_j+N)^2} \sum_{l=1}^{N} \left(\frac{\lambda_l+N}{\lambda_l}\right)^2 \epsilon}$$

$$= \frac{\rho/N}{\sum_{j=1}^{N} \frac{\rho^{-\alpha_j}}{(\rho^{-\alpha_j}+N)^2} \sum_{l=1}^{N} \left(\frac{\rho^{-\alpha_l}+N}{\rho^{-\alpha_l}}\right)^2 \epsilon}$$

$$\doteq \frac{\rho/N}{\sum_{j=1}^{N} \rho^{-\alpha_j} \sum_{l=1}^{N} \epsilon \rho^{2\alpha_l}}$$

$$\leqslant \frac{\rho/N}{\rho^{-\alpha_j} \sum_{l=1}^{N} \epsilon \rho^{2\alpha_l}} \qquad \text{for arbitrary } j$$

$$\doteq \frac{\rho/N}{\epsilon \, \rho^{-\alpha_j} \rho^{2\alpha_{\max}}}$$

$$= \frac{\rho/N \lambda_{\min}^2}{\epsilon \, \lambda_j}$$

$$\triangleq \breve{\gamma}. \tag{6.37}$$

Let $C_1 = (2^{\frac{R}{N}} - 1) \, \epsilon \, N$, $C_2 = C_1 \xi$ where ξ is a fixed positive constant (independent of ρ), we have

$$P_{\text{out}} \dot{\geqslant} \mathbb{P}\left(\check{\gamma} \leqslant 2^{\frac{R}{N}} - 1\right)$$

$$\dot{\geqslant} \mathbb{P}\left(\breve{\gamma} \leqslant 2^{\frac{R}{N}} - 1\right)$$

$$\dot{\geqslant} \mathbb{P}\left(\frac{\rho \lambda_{\min}^2}{\lambda_j} \leqslant C_1\right)$$

$$\geqslant \mathbb{P}\left(\frac{\rho \lambda_{\min}^2}{\lambda_j} \leqslant C_1 \,\Big|\, \lambda_j \geqslant \xi\right) \mathbb{P}(\lambda_j \geqslant \xi)$$

$$\geqslant \mathbb{P}\left(\rho \lambda_{\min}^2 \leqslant C_2\right) \mathbb{P}(\lambda_j \geqslant \xi)$$

$$\doteq \mathbb{P}\left(\rho \lambda_{\min}^2 \leqslant C_2\right). \tag{6.38}$$

The exponential inequality (6.38) holds because $\mathbb{P}(\lambda_j \geqslant \xi) = O(1)$, as proved in Appendix 6.5.1. We thus conclude:

$$d^{RZFP-ZF} = \frac{1}{2}(M - N + 1).$$

Remark 6.2.1 *We note that the diversity of regularized zero-forcing precoder together with a zero-forcing equalizer can be fractional. To our knowledge this is the first instance of fractional diversity uncovered in the literature.*

Matched Filter Precoding

In this case, the composite channel is

$$\mathbb{H} = \mathbf{HT} = \beta\mathbf{HH}^H.$$

The noise correlation matrix is given by

$$R_{\tilde{n}} = [\mathbb{H}^H\mathbb{H}]^{-1} = \frac{1}{\beta^2}\left[(\mathbf{HH}^H)^2\right]^{-1} = \frac{1}{\beta^2}\left(\mathbf{U\Lambda^2U}^H\right)^{-1}.$$

Thus

$$R_{\tilde{n}}(k,k) = \frac{1}{\beta^2}\sum_{l=1}^{B}\frac{1}{\lambda_l^2}|u_{kl}|^2 \tag{6.39}$$

The precoder normalization factor $\beta = 1/\sqrt{\eta}$, where η is given by

$$\eta = \text{tr}\left[\mathbf{HH}^H\right] = \sum_{l=1}^{N}\lambda_l$$

The signal to noise ratio of the k-th symbol of the ZF filter is

$$\gamma_k = \frac{\rho}{N\,R_{\tilde{n}}(k,k)}$$
$$= \frac{\rho/N}{\sum_{j=1}^{N}\lambda_j\sum_{l=1}^{N}\frac{1}{\lambda_l^2}|u_{kl}|^2}. \tag{6.40}$$

Notice that the SINR γ_k in (6.40) is similar to the SINR γ_k of the RZF precoding system with ZF equalizer given by (6.31). The only difference is the term $\lambda_k + N$ which, when applying the transformation of $\lambda_k = \rho^{-\alpha_k}$, has no effect on the diversity analysis as detailed in the previous section. We then conclude that the diversity of the MIMO system applying MF precoder and ZF equalizer is the same as the diversity of the RZF precoder with ZF equalizer. Thus:

$$d^{MFP-ZF} = \frac{1}{2}(M - N + 1). \tag{6.41}$$

6.2.2 MMSE equalizer

The MMSE equalizer has better performance compared to ZF and is therefore widely popular. We investigate the diversity of MIMO systems that deploy different precoders at the transmitter and MMSE equalizer at the receiver.

MFTx Precoding

The MFTx precoder, $\mathbf{T}_{MFP}$, is given by (5.64). The MMSE equalizer for the precoded channel is given by

$$\mathbf{W}_{MMSE} = \left[\mathbb{H}^H \mathbb{H} + N\rho^{-1}\mathbf{I}\right]^{-1} \mathbb{H}^H \tag{6.42}$$

where $\mathbb{H} = \mathbf{H}\,\mathbf{T}_{MFP} = \beta_{MFP}\mathbf{HH}^H$ and β_{MFP} is given by (5.65).

The SINR at the output of the MMSE filter is given by [16]

$$\gamma_k = \frac{\rho}{N}\mathbf{h}_k\left[\mathbf{I} + \frac{\rho}{N}\mathbb{H}_k\mathbb{H}_k^H\right]^{-1}\mathbf{h}_k$$

$$= \frac{1}{\left[\mathbf{I} + \frac{\rho}{N}\mathbb{H}^H\mathbb{H}\right]^{-1}_{kk}} - 1 \tag{6.43}$$

where $\mathbb{H}_k$ is a submatrix of $\mathbb{H}$ obtained by removing the k-th column, $\mathbf{h}_k$.

The diversity analysis of the precoded system uses some results from the un-precoded MMSE MIMO equalizers [5], which we quote in the following lemma.

Lemma 6.2.1 *consider a quasi-static Rayleigh fading MIMO channel* $\bar{\mathbf{H}} \in \mathbb{C}^{M \times N}$ *($M \geqslant N$), the outage probability of the MMSE receiver satisfies*

$$P_{out} \doteq \mathbb{P}\left(tr(\mathbf{I} + \frac{\rho}{N}\bar{\mathbf{H}}^H\bar{\mathbf{H}})^{-1} \geqslant N2^{-\frac{R}{N}}\right) \tag{6.44}$$

$$= \mathbb{P}\left(\sum_{k=1}^{N} \frac{1}{1 + \frac{\rho}{N}\lambda_k'} \geqslant N2^{-\frac{R}{N}}\right) \tag{6.45}$$

$$\doteq \rho^{-d^{MMSE}} \tag{6.46}$$

where $\{\lambda_k'\}$ *are the eigenvalues of* $\bar{\mathbf{H}}$ *and* d^{MMSE} *is given by (5.75).*

Substituting $\lambda'_k = \rho^{-\alpha'_k}$, we have

$$\frac{1}{1 + \frac{\rho}{N}\lambda'_k} \doteq \begin{cases} \rho^{\alpha'_k - 1} & \alpha'_k < 1 \\ 1 & \alpha'_k > 1 \end{cases} \qquad (6.47)$$

thus the term $\frac{1}{1+\rho\lambda'_k/N}$ is either zero or one at high SNR, and therefore to characterize the sum in (6.45) at high SNR we count the number of ones, or equivalently the number of $\alpha'_k > 1$. Hence the outage probability reduces to [5]

$$P_{out} \doteq \mathbb{P}\left(\sum_{\alpha'_k > 1} 1 = \lceil N2^{-\frac{R}{N}} \rceil \right). \qquad (6.48)$$

Now we apply the matched filter precoder. Similarly to (6.44), the outage portability is given by

$$P_{out} \doteq \mathbb{P}\left(\text{tr}(\mathbf{I} + \frac{\rho}{N}\mathbb{H}\mathbb{H}^H)^{-1} \geqslant N2^{-\frac{R}{N}} \right) \qquad (6.49)$$

$$= \mathbb{P}\left(\sum_{k=1}^{N} \frac{1}{1 + \frac{\rho}{N\eta}\lambda_k^2} \geqslant N2^{-\frac{R}{N}} \right) \qquad (6.50)$$

where we have used $\mathbb{H}\mathbb{H}^H = \frac{1}{\eta}(\mathbf{H}\mathbf{H}^H)^2 = \frac{1}{\eta}\mathbf{U}\Lambda^2\mathbf{U}^H$ to obtain (6.50),and $\{\lambda_k\}$ are the eigenvalues of the Wishart matrix $\mathbf{H}\mathbf{H}^H$. The scaling factor $\eta = \text{tr}(\mathbf{H}\mathbf{H}^H) = \sum_{l=1}^{N}\lambda_l$.

We begin with a hypothetical precoder whose transmit power is not normalized, i.e., $\eta = 1$. The outage probability of this un-normalized precoder is similar to that of the MMSE receiver with no precoding at the transmitter, as given in (6.46), except that the eigenvalues are now squared. Thus similarly to (6.47), we have the exponential inequality

$$\frac{1}{1 + \frac{\rho}{N}\lambda_k^2} \doteq \begin{cases} \rho^{2\alpha_k - 1} & \alpha_k < 0.5 \\ 1 & \alpha_k > 0.5 \end{cases}. \qquad (6.51)$$

The analysis of [5] then follows and we have

$$d = \frac{1}{2}\left(\lceil N2^{-\frac{R}{N}} \rceil^2 + (M - N)\lceil M2^{-\frac{R}{N}} \rceil \right). \qquad (6.52)$$

We conclude that the un-normalized matched filter precoding with MMSE receiver results in 50% diversity loss compared to MMSE receiver with no transmit precoding.

For the normalized precoder, we begin with the outage probability in (6.50). Assume $\alpha_1 \geqslant \alpha_2 \cdots \geqslant \alpha_N$, the sum term in (6.50) is given by

$$
\begin{aligned}
\sum_{k=1}^{N} \frac{1}{1 + \frac{\rho}{N\eta}\lambda_k^2} &= \sum_{k=1}^{N} \frac{\eta}{\eta + \frac{\rho}{N}\lambda_k^2} \\
&= \sum_{k=1}^{N} \frac{\sum_l \rho^{-\alpha_l}}{\sum_l \rho^{-\alpha_l} + \frac{\rho}{N}\rho^{-2\alpha_k}} \\
&\doteq \sum_{k=1}^{N} \frac{\rho^{-\alpha_N}}{\rho^{-\alpha_N} + \rho^{1-2\alpha_k}}.
\end{aligned}
\tag{6.53}
$$

where we have used the fact that the $\sum_l \rho^{-\alpha_k}$ is dominated by the maximum element at high SNR. It is easy to see that the terms of (6.53) are either one or zero at high SNR, depending on whether $\rho^{-\alpha_N}$ asymptotically dominates $\rho^{1-2\alpha_k}$ or vice versa. These two cases are delineated with the threshold $\alpha_k \lesseqgtr 0.5 \max(1, \alpha_N+1)$, or, considering that α_N is positive, $\alpha_k \lesseqgtr 0.5(\alpha_N + 1)$. Thus at high SNR, the outage probability is evaluated by counting the ones

$$
\begin{aligned}
P_{out} &\doteq \mathbb{P}\left(\sum_{k=1}^{N} \frac{1}{1 + \frac{\rho}{N\eta}\lambda_k^2} \geqslant N2^{-\frac{R}{N}} \right) \\
&\doteq \mathbb{P}\left(\sum_{\alpha_k > 0.5\,(\alpha_N+1)} 1 \geqslant N2^{-\frac{R}{N}} \right) \\
&\doteq \mathbb{P}\left(\sum_{\alpha_k > 0.5\,(\alpha_N+1)} 1 = L \right)
\end{aligned}
\tag{6.54}
$$

where $L = \lceil N2^{-\frac{R}{N}} \rceil$. The conversion from inequality to equality in equation (6.54) follows from arguments developed in [5, Section III-A] .

Therefore, the outage probability is asymptotically evaluated by:

$$
P_{out} \doteq \int_{\mathcal{S}^+} \mathbb{P}(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}
\tag{6.55}
$$

where $\mathbb{P}(\boldsymbol{\alpha})$ is the joint distribution of the ordered $\alpha_1 \geqslant \cdots \geqslant \alpha_N$ and the region of integration is defined as $\mathcal{S}^+ = \mathcal{S} \cap \mathbb{R}^{N+}$, where $\mathcal{S}$ is given as follows:

- If $L = N$, then we seek the probability that $\alpha_k > \frac{1}{2}(\alpha_N + 1)$ for $k = 1, \ldots, N$, which implies $\alpha_N \in (1, \infty)$. Thus the integration region can be tightly represented as:

$$S = \left\{ \alpha_N > 1 \,, \, \min_{1 \le k < N} \alpha_k > 0.5(\alpha_N + 1) \right\}$$

- If $L < N$, then we seek the joint probability that $\alpha_k > \frac{1}{2}(\alpha_N + 1)$ for $k = 1, \ldots, L$ and $\alpha_k \le \frac{1}{2}(\alpha_N + 1)$ for $k = L + 1, \ldots, N$, implying $\alpha_N \in (0, 1)$. Thus the region of integration is represented as:

$$S = \left\{ \alpha_N < 1 \,, \, \min_{1 < k \le L} \alpha_k > 0.5(\alpha_N + 1) \,, \, \max_{L < k < N} \alpha_k < 0.5(\alpha_N + 1) \right\}$$

Using methods similar to [2] and [5, Eq (18) - (20)], exponential equality relations can be used to reduce the integrand to the following:

$$P_{out} \doteq \int_{S+} \prod_k \rho^{-(2k-1+M-N)\alpha_k} \, d(\boldsymbol{\alpha}) \tag{6.56}$$

First we consider $L = N$. The probability expression is evaluated by simply taking the integral over all variables except α_N, and then taking an integral over α_N.

$$P_{out} \doteq \int_{\alpha_N=1}^{\infty} \rho^{-(2N-1+M-N)\alpha_N} \prod_{k=1}^{N-1} \rho^{-(2k-1+M-N)(0.5+0.5\alpha_N)} d(\boldsymbol{\alpha}) \tag{6.57}$$

$$\doteq \prod_{k=1}^{N} \rho^{-(2k-1+M-N)}$$

$$= \rho^{\sum_{k=1}^{N} -(2k-1+M-N)} \tag{6.58}$$

$$= \rho^{-MN}. \tag{6.59}$$

When $L < N$, we repeat the same integration strategy.

$$P_{out} \doteq \int_{\alpha_N=0}^{1} \rho^{-(2N-1+M-N)\alpha_N} \prod_{l=L+1}^{N} \left(1 - \rho^{-(2l-1+M-N)(0.5+0.5\alpha_N)}\right)$$

$$\times \prod_{k=1}^{L} \rho^{-(2k-1+M-N)(0.5+0.5\alpha_N)} d(\boldsymbol{\alpha}) \tag{6.60}$$

$$\doteq \int_{\alpha_N=0}^{1} \rho^{-(2N-1+M-N)\alpha_N} \prod_{k=1}^{L} \rho^{-(2k-1+M-N)(0.5+0.5\alpha_N)} d(\boldsymbol{\alpha}) \tag{6.61}$$

$$\doteq \prod_{k=1}^{L} \rho^{-\frac{1}{2}(2k-1+M-N)}$$

$$= \rho^{\sum_{k=1}^{L} -\frac{1}{2}(2k-1+M-N)}$$

$$= \rho^{-\frac{1}{2}(L^2+(M-N)L)} \tag{6.62}$$

In deriving (6.60) and (6.61) we have used $\int_a^b \rho^{-c_k \alpha_k} d(\alpha_k) \doteq \rho^{-ac_k}$ [5]. Equations (6.59) and (6.62) show that the system exhibits two distinct diversity behaviors based on whether $L = \lceil N2^{-\frac{R}{N}} \rceil < N$. We can solve to find the boundary of the two regions $R = N \log \frac{N}{N-1}$. To summarize:

$$d^{MFP-MMSE} = \begin{cases} \frac{1}{2}\left(\lceil N2^{-\frac{R}{N}} \rceil^2 + (M-N)\lceil M2^{-\frac{R}{N}} \rceil\right) & R > N \log \frac{N}{N-1} \\ MN & \text{otherwise} \end{cases}. \tag{6.63}$$

Remark 6.2.2 *The outcome is interesting for its practical implications: An MMSE receiver working with matched-filter precoding will suffer a significant diversity loss compared to an MMSE receiver without precoding, except for very low rates corresponding to $R < N \log \frac{N}{N-1}$, where the combination of MMSE receiver with matched filter precoding has exactly the same diversity as the MMSE receiver alone.*

Remark 6.2.3 *Recall that $R = N \log \frac{N}{N-1}$ is exactly the same threshold below which matched filter precoding (without receiver-side equalization) achieves full diversity.*

WFTx Precoding

Using the Wiener filter precoding at the receiver results in the composite channel

$$\mathbb{H} = \mathbf{HT} = \beta \mathbf{HH}^H (\mathbf{HH}^H + \rho^{-1} N \mathbf{I})^{-1}.$$

Using the eigen decomposition $\mathbf{HH}^H = \mathbf{U\Lambda U}^H$, it can be shown that

$$\mathbb{H}^H \mathbb{H} = \beta^2 \mathbf{U}(\Lambda + \rho^{-1} N \mathbf{I})^{-2} \Lambda^2 \mathbf{U}^H \tag{6.64}$$

Similar to the case of MF precoder with MMSE receiver, the outage probability of WF precoder with MMSE receiver is given by (c.f. (6.49))

$$P_{out} \doteq \mathbb{P}\left(\text{tr}(\mathbf{I} + \frac{\rho}{N}\mathbb{H}\mathbb{H}^H)^{-1} \geqslant N 2^{-\frac{R}{N}} \right)$$

$$= \mathbb{P}\left(\sum_{k=1}^{N} \frac{1}{1 + \frac{\rho}{N\eta}\hat{\lambda}_k} \geqslant N 2^{-\frac{R}{N}} \right) \tag{6.65}$$

where $\{\hat{\lambda}_k\}$ are the eigenvalues of $\mathbb{H}^H \mathbb{H}$ and η is the scale factor. Using (6.64), $\{\hat{\lambda}_k\}$ are given by

$$\hat{\lambda}_k = \frac{\lambda_k^2}{(\lambda_k + \rho^{-1}N)^2}, \quad k = 1, \dots, N \tag{6.66}$$

The scale factor η is calculated as in (5.35)

$$\eta = \sum_{l=1}^{N} \frac{\lambda_l}{(\lambda_l + \rho^{-1}N)^2}.$$

Thus the outage probability can be written as

$$P_{out} \doteq \mathbb{P}\left(\sum_{k=1}^{N} \gamma_k \geqslant N 2^{-\frac{R}{N}} \right) \tag{6.67}$$

where

$$\gamma_k \triangleq \frac{1}{1 + \frac{\rho}{N\eta}\hat{\lambda}_k} = \frac{\rho^{-1}\eta}{\rho^{-1}\eta + \frac{1}{N}\hat{\lambda}_k} = \frac{\rho^{-1}\eta}{\rho^{-1}\eta + \upsilon_k}$$

where we define $v_k = \frac{1}{N}\hat{\lambda}_k$. We now proceed to express both $\rho^{-1}\eta$ and v_k in terms of $\{\alpha_k\}$, the exponential orders of $\{\lambda_k\}$.

$$\rho^{-1}\eta = \sum_{l=1}^{N} \frac{\rho^{-1}\lambda_l}{(\rho^{-1}\lambda_l + N)^2} = \sum_{l=1}^{N} \frac{\rho^{1-\alpha_l}}{(\rho^{1-\alpha_l} + N)^2}$$

$$\doteq \sum_{\alpha_l>1} \rho^{1-\alpha_l} + \sum_{\alpha_l<1} \rho^{\alpha_l-1} \tag{6.68}$$

observe that all the terms in (6.68) have negative exponent. Using (6.66),

$$v_k = \frac{1}{N} \frac{\rho^{-2\alpha_k}}{(\rho^{-\alpha_k} + \rho^{-1}N)^2}$$

$$= \frac{1}{N} \frac{\rho^{2(1-\alpha_k)}}{(\rho^{1-\alpha_k} + N)^2}$$

$$\doteq \begin{cases} 1 & \alpha_k < 1 \\ \rho^{2(1-\alpha_k)} & \alpha_k > 1 \end{cases}. \tag{6.69}$$

From (6.68) and (6.69), we see that when $\alpha_k < 1$ then $v_k + \rho^{-1}\eta \doteq v_k \doteq 1$. On the other hand, when $\alpha_k > 1$ then

$$v_k + \rho^{-1}\eta \doteq \rho^{2(1-\alpha_k)} + \sum_{\alpha_l>1} \rho^{1-\alpha_l} + \sum_{\alpha_l<1} \rho^{\alpha_l-1}$$

$$= \rho^{2(1-\alpha_k)} + \rho^{1-\alpha_k} + \sum_{\substack{\alpha_l>1 \\ l\neq k}} \rho^{1-\alpha_l} + \sum_{\substack{\alpha_l<1 \\ l\neq k}} \rho^{\alpha_l-1}$$

$$\doteq \rho^{1-\alpha_k} + \sum_{\substack{\alpha_l>1 \\ l\neq k}} \rho^{1-\alpha_l} + \sum_{\substack{\alpha_l<1 \\ l\neq k}} \rho^{\alpha_l-1} \tag{6.70}$$

$$\doteq \rho^{-1}\eta \tag{6.71}$$

where (6.70) follows because $\alpha_k > 1$. Thus we have

$$\gamma_k = \frac{\rho^{-1}\eta}{\rho^{-1}\eta + v_k} \doteq \begin{cases} \rho^{-1}\eta & \alpha_k < 1 \\ 1 & \alpha_k > 1 \end{cases} \tag{6.72}$$

and $\rho^{-1}\eta$ has negative exponent thus vanishes at high SNR.

Observe that (6.72) is similar to (6.47) which corresponds to the case of the MMSE-only system (i.e. with no precoding). Thus substituting (6.72) in the outage probability (6.67) and repeating the same analysis of the MMSE-only system as in [5], we conclude that the diversity of the MMSE receiver when using WFTx precoding is the same as the diversity of the MMSE receiver with no linear precoding, which is given by (5.75).

RZF Precoding

Using the Regularized Zero Forcing precoding at the receiver results in the composite channel

$$\mathbb{H} = \mathbf{HT} = \beta \mathbf{HH}^H (\mathbf{HH}^H + c\,\mathbf{I})^{-1}.$$

where c is a fixed constant, $\beta = 1/\eta$ and η is given by (5.35)

$$\eta = \sum_{l=1}^{N} \frac{\lambda_l}{(\lambda_l + c)^2} = \sum_{l=1}^{N} \frac{\rho^{-\alpha_l}}{(\rho^{-\alpha_l} + c)^2}. \tag{6.73}$$

Similar to (6.65), the outage probability of RZF precoder with MMSE receiver is given by

$$P_{out} \doteq \mathbb{P}\left(\sum_{k=1}^{N} \gamma_k \geqslant N 2^{-\frac{R}{N}} \right)$$

and

$$\gamma_k \triangleq \frac{\eta}{\eta + \frac{\rho}{N} \bar{\lambda}_k}$$

where $\{\bar{\lambda}_k\}$ are the eigenvalues of $\mathbb{H}^H \mathbb{H}$ given by

$$\bar{\lambda}_k = \frac{\lambda_k^2}{(\lambda_k + c)^2} = \frac{\rho^{-2\alpha_k}}{(\rho^{-\alpha_k} + c)^2}, \quad k = 1, \dots, N \tag{6.74}$$

Notice that at high SNR we have

$$\eta \doteq \sum_{l=1}^{N} \frac{\rho^{-\alpha_l}}{c^2}$$

$$\bar{\lambda}_k \doteq \frac{\rho^{-2\alpha_k}}{c^2}.$$

Thus the SINR is given by (c.f. (6.53))

$$\gamma_k \doteq \frac{\sum_{l=1}^{N} \rho^{-\alpha_l}}{\sum_{l=1}^{N} \rho^{-\alpha_l} + \rho^{-2\alpha_k}} \doteq \frac{\rho^{-\alpha_N}}{\rho^{-\alpha_N} + \rho^{1-2\alpha_k}},$$

$$k = 1, \ldots, N$$

which are the same terms as in (6.53), implying that the outage probability of the MMSE receiver working with the regularized zero-forcing precoder is asymptotically the same as the outage probability of the MMSE receiver working with the matched filter precoder. This means:

$$d^{RZFP-MMSE} = d^{MFP-MMSE}$$

$$= \begin{cases} \frac{1}{2}\left(\lceil N2^{-\frac{R}{N}}\rceil^2 + (M-N)\lceil M2^{-\frac{R}{N}}\rceil\right) & R > N \log \frac{N}{N-1} \\ MN & \text{otherwise} \end{cases} \tag{6.75}$$

6.3 Simulation Results

This section produces numerical results for the outage probabilities of ZF, regularized ZF (RZF), matched filter (MF) and Wiener precoding systems under ZF and MMSE equalization. Figure 6.2 shows outage probabilities for a 2×2 and a 3×2 MIMO system with matched filter precoding and ZF equalization. The observed diversity values are consistent with Eq. (6.41). Figure 6.3 shows outage probabilities for a 2×2 and a 3×2 MIMO system with Wiener filter precoding and ZF equalization. Figure 6.4 and Figure 6.5 show outage probabilities for a 2×2 and a 3×3 MIMO system, respectively, with Wiener filter precoding and MMSE equalization. The diversity for the 3×3 system is the same as the diversity of the Wiener filtering precoding-only (c.f. Figure 5.3).

Figure 6.6 shows the outage probability of a 2×2 MIMO system with matched filter precoding and MMSE equalization, which is consistent with Eq. (6.63). We also plot the outage probability of the MMSE MIMO equalizer (without any precoding) for comparison.
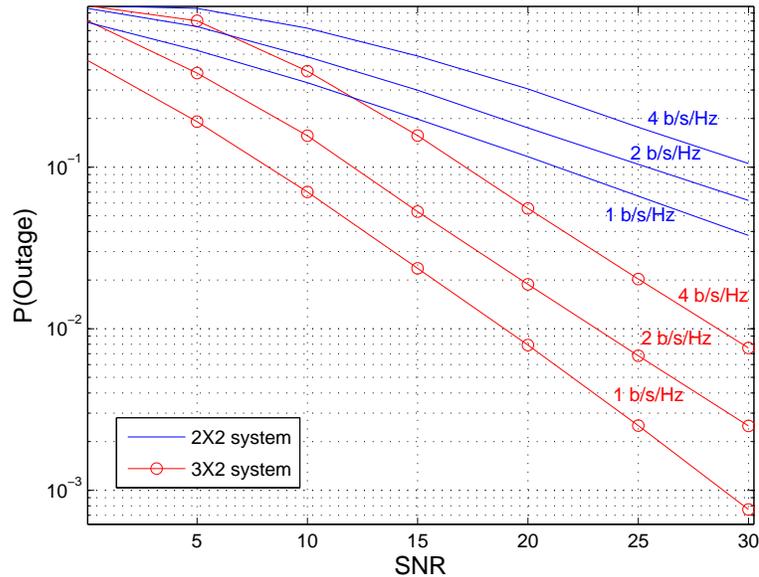
Figure 6.2. MIMO system with matched filtering precoding and ZF equalization for rates (left to right): $R = 1, 2,$ and 4 b/s/Hz. The diversity is $d = 0.5$ for a 2×2 MIMO system and $d = 1$ for a 3×2 MIMO system.

6.4 Conclusion

This chapter obtains various results for the diversity in the presence of *both* precoding and equalization.

6.5 Appendix

6.5.1 Proof of $\mathbb{P}(\lambda_l \geqslant \xi) = O(1)$ for any l

Define a Wishart matrix $\mathbf{W}$ using the Gaussian matrix $\mathbf{H}$.

$$
\mathbf{W} = \begin{cases} \mathbf{H}\,\mathbf{H}^H & M > N \\ \mathbf{H}^H \mathbf{H} & N \leqslant N \end{cases} .
$$

Let $n = \max(M, N)$ and $m = \min(M, N)$. The matrix $\mathbf{W}$ is $m \times m$ random non-negative definite that has real, non-negative eigenvalues with $\lambda_1 \geqslant \cdots \geqslant \lambda_m 0$. The joint density of
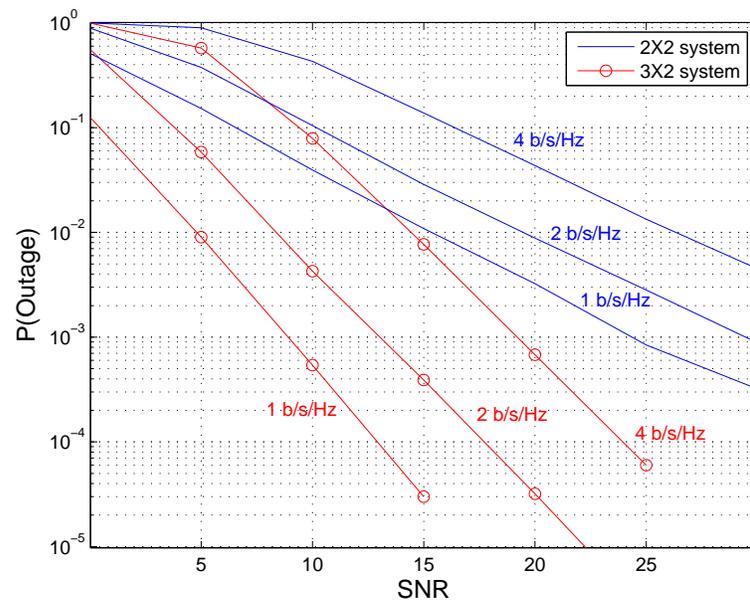
Figure 6.3. Outage probability of MIMO system with Wiener filtering precoding and ZF equalization for rates (left to right): $R = 1, 2$, and 4 b/s/Hz. The diversity is $d = 1$ for a 2×2 MIMO system and $d = 2$ for a 3×2 MIMO system.
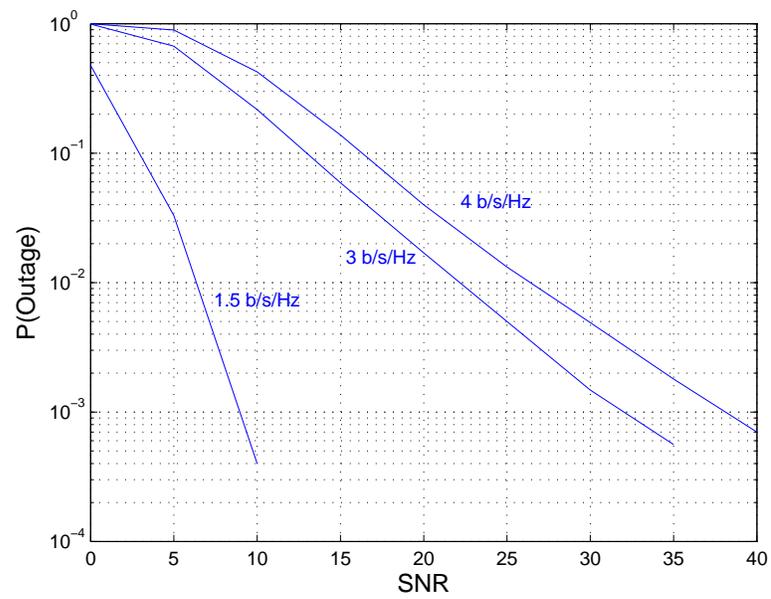
Figure 6.4. 2X2 MIMO system with Wiener filtering precoding and MMSE equalization for rates (left to right): $R = 1.5, 3$, and 4 b/s/Hz. The diversity $d = 1$ for $R = 3$ and 4 b/s/Hz, and the diversity $d = 4$ for $R = 1.5$ b/s/Hz
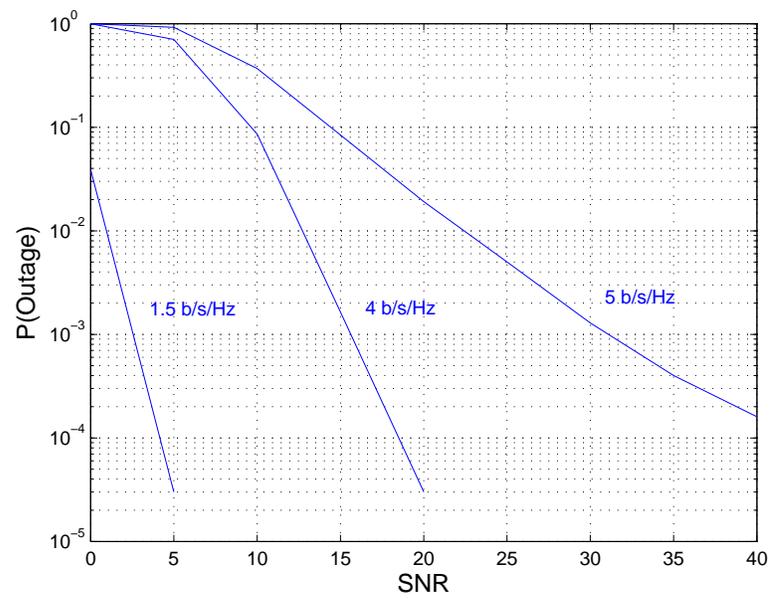
Figure 6.5. 3X3 MIMO system with Wiener filtering precoding and MMSE equalization for rates (left to right): $R = 1.5, 4$, and 5 b/s/Hz. The diversity is the same as the diversity of Wiener precoding without MMSE equalization which is given by (5.75).
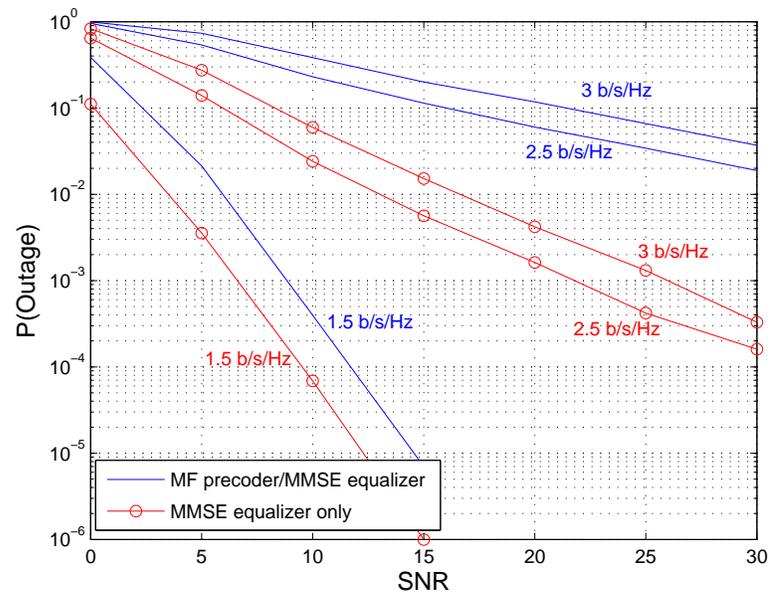
Figure 6.6. 2X2 MIMO system with MF precoding and MMSE equalization system for rates (left to right): $R = 1.5, 2.5,$ and 3 b/s/Hz. The diversity can be easily verified from Eq. (6.63).

the ordered eigenvalues is [17]

$$f(\boldsymbol{\lambda}) = K_{m,n}^{-1} e^{-\sum_i \lambda_i} \prod_i \lambda_i^{n-m} \prod_{i<j} (\lambda_i - \lambda_j)^2. \tag{6.76}$$

Thus the marginal distribution of λ_l is given by [17]

$$f_{\boldsymbol{\lambda}_l}(\lambda_l) = \int \dots \int f(\boldsymbol{\lambda}) \, d\lambda_2 \dots d\lambda_m$$

$$= \frac{1}{m} \sum_{i=1}^{m} \varphi_i(\lambda_l)^2 \lambda_l^{n-m} e^{-\lambda_1}$$

where

$$\varphi_{k+1}(\lambda) = \left[\frac{k!}{(k+n-m)!} \right]^{1/2} L_k^{n-m}(\lambda), \quad k = 0, \dots, m-1$$

where $L_k^{n-m}(x) = \frac{1}{k!} e^x x^{m-n} \frac{d^k}{dx^k}(e^{-x} x^{n-m+k})$ (with $L_0 = 1$) is the associated Laguerre polynomial of order k.

We now compute $\mathbb{P}(\lambda_l \geqslant \xi)$,

$$\mathbb{P}(\lambda_l \geqslant \xi) = \int_\xi^\infty \frac{1}{m} \sum_{i=1}^m \varphi_i(\lambda_l)^2 \lambda_1^{n-m} e^{-\lambda_l} d\lambda_l$$

$$\geqslant \int_\xi^\infty \frac{1}{m} \varphi_1(\lambda_l)^2 \lambda_l^{n-m} e^{-\lambda_l} d\lambda_l$$

$$= \int_\xi^\infty \frac{1}{m(n-m)!} \lambda_l^{n-m} e^{-\lambda_l} d\lambda_l$$

$$= \frac{1}{m(n-m)!} \left(-e^{-\lambda_l} \lambda_l^{n-m} - e^{-\lambda_l} \sum_{k=1}^{n-m} n(n-1)\dots(n-k+1)\lambda^{n-m-k} \right) \Big|_\xi^\infty$$

$$\tag{6.77}$$

$$= \frac{1}{m(n-m)!} \left(e^{-\xi} \xi^{n-m} + e^{-\xi} \sum_{k=1}^{n-m} n(n-1)\dots(n-k+1)\xi^{n-m-k} \right) \tag{6.78}$$

where (6.77) follows from [27, Section 2.32]. The right hand side of Eq. (6.78) is a non-zero constant bounded away from zero. This concludes the proof.

CHAPTER 7
SINGLE-CARRIER EQUALIZER WITH MULTI-ANTENNA TRANSMIT
DIVERSITY

7.1   Introduction

Single-Carrier Frequency Domain Equalization (SC-FDE) is an alternative to OFDM that avoids several OFDM drawbacks, including peak-to-average power ratio and the high sensitivity to carrier frequency offset [44]. SC-FDE has been adopted for the LTE uplink [44, 45].

In this chapter, we analyze the performance of SC-FDE in conjunction with either cyclic delay diversity (CDD) or Alamouti signaling, fully characterizing the diversity as a function of transmission-block length, data rate, channel memory, and number of antennas. In the process, we obtain a threshold rate (as a function of data-block length, channel memory, and number of antennas) below which the full spatial-temporal diversity is achieved, while at higher rates the diversity of both schemes diminishes, albeit not in quite the same way. Our analysis shows that at high rates the CDD diversity degenerates to the diversity of the SISO SC-FDE, while Alamouti signaling provides twice the diversity of SISO SC-FDE.

We find that beyond a certain rate threshold in either CDD or Alamouti signaling, an increase in transmission rate can reduce the diversity, but this diversity can be recovered by increasing the FFT block length. Specifically, in this operating regime, the diversity can be maintained if every additional bit/s/Hz of transmission rate is accompanied by a doubling of FFT block length. Naturally the block length cannot exceed the coherence time of the channel, therefore equalizer performance is in practice also limited by the coherence time.

A brief survey of related literature is as follows. It has been known that SC-FDE in single-antenna (SISO) systems displays a diversity that is a function of data rate and transmission block length (hence the FFT size) [15]. The behavior of SC-FDE has also been analyzed

in multi-stream (BLAST-type) MIMO systems, where its diversity multiplexing tradeoff (DMT) and bounds on its diversity have been obtained [46]. Al-Dhahir [20] proposed the Alamouti SC-FDE, but [20] only went so far as to show that the effective channel gain of Alamouti SC-FDE is a sum of two independent components, which only suggests that the diversity is at least two. The present work conclusively settles the question of the diversity of Alamouti SC-FDE.

Design rules are provided in [47] for achieving maximum diversity gains with linearly pre-coded OFDM but it requires ML decoding. Tepedelenlioglu [31] showed that linear equalizers achieve the maximum multipath diversity in linearly precoded OFDM systems. The zero-padded SC system with linear equalization was analyzed in [48] where it was shown that the full diversity is achievable by ZF equalizer. Muquet *et al.* [22] compared the performance of ZP-OFDM and CP-OFDM. Coded OFDM (COFDM) schemes were considered in [49], showing that COFDM achieves the maximum channel diversity with ML decoding. It was shown that the zero-padded and cyclic-prefix single-carrier system are special cases of the COFDM of [49] and thus achieve the maximum diversity with ML.

This chapter is organized as follows. Section 7.2 provides the system model for cyclic-prefix transmission and reviews the recent result for the SC MMSE-FDE receiver diversity. Section 7.3 provides the performance analysis for the CDD systems. Section 7.4 provides the performance analysis for the Alamouti (orthogonal-space time coded) systems. Section 7.5 provides simulations that illuminate our results.

7.2 Cyclic-Prefix Transmission

7.2.1 System Model

We consider a frequency selective quasi-static wireless fading channel. The equivalent base-band model for this inter-symbol interference (ISI) channel is given by a multipath model with ν paths. The channel vector is denoted by $\mathbf{h} = [h_0, \ldots, h_\nu]$ and the channel coefficients are assumed independent and identically distributed $\sim \mathcal{CN}(0, 1)$. We assume a block-fading

model where the channel is fixed over the transmission block. To remove the inter-block interference at the receiver, a cyclic-prefix (CP) with length at least ν is inserted at the beginning of each data-block of length L. The CP also transforms linear convolution into circular convolution and thus permits channel diagonalization. The input-output system model for a block transmission scheme with length-ν CP is

$$\mathbf{y} = \sqrt{\rho}\,\mathbf{H}\mathbf{x} + \mathbf{n} \tag{7.1}$$

where $\mathbf{x} \in \mathbb{C}^{(L+\nu)\times 1}$, $\mathbf{y} \in \mathbb{C}^{L\times 1}$ and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ denote the transmitted, received and noise vectors respectively. The noise vector is assumed white, Gaussian, and zero-mean with covariance matrix $\sigma_n^2 \mathbf{I}$ and ρ is the transmitted signal power. Without loss of generality we assume $\sigma_n^2 = 1$. The matrix $\mathbf{H} \in \mathbb{C}^{L\times(L+\nu)}$ is the convolution channel matrix. The linear data extension operation that maps the data block, denoted by $\bar{s}$, of length L to the transmitted vector of length $L+\nu$ can be expressed by $\mathbf{x} = \mathbf{U}_{cp}\,\bar{s}$ where $\mathbf{U}_{cp}$ is the CP matrix given by

$$\mathbf{U}_{cp} = \begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_\nu & \mathbf{0}_{\nu\times(L-\nu)} \end{bmatrix}.$$

Hence the model in (7.1) is equivalent to:

$$\mathbf{y} = \sqrt{\rho}\,\mathbf{H}\mathbf{U}_{cp}\bar{s} + \mathbf{n} = \sqrt{\rho}\,\mathbf{H_e}\bar{s} + \mathbf{n} \tag{7.2}$$

where $\mathbf{H_e} = \mathbf{H}\mathbf{U}_{cp}$ is the equivalent $L \times L$ circulant channel matrix. given by

$$\mathbf{H_e} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\nu & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_1 & h_2 & \cdots & h_\nu & 0 & \cdots & h_0 \end{bmatrix}. \tag{7.3}$$

Note that $\mathbf{H_e}$ has eigen decomposition $\mathbf{H_e} = \mathbf{Q}^H \Lambda \mathbf{Q}$ where $\mathbf{Q}$ is the unitary discrete Fourier transform matrix. The diagonal elements of Λ are given by [15]

$$\lambda_k = \sum_{i=0}^{\nu} h_i e^{-j\frac{2\pi i(k-1)}{L}} \quad \text{for } k = 1, \ldots, L.$$

In single-carrier frequency domain equalizer (SC-FDE), the DFT/IDFT operation is performed at the receiver. The DFT/IDFT in the SC-FDE diagonalizes the channel thus a single-tap equalizer can be used, reducing the complexity of equalization. The DFT-domain version of Equation (7.2) is

$$\mathbf{Y} = \mathbf{Q}\,\mathbf{y} = \sqrt{\rho}\,\Lambda\mathbf{S} + \mathbf{N} \tag{7.4}$$

where $\mathbf{S}$, $\mathbf{Y}$ and $\mathbf{N}$ are the DFT of the transmitted, received and noise vectors respectively.

Assuming perfect channel state information at the receiver (CSIR), the linear zero-forcing (ZF) and MMSE equalizers are given by [13]:

$$\mathbf{W} = (c\,\rho^{-1}\mathbf{I} + \mathbb{H}^H\mathbb{H})^{-1}\mathbb{H}^H \tag{7.5}$$

where the constant $c = 1$ for MMSE equalizer and $c = 0$ for ZF equalizer. The matrix $\mathbb{H}$ is the channel to be equalized. The corresponding unbiased decision-point SINR is

$$\gamma_k = \frac{1}{(c\,\mathbf{I} + \rho\mathbb{H}^H\mathbb{H})^{-1}_{kk}} - c \qquad k = 1, \ldots, L.$$

For completeness we mention the definition of the diversity gain

$$d \triangleq -\lim_{\rho \to \infty} \frac{\log P_{\mathrm{e}}}{\log \rho} \tag{7.6}$$

and the outage diversity

$$d_{out} \triangleq -\lim_{\rho \to \infty} \frac{\log P_{out}}{\log \rho} \tag{7.7}$$

where P_{e} is the pairwise error probability, P_{out} is the outage probability given by $P_{out} \triangleq \mathbb{P}(I(\mathbf{x}; \mathbf{y}) < R)$.

7.2.2 The Diversity of the MMSE Receiver in SC-FDE System

In [15] the linear MMSE receiver is analyzed for SC-FDE, a result that we shall refer to later in this chapter and therefore we review it here briefly. In [15] the received signal after equalization is given by

$$\tilde{\mathbf{y}} = \mathbf{W}\mathbf{y} = \sqrt{\rho}\mathbf{W}\mathbf{H}_{\mathbf{e}}\bar{s} + \mathbf{W}\mathbf{n}. \tag{7.8}$$

and subject to this model it was shown that MMSE SC-FDE can achieve full diversity for certain values of block length and operating rate R b/s/Hz. The process of developing this result was as follows.

The analysis performed in [15] consists of two main steps. The system outage diversity is first characterized. Then lower and upper bounds on the error probability via outage are provided. It is shown in [15] that these two bounds are tight and thus the diversity is fully characterized. This two-steps approach was first proposed in [2] due to the intractability of the direct pairwise error probability (PEP) analysis for many MIMO architectures.

The diversity of the MMSE SC-FDE is [15]

$$
d_{out} = \begin{cases} \nu + 1 & R \leq \log \frac{L}{\nu} \\ \lfloor 2^{-R}L \rfloor + 1 & R > \log \frac{L}{\nu} \end{cases}
\tag{7.9}
$$

7.3 Cyclic-Delay Diversity

One common transmit diversity technique used for single carrier and multicarrier systems is antenna delay diversity, which can take the form of time delay, cyclic delay and phase delay [50, 51]. Among them, cyclic delay diversity (CDD) is more widely adopted for single carrier and multicarrier applications as CDD can be applied to any number of transmit antennas without any rate loss or change in the receiver structure [51–53]. In this section we show that linear MMSE receivers can achieve the maximal spatio-temporal diversity provided that the equalizer and the cyclic delay taps are properly designed.

System Model

Consider a MISO system with M transmit antennas and a block fading channel model where the channel remains unchanged during the block transmission. The channel impulse response from the transmit antenna i to the received antenna is given by $\mathbf{h}_i = [h_{i,0}, \ldots, h_{i,\nu_i}]$ with channel memory length denoted by ν_i. We also define $\nu = \max_i \nu_i$. We adopt the system model of [51]. The model is shown in Figure 7.1 which displays the front end of a single
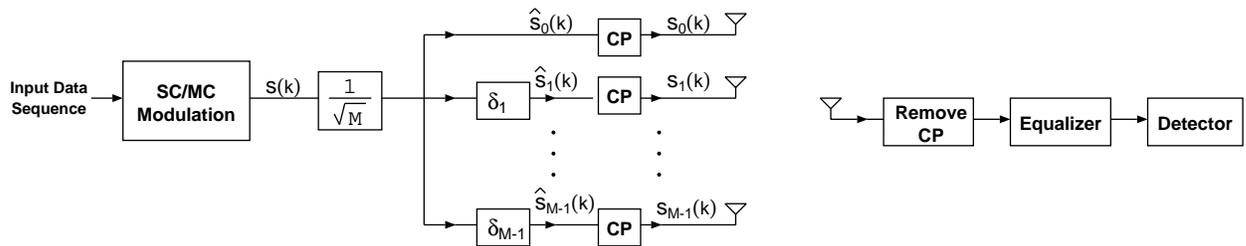
Figure 7.1. Single-carrier and multicarrier MISO system with transmitter sided CDD scheme and the proposed MMSE receiver.

carrier and multicarrier MISO system with CDD. In vector form, the received signal can be written as

$$\mathbf{y} = \sum_{i=0}^{M-1} \sqrt{\rho}\, \mathbf{H}_i\, \hat{\bar{s}}_i + \mathbf{n} \tag{7.10}$$

where $\mathbf{H}_i$ is an $L \times L$ circulant channel matrix whose first row is $[h_{i,0}, \ldots, h_{i,\nu_i}, 0, \ldots, 0]$, $\hat{\bar{s}}_i$ is the $L \times 1$ transmitted data-block (without the CP) from transmit antenna i. CDD converts the MISO channel into a SISO channel with increased channel selectivity. The model can be written as [52]

$$\mathbf{y} = \sqrt{\rho}\, \mathbf{H}_{\text{cir}}\, \bar{s} + \mathbf{n} = \sqrt{\rho}\, \mathbf{Q}^H \Lambda \mathbf{Q}\, \bar{s} + \mathbf{n} \tag{7.11}$$

where $\mathbf{H}_{\text{cir}}$ is $L \times L$ circulant matrix, $\bar{s}$ is the $L \times 1$ modulated symbols (cf. Figure 7.1), $\mathbf{Q}$ is the $L \times L$ normalized DFT matrix, and Λ is a diagonal matrix whose diagonal entries are the DFT point of the first row of $\mathbf{H}_{\text{cir}}$ which are given by $[\hat{h}(0), \ldots, \hat{h}(L-1)]$, and

$$\hat{h}(l) \triangleq \frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} h_i(l - \delta_i)_{\text{mod } L}. \tag{7.12}$$

The selection of the delay samples $\{\delta_i\}$ and its impact on the data rate, the signal-to-interference-and-noise ratio (SINR) and maximum achievable diversity is studied in [52,54]. While the CP length is independent of δ_i, it must be no less than the maximum channel delay spread ν [54]. Also for the receiver to exploit the full diversity the delays can be chosen as $\delta_i > \delta_{i-1} + \nu$ [52] or simply $\delta_i = i\delta$ with $\delta > \nu$ [53].

Diversity Analysis of MMSE Receiver

We first consider the case where $\nu_i = \nu = 0 \, \forall i$ (i.e. flat MISO channel) and the symbol delays $\delta_i = i$. In this case, the system model is equivalent to a SISO ISI channel under CP transmission. If the equalizer is designed according to $\mathbf{H}_{\mathrm{cir}}$, it is known that in the SISO ISI CP transmission a rate-dependent diversity is observed [15], and due to equivalence of channel models this result can be directly lifted to the flat MISO CDD system. This result will be extended to the general case of the multipath MISO channel under CDD.

For a flat channel ($\nu_i = 0 \, \forall i$) and the delays $\delta_i = i$, it can be shown that the first row of the $L \times L$ circulant channel matrix $\mathbf{H}_{\mathrm{cir}}$ is $[h_0, h_1, \cdots, h_{M-1}, 0, \cdots, 0]$ where the channel entries $\{\hat{h}(l)\}$ are given by (7.12). The equalized signal is

$$\tilde{\mathbf{y}} = \sqrt{\rho} \, \mathbf{W} \, \mathbf{H}_{\mathrm{cir}} \, \bar{s} + \mathbf{W} \mathbf{n}. \tag{7.13}$$

The system model in (7.13) is equivalent to (7.8) and we have the following lemma.

Lemma 7.3.1 *Consider the $M \times 1$ MISO flat-fading channel. The diversity of the MMSE receiver under* uncoded *CDD transmission and L data-blocks is given by*

$$d = \begin{cases} M & R \le \log \frac{L}{(M-1)} \\ \lfloor 2^{-R} L \rfloor + 1 & R > \log \frac{L}{(M-1)}. \end{cases} \tag{7.14}$$

Remark 7.3.1 *Note that the uncoded CP systems can suffer from loss of multipath diversity [55]. However if a linear receiver is used and the system parameters (rate, transmit antenna delays, and FFT block length) are appropriately designed, Lemma 7.3.1 indicates that the maximum diversity can be achieved. Specifically, the maximum diversity is achieved when*

$$\lfloor L 2^{-R} \rfloor + 1 \geqslant M \quad \textit{or} \quad R \leqslant \log \frac{L}{M-1} \triangleq R_{th} \tag{7.15}$$

For any given values of R and M, the data block length L can be chosen such that $R \leqslant R_{th}$. Therefore if we have flexibility in assigning data block length, maximum diversity can always be achieved.

Remark 7.3.2 *The developments in this chapter do not depend on whether the equalization matrix* **W** *is multiplied by the received data in the time domain, or is applied in the frequency domain (via FFT/IFFT). Both approaches lead to the same SINR and outage. Therefore, the results of this chapter are valid for single-carrier systems regardless of whether the receiver operates in the time domain or frequency domain.*

We now consider the second case: the frequency selective channel, i.e., $\nu_i \neq 0$. The delay taps are chosen such that $\delta_i = n_i \delta$ with $\delta > \nu$ and $\{n_i\}$ are distinct integers, and the block length L is chosen such that the transmitted blocks $\bar{s}_i$ are distinct. Notice that the condition on the block length guarantees that the channel coefficients seen at the receiver are independent. Thus the delay taps and the block length L must be chosen to satisfy two conditions

$$n_i \delta_0 \bmod L = n_j \delta_0 \bmod L \quad \forall i \neq j \tag{7.16}$$

$$L \geqslant M(\nu + 1). \tag{7.17}$$

We now consider the following case[1]: $\nu_i = \nu \,\forall i$, $\delta_i = i(\nu + 1)$. It can be shown that this system is equivalent to (7.13) where the variable M in (7.13) is replaced by $M(\nu + 1)$. Thus we obtain the diversity

$$d = \begin{cases} M(\nu + 1) & R \leq \log \frac{L}{(M(\nu+1)-1)} \\ \lfloor 2^{-R}L \rfloor + 1 & R > \log \frac{L}{(M(\nu+1)-1)}. \end{cases} \tag{7.18}$$

The maximum diversity is achieved when

$$R \leqslant \log \frac{L}{M(\nu + 1) - 1} \triangleq R_{th}. \tag{7.19}$$

Other choices of $\{\delta_i\}$ that satisfy (7.16) exist. Each of these choices yields a new $\mathbf{H}_{\text{cir}}$ whose first row is a permutation of the first row of $\mathbf{H}_{\text{cir}}$ under $\delta_i = i(\nu + 1)$. We thus have $L!$

---

[1]We assume $\nu_i = \nu$ for simplicity. The general case can be obtained in a straightforward manner.

different choices for the set $\{\delta_i\}$. Since these circulant matrices have different structures, the diversity analysis of [15] does not directly follow. We study the diversity of these systems and show that the $L! - 1$ remaining choices $\{\delta_i\}$ yield the same diversity as shown in (7.18).

The outage probability of the system model in (7.13) is given by [15]

$$P_{out} \doteq \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1 + \rho|\lambda_k|^2} > L2^{-R}\right)$$

where λ_k are the eigenvalues of $\mathbf{H}_{\text{cir}}$. Since each eigenvalue is a linear combination (DFT points) of channel coefficients $\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{M(\nu+1)-1}]$ then the eigenvalues $\{\lambda_k\}$ obey a zero-mean complex Gaussian distribution. In the special case $L = M(\nu+1)$ the eigenvalues $\{\lambda_k\}$ are independent and it can be shown that [15]

$$P_{out} \doteq \mathbb{P}\left(\sum_{k=1}^{M(\nu+1)} \frac{1}{1 + \rho|\lambda_k|^2} > L2^{-R}\right) \doteq \rho^{-d} \tag{7.20}$$

where d is given by (7.18).

Now let $L > M(\nu+1)$ and $\delta_i = i(\nu+1)$. The channel matrix is circulant with first-row vector $\mathbf{h}^{(1)} = [h_0 \cdots h_{M(\nu+1)-1} \, 0 \ldots 0]$ and the corresponding eigenvalues $\{\lambda_k^{(1)}\}_{k=1}^{L}$ are the DFT points of this zero-padded channel vector $\mathbf{h}^{(1)}$. In this case, [15, Lemma 2] applies and we have

$$\mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1 + \rho|\lambda_k^{(1)}|^2} > m\right) \doteq \mathbb{P}\left(\sum_{k=1}^{M(\nu+1)} \frac{1}{1 + \rho|\lambda_k|^2} > m\right)$$

$$\doteq \rho^{-d} \tag{7.21}$$

where $m = L2^{-R}$ and d is given by (7.18).

We now continue with $L > M(\nu+1)$ but we discard the assumption $\delta_i = i(\nu+1)$. Instead, the delays $\{\delta_i\}$ are only required to satisfy (7.16). In this case the zero-padded channel vector is

$$\mathbf{h}^{(2)} = [0 \cdots h_0 \cdots 0 \cdots h_1 \cdots 0 \cdots h_{M(\nu+1)-1} \cdots 0]$$

where the locations of zeros depend on the particular choice of the $\{\delta_i\}$ via Eq. (7.12). We have the following lemma.

Lemma 7.3.2 *The DFT of $\mathbf{h}^{(2)}$ and the DFT of $\mathbf{h}^{(1)}$ satisfy the following equality*

$$\mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1+\rho|\lambda_k^{(2)}|^2} > m\right) = \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1+\rho|\lambda_k^{(1)}|^2} > m\right) \tag{7.22}$$

where m is a positive number.

Proof Please refer to Appendix.

Setting $m = L2^{-R}$, we conclude from (7.21) and (7.22) that

$$\mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1+\rho|\lambda_k^{(2)}|^2} > m\right) \doteq \rho^{-d}$$

where d is given by (7.18).

Diversity Analysis of Zero-Forcing Receiver

The SINR of the ZF receiver is given by (7.5)

$$\gamma_k = \frac{\rho}{[(\mathbf{H_e}^H \mathbf{H_e})^{-1}]_{kk}}$$

$$= \frac{L\,\rho}{\mathrm{tr}[(\mathbf{H_e}^H \mathbf{H_e})^{-1}]} \tag{7.23}$$

$$= \frac{L\,\rho}{\sum_{k=1}^{L} \frac{1}{|\lambda_k|^2}} \tag{7.24}$$

where (7.23) follows since $\mathbf{H_e}$ is circulant thus the diagonal elements of $(\mathbf{H_e}^H \mathbf{H_e})^{-1}$ are equal, and (7.24) follows from the equivalence of the trace and the sum of eigenvalues.

Let us first consider the case of $L = M(\nu+1)$. In this case the eigenvalues are independent Gaussian random variables. We perform the outage analysis for this case and then show that the result holds for the more general case $L > M(\nu + 1)$. Let α_k be the SNR exponent of the channel eigenvalues, i.e.

$$\alpha_k \triangleq -\frac{\log |\lambda_k|^2}{\log \rho}.$$

The outage probability is

$$
\begin{aligned}
P_{out} &= \mathbb{P}\left(\frac{1}{L} \sum_{\ell=1}^{L} \log(1 + \gamma_\ell) < R \right) \\
&= \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{\rho|\lambda_k|^2} > \frac{L}{2^R - 1} \right) \\
&> \mathbb{P}\left(\frac{1}{\rho|\lambda_k|^2} > \frac{L}{2^R - 1} \right) \\
&= \mathbb{P}\left(\rho^{\alpha_k - 1} > \frac{L}{2^R - 1} \right) \\
&\doteq \mathbb{P}(\alpha_k - 1 > 0) \\
&= \mathbb{P}(\alpha_k > 1) \\
&= \rho^{-1}
\end{aligned}
\tag{7.25}
$$

Eq. (7.25) follows because $|\lambda_k|^2$ is exponential random variable and thus we have $\mathbb{P}(\alpha_k > 1) \doteq 1 - e^{-\rho^{-1}}$ via arguments similar to [15, Lemma 1]. We now obtain an upper bound on outage. The outage probability can be bounded as follows.

$$
\begin{aligned}
P_{out} &= \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{\rho|\lambda_k|^2} > \frac{L}{2^R - 1} \right) \\
&\leqslant \mathbb{P}\left(\frac{L}{\min_k \rho|\lambda_k|^2} > \frac{L}{2^R - 1} \right) \\
&= \mathbb{P}\left(\rho^{\alpha_{\max} - 1} > \frac{1}{2^R - 1} \right) \\
&\doteq \mathbb{P}(\alpha_{\max} > 1) \\
&= 1 - \mathbb{P}\left(\bigcap_{k=1}^{L} \left\{ \alpha_k < 1 \right\} \right) \\
&= 1 - [\mathbb{P}(\alpha_k < 1)]^L
\end{aligned}
\tag{7.26}
$$

where (7.26) follows since $L = M(\nu + 1)$ and thus $\{\alpha_k\}$ are independent. Since $\mathbb{P}(\alpha_k > 1) \doteq 1 - e^{-\rho^{-1}} \doteq \rho^{-1}$ then $\mathbb{P}(\alpha_k < 1) \doteq 1 - \rho^{-1}$ and thus

$$
1 - [\mathbb{P}(\alpha_k < 1)]^L \doteq \rho^{-1}
\tag{7.27}
$$

where we have used the fact that $\rho^{-i} + \rho^{-ni} \doteq \rho^{-i}$ for any positive integers i and n. We then conclude from (7.25) and (7.27) that, when $L = M(\nu + 1)$,

$$P_{out} = \mathbb{P}\bigg(\sum_{k=1}^{L} \frac{1}{\rho|\lambda_k|^2} > \frac{L}{2^R - 1} \bigg)$$

$$\doteq \rho^{-1} \tag{7.28}$$

We now consider the case $L > M(\nu + 1)$. We have the following lemma.

Lemma 7.3.3 *Consider the MISO CDD under two transmission scenarios that are similar in all respects except their data-block lengths $L, \tilde{L}$ where we assume $\tilde{L} > L = M(\nu + 1)$. The eigenvalues of the channel are denoted $\{\lambda_k\}$ when data block length is L, and are denoted $\{\tilde{\lambda}_k\}$ when data block length is $\tilde{L}$. Let m be a positive number. We have the following property*

$$\mathbb{P}\bigg(\sum_{k=1}^{L} \frac{1}{\rho|\lambda_k|^2} > m \bigg) \doteq \mathbb{P}\bigg(\sum_{k=1}^{\tilde{L}} \frac{1}{\rho|\tilde{\lambda}_k|^2} > m \bigg)$$

Proof We start with [15, Lemma 2], which states:

$$\mathbb{P}\bigg(\sum_{k=1}^{L} \frac{1}{1 + \rho|\lambda_k|^2} > m \bigg) \doteq \mathbb{P}\bigg(\sum_{k=1}^{\tilde{L}} \frac{1}{1 + \rho|\tilde{\lambda}_k|^2} > m \bigg)$$

We note the difference with our desired result is only a constant term in the denominators. Close inspection reveals that the proof outlined in [15, Appendix A] goes through even without these constant (identity) terms, therefore the proof of our desired result parallels step-by-step the proof of [15, Lemma 2].

Thus setting $m = \frac{L}{2^R - 1}$ in Lemma 7.3.3 and using (7.28) we conclude that the ZF receiver achieves diversity order *one* irrespective of the block length.

(Non-Cyclic) Delay Diversity

In this section we consider the more traditional delay diversity without cyclic prefix. We establish the equivalence between the delay diversity (DD) single carrier or multicarrier

MISO system[2] and the zero-padding single carrier or multicarrier SISO system. We then apply the result of [48] which shows that linear receiver achieves full multipath diversity for the zero-padding SISO system. In order to transform DD MISO channel into SISO ISI channel, the delays are set to $\delta_i = i\delta$ with $\delta > \nu$ symbols. For simplicity, let $\delta = \nu + 1$ (the proof is easily extended to other cases). It can be shown that the received signal after removing the padding is[3]

$$\mathbf{y} = \sqrt{\rho}\,\bar{\mathbf{H}}\,\bar{s} + \mathbf{n}. \tag{7.29}$$

In the case of single carrier system with no CP extension [50], $\bar{s} \in \mathbb{C}^{L \times 1}$ is the data block and $\bar{\mathbf{H}} \in \mathbb{C}^{L \times L}$ is the truncated Toeplitz channel matrix and the model is equivalent to the SISO ISI channel under single carrier zero-padding transmission. Using linear equalizers in (7.29) achieves the full diversity $d = M(\nu + 1)$ [26, 48]. In the case of multicarrier system, the channel $\bar{\mathbf{H}} \in \mathbb{C}^{L \times L}$ is a circulant channel matrix. Hence the result of Section 7.3 applies. However, the DD multicarrier system incurs a rate loss compared to the CDD system

$$\frac{R_{CDD}}{R_{DD}} = \frac{1}{\eta} = \frac{N_{FFT} + \delta_{\max} + \nu}{N_{\mathrm{FFT}} + \nu}$$

where N_{FFT} is the DFT size [54]. The diversity gain obtained in Section 7.3 is modified to

$$d = \begin{cases} M(\nu + 1) & R' \le \eta \log \frac{L}{(M(\nu+1)-1)} \\ \lfloor 2^{-R'} L \rfloor + 1 & R' > \eta \log \frac{L}{(M(\nu+1)-1)}. \end{cases} \tag{7.30}$$

7.4 Alamouti Signaling

The Alamouti method of space-time signaling can also be characterized as a transmit diversity scheme. Unlike the CDD system, our analysis shows that Alamouti signaling preserves the transmit diversity and thus provides a larger diversity gain compared with the CDD

---

[2]For DD SC system, we adopt the model of [50] where there is no CP insertion.

[3]In the DD multicarrier system, the received signal has a CP extension (added by the multicarrier modulator) and a ZP extension (due to the delays δ_i). In the DD single carrier system the received signal only has a ZP extension.

scheme above a rate threshold R_{th}. We consider single-carrier block transmission over an additive-noise frequency-selective channel with memory ν, similar to [20]. The model supports a 2×1 system and can be extended to $2 \times N$ system.

Each data-block of length L is appended with a CP of length ν to eliminate interblock interference. $\mathbf{x}_i^{(k)}(n)$ denotes the symbol n of the transmitted block k from antenna i. At even time slots, pairs of length-N blocks $\mathbf{x}_1^{(k)}(n)$ and $\mathbf{x}_2^{(k)}(n)$ are generated. The transmission scheme proposed by [20] is

$$\mathbf{x}_1^{(k+1)}(n) = -\mathbf{x}_2^{*(k)}((-n)_N)$$
$$\mathbf{x}_2^{(k+1)}(n) = -\mathbf{x}_1^{*(k)}((-n)_N) \tag{7.31}$$

for $n = 0, 1, \ldots, N-1$ and $k = 0, 2, 4 \ldots$, $(\cdot)^*$ denotes conjugate. A length-ν CP is added to each transmitted block. The total transmit power is divided equally among the antennas. The transmission scheme is shown in Figure 7.2.

The received blocks at time k and $k+1$ are given by

$$\mathbf{y}^{(j)} = \sqrt{\rho}\,\mathbf{H}_1^{(j)}\mathbf{x}_1^{(j)} + \sqrt{\rho}\,\mathbf{H}_2^{(j)}\mathbf{x}_2^{(j)} + \mathbf{n}^{(j)} \quad \text{for } j = k, k+1$$

where $\mathbf{H}_1^{(j)}$ and $\mathbf{H}_2^{(j)}$ are both circulant, and $\mathbf{n}^{(j)}$ is the noise vector for block j. A DFT is then applied to $\mathbf{y}^{(j)}$ to diagonalize the channels as follows

$$\mathbf{Y}^{(j)} = \sqrt{\rho}\,\Lambda_1^{(j)}\mathbf{X}_1^{(j)} + \sqrt{\rho}\,\Lambda_2^{(j)}\mathbf{X}_2^{(j)} + \mathbf{N}^{(j)} \quad \text{for } j = k, k+1$$

where $\mathbf{Y}^{(j)}$, $\mathbf{X}^{(j)}$ and $\mathbf{N}^{(j)}$ are the DFT vectors of $\mathbf{y}^{(j)}$, $\mathbf{x}^{(j)}$ and $\mathbf{n}^{(j)}$ respectively, and Λ_i (for $i = 1, 2$) are diagonal matrices containing the DFT coefficients of the channel impulse responses. Using (7.31) and assuming the channels are fixed over two consecutive blocks (indexed by k and $k+1$), it can be shown that

$$\mathbf{Y} \triangleq \begin{pmatrix} \mathbf{Y}^{(k)} \\ \mathbf{Y}^{*(k+1)} \end{pmatrix} = \underbrace{\begin{pmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_2^* & -\Lambda_1^* \end{pmatrix}}_{\triangleq \Lambda} \begin{pmatrix} \sqrt{\rho}\,\mathbf{X}_1^{(k)} \\ \sqrt{\rho}\,\mathbf{X}_2^{*(k)} \end{pmatrix} + \begin{pmatrix} \mathbf{N}^{(k)} \\ \mathbf{N}^{*(k+1)} \end{pmatrix}. \tag{7.32}$$
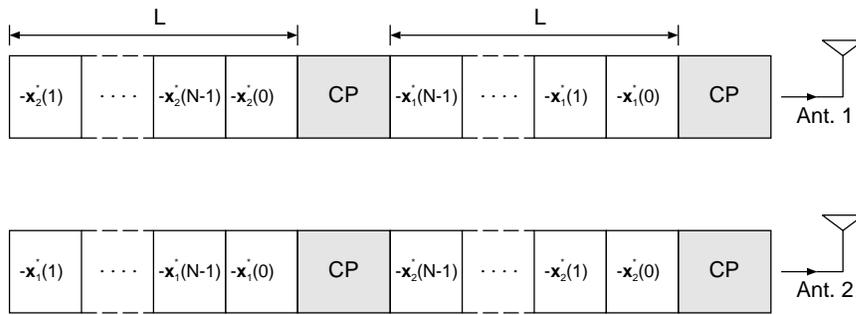
Figure 7.2. Transmission scheme proposed by Al-Dhahir for communication over frequency-selective fading channels.

By multiplying both sides of (7.32) by the orthogonal matrix Λ^* defined in (7.32)

$$\tilde{\mathbf{Y}} \triangleq \Lambda^* \mathbf{Y} = \begin{pmatrix} \tilde{\Lambda} & 0 \\ 0 & \tilde{\Lambda} \end{pmatrix} \begin{pmatrix} \sqrt{\rho}\, \mathbf{X}_1^{(k)} \\ \sqrt{\rho}\, \mathbf{X}_2^{*(k)} \end{pmatrix} + \tilde{\mathbf{N}} \tag{7.33}$$

where $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{N}}$ are the transformed receive vector $\mathbf{Y}$ and noise vector $\mathbf{N}$ respectively, and $\tilde{\Lambda} \triangleq \Lambda_1^H \Lambda_1 + \Lambda_2^H \Lambda_2$ is a $N \times N$ diagonal matrix whose diagonal element i is

$$|\Lambda_1(i, i)|^2 + |\Lambda_2(i, i)|^2. \tag{7.34}$$

7.4.1   MMSE Receiver

We now show that the $2 \times N$ Alamouti SC-FDE can achieve full diversity $2N(\nu + 1)$ as long as the transmission rate is below a certain threshold, and otherwise full spatio-temporal diversity is not achieved. We fully characterize the diversity in all cases. We start by considering $N = 1$ receive antenna.

The received signal for two blocks indexed by k and $k + 1$ is given by (7.32) which can be written as

$$\mathbf{Y} = \sqrt{\rho}\, \Lambda \mathbf{X} + \mathbf{N}.$$

The MMSE equalizer is

$$\mathbf{W} = \left(\Lambda^H \Lambda + \rho^{-1} \mathbf{I} \right)^{-1} \Lambda^H.$$

In other words, the coefficients of the MMSE FDE are given by

$$\mathbf{W}_{k,k} = \frac{\Lambda_{k,k}^*}{|\Lambda_{k,k}|^2 + \rho^{-1}}.$$

Performing the equalization process followed by the IDFT operation yields

$$\tilde{\mathbf{y}}_1 \triangleq \sqrt{\rho}\,\mathbf{Q}^H\left(\Lambda^H\Lambda + \rho^{-1}\mathbf{I}\right)^{-1}\Lambda^H\Lambda\mathbf{Q}\mathbf{x}_1 + \tilde{\mathbf{n}}_1$$

$$\tilde{\mathbf{y}}_2 \triangleq \sqrt{\rho}\,\mathbf{Q}^H\left(\Lambda^H\Lambda + \rho^{-1}\mathbf{I}\right)^{-1}\Lambda^H\Lambda\mathbf{Q}\mathbf{x}_2 + \tilde{\mathbf{n}}_2$$

where $\tilde{\mathbf{n}}_1$ and $\tilde{\mathbf{n}}_2$ are the filtered noise signals. Assuming the transmitted vectors have equal power, the unbiased decision-point SINR of the MMSE SC-FDE for detecting the symbol k of the vector $\mathbf{x}_1$ and $\mathbf{x}_2$ are denoted by $\gamma_{1,k}$ and $\gamma_{2,k}$ and are given by

$$\gamma_{1,k} = \gamma_{2,k} \triangleq \gamma_k = \frac{\rho}{[(\rho^{-1}\mathbf{I} + \mathbf{Q}^H\tilde{\Lambda}\mathbf{Q})^{-1}]_{kk}} - 1,$$

Observe that all γ_k are equal because the matrix $\mathbf{Q}^H\tilde{\Lambda}\mathbf{Q}$ is circulant. Thus γ_k can be written as

$$\begin{aligned}
\gamma_k &= \frac{\rho}{\frac{1}{L}\mathrm{tr}(\rho^{-1}\mathbf{I} + \mathbf{Q}^H\tilde{\Lambda}\mathbf{Q})^{-1}} - 1 \\
&= \frac{\rho}{\frac{1}{L}\mathrm{tr}(\rho^{-1}\mathbf{I} + \tilde{\Lambda})^{-1}} - 1 \\
&= \frac{1}{\frac{1}{L}\mathrm{tr}(\mathbf{I} + \rho\tilde{\Lambda})^{-1}} - 1.
\end{aligned} \tag{7.35}$$

The mutual information is given by

$$\begin{aligned}
I_{MMSE} &= \sum_{k=1}^{L}\log(1 + \gamma_k) \\
&= L\log(1 + \gamma_k) \tag{7.36} \\
&= L\log\left(\frac{1}{\frac{1}{L}\mathrm{tr}(\mathbf{I} + \rho\tilde{\Lambda})^{-1}}\right) \\
&= -L\log\left(\frac{1}{L}\sum_{k=1}^{L}[(\mathbf{I} + \rho\tilde{\Lambda})^{-1}]_{kk}\right) \\
&= -L\log\left(\frac{1}{L}\sum_{k=1}^{L}\frac{1}{1 + \rho\tilde{\lambda}_k}\right)
\end{aligned}$$

the eigenvalues $\tilde{\lambda}_k$ are the diagonal elements of $\tilde{\Lambda}$, and are given by $\tilde{\lambda}_k = |\lambda_{1,k}|^2 + |\lambda_{2,k}|^2$ where $\lambda_{i,k}$ are the eigenvalues of the channel $\mathbf{H}_i$ for $i = 1, 2$ (cf. Equation (7.34)). The outage probability of the MMSE receiver the is given by

$$P_{out} = \mathbb{P}\left(\frac{1}{L}I_{MMSE} < R\right)$$

$$= \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1 + \rho\tilde{\lambda}_k} > L2^{-R}\right). \tag{7.37}$$

Similar to the SISO analysis presented in [15], we first consider the case when $L = \nu + 1$ and later extend it to the other cases when $L > \nu + 1$. When $L = \nu + 1$ all the elements of $\{\Lambda_1(k,k)\}$ and $\{\Lambda_2(k,k)\}$ are i.i.d. Gaussian variables and hence the eigenvalues $\{\tilde{\Lambda}_{kk}\}$ obey the Gamma distribution with shape parameter $M = 2$ and scale parameter 1, i.e. $\tilde{\lambda}_k \sim \Gamma(M, 1)$. When $L > \nu + 1$ the elements $\{\Lambda_i(k,k)\}_{i=1,2}$ are no longer independent and thus analyzing this case requires the unknown distribution of $\{\lambda_k\}$. Instead, we indirectly show that the diversity of $L = \nu + 1$ also holds for $L > \nu + 1$. We continue with the case $L = \nu + 1$

Let $\tilde{\alpha}_k \triangleq -\frac{\log \tilde{\lambda}_k}{\log \rho}$, we have

$$\frac{1}{1 + \rho\tilde{\lambda}_k} = \frac{1}{1 + \rho^{1-\tilde{\alpha}_k}}.$$

Observe that the term $\frac{1}{1+\rho^{1-\tilde{\alpha}_k}}$ is either zero or one at high SNR depending on the value of $\tilde{\alpha}_k$ [15]

$$\lim_{\rho \to \infty} \frac{1}{1 + \rho\tilde{\lambda}_k} = \begin{cases} \rho^{\tilde{\alpha}_k - 1} & \bar{\alpha}_k < 1 \\ 1 & \bar{\alpha}_k > 1 \end{cases}.$$

therefore to characterize the sum $\sum_k \frac{1}{1+\rho\lambda_k}$ in (7.37) at high SNR, we basically count the ones. The outage probability can be written as

$$P_{out} = \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{(1 + \rho\tilde{\lambda}_k)} > L2^{-R}\right)$$

$$\doteq \mathbb{P}\left(\sum_{\tilde{\alpha}_k > 1} 1 > L2^{-R}\right). \tag{7.38}$$

A rigorous proof for (7.38) follows similarly to [5, Section III-A]. We thus need to evaluate $\mathbb{P}(\tilde{\alpha} > 1)$. The probability density function of $\tilde{\lambda}_k$ is

$$f_{\tilde{\lambda}_k}(x) = \frac{1}{\Gamma(M)} x^{M-1} e^{-x}. \tag{7.39}$$

where $\Gamma(M) = (M-1)!$ is the Gamma function.

We want to evalaute the probability $P(\tilde{\alpha}_k > 1)$ which is the same as $P(\tilde{\lambda}_k < \rho^{-1})$. Using (7.39) we have

$$
\begin{aligned}
P(\tilde{\lambda}_k < \rho^{-1}) &= \int_0^{\rho^{-1}} f_{\tilde{\lambda}_k}(x) \, dx \\
&= \frac{1}{\Gamma(M)} \int_0^{\rho^{-1}} x^{M-1} e^{-x} dx \tag{7.40} \\
&= \frac{1}{\Gamma(M)} \left(\Gamma(M) - e^{-\rho^{-1}} \sum_{k=0}^{M-1} \frac{\Gamma(M) \rho^{-1}}{k!} \right) \tag{7.41} \\
&= 1 - e^{-\rho^{-1}} \sum_{k=0}^{M-1} \frac{\rho^{-1}}{k!} \tag{7.42}
\end{aligned}
$$

where we have evaluated the integral in (7.40) according to [27, P.336]. Thus we have

$$
\begin{aligned}
P(\tilde{\alpha}_k > 1) &= P(\tilde{\lambda}_k < \rho^{-1}) \\
&= 1 - e^{-\rho^{-1}} \sum_{k=0}^{M-1} \frac{\rho^{-1}}{k!} \\
&\doteq 1 - \left(1 - \frac{1}{M!} \rho^{-M}\right) \tag{7.43} \\
&\doteq \rho^{-M} \tag{7.44}
\end{aligned}
$$

where (7.43) follows from the Taylor expansion of the exponential function in (7.42) and the fact that $\rho^{-i} + \rho^{-ni} \doteq \rho^{-i}$ for any positive integers i and n.

From the independence of $\{\tilde{\lambda}_k\}$, and subsequently the independence of $\{\tilde{\alpha}_k\}$, we conclude that $M(\tilde{\alpha})$ in (7.38) is binomially distributed with parameter ρ^{-M}. Hence, similar to [15],

we have

$$\mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1+\rho\tilde{\lambda}_k} > L2^{-R}\right) \doteq \mathbb{P}(M(\tilde{\alpha}) > L2^{-R})$$

$$= \sum_{i=\lfloor L2^{-R}\rfloor+1}^{L} \mathbb{P}(M(\tilde{\alpha}) = i)$$

$$\doteq \sum_{i=\lfloor L2^{-R}\rfloor+1}^{L} \binom{L}{i} \rho^{-Mi} \underbrace{(1-\rho^{-M})^{n-i}}_{\doteq 1}$$

$$\doteq \rho^{-M(\lfloor L2^{-R}\rfloor+1)}.$$

which concludes the proof for $L = \nu + 1$

For the case of $P > \nu + 1$, we will need the following lemma.

Lemma 7.4.1 *Consider the MISO Alamouti signaling given by (7.31) under two transmission scenarios that are similar in all respects except their data-block lengths L_1, L_2 where we assume $L_2 > L_1 = \nu + 1$. The eigenvalues of channels $\mathbf{H}_1$ and $\mathbf{H}_2$ are denoted $\{\lambda_{1,k}\}$ and $\{\lambda_{2,k}\}$ respectively when data block length is L_1, and are denoted $\{\breve{\lambda}_{1,k}\}$ and $\{\breve{\lambda}_{2,k}\}$ when data block length is L_2. We have the following property*

$$\mathbb{P}\left(\sum_{k=1}^{L_1} \frac{1}{(1+\rho\,(|\lambda_{1,k}|^2 + |\lambda_{2,k}|^2)} > L2^{-R}\right)$$

$$\doteq \mathbb{P}\left(\sum_{k=1}^{L_2} \frac{1}{(1+\rho\,(|\breve{\lambda}_{1,k}|^2 + |\breve{\lambda}_{2,k}|^2)} > L2^{-R}\right)$$

Proof The proof is similar to [56, Lemma 11].

Thus the data block length does not affect the diversity of 2×1 Alamouti scheme under MMSE SC-FDE is

$$d = \begin{cases} 2(\nu+1) & R \le \log\frac{L}{\nu} \\ 2(\lfloor 2^{-R}L\rfloor + 1) & R > \log\frac{L}{\nu}. \end{cases} \tag{7.45}$$

where L is the data block length.

Remark 7.4.1 *The result above shows that the Alamouti SC-FDE provides twice the diversity of SISO-SC-FDE. At first sight this result seems to be a straight forward product decomposition of a so-called "transmit diversity," i.e., a factor of two due to Alamouti signaling, and a term due to SC-FDE [15]. However, it is important to note that the problem was not fundamentally separable, therefore the results could not be deduced from [15, 20] without a proof, because the distribution of the summation of eigenvalues squared in [20, Eq. (11)] is needed before one can make a conclusive statement about the overall diversity of the system. This subtle point can be further appreciated by noting that diversity in frequency-selective channels cannot in general be decomposed into separate components, e.g., due to the transmitter signaling and otherwise. For example in the earlier case of CDD, the diversity was not a multiple of the diversity of SISO SC-FDE.[4] Thus, the result above was not preordained, even though its form is unsurprising.*

The analysis can easily be generalized for $N > 1$ receive antennas. The outage in the case of N receive antennas will depend on eigenvalues $\tilde{\lambda}_k \triangleq \sum_{i=1}^{2N} |\lambda_{i,k}|^2$ that have distribution $\tilde{\lambda}_k \sim \Gamma(2N, 1)$. Following a similar line of reasoning as earlier, full diversity is achieved when $R \leq \log \frac{L}{\nu}$. More broadly, the diversity for all spectral efficiencies is given by the following theorem.

Theorem 7.4.2 *In a $2 \times N$ quasi-static frequency-selective channel with channel memory ν, using Alamouti signaling given by (7.31) the diversity of the MMSE-SC-FDE is given by*

$$
d = \begin{cases} 2N(\nu + 1) & R \leq \log \frac{L}{\nu} \\ 2N\big(\lfloor 2^{-R}L \rfloor + 1\big) & R > \log \frac{L}{\nu}. \end{cases}
\tag{7.46}
$$

where L is the data block length.

---

[4]There are also other examples showing the non-decomposability of diversity in frequency-selective channels, e.g. [56].

7.4.2 Zero-Forcing Receiver

We now analyze the zero-forcing equalization for Alamouti transmission. It can be shown that the outage probability is

$$P_{out} = \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{\rho \tilde{\lambda}_k} > \frac{L}{2^R - 1} \right) \tag{7.47}$$

with $\tilde{\lambda}_k = |\lambda_{1,k}|^2 + |\lambda_{2,k}|^2$, and $\lambda_{i,k}$ is the k-th eigenvalue of the channel $\mathbf{H}_i$ $(i = 1, 2)$ and $\tilde{\lambda}_k$ obey the Gamma distribution. Similarly to Section 7.3, it is straightforward to show that the diversity of the ZF receiver is only $d_{zf} = 2$. The analysis details are omitted for brevity.

7.5 Simulation Results

Figure 7.3 shows the outage probability P_{out} for the MMSE receiver in the CDD CP MISO flat fading channel with 3 transmit antennas, under various choices of the cyclic delay taps. The rate is $R = 2$ b/s/Hz and $L = 5$. In this case, the MMSE diversity is two (as predicted from (7.18)) since this rate is greater than R_{th} given by (7.19). Figure 7.4 compares CDD-CP and DD-without-CP systems in a 2×1 MISO flat fading channel. The latter system is equivalent to zero-padding transmission over a SISO ISI channel with three channel coefficients and thus achieves the full diversity for all rates [48]. However, the CDD CP-system only achieves full diversity for the rates that satisfy (7.19).

Figure 7.5 compares the performance of zero-forcing and MMSE receivers in 2×1 Alamouti transmission for block-length $L = 4$. The diversity of the ZF is *two* for all rates R, whereas the diversity of the MMSE is greater than or equal to two depending on the value of rate R (cf. Eq. (7.45)).

Figure 7.6 compares the performance of zero-forcing receiver in 2×1 CDD and 2×1 Alamouti transmission wth $\nu = 1$. The diversity of the ZF-CDD is one whereas the diversity of the ZF-Alamouti is two. The diversities of both systems are independent of R.
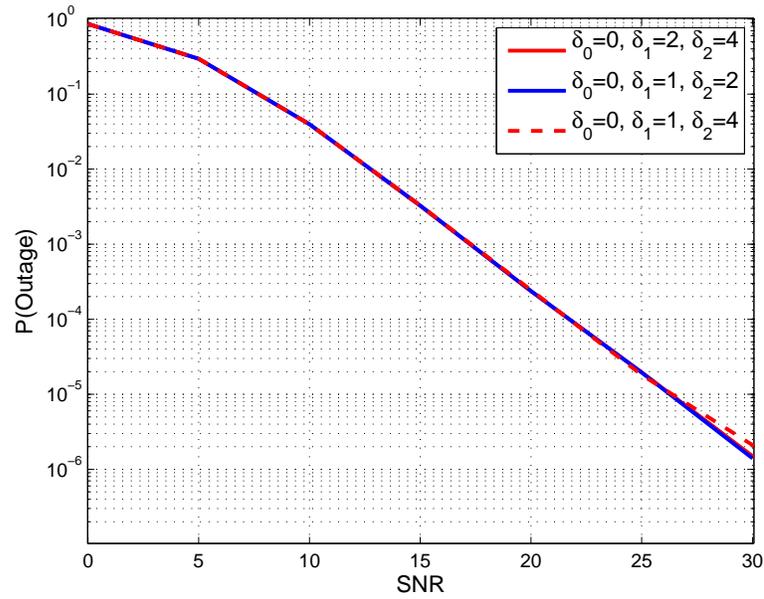
Figure 7.3. The Outage probability of SC/MC CDD MISO under flat fading with three transmit antennas and $R = 2$ b/s/Hz.
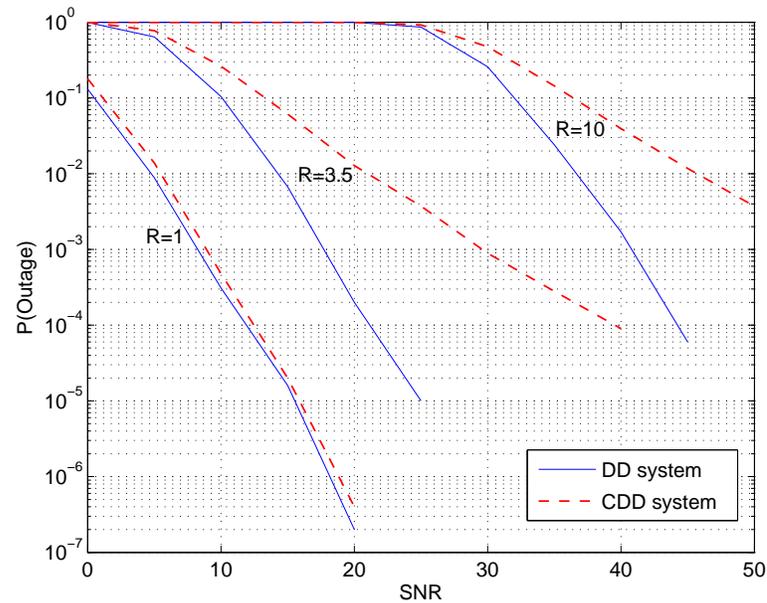
Figure 7.4. Performance of single-carrier delay diversity (DD) and cyclic delay diversity (CDD) in 2×1 MISO under flat fading and $R = 1, 3.5,$ and 10 b/s/Hz.
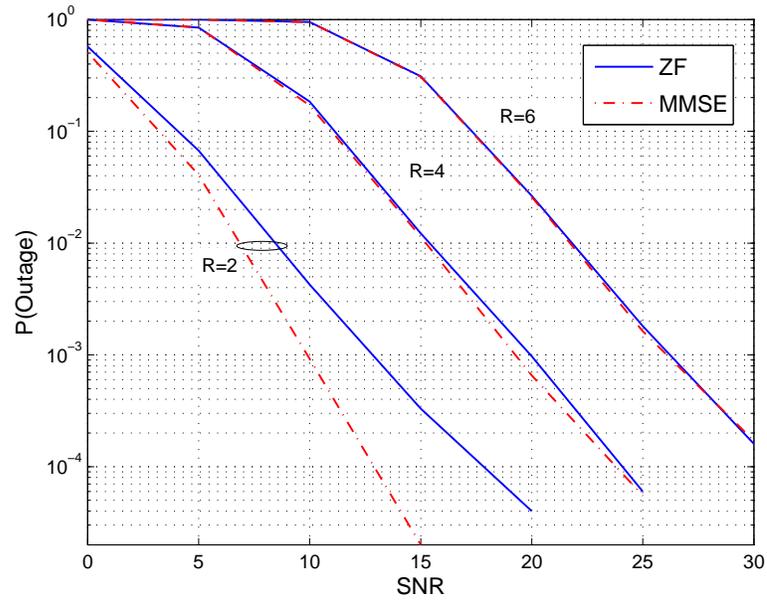
Figure 7.5. The outage probability of zero-forcing and MMSE receiver in 2×1 Alamouti system, $L = 4$, $\nu = 1$ and $R = 2, 4$, and 6 b/s/Hz.
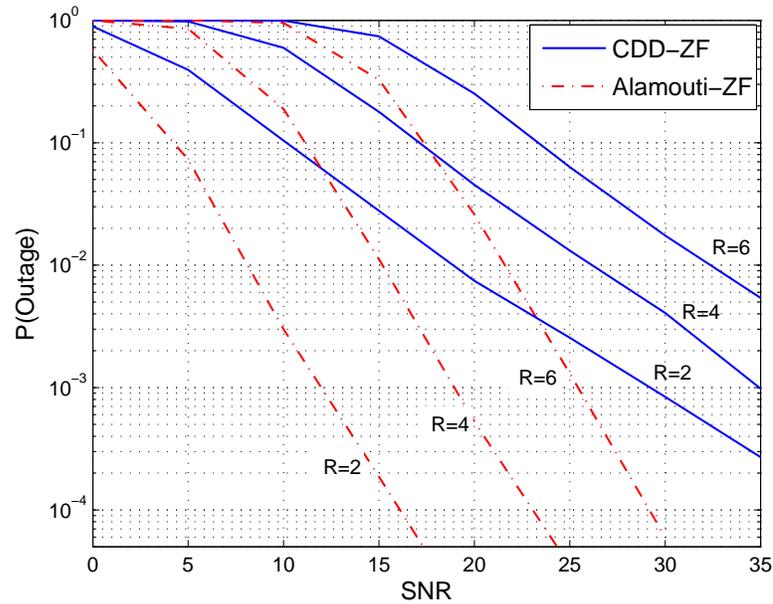
Figure 7.6. The outage probability of zero-forcing receiver in 2×1 single-carrier CDD and Alamouti signaling with $\nu = 1$ and $R = 2, 4, 6$ b/s/Hz .

7.6 Conclusion

This chapter analyzes the single-carrier frequency domain equalizer (SC-FDE) for two common transmit diversity schemes: cyclic delay diversity (CDD) and Alamouti signaling. We characterize the diversity for both schemes at all spectral efficiencies. In the process, we obtain a threshold rate (as a function of data-block length, channel memory, and number of antennas) below which the full spatial-temporal diversity is achieved. Our analysis shows that at high rates the CDD diversity degenerates to the diversity of the SISO SC-FDE, while Alamouti signaling provides twice the diversity of SISO SC-FDE.

7.7 Appendix

7.7.1 Proof of Lemma 7.3.2

Recall that the channel vectors $\mathbf{h}^{(1)}$ and $\mathbf{h}^{(2)}$ are given by

$$\mathbf{h}^{(1)} = [h_0 \ h_1 \ \cdots \ h_{M(\nu+1)-1}, 0, \cdots, 0]$$

$$\mathbf{h}^{(2)} = [0 \cdot \cdot h_0 \cdot \cdot 0 \cdot \cdot h_1 \cdot \cdot 0 \cdot \cdot h_{M(\nu+1)-1} \cdot \cdot 0]$$

and the corresponding DFT vectors are respectively given by

$$\lambda_k^{(1)} = \sum_{i=0}^{\nu} \mathbf{h}^{(1)}(i) \ e^{-j\frac{2\pi i(k-1)}{L}} \quad \text{for } k = 1, \ldots, L$$

$$\lambda_k^{(2)} = \sum_{i=0}^{\nu} \mathbf{h}^{(2)}(i) \ e^{-j\frac{2\pi i(k-1)}{L}} \quad \text{for } k = 1, \ldots, L. \tag{7.48}$$

Note that $\{h_0, \ldots, h_{M(\nu+1)-1}\}$ comprise the non-zero elements of both $\mathbf{h}^{(1)}$ and $\mathbf{h}^{(2)}$. The channel coefficients $\{h_i\}$ are assumed independent and identically distributed circular complex normal random variable with zero mean. Thus the vector $\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{M(\nu+1)-1}]$ obeys the complex normal distribution $\mathcal{CN}(\mathbf{0}, \Gamma)$ given by

$$f(\mathbf{h}) = \frac{1}{\pi^{M(\nu+1)} \det(\Gamma)} \ e^{-\mathbf{h}\Gamma^{-1}\mathbf{h}^H}. \tag{7.49}$$

where the covariance matrix Γ, considering that $\mathbf{h}$ is a row vector, is given by $\Gamma = \mathbb{E}[\mathbf{h}^H \mathbf{h}]$.

We now construct a vector $\bar{\mathbf{h}}$ that has the same distribution as $\mathbf{h}$

$$\bar{\mathbf{h}} = [h_0 e^{j\theta_0} \quad h_1 e^{j\theta_1} \quad \cdots \quad h_{M(\nu+1)-1} e^{j\theta_{M(\nu+1)}}]$$

$$= \mathbf{h}\Theta \tag{7.50}$$

where the matrix Θ is a diagonal matrix that has $\{e^{j\theta_i}\}$ on its diagonal, and $\{\theta_i\}$ are arbitrary real-valued constants. Eq. (7.50) shows that $\bar{\mathbf{h}}$ is a linear transform of $\mathbf{h}$. Thus $\bar{\mathbf{h}}$ obeys $\mathcal{CN}(\mathbf{0}, \Theta^H \Gamma \Theta)$. However, since the coefficients $\{h_i\}$ are independent, the covariance matrix Γ is diagonal and we have

$$\Theta^H \Gamma \Theta = \Theta^H \Theta \Gamma$$

$$= \Gamma. \tag{7.51}$$

Therefore

$$\bar{\mathbf{h}} \sim \mathcal{CN}(\mathbf{0}, \Gamma). \tag{7.52}$$

We thus have

$$\bar{\mathbf{h}} \stackrel{d}{=} \mathbf{h} \tag{7.53}$$

where $\stackrel{d}{=}$ denotes equality in distribution.

Define a vector $\bar{\mathbf{h}}^{(2)}$ that is of length L via appropriate number of zero-padding in arbitrary locations of vector $\mathbf{h}$, for example:

$$\bar{\mathbf{h}}^{(2)} \triangleq [0 \quad h_0 e^{j\theta_0} \quad 0 \quad 0 \quad h_1 e^{j\theta_1} \quad 0 \quad \cdots \quad 0 \quad h_{M(\nu+1)-1} e^{j\theta_{M(\nu+1)}} \quad 0].$$

Let $\{\bar{\lambda}_k^{(2)}\}$ be the DFT of $\bar{\mathbf{h}}^{(2)}$

$$\bar{\lambda}_k^{(2)} = \sum_{i=0}^{\nu} \bar{\mathbf{h}}^{(2)}(i)\, e^{-j\frac{2\pi i(k-1)}{L}} \quad \text{for } k = 1, \ldots, L. \tag{7.54}$$

Note that

$$\bar{\mathbf{h}}^{(2)} \stackrel{d}{=} \mathbf{h}^{(2)}$$

and therefore

$$\{\bar{\lambda}_k^{(2)}\} \stackrel{d}{=} \{\lambda_k^{(2)}\}. \tag{7.55}$$

The phases $\{\theta_i\}$ can be chosen such that for $k = 1, \ldots, L$,

$$\sum_{i=0}^{\nu} \bar{\mathbf{h}}^{(2)}(i) \, e^{-j\frac{2\pi i (k-1)}{L}} = \sum_{i=0}^{\nu} \mathbf{h}^{(1)}(i) \, e^{-j\frac{2\pi i (k-1)}{L}}$$

$$= \lambda_k^{(1)} \tag{7.56}$$

Using (7.55) and (7.56) we get

$$\{\bar{\lambda}_k^{(2)}\} \stackrel{d}{=} \{\lambda_k^{(2)}\}$$

$$= \{\lambda_k^{(1)}\}.$$

Therefore

$$\mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1 + \rho|\lambda_k^{(2)}|^2} > m \right) = \mathbb{P}\left(\sum_{k=1}^{L} \frac{1}{1 + \rho|\lambda_k^{(1)}|^2} > m \right)$$

CHAPTER 8

LATTICE-REDUCTION AIDED EQUALIZATION

8.1   Introduction

The main problem with zero-forcing equalization, which also manifests itself in the system diversity, is noise enhancement. When $\mathbf{H_e}$ is near-singular, zero-forcing amplifies the noise strongly along the smaller eigenvectors of the channel, causing difficulties for the detector. One way to address this problem is Lattice Reduction (LR) [57,58]. In this section we show that lattice reduction aided equalization achieves full diversity. We begin with a short review of LR-aided detection.

The orthogonality of a matrix $\mathbf{H}$ can be quantified using the notion of *orthogonality defect* defined as [59]

$$\delta = \frac{(||\mathbf{b}_1||^2 ||\mathbf{b}_2||^2 \dots ||\mathbf{b}_{ML_d}||^2)}{\det \mathbf{H}^H \mathbf{H}} \tag{8.1}$$

where $\mathbf{H} = [\mathbf{b}_1 \dots \mathbf{b}_{ML_d}]$. Moreover, the orthogonality defect can be bounded as follows [59]

$$1 \leqslant \delta \leqslant c = 2^{2ML_d(ML_d-1)}. \tag{8.2}$$

If the matrix $\mathbf{H}$ has orthogonal columns then $\mathbf{H}^H \mathbf{H}$ is diagonal and thus using ZF equalizer yields same performance as the ML detector, since the decision regions in both cases are the same. However, since $\mathbf{H}$ is in general not orthogonal, linear equalization does not perform as well as ML.

Lattice reduction finds a mapping of the lattice onto itself so that from the viewpoint of the detector, the columns of the transformed equivalent channel are near-orthonormal. More specifically, lattice reduction finds a unimodular[1] matrix $\mathbf{T}$ such that $\mathbf{H} = \mathbf{H_e T}$ is

---

[1]A unimodular matrix is has the following property: the entries of both the matrix and its inverse are complex integers and $\det(\mathbf{T}) = \pm 1$ or $\pm j$.

approximately orthogonal.[2] The process is as follows:

$$\mathbf{y} = \mathbf{H_e x} + \mathbf{n} = (\mathbf{H_e T})\underbrace{\mathbf{T}^{-1}\mathbf{x}}_{\hat{\mathbf{x}}} + \mathbf{n} \qquad (8.3)$$

Denoting with $\mathbf{W}_{ZF}$ the ZF equalizer for the reduced channel $\mathbf{H} = \mathbf{H_e T}$, we have

$$\mathbf{z} = \mathbf{W}_{ZF}\,\mathbf{y} = \hat{\mathbf{x}} + \mathbf{W}_{ZF}\,\mathbf{n} \qquad (8.4)$$

Now $\hat{\mathbf{x}}$ is decoded and $\mathbf{x}$ is recovered by multiplying with $\mathbf{T}$.

Lattice reduction aided detection has been proposed and analyzed in several scenarios. A quantative analysis was presented in [57] where it shown that the error rate performance of LR is within 3 dB of ML for a 2×2 Gaussian MIMO channel model. The LR-aided detection (LRAD) is known to achieve full receive diversity in uncoded flat MIMO broadcast channel [61]. It is also shown [62] that LRAD (with its modified version complex LLL [63]) achieves full receive diversity in MIMO flat V-BLAST channel. LRAD was also studied for OFDM SIOSO channel [64]. The diversity of lattice-reduction-aided receivers in the frequency-selective MIMO channel has been an open problem, which is addressed in this section.

8.1.1   Diversity Analysis

An n-dimensional lattice $\mathcal{L}$ in the m-dimensional space is generated by linear combination of n linearly independent vectors $\{\mathbf{b}_i\}$

$$\mathcal{L} = \left\{ \sum_{i=1}^{n} x_i \mathbf{b}_i \,\big|\, x_i \in \mathbb{Z}, \mathbf{b}_i \in \mathbb{C}^m \right\}$$

Each lattice can be represented by (infinitely) many different bases. Lenstra et al. [60] proposed the first polynomial time algorithm (LLL) that finds a near-orthogonal basis whose vectors are all roughly the same size, specifically, the ratio of the second norm of any two vectors of the basis is no bigger than $2^{(n-1)/2}$. The LLL algorithm was originally introduced

---

[2]Various algorithms exist for efficiently finding this transformation, among them the LLL algorithm [60].
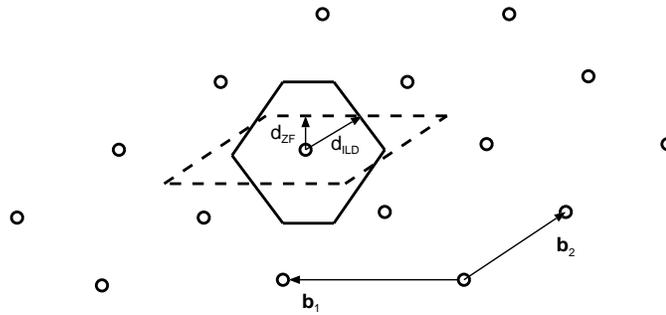
Figure 8.1. Decision regions of ZF (dashed) and infinite lattice decoding (solid) with the corresponding minimum distances in a 2-dimensional lattice

for real lattice bases and was shown to require $O(n^4)$ arithmetic operations. Complex LLL (CLLL) algorithms have been proposed in [62, 63] that yield similar performance as LLL when applied to complex channel but with reduced complexity.

We start with outlining a recent result from [65] which characterizes the performance of lattice-reduction-aided detection in infinite lattice decoding (ILD). Let d_{ILD} denote the Euclidean distance between a lattice point to the closest boundary of the corresponding Voronoi cell, and d_{ZF} denote the distance between the same lattice point and the closest boundary of the decision region of the ZF detector (c.f. Figure 8.1). The term "proximity factor" is defined in [65] as follows

$$\kappa_{zf} \triangleq \sup_{\mathbf{H} \in \mathcal{H}_{\text{reduced}}} \frac{d_{ILD}^2}{d_{ZF}^2} \tag{8.5}$$

where the supremum is taken over the set $\mathcal{H}_{\text{reduced}}$ of basis matrices satisfying a certain reduction criterion.

The factor κ_{zf} is a function of not only the lattice, but also the basis that is used to represent it, and without lattice reduction κ_{zf} is unbounded. With reduction, this factor is upper bounded by a constant that is a function of the cardinality of the basis of the lattice [65]. Although [65] is based on a real channel model, the analysis is easily extended to the complex case by rewriting the system model in (8.3) as

$$\begin{bmatrix} \text{Re}\{\mathbf{y}\} \\ \text{Im}\{\mathbf{y}\} \end{bmatrix} = \begin{bmatrix} \text{Re}\{\mathbf{H}\} & -\text{Im}\{\mathbf{H}\} \\ \text{Im}\{\mathbf{H}\} & \text{Re}\{\mathbf{H}\} \end{bmatrix} \begin{bmatrix} \text{Re}\{\mathbf{x}\} \\ \text{Im}\{\mathbf{x}\} \end{bmatrix} + \begin{bmatrix} \text{Re}\{\mathbf{n}\} \\ \text{Im}\{\mathbf{n}\} \end{bmatrix}.$$

Similar to the analysis of [65], by using (8.5) for a complex lattice it can be shown that the error probability of the LR-aided ZF detector is upper bounded as

$$P_{e,\text{ZF}} \leqslant 2ML_d P_e \tag{8.6}$$

where P_e is the error probability of the infinite lattice decoding. It is known that for sufficiently large block lengths the outage and error probabilities decay at the same rate with increasing SNR and we thus have [1,2]

$$P_e \doteq P(\mathcal{O}). \tag{8.7}$$

Therefore we concentrate on outage calculations. The outage probability $\mathbb{P}(\mathcal{O})$ is given by

$$P(\mathcal{O}) \triangleq \mathbb{P}\left(\frac{1}{ML_d} I(\mathbf{x};\mathbf{y}) < R\right) \tag{8.8}$$

where $I(\mathbf{x};\mathbf{y}) = \log \det(\mathbf{I} + \rho^{-1}\mathbf{H}_e^H\mathbf{H}_e)$ [66]. We now bound $\det(\mathbf{I} + \rho^{-1}\mathbf{H}_e^H\mathbf{H}_e)$

$$\det(\mathbf{I} + \rho^{-1}\mathbf{H}_e^H\mathbf{H}_e) = \prod_{k=1}^{ML_d}(1 + \rho\lambda_k)$$

$$\geqslant 1 + \rho\lambda_{\max}$$

$$\geqslant 1 + \rho\frac{1}{ML_d}\sum_{k=1}^{ML_d}\lambda_k \tag{8.9}$$

where $\{\lambda_k\}$ are the eigenvalues of $\mathbf{H}_e^H\mathbf{H}_e$. We can thus bound the outage probability in (8.8)

$$P(\mathcal{O}) \leqslant \mathbb{P}\left(\frac{1}{ML_d}\log\left(1 + \rho\frac{1}{ML_d}\sum_{k=1}^{ML_d}\lambda_k\right) < R\right)$$

$$= \mathbb{P}\left(\rho\frac{1}{ML_d}\sum_{k=1}^{ML_d}\lambda_k < 2^{RML_d} - 1\right)$$

$$= \mathbb{P}\left(\rho\sum_{k=1}^{ML_d}\lambda_k < ML_d(2^{RML_d} - 1)\right)$$

$$\doteq \mathbb{P}\left(\sum_{k=1}^{ML_d}\lambda_k < \rho^{-1}\right) \tag{8.10}$$

Recall from the definition of the Frobenius norm

$$||\mathbf{H_e}||_F^2 = \sqrt{\sum_{i,j} |h_{ij}|^2} = \sqrt{\text{tr}\,(\mathbf{H_e}^H \mathbf{H_e})} = \sqrt{\sum_{k=1}^{ML_d} \lambda_k}$$

where $h_{i,j}$ is the entry (i,j) of the matrix $\mathbf{H_e}$. Since $\mathbf{H_e}$ is block Toeplitz it can be easily verified that

$$\sum_{i,j} |h_{ij}|^2 = L_d \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{k=0}^{\nu} |h_{mn,k}|^2 \tag{8.11}$$

Using (8.10) and (8.11) we have

$$P(\mathcal{O}) \,\dot{\leqslant}\, \mathbb{P}\left(L_d \sum_{m,n,k} |h_{mn,k}|^2 < \rho^{-1} \right) \tag{8.12}$$

Since $\sum_{m,n,k} |h_{mn,k}|^2$ is a chi-square random variable with $2MN(\nu+1)$ degree of freedom, evaluating (8.12) yields [2,6]

$$P(\mathcal{O}) \,\dot{\leqslant}\, \rho^{-MN(\nu+1)}. \tag{8.13}$$

Thus the outage diversity of the LR-aided ZF is lower bounded $d_{out} \geq MN(\nu+1)$. We also know $d_{out} \leq MN(\nu+1)$ via the diversity of the optimum ML receiver. Therefore we conclude that the LR-aided ZF outage diversity is $d_{out}^{LR} = MN(\nu+1)$.

Remark 8.1.1 *We note that the identical proof easily follows for the MMSE equalizer, therefore it is established that the lattice-reduction MMSE equalizer also enjoys full diversity.*

CHAPTER 9

FINITE LENGTH DECISION FEEDBACK EQUALIZER

9.1 Introduction

Single-carrier frequency domain equalizer (SC-FDE) is attractive for broadband wireless communication and has been proposed for 3GPP long term evolution (LTE) standard. Compared to OFDM, the SC-FDE is robust with respect to peak-to-average power ratio and has lower sensitivity to frequency offset [44]. The performance of the FIR SC-MMSE-LE has been analyzed in [15] where it is shown that the diversity is a function of spectral efficiency in the fixed-rate regime. In this section we provide an in-depth analysis for the SC-MMSE-DFE.

Belfiore and Park [67] produced a slightly different DFE structure known as *noise predictive* (NP) DFE. This method is motivated by the fact that the feed-forward filter colors the noise which can be partially predicted. Structurally, this method adds a feed-forward loop (see Figure 9.1). Operationally, in the MMSE-DFE the two filters are jointly optimized, but in the noise-predictive DFE they are sequentially optimized. It is known that when the feed-forward filter is implemented in a single-carrier mode, the MSE of the two DFE
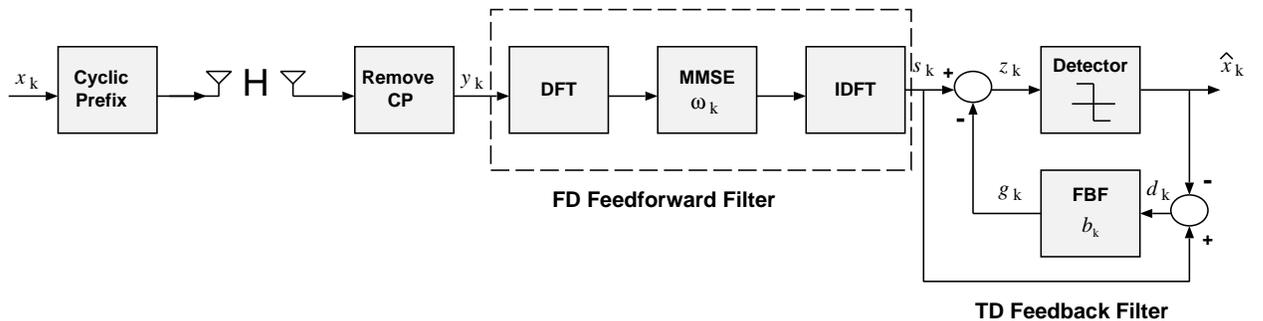
Figure 9.1. Block Diagram for the SC DFE NP

.

146

methods are identical [68] (subject to assumption of no error propagation). Therefore the MMSE-DFE and noise-predictive DFE, when using single-carrier feed-forward filters, have identical diversity. In this chapter we choose to analyze the noise-predictive DFE, noting that the results apply directly to MMSE-DFE.

9.2 System Model

We consider an ISI wireless channel (with channel memory ν) that is assumed fixed over the transmission block. To remove the inter-block interference at the receiver, a cyclic-prefix (CP) of at least ν-length is inserted at the beginning of each data-block of length L and discarded at the receiver. The input-output system model for block transmission scheme is

$$\mathbf{y} = \mathbf{H}\mathbf{U}_{cp}\mathbf{x} + \mathbf{n} \tag{9.1}$$

where $\mathbf{x} \in \mathbb{C}^{(L+\nu)\times 1}$, $\mathbf{y} \in \mathbb{C}^{L\times 1}$ and $\mathbf{n} \in \mathbb{C}^{L\times 1}$ denote the transmitted, the received and the noise vectors respectively. The noise vector is assumed white and Gaussian with zero mean and covariance matrix $\sigma^2\mathbf{I}$ (where σ^2 is the noise variance). The matrix $\mathbf{H} \in \mathbb{C}^{L\times(L+\nu)}$ is the convolution channel matrix. The linear data extension operation that maps the data block of length L to the transmitted vector of length $L + \nu$ can be expressed by $\bar{\mathbf{x}} = \mathbf{U}_{cp}\mathbf{x}$ where $\mathbf{U}_{cp}$ is the the CP matrix given by

$$\mathbf{U}_{cp} = \begin{bmatrix} \mathbf{I}_L \\ \mathbf{I}_\nu & \mathbf{0}_{\nu\times(L-\nu)} \end{bmatrix}.$$

Hence we can rewrite the model in (9.1) as

$$\mathbf{y} = \mathbf{H}\mathbf{U}_{cp}\mathbf{x} + \mathbf{n} = \mathbf{H_e}\mathbf{x} + \mathbf{n} \tag{9.2}$$

where $\mathbf{H_e} = \mathbf{H}\mathbf{U}_{cp}$ is the equivalent $L \times L$ circulant channel matrix given by

$$\mathbf{H_e} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\nu & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_\nu & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_1 & h_2 & \cdots & h_\nu & 0 & \cdots & h_0 \end{bmatrix}. \tag{9.3}$$

Note that $\mathbf{H_e}$ has eigen decomposition $\mathbf{H_e} = \mathbf{Q}^H \Lambda \mathbf{Q}$ where $\mathbf{Q}$ is the unitary discrete Fourier transform matrix. The diagonal elements of Λ are given by [69]

$$\lambda_k = \sum_{i=0}^{\nu} h_i e^{-j\frac{2\pi i(k-1)}{L}} \quad \text{for } k = 1, \ldots, L. \tag{9.4}$$

9.2.1   Optimum Filters – Methodology

It is well-known that MMSE estimation of one random variable from another (including random vectors) is made possible via manipulating the second-order properties of the variables, i.e., the power- and cross-spectral densities, and then applying spectral factorization. However, in our case (as we shall see shortly) one must obtain multiple filters that depend not only on "observation" random variables, but also on other random variables that produce affine terms and do not participate in the application of the orthogonality principle. In such complicated scenarios, the power spectral density (Fourier transform of the auto-correlation) is not sufficient to derive the resulting filters, and a Fourier-like representation of the random processes themselves is needed. This was recognized, e.g., in [70], where the notion of the D-transform was used on a discrete random process s_k to arrive at the following transform-domain random process:

$$X(D) = \sum_k x_k \, D^k \tag{9.5}$$

where D is a complex number, and when $D = e^{j\omega}$ we have the Fourier transform. This approach has been widely followed in the DFE literature since the appearance of [70].

Unfortunately, from an analytical viewpoint this approach is not completely satisfactory: the infinite sum in (9.5) does not converge in any sense (e.g. in quadratic mean or with probability 1) for any value of D on the unit circle. In fact it is well known that stationary random processes do not admit a Fourier transform in the usual sense [71].

The authors of [70] were aware that a power spectral density cannot be directly calculated from (9.5) and defined the spectral density via a truncation and normalization [70, Page 2583, footnote 1]:

$$S_x(D) = \lim_{M \to \infty} \frac{1}{2M+1} E\big[X_M(D)X_M^*(D^{-*})\big]$$

where $X_M(D) = \sum_{k=-M}^{M} x_k D^k$. By (effectively) allowing the exchange of $X_M(D)$ and $X(D)$ freely, [70] arrives at a semantic sequence of equations that are manipulated to arrive at an answer. Needless to say these manipulations are not problem-free, not least because they require $X(D)$ which does not exist. Thus the intermediate steps in [70], especially those involving filtering, are not on solid mathematical ground, albeit the *correct* end results provided in [70].

We are thus motivated to use a more robust approach based on the classical *spectral representation* of random processes, itself based on the generalized harmonic analysis devised by Wiener [72]. This classical branch of the theory of random processes defines a Fourier-like representation of the sample-paths of the random process, and links the behavior of this Fourier-like representation to the autocorrelation function and its Fourier transform (power spectral density). Unfortunately, the formal literature on random processes can be somewhat dense, and even the friendlier developments such as [71] have not found wide favor among information theorists and communication professionals. We suspect one reason is that the results are often developed in the continuous time and are presented in a way that does not immediately suggest the sampled-time counterpart. Motivated by the above considerations, this chapter also produces a compact and self-contained introduction to spectral representations, developed completely in sampled-time domain.

The spectral representation is defined using the inverse Fourier transform integral. Let x_k, $k = 0, \pm 1, \cdots$, denote a stationary random process with discrete time parameter. The spectral representation of x_k is defined via[1]

$$x_k = \int_{-\pi}^{\pi} e^{j\omega k} \, dX(\omega). \tag{9.6}$$

where ω is the variable representing frequency and $X(\omega)$ is called the spectral process of x_k. Note that (9.6) is a Stieltjes integral as $X(\omega)$ is not differentiable for stationary random

---

[1]In the literature on stochastic processes and some other fields, the Fourier transform appears with the kernel $e^{j\omega}$ while in engineering the kernel $e^{-j\omega}$ is common. This makes no difference in the substance of the arguments and is equivalent to exchanging negative and positive frequencies. In this chapter, we adopt the inverse Fourier transform as used in engineering.

processes. A basic but reasonably rigorous grounds-up treatment of spectral representation is given in Appendix 9.5.1.

9.2.2 Optimum Filters – Design

In the NP-DFE, the feedforward filter and the feedback filter are designed sequentially. The optimum feedforward MMSE filter coefficients, denoted by w_k, are given by

$$w_k \triangleq \mathbf{W}_k = \frac{\lambda_k^*}{|\lambda_k|^2 + \rho^{-1}} \tag{9.7}$$

In order to find the MSE we define the following signals. Let B_n be the L-size DFT of the feedback filter coefficients

$$B_n = \sum_{k=1}^{\nu} b_k e^{-j2\pi \frac{nk}{L}}, \qquad n = 1, 2 \cdots L - 1. \tag{9.8}$$

We also define $H(\omega)$ and $B(\omega)$ to be the the discrete-time Fourier transforms of the channel and the feedback filters respectively. Let $W(\omega)$ be the discrete-time Fourier transform of the IDFT(w_k), that is w_k is the sampling of the DTFT

$$w_k = W(\omega)|_{\omega = \frac{2\pi k}{L}} \tag{9.9}$$

where

$$W(\omega) \triangleq \frac{H^*(\omega)}{|H(\omega)|^2 + \rho^{-1}}. \tag{9.10}$$

Following the notation of [70] we derive the MSE of the DFE-NP filter. The error signal, defined as the noise plus distortion, is given by

$$e_k = z_k - x_k.$$

Using the spectral representation of the above equation

$$\int_{-\pi}^{\pi} e^{j\omega k} \, dE(\omega) = \int_{-\pi}^{\pi} e^{j\omega k} \, dZ(\omega) - \int_{-\pi}^{\pi} e^{j\omega k} \, dX(\omega) \qquad \forall k \in \mathbb{Z}$$

Using the completeness of the exponential basis, we have:

$$dE(\omega) = dZ(\omega) - dX(\omega)$$

From the system model given by Figure 9.1, we have

$$dZ(\omega) = \left(1 - B(\omega)\right)\left[W(\omega)H(\omega)\,dX(\omega) + W(\omega)\,dN(\omega)\right]$$

and thus

$$dE(\omega) = \left(1 - B(\omega)\right)\left[\left(W(\omega)H(\omega) - 1\right)dX(\omega) + W(\omega)\,dN(\omega)\right]. \tag{9.11}$$

Due to the orthogonal increment property of the spectral process (see Appendix 9.5.1) the power spectral densities[2] of the error, the input, and the noise signal can be found as follows

$$\mathbb{E}\left[dE(\omega)dE^*(\omega)\right] = S_e(\omega)d\omega \tag{9.12}$$

$$\mathbb{E}\left[dX(\omega)dX^*(\omega)\right] = S_x(\omega)d\omega$$

$$\mathbb{E}\left[dN(\omega)dN^*(\omega)\right] = S_n(\omega)d\omega$$

Define $\bar{B}(\omega) \triangleq 1 - B(\omega)$, we have

$$\mathbb{E}\left[|dE(\omega)|^2\right] = \mathbb{E}\left[dE(\omega)dE^*(\omega)\right]$$

$$= \left|\bar{B}(\omega)\right|^2\left(\left|W(\omega)H(\omega) - 1\right|^2 S_x(\omega)d(\omega) + |W(\omega)|^2 S_n(\omega)d(\omega)\right)$$

$$= \frac{\left|\bar{B}(\omega)\right|^2}{|H(\omega)|^2 + \rho^{-1}}\left[\frac{\rho^{-2}}{|H(\omega)|^2 + \rho^{-1}}S_x(\omega)d(\omega) + \frac{|H(\omega)|^2}{|H(\omega)|^2 + \rho^{-1}}S_n(\omega)d(\omega)\right]$$
$$\tag{9.13}$$

$$= \frac{\left|\bar{B}(\omega)\right|^2}{|H(\omega)|^2 + \rho^{-1}}\left[\frac{\rho^{-2}}{|H(\omega)|^2 + \rho^{-1}}\sigma_x^2\,d(\omega) + \frac{|H(\omega)|^2}{|H(\omega)|^2 + \rho^{-1}}\sigma_n^2\,d(\omega)\right] \tag{9.14}$$

$$= \frac{\left|\bar{B}(\omega)\right|^2}{|H(\omega)|^2 + \rho^{-1}}\,\sigma_n^2\,d(\omega). \tag{9.15}$$

---

[2]Denoted "spectral density function" in the random process literature

where (9.13) follows by using (9.10), Eq. (9.14) assumes that both the input and noise signals possess a white spectrum[3], and Eq. (9.15) uses $\rho = \frac{\sigma_x^2}{\sigma_n^2}$. Now, using (9.12) and (9.15) we conclude

$$S_e(\omega) = \frac{\left|\bar{B}(\omega)\right|^2}{|H(\omega)|^2 + \rho^{-1}}\, \sigma_n^2. \tag{9.16}$$

which can also be written in the z domain

$$S_e(z) = \frac{\bar{B}(z)\bar{B}^*(z^{-*})}{H(z)H^*(z^{-*}) + \rho^{-1}}\, \sigma_n^2 \tag{9.17}$$

We now use a celebrated result form discrete-time spectral factorization theory [70, 73].

Lemma 9.2.1 *Let $S(z)$ denote a power spectral density with finite power, i.e.,*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} S(e^{j\omega})\, d\omega < \infty \tag{9.18}$$

and it satisfies the Paley-Wiener condition

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left[S(e^{j\omega})\right] d\omega > -\infty \ . \tag{9.19}$$

Then a unique spectral factorization for $S(z)$ exists:

$$S(z) = g_0\, v(z)v^*(z^{-*})$$

where $v(z)$ is canonical[4] and g_0 is given by the so-called Szegö formula

$$g_0 = \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln S(e^{j\omega})\, d\omega\right]$$

---

[3]Using an equalizer implies that residual interference is treated as noise (otherwise one may as well use an ML decoder). Thus the equalized channel is treated as a AWGN channel, and for that channel the optimal codebook is white. Incidently, the whiteness of the input signal was also implicit in [70, Eq. (26) and (27)]. With our method, it is also possible to design DFE or linear equalizers for any non-white input power spectral density: it is sufficient to use the appropriate $S_x(\omega)$ in Eq. (9.13). The following spectral factorization steps will naturally be affected.

[4]A filter response is canonical if it is causal, monic ($v(0) = 1$) and minimum phase (all the poles are outside the unit circle and all the zeros are on or outside the unit circle).

Lemma 9.2.1 together with Eq. (9.17) imply that for the error to be white, the feedback filter must satisfy the following factorization

$$H(z)H^*(z^{-*}) + \rho^{-1} = g_0 \, \bar{B}(z) \, \bar{B}^*(z^{-*}) \tag{9.20}$$

which also means

$$S_e(z) = \frac{\sigma_n^2}{g_0}. \tag{9.21}$$

Thus the MSE is σ_n^2/g_0. The constant g_0 is given by

$$g_0 = \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left(|H(\omega)|^2 + \rho^{-1}\right) \, d\omega\right] \tag{9.22}$$

The unbiased SINR at the output of the DFE-NP filter is

$$\gamma^{\text{DFE-NP}} = \frac{\sigma_x^2}{MSE} - 1 = g_0\rho - 1 \tag{9.23}$$

This expression does not include the power loss due to boundary effects (e.g. cyclic prefix or zero padding), therefore it is only accurate at $L \to \infty$. We shall evaluate the outage probability for asymptotically large L, and then show that the result holds for arbitrary block length L.

9.3   Outage Analysis

The outage probability is

$$
\begin{aligned}
P_{out} &= \mathbb{P}\left(\log(\gamma^{\text{DFE-NP}} + 1) < R\right) \\
&= \mathbb{P}\left(\rho \, \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left(|H(\omega)|^2 + \rho^{-1}\right) \, d\omega\right] < 2^R\right) \\
&< \mathbb{P}\left(\rho \, \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left|H(\omega)\right|^2 \, d\omega\right] < 2^R\right) \\
&= \mathbb{P}\left(\exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left|H(\omega)\right|^2 \, d\omega\right] < \rho^r\right) \\
&\doteq \rho^{-(\nu+1)(1-r)}
\end{aligned}
\tag{9.24}
$$

The proof of (9.24) ignores error propagation and closely follows the proof of the infinite-length DFE presented in [26, Section IV-E] and is omitted here for brevity. Moreover one can show that (9.24) also holds for the fixed rate regime $r = 0$ thus (9.24) is valid for $r \in [0, 1]$ (the closed interval, instead of a semi-open interval).

We now proceed to consider a finite L, where we will show with a bounding technique that the same diversity as Eq. (9.24) is obtained. Recall that $\{\lambda_k\}$ are the DFT of the channel coefficients (cf. (9.4)). The mean square error for finite L is given by [74]

$$J = \frac{\sigma_n^2}{L} \sum_{k=0}^{L-1} \frac{|1 - B_k|^2}{|\lambda_k|^2 + \rho^{-1}} \tag{9.25}$$

where the coefficients B_k are optimized for the specific block length L. The unbiased SINR of the noise-predictive DFE with block length L is

$$\gamma = \frac{\sigma_x^2}{MSE} - 1 = \left(\frac{1}{L} \sum_{k=0}^{L-1} \frac{|1 - B_k|^2}{\rho|\lambda_k|^2 + 1} \right)^{-1} - 1.$$

So the outage of the system with length L is:

$$P_{out,L} = \mathbb{P}(\log(\gamma + 1) < R)$$
$$= \mathbb{P}\left(\sum_{k=0}^{L-1} \frac{|1 - B_k|^2}{\rho|\lambda_k|^2 + 1} > L2^{-R} \right) \tag{9.26}$$

We now consider *another* block length $\tilde{L} = TL$, where T is a positive integer, and recalculate the outage. Using the same arguments as above, the outage for length $\tilde{L}$ is:

$$P_{out,\tilde{L}} = \mathbb{P}\left(\sum_{k=0}^{\tilde{L}-1} \frac{|1 - \tilde{B}_k|^2}{\rho|\tilde{\lambda}_k|^2 + 1} > \tilde{L}2^{-R} \right) \tag{9.27}$$

where now $\tilde{B}_k$ are the coefficients optimized specifically for block length $\tilde{L}$. To find the relation of $P_{out,\tilde{L}}$ and $P_{out,L}$, we first concentrate on the respective channel eigenvalues. Recall that $\{\lambda_k\} = \text{DFT}(\mathbf{h})$ where $\mathbf{h} = [h_0, \ldots, h_\nu, 0, \ldots, 0]$ is a vector of length-L. The eigenvalues $\{\tilde{\lambda}\} = \text{DFT}(\tilde{\mathbf{h}})$, where $\tilde{\mathbf{h}}$ is a zero-padded version of $\mathbf{h}$:

$$\tilde{\mathbf{h}} \triangleq [h_0, h_1, \cdots, \underbrace{0, \cdots, 0}_{\tilde{L}-\nu-1}]$$

Using the DFT properties, zero-padding in the time domain is equivalent to interpolation in the DFT domain:

$$\tilde{\lambda}_{Tk} = \lambda_k \qquad k = 0, \ldots, L-1 \tag{9.28}$$

Let us denote by $\mathcal{A}$ the subset of $\{\tilde{\lambda}\}$ that are identical to $\{\lambda_k\}$, i.e., $\mathcal{A} = \{\tilde{\lambda}_i, i = 1, T+1, \cdots, T(L-1)+1\}$. It immediately follows:

$$P_{out,L} = \mathbb{P}\left(\sum_k \frac{|1-B_k|^2}{\rho|\lambda_k|^2 + 1} > L2^{-R} \right)$$

$$\leqslant \mathbb{P}\left(\sum_k \frac{|1-\tilde{B}_{Tk}|^2}{\rho|\lambda_k|^2 + 1} > L2^{-R} \right) \tag{9.29}$$

$$\doteq \mathbb{P}\left(\sum_{k=0}^{L-1} \frac{|1-\tilde{B}_{Tk}|^2}{\rho|\lambda_k|^2 + 1} > \tilde{L}2^{-R} \right) \tag{9.30}$$

$$= \mathbb{P}\left(\sum_{k=0}^{L-1} \frac{|1-\tilde{B}_{Tk}|^2}{\rho|\tilde{\lambda}_{Tk}|^2 + 1} > \tilde{L}2^{-R} \right) \tag{9.31}$$

$$\leqslant \mathbb{P}\left(\sum_{\ell=0}^{TL-1} \frac{|1-\tilde{B}_\ell|^2}{\rho|\tilde{\lambda}_\ell|^2 + 1} > \tilde{L}2^{-R} \right)$$

$$= P_{out,\tilde{L}} \tag{9.32}$$

where (9.29) holds because $\tilde{B}_{Tk}$ may not be the same as the optimized coefficients for block-length L, thus they increase outage probability; Eq. (9.30) holds because this exponential relationship is not affected by the magnitude of a constant; (9.31) holds by noting $\tilde{\lambda}_{Tk} = \lambda_k$, and the following inequality holds due to addition of positive terms inside the probability expression. Overall, we conclude $P_{out,L} \dot{\leqslant} P_{out,\tilde{L}}$

Recall that $\tilde{L} = TL$. When T is sufficiently large, Eq. (9.24) indicates that $P_{out,\tilde{L}} \doteq \rho^{-(\nu+1)(1-r)}$, therefore $P_{out,L} \dot{\leqslant} \rho^{-(\nu+1)(1-r)}$. Since $(\nu+1)(1-r)$ is the best possible SNR exponent in this system, we must have $P_{out,L} \doteq \rho^{-(\nu+1)(1-r)}$ for all L.

Note that the analysis so far ignores error propagation. We now complete the analysis by accounting for error propagation. For the initial symbol there is no ISI from the past (and hence no error propagation) therefore from (9.24) we have an error probability

$$\mathbb{P}(E_0) \doteq \rho^{-(\nu+1)(1-r)} .$$

For the next symbol

$$\mathbb{P}(E_1) = \mathbb{P}(E_1|E_0)\mathbb{P}(E_0) + \mathbb{P}(E_1|\bar{E}_0)\mathbb{P}(\bar{E}_0) \tag{9.33}$$

$$\dot{=} \quad \rho^{-(\nu+1)(1-r)} + \rho^{-(\nu+1)(1-r)} \tag{9.34}$$

$$\dot{=} \rho^{-(\nu+1)(1-r)} \tag{9.35}$$

where we have used $\mathbb{P}(E_1|E_0) \dot{=} 1$ and $\mathbb{P}(\bar{E}_0) \dot{=} 1$. Using mathematical induction it can be shown that $\mathbb{P}(E_k) = \rho^{-(\nu+1)(1-r)}$ for all k. We thus conclude that the DMT of the FIR MMSE-DFE is

$$d_{\text{DFE MMSE}}(r) = (\nu+1)(1-r), \quad r \in [0,1] \tag{9.36}$$

Remark 9.3.1 *As far as diversity is concerned, avoiding error propagation effects is relatively easy. If ν successive symbols are known at the receiver with error probability $\dot{\leqslant}\rho^{-(\nu+1)(1-r)}$, the subsequent symbols will be detected with full diversity. This condition is easily satisfied, e.g., if the system starts from the zero state, or if the transmitter sends at least ν consecutive known symbols (pilots).*

Remark 9.3.2 *The coefficients $\{B_k\}$ are the DFT of the coefficients $\{b_k\}$ (cf. Equation (9.8)). Considering that $\{b_k\}$ are complex-valued, the optimal J must satisfy:*

$$\frac{\partial J}{\partial Re(b_k)} = 0$$
$$\frac{\partial J}{\partial Im(b_k)} = 0 \tag{9.37}$$

The vector $\mathbf{b} = \{b_k\}$ is then obtained by solving the linear system of equations [74]

$$\mathbf{Ab} = \bar{q} \tag{9.38}$$

where

$$\mathbf{A}_{m,l} = \sum_{n=0}^{L-1} \frac{e^{-j2\pi\frac{n(l-m)}{L}}}{|\lambda_n|^2 + \rho^{-1}} \quad 1 \leqslant m, l \leqslant \nu \tag{9.39}$$

$$\bar{q}_{m,l} = \sum_{n=0}^{L-1} \frac{e^{j2\pi\frac{nm}{L}}}{|\lambda_n|^2 + \rho^{-1}} \quad 1 \leqslant m \leqslant \nu \tag{9.40}$$

9.4   Conclusion

This chapter addressed several open problems in the area of the asymptotic performance of linear- and decision-feedback equalizers. In the area of linear equalization, the DMT of the IIR zero-forcing equalizer as well as the FIR MMSE equalizer were studied. The diversity of the MMSE FIR equalizer in the fixed rate regime was calculated. The performance of decision-feedback equalizers was also studied under both IIR and FIR assumptions. In addition to the new results, this chapter also provides a self-contained introduction to spectral representations of random processes in the sampled-time domain and uses this technique to put the analysis of DFEs on a much firmer footing than was earlier available.

9.5   Appendix

9.5.1   Spectral Representation of Random Process

Spectral analysis of stochastic processes, and in particular the spectral representation, has a long and distinguished history. This appendix presents a brief and self-contained introduction, aimed at producing only the needed results for a communications engineer with a minimal amount of machinery. It is hoped that this approach leads to a more accessible treatment of the subject. The authors must acknowledge their debt to two excellent texts [71, 75] whose sharpness and economy of expression has been very inspiring.

A deterministic, periodic discrete-time signal $x(k)$ under certain conditions admits a Fourier series:

$$x(k) = \sum_{r=0}^{\infty} \alpha_r \cos(2\pi k/N) + \beta_r \sin(2\pi k/N)$$

Non-periodic functions, on the other hand, are more suited to the Fourier transform

$$x(k) = \int_{-\pi}^{\pi} \{A(\omega)\cos(\omega k) + B(\omega)\sin(\omega k)\}\, d\omega$$

The two forms were combined and generalized in 1930 by Wiener [72] into a *Fourier-Stieltjes integral*[5] to produce a generalized harmonic analysis:

$$x(k) = \int_{-\pi}^{\pi} e^{j\omega k} \, dX(\omega). \tag{9.41}$$

Clearly, if $X(\omega)$ is a piecewise constant function, we have a Fourier series while a continuous and differentiable $X(\omega)$ yields the usual Fourier transform integral, where then $X(\omega)$ above would be the integral of the Fourier spectrum. The beauty of this representation is that it also applies to cases where $X(\omega)$ is neither piece-wise constant nor differentiable, and in that way provides a perfect vehicle for a spectral characterization of stationary random processes.

Spectral Representation of Stationary Random Process

Now, with an abuse of notation, consider a stationary random process $x(k)$. The sample paths of this process do not admit a Fourier series because they are not periodic with probability 1, and also do not admit a Fourier integral since with probability 1 they are not absolutely integrable. However, a *spectral process* $X(\omega)$ exists [76] such that

$$x(k) = \int_{-\pi}^{\pi} e^{j\omega k} \, dX(\omega). \tag{9.42}$$

It is easy to show then that among random processes, linear operations such as

$$z(k) = x(k) + y(k)$$

are conserved in the spectral domain, i.e.,

$$dZ(\omega) = dY(\omega) + dX(\omega).$$

If $x(k)$ and $y(k)$ are respectively the input and output of a discrete-time linear time-invariant filter $h(k)$, that is,

$$x(k) = \sum_{-\infty}^{\infty} h(\ell) x(k - \ell). \tag{9.43}$$

---

[5]In fact Wiener's generalization did not even require $X(\omega)$ to be of bounded variation, but for our purposes this statement suffices.

it is straight forward to show:

$$
\begin{aligned}
y(k) &= \sum_{-\infty}^{\infty} h(\ell) \int_{-\pi}^{\pi} e^{j\omega(k-\ell)} \, dX(\omega) \\
&= \int_{-\pi}^{\pi} e^{j\omega k} \left\{ \sum_{-\infty}^{\infty} h(\ell) e^{j\omega \ell} \right\} dX(\omega) \\
&= \int_{-\pi}^{\pi} e^{j\omega k} H(\omega) \, dX(\omega).
\end{aligned}
\tag{9.44}
$$

Thus, we have the useful property of filtering in the spectral representation:

$$
dY(\omega) = H(\omega) dX(\omega)
\tag{9.45}
$$

We now proceed to describe the connection of this spectral representation to the Fourier transform of the autocorrelation, which will enable us to use spectral factorization and design the needed filters for estimation.

Definition 9.5.1 *A continuous-parameter random process $X(\omega)$ is an orthogonal increments process if it satisfies:*

$$
\mathbb{E}\left[\left(X(\omega_1) - X(\omega_2) \right) \left(X^*(\omega_3) - X^*(\omega_4) \right) \right] = 0.
\tag{9.46}
$$

whenever $[\omega_1, \omega_2) \cap [\omega_3, \omega_4) = \emptyset$

Theorem 9.5.1 *(Doob [76, Chapter X, Theorem 4.1]) For every discrete-time wide-sense stationary random process $x(k)$ there exists an essentially unique spectral representation*

$$
x(k) = \int_{-\pi}^{\pi} e^{j\omega k} \, dX(\omega)
$$

where $X(\omega)$ process has orthogonal increments.

This result is surprisingly fussy to show, especially when $X(\omega)$ is discontinuous at the boundaries. However, once established, the orthogonal increment property of the spectral process

has interesting consequences for the autocorrelation function:

$$
\begin{aligned}
R_x(n) &= \mathbb{E}\big[x(k)x^*(k-n)\big] \\
&= \mathbb{E}\bigg[\int_{-\pi}^{\pi} e^{j\omega k}dX(\omega) \ \int_{-\pi}^{\pi} e^{-j\omega'(k-n)}dX^*(\omega')\bigg] \\
&= \mathbb{E}\bigg[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega-\omega')k}e^{j\omega' n}dX(\omega)dX^*(\omega')\bigg] \\
&= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{j(\omega-\omega')k}e^{j\omega' n}\,\mathbb{E}\big[dX(\omega)dX^*(\omega')\big]
\end{aligned}
\tag{9.47}
$$

Since $\mathbb{E}\big[dX(\omega)dX^*(\omega')\big]$ has orthogonal increment then $\mathbb{E}\big[dX(\omega)dX^*(\omega')\big] = 0$ unless $\omega = \omega'$ and thus

$$
\mathbb{E}\big[dX(\omega)dX^*(\omega')\big] = \delta_{\omega\omega'}d\bar{S}_x(\omega),
\tag{9.48}
$$

where $\delta_{\omega\omega'}$ is the Kroenecker delta function. Equivalently:

$$
\mathbb{E}\big[|dX(\omega)|^2\big] = d\bar{S}_x(\omega).
$$

The properties of $\bar{S}_x$ are clarified by the celebrated Wiener-Kintchine theorem:

Theorem 9.5.2 *(Wiener-Kintchine) A necessary and sufficient condition for R_x to be the autocorrelation function of some stationary process is that there exists a bounded, positive non-decreasing function $\bar{S}_x(\omega)$ such that for all n,*

$$
R_x(n) = \int_{-\pi}^{\pi} e^{j\omega n}d\bar{S}_x(\omega)
$$

Remark 9.5.1 *When $R_x(n)$ dies fast enough, $\bar{S}_x(\omega)$ is absolutely continuous and a spectral density may be defined.*

$$
R_x(n) = \int_{-\pi}^{\pi} e^{j\omega n}S_x(\omega)d\omega
\tag{9.49}
$$

The latter is the case that is often represented in the engineering literature. However, the main point of the Wiener-Kintchine theorem is the generality of a spectral representation of the autocovariance, whose existence is guaranteed via Bochner's theorem.

To summarize, for each stationary random process $x(n)$ there exists a spectral process $X(\omega)$ with the usual linearity and convolution properties, and furthermore it is related to the spectral density function via $E[|dX(\omega)|^2] = S_x(\omega)d\omega$. This machinery is sufficient to fully support the developments of this chapter.

CHAPTER 10
CONCLUSIONS AND FUTURE WORK

This chapter summarizes the contributions of this dissertation and provides some possible avenues for future directions. The findings also appear in $[5, 46, 56, 77–81]$

Chapter 3 provides the first contribution in this dissertation. The work settles the long standing problem of the diversity of the MMSE MIMO receivers under all fixed rates for any number of transmit (M) and receive (N) antennas, giving the result as $d = \lceil M2^{-\frac{R}{M}} - \kappa \rceil^2 + |N - M| \lceil M2^{-\frac{R}{M}} - \kappa \rceil$, where $\kappa = max(0, M - N)$. The analysis confirms the earlier approximate results $[6, 7]$ showing that the system diversity can be as high as MN for low spectral efficiency and as low as $N - M + 1$ for high spectral efficiency. The result is easily extended to the multiple access channel (MAC).

Chapter 4 extends the results of the previous chapter and studies the performance of the MIMO MMSE receiver in the frequency selective channel under two common transmission schemes, the zero-padding and the cyclic-prefix tranmission. The explicit tradeoff between rate and diversity in these two cases are provided.

Chapter 5 studies the high-SNR performance of MIMO linear precoding. It is shown that the zero-forcing precoder under two common design approaches, maximizing the throughput and minimizing the transmit power, achieves the same diversity function as that of MIMO systems with ZF equalizer. When a regularized ZF (RZF) precoder (for a fixed regularization term that is independent of the signal-to-noise ratio) or matched filter (MF) precoder is used, we have $d(r) = 0$ for all r, implying an error floor under all conditions. However, in the fixed rate regime, RZF and MF filtering achieve full diversity up to a certain spectral efficiency, while at higher spectral efficiencies they produce an error floor. If the regularization parameter in the RZF is optimized in the MMSE sense (also known as Wiener filter precoding), the

RZF precoded MIMO system exhibits a complex rate-dependent behavior. In particular, the diversity of this system is characterized by $d(R) = \lceil N2^{-\frac{R}{N}} \rceil^2 + (M - N)\lceil N2^{-\frac{R}{N}} \rceil$ where M and N are the number of transmit and receive antennas. This is the same behavior observed in linear MMSE MIMO receivers in Chapter 3.

Chapter 6 extends the results of the previous chapters and analyze the case of MIMO systems with linear transmit and receive filters in the flat fading channel. It is shown that using linear equalization and precoding can remove the error floor observed when MF or RZF precoding is used.

Chapter 7 provides the analysis for common transmit diversity techniques when used with MMSE receiver. Alamouti and cyclic delay diversity schemes are analyzed and compared.

In order to improve the performance of the MIMO linear receivers, lattice-reduction aided equalization has been proposed in the literature for MIMO flat fading and SISO frequency selective channels. Chapter 8 analyzes the LR-aided equalization for MIMO frequency selective channel. It is shown that the maximum diversity is obtained for all spectral efficiencies.

Chapter 9 provides the analysis for the decision-feedback equalizer. As part of the developments of this chapter, the notion of the *spectral representation* of random processes is used to provide a rigorous analytical-framework for decision feedback equalizers.

Future work may be pursued for projects related to the work discussed in this dissertation. We provide some of these projects as well as other possible future work below.

- The current research mainly focused on linear filtering. Similar analysis is also needed for non-linear operations such as decision-feedback and successive interference cancellation techniques. The diversity analysis of these feedback systems is interesting and it is shown that error propagation which is one of the main challenges in analyzing feedback systems, can be taken into consideration and exact results are obtainable.

- It is also interesting to investigate the DMT (as well as the diversity in the fixed rate regime) of infinite impulse response (IIR) and finite (FIR) equalizers in SISO

channel as well as zero-padding and cyclic-prefix block transmission in MIMO channel. Preliminary results obtain the following:

$$d_{\text{MMSE}}^{\text{FIR}}(r) = \begin{cases} 1 - r & r \in (0, 1] \\ \nu + 1 & r = 0 \end{cases} \tag{10.1}$$

$$d_{\text{ZF}}^{\text{FIR}}(r) = 1 - r \quad r \in [0, 1] \tag{10.2}$$

$$d_{\text{MMSE}}^{\text{CP}}(r) = \begin{cases} 1 - r & r \in (0, 1] \\ \min\left(\nu, \lfloor N2^{-R} \rfloor\right) + 1 & r = 0 \end{cases} \tag{10.3}$$

$$d_{\text{ZF}}^{\text{CP}}(r) = (1 - r) \quad r \in [0, 1] \tag{10.4}$$

REFERENCES

[1] Z. Wang and G. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389 – 1398, Aug. 2003.

[2] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[4] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[5] A. Hesham Mehana and A. Nosratinia, "Diversity of MMSE MIMO receivers," in *Proc. IEEE ISIT*, June 2010.

[6] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Signal Processing*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007.

[7] K. R. Kumar, G. Caire, and A. L. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.

[8] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 271–286, Jan. 2006.

[9] A. L. Moustakas, K. R. Kumar, and G. Caire, "Performance of MMSE MIMO receivers: A large n analysis for correlated channels," in *Proc. IEEE Vehicular Technology Conference (VTC)*, Apr. 2009, pp. 1–5.

[10] E. A. Jorswieck and H. Boche, "Information theory outage probability in multiple antenna system," *European Trans. on Telecommun.*, vol. 18, no. 3, pp. 217–233, June 2006.

[11] Y. Jiang, M. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An In-Depth study of the high SNR regime," *IEEE Trans. Inform. Theory*, vol. 57, no. 4, pp. 2008–2026, Apr. 2011.

[12] E. N. Onggosanusi, A. G. Dabak, T. Schmidl, and T. Muharemovic, "Capacity analysis of frequency-selective MIMO channels with sub-optimal detectors," in *Proc. IEEE ICASSP*, vol. 3, May 2002, pp. 2369–2372.

[13] H. Gao, P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in rayleigh-fading additive interference channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 666–672, May 1998.

[14] D. Gore, A. Gorokhov, and A. Paulraj, "Joint MMSE versus V-BLAST and antenna selection," in *Proc. Asilomar Conference on Signals, Systems and Computers*, vol. 1, Nov. 2002, pp. 505–509.

[15] A. Tajer and A. Nosratinia, "Diversity order in ISI channels with single-carrier frequency-domain equalizer," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1022 –1032, Mar. 2010.

[16] S.Verdu, *Multiuser Detection.* Cambridge University Press, 1998.

[17] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[18] A. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475–501, 1964.

[19] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley and Sons, 1991.

[20] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304 –306, July 2001.

[21] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial tv broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100 –109, Feb. 1995.

[22] B. Muquet, Z. Wang, G. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136 – 2148, Dec. 2002.

[23] Z. Wang, X. Ma, and G. Giannakis, "Optimality of single-carrier zero-padded block transmissions," in *Proc. IEEE Wireless Communcations and Networking Conference (WCNC)*, vol. 2, Mar. 2002, pp. 660 – 664.

[24] A. Scaglione, P. Stoica, S. Barbarossa, G. 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.

[25] C. Tepedelenlioglu, "Low complexity linear equalizers with maximum multipath diversity for zero-padded transmissions," in *Proc. IEEE ICASSP*, vol. 4, Apr. 2003, pp. 636–639.

[26] L. Grokop and D. Tse, "Diversity-multiplexing tradeoff in ISI channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 1, pp. 109 –135, Jan. 2009.

[27] I. Gradshteyn and I. Ryzbik, *Tables of Integrals, Series, and Products, 6th ed.* Academic Press., 2000.

[28] A. Stamoulis, S. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2451 – 2464, Oct. 2002.

[29] A. Kaveh and H. Rahami, "Block circulant matrices and applications in free vibration analysis of cyclically repetitive structures," *Acta Mechanica*, pp. 1–12, 2010. [Online]. Available: http://dx.doi.org/10.1007/s00707-010-0382-x

[30] S. Bradely, A. Hax, and T. Magnanti, *Applied Mathematical Programming.* Addison-Wesley, 1977, Ch 9, also available online at: web.mit.edu/15.053/www/.

[31] C. Tepedelenlioglu, "Maximum multipath diversity with linear equalization in precoded OFDM systems," *IEEE Trans. Inform. Theory*, vol. 50, pp. 232–235, Jan. 2004.

[32] S. Jayaweera and H. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2697 – 2709, Oct. 2003.

[33] M. Joham, W. Utschick, and J. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2700 – 2712, Aug. 2005.

[34] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 2007 –2019, Sept. 1999.

[35] M. Skoglund and G. Jongren, "On the capacity of a multiple-antenna communication link with channel side information," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 395 – 405, Apr. 2003.

[36] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 528 – 541, Mar. 2006.

[37] R. Esmailzadeh and M. Nakagawa, "Pre-RAKE diversity combination for direct sequence spread spectrum communications systems," in *IEEE International Conference on Communications, ICC*, vol. 1, May 1993, pp. 463 –467 vol.1.

[38] C. Peel, B. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195 – 202, Jan. 2005.

[39] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Processing*, vol. 56, no. 9, pp. 4409 –4418, Sept. 2008.

[40] D. Gore, J. Heath, R.W., and A. Paulraj, "On performance of the zero forcing receiver in presence of transmit correlation," in *IEEE International Symposium on Information Theory*, 2002, p. 159.

[41] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590 –3600, Nov. 2010.

[42] E. Sengul, E. Akay, and E. Ayanoglu, "Diversity analysis of single and multiple beamforming," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 990 –993, June 2006.

[43] D. S. Bernstein, *Matrix Mathematics: Theory, Facts and Formula.* Princeton University Press, 2009.

[44] F. Pancaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Processing Mag.*, vol. 25, no. 5, pp. 37 –56, Sept. 2008.

[45] D. Falconer, S. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58 –66, Apr. 2002.

[46] A. Hesham Mehana and A. Nosratinia, "The diversity of MMSE receiver over frequency-selective MIMO channel," in *Proc. IEEE ISIT*, July 2011.

[47] Z. Wang and G. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades?" in *Proc. IEEE Signal Processing Advances in Wireless Communication (SPAWC)*, 2001, pp. 267 –270.

[48] C. Tepedelenlioglu and Q. Ma, "On the performance of linear equalizers for block transmission systems," in *Proc. IEEE GLOBECOM*, vol. 6, Dec. 2005.

[49] Z. Wang and G. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 707 – 720, Mar. 2003.

[50] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE ICC*, vol. 3, May 1993, pp. 1630 –1634 vol.3.

[51] A. Dammann and S. Kaiser, "On the equivalence of space-time block coding with multipath propagation and/or cyclic delay diversity in OFDM," in *IEEE European Wireless*, 2002.

[52] G. Bauch and J. Malik, "Cyclic delay diversity with bit-interleaved coded modulation in orthogonal frequency division multiple access," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2092 –2100, Aug. 2006.

[53] U.-K. Kwon and G.-H. Im, "Cyclic delay diversity with frequency domain turbo equalization for uplink fast fading channels," *IEEE Commun. Lett.*, vol. 13, no. 3, pp. 184 –186, Mar. 2009.

[54] A. Dammann, "On antenna diversity techniques for OFDM systems," Ph.D. dissertation, Universität Ulm, Germany, June 2005, VDI Verlag Düsseldorf, Series 10, No. 766.

[55] S. Zhou and G. Giannakis, "Single-carrier space-time block-coded transmissions over frequency-selective fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 1, pp. 164 – 179, Jan. 2003.

[56] A. H. Mehana and A. Nosratinia, "Diversity of MMSE MIMO receivers," *IEEE Trans. Inform. Theory*, no. 10, pp. 6788–6805, Oct. 2012. [Online]. Available: http://arxiv.org/abs/1102.1462

[57] H. Yao and G. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE GLOBECOM*, vol. 1, Nov. 2002, pp. 424 – 428 vol.1.

[58] C. Windpassinger and R. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for mimo systems using lattice reduction," in *Proc. IEEE ITA*, Mar. 2003, pp. 345 – 348.

[59] M. Taherzadeh, A. Mobasher, and A. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4801 –4805, Dec. 2007.

[60] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, pp. 515–534, 1982.

[61] M. Taherzadeh, A. Mobasher, and A. Khandani, "Communication over MIMO broadcast channels using lattice-basis reduction," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4567 –4582, Dec. 2007.

[62] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 309 –318, Feb. 2008.

[63] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2701 –2710, July 2009.

[64] X. Ma, W. Zhang, and A. Swami, "Lattice-reduction aided equalization for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1608 –1613, Apr. 2009.

[65] C. Ling, "On the proximity factors of lattice reduction-aided decoding," *IEEE Trans. Signal Processing*, vol. 59, no. 6, pp. 2795 –2808, June 2011.

[66] N. Al-Dhahir and J. Cioffi, "Block transmission over dispersive channels: transmit filter optimization and realization, and MMSE-DFE receiver performance," *IEEE Trans. Inform. Theory*, vol. 42, no. 1, pp. 137 –160, Jan. 1996.

[67] C. Belfiore and J. Park, J.H., "Decision feedback equalization," *Proceedings of the IEEE*, vol. 67, no. 8, pp. 1143 – 1156, Aug. 1979.

[68] Y. Zhu and K. Letaief, "Single carrier frequency domain equalization with noise prediction for broadband wireless systems," in *Proc. IEEE GLOBECOM*, vol. 5, nov.-dec. 2004, pp. 3098 – 3102.

[69] A. Tajer, A. Nosratinia, and N. Al-Dhahir, "Diversity order of infinite-length symbol-by-symbol linear equalization," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2343–2348, Sept. 2011.

[70] J. Cioffi, G. Dudevoir, M. Vedat Eyuboglu, and G. D. Forney Jr, "MMSE decision-feedback equalizers and coding- part I: Equalization results," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2582 –2594, Oct. 1995.

[71] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*. Springer, 1984.

[72] N. Wiener, *Generalized Harmonic Analysis*. Acta Math., 1930.

[73] A. Sayed and T. Kailath, "A survey of spectral factorization methods," *Numerical Linear Algebra With Applications*, vol. 8, pp. 467 –1 182 496, 2001.

[74] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 947 –955, June 2002.

[75] M. B. Priesley, *Spectral Analysis and Time Series*. Academic Press, 1983.

[76] J. L. Doob, *Stochastic Processes*. John Wiley and Sons, 1953.

[77] A. Hesham Mehana and A. Nosratinia, "On the strange error behavior of MMSE precoding and equalizatio," in *Proc. IEEE ITA*, 2012.

[78] ——, "DMT of MMSE receiver in the frequency-selective MIMO channel," in *Proc. IEEE ICC*, June 2012.

[79] ——, "High-snr analysis of MIMO linear precoders," in *Proc. IEEE ISIT*, 2012.

[80] ——, "Single-carrier frequency-domain equalizer with multi-antenna transmit diversity," *IEEE Trans. Commun.*, 2012.

[81] ——, "Diversity of MIMO linear precoding," 2010, submitted to Trans. Inform. Theory. [Online]. Available: http://arxiv.org/pdf/1202.4736v2.pdf

VITA

Ahmed Hesham Mehana was born in Cairo, Egypt, in 1981. He received his B.S. and M.S. degrees from Cairo University in 2004 and 2007, respectively, both in Electrical Engineering. He served as teaching assistant in Cairo University and American University in Cairo during 2004-2007. In 2007-2008 he was Research and Teaching Assistant at the University of South Florida and in the summer 2008 he was a Research Assistant in Texas A&M at Qatar. In August 2008, he started his Ph.D. studies in electrical engineering at The University of Texas at Dallas. He was an intern at Research in Motion Co. Ltd during 2010. His interests include MIMO precoding, linear receivers, and interference management.