Transmitter Strategies for Closed-Loop MIMO-OFDM

A Thesis
Presented to
The Academic Faculty

by

Joon Hyun Sung

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

School of Electrical and Computer Engineering
Georgia Institute of Technology
July 2004

Transmitter Strategies for Closed-Loop MIMO-OFDM

Approved by:

Professor John R. Barry, Adviser

Professor J. Stevenson Kenney

Professor Ye (Geoffrey) Li

Professor Thomas Morley

Professor Steven W. McLaughlin

Date Approved: 6 July 2004

To my parents.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Dr. John Barry for guiding me during my stay at Georgia Tech. Without his guidance and encouragement, I would never completed this thesis. I have learned much from working with him. I am also grateful to Dr. McLaughlin and Dr. Li for serving in the reading committee. I thank Dr. Kenney and Dr. Morley for helpful suggestions and comments as members of the defense committee.

Many thanks to my colleagues in the Communications Theory Research Group: Renato, Abdallah, Badri, Aravind, Piya, Deric, and Arumugam. I really enjoyed discussions and arguments with them. I also thank my friends in our basketball team. Special thanks to Sangwoo, Seungyup, and Seungho, who encouraged me when I was discouraged by unfamiliar life abroad. I also would like to express my gratitude to the members in our tennis team, particularly to Kyuwon, Hyungseop, and Hyeunjun, whose leadership led to great accomplishments. My sincere thanks to all my friends in Atlanta for their support and friendship.

I thank Dr. Choong Woong Lee for his guidance during my master degree years in Seoul Nation University, Korea. I thank all my friends in Korea.

Most of all, I thank my family for their constant encouragement throughout my studies. My mother and father have provided me with endless support and love. I would thank my sister, brother-in-law, and niece.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis concerns communication across channels with multiple inputs and multiple outputs. Specifically, we consider the closed-loop scenario in which knowledge of the state of the multiple-input multiple-output (MIMO) channel is available at the transmitter. We show how this knowledge can be exploited to optimize performance, as measured by the zero-outage capacity, which is the capacity corresponding to zero outage probability. On flat-fading channels, a closed-loop transmitter allocates different powers and rates to the multiple channel inputs so as to maximize zero-outage capacity. Frequency-selective fading channels call for a combination of orthogonal-frequency-division multiplexing (OFDM) and MIMO known as MIMO-OFDM. This exacerbates the allocation problem because it multiplies the number of allocation dimensions by the number of OFDM tones. Fortunately, this thesis demonstrates that simple allocations are sufficient to approach the zero-outage capacity. These simple strategies exploit the tendency for random MIMO channels to behave deterministically as the number of inputs becomes large.

We propose two simple allocation strategies: the frequency-uniform-spectral-efficiency (FUSE) allocation and the fixed-rate (FIX) allocation. The FUSE allocation simplifies the power allocation by forcing each OFDM tone to have the same spectral frequency, so that the scope of the power allocation reduces to spatial channels at each tone. In the FIX allocation, the achievable rate of each scalar channel is fixed irrespective of channel, and the fixed rates are predetermined to match the fading statistics. As the number of antennas tends to infinity, we analytically show that the proposed allocations approach the channel's zero-outage capacity. We also show that

the convergence is fast so that the FIX and FUSE allocations closely approach the channel capacity for a finite number of antennas. Experimental results are provided to support the theoretical analysis.

We also consider the bit-allocation problem for the case where granularity considerations require that the rate be drawn from a discrete and finite set. The best allocation is based on the exhaustive search over all possible candidates for the bit allocation satisfying the granularity constraint. However, an exhaustive search is unnecessary. In fact, we demonstrate that the search can be restricted to a small set containing only a few well-chosen candidates, without significantly affecting the optimality of the search. In particular, on a flat-fading channel, a *binary search* (only two candidates) and a *fixed allocation* (only one candidate) perform very close to the optimal allocation, as shown by simulation. The binary-search and fixed-allocation strategies extend to a MIMO-OFDM system by applying them on a tone-by-tone basis. We provide the bit-error rate results on a MIMO-OFDM system with the proposed bit-allocation strategies to show that the performance promised by theoretical analysis is actually achieved.

CHAPTER 1

INTRODUCTION

1.1 MIMO Wireless Communications

Recently, there has been a dramatic and rapid growth in wireless communications from cellular phone service to high-definition television broadcasting. More and more information is sent through wireless channels, and the demand for data rate is getting higher and higher. Wireless channels are open to everybody, but this openness puts a strict limit on the bandwidth and transmit power. To support high data rate for limited bandwidth, we desire higher spectral efficiency.

In a multiple-input multiple-output (MIMO) channel, created by employing multiple antennas at the transmitter and receiver, the spectral efficiency dramatically increases. Pioneering work by Foschini [24] and Telatar [60] showed that the capacity of M_T-input M_R-output MIMO can be $\min(M_T, M_R)$ times larger than the single-input single-output (SISO) capacity. In other words, the multiple transmit antennas are used to multiplex data in space, where the gain by *spatial multiplexing* can be as large as $\min(M_T, M_R)$. A simplest form of spatial multiplexing is known as V-BLAST (vertical Bell Labs layered space-time) transmission [27], where independent layers of data are transmitted form each antenna.

For reliable communications, high spectral efficiency must be accompanied with low error rate. A major obstacle to reliability on wireless channels is fading, which refers to deep attenuation of channel amplitude due to the mobility of users and surrounding obstacles [53]. Traditionally, to mitigate the effects of fading, the receiver uses multiple antennas, a technology known as antenna diversity [53]. In a MIMO

system, we can obtain diversity from the $M_T M_R$ links between transmitter and receiver. If the links are statistically independent, the diversity gain can be as large as $M_T M_R$ compared to a SISO channel. However, achieving the diversity is not quite straightforward with multiple transmit antennas. Alamouti introduced a clever way to achieve the maximum ($M_T M_R$) diversity gain for two transmit antennas [1]. Since then, many space-time codes have been proposed, which effectively provide spatial diversity on a MIMO channel [56, 57].

Another source for higher spectral efficiency and stronger reliability is frequency selectivity. In wideband transmission, the channel response is frequency selective, and multiple copies of a transmitted symbol arrive at the receiver over several signaling intervals. Analogous to multiple antennas, we can harvest the diversity from frequency selectivity [8]. Hence, the diversity gain is huge in a frequency-selective MIMO channel.

Unfortunately, all the advantages of MIMO are not free. A dearest penalty would be the *interference* between signals simultaneously emitted from the multiple transmit antennas [6]. This interference considerably increases the detection complexity. For example, a maximum-likelihood (ML) detector suffers a significant increase in complexity, which exponentially grows with the number of transmit antennas [59]. To avoid the exponential growth in complexity, simpler solutions have been explored, such as spatial equalizers or space-time codes specially designed to simplify the detection process.

Resolving interference in MIMO is analogous to a traditional problem of equalizing the effects of inter-symbol interference (ISI) on a frequency-selective channel. We can hence extend well-known solutions for the equalizer to MIMO [6]. Particularly, a decision-feedback equalizer is widely used in MIMO detection. For example, BLAST uses a zero-forcing decision-feedback detector, a special case of successive interference cancellation [64]. However, spatial equalizer can cost a significant loss in diversity. If

the zero-forcing decision-feedback detection is used without any help from an outer channel code, the diversity gain is only at most M_R.

Another example of simpler MIMO techniques is orthogonal space-time block codes (STBC) [57], a generalization of Alamouti's code to any number of transmit antennas. Orthogonal codes reduce the detection complexity since a symbol can be detected without interference from others owing to the orthogonality. To maintain orthogonality, however, orthogonal STBC sacrifices spatial multiplexing gain, which is as large as unity (when $M_T = 2$) or less (when $M_T > 2$) rather than the full multiplexing gain $\min(M_T, M_R)$.

Generally speaking, there is a tradeoff among spatial multiplexing (increasing spectral efficiency), diversity (mitigating fading), and complexity (mitigating interference). For example, if a simple decision-feedback detector is used, the diversity gain is no greater than M_R, smaller than maximum gain $M_T M_R$, while it achieves full spatial multiplexing gain. Orthogonal STBC sacrifices multiplexing gain for simple detection and full diversity gain. In both cases, either multiplexing gain or diversity gain is sacrificed to reduce the detection complexity. On the contrary, if an ML detector is employed, the complexity is highest, but we can achieve full gains for diversity and spatial multiplexing [63].

1.2 Closed-Loop MIMO

We consider the closed-loop scenario in which knowledge of channel state information (CSI) is available at the transmitter. By exploiting CSI at the transmitter, eigenbeamforming converts a MIMO channel into a bank of scalar channels, with no crosstalk from one scalar channel to next [7, 12, 38]. Thus, the complexity for detection only linearly increases with the number of antennas. Eigenbeamforming is the optimal space-time processing in the sense that it achieves the capacity of a MIMO channel, attaining the full multiplexing gain. Furthermore, the diversity is also fully achieved,

as will be discussed in depth later. Surprisingly, the transmitter CSI magically makes the above-mentioned tradeoff ineffective and three goals of MIMO communications (multiplexing, diversity, and simple detection) are attained at the same time.

Clearly, a new problem arising in a closed-loop system is the availability of CSI at the transmitter. An intuitive way to furnish the transmitter with CSI is to send the estimated CSI to the transmitter. However, the feedback is redundant information, and the feedback delay might cause a mismatch problem. On a slowly varying channel, the delay is not a big issue, but the additional transmission load could be a problem. Thus, we have a new tradeoff between the availability of CSI feedback and the optimality of MIMO. In this work, we do not directly investigate this tradeoff problem, but try to answer in part by showing that the advantage of knowing CSI at the transmitter is significant to allow the redundancy by the feedback.

We base our research on information theory for fading channels [8]. On fading channels, the mutual information between transmitter and receiver is a random variable, and the capacity is either the expected value of mutual information (*average capacity*) or the rate that achieves a target outage probability (*outage capacity*), depending on the channel generation process [60]. CSI at the transmitter does not dramatically impact the average capacity, as reported for a SISO system [28]. The same is true for a MIMO channel, where there is a distinct advantage of knowing CSI at low signal-to-noise ratio (SNR), but the advantage almost disappears at high SNR [18, 24, 30, 60].

Meanwhile, the transmitter CSI helps improve the outage performance significantly, not only for the single-input case [8, 13] but also for MIMO [7]. We use the *zero-outage capacity*, also known as the delay-limited capacity [8, 32], to measure the outage performance, which is the maximum achievable rate while maintaining zero outage probability. If CSI is unknown to the transmitter, the outage probability cannot be made zero and thus the zero-outage capacity is zero [8]. In stark contrast, if

the transmitter knows CSI, it is possible to achieve a positive zero-outage capacity by controlling power at the transmitter to avoid any outage [13]. For some fading statistics, the zero-outage capacity is zero even when CSI is known at the transmitter, such as on Rayleigh-flat-fading SISO channel [28]. However, when there is diversity, such as from frequency selectivity or from multiple receive antennas, we can achieve a positive zero-outage capacity [2, 8].

1.3 Objective and Contributions

We use the zero-outage capacity as a performance criterion to develop efficient and low-complexity transmitter strategies for a closed-loop frequency-selective MIMO system with orthogonal-frequency-division multiplexing (OFDM) [47]. Contributions and organization of this thesis are as follows.

We begin by describing a MIMO-OFDM system in Chapter 2. OFDM transforms a MIMO channel with memory into a set of memoryless channels $\{\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_N\}$ over N tones for both the cases of spatially correlated and uncorrelated fading. The eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$ are crucial in the information-theoretical analysis. We summarize the properties of eigenvalues.

In Chapter 3, we consider a closed-loop MIMO system. A combination of OFDM and eigenbeamforming creates a bank of MN scalar channels. Each tone has M spatial channels with the eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$ as squared channel gains, where M is the rank of each memoryless MIMO channel. We also address efficient implementation of eigenbeamforming in a time-division duplex (TDD) system, and develop adaptive algorithms for updating receive filters.

Chapter 4 reviews previous results on average and outage capacity. We have two goals in this chapter: (1) to confirm the increased capacity and improved outage performance on a MIMO channel; (2) to show the advantage of knowing CSI at the transmitter. We explicitly show that CSI at the transmitter does not dramatically

increase the average capacity. In terms of the outage capacity, however, we can achieve a substantial gain from the transmitter CSI. At the end of Chapter 4, we also propose the procedures for optimizing outage region in a closed-loop MIMO system.

Chapter 5 summarizes our contributions to the high-SNR analysis on average and outage capacity. More specifically, we prove that a closed-loop MIMO system achieves full orders of diversity and multiplexing. We also rigorously derive the asymptotes of capacity in terms of the geometric mean of the eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$. The geometric-mean representation enables simple and insightful analysis on average and outage capacity. Most importantly, we analyze the zero-outage capacity with respect to channel memory L, showing that it is a nondecreasing function of L and quantifying the increase in terms of M_T, M_R, and L.

We consider the power-allocation problem in Chapter 6. To achieve the zero-outage capacity, the transmitter must find the power allocation, distributing power to each scalar channel. The optimal power allocation is based on water-filling over MN scalar channels, but requires high complexity. In this work, we propose two simpler allocation strategies: the *frequency-uniform-spectral-efficiency* (FUSE) allocation and the *fixed-rate* (FIX) allocation. The FUSE allocation reduces the complexity by restricting the scope of water-filling over M spatial channels for each tone, such that each tone achieves the same spectral efficiency. In the FIX allocation, we furthermore abandon water-filling by fixing the achievable rate for each scalar channel. Thus, power allocation can be calculated by a simple closed-form formula. The proposed allocation strategies significantly reduce the allocation complexity, but inevitably incur capacity penalties. However, we prove that the penalties of both allocations converge to zero as the number of antenna array size tends to infinity. We also quantify the penalties for a finite antenna array by high-SNR analysis and show that the penalties quickly converge to zero.

In Chapter 7, we investigate the bit-allocation problem, where a granularity constraint is forced on supportable rates when solving power allocation. The analysis in Chapter 6 provides useful insight to the bit-allocation problem. The best bit allocation is based on a full search over all possible candidates for the allocation, which is usually intractable for the complexity reason. We propose simple bit-allocation strategies for flat-fading MIMO channels, *Binary Search* and *Fixed Allocation*, by exploiting the properties of MIMO channels. Compared to the full-search allocation, the proposed strategies exhibit nearly optimal performance, while the complexity reduction is remarkable. We also extend the proposed strategies to MIMO-OFDM by treating each tone equally similar to the FUSE allocation in Chapter 6.

Finally, in Chapter 8, we evaluate the bit-error rate (BER) performance with the proposed bit-allocation strategies. We confirm that theoretical results in Chapter 6 and Chapter 7 also hold with quadrature amplitude modulation (QAM) constellations and practical outer channel codes. From the BER results, we will see that CSI at the transmitter indeed improves the performance critically.

To summarize, the main contributions of this work are as follows.

- We derive high-SNR asymptotes of capacity in terms of the geometric mean of the eigenvalues of the MIMO channel, and analyze the zero-outage capacity by using the properties of the geometric mean. (Chapter 5)

- We propose simple power-allocation strategies: the FUSE and Fixed allocations. We prove that the proposed allocations are optimal in terms of the zero-outage capacity as the number of antennas tends to infinity. We also show that the proposed allocations perform well at a moderate number of antennas by high-SNR analysis. (Chapter 6)

- We propose practical bit-allocation strategies, Binary Search and Fixed Allocation, for MIMO flat-fading channels, which have remarkably low complexity.

We show that both strategies are nearly optimal as the number of antennas grows. Extension to MIMO-OFDM is also considered. (Chapter 7)

1.4 Notation Summary

We summarize acronyms and mathematical notations.

Notation	**Description**
BER	Bit-error rate
CDF	Cumulative distribution function
CSI	Channel state information
DMT	Discrete multitone
FUSE	Frequency uniform spectral efficiency
i.i.d.	Independent and identically distributed
ISI	Inter-symbol interference
MIMO	Multiple input multiple output
MISO	Multiple input single output
MRC	Maximum-ratio combining
OFDM	Orthogonal-frequency-division multiplexing
PDF	Probability density function
RCSI	Receiver channel state information
SIMO	Single input multiple output
SISO	Single input single output
SNR	Signal-to-noise ratio
STBC	Space-time block code
SVD	Singular-value decomposition
TDD	Time-division duplex
TRCSI	Transmitter and receiver channel state information
$\|A\|$	Cardinality of a set A

$\|\|\mathbf{A}\|\|_F$	Frobenius norm of a matrix $\mathbf{A}$
$[\mathbf{A}]_{p,q}$	Element of a matrix $\mathbf{A}$ at the pth row and qth column
$1\{\cdot\}$	Indicator function
$A_{\mathcal{L}}$	Arithmetic mean over index set $\mathcal{L}$
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian random variable with i.i.d. $\mathcal{N}(\mu, \sigma^2/2)$ entries
$\det(\mathbf{A})$	Determinant of a square matrix $\mathbf{A}$
$\mathbb{E}[X]$	Expectation of a random variable X
$E_n(x)$	Generalized exponential integral
$G_{\mathcal{L}}$	Geometric mean over index set $\mathcal{L}$
$\mathbf{I}_M$	$M \times M$ identity matrix
L	Number of channel memory
M	Rank of MIMO channel matrix
$\tilde{M}$	Number of available channels in FIX Allocation (Chapter 6)
M_T	Number of transmit antennas
M_R	Number of receive antennas
$\mathcal{N}_s$	Spatial index set at any tone in MIMO-OFDM
$\mathcal{N}_u$	Universe index set for all scalar channels of MIMO-OFDM
$\{s_n^{(m)}\}$	A set of squared singular values of channel matrix
$\mathrm{tr}(\mathbf{A})$	Sum of diagonal elements of a square matrix $\mathbf{A}$
γ	Euler constant $\gamma \approx 0.577215665$
Γ	SNR gap for bit allocation
$\Gamma(x)$	Gamma function
$\Gamma(x, y)$	Incomplete complementary Gamma function
ρ	SNR per receive antenna
$\Psi_{\mathcal{L}}(x)$	Empirical distribution over index set $\mathcal{L}$

CHAPTER 2

MIMO CHANNELS AND PROPERTIES

This chapter describes a channel model for MIMO wireless channels. Among many factors that characterize wireless channels, we mainly consider two factors: fading and frequency selectivity. Fading refers to the severe attenuation in the channel amplitude caused by the combination of multipath propagation and receiver movement [53]. Unlike the time-invariant channel for wired communications, the wireless channel is time-varying and its amplitude is often too small to deliver information reliably. Frequency-selectivity is a typical phenomenon for wideband transmission. If the channel is frequency selective, the received signals are impaired by inter-symbol interference (ISI) [6]. Orthogonal-frequency-division multiplexing (OFDM) is an efficient solution to removing the distortion by ISI [9].

We first describe a MIMO channel model with memory, which reflects fading and frequency selectivity. Then, we show how OFDM removes ISI and converts a frequency-selective channel into a bank of flat-fading (memoryless) MIMO channels. Finally, we present important properties of MIMO fading channels.

2.1 MIMO Wireless Channels

We consider a discrete-time baseband model. Suppose that there are M_T transmit antennas and M_R receive antennas, which create an $M_R \times M_T$ multiple-input multiple-output (MIMO) channel. Let $\boldsymbol{x}_k = [x_k^{(1)}, \ldots, x_k^{(M_T)}]$ be the transmit signal vector at the kth signaling interval. If the channel has memory L, the received signal vector,

$\boldsymbol{y}_k$, is a linear combination of $\{\boldsymbol{x}_k, \boldsymbol{x}_{k-1}, \ldots \boldsymbol{x}_{k-L}\}$ [46], such that

$$\boldsymbol{y}_k = \sum_{l=0}^{L} \mathbf{G}_l \boldsymbol{x}_{k-l} + \boldsymbol{n}_k. \tag{1}$$

The $M_R \times M_T$ matrix $\mathbf{G}_l$ represents the MIMO channel at the lth delay. We assume white Gaussian noise, such that the elements of $\boldsymbol{n}_k$ in (1) are circularly symmetric complex Gaussian random variable with zero mean and $\mathbb{E}[\boldsymbol{n}_k \boldsymbol{n}_{k'}^*] = N_0 \delta_{k-k'} \mathbf{I}_{M_R}$, where $(\cdot)^*$ denotes the Hermitian transpose, $\mathbf{I}_M$ is an $M \times M$ identity matrix, and the Kronecker delta function δ_k is unity when $k = 0$ and zero otherwise.

The frequency response of (1) is

$$\mathbf{G}(e^{j\theta}) = \sum_{l=0}^{L} \mathbf{G}_l e^{-jl\theta} \quad -\pi < \theta \leq \pi, \tag{2}$$

where $j = \sqrt{-1}$. The response in (2) is frequency selective when $L > 0$. When $L = 0$, $\mathbf{G}(e^{j\theta})$ is flat in the frequency domain (flat fading), and the channel in (1) reduces to a memoryless channel.

We assume Rayleigh fading, unless specified otherwise, throughout the thesis, so that each element of $\mathbf{G}_l$ is a complex Gaussian random variable [53]. Elements of each $\mathbf{G}_l$ can be spatially uncorrelated or correlated, but we assume for analytical simplicity that there is no correlation between channel taps, that is, elements of $\mathbf{G}_l$ and elements of $\mathbf{G}_{l'}$ are uncorrelated if $l \neq l'$.

2.1.1 Spatially Uncorrelated Channel

If the channel is spatially uncorrelated, the channel matrix is given by $\mathbf{G}_l = \sigma_l \mathbf{W}_l$, where $\{\sigma_l^2\}$ denote the power profile [53]. The $M_R \times M_T$ matrix $\mathbf{W}_l$ denotes a spatially uncorrelated matrix, whose elements are i.i.d. $\mathcal{CN}(0, 1)$, where $\mathcal{CN}(\mu, \sigma^2)$ denotes a circularly symmetric complex Gaussian random variable, whose real and imaginary parts are i.i.d. with mean μ and variance $\sigma^2/2$ for each. Without loss of generality, we normalize the channel, such that $\sum_{l=0}^{L} \sigma_l^2 = 1$.

Definition 2.1. The power profile $\{\sigma_l^2\}$ is said to be *uniform* when $\sigma_0^2 = \sigma_1^2 = \ldots = \sigma_L^2 = 1/(L+1)$. ♦

The spatially uncorrelated model is useful for mathematical analysis. Particularly, the random matrix theory for i.i.d. Gaussian elements [23] provides powerful tools for analysis.

2.1.2 Spatial Correlated Channel

For a spatially correlated channel, we introduce correlation matrices $\mathbf{R}_{T_l}$ $(M_T \times M_T)$ and $\mathbf{R}_{R_l}$ $(M_R \times M_R)$, which represent correlation at the transmitter and receiver, respectively, for the lth delay, such that $\mathbb{E}\left[[\mathbf{H}_l]_{p,j}[\mathbf{H}_l]_{q,k}^*\right] = [\mathbf{R}_{T_l}]_{j,k}[\mathbf{R}_{R_l}]_{p,q}$ [18], where $[\mathbf{A}]_{p,q}$ denotes the element of a matrix $\mathbf{A}$ at the pth row and the qth column. We assume that the correlation between the fading from transmit antennas j and k to a particular antenna is $[\mathbf{R}_{T_l}]_{j,k}$ and does not depend on the receive antenna. The same is true for $[\mathbf{R}_{R_l}]_{p,q}$. Then, $\mathbf{G}_l$ can be factored in the form:

$$\mathbf{G}_l = (\mathbf{R}_{R_l})^{\frac{1}{2}}\mathbf{W}_l(\mathbf{R}_{T_l})^{\frac{1}{2}}, \tag{3}$$

where $(\cdot)^{\frac{1}{2}}$ denotes the matrix square root, and $\mathbf{W}_l$ is an $M_R \times M_T$ spatially uncorrelated matrix with i.i.d. $\mathcal{CN}(0,1)$ entries. We normalize the channel, such that $\sum_{l=0}^{L} \mathrm{tr}\{\mathbf{R}_{T_l}\} = M_T$ and $\sum_{l=0}^{L} \mathrm{tr}\{\mathbf{R}_{R_l}\} = M_R$, where $\mathrm{tr}\{\cdot\}$ denotes the trace of diagonal elements of a square matrix [33]. If the channel is uncorrelated, the correlation matrices reduce to $\mathbf{R}_{T_l} = \sigma_l^2\mathbf{I}_{M_T}$ and $\mathbf{R}_{R_l} = \sigma_l^2\mathbf{I}_{M_R}$

The degree of correlation is measured by $\sigma_{\theta_l}^2$, the variance of the angle spread for the lth path with its departure or arrival angle θ_l [10]. For the lth path, $\bar{\theta}_{T_l}$ and $\bar{\theta}_{R_l}$ denote average departure and arrival angles, respectively, and $\boldsymbol{a}$ is the array response vector, defined as

$$\boldsymbol{a}(\theta) = [1, e^{j2\pi\Delta cos(\theta)}, \ldots, e^{j2\pi(K-1)\Delta cos(\theta)}]^T, \tag{4}$$

where Δ is the antenna spacing relative to wavelength and where K is the number of antennas ($K = M_T$ for the transmitter and $K = M_R$ for the receiver). If there is no angle spread ($\sigma_{\theta_l}^2 = 0$) [10], the correlation matrices then become $\mathbf{R}_{T_l} = \sigma_l^2 \boldsymbol{a}(\bar{\theta}_{T_l})\boldsymbol{a}^T(\bar{\theta}_{T_l})$ and $\mathbf{R}_{R_l} = \sigma_l^2 \boldsymbol{a}(\bar{\theta}_{R_l})\boldsymbol{a}^T(\bar{\theta}_{R_l})$, such that, each tap $\mathbf{H}_l$ can be written as [47]

$$\mathbf{H}_l = \sigma_l^2 \alpha_l \boldsymbol{a}(\bar{\theta}_{R_l})\boldsymbol{a}^T(\bar{\theta}_{T_l}), \tag{5}$$

where α_l is i.i.d. $\mathcal{CN}(0,1)$, representing Rayleigh fading, and where $\{\sigma_l^2\}$ is the power profile.

For nonzero angle spread, we consider correlation only at the receiver, that is, $\mathbf{R}_{T_l} = \mathbf{I}_{M_T}$ [10]. The receive correlation matrix can be approximated as [3]

$$[\mathbf{R}_{R_l}]_{p,q} \approx \sigma_l^2 e^{-j2\pi|p-q|\Delta\cos(\overline{\theta}_{R_l})} e^{-0.5(2\pi|p-q|\Delta\sin(\overline{\theta}_{R_l})\sigma_{\theta_l})^2}. \tag{6}$$

In fact, the approximation in (6) is accurate only for small angle spread, but it provides the correct trend for large spread [10]. Note that each $\mathbf{R}_{R_l}$ collapses to a rank-1 matrix, $\mathbf{R}_{R_l} = \sigma_l^2 \boldsymbol{a}(\bar{\theta}_{R_l})\boldsymbol{a}^T(\bar{\theta}_{R_l})$, when $\sigma_{\theta_l} = 0$.

2.2 MIMO-OFDM

If the channel is frequency selective, the received signals are distorted by ISI, which makes the detection of transmitted signals difficult [6]. OFDM has emerged as one of most efficient ways to remove such ISI [9]. In this section, we briefly review how a MIMO channel with memory in (1) becomes a set of memoryless MIMO channels by OFDM.

Figure 1 illustrates the block diagram for MIMO-OFDM. Suppose that the transmitter collects N symbol vectors: $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N\}$, where $\boldsymbol{u}_n = [u_n^{(1)}, \ldots, u_n^{(M_T)}]^T$. We regroup N signals $\{u_1^{(m)}, \ldots, u_N^{(m)}\}$ for $m = 1, 2, \ldots, M_T$ by the $NM_T \times NM_T$ permutation matrix $\hat{\mathbf{P}}_T$ in Figure 1. Each group of N signals is fed into the inverse discrete Fourier transform (IDFT) block, which produces $\{x_1^{(m)}, \ldots, x_N^{(m)}\}$. Then, we add a

cyclic prefix of length L, such that

$$\{\underbrace{x_{N-L+1}^{(m)}, \ldots, x_N^{(m)}}_{\text{cyclic prefix}}, x_1^{(m)}, \ldots, x_N^{(m)}\} \tag{7}$$

is the set of signals transmitted from the mth antenna after digital-to-analog (D/A) conversion and upconversion.

At the receiver, we remove the cyclic prefix from the received signals after down-conversion and analog-to-digital conversion (A/D), producing a group of N signals: $\{y_1^{(m)}, \ldots, y_N^{(m)}\}$ for $m = 1, 2, \ldots, M_R$. Discrete Fourier transform (DFT) block transforms each group into $\{v_1^{(m)}, \ldots, v_N^{(m)}\}$. After the permutation by $\hat{\mathbf{P}}_R$, the received signals are recollected such that $\boldsymbol{v}_n = [v_n^{(1)}, \ldots, v_n^{(M_R)}]^T$ for $n = 1, 2, \ldots, N$.

Equivalently, as shown in Figure 1, the received signals after DFT can be written as

$$\begin{aligned} \boldsymbol{v}_1 &= \mathbf{H}_1 \boldsymbol{u}_1 + \tilde{\boldsymbol{n}}_1 \\ \vdots \quad &\vdots \quad \vdots \qquad , \\ \boldsymbol{v}_N &= \mathbf{H}_N \boldsymbol{u}_N + \tilde{\boldsymbol{n}}_N \end{aligned} \tag{8}$$

where $\{\tilde{\boldsymbol{n}}_n\}$ are additive noise. Since DFT/IDFT and permutations are unitary processes, $\{\tilde{\boldsymbol{n}}_n\}$ are statistically identical to $\{\boldsymbol{n}_k\}$ in (1). The $M_R \times M_T$ matrix $\mathbf{H}_n$ represents the memoryless channel at the nth tone, where

$$\mathbf{H}_n = \mathbf{G}(e^{j2\pi n/N}) = \sum_{l=0}^{L} \mathbf{G}_l e^{-j2\pi ln/N}. \tag{9}$$

Note that $\mathbf{H}_n$ is identical for all n when the channel is memoryless ($L = 0$).

For the rest of the thesis, we will consider the MIMO-OFDM model of (8) instead of the underlying channel of (1).

Lemma 2.1. Given our assumption that the channel taps are uncorrelated, all $\{\mathbf{H}_n\}$ are statistically identical to each other. If fading is spatially uncorrelated, each entry of $\mathbf{H}_n$ is i.i.d. $\mathcal{CN}(0, 1)$, that is, all $\{\mathbf{H}_n\}$ is statistically identical to spatially uncorrelated Rayleigh flat fading.

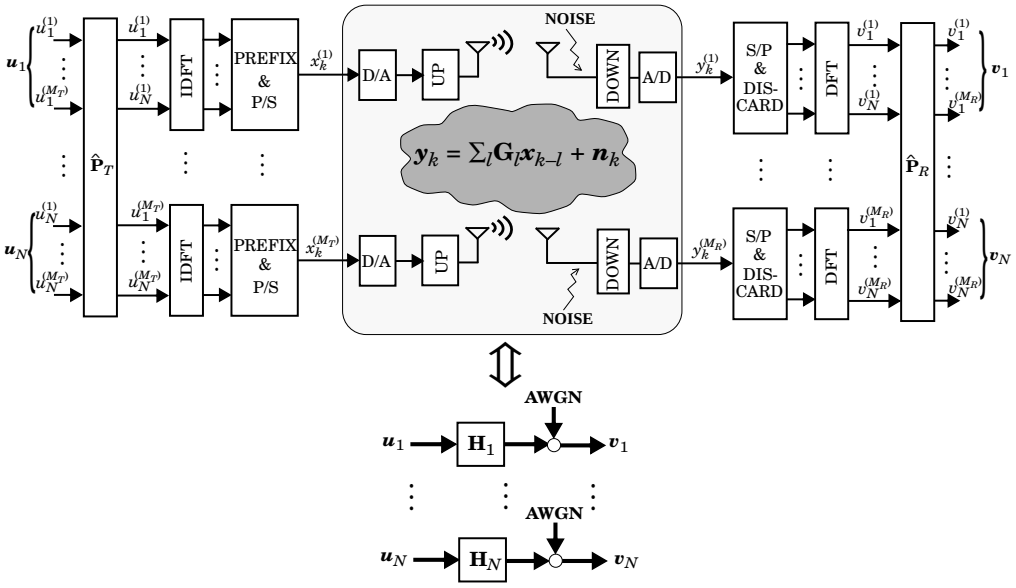

Figure 1: Block diagram for MIMO-OFDM and its equivalent discrete-time channel model.

Proof. From (3), the channel matrix at each tap is $\mathbf{G}_l = (\mathbf{R}_{R_l})^{\frac{1}{2}}\mathbf{W}_l$, where $\mathbf{R}_{R_l}^{1/2} = [\boldsymbol{r}_{l,1}, \cdots, \boldsymbol{r}_{l,M_R}]^T$ is the receiver correlation matrix and $\mathbf{W}_l = [\boldsymbol{w}_{l,1}, \cdots, \boldsymbol{w}_{l,M_T}]$ is an $M_R \times M_T$ matrix with i.i.d. $\mathcal{CN}(0,1)$ elements, with $\boldsymbol{r}_{l,i}$ and $\boldsymbol{w}_{l,i}$ are $M_R \times 1$ vectors. Then, the element of $\mathbf{H}_n$ at the p-th row and q-th column is

$$[\mathbf{H}_n]_{p,q} = \sum_{l=0}^{L} \boldsymbol{r}_{l,p}^T \boldsymbol{w}_{l,q} e^{-j2\pi ln/N}, \tag{10}$$

where superscript T denotes the transpose of a vector. Since $[\mathbf{H}_n]_{p,q}$ is a linear combination of zero-mean unit-variance complex Gaussian random variables in $\boldsymbol{w}_{l,p}$, $[\mathbf{H}_n]_{p,q}$ is also a complex Gaussian random variable. Clearly, its mean and variance $\mathbb{E}\left[[\mathbf{H}_n]_{p,q}\right] = 0$ and $\mathbb{E}\left[|[\mathbf{H}_n]_{p,q}]\right|^2 = \sum_{l=0}^{L} ||\boldsymbol{r}_{l,i}||^2$ are independent of n, where $||\cdot||$ is the Euclidean norm [33]. Since a Gaussian random variable is fully described by its mean and variance, it is sufficient to say that $\mathbf{H}_n$ has identical statistics for all n.

For spatially uncorrelated fading, the correlation matrix is $\mathbf{R}_{R_l} = \sigma_l^2 \mathbf{I}_{M_R}$. Thus, each element of $\mathbf{H}_n = \sum_{l=0}^{L} \sigma_l^2 \mathbf{W}_l e^{-j2\pi ln/N}$ is a Gaussian random variable with zero mean and variance $\mathbb{E}\left[\left|[\mathbf{H}_n]_{p,q}\right|^2\right] = \sum_{l=0}^{L} \sigma_l^2 = 1$. Also $\mathbf{H}_n$ inherits the spatially

uncorrelated property from $\{\mathbf{W}_l\}$. Thus, each $\mathbf{H}_n$ is statistically identical to a flat-fading MIMO channel ($L = 0$). ■

Corollary 2.1. Let M be the rank of $\mathbf{H}_n$, such that $M \leq \min(M_T, M_R)$. Then, M is constant for all n with probability one. If the channel is spatially uncorrelated, $M = \min(M_T, M_R)$ with probability one.

Proof. The proof is straightforward from Lemma 2.1. ■

2.3 Functions for Eigenvalues

Eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$ are crucial in the information-theoretical analysis. As $\mathbf{H}_n$ is a random matrix, the eigenvalues are also random variables. This section introduces some useful functions of eigenvalues for the future use in analysis.

Eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$ are closely related to *singular-value decomposition* (SVD) of $\mathbf{H}_n$.

Theorem 2.1 (SVD). For any complex $M_R \times M_T$ matrix $\mathbf{A}$ with rank M, there exists an SVD of the form:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^*, \tag{11}$$

where $\mathbf{U}$ ($M_R \times M_R$) and $\mathbf{V}$ ($M_T \times M_T$) are unitary. The $M_R \times M_T$ matrix $\mathbf{D} = [d_{i,j}]$ has $d_{i,j} = 0$ for all $i \neq j$, and $d_{1,1} \geq d_{2,2} \geq \ldots \geq d_{M,M} > d_{M+1,M+1} = \ldots = d_{a,a} = 0$, where $a = \min(M_T, M_R)$. Then, the *singular values* of $\mathbf{A}$, $\{d_{i,i}; i = 1, 2, \ldots, M\}$, are the positive square roots of eigenvalues of $\mathbf{A}\mathbf{A}^*$, and hence are uniquely determined.

Proof. See [33]. ■

The unitary matrices $\mathbf{U}$ and $\mathbf{V}$ are not unique. If $M_R \leq M_T$ and if $\mathbf{A}\mathbf{A}^*$ has distinct eigenvalues, then $\mathbf{V}$ is determined up to a right diagonal factor $\mathbf{T} = \text{diag}(e^{j\theta_1}, \ldots, e^{j\theta_n})$ with real-valued θ_i; that is, if $\mathbf{A} = \mathbf{U}_1\mathbf{D}\mathbf{V}_1^* = \mathbf{U}_2\mathbf{D}\mathbf{V}_2^*$, then $\mathbf{V}_2 = \mathbf{V}_1\mathbf{T}$. If $M_R < M_T$, then $\mathbf{V}$ is never uniquely determined; if $M_R = M_T = M$

and $\mathbf{V}$ is given, then $\mathbf{U}$ is uniquely determined. If $M_T \leq M_R$, the uniqueness of $\mathbf{U}$ and $\mathbf{V}$ is determined by considering $\mathbf{A}$ in a similar way of the case where $M_R \leq M_T$.

Corollary 2.2. Let $\{d_{m,m}\}$ be M nonzero singular values of $\mathbf{H}_n$. Then, $\{s_n^{(m)} = d_{m,m}^2\}$ are M nonzero eigenvalues of either $\mathbf{H}_n\mathbf{H}_n^*$ or $\mathbf{H}_n^*\mathbf{H}_n$.

Proof. See [33]. ■

2.3.1 Arithmetic and Geometric Means of Eigenvalues

We begin with well-known arithmetic and geometric means of $\{s_n^{(m)}\}$.

Definition 2.2. For a given index set $\mathcal{L}$ for (m, n), we define the arithmetic and geometric means of $\{s_n^{(m)}; (m, n) \in \mathcal{L}\}$ as:

$$A_{\mathcal{L}} = \frac{1}{|\mathcal{L}|} \sum_{(m,n)\in\mathcal{L}} s_n^{(m)}, \tag{12}$$

and

$$G_{\mathcal{L}} = \left(\prod_{(m,n)\in\mathcal{L}} s_n^{(m)} \right)^{1/|\mathcal{L}|}, \tag{13}$$

respectively, where $|\mathcal{L}|$ denotes the cardinality of $\mathcal{L}$. ♦

A well-known inequality is that $A_{\mathcal{L}} \geq G_{\mathcal{L}}$ with equality if and only if $s_n^{(m)}$ is independent of m and n for any index set $\mathcal{L}$ of m and n. Both $A_{\mathcal{L}}$ and $G_{\mathcal{L}}$ are random variables since they are just sum or product of $\{s_n^{(m)}\}$. We can relate $A_{\mathcal{L}}$ with the channel taps $\{\mathbf{G}_l\}$ of (1), as follows.

Lemma 2.2. If $\mathcal{N}_u$ is the universe index set:

$$\mathcal{N}_u = \{(m, n); m = 1, \ldots, M \text{and } n = 1, \ldots, N\}, \tag{14}$$

which encompasses all m and n, then we have

$$A_{\mathcal{N}_u} = \frac{1}{M} \sum_{l=0}^{L} ||\mathbf{G}_l||_F^2 = \frac{1}{M(L+1)} \sum_{l=0}^{L} ||\mathbf{W}_l||_F^2, \tag{15}$$

where $\mathbf{W}_l$ is a $M_R \times M_T$ matrix with i.i.d. $\mathcal{CN}(0,1)$ elements, and where $||\mathbf{A}||_F$ denotes the Frobenius norm of $\mathbf{A}$ defined as

$$||\mathbf{A}||_F = \sqrt{\mathrm{tr}\{\mathbf{A}\mathbf{A}^*\}}. \tag{16}$$

Proof. From the definition of arithmetic mean,

$$A_{\mathcal{N}_u} = \frac{1}{MN}\sum_{n=1}^{N}\sum_{m=1}^{M} s_n^{(m)} = \frac{1}{MN}\sum_{n=1}^{N} ||\mathbf{H}_n||_F^2. \tag{17}$$

Since $||\mathbf{H}_n||_F^2 = \sum_{p=1}^{M_R}\sum_{q=1}^{M_T} |[\mathbf{H}_n]_{p,q}|^2$, we have

$$\begin{aligned}
\sum_{n=1}^{N} ||\mathbf{H}_n||_F^2 &= \sum_{n=1}^{N}\sum_{p=1}^{M_R}\sum_{q=1}^{M_T} |[\mathbf{H}_n]_{p,q}|^2 \\
&= \sum_{p=1}^{M_R}\sum_{q=1}^{M_T}\sum_{n=1}^{N} \left|\sum_{l=0}^{L} [\mathbf{G}_l]_{p,q} e^{j2\pi ln/N}\right|^2 \\
&= \sum_{p=1}^{M_R}\sum_{q=1}^{M_T}\sum_{n=1}^{N} \Big\{ \sum_{l=0}^{L} |[\mathbf{G}_l]_{p,q}|^2 \\
&+ \sum_{l\neq l'} 2\mathrm{Re}\left[[\mathbf{G}_l]_{p,q}[\mathbf{G}_l]_{p,q}^* e^{-j2\pi(l-l')n/N}\right] \Big\}.
\end{aligned} \tag{18}$$

Since $\sum_{n=1}^{N} e^{-j2\pi(l-l')n/N} = 0$, (18) reduces to

$$\sum_{n=1}^{N} ||\mathbf{H}_n||_F^2 = \sum_{l=0}^{L}\sum_{p=1}^{M_R}\sum_{q=1}^{M_T}\sum_{n=1}^{N} |[\mathbf{G}_l]_{p,q}|^2 = N\sum_{l=0}^{L} ||\mathbf{G}_l||_F^2. \tag{19}$$

Substituting (19) into (17), we obtain (15). ■

Corollary 2.3. If the channel is spatially uncorrelated and the power profile is uniform, $M(L+1)A_{\mathcal{N}_u}$ is a chi-square random variable with a degree of freedom $M_T M_R(L+1)$ [39]. The probability density function (PDF) is

$$f_{M(L+1)A_{\mathcal{N}_u}}(x) = \frac{x^{M_T M_R(L+1)-1} e^{-x}}{\Gamma(M_T M_R(L+1))} \quad x \geq 0, \tag{20}$$

where

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \tag{21}$$

is the Gamma function [29].[1]

[1]When x is an integer, $\Gamma(x) = (x-1)!$, where ! denotes the factorial.

Proof. With the uniform power profile, $A_{\mathcal{N}_u} = \frac{1}{M(L+1)} \sum_{l=0}^{L} ||\mathbf{W}_l||_F^2$, where $\mathbf{W}_l$ has i.i.d. $\mathcal{CN}(0,1)$ entries. From the definition of the Frobenius norm, $A_{\mathcal{N}_u}$ is a chi-square random variable with a degree of freedom $M_T M_R(L+1)$. ■

2.3.2 Empirical Distribution

Another important function for $\{s_n^{(m)}\}$ is the empirical distribution function [30].

Definition 2.3. Let $\theta_n^{(m)} = s_n^{(m)}/M$. Then, the empirical distribution of $\{\theta_n^{(m)}\}$ over an index set $\mathcal{L}$ is given by

$$\Psi_{\mathcal{L}}(x) = \frac{1}{|\mathcal{L}|} \sum_{(m,n)\in\mathcal{L}} 1\{\theta_n^{(m)} \leq x\}, \tag{22}$$

where $1\{\cdot\}$ is an indicator function, such that $1\{A\}$ is unity if the condition A is satisfied or zero otherwise. ♦

Empirical distribution is a random variable when M is finite, but it is known to converge to a non-random limit as $M \to \infty$. For a spatially uncorrelated channel with $L = 0$ memory, we can explicitly evaluate the limit of $\Psi_{\mathcal{N}_u}(x)$, where $\mathcal{N}_u$ is the universe index set in (14), as follows.

Theorem 2.2. We consider a spatially uncorrelated channel with $L = 0$ memory. Suppose that M tends to infinity such that $\frac{\min(M_T,M_R)}{\max(M_T,M_R)} \to \beta \leq 1$. Then,

$$\begin{aligned} \theta_n^{(1)} &\to a \triangleq (1+\sqrt{\beta})^2 \\ \theta_n^{(M)} &\to b \triangleq (1-\sqrt{\beta})^2 \end{aligned}, \tag{23}$$

with probability one, for any n. In fact, the entire empirical distribution of a randomly selected eigenvalue converges. Also, the empirical distribution converges to a non-random limit, such that

$$d\Psi_{\mathcal{N}_u}(x) = \begin{cases} \frac{1}{2\pi\beta x}\sqrt{(x-a)(b-x)} & a \geq x \geq b \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

Proof. For a tutorial, see [30]. The proof for this theorem can be found in [5] for the smallest eigenvalue and [4] for the largest eigenvalue, respectively. For the last statement, see for example [35]. ■

From Lemma 2.1, we deduce that $\Psi_{\mathcal{N}_u}(x)$ also converges to a non-random limit when $L > 0$.

2.3.3 Joint Distribution of Eigenvalues

When the channel is spatially uncorrelated, Lemma 2.1 states that each $\mathbf{H}_n$ is statistically identical to $\mathbf{W}$, a random matrix with i.i.d. $\mathcal{CN}(0,1)$ in its entries. In mathematics, $\mathbf{W}\mathbf{W}^*$ is called a Wishart matrix [23, 30]. The eigenvalues of a Wishart matrix has been extensively studied. Particularly, the joint PDF of $\{s_n^{(m)}\}$ is known [23, 30]. For notational simplicity, we discard the tone index n in this section, since $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ are statistically identical for any n.

Theorem 2.3 (Ordered Eigenvalue Distribution). The joint distribution of $\boldsymbol{s} = [s^{(1)}, \ldots, s^{(M)}]$ is

$$f_{\boldsymbol{s}}(x_1, x_2, \ldots, x_M) = \frac{\pi^{M(M-1)}}{\Gamma_{M_T}(M_R)\Gamma_{M_R}(M_T)} \exp\left(\sum_{u=1}^{M} x_u\right) \prod_{u=1}^{M} x_u^D \prod_{u<v} (x_u - x_v)^2, \quad (25)$$

where $D = \max(M_T, M_R) - \min(M_T, M_R)$, and where

$$\Gamma_m(x) = \pi^{m(m-1)/x} \prod_{i=1}^{m} \Gamma(x - i + 1). \quad (26)$$

Proof. See [30]. ■

Corollary 2.4. Let $\mathbf{A}$ and $\mathbf{B}$ be $a \times b$ and $b \times a$ matrices, respectively, with $\mathcal{CN}(0,1)$ in their entries. The eigenvalues of $\mathbf{A}\mathbf{A}^*$ and $\mathbf{B}\mathbf{B}^*$ have identical joint PDF.

Proof. The proof is straightforward from the symmetry of (25) with respect to M_T and M_R. ■

If we randomize the order of $\{s^{(m)}\}$, we have a simpler distribution function.

Theorem 2.4 (Unordered Eigenvalue Distribution). Let s be a randomly selected eigenvalue from $\{s^{(m)}\}$. Then, its PDF is given by

$$f_s(x) = \frac{1}{M} \sum_{k=0}^{M-1} \frac{k! x^D e^{-x}}{(k+D)!} \left\{ L_k^D(x) \right\}, \tag{27}$$

where $L_k^m(x)$ is the Laguerre polynomial of order k [29], defined as

$$L_k^m(x) = \frac{e^x x^{-m}}{k!} \frac{d^k}{dx^k} \left\{ e^{-x} x^{k+m} \right\} = \sum_{i=0}^{k} (-1)^i \binom{k+m}{k-i} \frac{x^i}{i!}. \tag{28}$$

Proof. See [52]. ■

Theoretically, it is possible to derive the marginal distribution by evaluating $(M-1)$-fold integration of (25). The required integrals, however, quickly become intractable. To our knowledge, marginal PDF is known only for small M_T and M_R.

Example 2.1. When $M_T = M_R = 2$ and $M_T = M_R = 3$, marginal distribution functions are:

$$\begin{aligned} f_{s^{(1)}}(x_1) &= (2 - 2x_1 + x_1^2) e^{-x_1} \\ f_{s^{(2)}}(x_2) &= 2e^{-2x_2}, \end{aligned} \tag{29}$$

and,

$$\begin{aligned} f_{s^{(1)}}(x_1) &= \frac{1}{4}(12 - 24x_1 + 24x_1^2 - 8x_1^3 + x_1^4) e^{-x_1} \\ &- \frac{1}{2}(12 - 12x_1 + 6x_1^2 + 2x_1^3 + x_1^4) e^{-2x_1} + 3e^{-3x_1} \\ f_{s^{(2)}}(x_2) &= \frac{1}{2}(12 - 12x_2 + 6x_2^2 + 2x_2^3 + x_2^4) e^{-2x_2} - 6e^{-3x_2} \\ f_{s^{(3)}}(x_3) &= 3e^{-3x_3}, \end{aligned} \tag{30}$$

respectively [65]. ♦

In communications, the largest eigenvalue $s^{(1)}$ is of particular interest since it carries the largest amount of information. A general form for the marginal CDF of $s^{(1)}$ is known [30].

Theorem 2.5. The marginal CDF of the largest eigenvalue (s_1) is given by

$$\text{Prob}[s^{(1)} < x] = \frac{\Gamma_{M_R}(M_R)}{\Gamma_{M_R}(M_T + M_R)} x^{M_T M_R} {}_1F_1(M_T; M_T + M_R; -x\mathbf{I}), \tag{31}$$

where ${}_1F_1(;;)$ is the hypergeometric function of matrix argument [29].

Proof. See [30]. ■

The hypergeometric function of matrix argument is extremely difficult to compute as it is represented by slowly-converging zonal polynomials [30]. Instead of (31), we can directly derive the marginal distribution functions for some special cases.

Example 2.2. When $M_T = 1$ and $M_R \geq 1$, namely a single-input multiple-input (SIMO) channel and fading is spatially uncorrelated, there is only one nonzero eigenvalue ($M = 1$), and (31) reduces to

$$f_{s^{(1)}}(x) = \frac{x^{M_R - 1} e^{-x}}{\Gamma(M_R)}. \tag{32}$$

When $M_T \geq 1$ and $M_R = 1$, namely a multiple-input single-output (MISO) channel, (32) is valid with M_T replacing M_R. ♦

Example 2.3. When $M_T = M_R = M = 4$, the marginal PDF of the largest eigenvalue is

$$f_{s^{(1)}}(x) = \sum_{k=1}^{M} \varphi_k(x) e^{-kx}, \tag{33}$$

where

$$\begin{aligned}
\varphi_1(x) &= 4 - 12x + 18x^2 - \frac{34}{3}x^3 + \frac{7}{2}x^4 - \frac{1}{2}x^5 + \frac{1}{36}x^6 \\
\varphi_2(x) &= -12 + 24x - 24x^2 + \frac{8}{3}x^3 - \frac{1}{2} \\
\varphi_3(x) &= 12 - 12x + 6x^2 + \frac{14}{3}x^3 + \frac{23}{6}x^4 + \frac{5}{6}x^5 + \frac{1}{12}x^6 \\
\varphi_4(x) &= -4.
\end{aligned} \tag{34}$$

♦

On the other hand, the marginal distribution of the smallest eigenvalue is known for $M_T = M_R$.

Example 2.4. When $M_T = M_R = M$, the marginal PDF of the smallest eigenvalue $s^{(M)}$ is [23]

$$f_{s^{(M)}}(x) = Me^{-Mx}. \tag{35}$$

Note that the smallest eigenvalue is exponentially distributed, which means that $\mathbb{E}[1/s^{(M)}] \to \infty$. ♦

CHAPTER 3

CLOSED-LOOP MIMO-OFDM

This chapter considers a closed-loop MIMO-OFDM system, where the transmitter has CSI. We address the transmit beamforming technique for a MIMO channel by exploiting CSI. If the transmitter uses the optimal beamforming, known as *eigenbeamforming*, to maximize the achievable rate, a MIMO channel is transformed into a bank of scalar channels with no crosstalk from one scalar channel to next. Next, we briefly mention the feasibility of having CSI at the transmitter and propose an filter-reuse scheme for a time-division duplex (TDD) system. Finally, we survey adaptive algorithms for the receive filter of eigenbeamforming.

3.1 Transmit Beamforming

We consider the effective MIMO-OFDM channel in Figure 1:

$$\begin{aligned} \boldsymbol{v}_1 &= \mathbf{H}_1 \boldsymbol{u}_1 + \tilde{\boldsymbol{n}}_1 \\ \vdots \quad &\quad \vdots \qquad . \\ \boldsymbol{v}_N &= \mathbf{H}_N \boldsymbol{u}_N + \tilde{\boldsymbol{n}}_N \end{aligned} \tag{36}$$

For each $\mathbf{H}_n$, we set a transmit filter and a receive filter, as illustrated in Figure 2. Joint optimization of the transmit and receiver filters for MIMO is already a familiar topic in signal processing [45, 49, 51, 67]. In this thesis, we focus on maximizing the achievable rate.

3.1.1 Principal Eigenmode

First, we consider the transmit beamforming with traditional weight vectors [26], where we set $\boldsymbol{v}_n$ as a transmit filter and $\boldsymbol{u}_n^*$ as a receive filter for each $\mathbf{H}_n$ in Figure 2.

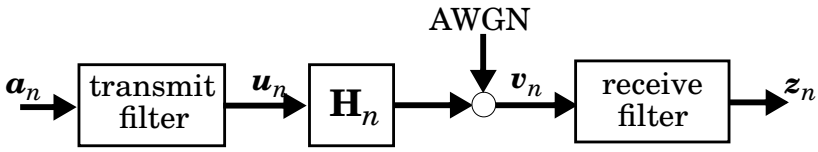

Figure 2: Block diagram for transmit beamforming.

Then, the output of the receive filter is

$$z_n = \boldsymbol{u}_n^* \mathbf{H}_n \boldsymbol{v}_n a_n + \boldsymbol{u}_n^* \tilde{\boldsymbol{n}}_n \quad n = 1, 2, \ldots, N, \tag{37}$$

where a_n is the data signal and $\tilde{\boldsymbol{n}}_n$ is the noise vector. The receive SNR is maximized by choosing $\boldsymbol{v}_n$ and $\boldsymbol{u}_n^*$ as the *principal* (corresponding to the largest eigenvalue) eigenvectors to $\mathbf{H}_n \mathbf{H}_n^*$ and $\mathbf{H}_n^* \mathbf{H}_n$, respectively, under the unit-norm constraint [66]. Then, (37) becomes:

$$z_n = \sqrt{s_n^{(1)}} a_n + w_n \quad n = 1, 2, \ldots, N, \tag{38}$$

where $w_n = \boldsymbol{u}_n^* \tilde{\boldsymbol{n}}_n$ and where $s_n^{(1)}$ is the largest eigenvalue of $\mathbf{H}_n \mathbf{H}_n^*$. For this reason, we call the transmit beamforming scheme in (37) the *principal-eigenmode* transmission.

We will later see that the principal eigenmode achieves a full diversity on a spatially uncorrelated channel. However, since there is only one scalar channel, the spatial multiplexing gain of MIMO disappears and the principal eigenmode can suffer a significant loss in rate. Practically, however, the principal eigenmode is considered to be competitive since its implementation is quite simple. Especially, it is suitable to the outdoor environment, where the largest eigenvalue is often dominantly larger than others [25]. An extreme case is when there is only one path ($L = 0$) in the model of (5). Then, the rank drops to $M = 1$, meaning that the largest eigenvalue is the only nonzero eigenvalue. In such a case, the principal eigenmode is optimal.

3.1.2 Eigenbeamforming

Beamforming weight vectors at both ends create a principal-eigenmode transmission, in which only the largest eigenvalue of $\mathbf{H}_n\mathbf{H}_n^*$ is used for sending data. In this section, we examine the optimal beamforming, where all eigenvalues are used for transmission, so that it maximizes the achievable rate of a MIMO channel.

Instead of weight vectors, we use matrices for transmit and receive filters. For MIMO-OFDM, the optimal (capacity-maximizing) beamforming is based on the SVD of each $\mathbf{H}_n$ [11, 48]. Let $\mathbf{H}_n = \mathbf{U}_n\mathbf{D}_n\mathbf{V}_n^*$ be an SVD of $\mathbf{H}_n$, where $\{s_n^{(m)}\}$ are M nonzero eigenvalues of $\mathbf{H}_n\mathbf{H}_n^*$. Having CSI at the transmitter as well as at the receiver, we set $\mathbf{V}_n$ as an eigenbeamforming filter at the transmitter and $\mathbf{U}_n^*$ as a matched filter at the receiver for each OFDM tone, as shown in Figure 3, which transforms the MIMO-OFDM channel in (8) into a bank of MN memoryless scalar channels over space (m) and frequency (n) [47]:

$$\left.\begin{aligned} z_1^{(1)} &= \sqrt{s_1^{(1)}}a_1^{(1)} + w_1^{(1)} \\ &\vdots \quad \vdots \\ z_1^{(M)} &= \sqrt{s_1^{(M)}}a_1^{(M)} + w_1^{(M)} \end{aligned}\right\} \text{1st tone}$$
$$\vdots \quad \vdots \quad \vdots \qquad , \tag{39}$$
$$\left.\begin{aligned} z_N^{(1)} &= \sqrt{s_N^{(1)}}a_N^{(1)} + w_N^{(1)} \\ &\vdots \quad \vdots \\ z_N^{(M)} &= \sqrt{s_N^{(M)}}a_N^{(M)} + w_N^{(M)} \end{aligned}\right\} N\text{th tone}$$

as illustrated in Figure 3, where $\{a_n^{(m)}\}$ are the input signals to the eigenbeamforming filters at the transmitter and $\{z_n^{(m)}\}$ are the output signals to the matched filters at the receiver. The noise $\{w_n^{(m)}\}$ has the same statistics as the noise vector in (8), that is, $\{w_n^{(m)}\}$ are i.i.d. Gaussian random variables with zero mean and $\mathbb{E}[|w_n^{(m)}|^2] = N_0$, since $\mathbf{V}_n$ and $\mathbf{U}_n$ are unitary. For each tone, eigenbeamforming creates M spatial channels (eigenmodes).

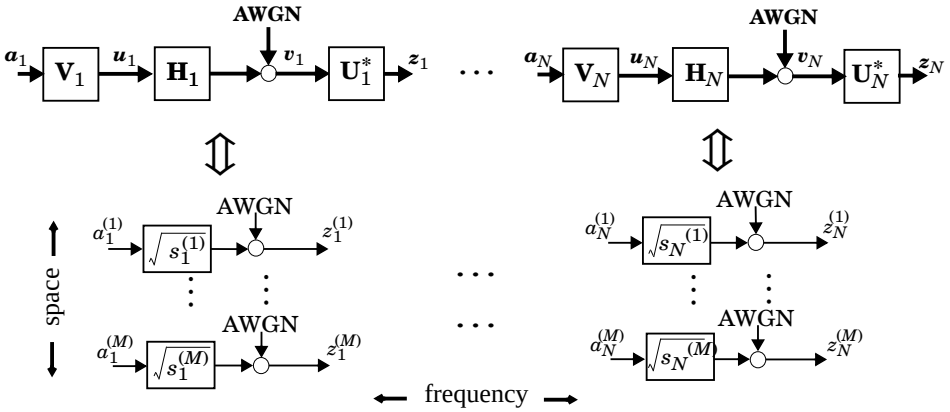

Figure 3: Eigenbeamforming transforms a MIMO-OFDM channel into a bank of scalar channels over space and frequency.

Remark 3.1. In the case of eigenbeamforming, an a-input b-output channel is equivalent with a b-input a-output channel in the sense that both have identical eigenvalues $\{s_n^{(m)}\}$. ♦

Example 3.1. Let us consider the case where $M_T = 1$ but $M_R > 1$ (SIMO). For simplicity, we assume that flat fading ($L = 0$). A flat-fading SIMO channel is represented by:

$$\mathbf{H} = \begin{bmatrix} h_{1,1} \\ \vdots \\ h_{M_R,1} \end{bmatrix}. \tag{40}$$

Let $\mathbf{H} = \mathbf{UDV}^*$ be an SVD of $\mathbf{H}$, where $\mathbf{D} = [||\mathbf{H}||_F, 0, \ldots, 0]^T$; $\mathbf{U} = [\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \ldots, \hat{\boldsymbol{u}}_{M_R}]$ such that $\hat{\boldsymbol{u}}_1 = \mathbf{H}/||\mathbf{H}||_F$ and $\hat{\boldsymbol{u}}_m^* \mathbf{H} = 0$ satisfying $||\hat{\boldsymbol{u}}_m||^2 = 1$ for $m = 2, 3, \ldots, M_R$; and $\mathbf{V} = 1$. Let $\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{n}$ be the received signal. Then

$$z = \mathbf{U}^* \boldsymbol{y} = ||\mathbf{H}||_F a + \mathbf{U}^* \boldsymbol{n}. \tag{41}$$

Thus, eigenbeamforming reduces to the maximum-ratio combining (MRC) in a SIMO channel [53]. ♦

Example 3.2. Suppose that $M_T > 1$ and $M_R = 1$ (MISO). Let $\mathbf{H} = \mathbf{UDV}^*$ be an SVD of the channel matrix $\mathbf{H} = [h_{1,1}, \ldots, h_{1,M_T}]$, where $\mathbf{D} = [||\mathbf{H}||_F, 0, \ldots, 0]$;

$\mathbf{U} = [1]$; $\mathbf{V} = [\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \ldots, \hat{\boldsymbol{v}}_{M_T}]$ such that $\hat{\boldsymbol{v}}_1 = \mathbf{H}^*/||\mathbf{H}||_F$ and $\hat{\boldsymbol{v}}_m^* \mathbf{H}^* = 0$ satisfying $||\hat{\boldsymbol{v}}_m||^2 = 1$ for $m = 2, 3, \ldots, M_T$. Note that $1 \times M$ MIMO is equivalent to $M \times 1$ MIMO after eigenbeamforming since both have the same effective channel $\mathbf{D}$. ♦

The advantages of eigenbeamforming include

- There is no crosstalk between spatial channels (eigenmodes). Thus, conventional channel codes are readily used, and the complexity of decoding only linearly grows with the number of transmit antennas.
- Eigenbeamforming can be applied to any size and any rank of channel matrix since SVD exists for any matrix, that is, eigenbeamforming is readily applicable to spatially correlated channels.
- Finally and most importantly, eigenbeamforming is optimal in the information-theoretical sense since unitary filters preserve information. We will discuss this advantage in depth in the next chapter.

However, there are also drawbacks:

- It requires perfect CSI at the transmitter as well as at the receiver.
- SVD requires a high complexity with an order O^3.

As the answers to the above drawbacks, we will address the availability of CSI at the transmitter in Section 3.2 and develop low-complexity adaptive eigenbeamforming in Section 3.3.

3.2 Availability of CSI at the Transmitter

In this section, we briefly review how a transmitter can obtain CSI. Generally speaking, there are two ways: (1) sending back the estimated CSI from receiver to transmitter via a dedicated feedback; (2) using reciprocity of a time-division duplex (TDD) channel.

3.2.1 Feedback of CSI

An intuitive way to furnish the transmitter with CSI is to feed back the estimated CSI from receiver to transmitter. However, the feedback information is redundant information, which has nothing to do with information data. Moreover, the delay due to the feedback might cause a mismatch problem if the channel changes during feedback. The latter is not a big problem on a slowly varying channel. However, the redundant load of the feedback could be a serious problem, particularly when the number of tones is large in MIMO-OFDM.

The question is whether the feedback of CSI is worthwhile despite the side information. We will in part answer this question by showing that a transmitter with CSI performs significantly better than a transmitter ignorant of CSI in terms of outage capacity in Chapter 4.

In some cases, a partial feedback of CSI would be sufficient. An example is a transmit beamforming strategy using covariance matrix feedback [34]. On the other hand, only the quality of channel is sent back to the transmitter. In this case, eigenbeamforming is not possible, but adaptive transmission using partial CSI can improve performance.

3.2.2 Time-Division Duplex

In a TDD system, we can avoid a feedback of CSI to the transmitter if the reciprocity of wireless channels is exploited [37, 54].

Despite the reciprocity of the propagation channel, there are nontrivial implementation issues that make the problem harder in a system with multiple transmit antennas, namely the effects of transmission and reception electronics [37]. More precisely, the channel responses are cascades of the physical channel and the responses of the transmission and reception electronics. Since the transmitter and receive electronics do not have the same response, they must be estimated using self-calibrating

circuits. This is less of an issue in a single-input system, since estimates of the cascaded channels phase responses are not required. With multiple antennas, accurate phase estimates are required to use techniques such as beamforming. Some important practical issues are discussed in [41].

In this work, we assume that the reciprocity is perfect. Then, the receive filter may be estimated while receiving signals from the other end, and the estimated receive filter may be used as the transmit filter during transmission in the opposite direction. As shown in Figure 4, where we assume flat fading for simplicity, there exist two links: a forward link on the top and a reverse link on the bottom. Let $\mathbf{H}$ be the channel for the forward link and $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^*$ be an SVD of $\mathbf{H}$. In the forward link, we set $\mathbf{V}$ as a transmit beamforming filter and $\mathbf{U}^*$ as a receive beamforming filter, which diagonalizes the channel matrix $\mathbf{H}$:

$$\boldsymbol{z}_F = \mathbf{U}^*\mathbf{H}\mathbf{V}\boldsymbol{a}_F + \mathbf{U}^*\boldsymbol{n}_F = \mathbf{D}\boldsymbol{a}_F + \boldsymbol{w}_F, \tag{42}$$

where the subscript F denotes signals in the forward link. In the reverse link, due to reciprocity, the channel matrix is $\mathbf{H}^T$, and the eigenbeamforming diagonalizes the channel matrix, such that

$$\boldsymbol{z}_R = \mathbf{V}^T\mathbf{H}^T(\mathbf{U}^*)^T\boldsymbol{a}_R + \mathbf{V}^T\boldsymbol{n}_R = \mathbf{D}\boldsymbol{a}_R + \boldsymbol{w}_R, \tag{43}$$

where we use $(\mathbf{U}^*)^T$ as a transmit filter and $\mathbf{V}^T$ as a receive filter.

Note that both transmitters see the same effective channel $\mathbf{D}$ in (42) and (43). If $\mathbf{V}$ is used for a transmit filter, its transpose is used as a receiver filter on the left-handed side in Figure 4. The same is true for the right-handed side, implying that no feedback is necessary.

3.3 Adaptive Eigenbeamforming

An intuitive method to estimate the receive filter $\mathbf{U}$ is to compute the SVD from an estimate of $\mathbf{H}$ directly. However, as long as we know $\mathbf{U}$ and $\mathbf{D}$, the explicit knowledge

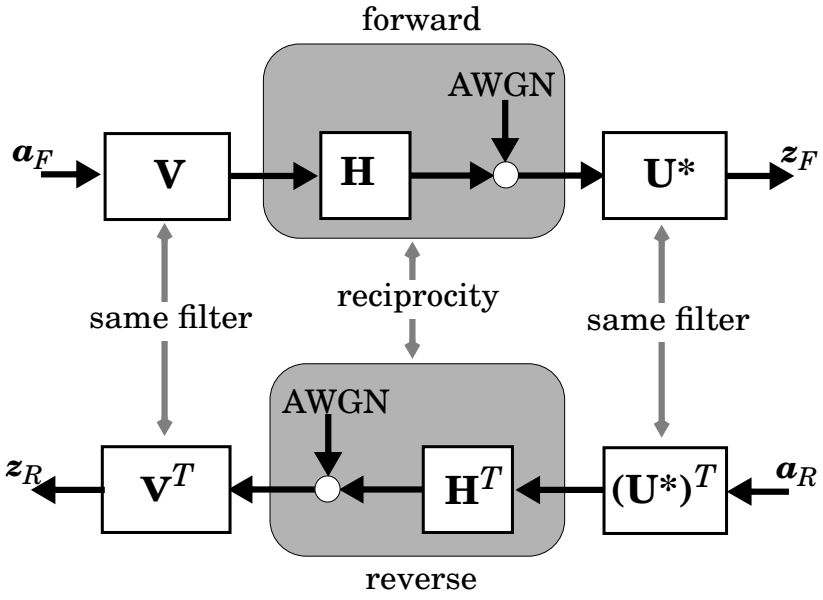

Figure 4: Filter reuse scheme by exploiting reciprocity of TDD.

of $\mathbf{H}$ is not necessary. In this section, we investigate adaptive methods for estimating $\mathbf{U}$ and $\mathbf{D}$ at the receiver. We set either $\mathbf{V}$ (correct eigenbeamforming filter), or $\mathbf{I}_{M_T}$ (arbitrary unitary filter) as the transmit filter and estimate $\mathbf{U}$ either blindly or with a pilot sequence. We consider three different ways.

First, we use multidimensional phase-locked loop (MPLL) of [15] to solve the problem of estimating an unitary filter $\mathbf{U}$. Following [15], we define a partial rotation from $\boldsymbol{x}$ to $\boldsymbol{y}$ as

$$\mathcal{R}^\lambda(\boldsymbol{x} \to \boldsymbol{y}) = \mathbf{I} + \begin{bmatrix} \boldsymbol{u}, \boldsymbol{v} \end{bmatrix} \begin{bmatrix} p-1 & \frac{-p}{|p|}\sqrt{1-|p|^2} \\ \sqrt{1-|p|^2} & |p|-1 \end{bmatrix} \begin{bmatrix} \boldsymbol{u}^* \\ \boldsymbol{v}* \end{bmatrix}, \tag{44}$$

where p is the normalized inner product, $p = \frac{\boldsymbol{x}^*\boldsymbol{z}}{||\boldsymbol{x}||||\boldsymbol{z}||}$, with $\boldsymbol{z} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{x}$, and where $\{\boldsymbol{u}, \boldsymbol{v}\}$ is a basis for the two-dimensional subspace spanned by $\boldsymbol{x}$ and $\boldsymbol{y}$: $\boldsymbol{u} = \frac{\boldsymbol{x}}{||\boldsymbol{x}||}$ and $\boldsymbol{v} = \frac{\boldsymbol{z}/||\boldsymbol{z}|| - p\boldsymbol{u}}{\sqrt{1-|p|^2}}$.[1] Roughly stated, $\mathcal{R}$ is a unitary matrix that rotates all of the way from $\frac{\boldsymbol{x}}{||\boldsymbol{x}||}$ to $\frac{\boldsymbol{y}}{||\boldsymbol{y}||}$, so that $\mathcal{R}^\lambda$ rotates only a fraction λ of the way, where $0 < \lambda < 1$. More precisely, $\mathcal{R}^\lambda$ is a unitary matrix satisfying $\mathcal{R}^\lambda \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = \frac{\boldsymbol{z}}{||\boldsymbol{z}||}$, where $\boldsymbol{z}$

[1] For the case where $\boldsymbol{x}$ and $\boldsymbol{y}$ are collinear ($|p| = 1$), we take $\boldsymbol{v} = \mathbf{0}$ in (44).

is an intermediate vector lying between $\boldsymbol{x}$ and $\boldsymbol{y}$.

For a blind scheme, the choice of transmit filter does not matter as long as it is unitary. Let $\hat{\mathbf{U}}_k$ be an estimate for $\mathbf{U}$ at the kth signaling interval. The update for $\hat{\mathbf{U}}_k$ is given by

$$\hat{\mathbf{U}}_k = \hat{\mathbf{U}}_{k-1}\mathcal{R}^{\lambda}(\mathbf{G}\boldsymbol{z} \rightarrow \boldsymbol{z}), \tag{45}$$

where $\mathbf{G} = \text{diag}[g_1, g_2, \ldots, g_K]$ is a diagonal matrix with $g_1 > g_2 > ... > g_K > 0$ [16]. We emphasize that because the left factor $\mathbf{U}$ of an SVD is not unique, (45) does not converge to a unique solution, but rather only to a $\mathbf{U}$ corresponding to one of many possible singular-value decompositions. This uniqueness problem is typical in blind algorithms, and the phase of diagonal elements of $\hat{\mathbf{U}}^*\mathbf{HV}$, where $\mathbf{V}$ is a valid right factor of an SVD, cannot be completely resolved. For this reason, training is necessary to resolve the rotation [16], or differential coding is used [21].

MPLL can also be used if the correct eigenbeamforming filter $\mathbf{V}$ is used. In such a case, the update is given by

$$\hat{\mathbf{U}}_k = \hat{\mathbf{U}}_{k-1}\mathcal{R}^{\lambda}(\mathbf{D}\boldsymbol{a} \rightarrow \boldsymbol{z}), \tag{46}$$

where $\boldsymbol{a}$ is the transmitted signal vector known at the receiver. At high SNR, the estimation is almost perfect, but performance degrades as SNR decreases.

We can use traditional equalization methods to obtain $\mathbf{U}$ if $\mathbf{V}$ is used for the transmit filter. The receive filter can be estimated by a stochastic gradient algorithm [6], such that

$$\mathbf{C}^* = \mathbf{C}^* - \mu(\boldsymbol{z} - \boldsymbol{a})\boldsymbol{y}^*, \tag{47}$$

where μ is a step size for the update, if the receiver knows the transmitted signal vector $\boldsymbol{a}$, where we use $\mathbf{C} = [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{M_R}]$ instead of $\mathbf{U}$ to emphasize that $\mathbf{C}$ is not necessarily unitary. If the noise is absent, $\mathbf{C}$ converges to

$$\boldsymbol{c}_m = \frac{\boldsymbol{u}_m}{\sqrt{s^{(m)}}}. \tag{48}$$

The stochastic gradient algorithm works well for the detection of $\boldsymbol{a}$. However, in a TDD system, where we need to know an unitary matrix $\mathbf{U}$, which will be used a transmit filter in the opposite-direction transmission, a problem arises since there is no constraint on the unitary property of the estimated filter in the update of (47) and a slight error in $\boldsymbol{c}_m$ leads to a large error in estimated $\mathbf{U}$.

Finally, we introduce a method to estimate $\mathbf{U}$ from the eigenvectors of the covariance matrix of the received signal vector $\boldsymbol{v}$:

$$\mathbf{K}_{\boldsymbol{vv}} = \mathbb{E}[\boldsymbol{v}\boldsymbol{v}^*]. \tag{49}$$

If $\mathbb{E}[\boldsymbol{a}\boldsymbol{a}^*] = E_a\mathbf{I}_{M_T}$, (49) reduces to $\mathbf{K} = E_a\mathbf{U}\mathbf{D}^2\mathbf{U}^* + N_0\mathbf{I}_{M_R}$, from which we can estimate $\mathbf{U}$. Another case is when we use a valid right factor of an SVD as the transmit filter, $\mathbf{V}$, then $\mathbf{K} = \mathbf{U}\hat{\mathbf{D}}\mathbf{U}^* + N_0\mathbf{I}_{M_R}$, where $\hat{\mathbf{D}} = \mathrm{diag}[s^{(1)}e^{(1)}, \ldots, s^{(M)}e^{(M)}]$. Adaptive algorithms to estimate the eigenvectors form the covariance matrix has been studied for a long time, such as [36]. Both blind or data-aided update is possible in this case.

Recently, new algorithms for estimating transmit and receiver filters have been proposed based on power method [21] and subspace estimation [22], both for a TDD system. The algorithm based on power method can estimate the transmit and receive filters as well as eigenvalues blindly. The subspace method basically uses the covariance matrix in (49) and calculate the receive filter from the eigenvectors of the covariance matrix. Both algorithms work when only K ($\leq M$) eigenmodes are used for transmission.

CHAPTER 4

INFORMATION THEORY FOR MIMO

Information theory for MIMO fading channels has drawn considerable attention. As a fading channel is a random variable, so is its mutual information between transmitter and receiver. Thus, we need to map the random mutual information to a meaningful deterministic value, such as *average* or *outage* capacity, depending on the channel generation process [60].

In this chapter, we review the average and outage capacity of MIMO-OFDM in Figure 1 assuming that CSI is known at the receiver. We consider both the cases where perfect CSI is known and unknown at the transmitter. We have two goals in this chapter: (1) to review increased capacity (spatial multiplexing) and improved outage performance (diversity) in MIMO; (2) to show that the transmitter CSI considerably improves the outage performance. We also present a contribution on *outage-region capacity* at the end of this chapter. First, we begin by defining average and outage capacity.

4.1 Definition of Average and Outage Capacity

In a wireless communication system, a transmitter is usually constrained in its power. We consider the MIMO-OFDM model in Figure 1, where $\boldsymbol{v}_n = \mathbf{H}_n \boldsymbol{u}_n + \tilde{\boldsymbol{n}}_n$ for $n = 1, 2, \ldots, N$. If $\hat{\mathbf{H}} = \mathrm{diag}[\mathbf{H}_1, \ldots, \mathbf{H}_N]$, let $\mathbf{Q}_n = \mathbb{E}\left[\boldsymbol{u}_n \boldsymbol{u}_n^* \middle| \hat{\mathbf{H}}\right]$ be the covariance matrix of the input signal vectors $\boldsymbol{u}_n$ to the nth memoryless channel $\mathbf{H}_n$. We consider two energy constraints: the total energy must satisfy either a *long-term* (average) constraint

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\mathrm{tr}(\mathbf{Q}_n)\right] = \bar{E}, \tag{50}$$

where the expectation is over $\hat{\mathbf{H}}$, or a *short-term* (instantaneous) constraint:

$$\frac{1}{N}\sum_{n=1}^{N}\mathrm{tr}(\mathbf{Q}_n) = \bar{E}. \tag{51}$$

We note that the short-term constraint is a special case of the long-term constraint. For both cases, the SNR per receive antenna is defined as

$$\rho = \bar{E}/N_0. \tag{52}$$

With the long-term constraint, the transmitter can control power according to the current channel status, known as *power control* for the single-user communications [13].

In fading environment, $\hat{\mathbf{H}}$ is a random matrix. Thus, the maximum mutual information[1] for a particular channel realization $\hat{\mathbf{H}}$:

$$I(\hat{\mathbf{H}}) = \frac{1}{N}\sum_{n=1}^{N}\log_2\det\left(\mathbf{I}_{M_R} + \frac{\mathbf{H}_n\mathbf{Q}_n\mathbf{H}_n^*}{N_0}\right) \tag{53}$$

is also a random variable. The units of $I(\hat{\mathbf{H}})$ are bits per signaling interval, which reduce to bits/sec/Hz when the rate loss due to the cyclic prefix in (7) is negligible. To characterize the information-theoretical aspects of (53), it is necessary to transform the mutual information into a non-random quantity, such as ensemble average or cumulative distribution function (CDF).

Average capacity is obtained by taking expectation of $I(\hat{\mathbf{H}})$, as defined below.

Definition 4.1 (Average Capacity). The average capacity of MIMO-OFDM is

$$C = \sup_{\{\mathbf{Q}_n\}}\mathbb{E}[I(\hat{\mathbf{H}})], \tag{54}$$

where the supremum is over all $\{\mathbf{Q}_n\}$ satisfying the energy constraint. ♦

If the process that generates the channel is ergodic [39], such that the time average of $I(\hat{\mathbf{H}})$ converges to the ensemble average $\mathbb{E}[I(\hat{\mathbf{H}})]$ as the window for the time average becomes large, there exists a channel code that is able to achieve (54) [60].

[1] Maximized for Gaussian noise by Gaussian distributed transmitted signals [20].

However, the ergodic assumption is not necessarily satisfied, for example, when the channel is chosen randomly at the beginning of transmission and remains fixed for all channel uses [60]. In this case, the average capacity in (54) is not the achievable rate any longer, but we use the outage probability. Let R be the transmission rate in bits per signaling interval. Then, we declare an *outage* occurs when $I(\hat{\mathbf{H}})$ of (53) is smaller than R. In other words, the transmission rate exceeds its limit, and no code can achieve arbitrarily small error probability. We measure how often an outage occurs by the probability of outage.

Definition 4.2 (Outage Probability). Let R be the fixed transmission rate. We define the outage probability:

$$P_{\text{OUT}} = F_{I(\hat{\mathbf{H}})}(R) = \text{Prob}[I(\hat{\mathbf{H}}) < R], \tag{55}$$

where $F_{I(\hat{\mathbf{H}})}(x) = \text{Prob}[I(\hat{\mathbf{H}}) < x]$ denotes the CDF of $I(\hat{\mathbf{H}})$.[2] ♦

In some cases, it is impossible to make $P_{\text{OUT}} = 0$ for any nonzero R, and there is a tradeoff between outage probability and supportable rate.

Definition 4.3 (Outage Capacity). We define ϵ-*achievable rate* [7] as

$$C_\epsilon = \sup_{\{\mathbf{Q}_n\}} \sup\{R : \text{Prob}[I(\hat{\mathbf{H}}) < R] < \epsilon\}, \tag{56}$$

where the first supremum is over all $\{\mathbf{Q}_n\}$ satisfying the energy constraint. ♦

Conventionally, 1% outage capacity means $C_{\epsilon=10^{-2}}$. Clearly, C_ϵ is reduced as ϵ decreases.

4.1.1 Spatial Multiplexing and Diversity Orders

We define two asymptotic measures for capacity C and outage probability P_{OUT} at high SNR: *spatial multiplexing order* and *diversity order*, which measure the advantages of using multiple antennas on a MIMO channel.

[2] We use the definition $\text{Prob}[X < x]$ for CDF instead of frequently used $\text{Prob}[X \leq x]$. For a continuous CDF, this difference is unimportant.

Definition 4.4 (Spatial Multiplexing Order). When C denotes the capacity, either average or outage, the spatial multiplexing order is defined as:

$$\text{spatial multiplexing order } = \lim_{\rho\to\infty} \frac{C}{\log_2(\rho)}, \tag{57}$$

where ρ is SNR. ⧫

Definition 4.5 (Diversity Order). The diversity order quantifies how sharply the outage probability decays as SNR grows [62], defined as:

$$\text{diversity order } = -\lim_{\rho\to\infty} \frac{\log(P_{\text{OUT}})}{\log(\rho)}, \tag{58}$$

where ρ is SNR.[3] ⧫

The spatial multiplexing order is the asymptotic rate at which C increases with log-scale SNR. Graphically, the spatial multiplexing order measures the asymptotic slope of C versus log-scale SNR. A MIMO channel offers a spatial multiplexing order of as large as M, where M is the rank of $\mathbf{H}_n$ [62].

On the other hand, the diversity order measures the asymptotic slope of outage probability versus SNR on a log-log scale, namely how sharply the outage probability decays as SNR grows. As the outage probability is a lower bound for the probability of error in a practical system, the diversity order is related to the reliability of communications on fading channels. Intuitively, a MIMO channel has $M_T M_R$ links between transmitter and receiver, and thus provides $M_T M_R$ times more protection against the effects of fading than a SISO channel. If $P_{\text{OUT}} = 0$ is possible at a finite SNR, the outage probability curve drops vertically and its diversity order is infinite, as on non-fading channels. We will see that an infinite diversity order can be achieved with CSI at the transmitter when the long-term energy constraint in (50) is used.

[3] An alternative definition for diversity order, $\lim_{\rho\to\infty} \frac{\log(P_e)}{\log(\rho)}$ is also used as a design criterion, where P_e is the pairwise error probability. This alternative definition is used in well-known space-time code design [56, 57, 58, 59]

4.2 CSI Unknown at Transmitter

When the transmitter is ignorant of CSI, it has no choice but to distribute power equally over space, frequency, and channel realizations, such that:

$$\mathbf{Q}_n = \frac{\bar{E}}{M_T}\mathbf{I}_{M_T}, \tag{59}$$

so that it satisfies the short-term constraint:

$$\frac{1}{N}\sum_{n=1}^{N}\mathrm{tr}\{\mathbf{Q}_n\} = \bar{E}. \tag{60}$$

4.2.1 Average Capacity

With $\mathbf{Q}_n = \frac{\bar{E}}{M_T}\mathbf{I}_{M_T}$, it is not hard to show that the mutual information in (53) becomes

$$\begin{aligned} I_{\text{RCSI}} &= \frac{1}{N}\sum_{n=1}^{N}\log_2\det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}_n\mathbf{H}_n^*\right) \\ &= \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(1+\frac{\rho}{M_T}s_n^{(m)}\right), \end{aligned} \tag{61}$$

where we use $\mathbf{H}_n\mathbf{H}_n^* = \mathbf{U}_n\mathbf{S}_n\mathbf{U}_n^*$ from Definition 2.1 with $\mathbf{S}_n = \mathrm{diag}[s_n^{(1)},\ldots,s_n^{(M)}]$. Then, the average capacity, constrained by the transmitter's ignorance of CSI, is

$$\begin{aligned} C_{\text{RCSI}} &= \mathbb{E}\left[I_{\text{RCSI}}\right] \\ &= \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(1+\frac{\rho}{M_T}s_n^{(m)}\right)\right]. \end{aligned} \tag{62}$$

Lemma 4.1. The average capacity in (62) is independent of n in the sense that

$$C_{\text{RCSI}} = \mathbb{E}\left[\sum_{m=1}^{M}\log_2\left(1+\frac{\rho}{M_T}s_n^{(m)}\right)\right], \tag{63}$$

for all n.

Proof. Straightforward from Lemma 2.1. ■

A distinct difference from the single-input case lies in the summation over M spatial channels in (63), which implies the spatial multiplexing gain of a MIMO channel can be at most M times larger, as will be shown in Chapter 5.

4.2.1.1 Spatially Uncorrelated Channel

When the channel is spatially uncorrelated, the rank is $M = \min(M_T, M_R)$ for all n, and we can rewrite (62) as

$$C_{\text{RCSI}} = \mathbb{E}\left[\sum_{m=1}^{M} \log_2\left(1 + \frac{\rho}{M_T} s^{(m)}\right)\right], \tag{64}$$

where the elements of $\mathbf{W}$ are i.i.d. $\mathcal{CN}(0,1)$ and $\{s^{(m)}; m = 1, 2, \ldots, M\}$ are the eigenvalues of $\mathbf{WW}^*$. From (64), C_{RCSI} is independent of L.

Using the PDF of unordered eigenvalues in (27), C_{RCSI} becomes

$$C_{\text{RCSI}} = \int_0^\infty \log_2\left(1 + \frac{\rho x}{M_T}\right) \sum_{k=0}^{M} \frac{k! x^D e^{-x}}{(k+D)!} \left\{L_k^D(x)\right\}^2 dx, \tag{65}$$

where $D = \max(M_T, M_R) - \min(M_T, M_R)$ and $L_k^m(x)$ is the Laguerre polynomial of order k defined in (28).

Theorem 4.1. The average capacity of a spatially uncorrelated MIMO Rayleigh-fading channel can be expressed in a semi-analytic form as

$$C_{\text{RCSI}} = \frac{e^{M_T/\rho}}{\log(2)} \sum_{k=0}^{M-1} \sum_{l=0}^{k} \sum_{i=0}^{2l} \left\{ \frac{(-1)^i (2l)!(D+i)!}{2^{2k-i} l! i! (D+l)!} \right. \tag{66}$$

$$\times \binom{2k-2l}{k-l} \binom{2l+2b-2a}{2l-i} \left. \sum_{j=0}^{D+i} E_{j+1}\left(\frac{M_T}{\rho}\right) \right\}, \tag{67}$$

where

$$E_n(x) = \int_1^\infty e^{-xt} t^n dt. \tag{68}$$

Proof. See [52]. ■

We consider the SIMO and MISO cases.

Corollary 4.1. When $M_T \geq 1$ and $M_R = 1$ (MISO),

$$C_{\text{RCSI}} = \frac{e^{M_T/\rho}}{\log(2)} \sum_{j=0}^{M_T-1} E_{j+1}\left(\frac{M_T}{\rho}\right), \tag{69}$$

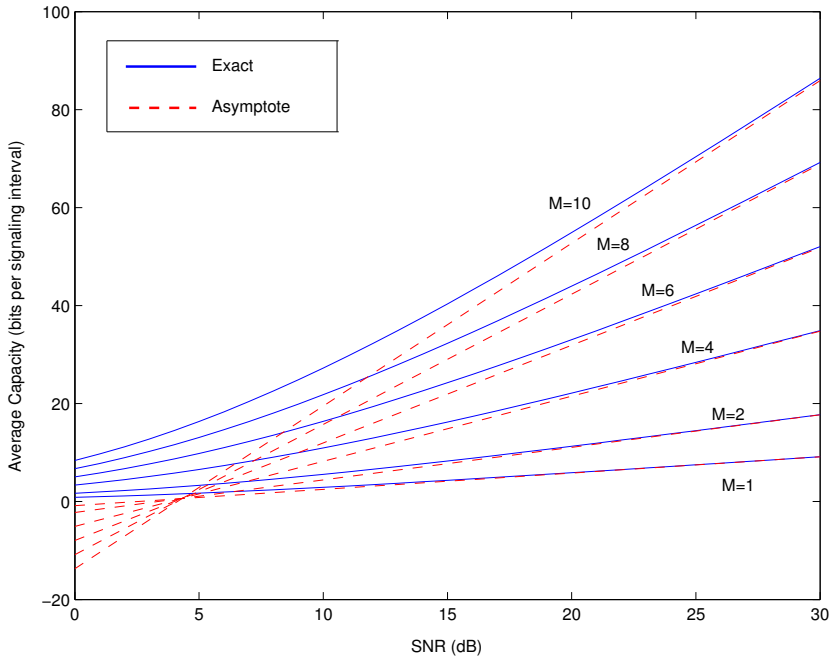

Figure 5: Average capacity of a $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel when $M \in \{1, 2, 4, 6, 8, 10\}$ when CSI is only known at the receiver. Also plotted are the high-SNR asymptotes of the average capacity.

and when $M_T = 1$ and $M_R \geq 1$ (SIMO),

$$C_{\text{RCSI}} = \frac{e^{1/\rho}}{\log(2)} \sum_{j=0}^{M_R-1} E_{j+1}\left(\frac{1}{\rho}\right). \tag{70}$$

Example 4.1. Figure 5 plots the average capacity in (66) of an $M \times M$ Rayleigh-flat-fading channel for $M \in \{1, 2, 4, 6, 8, 10\}$. We can see the increased spatial multiplexing order as the slope of C_{RCSI} in Figure 5 at high SNR increases with M. For emphasis, Figure 6 illustrates that the ratio of $C_{\text{RCSI}}/\log_2(\rho)$, spatial multiplexing order, converges to M for each M. ♦

4.2.1.2 Spatially Correlated Channel

Now we consider the case where the channel is spatially correlated. Following [10], we assume that fading at the transmitter is spatially uncorrelated ($\mathbf{R}_{T_l} = \mathbf{I}_{M_T}$) but fading is correlated at the receiver ($\mathbf{R}_{R_l} \neq \mathbf{I}_{M_T}$). In such a case, the average capacity is given by the following theorem.

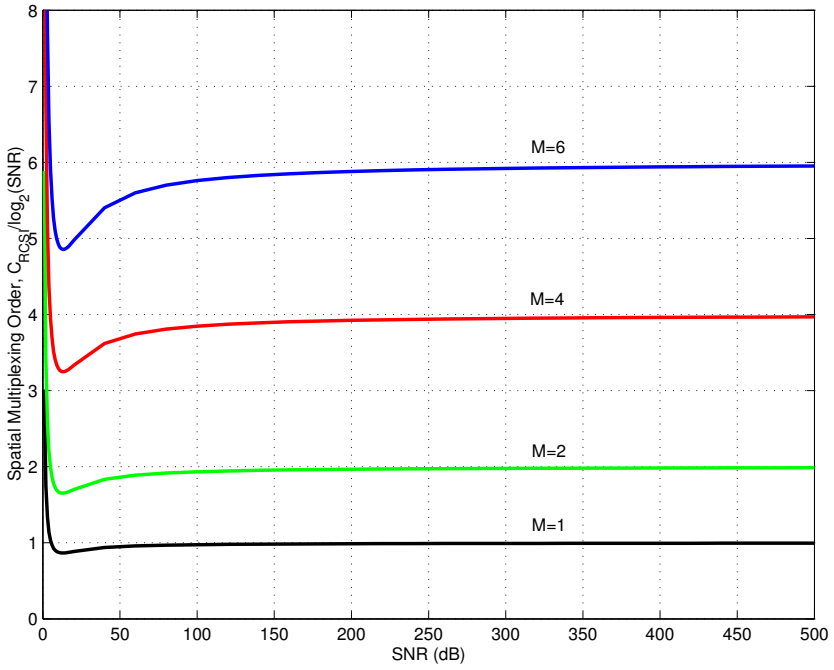

Figure 6: Ratio of $C_{\text{RCSI}}/\log_2(\text{SNR})$ for $M \in \{1, 2, 4, 6\}$ on a $M \times M$ spatially uncorrelated channel.

Theorem 4.2 (Average Capacity).

$$C_{\text{RCSI}} = \mathbb{E}\left[\log_2 \det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{\Lambda}\mathbf{W}\mathbf{W}^*\right)\right], \tag{71}$$

where

$$\mathbf{\Lambda} = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_{M_R}] \tag{72}$$

with λ_m denoting the mth eigenvalue of $\mathbf{R} = \sum_{l=0}^{L} \mathbf{R}_{R_l}$ satisfying $\sum_{m=1}^{M} \lambda_m = M_R$, and where $\mathbf{W}$ is a $M_R \times M_T$ matrix, whose elements are i.i.d. $\mathcal{CN}(0, 1)$.

Proof. See [10]. ■

We remark that C_{RCSI} is independent of the tone index n in (71), which agrees with Lemma 4.1. However, unlike the uncorrelated channel, C_{RCSI} increases with L since the number of nonzero λ_m is dependent on L.

The average capacity C_{RCSI} is maximized when $\{\lambda_m\}$ are equal for all m, namely when fading at the receiver is spatially uncorrelated, $\mathbf{R}_{R_l} = \mathbf{I}_{M_R}$ [10] . Therefore,

Table 1: $\bar{\theta}_l$ used in Example 4.2.

L	$\{\theta_l\}$
0	$\pi/2$
1	$\pi/4$, $3\pi/4$
2	$\pi/4$, $\pi/2$, $3\pi/4$
3	$\pi/4$, $5\pi/12$, $7\pi/12$, $3\pi/4$
4	$\pi/4$, $3\pi/8$, $\pi/2$, $5\pi/8$, $3\pi/4$
5	$\pi/4$, $7\pi/20$, $9\pi/20$, $11\pi/20$, $13\pi/20$, $3\pi/4$

C_{RCSI} of uncorrelated channel (62) is the upper limit of C_{RCSI} with correlation (71), and C_{RCSI} of correlated channel approaches C_{RCSI} of uncorrelated channel as $L \to \infty$. We show the increase of C_{RCSI} of correlated channel with respect to L by the following example.

Example 4.2. We assume that fading at the receiver is spatially correlated with the correlation matrix in (6). For $L \in \{0, 1, 2, 3, 4, 5\}$, Figure 7 and Figure 8 illustrate the average capacity in (71) for $\sigma_\theta = 0.25$ (large spread) and $\sigma_\theta = 0$ (small spread), respectively. The average arrival angles, $\bar{\theta}_l$ is summarized in Table 1. Recall that the spatial multiplexing order is dependent of the rank of $\sum_{l=0}^{L} \mathbf{R}_{R_l}$. In Figure 7 and Figure 8, we notice that C_{RCSI} increases as L grows. Especially, when $\sigma_\theta = 0$, where each $\mathbf{R}_{R_l}$ collapses to a rank-1 matrix, a significant increase can be observed as L grows in Figure 8. It is not surprising that the increase saturates from $L = 3$ since the rank of $\sum_{l=0}^{L} \mathbf{R}_{R_l}$ becomes full ($M = 4$), and the rank is limited to $M = 4$ even for $L > 3$. On the other hand, when $\sigma_\theta = 0.25$, the rank of $\sum_{l=0}^{L} \mathbf{R}_{R_l}$ is already full when $L = 0$. That is why the increase is less significant than when $\sigma_\theta = 0$. But, considering that there is no increase at all when the channel is spatially uncorrelated, there is an increase in C_{RCSI} when $\sigma_\theta = 0.25$ as L grows as shown in Figure 7. ♦

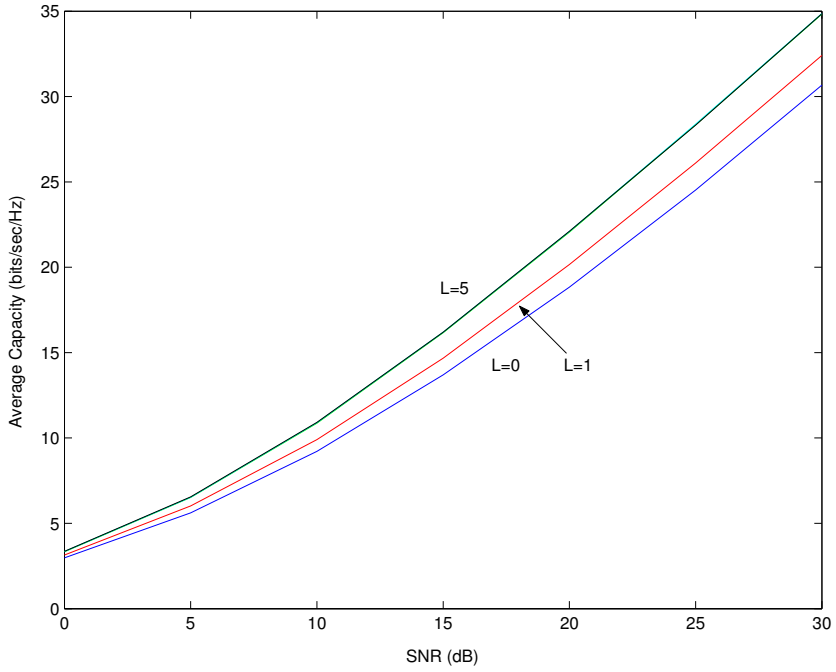

Figure 7: Average capacity with CSI only known at the receiver for $L \in \{0, 1, 2, 3, 4, 5\}$ when the channel is spatially correlated with $\sigma_\theta = 0.25$.

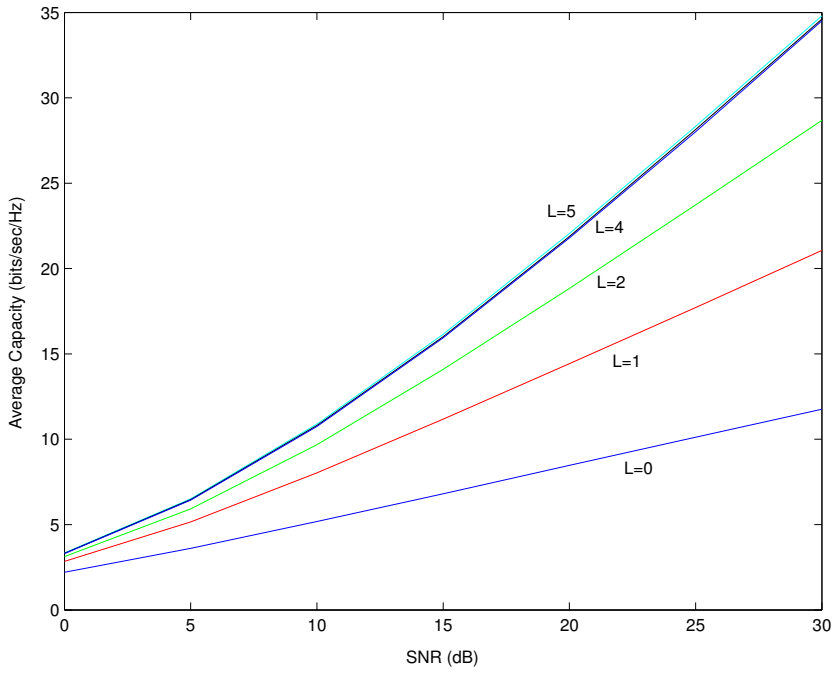

Figure 8: Average capacity with CSI only known at the receiver for $L \in \{0, 1, 2, 3, 4, 5\}$ when the channel is spatially correlated with $\sigma_\theta = 0$.

4.2.2 Outage Performance

We investigate the outage performance of MIMO channels in terms of either the outage probability in (55) or the outage capacity in (56).

From (61), the outage probability when CSI is unknown to the transmitter is

$$\begin{aligned} P_{\text{OUT}} &= \text{Prob}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(1+\frac{\rho}{M_T}s_n^{(m)}\right) < R\right] \\ &= \text{Prob}\left[\frac{1}{N}\sum_{n=1}^{N} I_n < R\right], \end{aligned} \tag{73}$$

where

$$I_n = \sum_{m=1}^{M}\log_2\left(\ 1+\frac{\rho}{M_T}s_n^{(m)}\right) \tag{74}$$

is the mutual information for the nth tone. From Lemma 2.1, I_n is identically distributed for all n, whether fading is spatially correlated or not, and is correlated with each other. Recall that average capacity is independent of n, such that

$$C_{\text{RCSI}} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} I_n\right] = \mathbb{E}\left[I_n\right], \tag{75}$$

since the expectation and summation are commutable. Therefore, the correlation of I_n has no effect on average capacity [10]. For the outage probability in (73), however, the probability and summation cannot commute, implying that the correlation from frequency selectivity affects the outage probability.

In fact, frequency selectivity significantly improves the outage performance in a SISO system [8]. For MIMO, there is a brief discussion in [10], saying that the number of degrees of freedom in the channel is as large as $M_R M_T(L+1)$ when each correlation matrix is full rank (spatially uncorrelated). Thus, the diversity advantage of MIMO can be $M_T M_R(L+1)$ times higher than on flat-fading SISO channels, which will be rigorously proved in Chapter 5. The following example illustrates the well-known antenna diversity at the receiver [53].

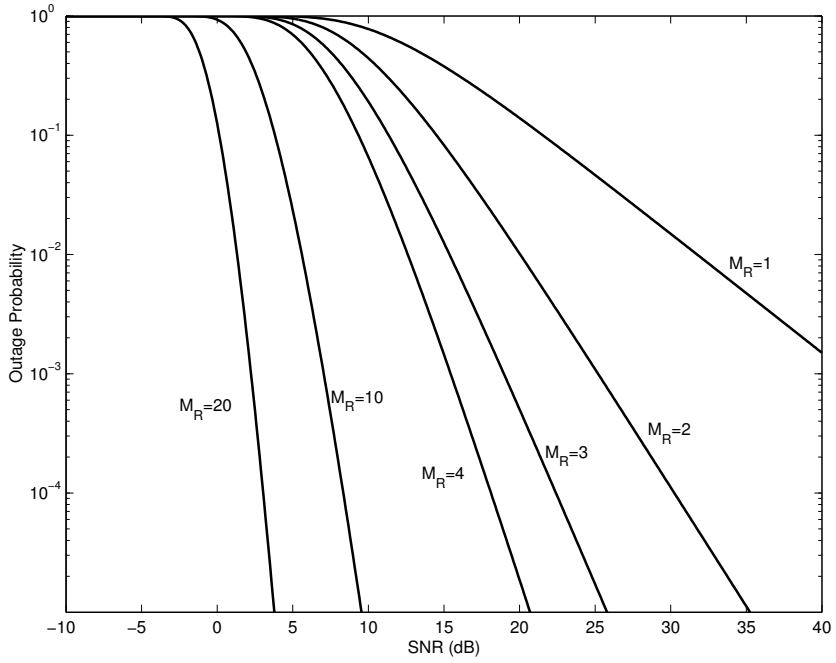

Figure 9: Outage probability on a spatially uncorrelated flat-fading channel with $M_T = 1$ and $M_R \geq 1$ in (76) at $R = 4$ bits per signaling interval.

Example 4.3. For a spatially uncorrelated flat-fading channel with $M_T = 1$ and $M_R \geq 1$ (SIMO),

$$P_{\text{OUT}} = 1 - \frac{\Gamma(M_R, \frac{2^R - 1}{\rho})}{\Gamma(M_R)}, \tag{76}$$

where

$$\Gamma(x, y) = \int_y^\infty t^{x-1} e^{-t} dt \tag{77}$$

is the incomplete complementary Gamma function. Figure 9 plots P_{OUT} in (76) for $M_R \in \{1, 2, 3, 4, 10, 20\}$ at $R = 4$ bits per signaling interval. Clearly, the slope of P_{OUT} becomes more steeper as M_R grows. In other words, the diversity order increases with M_R. With multiple antennas at the receiver, the outage performance remarkably improves. ♦

With multiple antennas employed at the transmitter in a MIMO system, we can obtain transmit diversity gain as well, expecting even better outage performance on a MIMO channel.

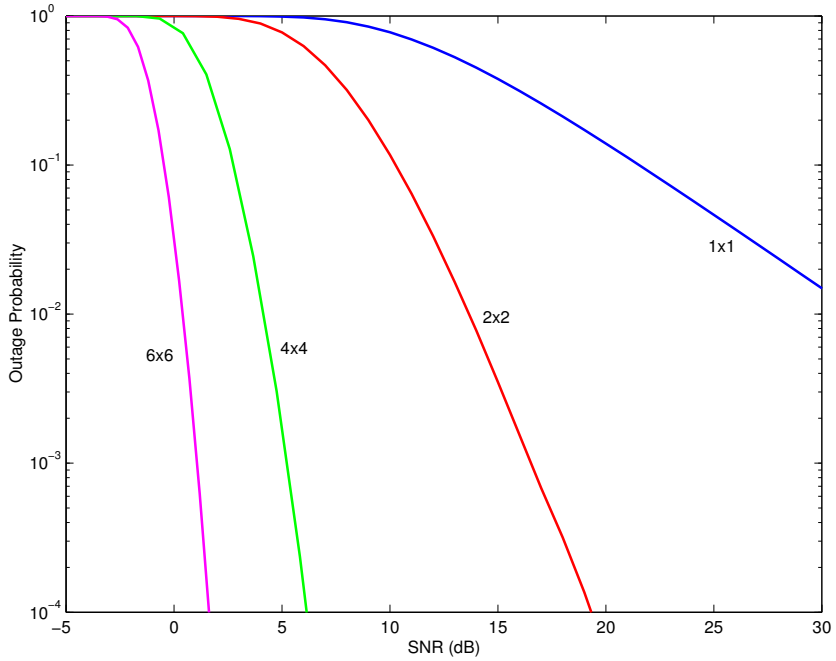

Figure 10: Outage probability on a spatially uncorrelated channel with $M_T = M_R = M$ for $M \in \{1, 2, 4, 6\}$ at $R = 4$ bits per signaling interval.

Example 4.4. Figure 10 plots the outage probability at $R = 4$ bits per signaling interval on an $M \times M$ spatially uncorrelated flat-fading MIMO channel for $M \in \{1, 2, 4, 6\}$. We can see the diversity order increasing with M. The improvement on outage performance is drastic. In terms of SNR at $P_{\text{OUT}} = 10^{-2}$, the advantage of $M = 2$ over $M = 1$ is more than 15 dB. From Figure 10, we confirm that the diversity advantage is incredibly increased on a MIMO channel. ♦

This diversity gain from frequency selectivity is a well-known fact on a SISO channel [8]. We investigate the outage probability of a MIMO-OFDM system in the following example and show that diversity order increases with L in the following example.

Example 4.5. Figure 11 illustrates the outage probability of a 4×4 spatially uncorrelated channel at $R = 10$ bits per signaling interval and $L \in \{0, 1, 2, 3, 4, 5\}$.

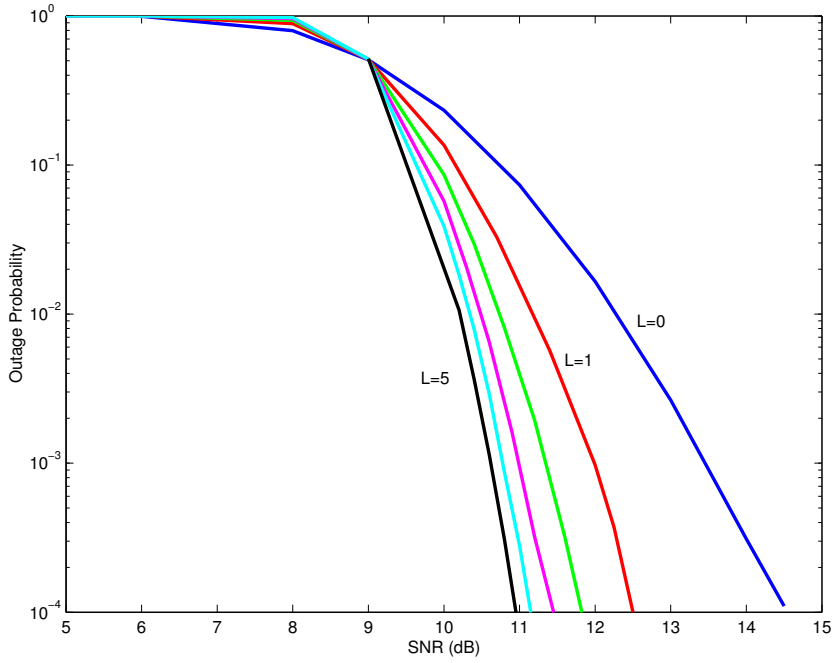

Figure 11: Outage probability of a spatially uncorrelated 4×4 Rayleigh-fading channel with memory $L \in \{0, 1, 2, 3, 4, 5\}$ at $R = 10$ bits per signaling interval.

Clearly, the outage probability drops more sharply as L grows, implying higher diversity order. ♦

From Definition 4.3, when CSI is available only at the receiver, outage capacity C_ϵ can be written as

$$C_\epsilon = F^{-1}_{I_{\text{RCSI}}}(\epsilon), \tag{78}$$

where $I_{\text{RCSI}} = \frac{1}{N}\sum_{n=1}^{N} I_n$ and $F_{I_{\text{RCSI}}}(x)$ is its CDF. Clearly, C_ϵ is affected by the correlation of I_n. A well-known result is that C_ϵ achieves M spatial multiplexing order regardless of $\epsilon > 0$ [62]. However, C_ϵ is zero when $\epsilon = 0$.

Example 4.6. For a spatially uncorrelated Rayleigh-flat-fading channel with $M_T = M_R = 1$ (SISO), the outage capacity is given by:

$$C_\epsilon = \log_2\left(1 + \rho \log((1-\epsilon)^{-1})\right). \tag{79}$$

Figure 12 illustrates C_ϵ for $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-5}\}$. It is not surprising that $C_{\epsilon=10^{-1}}$ is the largest. However, as SNR tends to infinity, the spatial multiplexing orders (or the

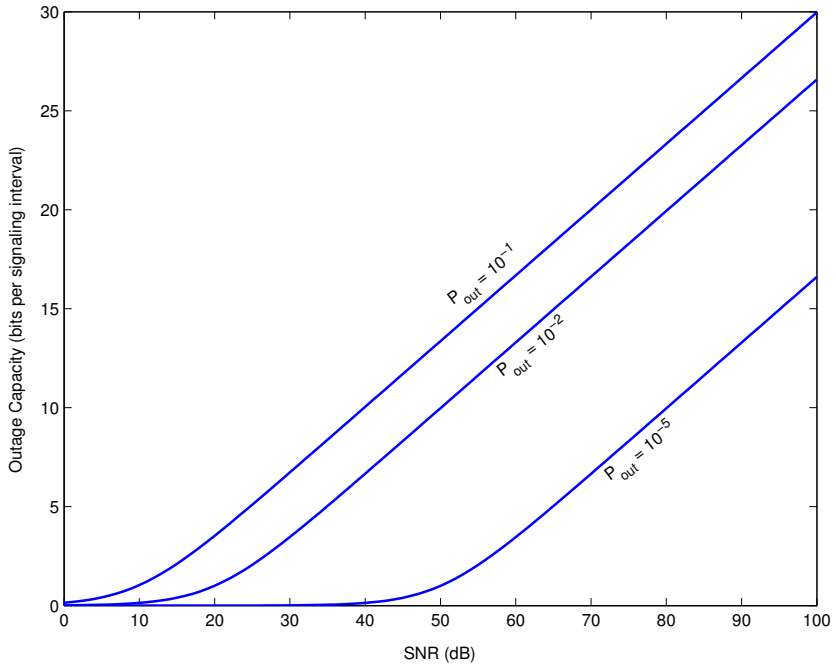

Figure 12: Outage capacity for 1×1 spatially uncorrelated Rayleigh flat fading with $P_{\text{OUT}} \in \{10^{-1}, 10^{-2}, 10^{-5}\}$.

slopes of the curves) are identical, though there is a difference in convergence speed.

♦

For MIMO flat fading, calculating $F_{I_{\text{RCSI}}}(x)$ is difficult, so we resort to Monte-Carlo simulations to obtain C_ϵ.

Example 4.7. Consider an $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel. Figure 13 illustrate the outage capacity C_ϵ for $\epsilon = 0.01$ and $M \in \{1, 2, 4, 6\}$. Clearly, the spatial multiplexing order increases as M grows. The capacity improvement of MIMO, when compared to $M = 1$, is remarkable. For emphasis, the spatial multiplexing order is plotted in Figure 14. As SNR gets larger, the spatial multiplexing order converges to M for each M. ♦

Since C_ϵ is a function of $F_{I_{\text{RCSI}}}(x)$, it also depends on n. On a spatially uncorrelated channel, C_ϵ increases with L, but the spatial multiplexing order remains unchanged, $M = \min(M_T, M_R)$.

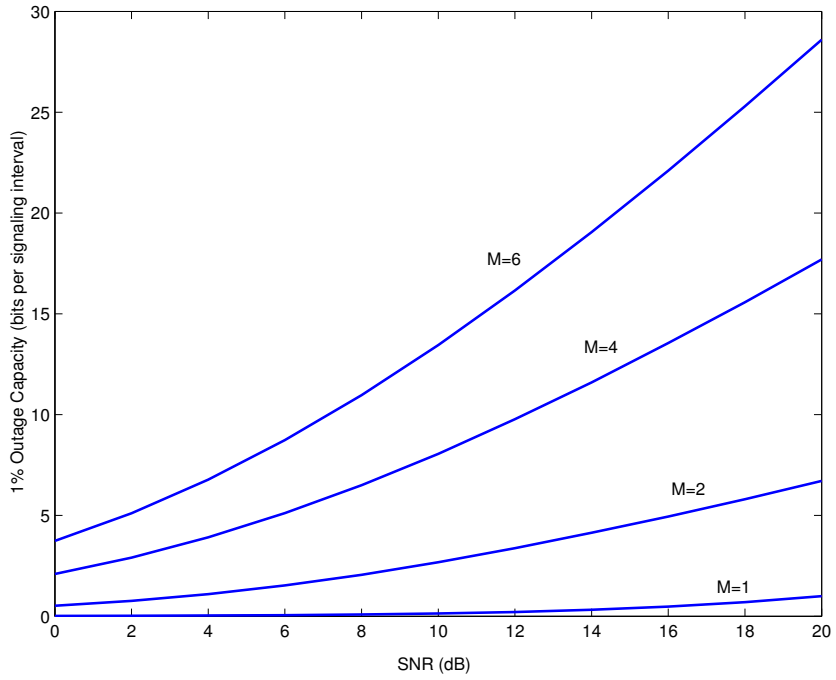

Figure 13: Outage capacity $C_{\epsilon=0.01}$ on an $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel with $M \in \{1, 2, 4, 6\}$.

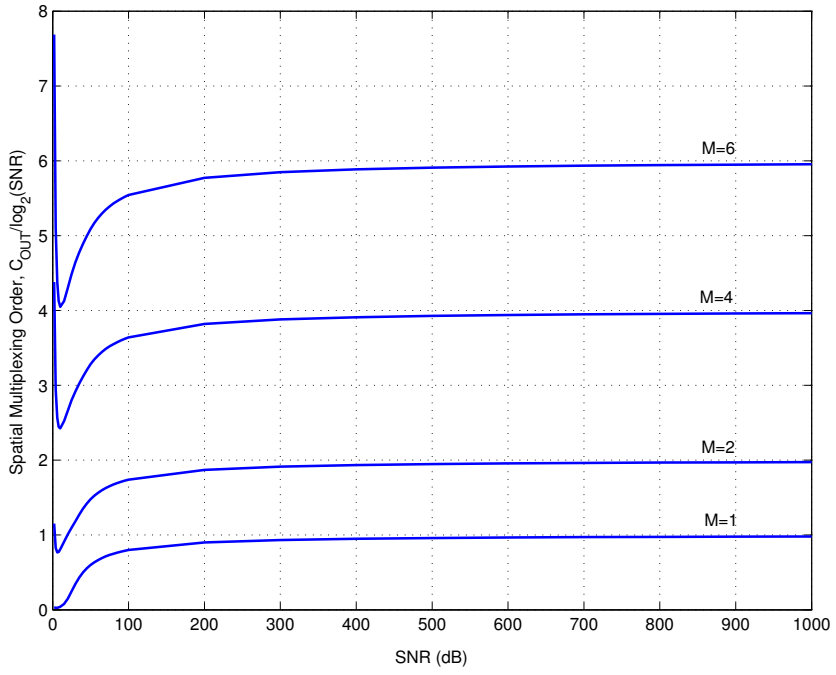

Figure 14: Spatial multiplexing order on an $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel with $M \in \{1, 2, 4, 6\}$.

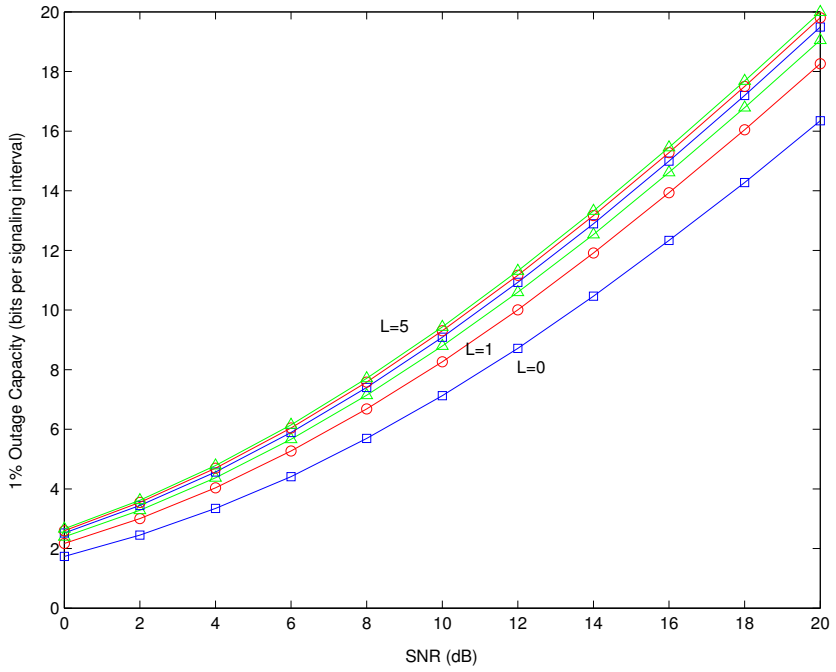

Figure 15: 1% outage capacity on a 4×4 spatially uncorrelated Rayleigh-fading channel with memory $L \in \{0, 1, 2, 3, 4, 5\}$.

Example 4.8. We consider a 4×4 spatially uncorrelated Rayleigh-fading channel with memory L and $N = 128$. In Figure 15, we plot the outage capacity at $P_{\text{OUT}} = 0.01$ for $L \in \{0, 1, 2, 3, 4, 5\}$. Unlike the average capacity, which is independent of L, there is a distinct increase as L grows form 0 to 5. To further understand this behavior, Figure 16 shows the PDF of I at $\rho = 8$ dB as L ranges from 0 to 5. The PDF curves are centered at the same point, meaning that the average capacity is identical regardless of L. However, as L grows, the PDF curves have narrower widths, meaning that the outage capacity increases with L. ♦

4.3 CSI Known at Transmitter

The previous section reviewed the results on average and outage capacity assuming that CSI is available only at the receiver. In this section, we assume a closed-loop system in Chapter 3, where CSI is also known to the transmitter, and examine the

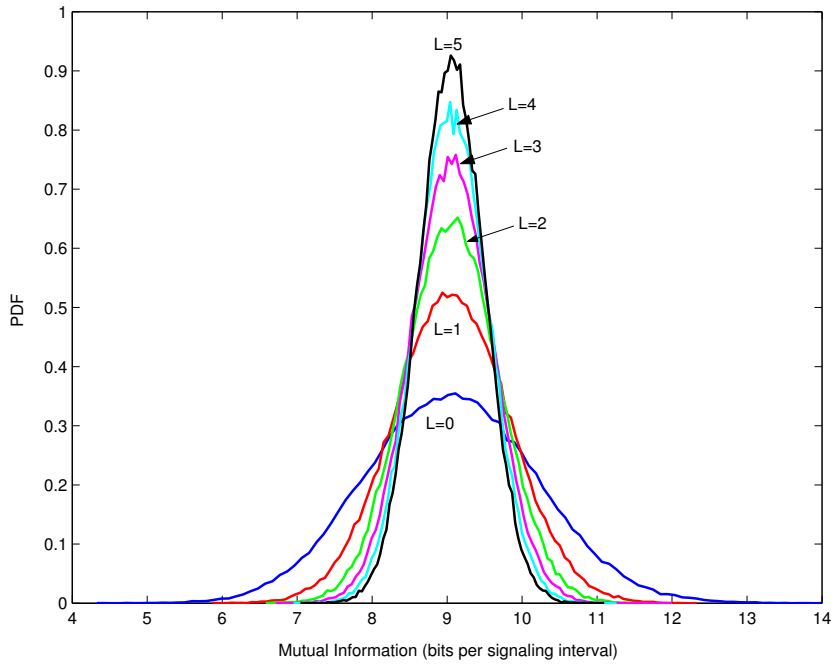

Figure 16: PDF of mutual information of a 4×4 spatially uncorrelated Rayleigh-fading channel with $L \in \{0, 1, 2, 3, 4, 5\}$ at $\rho = 8$ dB.

information-theoretical results. From Chapter 3, we know that eigenbeamforming is optimal in the sense that there is no information loss. We consider the parallel channels in Figure 3:

$$\left.\begin{aligned} z_1^{(1)} &= \sqrt{s_1^{(1)}}a_1^{(1)} + w_1^{(1)} \\ &\vdots \quad \vdots \\ z_1^{(M)} &= \sqrt{s_1^{(M)}}a_1^{(M)} + w_1^{(M)} \end{aligned}\right\} \text{1st tone}$$
$$\vdots \quad \vdots \quad \vdots \qquad , \tag{80}$$
$$\left.\begin{aligned} z_N^{(1)} &= \sqrt{s_N^{(1)}}a_N^{(1)} + w_N^{(1)} \\ &\vdots \quad \vdots \\ z_N^{(M)} &= \sqrt{s_N^{(M)}}a_N^{(M)} + w_N^{(M)} \end{aligned}\right\} N\text{th tone}$$

First, we address the classical problem of allocating power to parallel channels. Then, we examine average and outage capacity mainly with the long-term energy constraint in (50). We also provide the results with the short-term energy constraint in (51) for comparison.

4.3.1 Power Allocation Problem

For the parallel channels in (80), we first consider the mutual information. For the nth tone, let $\mathbf{H}_n = \mathbf{U}_n\mathbf{D}_n\mathbf{V}_n^*$ be an SVD of $\mathbf{H}_n$. As described in Figure 3, the input vector to each $\mathbf{H}_n$ is $\boldsymbol{u}_n = \mathbf{V}_n\boldsymbol{a}_n$, where $\boldsymbol{a}_n = [a_n^{(1)}, \ldots, a_n^{(m)}]^T$. Given $\hat{\mathbf{H}} = \mathrm{diag}[\mathbf{H}_1, \ldots, \mathbf{H}_N]$, the covariance matrix of $\boldsymbol{x}_n$ can be written as

$$\mathbf{Q}_n = \mathbb{E}\left[\boldsymbol{u}_n\boldsymbol{u}_n^* \middle| \hat{\mathbf{H}}\right] = \mathbf{V}_n\mathbf{E}_n\mathbf{V}_n^*, \tag{81}$$

where $\mathbf{E}_n = \mathrm{diag}\{e_n^{(1)}, \ldots, e_n^{(M)}\} = \mathbb{E}\left[\boldsymbol{a}_n\boldsymbol{a}_n^* \middle| \hat{\mathbf{H}}\right]$. Then, the mutual information of (53) becomes

$$\begin{aligned} I_{\text{TRCSI}} &= \frac{1}{N}\sum_{n=1}^{N} \log_2 \det\left(\mathbf{I}_{M_R} + \frac{\mathbf{H}_n\mathbf{Q}_n\mathbf{H}_n^*}{N_0}\right) \\ &= \frac{1}{N}\sum_{n=1}^{N} \log_2 \det\left(\mathbf{U}_n\left(\mathbf{I}_{M_R} + \frac{\mathbf{D}_n\mathbf{E}_n\mathbf{D}_n}{N_0}\right)\mathbf{U}_n^*\right) \\ &= \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left(1 + \frac{s_n^{(m)}e_n^{(m)}}{N_0}\right), \end{aligned} \tag{82}$$

which is a function of $\{e_n^{(m)}\}$ as well as $\{s_n^{(m)}\}$.

Unlike equal distribution of energy when CSI is unknown, the transmitter with CSI has an additional task, *power allocation*. With either short-term or long-term energy constraint, the transmitter must decide $\{e_n^{(m)}\}$ in order to achieve a certain goal, such as maximizing the average capacity $\mathbb{E}[I_{\text{TRCSI}}]$ or minimizing the outage probability $\mathrm{Prob}[I_{\text{TRCSI}} < R]$. Power allocation is a common problem in a closed-loop system. The solution for the optimal allocation is in the form of the water-filling procedure [20].

4.3.2 Average Capacity

We consider the average capacity with the long-term constraint in (50). The power-allocation problem for average capacity is described by

$$C_{\text{TRCSI}} = \sup_{\{e_n^{(m)}\}} \mathbb{E}\left[I_{\text{TRCSI}}\right], \tag{83}$$

where the supremum is over all $\{e_n^{(m)}\}$ satisfying the long-term constraint. Note that the optimization is not only over MN scalar channels but also over all realizations of $\{s_n^{(m)}\}$ because of the expectation in (83). Goldsmith *et al.* [28] presents a solution for the problem in (83) in the case of SISO. The solution can be generalized to the MIMO case [7], as follows:

Theorem 4.3. The power allocation $\{e_n^{(m)}\}$ that achieves C_{TRCSI} is

$$e_n^{(m)} = N_0 \left\{ \lambda - \frac{1}{s_n^{(m)}} \right\}^+, \tag{84}$$

where $\{x\}^+ = \max(x, 0)$ and λ ensures that the long-term energy constraint in (50) is satisfied. Then, the resulting average capacity is

$$C_{\text{TRCSI}} = \mathbb{E}\left[\frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \left\{ \log_2 \left(\lambda s_n^{(m)} \right) \right\}^+ \right], \tag{85}$$

at an SNR of

$$\rho_{\text{TRCSI}} = \mathbb{E}\left[\frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \left\{ \lambda - \frac{1}{s_n^{(m)}} \right\}^+ \right]. \tag{86}$$

Proof. See [28]. ■

Remark 4.1. The average capacity C_{TRCSI} is independent of n in the sense that (85) and (86) are equivalent to

$$C_{\text{TRCSI}} = \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \log_2 \left(\lambda s_n^{(m)} \right) \right\}^+ \right], \tag{87}$$

and

$$\rho_{\text{TRCSI}} = \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \lambda - \frac{1}{s_n^{(m)}} \right\}^+ \right], \tag{88}$$

for any n since $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ are identically distributed for all n from Lemma 2.1. If the channel is spatially uncorrelated, C_{TRCSI} is independent of L. ♦

In Theorem 4.3, the optimal power allocation is linked with SNR via the water-level parameter λ. Solving (85) or (86) to obtain λ is quite difficult due to the expectation, which requires multiple integrals. For the case of single transmit antenna,

it can be represented as a semi-analytic form, but no analytic solution is known for MIMO.

Example 4.9 (From [2]). When $M_T = 1$ and $M_R \geq 1$ (SIMO) on a spatially uncorrelated Rayleigh fading channel,[4]

$$C_{\text{TRCSI}} = \frac{J_{M_R}(1/\lambda)}{\Gamma(M_R)\lambda^{M_R}\log(2)}, \tag{89}$$

where

$$\begin{aligned} J_n(x) &= \int_1^\infty t^{n-1}\log(t)e^{-xt}dt & (90)\\ &= \frac{(n-1)!}{x^n}\sum_{k=0}^{n-1}\frac{\Gamma(k,x)}{k!}, & (91)\end{aligned}$$

where $\Gamma(k,x)$ is the incomplete complementary Gamma function defined in (77). The SNR is:

$$\rho = \frac{\lambda}{\Gamma(M_R)}\Gamma\left(M_R, \frac{1}{\lambda}\right) - \frac{1}{\Gamma(M_R)}\Gamma\left(M_R - 1, \frac{1}{\lambda}\right). \tag{92}$$

If $M_R = 1$ (SISO),

$$C_{\text{TRCSI}} = \frac{E_1(1/\lambda)}{\log(2)} \tag{93}$$

and

$$\rho = \lambda\exp(-1/\lambda) - E_1(1/\lambda), \tag{94}$$

where $E_1(x)$ is defined in (68). ♦

We emphasize that C_{TRCSI} is the average capacity with the long-term constraint. If the short-term constraint of (51) is used, the average capacity at an SNR of ρ is [18]

$$C_{\text{TRCSI,ST}} = \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left\{\log_2(\lambda s_n^{(m)})\right\}^+, \tag{95}$$

where λ satisfies

$$\rho = \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left\{\lambda - \frac{1}{s_n^{(m)}}\right\}^+. \tag{96}$$

[4]Also true for $M_T \geq 1$ and $M_R = 1$ (MISO).

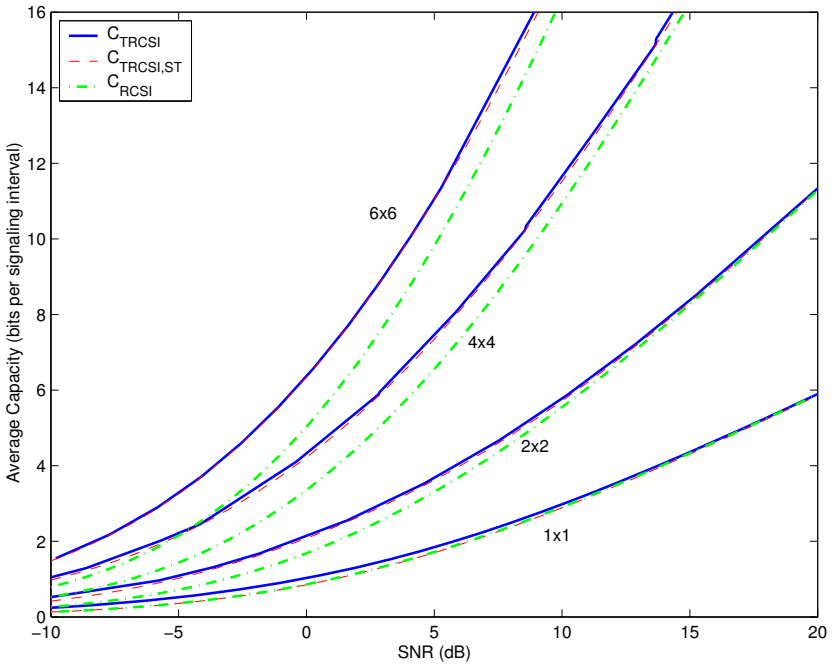

Figure 17: Comparison between C_{TRCSI} and C_{RCSI} on a $M \times M$ spatially uncorrelated Rayleigh fading with $M \in \{1, 2, 4, 6\}$.

Clearly, we have $C_{\text{TRCSI}} \geq C_{\text{TRCSI,ST}} \geq C_{\text{RCSI}}$. It can be shown that $C_{\text{TRCSI,ST}}$ is 4 times larger than C_{RCSI} as $\rho \to 0$, but the advantage vanishes as SNR grows [18].

However, the difference between C_{TRCSI} and C_{RCSI} is not dramatically large, as discussed for SISO [28]. In fact, as will be shown in Chapter 5, the difference converges to zero at high SNR if $M_T \geq M_R$. To see how M affects the difference, we consider the following example.

Example 4.10. Figure 17 illustrates C_{TRCSI} and C_{RCSI} for an spatially uncorrelated $M \times M$ channel with $M \in \{1, 2, 4, 6\}$. The capacity gap between C_{TRCSI} and C_{RCSI} increases with M. The gap is more conspicuous at low SNR, and C_{RCSI} seems to converges to C_{TRCSI} at high SNR. We also plot $C_{\text{TRCSI,ST}}$ in (95). From Figure 17, $C_{\text{TRCSI,ST}}$ is almost as large as C_{TRCSI}, especially when M is large, pointing out that water-filling over channel realizations (for C_{TRCSI}) is unnecessary as M grows. ♦

4.3.3 Zero-Outage Capacity

In a closed-loop system with the long-term constraint, the concept of the outage probability is slightly different. With the knowledge of CSI, the transmitter knows whether the channel can support the transmission rate, and it would attempt to reduce the rate or to stop transmission if the current rate is not supportable. Hence, we declare an outage occurrence when the transmitter reduces the rate to zero. Let $\mathcal{R}_{\text{OFF}}$ denote a set of channel realizations $\{s_n^{(m)}\}$ for which the transmitter decides not to transmit. Then, we define the outage probability as

$$P_{\text{OUT}} = \text{Prob}[\{s_n^{(m)}\} \in \mathcal{R}_{\text{OFF}}]. \tag{97}$$

We mean by *zero outage* that $\mathcal{R}_{\text{OUT}}$ is a null set, such that data is transmitted at a nonzero rate for any realization of channel.

We assume that the transmission rate is fixed as R. By using the ϵ-capacity, defined in (56), we define zero-outage capacity.

Definition 4.6 (Zero-Outage Capacity). Given a nonzero rate R, the zero-outage capacity is defined as

$$C_0 = \lim_{\epsilon \to 0} \sup_{\{e_n^{(m)}\}} \sup \{R; P_{\text{OUT}} \leq \epsilon\}, \tag{98}$$

where the first supremum is over all $\{e_n^{(m)}\}$ satisfying the long-term constraint of (50) [7].[5] ♦

The optimization in (98) is also a power-allocation problem, where we find $\{e_n^{(m)}\}$ such that C_0 is maximized. The optimal solution is obtained by finding $\{e_n^{(m)}\}$ that minimize

$$\bar{E} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} e_n^{(m)}\right], \tag{99}$$

[5] If the instantaneous energy constraint is used instead of the average energy constraint, the zero-outage capacity is zero even though the channel is known at the transmitter [7].

subject to

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(1+\frac{s_n^{(m)}e_n^{(m)}}{N_0}\right)=R. \tag{100}$$

Then, $\bar{E}$ is a function of R, that is $\bar{E}=f(R)$. The zero-outage capacity at an SNR of $\rho=\bar{E}/N_0$ is then obtained by inverting this function, $C_{_0}=f^{-1}(\bar{E})$. The optimal solution to the power-allocation problem is based on the water-filling procedure [20].

Theorem 4.4 (Optimal Allocation). The optimal power allocation is given by $e_n^{(m)}=N_0\{\lambda-1/s_n^{(m)}\}^+$, where $\{x\}^+=\max(0,x)$, and where

$$\lambda=\frac{2^{RN/|\mathcal{M}|}}{G_{\mathcal{M}}} \tag{101}$$

ensures that the constraint in (100) is satisfied. The index set $\mathcal{M}$ in (101) identifies the *used* channels according to $\mathcal{M}=\{(m,n);\lambda s_n^{(m)}\geq 1\}$. With this optimal allocation, the minimum average energy required for $P_{\text{OUT}}=0$ is:

$$\rho_{_0}=\mathbb{E}\left[\frac{1}{N}\sum_{(m,n)\in\mathcal{M}}\left(\lambda-\frac{1}{s_n^{(m)}}\right)\right]. \tag{102}$$

Proof. See [7]. ■

From Theorem 4.4, the zero-outage capacity can be written as

$$C_{_0}=\mathbb{E}\left[\frac{1}{N}\sum_{(m,n)\in\mathcal{M}}\log_2\left(\lambda s_n^{(m)}\right)\right]. \tag{103}$$

When $M_T=1$ on flat fading ($L=0$), only one scalar channel is available, and $C_{_0}$ can be analytically evaluated as the power allocation is trivial.

Example 4.11. When $M_T=1$ on a flat-fading channel, the strategy in Theorem 4.4 is often called *channel inversion* because the transmitter inverts channel by

$$e=N_0\frac{2^R-1}{s^{(1)}}. \tag{104}$$

Then the SNR required for zero outage probability is:

$$\rho=\frac{\mathbb{E}[e]}{N_0}=(2^R-1)\mathbb{E}[1/s^{(1)}], \tag{105}$$

and therefore the zero-outage capacity is:

$$C_0 = \log_2\left(1 + \frac{\rho}{\mathbb{E}[1/s^{(1)}]}\right). \tag{106}$$

When $M_T = M_R = 1$ (SISO) on a spatially uncorrelated Rayleigh-fading channel, $\mathbb{E}[1/s^{(1)}]$ diverges since

$$\mathbb{E}[1/s^{(1)}] = \int_0^\infty \frac{e^{-x}}{x} dx = E_1(0) \to \infty, \tag{107}$$

where e^{-x} is the PDF of $s^{(1)}$.

However, when $M_T = 1$ but $M_R > 1$ (SIMO), there is only one nonzero eigenvalue ($s^{(1)}$) and its density function is:

$$f_{s^{(1)}}(x) = \frac{x^{M_R-1}e^{-x}}{\Gamma(M_R)} \quad x \geq 0. \tag{108}$$

Then

$$\mathbb{E}[1/s^{(1)}] = \int_0^\infty \frac{x^{M_R-2}e^{-x}}{\Gamma(M_R)} dx = \frac{1}{M_R - 1}. \tag{109}$$

Therefore, the zero-outage capacity for $M_T = 1$ and $M_R > 1$ is positive. In fact, the zero-outage capacity is

$$C_0 = \log_2(1 + \rho(M_R - 1)), \tag{110}$$

which is nonzero for $M_R > 1$ [2]. ♦

Another example is the principal eigenmode in Section 3.1.1.

Example 4.12. For the principal eigenmode transmission, the zero-outage capacity is:

$$C_0 = \log_2\left(1 + \frac{\rho}{\mathbb{E}[1/s^{(1)}]}\right). \tag{111}$$

Table 2 summarizes $\mathbb{E}[1/s^{(1)}]$ for spatially uncorrelated channels. ♦

For $M \geq 2$, evaluating (103) is usually difficult, even for 2×2 MIMO as illustrated in the following example.

Table 2: $\mathbb{E}[1/s^{(1)}]$ for an $M \times M$ spatially uncorrelated Rayleigh fading with $M \in \{1, 2, 3, 4, 6, 8, 10\}$.

	1×1	2×2	3×3	4×4	6×6	8×8	10×10
$\mathbb{E}[1/s^{(1)}]$	∞	0.3864	0.1755	0.1110	0.0627	0.0433	0.0329

Example 4.13. When $M_T = M_R = 2$, so as to achieve C_0, the required SNR is

$$\begin{aligned}
\rho &= (2^{C_0} - 1) \int_0^{\infty} \int_0^{2^{-C_0} x_1} \frac{p(x_1, x_2)}{x1} dx_2 dx_1 \\
&\quad + 2\sqrt{2^{C_0}} \int_0^{\infty} \int_{2^{-C_0} x_1}^{x_1} \frac{p(x_1, x_2)}{\sqrt{x_1 x_2}} dx_2 dx_1 \\
&\quad - \int_0^{\infty} \int_{2^{-C_0} x_1}^{x_1} \left(\frac{1}{x_1} + \frac{1}{x_2} \right) p(x_1, x_2) dx_2 dx_1 \\
&= \left[2 \log(1 + 2^{-C_0}) - \left(\frac{2}{1 + 2^{C_0}} \right)^2 \right] (2^{C_0} - 1) \\
&\quad + 2\sqrt{2^{C_0}} \int_0^{\infty} \int_{2^{-C_0} x_1}^{x_1} \frac{p(x_1, x_2)}{\sqrt{x_1 x_2}} dx_2 dx_1 \\
&\quad - \left[\left(\frac{2^{-C_0} - 1}{2^{-C_0} + 1} \right)^2 + 2 \left(\frac{2^{-C_0} - 1}{2^{-C_0} + 1} \right) + 2 \log \left(\frac{2}{2^{-C_0} + 1} \right) \right. \\
&\quad \left. + 1 + \frac{2^{-C_0} - 3}{(2^{-C_0} + 1)^2} - \frac{1}{2^{-C_0} + 1} + 2 \log \left(\frac{2^{-C_0} + 1}{2} \right) + 2C_0 \log(2) \right], \quad (112)
\end{aligned}$$

where

$$p(x_1, x_2) = (x_1 - x_2)^2 e^{-x_1} e^{-x_2} \quad \text{for } x_1 \geq x_2 \geq 0. \quad (113)$$

The second integral in (112) is difficult to evaluate. At high SNR, the above equation can be approximated as

$$C_0 \approx 2 \log_2(\rho) - 2 \log_2(\pi), \quad (114)$$

where we use

$$\lim_{C_0 \to \infty} (1 + 2^{-C_0})^{2^{C_0}} = e, \quad (115)$$

and

$$\lim_{C_0 \to \infty} 2 \int_0^{\infty} \int_{2^{-C_0} x_1}^{x_1} \frac{p(x_1, x_2)}{\sqrt{x_1 x_2}} dx_2 dx_1 = \pi. \quad (116)$$

♦

We have seen that C_0 is not always nonzero. Generally, the following theorem states under what conditions C_0 is nonzero.

Theorem 4.5. $C_0 > 0$ if $\mathbb{E}[1/G_{\mathcal{N}_u}] < \infty$, where $\mathcal{N}_u$ is the universe index set in (14).

Proof. See [7]. ■

We note that C_0 is always suboptimal to C_{TRCSI} since the water-filling for C_0 is performed over space (m) and frequency (n), while the water-filling for C_{TRCSI} is performed over all channel realizations as well. The omission of water-filling over channel realizations can greatly simplify the power allocation. In the following example, we compare C_0 and C_{TRCSI}.

Example 4.14. Consider a spatially uncorrelated $M \times M$ channel. Figure 18 shows C_0 and C_{TRCSI} for $M \in \{1, 2, 4, 6\}$. When $M = 1$ (SISO), C_0 is zero, as mentioned in Example 4.11, where C_0 incurs an infinite SNR penalty relative to C_{TRCSI}. When $M = 2$, however, C_0 is nonzero and not far from C_{TRCSI}. For $M = 4$ and $M = 6$, C_0 approaches C_{TRCSI} closely. ♦

In fact, the difference between C_{TRCSI} and C_0 vanishes as the number of antennas tends to infinity.

Theorem 4.6. C_0 asymptotically converges to C_{TRCSI} as $M \to \infty$.

Proof. See [7]. ■

A conclusion of this section is that the advantage of knowing CSI at the transmitter is tremendous with the long-term constraint. Theorem 4.6 suggests that the zero-outage capacity is nearly optimal, namely close to the average capacity. In stark contrast, without CSI at the transmitter, the outage probability cannot be made zero at a finite SNR. In other words, knowing CSI at the transmitter offers an infinite SNR advantage.

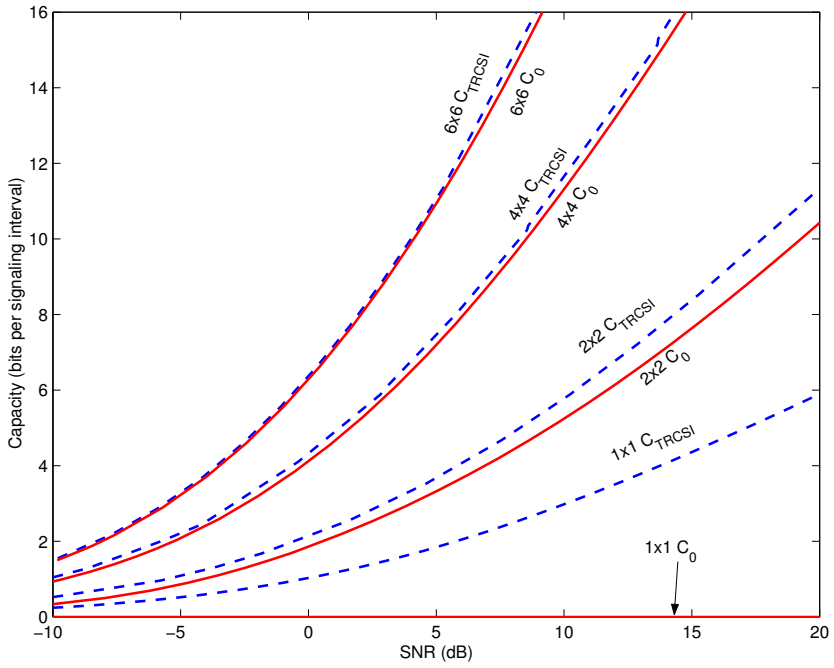

Figure 18: C_0 and C_{TRCSI} on a $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel for $M \in \{1, 2, 4, 6\}$.

4.3.4 Minimum Outage Probability

To achieve zero-outage capacity, the transmitter always makes sure that no outage occurs. In other words, however bad a channel is, the transmitter inverts the channel for each scalar channel, such that $e_n^{(m)} = (2^{r_n^{(m)}} - 1)/s_n^{(m)}$. In the average sense, this channel inversion is acceptable as long as the long-term constraint is satisfied. However, small $\{s_n^{(m)}\}$ can cause a very high instantaneously required energy:

$$E(\boldsymbol{s}) = \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} e_n^{(m)}, \tag{117}$$

where $\boldsymbol{s} = [s_1^{(1)}, \ldots, s_1^{(M)}, \ldots, s_N^{(1)}, \ldots, s_N^{(M)}]$.

We define *the peak-power probability* as

$$P_{\text{PEAK}} = \text{Prob}\left[E(\boldsymbol{s}) > E_{\text{TH}}\right], \tag{118}$$

for a given threshold E_{TH}. For any finite M_T and M_R and a finite channel memory L, P_{PEAK} is always nonzero since the joint distribution function of $\boldsymbol{s}$ is continuous and

nonzero for any $s_n^{(m)} > 0$. There is a tradeoff between P_{OUT} and P_{PEAK} in the sense that we never make both zero at the same time.

Example 4.15. When $M_T = 1$ and $M_R > 1$ (SIMO), we have seen that C_0 is positive for spatially uncorrelated Rayleigh flat fading ($L = 0$). In this case, the peak-power probability, at a rate of R, can be evaluated as

$$P_{\text{PEAK}} = 1 - \frac{\Gamma(M_R, \frac{2^R - 1}{\rho_{\text{TH}}})}{\Gamma(M_R)}, \tag{119}$$

where $\rho_{\text{TH}} = E_{\text{TH}}/N_0$. Figure 19 illustrates P_{PEAK} at $R = 4$ bits per signaling interval as M_R ranges from 2 to 10. We can see the tendency that P_{PEAK} decays more sharply as M_R increases. When $M_R = 2$, the probability that $E(\boldsymbol{s})$ exceeds 20 dB is 10^{-2}, but we expect a very small probability for $M_R = 10$. In Figure 19, we never attain $P_{\text{PEAK}} = 0$ for $R > 0$ however large M_R is. ♦

Example 4.16. Consider a spatially uncorrelated $M \times M$ MIMO channel with $L = 0$. Figure 20 illustrates P_{PEAK} at $R = 4$ bits per signaling interval for $M \in \{2, 4, 6\}$. Again, we can see a distinct reduction in P_{PEAK} as M grows, but P_{PEAK} cannot be made zero. ♦

Clearly, the above examples show that employing more antennas improves the peak-power performance. For a small size of antenna array, forcing zero outage probability results in very large instantaneous energy requirement at the transmitter with a relatively large probability, which is practically impossible. The peak-power viewpoint motivates us to use an outage region $\mathcal{R}_{\text{OFF}}$. With an outage region, however, the actual transmission rate reduces to $R(1 - P_{\text{OUT}})$ if R is the transmission rate when $\boldsymbol{s} \notin \mathcal{R}_{\text{OFF}}$.

We consider a power-allocation problem: find $\{e_n^{(m)}\}$ such that P_{OUT} is minimized subject to the long-term constraint when ρ and R are given. This problem is solved for a single transmit antenna in [13] and generalized to MIMO in [7].

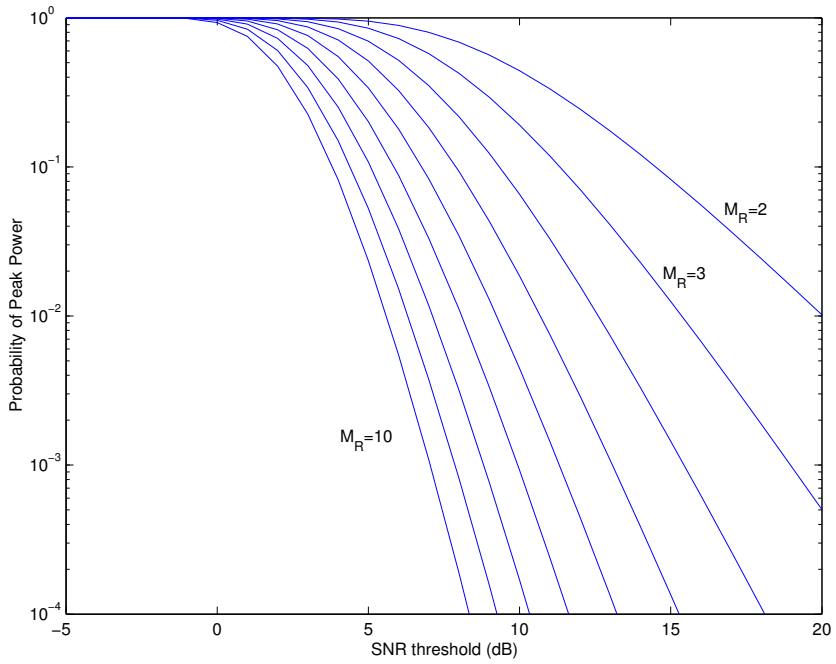

Figure 19: Peak-power probability at $R = 4$ bits per signaling interval for $M_T = 1$ and $M_R > 1$ on a spatially uncorrelated Rayleigh-flat-fading channel for $M_R \in \{2, 3, \ldots, 10\}$.

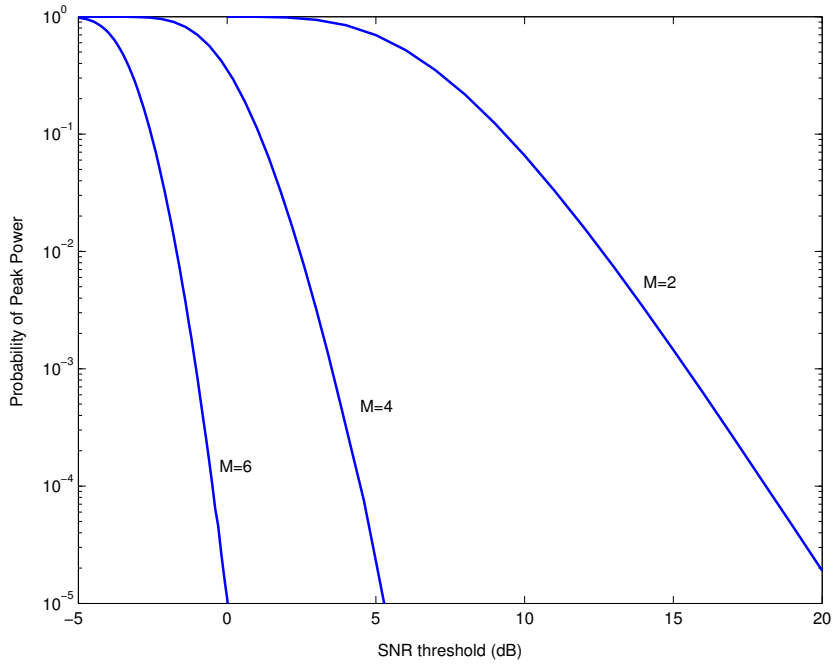

Figure 20: Peak-power probability at $R = 4$ bits per signaling interval on an $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel for $M \in \{2, 4, 6\}$.

Theorem 4.7 (Minimum-Outage-Probability Allocation). For given R and ρ, the outage probability is minimized when the power allocation is given by:

$$\frac{e_n^{(m)}}{N_0} = \begin{cases} \left\{\lambda - \frac{1}{s_n^{(m)}}\right\}^+ & \boldsymbol{s} \notin \mathcal{R}_{\text{OFF}} \\ 0 & \boldsymbol{s} \in \mathcal{R}_{\text{OFF}} \end{cases}, \tag{120}$$

where λ is determined by (101). The outage region $\mathcal{R}_{\text{OFF}}$ is determined such that

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left\{\lambda - \frac{1}{s_n^{(m)}}\right\}^+ 1\{\boldsymbol{s} \notin \mathcal{R}_{\text{OFF}}\}\right] = \rho. \tag{121}$$

Proof. See [7]. ■

Remark 4.2. The minimum-outage-probability allocation in Theorem 4.7 reduces to the allocation in Theorem 4.4 for the zero-outage capacity by setting $E_{\text{TH}} = 0$.

Example 4.17. Consider a SIMO channel ($M_T = 1$ and $M_R \geq 1$). The instantaneous SNR requirement is

$$\rho(s) = \begin{cases} \frac{2^R-1}{s} & \frac{2^R-1}{s} \geq \rho_{\text{TH}} \\ 0 & \frac{2^R-1}{s} < \rho_{\text{TH}} \end{cases}, \tag{122}$$

where ρ_{TH} is a threshold. The outage probability is

$$P_{\text{OUT}} = \text{Prob}\left[\frac{2^R-1}{s} > \rho_{\text{TH}}\right] = 1 - \frac{\Gamma(M_R, \rho_{\text{TH}})}{\Gamma(M_R)}, \tag{123}$$

when SNR is

$$\rho = \mathbb{E}[\rho(s)] = (2^R - 1)\frac{\Gamma(M_R - 1, \rho_{\text{TH}})}{\Gamma(M_R)}, \tag{124}$$

P_{OUT} and ρ are linked through ρ_{TH}. Figure 21 illustrates P_{OUT} at $R = 4$ bits per signaling interval for $M_R \in \{1, 2, 4, 10\}$. We confirm that P_{OUT} becomes zero at $\rho = (2^R - 1)\frac{\Gamma(M_R-1,0)}{\Gamma(M_R)}$ for $M_R > 1$ when $\rho_{\text{TH}} = 0$.

As a benchmark, we also plot P_{OUT} in (55) when CSI is unknown to the transmitter. See that the advantage knowing CSI at the transmitter is remarkable, especially for small M_R. However, when M_R is large, the advantage is reduced, but exists. For example, the SNR advantage is approximately 5 dB to achieve $P_{\text{OUT}} = 10^{-3}$ for $M_R = 10$. ♦

Example 4.18. Consider a MIMO channel with $M_T = M_R = M$. As no closed-form for P_{OUT} is known, we resort to Monte-Carlo simulations by generating 10^5 independent channels. Figure 22 illustrates P_{OUT} at $R = 5$ bits per signaling interval for $M \in \{2, 4, 6\}$, where we can confirm the infinite asymptotic slope of P_{OUT}. In other words, $C_0 = R = 5$ is achieved at the point where P_{OUT} drops vertically.

Also plotted is P_{OUT} with the short-term constraint for both cases where CSI is known at the transmitter (TRCSI) and where CSI is available only at the receiver (RCSI). Notice that the slopes of P_{OUT} for both the cases are finite and identical, implying that CSI at the transmitter cannot achieve an infinite diversity order with the short-term constraint. However, P_{OUT} with TRCSI is smaller than P_{OUT} with RCSI. Intriguingly, the gap between these two P_{OUT} curves becomes larger as M grows. This fact favors knowing CSI at the transmitter even with the short-term constraint. ♦

4.4 Optimal Outage Region in MIMO

In Theorem 4.7, we addressed the problem of finding minimum outage probability P_{OUT} for given R and ρ. At an SNR of ρ, however, it should be pointed out that $R(1 - P_{\text{OUT}})$ is not necessarily the maximum achievable rate with an outage region. For instance, there might be R' such that $R'(1 - P'_{\text{OUT}}) \geq R(1 - P_{\text{OUT}})$ at the same SNR. In other words, there are infinitely many combinations of (R, P_{OUT}), and the maximum achievable rate when outage region is allowed is the one that has largest $R(1 - P_{\text{OUT}})$ among them. In this section, we consider the optimal choice of outage region in the sense that it maximizes the achievable rate.

Definition 4.7 (Outage-Region Capacity). Let $\mathcal{R}_{\text{OFF}}$ denote outage region. The maximum achievable rate is achieved by solving

$$C_{\text{OR}} = \sup_{\mathcal{R}_{\text{OFF}}} \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} \left\{ \log_2(\lambda s_n^{(m)}) \right\}^+ (1 - P_{\text{OUT}}), \tag{125}$$

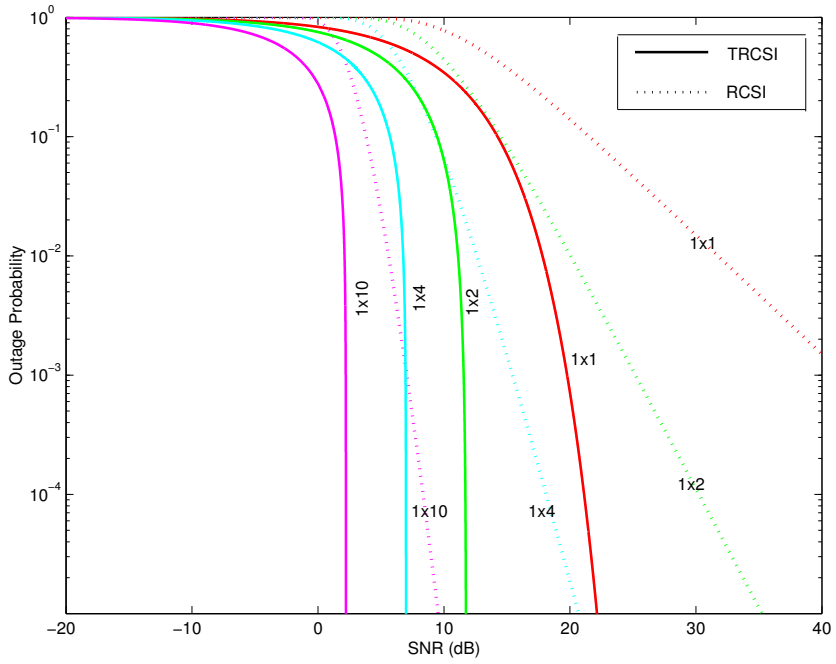

Figure 21: Outage probability for $M_T = 1$ and $M_R \geq 1$ spatially uncorrelated Rayleigh fading when CSI is known at the transmitter at $R = 4$ bits per signaling interval. As a benchmark, outage probability when CSI is unknown at the transmitter is also plotted.

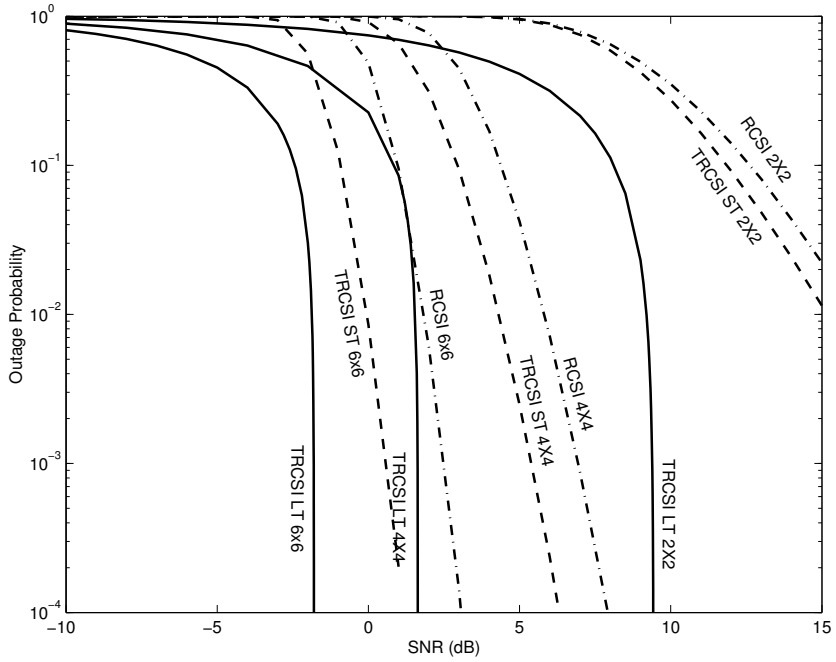

Figure 22: Outage probability on an $M \times M$ spatially uncorrelated Rayleigh fading for $M \in \{2, 4, 6\}$ when CSI is known at the transmitter [7].

called *outage-region (OR) capacity*, where λ and P_{OUT} are determined by

$$P_{\text{OUT}} = \text{Prob}\left[\{s_n^{(m)}\} \in \mathcal{R}_{\text{OFF}}\right], \tag{126}$$

and

$$\rho = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left\{\lambda - \frac{1}{s_n^{(m)}}\right\}^+ \middle| \{s_n^{(m)} \notin \mathcal{R}_{\text{OFF}}\right]. \tag{127}$$

♦

For a SIMO system ($M_T = 1$) on flat fading, the outage-region capacity was studied [2, 28], known as *truncated channel inversion*. Since there is only $M = 1$ nonzero eigenvalue ($s^{(1)}$), the instantaneous SNR requirement at the transmitter is

$$\rho(s^{(1)}) = \begin{cases} \frac{2^R-1}{s^{(1)}} & \frac{2^R-1}{s^{(1)}} \le \rho_{\text{TH}} \\ 0 & \text{otherwise} \end{cases}, \tag{128}$$

where the outage region is $\mathcal{R}_{\text{OFF}} = \left\{s^{(1)} : \frac{2^R-1}{s^{(1)}} \le \rho_{\text{TH}}\right\}$.[6] The transmitter truncates $s^{(1)} \in \mathcal{R}_{\text{OFF}}$ and stops transmission, which is the reason it is called truncated channel inversion. The optimal choice of ρ_{TH} that maximizes $R \cdot \text{Prob}[s^{(1)} \notin \mathcal{R}_{\text{OFF}}]$ is as follows.

Theorem 4.8. For $M_T = 1$ and $M_R \ge 1$ (SIMO) on a spatially uncorrelated Rayleigh-flat-fading channel,

$$C_{\text{OR}} = \max_{\rho_{\text{TH}}} \frac{\Gamma(M_R, \rho_{\text{TH}})}{\Gamma(M_R)} \log_2\left(1 + \frac{\rho}{\frac{\Gamma(M_R-1,\rho_{\text{TH}})}{\Gamma(M_R)}}\right). \tag{129}$$

When $M_R = 1$ (SISO), it can be more simplified:

$$C_{\text{OR}} = \max_{\rho_{\text{TH}}} e^{-\rho_{\text{TH}}} \log_2\left(1 + \frac{\rho}{E_1(\rho_{\text{TH}})}\right), \tag{130}$$

where $E_1(x)$ is defined in (68).

Proof. See [2]. ■

[6] We discard the tone index n.

Remark 4.3. As SNR tends to infinity, ρ_{TH} converges to zero. Therefore, the spatial multiplexing order of C_{OR}, $\frac{\Gamma(M_R,\rho_{\text{TH}})}{\Gamma(M_R)}$, is unity since $\Gamma(x,0)=\Gamma(x)$. ♦

Corollary 4.2. For any M_T and M_R, we have

$$C_{\text{TRCSI}} \geq C_{\text{OR}} \geq C_0. \tag{131}$$

Proof. The optimization in (125) includes the case where $\mathcal{R}_{\text{OFF}}$ is a null set corresponding to C_0. Therefore, we have $C_{\text{OR}} \geq C_0$. We view C_{OR} as a most primitive form of the rate adaptation, choosing between R and 0. As C_{TRCSI} has more freedom on the rate adaptation, we infer that $C_{\text{TRCSI}} \geq C_{\text{OR}}$. ■

Example 4.19. For $M_T = 1$ and $M_R = 2$, spatially uncorrelated Rayleigh flat fading, Figure 23 illustrates C_{OR} along with C_0. The figure shows the gap between them is relatively large, while it seems to converges to zero as SNR grows. Intuitively, ρ_{TH} becomes small as SNR grows, which implies that C_{OR} is close to C_0 at high SNR. In Figure 23, we also plot $R(1-P_{\text{OUT}})$ against ρ, where P_{OUT} is minimized according to Theorem 4.7 for $R \in \{4,6,8\}$. When SNR is large enough so that $P_{\text{OUT}} = 0$, $R(1-P_{\text{OUT}})$ approaches C_0. On its way to C_0, the curve $R(1-P_{\text{OUT}})$ once meets the curve C_{OR}. At the meeting point, it is $R(1-P_{\text{OUT}})$ is maximum in the sense that $R(1-P_{\text{OUT}}) \geq R'(1-P'_{\text{OUT}})$ for $R \neq R'$. ♦

For the MIMO case, to our knowledge, no closed-form formula for C_{OR}, such as (129) or (130), exists, and it can be only solved by a numerical method. In the following, we propose procedures to obtain the outage-region capacity. To simplify the notation, we assume flat fading and ignore the tone index n.

Given C_{OR}

For $R \in [C_{\text{OR}}, \infty)$, repeat:

1. Determine λ from $R = \sum_{m=1}^{M} \left\{\log_2(\lambda s^{(m)})\right\}^+$.

2. Set $P_{\text{OUT}}(R) = 1 - \frac{C_{\text{OR}}}{R}$.

3. Determine α such that $\mathbb{E}\left[1\left\{\sum_{m=1}^{M}\{\lambda - \frac{1}{s^{(m)}}\}^{+}\right\} > \alpha\right] = P_{\text{OUT}}(R)$.

4. $\gamma = \begin{cases} \sum_{m=1}^{M}\left\{\lambda - \frac{1}{s^{(m)}}\right\}^{+} & \sum_{m=1}^{M}\left\{\lambda - \frac{1}{s^{(m)}}\right\}^{+} > \alpha \\ 0 & \text{otherwise} \end{cases}$.

5. $\rho(R) = \mathbb{E}[\gamma]$.

Choose R that has minimum $\rho(R)$.

Then, R with minimum $\rho(R)$ is C_{OR} at an SNR of $\rho(R)$. These procedures can be interpreted as

$$\rho = \inf_{R \in [C_{\text{OR}}, \infty)} \rho(R, P_{\text{OUT}}(R)), \tag{132}$$

where $\rho(R, P_{\text{OUT}}(R))$ is the average SNR required by R and $P_{\text{OUT}}(R)$ in Theorem 4.7. Notice that $R(1 - P_{\text{OUT}}(R)) = R\left(1 - \left(1 - \frac{C_{\text{OR}}}{R}\right)\right) = C_{\text{OR}}$ for any R in the above procedures.

Example 4.20. Consider an $M \times M$ spatially uncorrelated Rayleigh-flat-fading channel. Figure 24 shows C_{OR} for $M \in \{1, 2, 4, 6\}$. As a benchmark, C_{TRCSI} and C_0 are also plotted from Figure 18. We confirm that the inequalities, $C_{\text{TRCSI}} \geq C_{\text{OR}} \geq C_0$, hold. As shown in Example 4.14, the gap between C_{TRCSI} and C_0 becomes very small as M grows. The gap between C_{OR} and C_0 is even smaller. Even for $M = 2$, the advantage of C_{OR} almost vanishes at high SNR, which is contrasted to when $M = 1$, where C_{OR} is nonzero while $C_0 = 0$, that is, an infinite SNR penalty. ♦

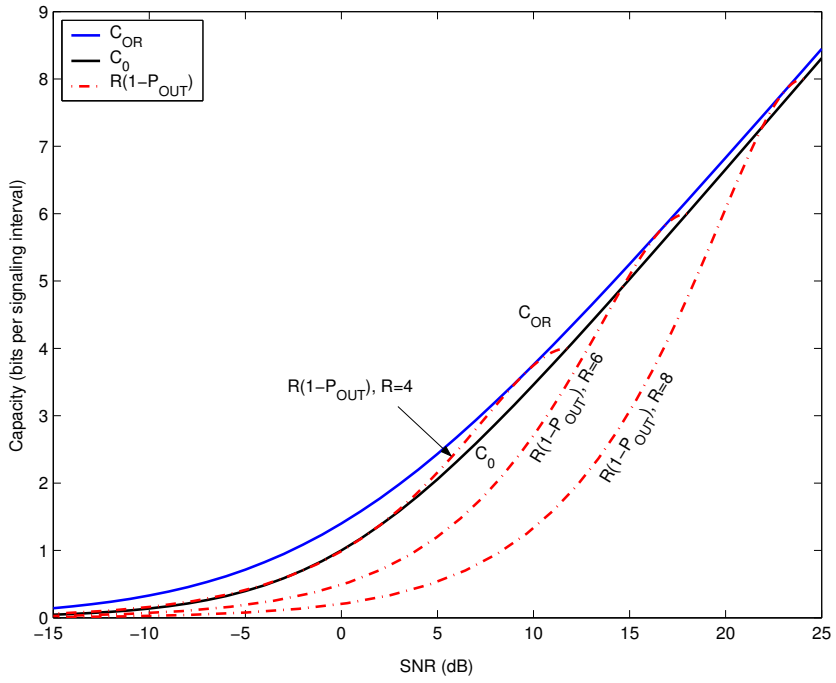

Figure 23: Comparison between C_{OR} and C_0 on a 2×1 spatially uncorrelated channel with $L = 0$ memory. Also plotted is $R(1 - P_{\text{OUT}})$ with P_{OUT} minimized according to Theorem 4.7 for $R \in \{4, 6, 8\}$.

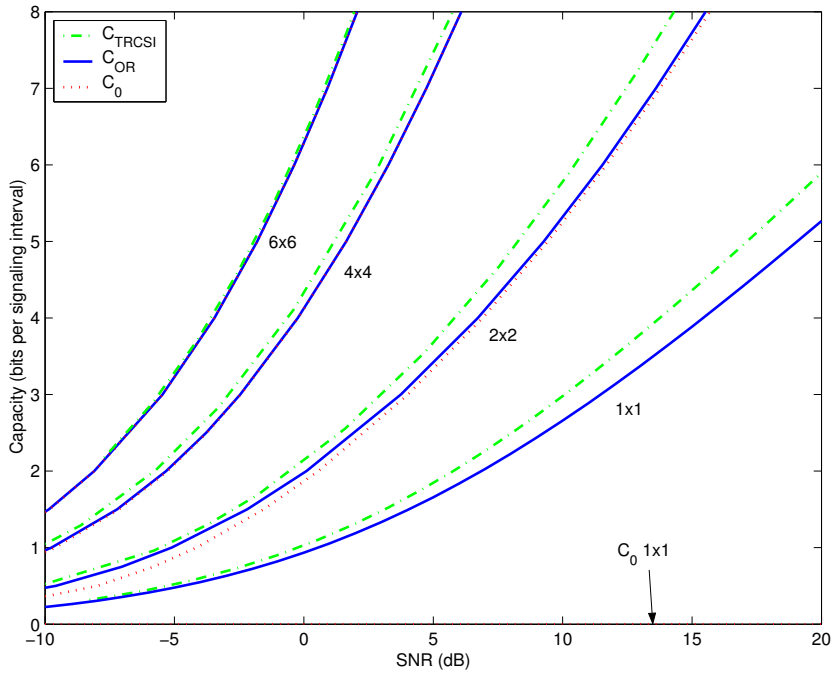

Figure 24: Comparison of C_{TRCSI}, C_{OR}, and C_0 for $M \times M$ Rayleigh flat fading with $M \in \{1, 2, 4, 6\}$.

CHAPTER 5

ASYMPTOTIC BEHAVIORS AT HIGH SNR

The previous chapter reviewed information-theoretical concepts of MIMO fading channels, such as average capacity or outage probability, depending on the availability of CSI at the transmitter. In this chapter, we examine the asymptotic behaviors of outage probability and capacity at high SNR. In the analysis of capacity, the high-SNR assumption is popular because it makes the analysis tractable, while analysis is difficult or impossible at normal SNR. Moreover, even when a closed-form formula exists at normal SNR, such as average capacity with CSI at the receiver in (66), the formula is sometimes too complicated to get insight from it. High-SNR analysis will provide simplicity and insight.

In the case of MIMO, high-SNR analysis is particularly useful to understand the behaviors at a finite number of antennas. For example, Theorem 4.6 showed that C_0 converges to C_{TRCSI} as $M \to \infty$, but the theorem says nothing about the convergence speed. We will use the high-SNR assumption to show whether the convergence is fast or slow. Also, the high-SNR assumption is directly related to the spatial multiplexing order in Definition 4.4 or the diversity order in Definition 4.5.

In this chapter, we present our contributions on the high-SNR analysis. More precisely, the contributions are as follows.

- We prove the diversity order of MIMO-OFDM is $M_T M_R(L+1)$ when CSI is only available at the receiver.
- With the short-term energy constraint in (51), we prove that a closed-loop MIMO system achieves full diversity and multiplexing orders.

- We derive the asymptotes for C_{RCSI}, C_{TRCSI}, C_ϵ, and C_0 in terms of the geometric means of $\{s_n^{(m)}\}$, and show that the spatial multiplexing order is $M = \min(M_T, M_R)$ for all the cases.

- We derive a closed-form formula for the approximation of C_{RCSI} at high SNR when the channel is spatially uncorrelated.

- We show that the outage capacity C_ϵ, when CSI is available only at the receiver, increases with L, but is finite when $L \to \infty$.

- We show that the convergence ($C_0 \to C_{\text{TRCSI}}$ as $M \to \infty$) in Theorem 4.6 is fast, and the difference between them is small for a moderate M.

- We investigate the impact of L on C_0, where an upper bound of C_0 is analytically derived.

We begin with the diversity order, and then move to the capacity.

5.1 Diversity Order

Diversity order, defined in Definition 4.5, measures the degree of protection against the effects of fading. When CSI is unknown to the transmitter, the diversity order is finite and limits the outage probability as we have seen in Section 4.2.2. On M_T-input M_R-output MIMO channels, the diversity order can be as large as $M_T M_R$. When the transmitter knows CSI, we showed that the diversity order can be infinite with the long-term energy constraint in Section 4.3.3. With the short-term energy constraint, however, the diversity order is finite even with the transmitter CSI.

5.1.1 CSI Unknown to Transmitter

On an M_T-input M_R-output MIMO channel, there are $M_T M_R$ links between transmitter and receiver. If the channel is spatially uncorrelated, the fading in each link is independent, and the probability that all links are in deep fades is far less probable

than in a SISO system. Intuitively, we expect $M_T M_R$ times better protection against the effects of fading in MIMO. On Rayleigh flat fading, we can derive the diversity order.

Theorem 5.1. On spatially uncorrelated Rayleigh flat fading ($L = 0$), the diversity order is $M_T M_R$.

Proof. See [62]. ■

Theorem 5.1 indicates a significant increase in diversity order, $M_T M_R$ times higher than SISO flat fading. We expect a further increase in diversity order when the channel has memory. In Section 4.2.2, we had a discussion that a MIMO channel with L memory has a degree of freedom of $M_T M_R(L+1)$ and hence significantly better outage performance is expected. In [43], matched-filter bound is used to show that a system with two transmit antenna has a diversity order of $2M_R(L+1)$. In the following, we derive the diversity order in a frequency-selective channel directly from its definition.

Proposition 5.1. On a spatially uncorrelated MIMO channel with memory L, the diversity order is at least $M_T M_R(L+1)$ when the power profile is uniform.

Proof. Deferred to Section 5.3. ■

From (61), the mutual information is bounded by

$$\begin{aligned} I_{\text{RCSI}} &\geq \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(\frac{\rho}{M_T}s_n^{(m)}\right) \\ &= \log_2\left[\left(\prod_{n=1}^{N}\prod_{m=1}^{M}s_n^{(m)}\right)^{1/N}\left(\frac{\rho}{M_T}\right)^{M}\right]. \end{aligned} \tag{133}$$

Then,

$$P_{\text{OUT}} \leq \text{Prob}\left[G_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M_T}\right], \tag{134}$$

where $G_{\mathcal{N}_u}$ is the universe index set in (14). At high SNR, the upper bound (134) is very tight since $\log(1+x) \approx \log(x)$ when $x \gg 1$. We find the lower bound useful as the PDF of $G_{\mathcal{N}_u}$ is independent of ρ, while the PDF of I_{RCSI} changes with ρ. Also, from $G_{\mathcal{N}_u} \geq A_{\mathcal{N}_u}$, we have the inequality:

$$\text{Prob}\left[G_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M_T}\right] \geq \text{Prob}\left[A_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M_T}\right]. \tag{135}$$

We remark that (135) is a lower bound of the high-SNR approximation of P_{OUT}, namely $\text{Prob}\left[G_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M_T}\right]$. The lower bound in (135) can be analytically evaluated for the uniform power profile, such that

$$\text{Prob}\left[A_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M}\right] = 1 - \frac{\Gamma\left(M_T M_R(L+1), (M(L+1)\frac{2^{R/M}}{\rho/M})\right)}{\Gamma(M_T M_R(L+1))}. \tag{136}$$

From (136), we conjecture that the diversity order is no greater than $M_T M_R(L+1)$ with the uniform power profile.

Example 5.1. We consider a 4×4 spatially uncorrelated MIMO channel with memory L. In Figure 25, we plot $\text{Prob}\left[A_{\mathcal{N}_u} < \frac{2^{R/M}}{\rho/M}\right]$ in (136) at $R = 10$ bits per signaling interval as L ranges from 0 to 10. Figure 25 clearly shows the increase in the diversity order as L grows, which agrees with Proposition 5.1. ♦

5.1.2 Eigenbeamforming with Short-Term Constraint

Now we turn our attention to the case where the transmitter knows CSI. With the short-term energy constraint in (51), achieving zero outage probability is impossible, and the diversity order is finite in contrast to the infinite diversity order with the long-term energy constraint. as illustrated in Figure 22. We saw that eigenbeamforming in Chapter 3 is optimal in the sense that it does not incur any information loss owing to the unitary filters. In this section, we examine the diversity order of the parallel channels in (39), created by eigenbeamforming, with the short-term constraint. We assume flat fading in this section, and ignore the tone index n for the notational convenience.

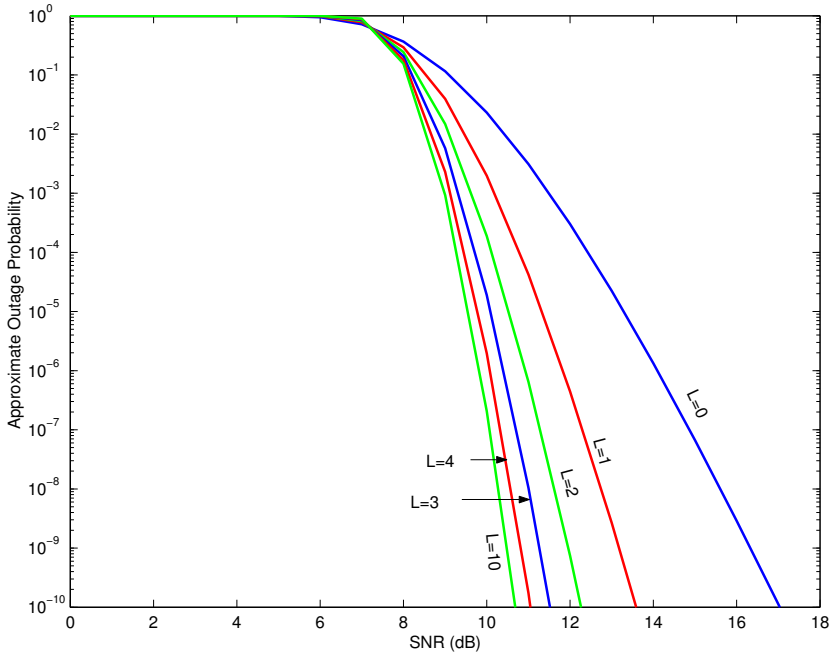

Figure 25: Lower bound in (136) of the outage probability at $R = 10$ bits per signaling interval on a 4×4 spatially uncorrelated Rayleigh-fading channel with memory $L \in \{0, 1, 2, 3, 4, 10\}$.

First, we consider the principal-eigenmode transmission discussed in Section 3.1.1, where only the largest eigenmode, corresponding to $s^{(1)}$, is used for transmission. From the marginal CDF of $s^{(1)}$ in (31), outage probability is given by

$$P_{\text{OUT}} = \frac{\Gamma_M(M)}{\Gamma_M(M_T + M_R)} \chi^{M_T M_R} {}_1F_1(M; M_T + M_R; -\chi \mathbf{I}_M), \tag{137}$$

where $\chi = (e^R - 1)/\rho$, assuming a spatially uncorrelated flat-fading channel.

Lemma 5.1. On a spatially uncorrelated flat-fading MIMO channel, the diversity order of the principal-eigenmode transmission is $M_T M_R$.

Proof. From (137), we have

$$\frac{\log(P_{\text{OUT}})}{\log(\rho)} = \frac{\log\left(\frac{\Gamma_M(M)}{\Gamma_M(M_T+M_R)}\right) + M_T M_R \log(\chi) + \log\left({}_1F_1(M; M_T + M_R; -\chi \mathbf{I}_M)\right)}{\log\left(\frac{2^R-1}{\chi}\right)}. \tag{138}$$

From the zonal representation of ${}_1F_1(;;)$ [30], ${}_1F_1(M; M_T + M_R; -\chi \mathbf{I}_M)$ with $\chi = 0$ ($\rho \to \infty$) is positive and finite. Therefore,

$$-\lim_{\rho\to\infty} \frac{\log(P_{\text{OUT}})}{\log(\rho)} = M_T M_R, \tag{139}$$

since the first and last terms in the numerator of (138) are finite. ■

Lemma 5.1 suggests that the diversity order is full even when the principal eigenmode is only used. Thus, if we use all eigenmodes (*i.e.* all scalar channels) of a MIMO system, we expect a diversity order of at least $M_T M_R$.

Proposition 5.2. On a spatially uncorrelated flat-fading channel, the diversity order of the parallel channels in (80) is $M_T M_R$.

Proof. First, the mutual information of (80) is always larger than or equal to the mutual information of the principal eigenmode. Thus, the diversity order of (80) is at least $M_T M_R$. To show that the diversity order cannot exceed $M_T M_R$, we consider a bank of M scalar channels $\{s^{(1)}, s^{(1)}, \ldots, s^{(1)}\}$, whose mutual information is clearly larger than or equal to that of (80). The outage probability of $\{s^{(1)}, s^{(1)}, \ldots, s^{(1)}\}$ is

$$P_{\text{OUT}} = \text{Prob}\left[s^{(1)} < \frac{2^{R/M} - 1}{\rho/M}\right]. \tag{140}$$

Since $\frac{2^{R/M}-1}{\rho/M}$ goes to zero as $\rho \to \infty$, the diversity order is $M_T M_R$ from Lemma 5.1. Therefore, $M_T M_R$ is also an upper limit for the diversity order of (80), which completes the proof. ■

The results in Proposition 5.2 is somewhat obvious. We already saw the outage probability with the short-term constraint in Figure 22, which illustrates that the diversity order of "TRCSI ST" (CSI known to both the transmitter and receiver with the short-term constraint) is equal to the diversity order of "RCSI" (CSI known only at the receiver). We also explained that the SNR gap between "TRCSI ST" and "RCSI" becomes wider as M grows.

5.2 Capacity at High SNR

The spatial multiplexing order in Definition 4.4 suggests that capacity is linearly proportional to log(SNR) as SNR tends to infinity. In this section, we derive the asymptotes of average and outage capacity in terms of the geometric mean of eigenvalues $\{s_n^{(m)}\}$, and investigate the spatial multiplexing orders.

5.2.1 CSI Unknown to Transmitter

First, we assume that CSI is available only at the transmitter.

5.2.1.1 Average Capacity

We begin with the average capacity in Lemma 4.1. We already saw in Figure 6 that the ratio $\frac{C_{\text{RCSI}}}{\log_2(\rho)}$ converges to M as SNR ρ tends to infinity when the channel is spatially uncorrelated. Generally, the asymptote of C_{RCSI} is as follows.

Proposition 5.3. At high SNR, the asymptote of C_{RCSI} is

$$C_{\text{RCSI}} \to M \log_2\left(\frac{\rho}{M_T}\right) + M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right], \tag{141}$$

where $\mathcal{N}_u$ is the universe index set in (14).

Proof. Deferred to Section 5.4. ■

In fact, $M \log_2\left(\frac{\rho}{M_T}\right) + M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right]$ is a lower bound on C_{RCSI} for any ρ, and Proposition 5.3 indicates that the bound asymptotically approaches C_{RCSI} as $\rho \to \infty$. An implication of Proposition 5.3 is that the spatial multiplexing order is M. The term $M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right]$ accounts for the parallel shift of the asymptote of (141). Note that

$$\begin{aligned} M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right] &= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[\sum_{m=1}^{M}\log_2(s_n^{(m)})\right] \\ &= \mathbb{E}\left[\sum_{m=1}^{M}\log_2(s_n^{(m)})\right] \quad \text{for any } n, \end{aligned} \tag{142}$$

from Lemma 2.1. Hence, C_{RCSI} at high SNR is independent of n, and it is not affected by L for spatially uncorrelated fading, which agrees with Lemma 4.1.

If the channel is spatially uncorrelated and square ($M_T = M_R = M$), we evaluate $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right]$ by averaging over 10,000 independent channels, resulting in $\{-0.8314, -0.2256, 2.9097, 7.5466\}$ for $M \in \{1, 2, 4, 6\}$, respectively. In fact, we can evaluate the approximation of C_{RCSI} analytically.

Proposition 5.4. At high SNR for a finite M_T and M_R,

$$C_{\text{RCSI}} \approx M \log_2\left(\frac{\rho}{M_T}\right) + K - \Gamma/\log(2), \tag{143}$$

where $\Gamma \approx 0.577215665$ is the Euler constant and

$$\begin{aligned} K &= \frac{1}{\log(2)} \sum_{k=0}^{M-1} \sum_{l=0}^{k} \sum_{i=0}^{2l} \left\{ \frac{(-1)^i (2l)!(D+i)!}{2^{2k-i} l! i! (D+l)!} \right. \\ &\times \left. \binom{2k-2l}{k-l} \binom{2l+2D}{2l-i} \sum_{j=0}^{D+i} \frac{1}{j} \right\}, \end{aligned} \tag{144}$$

with $D = \max(M_T, M_R) - \min(M_T, M_R)$.

Proof. Deferred to Section 5.5. ■

From Proposition 5.4, we infer that $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right] \approx K - \Gamma/\log(2)$ holds, but this does not imply $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right] = K - \Gamma/\log(2)$ since $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right]$ is from the exact asymptote as $\rho \to \infty$ in Proposition 5.3, while $K - \Gamma/\log(2)$ is from an approximation in Proposition 5.4. Nonetheless, it is certain that $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right] \approx K - \Gamma/\log(2)$ is very exact approximation, as shown in Example 4.1, where we used (66) to calculate the asymptote of C_{RCSI}.

5.2.1.2 Outage Capacity

We proceed to the outage capacity C_ϵ in (78) when the transmitter is ignorant of CSI. In contrast to C_{RCSI}, Proposition 5.1 shows that the channel memory L has an distinct impact on C_ϵ. Figure 15 showed that C_ϵ increases with L. What happens if

$L \to \infty$? According to Figure 15, C_ϵ seems to saturate. To investigate the impact of L, we use the high-SNR assumption.

Proposition 5.5. At high SNR, the outage capacity when $P_{\text{OUT}} = \epsilon$ can be approximated as

$$C_\epsilon \approx M \log_2 \left(\frac{\rho}{M_T} \right) + M \log_2 \left(F^{-1}_{G_{\mathcal{N}_u}}(\epsilon) \right), \tag{145}$$

where $F_{G_{\mathcal{N}_u}}(x)$ is the CDF of $G_{\mathcal{N}_u}$.

Proof. We use the approximation $\log(1+x) \approx \log(x)$ when $x \gg 1$. Then, from (61), $I_{\text{RCSI}} \approx M \log_2 \left(G_{\mathcal{N}_u} \frac{\rho}{M_T} \right)$, resulting in $P_{\text{OUT}} \approx F_{G_{\mathcal{N}_u}} \left(\frac{2^{R/M}}{\rho / M_T} \right)$. When $P_{\text{OUT}} = \epsilon$, $R \approx M \log_2 \left(\frac{\rho}{M_T} F^{-1}_{G_{\mathcal{N}_u}}(\epsilon) \right)$. By setting $R = C_\epsilon$, the proof completes. ■

Proposition 5.5 suggests that the spatial multiplexing order is M regardless of fading statistics. If the channel is uncorrelated, the spatial multiplexing order is independent of L. However, C_ϵ is still affected by L through the term $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon)$ in (145).

Example 5.2. Figure 26 illustrate $F_{G_{\mathcal{N}_u}}(x)$ on a 4×4 spatially uncorrelated channel for $L \in \{0, 1, 2, 3, 4, 5\}$. For small $F_{G_{\mathcal{N}_u}}(x) = \epsilon$ on the vertical axis, we can clearly see that $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon)$ (corresponding points on the horizontal axis) increases with L. Therefore, C_ϵ in (145) also increases with L. From Figure 26, we can also confirm that C_ϵ decreases with ϵ and becomes zero when $\epsilon = 0$, that is, $C_{\epsilon=0} = 0$. ♦

We conjecture that C_ϵ is bounded as L grows, and discuss the conjecture with an example.

Conjecture 5.1. When $L \to \infty$, $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon)$ is finite if $\epsilon < 1$. Therefore, C_ϵ is bounded.

Example 5.3. With the inequality $G_{\mathcal{N}_u} \leq A_{\mathcal{N}_u}$, we have $F_{G_{\mathcal{N}_u}}(x) \geq F_{A_{\mathcal{N}_u}}(x)$ or $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon) \leq F^{-1}_{A_{\mathcal{N}_u}}(\epsilon)$. For the uniform power profile in Definition 2.1, $F_{A_{\mathcal{N}_u}}(x)$ can be evaluated as

$$F_{A_{\mathcal{N}_u}}(x) = 1 - \frac{\Gamma(M_T M_R (L+1), M(L+1)x)}{\Gamma(M_T M_R (L+1))}. \tag{146}$$

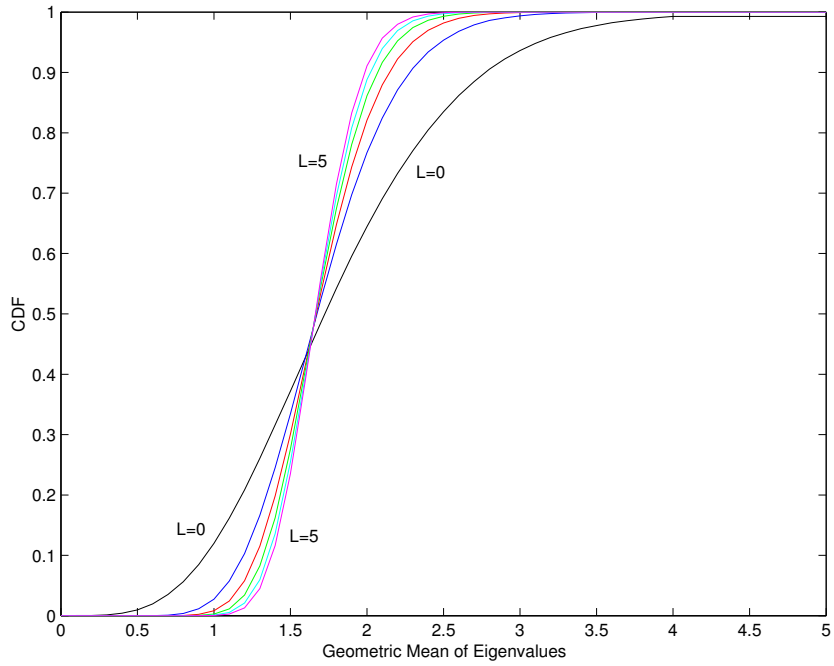

Figure 26: CDF of $G_{\mathcal{N}_u}$ on a 4×4 spatially uncorrelated channel for $L \in \{0, 1, 2, 3, 4, 5\}$.

Figure 27 plots (146) for $M_T = M_R = 4$ for $L \in \{0, 1, 2, 5, 10, 100, 1000\}$. When $L = 1000$ or lager, $F_{A_{\mathcal{N}_u}}(x)$ is nearly a step function, where there is a jump from 0 to 1 near $x = 4$ in Figure 27. Thus, $F^{-1}_{A_{\mathcal{N}_u}}(\epsilon)$ is approximately 4 for $0 < \epsilon < 1$. Therefore, $F^{-1}_{A_{\mathcal{N}_u}}(\epsilon)$ is finite as $L \to \infty$ for $0 < \epsilon < 1$. From $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon) \leq F^{-1}_{A_{\mathcal{N}_u}}(\epsilon)$, we conclude that $F^{-1}_{G_{\mathcal{N}_u}}(\epsilon)$ is also finite. ♦

5.2.2 CSI Known at Transmitter

When the transmitter knows CSI, the capacity is achieved by water-filling and its analysis is extremely difficult. High-SNR analysis gives us insights into the behaviors of capacity.

5.2.2.1 Average Capacity

We consider the average capacity with the transmitter CSI, namely, C_{TRCSI} in (85). We already saw that $C_{\text{TRCSI}} \geq C_{\text{RCSI}}$ always holds. From Figure 17, we observed that

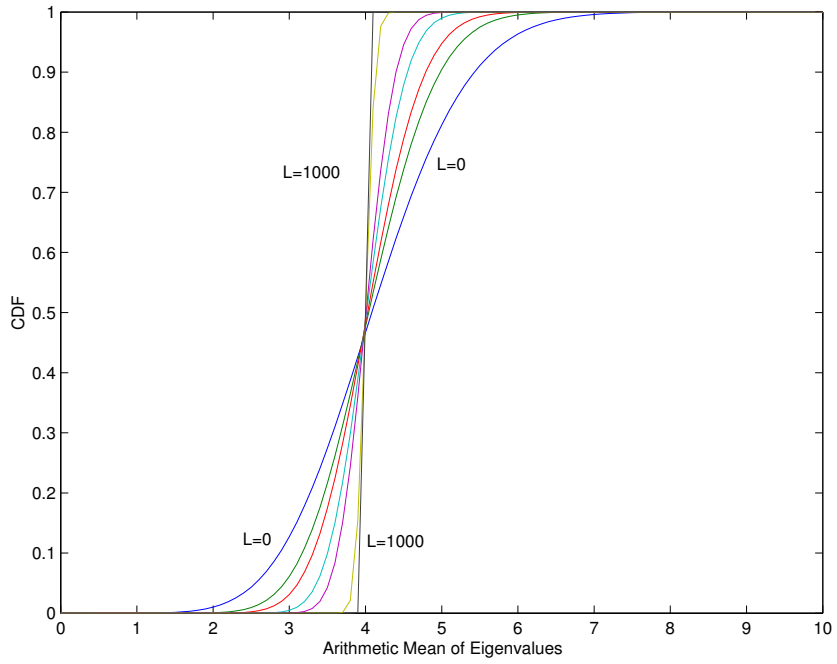

Figure 27: CDF of $A_{\mathcal{N}_u}$ on a 4×4 spatially uncorrelated channel for $L \in \{0, 1, 2, 5, 10, 100, 1000\}$.

the gap between C_{TRCSI} and C_{RCSI} is reduced as SNR grows. Thus, we expect C_{TRCSI} has the same spatial multiplexing order M as C_{RCSI}. In the following proposition, we derive the asymptote of C_{TRCSI}.

Proposition 5.6. At high SNR, the asymptote of C_{TRCSI} is given by

$$C_{\text{TRCSI}} \to M \log_2\left(\frac{\rho}{M}\right) + M\mathbb{E}[\log_2(G_{\mathcal{N}_u})]. \tag{147}$$

Proof. Deferred to Section 5.6. ■

Proposition 5.6 shows that C_{TRCSI} is equal to C_{RCSI} at high SNR except that M replaces M_T for C_{TRCSI}. Therefore, $C_{\text{TRCSI}} = C_{\text{RCSI}}$ for $M_T \leq M_R$, but $C_{\text{TRCSI}} > C_{\text{RCSI}}$ for $M_T < M_R$ at high SNR. Proposition 5.6 also implies that $\mathbb{E}[\log_2(G_{\mathcal{N}_u})]$ is independent of n, as discussed for C_{RCSI}. Therefore, C_{TRCSI} is not affected by L on spatially uncorrelated fading.

5.2.2.2 Zero-Outage Capacity

Finally and most importantly, we consider the zero-outage capacity C_0 in (164). The asymptote of C_0 is derived in [7], as follows.

Theorem 5.2. As SNR tends to infinity, C_0 asymptotically approaches:

$$C_0 \to M \log_2 \left(\frac{\rho}{M}\right) + M \log_2 \left(\frac{1}{\mathbb{E}[1/G_{\mathcal{N}_u}]}\right), \tag{148}$$

as long as $\mathbb{E}[1/G_{\mathcal{N}_u}]$ exists.

Proof. See [7]. ■

From Theorem 5.2, we can confirm that the spatial multiplexing order is M. Theorem 4.6 tells us that C_0 converges to C_{TRCSI} as $M \to \infty$, suggesting that C_0 is optimal for infinite M. However, Theorem 4.6 does not indicate how fast C_0 approaches C_{TRCSI}. To investigate the convergence speed, we use Proposition 5.6 and Theorem 5.2.

Corollary 5.1. At high SNR, $C_{\text{TRCSI}} \geq C_0$.

Proof. From Jensen's inequality [20],

$$\mathbb{E}\left[\log(X)\right] = \mathbb{E}\left[\log\left(\frac{1}{1/X}\right)\right] \geq \log\left(\frac{1}{\mathbb{E}\left[1/X\right]}\right), \tag{149}$$

since $\log(1/x)$ is a convex function. By applying (149) to $\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right]$, we have the assertion. ■

We define the SNR penalty as the additional SNR required by C_0 relative to C_{TRCSI}, which is

$$\text{SNR penalty} = \frac{\rho_0}{\rho_{\text{TRCSI}}} = 2^{\mathbb{E}[\log_2(G_{\mathcal{N}_u})]}\mathbb{E}[1/G_{\mathcal{N}_u}], \tag{150}$$

from $\frac{\rho_0}{M\mathbb{E}[1/G_{\mathcal{N}_u}]} = \frac{\rho_{\text{TRCSI}}}{M} 2^{\mathbb{E}[\log_2(G_{\mathcal{N}_u})]}$. The SNR penalty in dB corresponds to the horizontal separation between C_0 and C_{TRCSI} in a capacity-versus-log(SNR) plot. We

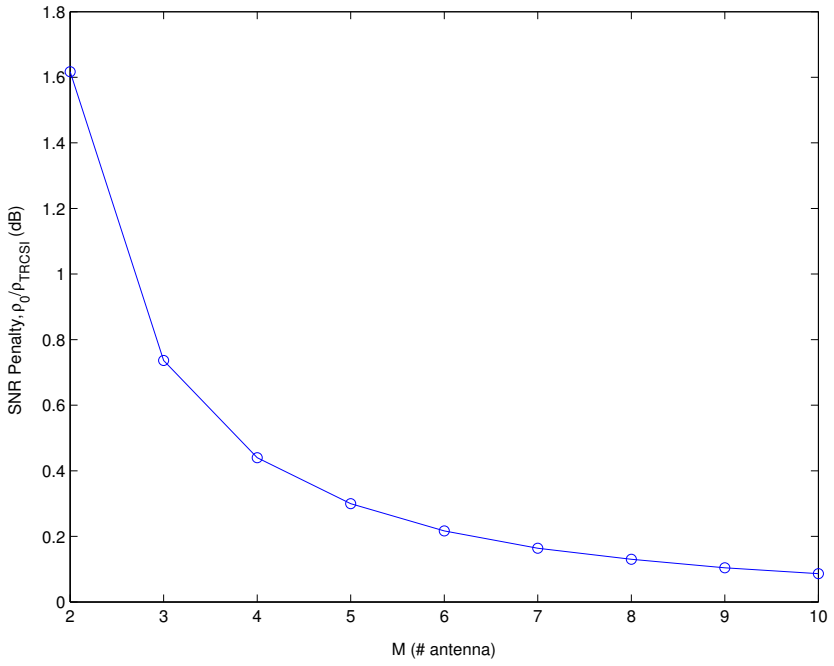

Figure 28: SNR penalty in (150) for $M \in \{2, 3, \ldots, 10\}$ on a $M \times M$ spatially uncorrelated Rayleigh flat fading.

investigate the SNR penalty as M ranges 2 from 10 in the following example in order to see the convergence speed of Theorem 4.6.

Example 5.4. We evaluate the SNR penalty as M ranges from 2 to 10 on a $M \times M$ spatially uncorrelated flat-fading channel by averaging over 10,000 independent channels. Figure 28 summarizes the SNR penalty in dB as M ranges from 2 to 10. We can see that the penalty rapidly reduces towards 0 dB as M grows. We emphasize that the SNR penalty is the worst case. The SNR penalties at finite SNR would be even smaller. ♦

Unlike $\mathbb{E}[\log_2(G_{\mathcal{N}_u})]$ in C_{TRCSI}, which is independent of n in the sense that we ignore the summation over n in $\mathbb{E}[\log_2(G_{\mathcal{N}_u})]$ such that $\mathbb{E}[\log_2(G_{\mathcal{N}_u})] = \sum_{m=1}^{M} \log_2(s_n^{(m)})$, we cannot ignore the product over n in $\mathbb{E}[1/G_{\mathcal{N}_u}]$. When the channel is spatially uncorrelated, $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ is an nondecreasing function of L, which implies that C_0 also increases with L. Since C_0 never exceeds C_{TRCSI}, we conjecture that C_0 increases

Table 3: $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ of $M \times M$ Rayleigh-fading channels with channel memory L by Monte-Carlo simulations.

L	0	1	2	3	4	5
$M = 2$	0.638	0.824	0.864	0.876	0.890	0.894
$M = 4$	1.496	1.604	1.628	1.636	1.640	1.644
$M = 6$	2.274	2.358	2.370	2.376	2.388	2.388

with L, but it is bounded as $L \to \infty$. We resort to Monte-Carlo simulations to evaluate $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ as summarized in Table 3 for $L \in \{0, 1, \ldots, 5\}$ and $M \in \{2, 4, 6\}$ by averaging over 10,000 $M \times M$ Rayleigh independent channels. Table 3 clearly shows that $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ increases with L.

However, it would still be nice to have a closed-form formula for $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ to understand the behaviors of C_0 better. We are also interested in what would happen when $L \to \infty$. Does C_0 increase without a bound when $L \to \infty$? How fast does the penalty converge to zero as $M \to \infty$? To answer these questions, we use arithmetic mean, instead of geometric mean, to derive a closed-form bound as follows

Proposition 5.7. The asymptote of C_0 in (148) is bounded by

$$C_0 \leq M \log_2 \left(\frac{\rho}{M\mathbb{E}[1/A_{\mathcal{N}_u}]} \right), \tag{151}$$

which is maximized when $\sigma_0^2 = \ldots = \sigma_L^2 = 1/(L+1)$, that is, when the power profile is uniform. For the uniform power profile, $1/\mathbb{E}[1/A_{\mathcal{N}_u}]$ in (151) can be evaluated as

$$\frac{1}{\mathbb{E}[1/A_{\mathcal{N}_u}]} = \frac{M_T M_R (L+1) - 1}{M(L+1)}, \tag{152}$$

on a spatially uncorrelated channel.

Proof. Deferred to Section 5.7. ■

From Proposition 5.7, we deduce that C_0 is bounded even when L is infinite, which can be seen from the limiting result: $\frac{1}{M\mathbb{E}[1/A_{\mathcal{N}_u}]}$ converges to $\frac{\max(M_T, M_R)}{\min(M_T, M_R)}$ as $L \to \infty$. On the other hand, when L is fixed and M tends to infinity such that

$\frac{\max(M_T,M_R)}{\min(M_T,M_R)} \to \beta \geq 1$, $\frac{1}{M\mathbb{E}[1/A_{\mathcal{N}_u}]}$ converges to β. Therefore, the upper bound of C_0 in (151) is no longer increased by L.

5.3 Proof of Proposition 5.1

First, we have from [44] the inequality:

$$\log_2 \det\left(\mathbf{I}_{M_R} + a\mathbf{H}\mathbf{H}^*\right) \geq \log_2\left(1 + a||\mathbf{H}||_F^2\right), \tag{153}$$

where $\mathbf{H}$ is an $M_R \times M_T$ matrix and $||\cdot||_F$ denotes the Frobenius norm. With this inequality, the mutual information in (61) becomes

$$\begin{aligned} I_{\text{RCSI}} &= \frac{1}{N}\sum_{n=1}^{N} \log_2 \det\left(\mathbf{I}_{M_R} + \frac{\rho}{M_T}\mathbf{H}_n\mathbf{H}_n^*\right) \\ &= \frac{1}{N}\log_2 \det\left(\mathbf{I}_{NM_R} + \frac{\rho}{M_T}\hat{\mathbf{H}}\hat{\mathbf{H}}^*\right) \\ &\geq \frac{1}{N}\log_2\left(1 + \frac{\rho}{M_T}||\hat{\mathbf{H}}||_F^2\right), \end{aligned} \tag{154}$$

for $\hat{\mathbf{H}} = \text{diag}[\mathbf{H}_1, \ldots, \mathbf{H}_N]$. Hence, the outage probability is upper-bounded by

$$\text{Prob}[I_{\text{RCSI}} < R] \leq \text{Prob}\left[\frac{||\hat{\mathbf{H}}||_F^2(L+1)}{N} < \frac{(2^{RN}-1)(L+1)}{\rho N/M_T}\right], \tag{155}$$

where we denote the upper bound $\text{Prob}\left[\frac{||\hat{\mathbf{H}}||_F^2(L+1)}{N} < \frac{(2^{RN}-1)(L+1)}{\rho N/M_T}\right]$ as $P_{\text{OUT,UB}}$. The arithmetic mean of the eigenvalues $\{s_n^{(m)}\}$ is

$$A_{\mathcal{N}_u} = \frac{1}{MN}\sum_{n=1}^{N}||\mathbf{H}_n||_F^2 = \frac{1}{MN}||\hat{\mathbf{H}}||_F^2. \tag{156}$$

From Lemma 2.2, $||\hat{\mathbf{H}}||_F^2(L+1)/N$ is a chi-square random variable with the degree of freedom $M_TM_R(L+1)$, implying that $-\lim_{\rho\to\infty}\frac{\log(P_{\text{OUT,UB}})}{\log(\rho)} = M_TM_R(L+1)$. Therefore, the diversity order is lower-bounded by $M_TM_R(L+1)$.

5.4 Proof of Proposition 5.3

We remove the index n for the notational simplicity since C_{RCSI} is independent of n in Lemma 4.1. From (62),

$$C_{\text{RCSI}} = \mathbb{E}\left[\sum_{m=1}^{M}\log_2\left(1 + \frac{\rho}{M_T}s^{(m)}\right)\right]$$

$$\geq \mathbb{E}\left[\sum_{m=1}^{M}\log_2\left(\frac{\rho}{M_T}s^{(m)}\right)\right]. \tag{157}$$

Thus, the inequality:

$$C_{\text{RCSI}} \geq M\log_2\left(\frac{\rho}{M_T}\right) + M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right], \tag{158}$$

holds for any ρ.

On the other hand, we can rewrite (62) as

$$\begin{aligned} C_{\text{RCSI}} &= \mathbb{E}\left[\sum_{m=1}^{M}\log_2\left(\frac{\rho}{M_T}s^{(m)}\left(1+\frac{M_T}{\rho s^{(m)}}\right)\right)\right] \\ &\leq M\log_2\left(\frac{\rho}{M_T}\right) + M\mathbb{E}\left[\log_2(G_{\mathcal{N}_u})\right] + M\mathbb{E}\left[\log_2\left(1+\frac{M_T}{\rho s^{(M)}}\right)\right], \end{aligned} \tag{159}$$

since $s^{(M)} \geq s^{(m)}$ for all m. We only need to prove $\mathbb{E}\left[\log_2\left(1+\frac{M_T}{\rho s^{(M)}}\right)\right]$ converges to 0 as $\rho \to \infty$. If $\mathbb{E}[1/s^{(M)}] < \infty$, the proof is simple. From Jensen's inequality,

$$\mathbb{E}\left[\log_2\left(1+\frac{M_T}{\rho s^{(M)}}\right)\right] \leq \log_2\left(1+\frac{M_T}{\rho}\mathbb{E}\left[1/s^{(M)}\right]\right). \tag{160}$$

Since $\frac{M_T}{\rho}\mathbb{E}\left[1/s^{(M)}\right] \to 0$ as $\rho \to \infty$, $\mathbb{E}\left[\log_2\left(1+\frac{M_T}{\rho s^{(M)}}\right)\right]$ also converges to zero. If $\mathbb{E}[1/s^{(M)}]$ is infinite, such as on Rayleigh fading,

$$\begin{aligned} &\mathbb{E}\left[\log_2\left(1+\frac{M_T}{\rho s^{(M)}}\right)\right] \\ =& \int_0^{\infty}\log_2\left(1+\frac{M_T}{\rho x}\right)f_{s^{(M)}}(x)dx \\ =& \int_0^{\frac{M_T}{a}}\log_2\left(1+\frac{M_T}{\rho x}\right)f_{s^{(M)}}(x)dx + \int_{\frac{M_T}{a}}^{\infty}\log_2\left(1+\frac{M_T}{\rho x}\right)f_{s^{(M)}}(x)dx \\ \leq& B\int_0^{\frac{M_T}{a}}\log_2\left(1+\frac{M_T}{\rho x}\right)dx + \log\left(1+\frac{a}{\rho}\right)\int_{\frac{M_T}{a}}^{\infty}f_{s^{(M)}}(x)dx, \end{aligned} \tag{161}$$

for any $a > 0$. The inequality in (161) is from (1) $f_{s^{(M)}} \leq B$ for $0 < x < M_T/a$ with a certain $B \geq 0$, since $f_{s^{(m)}}$ is a continuous function[1] and its integration from 0 to

[1]We consider Rayleigh fading, whose joint PDF is continuous. However, the proof can be generalized for the discontinuous case.

M_T/a is finite; (2) $\log_2\left(1+\frac{M_T}{\rho x}\right) \le \log_2\left(1+\frac{a}{\rho}\right)$ for $x \ge M_T/a$. The second term in (161) clearly converges to 0 as $\rho \to \infty$. The first term in the limit of $\rho \to \infty$ becomes

$$\lim_{\rho\to\infty} \int_0^{\frac{M_T}{a}} \log\left(1+\frac{M_T}{\rho x}\right) dx = \lim_{\rho\to\infty} M_T \left[\frac{\log\left(\frac{\rho+a}{a}\right)}{\rho} + \frac{1}{a}\log\left(1+\frac{a}{\rho}\right)\right] = 0, \quad (162)$$

where we use

$$\int \frac{\log(1+x)}{x^2} dx = \log\left(\frac{x}{1+x}\right) - \frac{\log(1+x)}{x}. \quad (163)$$

Combining two inequalities completes the proof.

5.5 Proof of Proposition 5.4

We can rewrite (66) as

$$\begin{aligned} C_{\text{RCSI}} &= \frac{e^{M_T/\rho}}{\log(2)} \sum_{k=0}^{M-1} \sum_{l=0}^{k} \sum_{i=0}^{2l} \left\{ \frac{(-1)^i (2l)!(D+i)!}{2^{2k-i} l! i! (D+l)!} \right. \\ &\times \left. \binom{2k-2l}{k-l} \binom{2l+2D}{2l-i} \left[E_1\left(\frac{M_T}{\rho}\right) + \sum_{j=1}^{D+i} E_{j+1}\left(\frac{M_T}{\rho}\right) \right] \right\}, \quad (164) \end{aligned}$$

where we use the relation:

$$E_n(x) = \frac{1}{n-1}\left\{e^{-x} - xE_{n-1}(x)\right\} \quad n > 1, \quad (165)$$

if x is real. As $x \to 0$, we can approximate $E_n(x)$ as

$$E_1(x) = -\log(x) - \sum_{k=1}^{\infty} \frac{(-x)^k}{k \cdot k!} - \Gamma \approx -\log(x) - \Gamma, \quad (166)$$

and

$$E_n(x) \approx \frac{1}{n-1} \quad n > 1, \quad (167)$$

since $\lim_{x\to 0} x\log(x) = 0$. If we use the following identity:

$$\begin{aligned} &\sum_{k=0}^{M-1} \sum_{l=0}^{k} \sum_{i=0}^{2l} \frac{(-1)^i (2l)!(D+i)!}{2^{2k-i} l! i! (D+l)!} \binom{2k-2l}{k-l} \binom{2l+2D}{2l-i} \\ = &\sum_{k=0}^{M-1} \frac{1}{4^k} \sum_{l=0}^{k} \frac{(2l)!}{l!(D+l)!} \binom{2k-2l}{k-l} \sum_{i=0}^{2l} \frac{(-1)^i (t-i)!}{2^{-i} i!} \binom{2l+2D}{2l-i} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{M-1} \frac{1}{4^k} \sum_{l=0}^{k} \frac{(2l)!}{l!(D+l)!} \binom{2k-2l}{k-l} \frac{(D+l)!}{l!} \\
&= \sum_{k=0}^{M-1} \frac{1}{4^k} \sum_{l=0}^{k} \frac{(2l)!}{l!l!} \binom{2k-2l}{k-l} \\
&= M. \qquad (168)
\end{aligned}
$$

Substituting (166) and (167) into (164) results in

$$
M \log_2 \left(\frac{\rho}{M_T} \right) + K - \Gamma / \log(2) \qquad (169)
$$

where K is defined in (144).

5.6 Proof of Proposition 5.6

The proof consists of two parts, similar to the proof of Proposition 5.3. As C_{TRCSI} is independent of n, we ignore the tone index n. First, we prove

$$
C_{\text{TRCSI}} \approx M \log_2 \left(\frac{\rho}{M} \right) + M\mathbb{E}[\log_2(G_{\mathcal{N}_u})], \qquad (170)
$$

and then proceed to the proof of (147).

From (85), we have the following inequality:

$$
\begin{aligned}
C_{\text{TRCSI}} &= \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \log_2(\lambda s^{(m)}) \right\}^+ \right] \\
&\geq \mathbb{E}\left[\sum_{m=1}^{M} \log_2(\lambda s^{(m)}) \right] \\
&= M \log_2(\lambda) + \mathbb{E}[M \log_2(G_{\mathcal{N}_u})]. \qquad (171)
\end{aligned}
$$

On the other hand, from (86), we have

$$
\begin{aligned}
\rho &= \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \lambda - \frac{1}{s^{(m)}} \right\}^+ \right] \\
&\leq \mathbb{E}\left[M \sum_{m=1}^{M} \lambda \right] \\
&= M\lambda. \qquad (172)
\end{aligned}
$$

Therefore, we have the inequality:

$$C_{\text{TRCSI}} \geq M \log_2 \left(\frac{\rho}{M} \right) + \mathbb{E}[M \log_2(G_{\mathcal{N}_u})]. \tag{173}$$

Then, we prove that

$$C \leq M \log_2 \left(\frac{\rho}{M} \right) + \mathbb{E}[M \log_2(G_{\mathcal{N}_u})] \tag{174}$$

at SNR tends to infinity, which will complete the proof. First, we note that the optimal λ in (84) is given by

$$\lambda = \frac{2^{C/M}}{G_{\mathcal{M}_{\text{all}}}}, \tag{175}$$

where $G_{\mathcal{M}_{\text{all}}}$ is the geometric mean over the index set

$$\mathcal{M}_{\text{all}} = \left\{ \text{all realizations of } s^{(m)}; \lambda s^{(m)} \geq 1 \right\}. \tag{176}$$

Clearly, $G_{\mathcal{M}_{\text{all}}}$ is deterministic, as it is averaged in a geometric sense over all realizations of $\{s^{(m)}\}$. Let $\mathcal{N}_{\text{all}} = \left\{ \text{all realizations of } s^{(m)} \right\}$. Then, the following equality holds: $\log_2 G_{\mathcal{N}_{\text{all}}} = \mathbb{E}[\log_2 G_{\mathcal{M}}]$. We define $s^{(M)} = \min \left\{ s^{(m)}, m \in \{1, 2, \ldots, M\} \right\}$. Then, the average SNR requirement is

$$\begin{aligned}
\rho &= \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \lambda - \frac{1}{s^{(m)}} \right\}^+ \right] \\
&\geq \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \lambda - \frac{1}{s^{(M)}} \right\}^+ \right] \\
&\geq \mathbb{E}\left[\sum_{m=1}^{M} \left\{ \frac{2^{C/M}}{G_{\mathcal{N}_{\text{all}}}} - \frac{1}{s^{(M)}} \right\} 1\left\{ s^{(M)} > G_{\mathcal{N}_{\text{all}}} 2^{-C/M} \right\} \right] \\
&= M 2^{C/M} \Bigg(\frac{\mathbb{E}\left[1\left\{ s^{(M)} > G_{\mathcal{N}_{\text{all}}} 2^{-C/M} \right\} \right]}{G_{\mathcal{N}_{\text{all}}}} \\
&- 2^{-C/M} \mathbb{E}\left[\frac{1}{s^{(M)}} \middle| s^{(M)} > G_{\mathcal{N}_{\text{all}}} 2^{-C/M} \right] \Bigg).
\end{aligned} \tag{177}$$

As SNR tends to infinity, so does C. As $C \to \infty$,

$$\mathbb{E}\left[1\left\{ s^{(M)} > G_{\mathcal{N}_{\text{all}}} 2^{-C/M} \right\} \right] \to 1. \tag{178}$$

Now we show that

$$2^{-C/M}\mathbb{E}\left[\frac{1}{s^{(M)}}\middle| s^{(M)} > G_{\mathcal{N}_{\text{all}}}2^{-C/M}\right] \to 0. \tag{179}$$

We assume that the PDF of $\{s^{(m)}\}$ is continuous. The marginal PDF of $s^{(M)}$ is also continuous and thus bounded: $f_{s^{(M)}}(x) \leq B$ for all x. Then,

$$\begin{aligned} & 2^{-C/M}\mathbb{E}\left[\frac{1}{s^{(m)}}\middle| s^{(m)} > G_{\mathcal{N}_{\text{all}}}2^{-C/M}\right] \\ = & \int_{x=G_{\mathcal{N}_{\text{all}}}2^{-C/M}}^{\infty} \frac{f_{s^{(m)}}(x)}{x}dx \\ \leq & B2^{-C/M}\left[\int_{x=G_{\mathcal{N}_{\text{all}}}2^{-C/M}}^{a} \frac{dx}{x} + \int_{a}^{\infty}\frac{f(x)}{a}dx\right], \end{aligned} \tag{180}$$

for any $G_{\mathcal{N}_{\text{all}}}2^{-C/M} < a < \infty$. Clearly, $\int_a^\infty \frac{f(x)}{a}dx$ is bounded, and

$$\lim_{C\to\infty} 2^{-C/M}\log(G_{\mathcal{N}_{\text{all}}}2^{-C/M}) = 0. \tag{181}$$

If we rewrite the above inequality with respect to C, we have

$$\begin{aligned} C &\leq M\log_2\left(\frac{\rho}{M}G_{\mathcal{N}_{\text{all}}}\right) \\ &= M\log_2\left(\frac{\rho}{M}\right) + \mathbb{E}[M\log_2(G_{\mathcal{N}_u})]. \end{aligned} \tag{182}$$

Therefore, by combining two inequalities, we have the assertion in (147).

5.7 Proof of Proposition 5.7

In this section, we prove Proposition 5.7. The upper bound in (151) comes from $A_{\mathcal{N}_u} \geq G_{\mathcal{N}_u}$. To prove that the uniform power profile maximizes the upper bound, we consider the following lemma.

Lemma 5.2. Suppose that $\boldsymbol{x} = [X_0, X_1, \ldots, X_L]^T$ is an $(L+1)\times 1$ vector, whose elements are i.i.d. random variables with $\Pr[X_l > 0] = 1$. Let $Z = \boldsymbol{s}^T\boldsymbol{x}$ be a scalar, where $\boldsymbol{s} = [\sigma_0^2, \sigma_1^2, \ldots, \sigma_L^2]^T$ is an $(L+1)\times 1$ deterministic vector. Then, $\{\sigma_l^2\}$ that minimize $\mathbb{E}[1/Z]$ subject to $\sum_{l=0}^L \sigma_l^2 = 1$ are:

$$\sigma_0^2 = \sigma_1^2 = \ldots = \sigma_L^2 = \frac{1}{L+1}. \tag{183}$$

Proof. Let $f(\boldsymbol{s}) = \mathbb{E}[1/Z]$ be a function from a vector space $\mathcal{S} = \{\boldsymbol{s}; \sum_{l=0}^{L} \sigma_l^2 = 1\}$ to a real scalar value. For any $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ belonging to $\mathcal{S}$, we have the inequality:

$$\begin{aligned} \lambda f(\boldsymbol{s}_1) + (1-\lambda) f(\boldsymbol{s}_2) &\overset{(a)}{=} \mathbb{E}\left[\frac{\lambda}{\boldsymbol{s}_1^T \boldsymbol{x}} + \frac{1-\lambda}{\boldsymbol{s}_2^T \boldsymbol{x}}\right] \\ &\overset{(b)}{\geq} \mathbb{E}\left[\frac{1}{\lambda \boldsymbol{s}_1^T \boldsymbol{x} + (1-\lambda)\boldsymbol{s}_2^T \boldsymbol{x}}\right] \\ &= f(\lambda \boldsymbol{s}_1 + (1-\lambda)\boldsymbol{s}_2), \end{aligned} \tag{184}$$

if $0 \leq \lambda \leq 1$, where (a) comes from the fact that $\mathbb{E}[\cdot]$ is a linear operator. The inequality (b) is because $g(x) = x^{-1} (x > 0)$ is convex:

$$\lambda g(\boldsymbol{s}_1^T \boldsymbol{x}) + (1-\lambda) g(\boldsymbol{s}_2^T \boldsymbol{x}) \geq g(\lambda \boldsymbol{s}_1^T \boldsymbol{x} + (1-\lambda)\boldsymbol{s}_2^T \boldsymbol{x}), \tag{185}$$

for $0 \leq \lambda \leq 1$. Therefore, from (184), $f(\boldsymbol{s})$ is convex.

Define an $(L+1) \times (L+1)$ permutation matrix:

$$\mathbf{P} = \begin{bmatrix} \mathbf{0}_{L\times 1}^T & 1 \\ \mathbf{I}_L & \mathbf{0}_{L\times 1} \end{bmatrix}, \tag{186}$$

where $\mathbf{0}_{L\times 1}$ denotes an $N \times 1$ all-zero vector. Note that $\mathbf{P}^0 = \mathbf{P}^{L+1} = \mathbf{I}_{L+1}$. Then we have:

$$\begin{aligned} f(\boldsymbol{s}) &\overset{(a)}{=} \frac{1}{L+1} \sum_{l=0}^{L} f(\mathbf{P}^l \boldsymbol{s}) \\ &\overset{(b)}{\geq} f\left(\frac{1}{L+1} \sum_{l=0}^{L} \mathbf{P}^l \boldsymbol{s}\right) \end{aligned}$$

where (a) is true because $\{X_l\}$ are i.i.d. and thus the order of $\{\sigma_l^2\}$ does not matter, that is, $f(\boldsymbol{s}) = f(\mathbf{P}^l \boldsymbol{s})$ for any l, and where (b) comes from Jensen's inequality [20] since $f(\boldsymbol{s})$ is convex. The equality in (b) holds if and only if $\{\mathbf{P}^l \boldsymbol{s}\}$ are constant. Therefore, from $\sum_{l=0}^{L} \sigma_l^2 = 1$, we prove that (183) minimizes $\mathbb{E}[1/Z]$. ■

From (187), the arithmetic mean is $A_{\mathcal{N}_u} = \frac{1}{M} \sum_{l=0}^{L} \sigma_l^2 ||\mathbf{W}_l||_F^2$ if the channel is uncorrelated. Then, by letting $X_l = ||\mathbf{W}_l||_F^2 / M$, we have the assertion that the

uniform power profile maximizes the upper bound, which completes the first part of the proof.

Now we derive (152). From Lemma 2.2, we have

$$A_{\mathcal{N}_u} = \frac{1}{M}\sum_{l=0}^{L} ||\mathbf{G}_l||_F^2 = \frac{1}{M(L+1)}\sum_{l=0}^{L} ||\mathbf{W}_l||_F^2, \tag{187}$$

by assuming uniform power profile, where the elements of $\mathbf{W}_l$ are i.i.d. $\mathcal{CN}(0,1)$. Then, $M(L+1)A_{\mathcal{N}_u}$ is a chi-square random variable with a degree of freedom $M_T M_R(L+1)$. We can evaluate $\mathbb{E}[1/A_{\mathcal{N}_u}]$ with the PDF in Corollary 2.3, which results in (152).

CHAPTER 6

APPROACHING ZERO-OUTAGE CAPACITY WITHOUT WATER-FILLING

We consider the power-allocation problem for the zero-outage capacity in a closed-loop MIMO-OFDM system in Figure 3. The optimal allocation in Theorem 4.4 is based on water-filling over all scalar channels over space (antennas) and frequency (tones). The computational complexity required by water-filling is high, especially when the number of scalar channels is large. In this chapter, we propose simpler allocation strategies. We will show that the proposed strategies have much less complexity, but performs nearly as well as the optimal allocation. The contributions of this chapter will be

- To propose reduced-complexity allocation strategies and to derive their optimal power allocation.
- To prove that the performance penalties of the proposed strategies relative to the optimal allocation, in terms of the zero-outage capacities they achieve, approach zero as the number of antennas tends to infinity.
- To show by the high-SNR analysis that the penalties of the proposed strategies are small for a moderate number of antennas.

6.1 Problem Statement

We assume that CSI is known at the transmitter and a MIMO-OFDM channel in Figure 1 converts into a bank of MN scalar channels in Figure 3. Recall from Theorem 4.4 that so as to achieve the zero-outage capacity in (98), the transmitter decides

the power allocation $\{e_n^{(m)}\}$ by solving the following problem:

$$\begin{array}{l} \text{Minimize } \bar{E} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} e_n^{(m)}\right] \\ \text{Subject to } \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\log_2\left(1+\frac{e_n^{(m)}s_n^{(m)}}{N_0}\right) = R \end{array}, \tag{188}$$

where SNR is a function of the rate, namely $\rho = \bar{E}/N_0 = f(R)$. By inverting this function, $C_{0,\text{OPT}} = f^{-1}(\rho)$ is the zero-outage capacity at SNR ρ.[1] The optimal solution for $\{e_n^{(m)}\}$ is presented in Theorem 4.4, which requires water-filling considering all MN scalar channels. In MIMO-OFDM, MN is often large, and the complexity required by water-filling can be very high. In the following, we propose the *frequency-uniform-spectral-efficiency* and *fixed-rate* allocations by applying stricter constraints than the constraint in (188).

6.2 FUSE Allocation

First, we introduce the frequency-uniform-spectral-efficiency (FUSE) constraint. Instead of the constraint in (188), the FUSE constraint forces M spatial channels $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ at each n to achieve the same spectral efficiency, so that the allocation problem becomes

$$\begin{array}{l} \text{Minimize } \bar{E} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} e_n^{(m)}\right] \\ \text{Subject to } \sum_{m=1}^{M}\log_2\left(1+\frac{s_n^{(m)}e_n^{(m)}}{N_0}\right) = R \quad \text{for all } n \end{array}. \tag{189}$$

The optimal solution to (189) is as follows.

Proposition 6.1 (FUSE Allocation). The power allocation that solves (189) is $e_n^{(m)} = N_0\{\mu_n - 1/s_n^{(m)}\}^+$, where

$$\mu_n = \frac{2^{R/|\mathcal{M}_n|}}{G_{\mathcal{M}_n}} \tag{190}$$

ensures that $\sum_{m=1}^{M}\log_2\left(\mu_n s_n^{(m)}\right) = R$ is satisfied for all n, and where

$$\mathcal{M}_n = \{(m,n); \mu_n s_n^{(m)} \geq 1\} \tag{191}$$

[1] For emphasis, we denote $C_{0,\text{OPT}}$, instead of C_0, as the zero-outage capacity with the optimal allocation.

is the index set for the used channels of the nth tone, such that $|\mathcal{M}_n| \leq M$.

Proof. From (189), the objective function can be rewritten as

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} e_n^{(m)}\right] = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[\sum_{m=1}^{M} e_n^{(m)}\right]. \tag{192}$$

Thus, the problem reduces to N independent smaller problems:

$$\begin{array}{l} \text{Minimize } \mathbb{E}\left[\sum_{m=1}^{M} e_n^{(m)}\right] \\ \text{Subject to } \sum_{m=1}^{M} \log_2\left(1+\frac{s_n^{(m)} e_n^{(m)}}{N_0}\right) = R \end{array} \quad \text{for all } n. \tag{193}$$

For each n, the optimal $\{e_n^{(m)}\}$ is in the same form of Theorem 4.4, where μ_n is the water-level parameter and $\mathcal{M}_n$ is the index set for the used channels. ■

Proposition 6.1 suggests that the FUSE constraint breaks the original power-allocation problem in (188) into a set of smaller problems in (193). Each problem in (193) is also solved by water-filling, but the scope of water-filling reduces to M scalar channels $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ for each n. Thus, the FUSE allocation performs water-filling over M spatial channels independently N times instead of performing water-filling over MN channels. In MIMO-OFDM, M is much smaller than MN, and hence the complexity can be considerably reduced.

The complexity reduction is not free, incurring a penalty in the zero-outage capacity. In other words, the FUSE allocation requires larger average energy than the optimal allocation, that is, $\bar{E}_{\text{FUSE}} \geq \bar{E}_{\text{OPT}}$ always holds. In terms of the zero-outage capacity, the zero-outage capacity with FUSE constraint, $C_{0,\text{FUSE}}$, is always less than or equal to $C_{0,\text{OPT}}$.

Note that the energy $E_n = \sum_{m=1}^{M}\{\mu_n - 1/s_n^{(m)}\}^+$ required by each tone is not constant across tones for a certain realization of $\{s_n^{(m)}\}$. In the average sense, however, the average energy $\mathbb{E}[E_n]$ required by each tone is uniform since the statistics of $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ are identical for all n from Lemma 2.1. Therefore, $C_{0,\text{FUSE}}$ is equal to the zero-outage capacity of the memoryless MIMO channel $\{s_n^{(m)}; m =$

$1, 2, \ldots, M\}$ for any n. If the channel is spatially uncorrelated, each $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ is statistically identical to a flat-fading ($L = 0$) channel, which yields the following remark:

Remark 6.1. On a spatially uncorrelated channel, $C_{0,\mathrm{FUSE}} = C_{0,\mathrm{OPT},L=0}$ since each channel matrix $\mathbf{H}_n$ is statistically equivalent to a flat-fading MIMO channel from Lemma 2.1. Thus, $C_{0,\mathrm{FUSE}}$ is constant independent of L ♦

From Remark 6.1, we can confirm the inequality $C_{0,\mathrm{FUSE}} \leq C_{0,\mathrm{OPT}}$ since $C_{0,\mathrm{OPT}}$ increases with L, while $C_{0,\mathrm{FUSE}}$ remains unchanged.

On a Rayleigh-fading SISO channel ($M_T = M_R = 1$), $C_{0,\mathrm{FUSE}}$ is zero regardless of L since $C_{0,\mathrm{OPT},L=0}$ is zero, meaning that the FUSE allocation has no benefit from frequency selectivity. In contrast, $C_{0,\mathrm{OPT}}$ can be nonzero for $L > 0$ owing to the diversity from frequency selectivity [55]. Hence, the FUSE allocation is grossly suboptimal to the optimal allocation in a SISO channel, incurring an infinite SNR penalty in the sense that an infinite SNR is required to achieve a positive zero-outage capacity.

With multiple antennas at the transmitter and receiver, the FUSE allocation benefits from the spatial diversity and $C_{0,\mathrm{FUSE}}$ increases to a nonzero value. Intuitively, as the number of antennas grows, the spatial diversity becomes more dominant than the frequency diversity, and $C_{0,\mathrm{FUSE}}$ ($= C_{0,\mathrm{OPT},L=0}$) would be close to $C_{0,\mathrm{OPT}}$. As an extreme case, when $M = \min(M_T, M_R)$ tends to infinity, we expect that the benefit from the frequency diversity is negligible, and $C_{0,\mathrm{FUSE}}$ would be as large as $C_{0,\mathrm{OPT}}$. In the following proposition, we prove that the gap between $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{OPT}}$ disappears for large M.

Proposition 6.2. On a spatially uncorrelated channel, the power allocation with the FUSE constraint is asymptotically optimal as the number of antennas gets large, in the sense that $C_{0,\mathrm{FUSE}} \to C_0$ as $M \to \infty$.

Proof. The proof is based on the fact that the power allocation depends on the empirical distribution in (22). In other words, if two systems have the same empirical distribution for all channel realizations, they achieve the same zero-outage capacity. Consider two index sets: the universe index set $\mathcal{N}_u$ in (14) and the spatial index set

$$\mathcal{N}_s = \{(m,n); m = 1, 2, \ldots, M \text{ for specific } n\}, \tag{194}$$

where $\Psi_{\mathcal{N}_u}(x)$ and $\Psi_{\mathcal{N}_s}(x)$ correspond to $C_{0,\text{OPT}}$ and $C_{0,\text{FUSE}}$, respectively. When M is finite, the empirical distribution functions for $\mathcal{N}_u$ and $\mathcal{N}_s$ are different, $\Psi_{\mathcal{N}_u}(x) \neq \Psi_{\mathcal{N}_s}(x)$. However, their expected values are equal [42], such that

$$\mathbb{E}[\Psi_{\mathcal{N}_u}(x)] = \mathbb{E}[\Psi_{\mathcal{N}_s}(x)]. \tag{195}$$

As $M \to \infty$, both $\Psi_{\mathcal{N}_u}(x)$ and $\Psi_{\mathcal{N}_s}(x)$ converge to non-random limits, as shown in Theorem 2.2, implying that $\Psi_{\mathcal{N}_u}(x) = \mathbb{E}[\Psi_{\mathcal{N}_u}(x)]$ and $\Psi_{\mathcal{N}_s}(x) = \mathbb{E}[\Psi_{\mathcal{N}_s}(x)]$. Therefore, from (195), we have $\Psi_{\mathcal{N}_u}(x) = \Psi_{\mathcal{N}_s}(x)$ at infinite M, and both $\Psi_{\mathcal{N}_u}(x)$ and $\Psi_{\mathcal{N}_s}(x)$ achieve the same zero-outage capacity. ■

Proposition 6.2 suggests that the penalty by FUSE constraint converges to zero as M grows. It is an encouraging result to justify the use of FUSE constraint in power allocation, but we are more interested in its performance at a finite M. We will show that the convergence is fast by high-SNR analysis (Section 6.4) and experimental results (Section 6.5).

6.3 Fixed-Rate Allocation

Another simplifying constraint for the allocation problem is the fixed-rate (FIX) constraint, where the achievable rate of each scalar channel, $r_n^{(m)} = \log_2(1 + s_n^{(m)} e_n^{(m)}/N_0)$, is fixed independent of $\{s_n^{(m)}\}$, namely a non-random constant. With the FIX constraint, the problem in (188) reduces to a simpler problem:

$$\begin{aligned} &\text{Minimize } \bar{E} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{2^{r_n^{(m)}}-1}{s_n^{(m)}/N_0}\right], \\ &\text{Subject to } \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = R \end{aligned} \tag{196}$$

where $\{r_n^{(m)}\}$ are deterministic and independent of $\{s_n^{(m)}\}$.

Once $\{r_n^{(m)}\}$ are determined, power allocation can be simply calculated from a closed-form formula, $e_n^{(m)} = N_0(2^{r_n^{(m)}} - 1)/s_n^{(m)}$, and thus no water-filling is needed. We emphasize that fixing $\{r_n^{(m)}\}$ does not mean an equal allocation, where $\{r_n^{(m)}\}$ are uniform for m and n, which would be a bad allocation, resulting in a very large average energy $\bar{E}$. For best results, the choice of $\{r_n^{(m)}\}$ must be optimized to the anticipated channel statistics so as to minimize the average energy $\bar{E}$.

First, we define the number of *available* channels for the FIX constraint.

Definition 6.1. Let $\tilde{M}$ be the largest integer satisfying $\mathbb{E}[1/s_{\tilde{M},n}] < \infty$ for each n, which is the number of scalar channels we can use for power allocation at each tone.

♦

If a nonzero rate were allocated to (m, n) with $\mathbb{E}[1/s_n^{(m)}] \to \infty$, the average energy requirement for (m, n) would be infinite. For this reason, we avoid using such a scalar channel. From Lemma 2.1, $\tilde{M}$ is identical for all n. For example, when the channel is spatially uncorrelated flat fading and square ($M_T = M_R = M$), $\tilde{M}$ is at most $M - 1$ since $s_n^{(M)}$ is exponential distributed and $\mathbb{E}[1/s_n^{(M)}]$ diverges [23]. We conjecture, from computer simulations, that $\tilde{M} = M - 1$ for any $M \times M$ channel though we can prove this only for small M by explicitly calculating $\mathbb{E}[1/s_n^{(m)}]$ from the marginal distribution of $\{s_n^{(m)}\}$ [65]. When the channel is non-square ($M_T \neq M_R$), we have $\tilde{M} = M = \min(M_T, M_R)$ as deduced from [23].

The optimal choice of $\{r_n^{(m)}\}$ is also given by water-filling, as illustrated in the following proposition.

Proposition 6.3 (FIX Allocation). The optimal choice of $\{r_n^{(m)}\}$ that minimizes $\bar{E}$ is

$$r_n^{(m)} = \begin{cases} \left\{\log_2\left(\frac{\nu}{\mathbb{E}[1/s_n^{(m)}]}\right)\right\}^+ & 1 \le m \le \tilde{M} \\ 0 & \text{otherwise} \end{cases}, \tag{197}$$

where

$$\nu = \frac{2^{RN/\left|\mathcal{M}_{\text{FIX}}\right|}}{\Gamma_{\mathcal{M}_{\text{FIX}}}} \tag{198}$$

ensures that $\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = R$ is satisfied, and where

$$\mathcal{M}_{\text{FIX}} = \{(m,n); \nu/\mathbb{E}[1/s_n^{(m)}] \geq 1\} \tag{199}$$

is the index set for used channels out of $\tilde{M}$ channels for each n. In (198) and (199), $\Gamma_{\mathcal{M}_{\text{FIX}}}$ denotes the geometric mean of $1/\mathbb{E}[1/s_n^{(m)}]$ over $\mathcal{M}_{\text{FIX}}$, namely,

$$\Gamma_{\mathcal{M}_{\text{FIX}}} = \left(\prod_{(m,n)\in\mathcal{M}_{\text{FIX}}} 1/\mathbb{E}[1/s_n^{(m)}]\right)^{1/\left|\mathcal{M}_{\text{FIX}}\right|}. \tag{200}$$

The optimal power allocation can be calculated from $e_n^{(m)} = N_0(2^{r_n^{(m)}} - 1)/s_n^{(m)}$.

Proof. Since $\{r_n^{(m)}\}$ are independent of $\{s_n^{(m)}\}$, the objective function in (196) becomes

$$\frac{\bar{E}}{N_0} = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{2^{r_n^{(m)}}-1}{s_n^{(m)}}\right] = \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{\tilde{M}}\frac{(2^{r_n^{(m)}}-1)}{u_n^{(m)}}, \tag{201}$$

where $u_n^{(m)} = 1/\mathbb{E}[1/s_n^{(m)}]$. In (201), $\{u_n^{(m)}\}$ act as the squared channel gains, and each tone has $\tilde{M}$ scalar channels. Since $\{u_n^{(1)}, \ldots, u_n^{(\tilde{M})}\}$ is identical for all n, the problem reduces to water-filling over $\tilde{M}$ scalar channels, such that ν is the water-level parameter and $\mathcal{M}_{\text{FIX}}$ is the index set for used channels. The geometric mean $\Gamma_{\mathcal{M}_{\text{FIX}}}$ replaces $G_{\mathcal{M}_n}$ in this case. ■

From (197), we can confirm that $\{r_n^{(m)}\}$ are deterministic, obtained from performing water-filling over deterministic channels, $\{1/\mathbb{E}[1/s_n^{(m)}]\}$. Thus, we only need $\{1/\mathbb{E}[1/s_n^{(m)}]\}$, not all statistics of $\{s_n^{(m)}\}$, to decide $\{r_n^{(m)}\}$. Once $\{r_n^{(m)}\}$ are predetermined according to Proposition 6.3, power allocation is trivial, done by a closed-form formula:

$$e_n^{(m)} = \frac{2^{r_n^{(m)}}-1}{s_n^{(m)}/N_0} = \begin{cases} N_0\left\{\frac{\nu}{s_n^{(m)}\mathbb{E}[1/s_n^{(m)}]} - \frac{1}{s_n^{(m)}}\right\}^+ & 1 \leq m \leq \tilde{M} \\ 0 & \text{otherwise} \end{cases}, \tag{202}$$

for each realization of $\{s_n^{(m)}\}$.

From Lemma 2.1, the effective channels at each tone $\{1/\mathbb{E}[1/s_n^{(m)}]; m = 1, 2, \ldots, M\}$ are identical for all n. Thus, the FIX allocation has no dependency on n. We already mentioned that $\tilde{M}$ is constant for all n. In Proposition 6.3, ν and $\mathcal{M}_{\text{FIX}}$ are also independent of n. As $\{1/\mathbb{E}[1/s_n^{(m)}]; m = 1, 2, \ldots, M\}$ are deterministic, so are ν and $\mathcal{M}_{\text{FIX}}$.

Corollary 6.1. The FIX constraint inherently implies the FUSE constraint in the sense that $\{r_n^{(m)}\}$ in Proposition 6.3 satisfy $\sum_{m=1}^{M} r_n^{(m)} = R$.

Proof. Straightforward from the proof of Proposition 6.3. ■

Complexity reduction is remarkable with the FIX constraint, but fixing $\{r_n^{(m)}\}$ could incur a significant penalty. Let $C_{0,\text{FIX}}$ denote the zero-outage capacity with the FIX constraint. Since the FIX constraint is stricter than the FUSE constraint, we have $C_{0,\text{OPT}} \geq C_{0,\text{FUSE}} \geq C_{0,\text{FIX}}$.

On a spatially uncorrelated SISO channel ($M = 1$), we have $\tilde{M} = 0$ for the FIX allocation, that is, no channel is available, since $1/\mathbb{E}[1/s_n^{(1)}] = 0$. Thus, $C_{0,\text{FIX}}$ is always zero. With multiple antennas, similar to $C_{0,\text{FUSE}}$, $C_{0,\text{FIX}}$ becomes nonzero as $\tilde{M} > 0$. However, the number of available channels ($\tilde{M}$) can be smaller than the maximum multiplexing order (M), such as on a $M \times M$ channel where $\tilde{M} \leq M - 1$, incurring a significant penalty due to the smaller spatial multiplexing order. As M grows, however, we have the following proposition.

Proposition 6.4. On a spatially uncorrelated channel, the power allocation with the FIX constraint is asymptotically optimal as the number of antennas gets large, in the sense that $C_{0,\text{FIX}} \to C_0$ as $M \to \infty$.

Proof. Let $\theta_n^{(m)} = s_n^{(m)}/M$ and let $\mathcal{N}_s$ be the spatial index set in (194). From Proposition 6.3, $C_{0,\text{FIX}}$ depends on the empirical distribution of $1/\mathbb{E}[1/\theta_n^{(m)}]$:

$$V_{\mathcal{N}_s}(x) = \frac{1}{|\mathcal{N}_s|} \sum_{m=1}^{M} 1\left\{1/\mathbb{E}[1/\theta_n^{(m)}] \leq x\right\}, \tag{203}$$

if including $1/\mathbb{E}[1/\theta_n^{(m)}] = 0$, which is identical for all n. The goal is to show that $V_{\mathcal{N}_s}(x)$ is equal to $\Psi_{\mathcal{N}_s}(x)$, the empirical distribution for $\{s_n^{(m)}\}$ in (22), as $M \to \infty$.

When $M \to \infty$, $\Psi_{\mathcal{N}_s}(x)$ becomes non-random, and thus $1\{\theta_n^{(m)} \leq x\} = \mathbb{E}[1\{\theta_n^{(m)} \leq x\}] = \text{Prob}[\theta_n^{(m)} \leq x]$, implying that the distribution function of $\theta_n^{(m)}$ is a delta function at $\theta_n^{(m)}$. Therefore, $1/\mathbb{E}[1/\theta_n^{(m)}] = \theta_n^{(m)}$. If $1/\mathbb{E}[1/\theta_n^{(m)}] = 0$, this corresponds to $\theta_n^{(m)} = 0$. Therefore, by substituting $1/\mathbb{E}[1/\theta_n^{(m)}] = \theta_n^{(m)}$ into (22), we obtain $V_{\mathcal{N}_s}(x) = \Psi_{\mathcal{N}_s}(x)$ for infinite M. Therefore, $C_{0,\text{FIX}}$ converges to $C_{0,\text{FUSE}}$ as $M \to \infty$ sine $V_{\mathcal{N}_s}(x)$ and $\Psi_{\mathcal{N}_s}(x)$ account for $C_{0,\text{FIX}}$ and $C_{0,\text{FUSE}}$, respectively. Proposition 6.2, we deduce that $C_{0,\text{FIX}}$ is asymptotically identical to $C_{0,\text{OPT}}$ as $M \to \infty$. ■

We showed in Proposition 6.2 that water-filling over spatial channels is sufficient to approach $C_{0,\text{OPT}}$. Furthermore, Proposition 6.4 illustrates that a nonadaptive allocation for $\{r_n^{(m)}\}$, if carefully chosen as described in Proposition 6.3, also achieves $C_{0,\text{OPT}}$. In other words, no water-filling is necessary to approach $C_{0,\text{OPT}}$ in MIMO. We remark that the FIX allocation must be matched to channel statistics, while the FUSE allocation only needs to know the current channel status.

6.4 Asymptotic Behaviors

In this section, we conduct high-SNR analysis on the FUSE and FIX allocations. There are two goals: (1) we can approximately quantify the SNR penalty of $C_{0,\text{FUSE}}$ relative to $C_{0,\text{OPT}}$ at high SNR, and show that the penalty quickly approach 0 dB (no penalty) as M grows; (2) we derive the asymptote of $C_{0,\text{FIX}}$ and discuss the ill effects of deficient $\tilde{M}$.

6.4.1 FUSE Allocation

Recall from Theorem 5.2, that the asymptote of $C_{0,\mathrm{OPT}}$ is

$$C_{0,\mathrm{OPT}} \to M \log_2 \left(\frac{\rho}{M} \right) + M \log_2 \left(\frac{1}{\mathbb{E}[1/G_{\mathcal{N}_u}]} \right), \tag{204}$$

at high SNR, and from Proposition 5.7 that the upper bound of $C_{0,\mathrm{OPT}}$ is

$$C_{0,\mathrm{OPT}} \leq M \log_2 \left(\frac{\rho}{M\mathbb{E}[1/A_{\mathcal{N}_u}]} \right), \tag{205}$$

where $1/\mathbb{E}[1/A_{\mathcal{N}_u}]$ can be evaluated for the uniform power profile as

$$\frac{1}{\mathbb{E}[1/A_{\mathcal{N}_u}]} = \frac{M_T M_R (L+1) - 1}{M(L+1)}, \tag{206}$$

on a spatially uncorrelated channel. Since $C_{0,\mathrm{FUSE}}$ is equal to the zero-outage capacity of the spatial channels $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$, (204) can be modified to obtain the asymptote of $C_{0,\mathrm{FUSE}}$.

Corollary 6.2. As SNR tends to infinity, $C_{0,\mathrm{FUSE}}$ asymptotically approaches

$$C_{0,\mathrm{FUSE}} \to M \log_2 \left(\frac{\rho}{M\mathbb{E}[1/G_{\mathcal{N}_s}]} \right) \leq M \log_2 \left(\frac{\rho}{M\mathbb{E}[1/A_{\mathcal{N}_s}]} \right), \tag{207}$$

where $\mathcal{N}_s$ is the spatial index set in (194).

Proof. The proof is straightforward from Proposition 6.1. ■

Corollary 6.2 shows that the spatial multiplexing order is also M. To measure the penalty by the FUSE constraint, we define the SNR penalty as

$$\text{SNR penalty } = \frac{1/\mathbb{E}[1/G_{\mathcal{N}_u}]}{1/\mathbb{E}[1/G_{\mathcal{N}_s}]}, \tag{208}$$

which accounts for the additional SNR required by the FUSE constraint at high SNR, relative to the optimal allocation. We can show that the SNR penalty is always greater than or equal to unity since

$$\mathbb{E}\left[\frac{1}{G_{\mathcal{N}_u}} \right] = \mathbb{E}\left[\left(\prod_{n=1}^{N} \left(\prod_{m=1}^{M} \frac{1}{s_n^{(m)}} \right)^{\frac{1}{M}} \right)^{\frac{1}{N}} \right] \leq \mathbb{E}\left[\frac{1}{N} \sum_{n=1}^{N} \left(\prod_{m=1}^{M} \frac{1}{s_n^{(m)}} \right)^{\frac{1}{M}} \right] = \mathbb{E}\left[\frac{1}{G_{\mathcal{N}_s}} \right], \tag{209}$$

where the inequality in the above equation comes from the inequality of arithmetic and geometric means of $\left\{(\prod_{m=1}^{M} 1/s_n^{(m)})^{1/M}; m = 1, 2, \ldots, M\right\}$.

If the channel is spatially uncorrelated, $\{s_n^{(m)}; m = 1, 2, \ldots, M\}$ is statistically identical to a flat-fading MIMO channel, so that $G_{\mathcal{N}_s}$ is equal to $G_{\mathcal{N}_{u,L=0}}$. The term $\frac{1}{\mathbb{E}[1/A_{\mathcal{N}_s}]}$ in (207) can be evaluated as $\frac{1}{\mathbb{E}[1/A_{\mathcal{N}_s}]} = (M_T M_R - 1)/M$. Then, we can explicitly evaluate the approximate SNR penalty as

$$\text{SNR penalty} \approx \frac{1/\mathbb{E}[1/A_{\mathcal{N}_u}]}{1/\mathbb{E}[1/A_{\mathcal{N}_{u,L=0}}]} = \frac{M_T M_R(L+1) - 1}{(M_T M_R - 1)(L+1)}, \tag{210}$$

for the uniform power profile. We are particularly interested in the uniform power profile because it maximizes the approximate SNR penalty, namely, (210) is the worst case penalty of the FUSE allocation.

The approximation in (210) is useful to see how M and L affect the SNR penalty of the FUSE allocation. We can confirm that the approximate SNR penalty in (210) becomes 0 dB (no penalty) when either $L = 0$ or $M \to \infty$. The result for $M \to \infty$ agrees with Proposition 6.2. When $L \to \infty$, it converges to $M_T M_R/(M_T M_R - 1)$, that is, the penalty is bounded. We take examples to verify the validness of the approximate SNR penalty and to investigate the impact of M and L on the SNR penalty.

Example 6.1. If the channel is square ($M_T = M_R = M$), Figure 29 illustrates the approximate SNR penalty in (210) for $L = 4$ as M grows from 1 to 10, in which we observe that the penalty rapidly decreases toward 0 dB. The actual penalty in (208) is also plotted in Figure 29, slightly larger than (210), which also decreases toward 0 dB as M grows. Figure 29 suggests that the convergence in Proposition 6.2 is fast.

♦

Example 6.2. Figure 30 plots the approximate SNR penalty in (210) for $M_T = M_R = M \in \{2, 3, \ldots, 10\}$ when $L \in \{0, 1, 2, 5, 10, 100, 1000\}$. Clearly, the approximate SNR

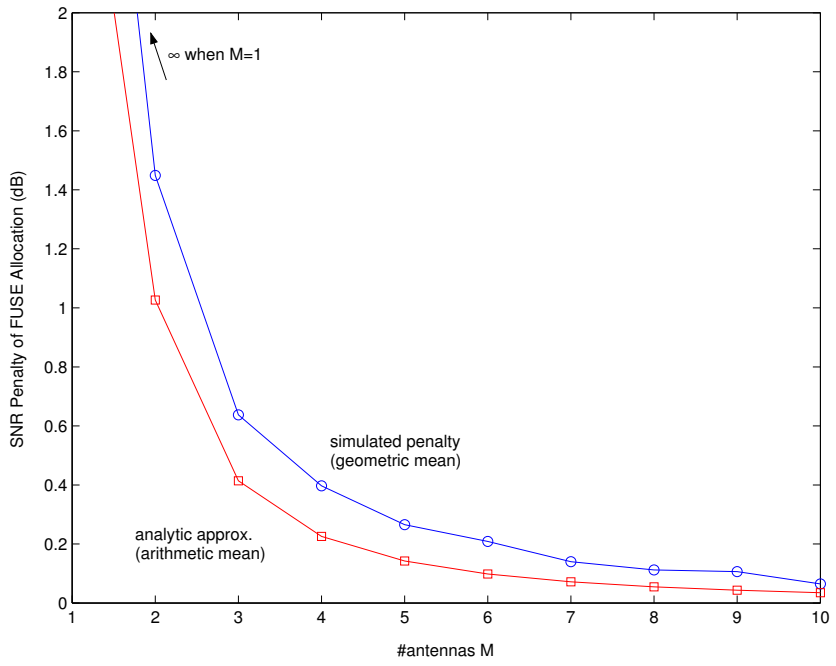

Figure 29: SNR penalty of the FUSE allocation and its approximation as M grows at high SNR.

penalty converges as L grows. The convergence of Proposition 6.2 is slower as L gets larger. ♦

Example 6.3. In this example, we consider three cases: (1) $M_T = 1$ and $M_R = k$; (2) $M_T = M_R = k$; and (3) $M_T = k$ and $M_R = 2k$ when k ranges from 2 to 10. Figure 31 plots the approximate SNR penalty for the three cases. Clearly, the convergence is slowest for the case (1) and fastest for the case (3). ♦

6.4.2 FIX Allocation

Now we proceed to deriving the asymptote of the FIX allocation at high SNR.

Proposition 6.5. As SNR tends to infinity,

$$C_{0,\mathrm{FIX}} \to \tilde{M} \log_2 \left(\frac{\rho}{\tilde{M}/\Gamma_{\tilde{\mathcal{N}}_u}} \right), \tag{211}$$

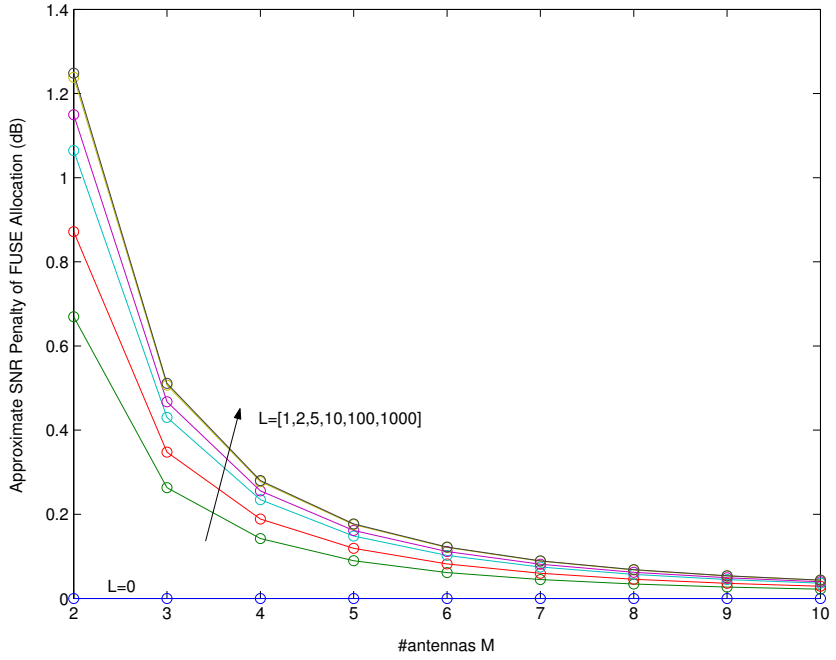

Figure 30: Approximate SNR penalty of the FUSE allocation for $L \in \{0, 1, 2, 5, 10, 100, 1000\}$.

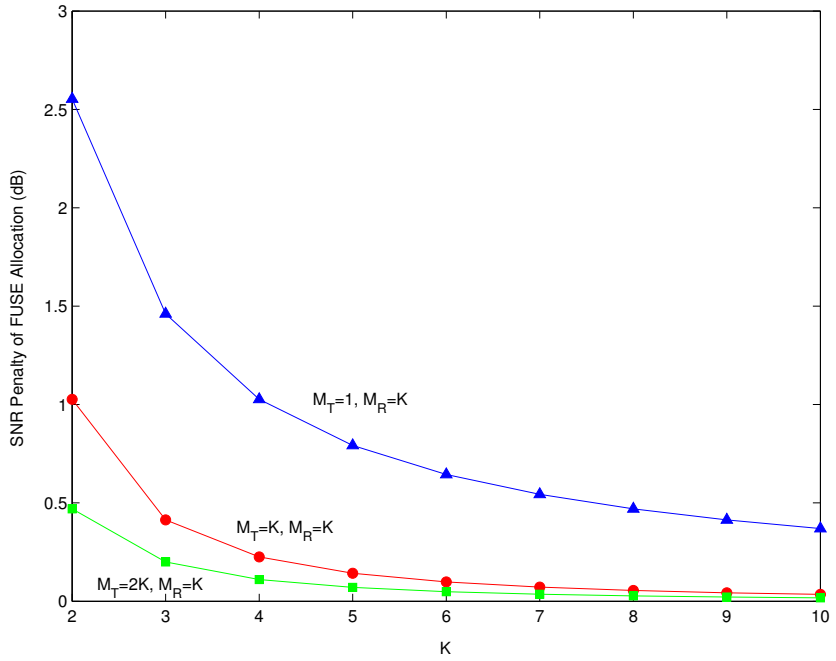

Figure 31: Approximate SNR penalty of the FUSE allocation for $(M_T = 1, M_R = k)$, $(M_T = k, M_R = k)$, and $(M_T = 2k, M_R = k)$ with $k \in \{2, 3, \ldots, 10\}$.

where $\Gamma_{\tilde{\mathcal{N}}_u}$ is the geometric mean of $1/\mathbb{E}[1/s_n^{(m)}]$ over the index set $\tilde{\mathcal{N}}_u$, and where $\tilde{\mathcal{N}}_u = \{(m,n); m = 1,2,\ldots,\tilde{M} \text{ and } n = 1,2,\ldots,N\}$ is the index set for available channels in Proposition 6.3.

Proof. Deferred to Section 6.6. ■

Proposition 6.5 shows that the spatial multiplexing order is $\tilde{M}$ rather than M. As already discussed, the deficiency in the spatial multiplexing order leads to a significant loss in $C_{0,\text{FIX}}$. As before, we define the SNR penalty as the additional SNR required by the FIX allocation to achieve the same capacity as the optimal allocation ($C_{0,\text{OPT}} = C_{0,\text{FIX}} = R$), that is,

$$\text{SNR penalty } = \frac{\rho_{\text{FIX}}}{\rho_{\text{OPT}}} = \frac{\tilde{M}}{M\Gamma_{\tilde{\mathcal{N}}_u}\mathbb{E}[1/G_{\mathcal{N}_u}]} 2^{R\left(\frac{1}{\tilde{M}}-\frac{1}{M}\right)}, \tag{212}$$

where ρ_{OPT} and ρ_{FIX} are the SNR required by the optimal and FIX allocations, respectively, to achieve R. From (212), the SNR penalty diverges if $M > \tilde{M}$ since (211) assumes SNR tends to infinity, which implies that $R \to \infty$. If $M = \tilde{M}$, the SNR penalty is finite and depends only on the term $\frac{\tilde{M}}{M\Gamma_{\tilde{\mathcal{N}}_u}\mathbb{E}[1/G_{\mathcal{N}_u}]}$. As L grows, $1/\mathbb{E}[1/G_{\mathcal{N}_u}]$ also increases, but $\Gamma_{\tilde{\mathcal{N}}_u}$ is unchanged on a spatially uncorrelated channel. Thus, the SNR penalty of the FIX allocation is proportional to L.

As an infinite penalty is too pessimistic for $M > \tilde{M}$, we consider large but finite SNR, for which $C_{0,\text{FIX}} \approx \tilde{M}\log_2\left(\frac{\rho}{\tilde{M}/\Gamma_{\tilde{\mathcal{N}}_u}}\right)$. In such a case, we take logarithm on both sides of (212), yielding

$$\log_2(\text{SNR penalty}) = \left(\frac{1}{\tilde{M}} - \frac{1}{M}\right) R + \log_2\left(\frac{\tilde{M}}{M\Gamma_{\tilde{\mathcal{N}}_u}\mathbb{E}[1/G_{\mathcal{N}_u}]}\right), \tag{213}$$

which is linearly proportional to R, with a slope of $\frac{1}{\tilde{M}} - \frac{1}{M}$. When M is large, however, the slope quickly approaches zero, and the SNR penalty is reduced. As $M \to \infty$, we have $\frac{1}{\tilde{M}} - \frac{1}{M} \to 0$ and $\frac{\tilde{M}}{M\Gamma_{\tilde{\mathcal{N}}_u}\mathbb{E}[1/G_{\mathcal{N}_u}]} \to 1$, implying that $\log_2(\text{SNR penalty})$, which agrees with Proposition 6.4. To check the convergence speed, we consider the following example.

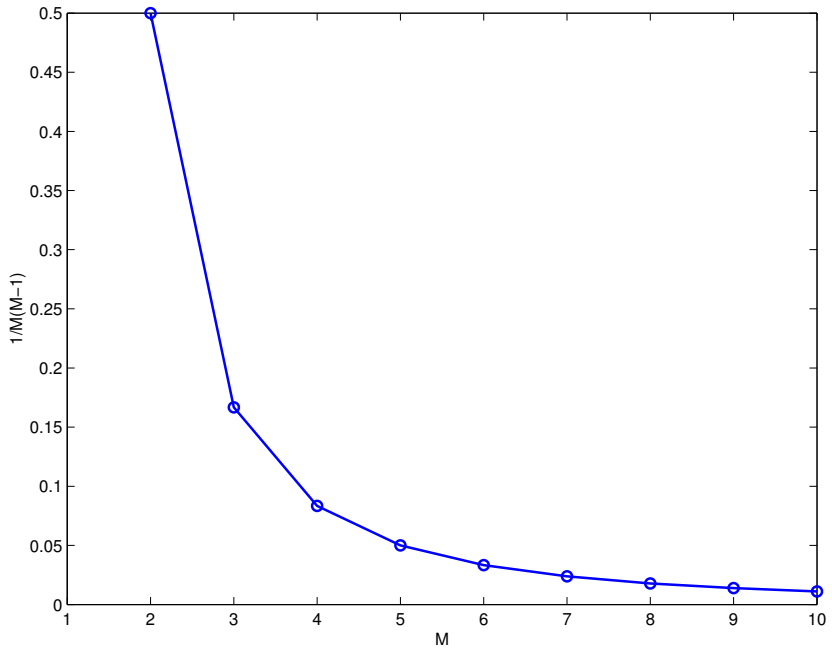

Figure 32: A plot of $\frac{1}{\tilde{M}} - \frac{1}{M}$ versus M on an $M \times M$ MIMO channel.

Example 6.4. On a $M \times M$ spatially uncorrelated channel with memory $L = 4$, we conjectured that $\tilde{M} = M - 1$. Figure 32 shows that $\frac{1}{\tilde{M}} - \frac{1}{M} = \frac{1}{M(M-1)}$ rapidly decreases to 0 as M grows. ♦

To calculate the SNR penalties at practical R, we must resort to Monte-Carlo simulations in the following section.

6.5 Numerical Results

We have proposed three power-allocation strategies for MIMO-OFDM with eigen-beamforming:

- Optimal allocation (Theorem 4.4)
- FUSE-constrained allocation (Proposition 6.1)
- FIX-constrained allocation (Proposition 6.3)

which achieve $C_{0,\mathrm{OPT}}$, $C_{0,\mathrm{FUSE}}$, and $C_{0,\mathrm{FIX}}$, respectively. It has been shown that both $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ converge to $C_{0,\mathrm{OPT}}$ as the antenna array size tends to infinity. In this section, we show via Monte-Carlo simulations that $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ are also nearly optimal for moderate antenna array sizes. Monte-Carlo simulations generated 10,000 independent sets of channels, $\{\mathbf{G}_l\}$, with the uniform power profile. We assume that OFDM has $N = 256$ tones.

6.5.1 Spatially Uncorrelated Channels

First, we consider a spatially uncorrelated channel.

6.5.1.1 Square Channels

We assume that the channel matrices are square ($M_T = M_R = M$). Figure 33 shows $C_{0,\mathrm{OPT}}$, $C_{0,\mathrm{FUSE}}$, and $C_{0,\mathrm{FIX}}$ against SNR in dB for $L = 4$ and $M \in \{1, 2, 4, 6\}$ when the channel is spatially uncorrelated. When $M = 1$ (SISO), both $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ are zero, as mentioned before. The optimal allocation has nonzero $C_{0,\mathrm{OPT}}$ because water-filling across tones exploits the diversity from frequency selectivity, otherwise it too would be zero.

In stark contrast, when $M = 2$, $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ become nonzero due to spatial (antenna) diversity. As explained in Corollary 6.2, $C_{0,\mathrm{OPT}}$ and $C_{0,\mathrm{FUSE}}$ have the same asymptotic slope ($M = 2$), *i.e.* spatial multiplexing order, while $C_{0,\mathrm{FIX}}$ has a lower slope ($\tilde{M} = 1$). Thus, it can be seen that the gap between $C_{0,\mathrm{FIX}}$ and $C_{0,\mathrm{OPT}}$ grows with SNR, and the FIX constraint incurs an unbounded penalty in the end as SNR tends to infinity.

As the antenna array sizes increase to $M = 4$ and $M = 6$, both $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ are very close to $C_{0,\mathrm{OPT}}$, as expected from high-SNR analysis. For $M = 4$ or $M = 6$, $C_{0,\mathrm{FIX}}$ has a lower asymptotic slope ($\tilde{M} = M - 1$) than $C_{0,\mathrm{OPT}}$ or $C_{0,\mathrm{FUSE}}$, and eventually incurs an unbounded penalty as SNR goes to infinity. However, for a range of practical SNR, Figure 33 illustrates the penalty of the FIX allocation is

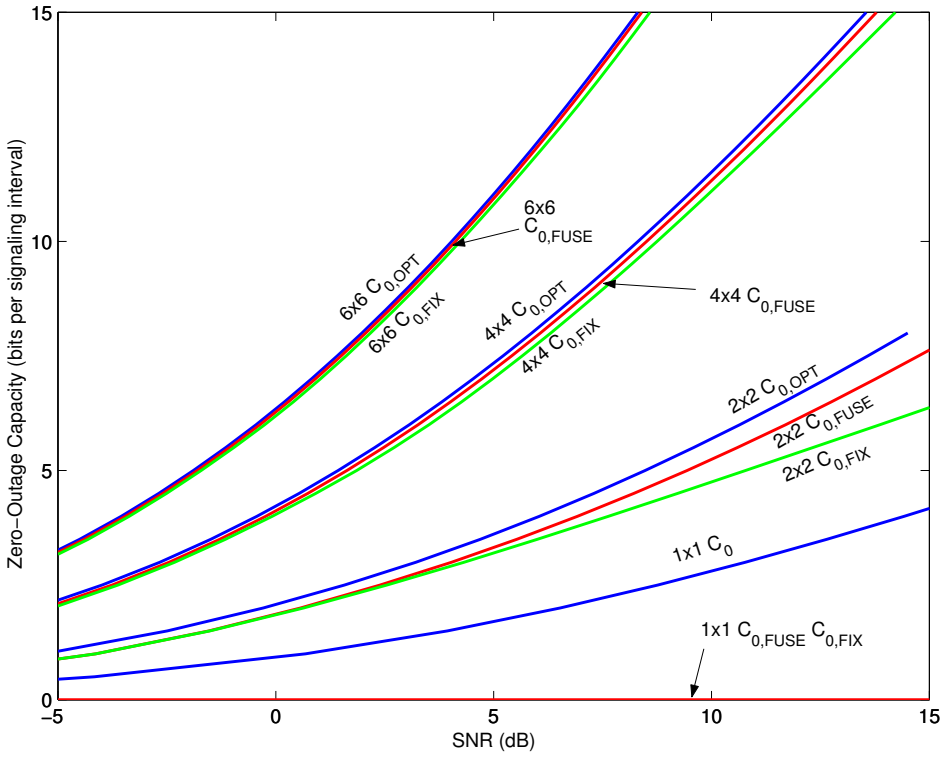

Figure 33: Zero-outage capacities of the FUSE and FIX allocations on an $M \times M$ spatially uncorrelated Rayleigh-fading channel with $L = 4$ memory.

small.

To emphasize how far $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ are separated from $C_{0,\mathrm{OPT}}$, Figure 34 and Figure 35 illustrate the SNR penalty, which is the additional SNR required by the FUSE and FIX allocations relative to the optimal allocation at a given target rate R ranging from 2 to 14. The SNR penalty corresponds to horizontal separation from $C_{0,\mathrm{OPT}}$ in Figure. 33.

The FUSE allocation incurs an SNR penalty of more than 0.8 dB for $M = 2$, but the penalty reduces to less than 0.3 dB for $M = 4$ and $M = 6$ in Figure 34. From Figure 29, the SNR penalties converge to $\{1.47, 0.41, 0.20\}$ in dB for $M \in \{2, 4, 6\}$, respectively, as the target rate goes to infinity.

For the FIX allocation, the SNR penalty at $M = 2$ is quite large due to a smaller multiplexing order ($\tilde{M} = 1$) in Figure 33. Figure 35 illustrates that the penalty is

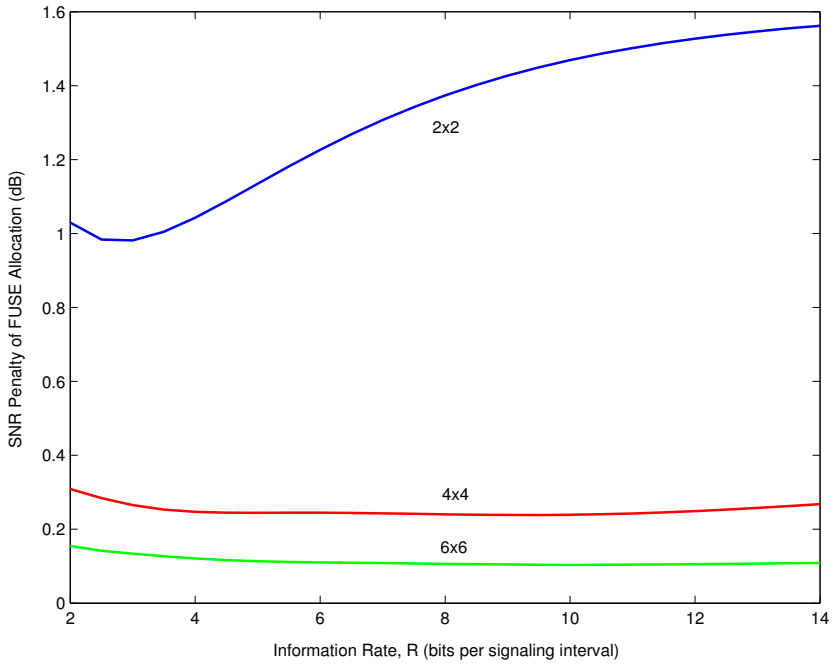

Figure 34: SNR penalty of the FUSE allocation on an $M \times M$ spatially uncorrelated Rayleigh-fading channel with $L = 4$ memory for $M \in \{2, 4, 6\}$.

more than 1 dB and quickly diverges. For $M = 4$ and $M = 6$, the FIX allocation also has a shallower slope in Figure 33 and the SNR penalties ultimately diverge. However, Figure 35 demonstrates that the SNR penalties for $2 \leq R \leq 14$ are surprisingly small, less than 0.5 and 0.3 dB for $M = 4$ and $M = 6$, respectively. For both the FUSE and FIX allocations, Figure 35 confirms that the penalties become small for a moderate M, implying that the convergence of $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ toward $C_{0,\mathrm{OPT}}$ is fast.

As observed before, the SNR penalty is a function of L. When $L = 0$, there is no penalty, while the SNR penalty increases with L. In Figure 36, we plot the SNR penalty of the FUSE-constrained allocation for various L. An interesting observation is that the increase step is getting smaller as L grows and looks to converge to a finite value. This agrees with Proposition 5.7, where we have shown the SNR penalty is finite.

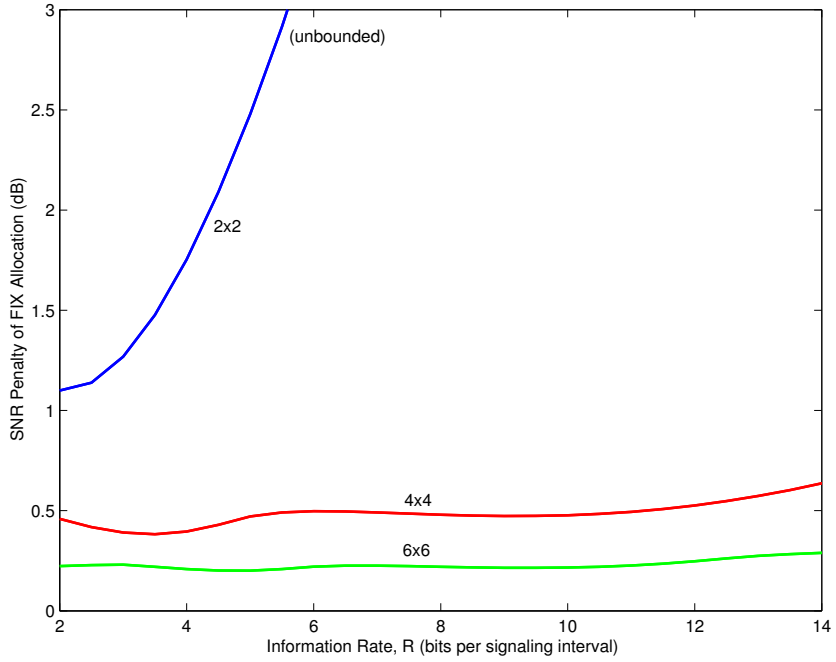

Figure 35: SNR penalty of the FIX allocation on an $M \times M$ spatially uncorrelated Rayleigh-fading channel with $L = 4$ memory for $M \in \{2, 4, 6\}$.

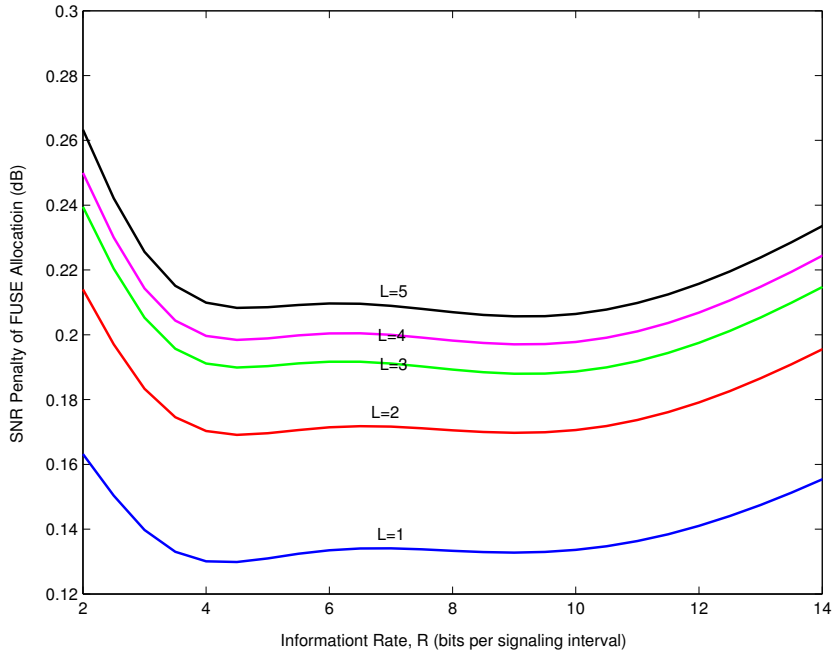

Figure 36: SNR penalty of the FUSE Allocation on a 4×4 spatially uncorrelated Rayleigh-fading channel with memory $L \in \{1, 2, 3, 4, 5\}$ grows.

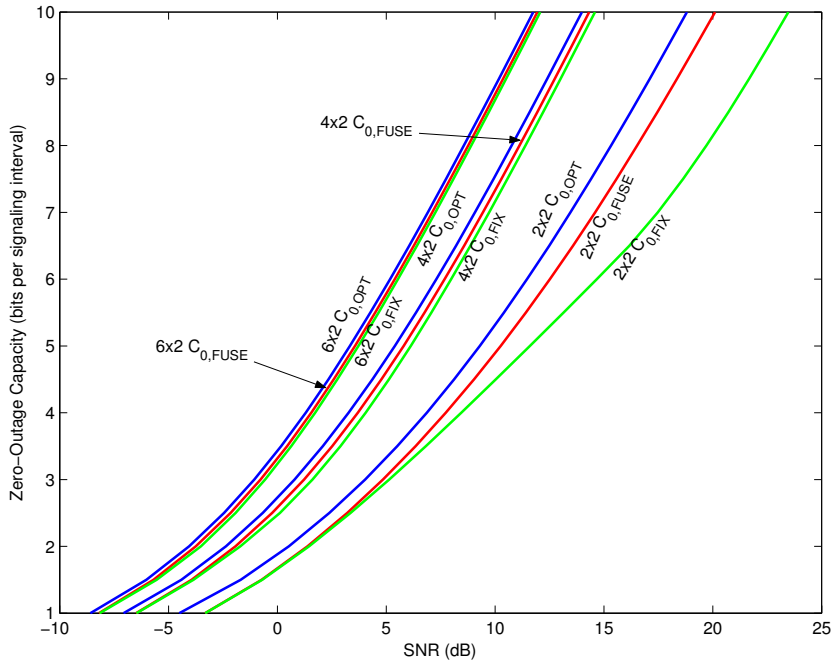

Figure 37: Zero-outage capacities of the FUSE and FIX allocations on spatially uncorrelated 2×2, 4×2, and 6×2 channels with $L = 4$ memory.

6.5.1.2 Non-square Channels

Now we consider non-square MIMO channels with $M_T = 2$ and $M_R \in \{2, 4, 6\}$ with channel memory $L = 4$. As observed in Figure 33, when $M_T = M_R = 2$, $C_{0,\mathrm{FIX}}$ has a lower slope ($\tilde{M} = 1$) than $C_{0,\mathrm{OPT}}$ and $C_{0,\mathrm{FUSE}}$, whose slopes are $M = 2$, and hence $C_{0,\mathrm{FIX}}$ can be significantly smaller than $C_{0,\mathrm{OPT}}$ and $C_{0,\mathrm{FUSE}}$, especially at high SNR. However, as the number of receive antennas grows to $M_R = 4$ (4×2) or $M_R = 6$ (6×2), as shown in Figure 37, $C_{0,\mathrm{FIX}}$ has the same slope as $C_{0,\mathrm{OPT}}$ and $C_{0,\mathrm{FUSE}}$, namely $M = \tilde{M} = 2$, and also the gaps between the curves are reduced owing to higher spatial diversity orders, namely $M_T M_R = 8$ for 4×2 and $M_T M_R = 12$ for 6×2. Note that the slopes (the multiplexing orders) for all $M_R \in \{2, 4, 6\}$ are identical as $M = \min(M_T, M_R) = 2$ in Figure 37.

6.5.2 Spatially Correlated Channels

In the section, we examine the case where the fading on the transmitter side is spatially uncorrelated, while spatial fading at receive antennas is correlated with $\mathbf{R}_{R_l}$ given in (6). Based on the correlation model in Section 2.1.2, we calculate $C_{0,\text{OPT}}$, $C_{0,\text{FUSE}}$, and $C_{0,\text{FIX}}$ via Monte-Carlo simulations. As the rank of $\mathbf{R}_{R_l}$ is increasing with L, $C_{0,\text{OPT}}$ also increases notably. Simulations show that this is true for $C_{0,\text{FUSE}}$ and $C_{0,\text{FIX}}$.

Figure 36 illustrates the SNR penalty of the FUSE constraint for $M = 4$ when cluster angle spread is either large ($\sigma_{\theta_l} = 0.25$) or small ($\sigma_{\theta_l} = 0$) as well as the SNR penalty of uncorrelated channels as a benchmark. We assume that there are $L = 4$ clusters, whose average angles $\{\bar{\theta}_l\}$ are $\{\pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4\}$. From Figure 38, it can be seen that the SNR penalties are 0.1 dB for small spread and 0.15 dB for large spread, both of which are less than the penalty on the uncorrelated channel. Figure 38 shows a tendency that the SNR penalty of a spatially correlated channel increases as the cluster angle spread gets larger and in the end reaches the SNR penalty of a spatially uncorrelated channel.

6.6 Proof of Proposition 6.5

The proof is similar to Theorem 5.2 [7], but much easier since $\mathbb{E}[1/s^{(m)}] < \infty$ for $m = 1, 2, \ldots, \tilde{M}$. Let $u_n^{(m)} = 1/\mathbb{E}[1/s_n^{(m)}]$. First, we derive a lower bound for $C_{0,\text{FIX}}$, which holds for all SNR. From Proposition 6.3, we have

$$\rho = \mathbb{E}\left[\sum_{m=1}^{\tilde{M}} \left\{\frac{\nu u_n^{(m)}}{s_n^{(m)}} - \frac{1}{s_n^{(m)}}\right\}^+\right] \leq \mathbb{E}\left[\sum_{m=1}^{\tilde{M}} \frac{\nu u_n^{(m)}}{s_n^{(m)}}\right] = \tilde{M}\nu. \tag{214}$$

Then,

$$\begin{aligned} C_{0,\text{FIX}} &= \sum_{m=1}^{\tilde{M}} \left\{\log_2\left(\nu u_n^{(m)}\right)\right\}^+ \\ &\geq \sum_{m=1}^{\tilde{M}} \log_2\left(\nu u_n^{(m)}\right) \end{aligned}$$

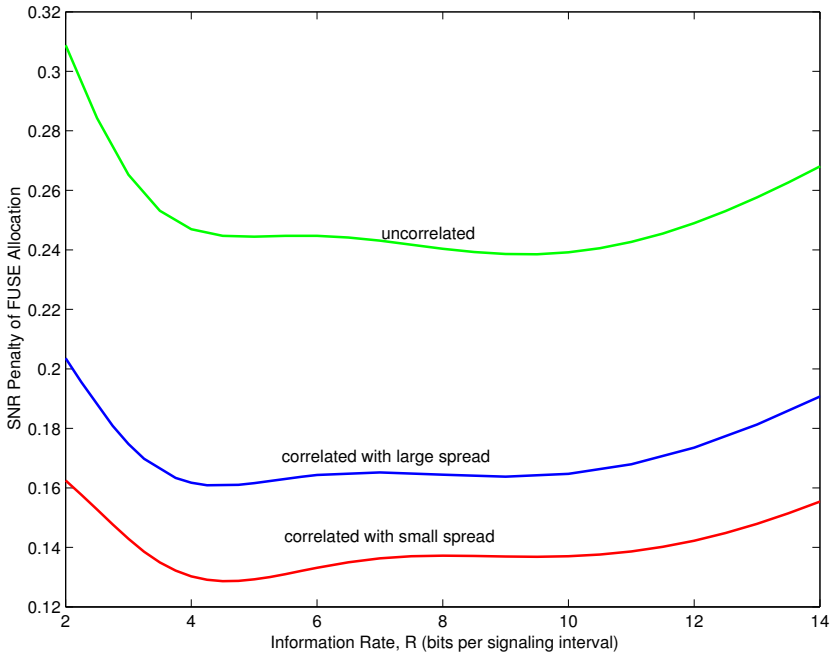

Figure 38: SNR penalty of FUSE constraint on a 4×4 spatially correlated channel with $L = 4$ memory when $\sigma_{\theta_l} = 0$ and $\sigma_{\theta_l} = 0.25$.

$$\geq \sum_{m=1}^{\tilde{M}} \log_2 \left(\frac{\rho u_n^{(m)}}{\tilde{M}} \right), \tag{215}$$

where the second inequality comes from (214). Thus, we have the inequality

$$C_{0,\text{FIX}} \geq M \log_2 \left(\frac{\rho}{\tilde{M}} \right) + M \log_2 \left(\Gamma_{\tilde{\mathcal{N}}_u} \right). \tag{216}$$

On the other hand, we consider a region:

$$\left\{ \left[u_n^{(1)}, \ldots, u_n^{(\tilde{M})} \right]; u_n^{(\tilde{M})} > \Gamma_{\tilde{\mathcal{N}}_u} 2^{-R/\tilde{M}} \right\}, \tag{217}$$

for which we use all $\tilde{M}$ channels at a rate of R, such that $|\mathcal{M}_{\text{FIX}}|$ in Proposition 6.3 is $\tilde{M}$. Then,

$$\begin{aligned} \rho &\geq \mathbb{E} \left[\sum_{m=1}^{\tilde{M}} \frac{2^{R/\tilde{M}} u_n^{(m)}}{\Gamma_{\tilde{\mathcal{N}}_u} s_n^{(m)}} - \frac{1}{s_n^{(m)}} \right] 1 \left\{ u_n^{(\tilde{M})} > \Gamma_{\tilde{\mathcal{N}}_u} 2^{-R/\tilde{M}} \right\} \\ &= 2^{R/\tilde{M}} \left(\frac{1}{\Gamma_{\tilde{\mathcal{N}}_u}} - 2^{-R/\tilde{M}} \mathbb{E} \left[\frac{1}{s_n^{(m)}} \middle| 1 \left\{ u_n^{(\tilde{M})} > \Gamma_{\tilde{\mathcal{N}}_u} 2^{-R/\tilde{M}} \right\} \right] \right). \end{aligned} \tag{218}$$

As $R \to \infty$, $\mathbb{E}\left[\frac{1}{s_n^{(m)}}\middle|1\left\{u_n^{(\tilde{M})} > \Gamma_{\tilde{\mathcal{N}}_u} 2^{-R/\tilde{M}}\right\}\right] \to \mathbb{E}\left[\frac{1}{s_n^{(m)}}\right] < \infty$. Therefore, (218) asymptotically becomes

$$C_{_{0,\mathrm{FIX}}} \leq M \log_2\left(\frac{\rho}{\tilde{M}}\right) + M \log_2\left(\Gamma_{\tilde{\mathcal{N}}_u}\right), \tag{219}$$

at high SNR. From (216) and (219), we have the assertion in (211).

CHAPTER 7

RATE ALLOCATION WITH GRANULARITY CONSIDERATIONS

With $\{e_n^{(m)}\}$ from power allocation, the corresponding achievable rate of the mth spatial channel at the nth tone is $r_n^{(m)} = \log_2\left(1 + \frac{s_n^{(m)} e_n^{(m)}}{N_0}\right)$. When calculating $\{e_n^{(m)}\}$, there is no restriction in $\{r_n^{(m)}\}$ except that it should be real and nonnegative. In practice, however, infinite-precision rate can hardly be realizable for the complexity reason, and thus a *granularity* constraint is applied to $\{r_n^{(m)}\}$, that is, $\{r_n^{(m)}\}$ must be chosen from a discrete and finite set.

In this chapter, we investigate the allocation problem with granularity considerations, often known as *bit-allocation problem*. Traditionally, the bit-allocation problem has been extensively studied for DSL applications [14, 17, 19], which basically modifies the water-filling allocation to meet the granularity constraint. In this work, we attack the bit-allocation problem with a different viewpoint, and propose very simple allocation strategies for MIMO flat-fading channels by exploiting the statistical properties of MIMO channels. The proposed bit-allocation strategies can extend the proposed strategies to MIMO-OFDM combined with the FUSE constraint in Proposition 6.1.

7.1 Bit-Allocation Problem and Full-Search Allocation

We restrict our attention to the flat-fading channel, which is converted by eigenbeamforming into a bank of M scalar channels:

$$\begin{aligned} z^{(1)} &= \sqrt{s^{(1)}}a^{(1)} + w^{(1)} \\ &\vdots \quad \vdots \\ z^{(M)} &= \sqrt{s^{(M)}}a^{(M)} + w^{(M)} \end{aligned}, \tag{220}$$

as illustrated in Figure 3, where $\boldsymbol{s} = [s^{(1)}, s^{(2)}, \ldots, s^{(M)}]$ are the nonzero eigenvalues of $\mathbf{HH}^*$ such that $s^{(1)} \geq s^{(2)} \geq \ldots \geq s^{(M)} > 0$, and where the noise $\{w^{(m)}\}$ are i.i.d. $\mathcal{CN}(0, N_0)$. Note that we ignore the tone index n for the notational simplicity.

So as to achieve a rate of $r^{(m)}$ bits per signaling interval across the mth scalar channel, its SNR $s^{(m)}e^{(m)}/N_0$ must be at least $\Gamma(2^{r^{(m)}}-1)$, where $e^{(m)} = \mathbb{E}\left[|a^{(m)}|^2 \left| \{s^{(m)}\}\right.\right]$, and where Γ is an SNR gap, which accounts for the additional SNR required for a practical code to achieve a given target error probability [19]. With an ideal capacity-achieving code, Γ reduces to unity. Then, the energy required by the transmitter to achieve a given set of rate $\{r^{(m)}\}$ is:

$$E(\boldsymbol{s}) = \sum_{m=1}^{M} e^{(m)} = \Gamma \sum_{m=1}^{M} \frac{2^{r^{(m)}} - 1}{s^{(m)}/N_0}. \tag{221}$$

Without the granularity restriction, the rate allocation that minimizes (221) to achieve a given total rate of $R = \sum_{m=1}^{M}$ bits per signaling interval is given by the water-filling solution:

$$r^{(m)} = \left\{ \log_2 \left(\frac{\lambda s^{(m)}}{\Gamma} \right) \right\}^+, \tag{222}$$

where λ ensures that $R = \sum_{m=1}^{M} r^{(m)}$.

In practice, complexity considerations require that $\{r^{(m)}\}$ be drawn from a discrete and finite set. Let the granularity β be the smallest incremental unit for $r^{(m)}$. Then, the rate of any scalar channel is given by $r^{(m)} = \beta B^{(m)}$, where $B^{(m)}$ is a non-negative integer and $B^{(m)}$ is limited to $B_{\max}$.

Definition 7.1 (Bit-Allocation Problem). With the granularity β and the upper limit $B_{\max}$, the bit-allocation problem, given $\{s^{(m)}\}$, is to find the $\{r^{(m)}\}$ with

$$r^{(m)} \in \{0, \beta, 2\beta, \ldots, B_{\max}\beta\},^{1} \tag{223}$$

that minimizes the energy requirement in (221) subject to a rate constraint, $R = \sum_{m=1}^{M} r^{(m)}$. ♦

Clearly, the best allocation is based on a full search that enumerates all possible candidates and chooses the candidate that has minimum energy requirement in (221). In other words, the best allocation compare all candidates in the full-search set, as follows.

Definition 7.2 (Full-Search Set). We define a full-search set:

$$\mathcal{B} = \left\{ [r^{(m)}]; \sum_{m=1}^{M} r^{(m)} = R, r^{(1)} \geq \ldots \geq r^{(M)} \geq 0, r^{(m)} \in \{0, \beta, 2\beta, \ldots, B_{\max}\beta\} \right\}, \tag{224}$$

where the ordering $r^{(1)} \geq r^{(2)} \geq \ldots \geq r^{(M)}$ is due to the increasing nature of $\{s^{(m)}\}$ since we never allocate more bits to $s^{(m)}$ than to $s^{(m')}$ for $m < m'$. ♦

We present a couple of examples of the full-search sets. As a short-hand notation, $[r^{(1)}, r^{(2)}, \ldots, r^{(M)}]$ is used to represent the bit allocation.

Example 7.1. Suppose that we use uncoded modulation based on complex-valued constellations, such as QAM. Then, the incremental unit is $\beta = 1$. If $B_{\max} = 6$, when $R = 4$ and $M = 4$, the full-search set is

$$\mathcal{B} = \{[4,0,0,0], [3,1,0,0], [2,2,0,0], [2,1,1,0], [1,1,1,1]\}, \tag{225}$$

whose cardinality is only $|\mathcal{B}| = 5$. When $R = 8$ and $M = 4$,

$$\mathcal{B} = \Big\{[6,2,0,0], [6,1,1,0], [5,3,0,0], [5,2,1,0], [5,1,1,1], [4,4,0,0], \tag{226}$$

$$[4,3,1,0], [4,2,2,0], [4,2,1,1], [3,3,2,0], [3,3,1,1], [2,2,2,2]\Big\}, \tag{227}$$

[1] We have freedom to choose the set with an irregular step size, such as $\{0, \beta, 4\beta\}$.

whose cardinality increases to 12. ♦

Example 7.2. We assume uncoded modulation as in Example 7.1, but only even constellations (*e.g.* 4QAM, 16QAM, 64QAM) are allowed, so that $\beta = 2$. When $R = 4$ and $M = 4$, the full-search set is

$$\mathcal{B} = \{[4, 0, 0, 0], [2, 2, 0, 0]\}, \tag{228}$$

whose cardinality is $|\mathcal{B}| = 2$. When $R = 8$ and $M = 4$,

$$\mathcal{B} = \Big\{[6, 2, 0, 0], [4, 4, 0, 0], [4, 2, 2, 0], [2, 2, 2, 2]\Big\}, \tag{229}$$

whose cardinality is to $|\mathcal{B}| = 4$. ♦

Example 7.3. In this example, we investigate the size of search set, $|\mathcal{B}|$, with respect to M and R. Suppose that $\beta = 1$ for uncoded modulation and $B_{\max} = 6$. Figure 39 illustrates $|\mathcal{B}|$ for $M \in \{2, 4, 6\}$ as R grows. For $M = 2$, $|\mathcal{B}|$ is relatively small. For $M = 4$ and $M = 6$, in contrast, $|\mathcal{B}|$ can be large. ♦

According to the above examples, the size of the full-search set is relatively small compared to that of MIMO-OFDM. However, the size can be still too large to consider all candidates of $\mathcal{B}$ for the bit allocation when M and R are large. In the following, we investigate the properties of the full-search allocation and show that only a few candidates from $\mathcal{B}$ are sufficient for the members of the search set.

We use the average energy requirement

$$\bar{E} = \mathbb{E}\left[E(\boldsymbol{s})\right] = \mathbb{E}\left[\Gamma \sum_{m=1}^{M} \frac{2^{r^{(m)}} - 1}{s^{(m)}/N_0}\right], \tag{230}$$

as the performance criterion. Let $\boldsymbol{b}_j = [b_j^{(1)}, \ldots, b_j^{(M)}]$ indicate the jth candidate (arbitrary ordering) in $\mathcal{B}$ and let A_j be the subregion in an M-dimensional space $\{\boldsymbol{s} = [s^{(1)}, \ldots, s^{(M)}]; s^{(1)} \geq s^{(2)} \geq \ldots \geq s^{(M)}\}$, in which $\boldsymbol{b}_j$ is the optimal choice in

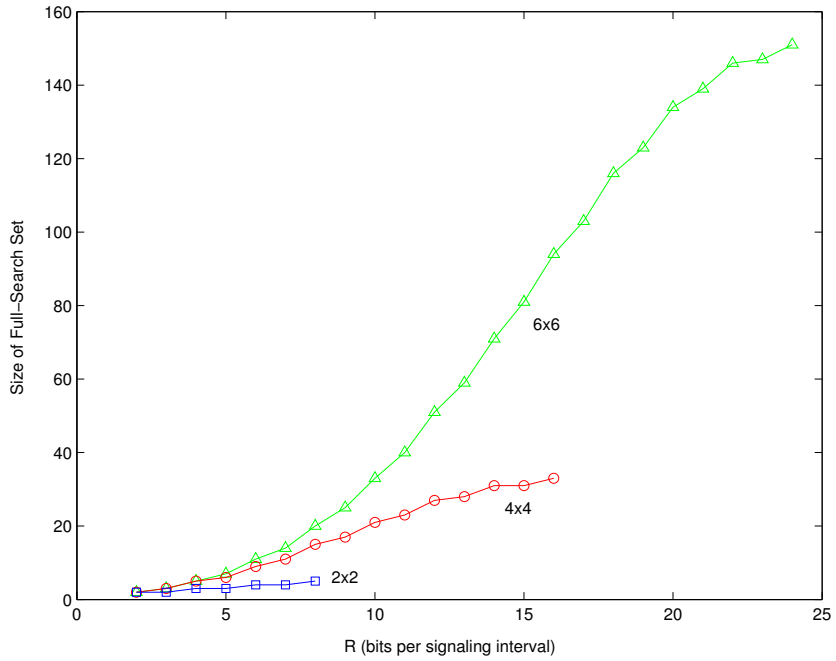

Figure 39: Size of full-search set for $M \in \{2, 4, 6\}$.

terms of (230). In other words, if $\boldsymbol{s} \in A_j$, we will choose $r^{(m)} = b_j^{(m)}$ for $m = 1, \ldots, M$. Then, the average energy requirement for the full search becomes

$$\begin{aligned} \bar{E} &= \mathbb{E}_{A_j}\left[\mathbb{E}_{\boldsymbol{s}}\left[\Gamma \sum_{m=1}^{M} \frac{2^{b_j^{(m)}} - 1}{s^{(m)}/N_0}\middle| A_j\right]\right] \\ &= \Gamma \sum_{j=1}^{|\mathcal{B}|} P_j \sum_{m=1}^{M} \left(2^{b_j^{(m)}} - 1\right) \mathbb{E}_{\boldsymbol{s}}\left[\frac{1}{s^{(m)}}\middle| A_j\right] \\ &= \Gamma \sum_{j=1}^{|\mathcal{B}|} P_j \epsilon_j, \end{aligned} \tag{231}$$

where $\epsilon_j = \sum_{m=1}^{M} \left(2^{b_j^{(m)}} - 1\right) \mathbb{E}_{\boldsymbol{s}}\left[1/s^{(m)} | A_j\right]$ is the partial energy requirement conditioned on A_j. The probability mass function of candidates is denoted by $P_j = \text{Prob}[\boldsymbol{s} \in A_j]$ for $j = 1, \ldots, |\mathcal{B}|$, which indicates how often $\boldsymbol{b}_j$ is selected over realizations of $\boldsymbol{s}$.

Example 7.4. Suppose $\beta = 3/4$ and $R = 9$ on a 6×6 Rayleigh-flat-fading channel. If $B_{\max} = 8$, there are 51 members (candidates) in $\mathcal{B}$. We use Monte-Carlo simulations

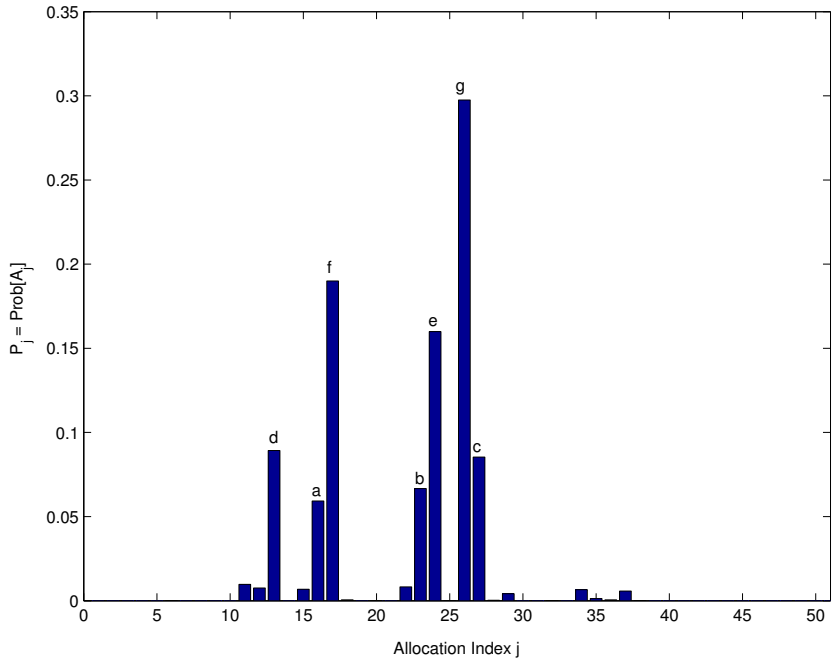

Figure 40: Probability that a certain allocation is used for $M = 6$ and $R = 9$.

by generating 10^5 independent $\{s^{(m)}\}$ to calculate P_j. Figure 40 plots P_j for $|\mathcal{B}| = 51$ candidates. Out of 51 candidates, only a few candidates have large P_j. We define the candidates with $P_j > 0.05$ as *dominant* candidates.[2] In Figure 40, there are seven dominant candidates. Alphabetical labels in Figure 40 identify seven dominant candidates, summarized in Table 4. ⧫

If an candidate has small P_j, its contribution to the average energy requirement in (231) is marginal unless ϵ_j is enormously large. Thus, small-P_j candidates are deleted from considerations without significantly increasing (231). A natural question is how many dominant candidates a full-search set has.

Observation 7.1. In a MIMO channel, the number of dominant candidates is small relative to the size of the full-search set $|\mathcal{B}|$. ⧫

[2]The definition of dominant candidates is arbitrary. To investigate the properties of P_j, we choose 0.05 as a threshold.

Table 4: Dominant allocation candidates for $M = 6$ and $R = 9$, which are denoted in the form of $[r^{(1)}/\beta, \ldots, r^{(M)}]/\beta$.

Label	Candidate
a	[4,4,2,2,0,0]
b	[5,3,2,2,0,0]
c	[5,4,3,0,0,0]
d	[4,3,3,2,0,0]
e	[5,3,3,1,0,0]
f	[4,4,3,1,0,0]
g	[5,4,2,1,0,0]

This observation can be explained in part by the fact that $\{s^{(m)}\}$ are more predictable in MIMO, especially when M is large. A justification for the predictability is that, $s^{(m)}/M$ converges with probability one to a non-random value from Theorem 2.2. To show the above observation more explicitly, we consider the following example.

Example 7.5. Consider 4×4 and 6×6 Rayleigh fading with $\beta = 3/4$ and $B_{\max} = 8$. We say that a candidate is dominant if $P_j > 0.05$. The number of dominant candidates is $|\mathcal{B}| \cdot \text{Prob}[P_j < 0.05]$. Figure 41 shows the number of dominant candidates for various $B = R/\beta$. We can see that the number of dominant candidates is small (less than or equal to 8) for all B, which justifies Observation 7.1. For small B, the number of dominant candidates is small since $|\mathcal{B}|$ is small though $\text{Prob}[P_j < 0.05]$ is relatively large. When B is large, $\text{Prob}[P_j < 0.05]$ is small, which compensates for large $|\mathcal{B}|$. ♦

From Observation 7.1, we know that most candidates, except only a few dominant candidates, have insignificant P_j. According to (231), insignificant P_j does not affect the average energy requirement $\bar{E}$, that is, deleting these insignificant candidates from considerations does not degrade the performance considerably.

Observation 7.2. If P_j is negligible, deleting its candidate $\boldsymbol{b}_j$ from $\mathcal{B}$ and using other candidate(s) for A_j will increase the average energy requirement in (231), but the increase is only marginal. ♦

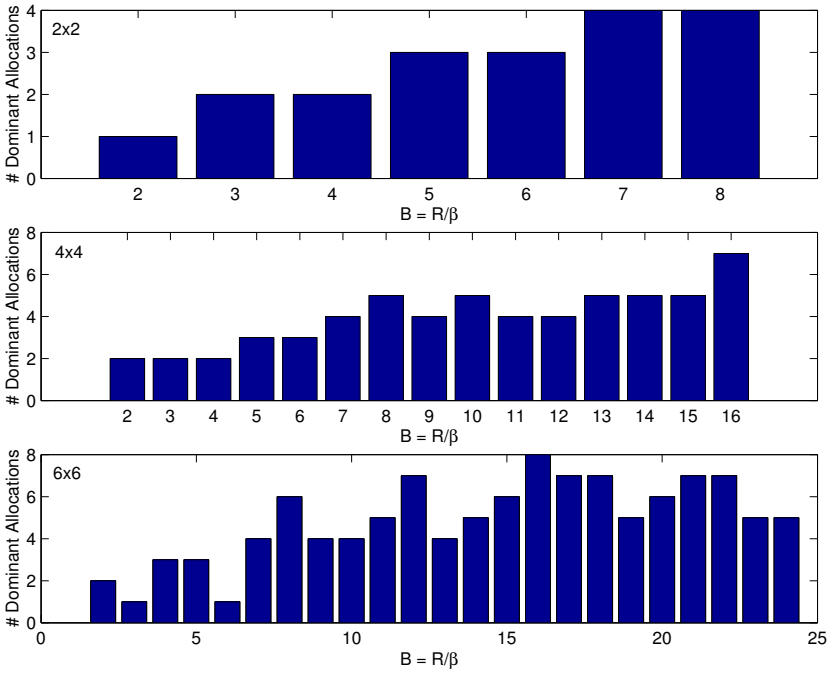

Figure 41: Number of dominant candidates ($P_j > 0.05$) for 2×2, 4×4, and 6×6 MIMO channels, respectively.

To illustrate Observation 7.2, we reduce the number of candidates for the bit-allocation search. Let $\mathcal{B}_k$ denote *reduced search set*, which contains k candidates, such that $\mathcal{B}_k = \mathcal{B}$ if $k = |\mathcal{B}|$. Instead of considering all candidates in $\mathcal{B}$, we restrict our search to $\mathcal{B}_k$. Clearly, as k becomes small, the complexity for the bit-allocation search is reduced, whereas the performance is more degraded. According to Observation 7.2, the degradation is insignificant as long as dominant candidates remain in the search set. Let $\rho_{\mathcal{B}_k}$ denote the average SNR, $\rho = \mathbb{E}[E(\boldsymbol{s})]/N_0$, required by restricting search to $\mathcal{B}_k$. Clearly, the inequality $\rho_{\mathcal{B}_k} \geq \rho_{\mathcal{B}}$ holds for any k. To measure the degradation, we examine the SNR penalty by restricting search to $\mathcal{B}_k$ instead of $\mathcal{B}$, defined as

$$\text{SNR penalty } = \frac{\rho_{\mathcal{B}_k}}{\rho_{\mathcal{B}}}. \tag{232}$$

We calculate the SNR penalty by deleting candidates in $\mathcal{B}$ one by one. The procedures are as follows:

- Initialize $k = |\mathcal{B}|$.

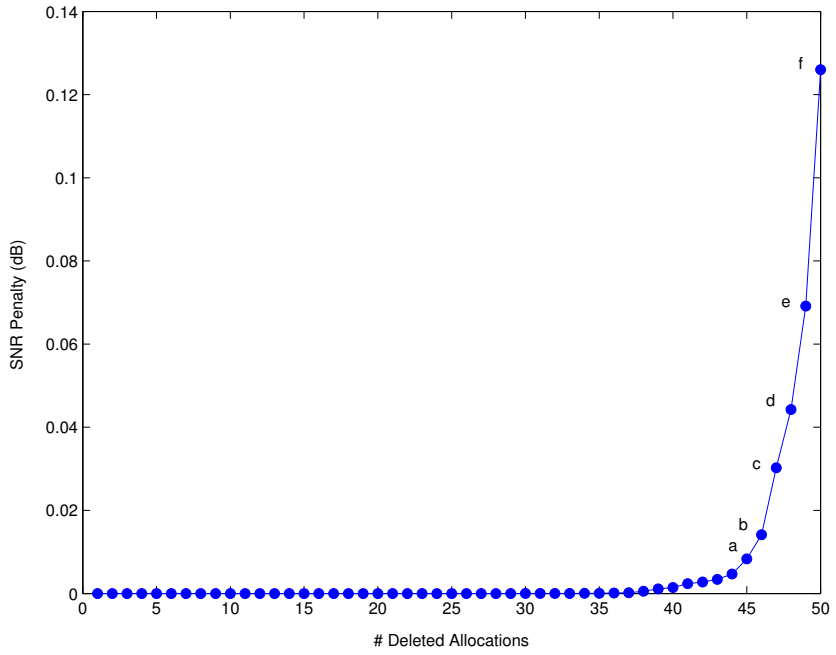

Figure 42: SNR penalty as we delete allocations one by one from a full-search set for $M = 6$ and $R = 9$.

- Repeat while $k > 1$

  1. Calculate P_j for $\mathcal{B}_k$.
  2. Delete the member with the smallest P_j to constitute $\mathcal{B}_{k-1}$.[3]
  3. Calculate $\rho_{\mathcal{B}_{k-1}}$.
  4. $k = k - 1$

Example 7.6. Suppose $\beta = 3/4$, $B_{\max} = 8$, and $R = 9$ on 6×6 Rayleigh fading. Figure 42 illustrates the SNR penalty of (232) in dB as the size of the search set is reduced. The alphabetical labels in Figure 42 matches the dominant candidates in Figure 40. Until deleting seven (labeled) dominant candidates, the SNR penalty is negligible, which agrees with Observation 7.2. ♦

In Figure 42, we notice that the SNR penalty begins to grow sharply by removing

[3]If more than one members have the same smallest P_j, choose randomly one of them.

'a' candidate because P_j is no longer negligible. However, when all but one candidate are removed from considerations ($\mathcal{B}_1$), the SNR penalty reaches up to 0.13 dB, which is reasonably small. From simulations for other M and R, we have the following observation:

Observation 7.3. The penalty by removing dominant candidates is not negligible any longer, but still small especially for large M. ♦

To illustrate Observation 7.3, we consider an extreme case where only one candidate, $\{b_{\text{fix}}^{(m)}, m = 1, 2, \ldots, M\}$, remains in the search set ($\mathcal{B}_1$), that is, the bit allocation is fixed independent of $\boldsymbol{s}$. Then, the average energy requirement becomes

$$\begin{aligned} \bar{E} &= \Gamma \sum_{j=1}^{|\mathcal{B}|} P_j \sum_{m=1}^{\tilde{M}} \left(2^{b_{\text{fix}}^{(m)}} - 1\right) \mathbb{E}_{\boldsymbol{s}} \left[\frac{1}{s^{(m)}} \middle| A_j \right] \\ &\approx \Gamma \sum_{\text{dominant} j} P_j \epsilon_j^{\text{fix}}, \end{aligned} \tag{233}$$

where $\tilde{M}$ is the largest integer satisfying $\mathbb{E}[1/s^{(\tilde{M})}] < \infty$, and where we denote the partial error as $\epsilon_j^{\text{fix}} = \sum_{m=1}^{M} \left(2^{b_{\text{fix}}^{(m)}} - 1\right) \mathbb{E}_{\boldsymbol{s}} \left[\frac{1}{s^{(m)}} \middle| A_j \right]$. Then, the difference between the average energy with the fixed allocation in (233) and the average energy of the full search in (231) lies in ϵ_j^{fix} replacing ϵ_j in (231), where $\epsilon_j^{\text{fix}} \geq \epsilon_j$.

Given channel statistics, the distance between dominant candidates is relatively small, where the distance of $\boldsymbol{r} = [r^{(1)}, \ldots, r^{(M)}]$ and $\boldsymbol{r}' = [r'^{(1)}, \ldots, r'^{(M)}]$ is defined as the norm of $\boldsymbol{r} - \boldsymbol{r}'$. In Table 4, we can confirm that each $r^{(m)}$ differs only by β if different. Usually, the distance between dominant candidates is not significantly large, which is also in part explained by the predictability of MIMO channels, that is, MIMO channels have more deterministic nature as M grows. Thus, ϵ_j^{fix} is usually only slightly greater than ϵ_j. We take an example to justify Observation 7.3.

Example 7.7. From Example 7.4, there are seven dominant candidates in $\mathcal{B}$ for $\beta = 3/4$, $M = 6$, and $R = 9$, which are alphabetically labeled in Figure 40. Figure 43 plots the excessive SNR requirement for each dominant candidate by restricting search

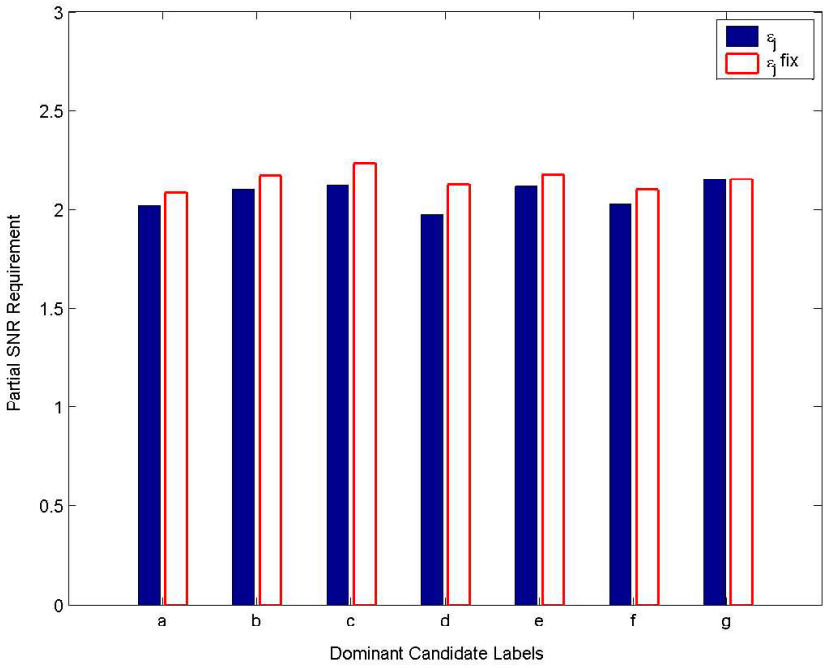

Figure 43: Partial SNR requirement when restricting search to $\mathcal{B}_1$ (ϵ_j^{fix}) in comparison with the partial SNR requirement of full search (ϵ_j) when $M = 6$ and $R = 9$.

to the candidate with label g, namely $\mathcal{B}_1 = \{\boldsymbol{b}_{\text{fix}}\}$. For the candidate with label g, we can confirm that $\epsilon_j^{\text{fix}} = \epsilon_j$. For other dominant candidates, ϵ_j^{fix} is clearly larger than ϵ_j, but none of them costs significantly large excessive SNR. ♦

Another justification comes from the fixed-rate allocation (without granularity constraint) in Proposition 6.3. Proposition 6.4 showed that the fixed-rate allocation is optimal as $M \to \infty$, and numerical results in Section 6.5.1 revealed that the SNR penalty of the fixed allocation is small for $M = 4$ and $M = 6$. These results guarantees that restricting search to $\mathcal{B}_1$ will not incur a significant penalty.

In this section, we showed by three observations that only a few candidates are sufficient for the members of search set without harming the performance significantly.

7.2 Proposed Bit-Allocation Strategies

In the previous section, we observed that reducing search-set size does not seriously degrade performance. Examples showed that we can reduce up to $\mathcal{B}_1$, which contains only one candidate, without harming the performance significantly when $M = 6$. However, when M is small, more candidates are necessary in the search set. In this section, we propose two bit-allocation strategies based on $\mathcal{B}_1$ and $\mathcal{B}_2$ and discuss the optimal choice of $\mathcal{B}_1$ and $\mathcal{B}_2$. Here are two proposed allocation strategies:

- **Binary Search:** Bit-allocation search is restricted over two candidates: $\mathcal{B}_2 = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$. For each realization of $\{s^{(m)}\}$, the transmitter chooses over $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, whichever has lower energy requirement of (230).

- **Fixed Allocation:** One allocation $\boldsymbol{b}_{\text{FIX}}$ is used independent of $\{s^{(m)}\}$, such that $\mathcal{B}_1 = \{\boldsymbol{b}_{\text{FIX}}\}$.

A remaining problem is how to choose $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ in Binary Search and $\boldsymbol{b}_{\text{FIX}}$ in Fixed Allocation among all candidates in $\mathcal{B}$. If the transmitter knows fading statistics, the choice can be made beforehand. One way is to delete the candidate which has the smallest P_j and to repeat this process until one or two candidates remain as in Example 7.6. However, this deleting process does not necessarily lead to the optimal choice that minimizes (230).

Instead, we compare average energy of (230) for all possible combinations that constitute $\mathcal{B}_k$, and choose the one that produces minimum average energy. For Fixed Allocation ($\mathcal{B}_1$), the optimal choice is very simple. For each $\boldsymbol{b}_j \in \mathcal{B}$, we calculate

$$\rho = \mathbb{E}\left[\sum_{m=1}^{M} \frac{2^{b_j^{(m)}} - 1}{s^{(m)}}\right] = \sum_{m=1}^{\tilde{M}} (2^{b_j^{(m)}} - 1)\mathbb{E}\left[\frac{1}{s^{(m)}}\right], \tag{234}$$

where $\tilde{M}$ is the largest integer satisfying $\mathbb{E}[1/s^{(\tilde{M})}] < \infty$, and then choose $\boldsymbol{b}_j$ with minimum ρ. If $b_j^{(m)} > 0$ for $m > \tilde{M}$, we remove the corresponding $\boldsymbol{b_j}$ from considerations. Notice that all it requires for calculating (234) is $\{\mathbb{E}[1/s^{(m)}]\}$. Also, note

that Γ and N_0 in (230) do not affect the optimal choice, which are hence removed in (234).

For Binary Search, the optimal choice is more complicated. For every possible pair of $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ from $\mathcal{B}$, we calculate

$$\rho = P_{\boldsymbol{b}_1} \sum_{m=1}^{M} (2^{b_1^{(m)}} - 1)\mathbb{E}\left[\frac{1}{s^{(m)}}\middle| A_{\boldsymbol{b}_1}\right] + P_{\boldsymbol{b}_1} \sum_{m=1}^{M} (2^{b_2^{(m)}} - 1)\mathbb{E}\left[\frac{1}{s^{(m)}}\middle| A_{\boldsymbol{b}_2}\right], \tag{235}$$

where $A_{\boldsymbol{b}_1}$ and $A_{\boldsymbol{b}_2}$ are the regions of $\boldsymbol{s}$, where $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are used, respectively, such that $A_{\boldsymbol{b}_1} \cap A_{\boldsymbol{b}_2}$ is a null space and $A_{\boldsymbol{b}_1} \cup A_{\boldsymbol{b}_2}$ is the entire space of $\boldsymbol{s}$, and where $P_{\boldsymbol{b}_1} = \text{Prob}[A_{\boldsymbol{b}_1}]$ and $P_{\boldsymbol{b}_2} = \text{Prob}[A_{\boldsymbol{b}_2}]$. For many cases, evaluating (235) is a difficult and tedious job. When $M = 2$, however, we can derive a closed-form formula for (235) without difficulty, and the optimal choice does not require Monte-Carlo simulations, which generate $\{s^{(m)}\}$ to calculate (235).

Example 7.8. On a 2×2 Rayleigh-fading channel, we consider Binary Search: $\mathcal{B} = \{\boldsymbol{b}_1 = [b_1^{(1)}, b_1^{(2)}], \boldsymbol{b}_2 = [b_2^{(1)}, b_2^{(2)}]\}$. For every realization of $\{s^{(m)}\}$, we choose either $\boldsymbol{b}_1$ or $\boldsymbol{b}_2$ by comparing

$$\frac{2^{b_1^{(1)}} - 1}{s^{(1)}} + \frac{2^{b_1^{(2)}} - 1}{s^{(2)}} \underset{\boldsymbol{b}_2}{\overset{\boldsymbol{b}_1}{<}} \frac{2^{b_2^{(1)}} - 1}{s^{(1)}} + \frac{2^{b_2^{(2)}} - 1}{s^{(2)}}. \tag{236}$$

The boundary of choice can be expressed as:

$$s^{(2)} = K s^{(1)}, \tag{237}$$

where

$$K = \frac{2^{b_2^{(2)}} - 2^{b_1^{(2)}}}{2^{b_1^{(1)}} - 2^{b_2^{(1)}}} \leq 1. \tag{238}$$

Then, the average SNR requirement is:

$$\rho = \mathbb{E}\left[\frac{2^{r^{(1)}} - 1}{s^{(1)}} + \frac{2^{r^{(2)}} - 1}{s^{(2)}}\right], \tag{239}$$

where $r^{(1)}$ is chosen between $\{b_1^{(1)}, b_1^{(2)}\}$ and $r^{(2)}$ is chosen between $\{b_2^{(1)}, b_2^{(2)}\}$. If the fading is Rayleigh, the joint PDF of $\{s^{(m)}\}$ is given in (113). Then, the average SNR

Table 5: Samples of Binary Search and Fixed Allocation strategies optimized to $M \times M$ Rayleigh fading for $M \in \{2, 4, 6\}$.

		Binary Search	Fixed Allocation
$M = 2$	$B = 2$	[2,0],[1,1]	[2,0]
	$B = 4$	[4,0],[3,1]	[4,0]
	$B = 6$	[6,0],[5,1]	[6,0]
	$B = 8$	[8,0],[6,2]	[8,0]
$M = 4$	$B = 4$	[3,1,0,0],[2,2,0,0]	[3,1,0,0]
	$B = 8$	[4,3,1,0],[5,3,0,0]	[4,3,1,0]
	$B = 12$	[5,4,3,0],[6,4,2,0]	[6,4,2,0]
	$B = 16$	[7,5,4,0],[7,6,3,0]	[7,6,3,0]
$M = 6$	$B = 8$	[3,3,2,0,0,0],[4,3,1,0,0,0]	[4,3,1,0,0,0]
	$B = 12$	[5,4,2,1,0,0],[4,4,3,1,0,0]	[5,4,2,1,0,0]
	$B = 18$	[6,5,3,2,0,0],[5,5,4,2,0,0]	[6,5,3,2,0,0]
	$B = 24$	[6,5,4,3,2,0],[7,5,4,3,1,0]	[7,5,4,3,1,0]

requirement becomes:

$$\begin{aligned}\rho &= (b_1^{(1)} - 1)\left\{2\log(K+1) - \left(\frac{2K}{K+1}\right)^2\right\} \\ &+ (b_2^{(1)} - 1)\left\{\left(\frac{K-1}{K+1}\right)^2 + 2\left(\frac{K-1}{K+1}\right) + 2\log\left(\frac{2}{K+1}\right)\right\} \\ &+ (b_2^{(2)} - 1)\left\{1 + \frac{K-3}{(K+1)^2} - \frac{1}{K+1} + 2\log\left(\frac{K+1}{2K}\right)\right\}. \qquad (240)\end{aligned}$$

The derivation of (240) is in Section 7.6. ♦

Table 5 summarizes the optimal choice for Binary Search and Fixed Allocation for $M \in \{2, 4, 6\}$ when $\beta = 3/4$ and $B_{\max} = 8$. In the table, a short-hand notation, $[r^{(1)}/\beta, \ldots, r^{(M)}/\beta]$, is used. For example, an allocation denoted as $[4, 3]$ means $r^{(1)} = 4\beta$ and $r^{(2)} = 3\beta$. We can confirm that the optimal choice in Table 5 coincides with the results by deleting most infrequently used candidates in Example 7.4, but there are a few exceptions especially for $M = 2$.

Advantages of proposed Binary Search and Fixed Allocation include

- A significant reduction in complexity.
- No increase in complexity as M grows.

- Applicable to any constraint on rate (*e.g.* any β or $B_{\max}$).

Complexity reduction is quite impressive when compared to the full-search strategy. For example, only two calculations of (221) are required for Binary Search in contrast to 51 calculations required for a full search when $M = 6$ and $R = 9$. Especially low in complexity is Fixed Allocation, where no search is required for bit allocation. The complexity reduction will be particularly valuable when these ideas extend to a frequency-selective channel, as will be discussed in Section 7.5.

On the other hand, there are drawbacks. The optimal choice of Binary Search and Fixed Allocation must be matched to fading statistics, otherwise a mismatch would incur a significant penalty. We will discuss the mismatch problem in Section 7.3. It is clear that both Binary Search and Fixed Allocation, though candidates in $\mathcal{B}_1$ and $\mathcal{B}_1$ are optimally selected, are suboptimal to the full-search allocation ($\mathcal{B}$). However, the suboptimality is not a serious problem in MIMO fading channels. We show this by the following example.

Example 7.9. We consider an $M \times M$ Rayleigh-flat-fading channel for $M \in \{2, 4, 6\}$. Suppose that each rate is restricted to discrete values with $\beta = 3/4$ and $B_{\max} = 6$. We evaluate the performance of Binary Search and Fixed Allocation strategies by using (230) with $\Gamma = 1$. Figure 44 plots the achievable rate versus SNR. As a benchmark, $C_{0,\mathrm{OPT}}$ and $C_{0,\mathrm{FIX}}$ from Chapter 6.[4] Note that Binary Search and Fixed Allocation are bounded by C_0 and $C_{0,\mathrm{FIX}}$, respectively, in Figure 44. We can see that the performance penalty of Binary Search and Fixed Allocation relative to the full-search strategy is nonzero, but very small. An exception is Fixed Allocation for $M = 2$, which incurs a large penalty, as expected from its lower bound $C_{0,\mathrm{FIX}}$, which has a lower multiplexing order ($\tilde{M} = 1$) than $C_{0,\mathrm{OPT}}$ or $C_{0,\mathrm{FUSE}}$, as discussed in Chapter 6. For $M = 4$ and $M = 6$, however, Fixed Allocation shows nearly optimal performance.

[4] As we assume flat fading, we thus have $C_{0,\mathrm{OPT}} = C_{0,\mathrm{FUSE}}$

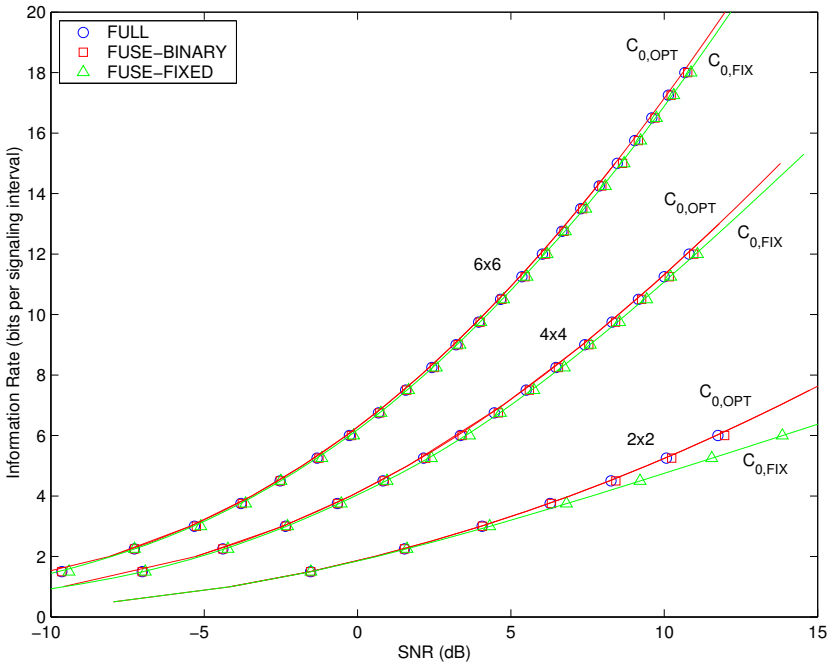

Figure 44: Achievable rates of Binary Search and Fixed Allocation on $M \times M$ flat fading for $M \in \{2, 4, 6\}$. Also plotted are $C_{0,\mathrm{OPT}}$, $C_{0,\mathrm{FUSE}}$, and $C_{0,\mathrm{FIX}}$ as references.

To emphasize the penalty, Figure 45 and Figure 46 illustrate the SNR penalty of Binary Search and Fixed Allocation compared to the full-search strategy, respectively. The Binary Search strategy is only worse by at most 0.2 dB, and the penalty decreases as M grows. In the case of Fixed Allocation, there is a distinct performance degradation for $M = 2$, but the SNR penalty is only less than 0.3 dB for $M = 4$ and $M = 6$. ♦

This section proposed Fixed Search and Fixed Allocation bit-allocation strategies, which approach $C_{0,\mathrm{OPT}}$ for a moderate M.

7.3 Robust Allocation Strategy

In the previous section, we have seen that a mismatch in fading statistics can incur a significant penalty when using Binary Search or Fixed Allocation. Now, we consider a bit-allocation strategy that shows robust performance for various types of fading. An

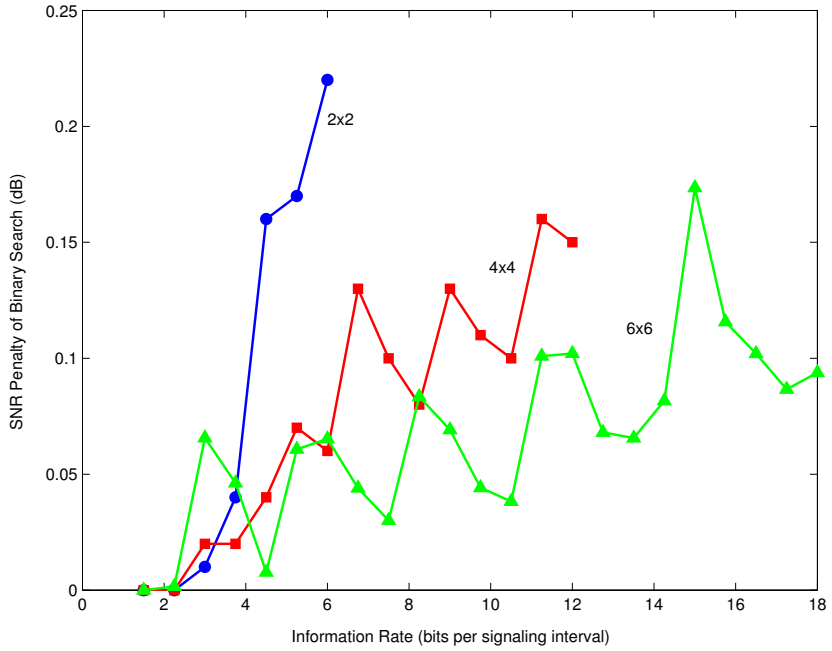

Figure 45: SNR penalty of Binary Search relative to the full-search strategy on $M \times M$ flat fading for $M \in \{2, 4, 6\}$.

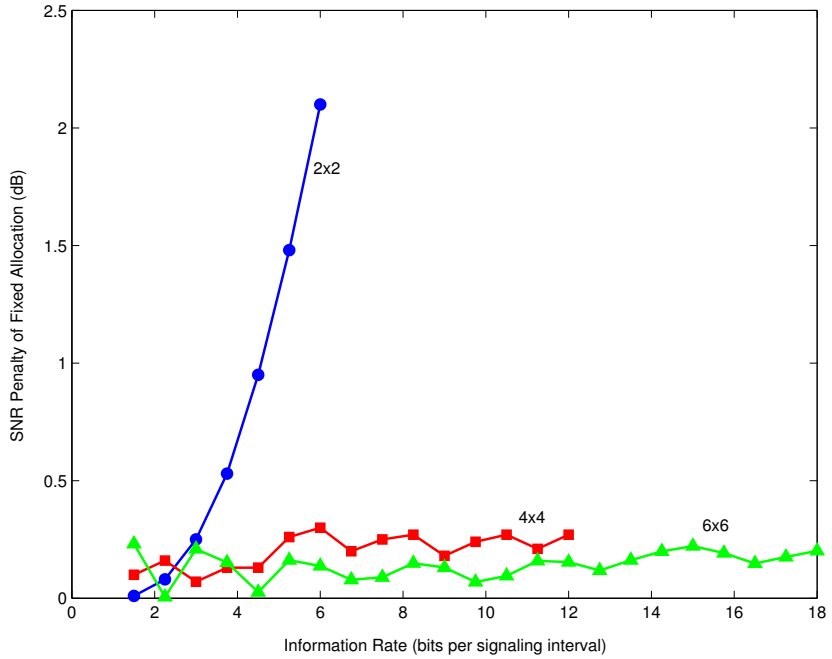

Figure 46: SNR penalty of Fixed Allocation relative to the full-search strategy on $M \times M$ flat fading for $M \in \{2, 4, 6\}$.

intuitive solution is increasing the size of search set, but there is a tradeoff between robustness and complexity. We will show that a few additions, incurring a small increase in complexity, is sufficient for a robust allocation.

To reflect the change in fading statistics, we consider Ricean fading. Given Ricean factor K, which accounts for the line-of-sight component, each channel is generated according to $\mathcal{N}(\mu,\sigma) + j\mathcal{N}(\mu,\sigma)$, where $\mu = \sqrt{\frac{K}{2(K+1)}}$ and $\sigma = \sqrt{\frac{1}{2(K+1)}}$ [53]. We consider an example on Ricean fading with $K = 4.45$.

Example 7.10. For a 4×4 spatially uncorrelated MIMO, channel generation is according to Ricean fading with $K = 4.45$. Suppose that Binary Search optimized to Rayleigh fading ($K = 0$) incurs a large SNR penalty, more than 2 dB for some R, as illustrated in Figure 47. We create a robust search set consisting of three members:

$$\mathcal{B}_{\text{robust}} = \{\boldsymbol{b}_{\text{FIX},K=0}, \boldsymbol{b}_{\text{FIX},K=2.41}, \boldsymbol{b}_{\text{FIX},K=6.46}\}, \tag{241}$$

where $\boldsymbol{b}_{\text{FIX},K}$ is the optimal choice of Fixed Allocation for Ricean fading with K. Therefore, for K ranging from 0 to 6.46, the allocation based on $\mathcal{B}_{\text{robust}}$ is expected to show consistent performance. Figure 47 illustrates the SNR penalty of $\mathcal{B}_{\text{robust}}$ when $K = 4.45$. As a benchmark, Figure 47 also shows the SNR penalty of Fixed Allocation and Binary Search optimized to actual fading statistics ($K = 4.45$). Clearly, the robust allocation performs very well, whose SNR penalty is less than 0.2 dB. ♦

Example 7.10 shows that a search set with three candidates is sufficient to cover Ricean fading from $K = 0$ to $K = 6.46$. Generally, a search set with three or four candidates shows consistent performance for various R and M. Thus, the mismatch problem can be solved with a slight sacrifice in complexity.

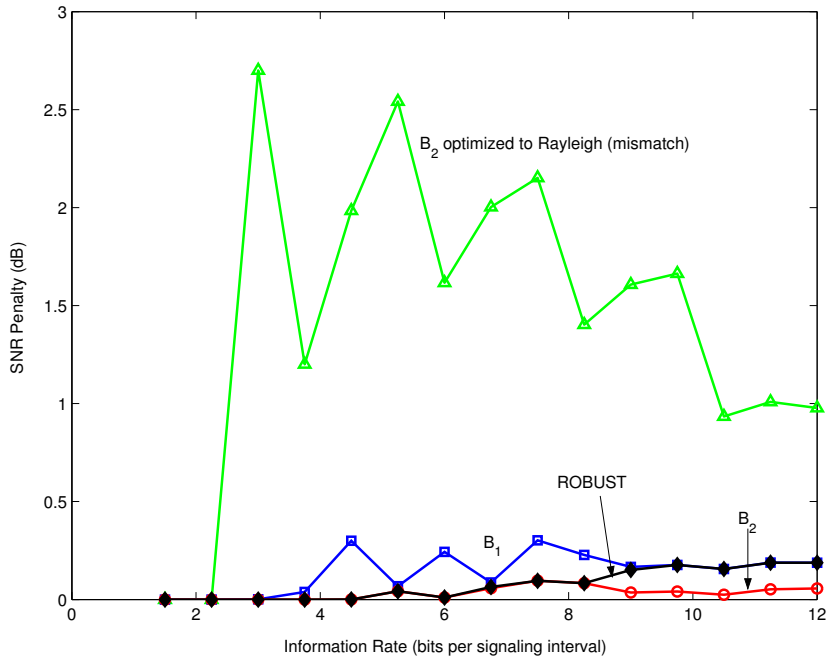

Figure 47: Fading mismatch of Binary Search optimized to Rayleigh fading and performance of a robust allocation on a 4×4 Ricean fading channel with $K = 4.45$.

7.4 Simple Allocation Algorithms for DMT

Before we move to MIMO-OFDM, we briefly review allocation algorithms for the discrete multitone (DMT) in the DSL applications.[5] In DMT, the full-search allocation is nearly impossible since the number of tones is usually large. To replace the full-search allocation, many practical algorithms have been proposed [14, 17, 19]. Among those, we describe the algorithm in [14].

When $\{s_n^{(m)}\}$ are given, the goal is to find $\{r_n^{(m)}\}$ with a finite step size β so as to satisfy $\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = R$. A good starting point for $\{r_n^{(m)}\}$ is in the form of

$$r_n^{(m)} = \left[\lfloor \log_2(g_n^{(m)}) \rfloor_\beta + i\right]_0^{B_{\max}}, \tag{242}$$

where $i \in \{\ldots, -2\beta, -\beta, 0, \beta, 2\beta \ldots\}$, and $\lfloor x \rfloor_\beta = \beta \lfloor x/\beta \rfloor$. We initialize by setting $i = 0$ where $g_n^{(m)} = s_n^{(m)}/(\Gamma N_0)$ and $[x]_a^b = \min(a, \max(x, b))$ when $a < b$. Depending on Γ, N_0, or $s_n^{(m)}$, $\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)}$ can be larger or smaller than R. We decrease

[5] DMT is a multicarrier system, similar to OFDM, developed for the DSL technology [19].

or increase i, respectively, until we find i_B such that

$$i_B = \max\left\{ i : \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left[\lfloor \log_2(g_n^{(m)})\rfloor_\beta + i\right]_0^{B_{\max}} \right\}. \tag{243}$$

In fact, an efficient way to initialize i by grouping $\log_2(g_n^{(m)})$ is proposed in [14].

Once i_B is found, we need to increment $r_n^{(m)}$ until $\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = R$ holds. We define energy increase for the (m,n) channel as

$$\Delta e_n^{(m)} = e_n^{(m)}(r_n^{(m)} + \beta) - e_n^{(m)}(r_n^{(m)}), \tag{244}$$

where $r_n^{(m)}$ is the current rate. We calculate $\Delta e_n^{(m)}$ for all (m,n) with $r_n^{(m)} + \beta < B_{\max}$ and pick up (m,n) that has the smallest $\Delta e_n^{(m)}$. Repeat this process until $\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = R$ is satisfied.

7.5 Bit Allocation for MIMO-OFDM

In Section 7.2, Binary Search and Fixed Allocation strategies are proposed for MIMO flat-fading channels. Now, we consider the bit allocation for MIMO-OFDM in (8). A distinct difference from the flat-fading case is that the number of scalar channels (MN) is usually much larger than M scalar channels for flat fading. Thus, it is MIMO-OFDM where the complexity reduction is really necessary.

Clearly, a full search would be practically impossible since the size of full-search set is immense. Unlike flat fading, we cannot simply apply Binary Search or Fixed Allocation strategies to MIMO-OFDM since scalar channels across frequency do not have the properties we have used for the spatial channels of MIMO flat-fading channel. Even if MIMO-OFDM had such properties, the optimal choice for Binary Search and Fixed Allocation would be too high in complexity.

For each tone, however, it has the same statistics as MIMO flat fading. Hence, Binary Search and Fixed Allocation are readily applicable to each tone with an independent rate constraint, namely the FUSE constraint in Section 6.2. We already

examined this case without granularity constraint in Proposition 6.1. Near optimality of the FUSE constraint in Section 6.4 and Section 6.5.1 guarantees good performance of Binary Search and Fixed Allocation for MIMO-OFDM. Actually, $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ serve as upper bounds for Binary Search and Fixed Allocation, respectively.

Example 7.11. Figure 48 illustrates the performance of the Binary Search and Fixed Allocation strategies with the FUSE constraint on a $M \times M$ spatially uncorrelated MIMO channel with $L = 4$, $N = 64$, and $M \in \{2, 4, 6\}$. We use (230) with Γ to calculate SNR required by Binary Search and Fixed Allocation. The Binary Search strategy with the FUSE constraint (marked as squares) incurs an SNR penalty of between 0.4 dB and 0.9 dB when $M = 2$ compared to the iterative algorithm by Campello (marked as circles), described in the previous section, whereas the penalty is negligible for $M = 4$ and $M = 6$. The Fixed Allocation strategy with the FUSE constraint (marked as triangles) is also nearly optimal for $M = 4$ and $M = 6$, while its penalty can be large for $M = 2$. As shown in Figure 48, the iterative algorithm of [14] is tightly bounded by $C_{0,\mathrm{OPT}}$, while Binary Search and Fixed Allocation are bounded by $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$, respectively. Thus, the infinite-precision water-filling is a good indicator for the performance of practical bit-allocation strategies. ⧫

This section showed that Binary Search and Fixed Allocation strategies combined with the FUSE constraint in Section 6.2 perform very well in MIMO-OFDM, but the complexity required by them is remarkably small.

7.6 Derivation of (240)

The joint PDF of s_1 and s_2 is $f_{\boldsymbol{s}}(x_1, x_2) = (x_1 - x_2)^2 e^{-x_1} e^{-x_2}$ for $x_1 \geq x_2$. Then, the SNR requirement in (230) is

$$\rho \;=\; \Gamma \int_0^\infty \int_0^x \left(\frac{2^{r^{(1)}} - 1}{x} + \frac{2^{r^{(2)}} - 1}{y} \right) (x - y)^2 e^{-x} e^{-y} dy dx$$

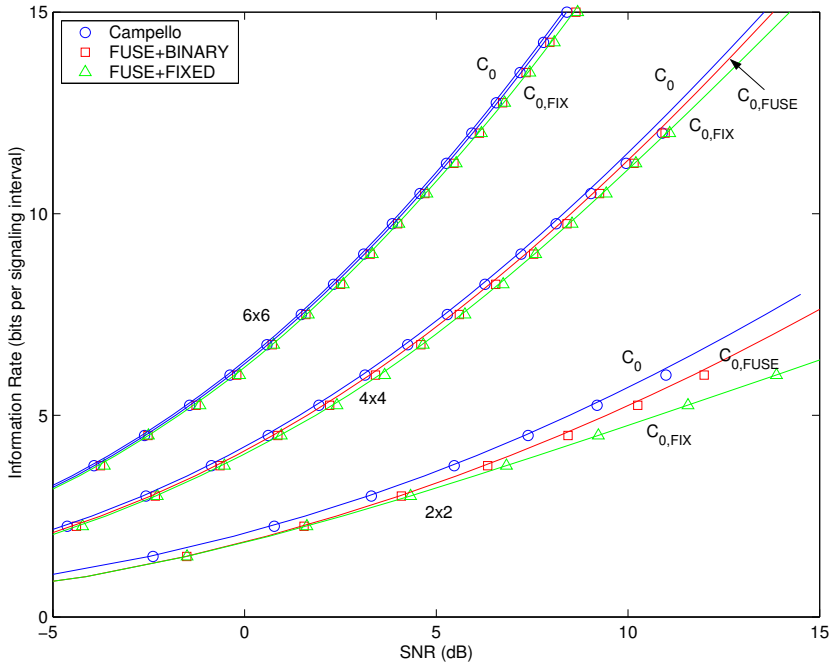

Figure 48: Performance of Binary-Search and Fixed Allocation combined with the FUSE constraint (denoted as FUSE-BINARY and FUSE-FIXED, respectively) on a $M \times M$ spatially uncorrelated Rayleigh fading channel with $M \in \{2, 4, 6\}$ when $L = 4$ and $N = 128$.

$$
\begin{aligned}
&= \Gamma \underbrace{\int_0^{\infty} \int_0^{Kx} \left(\frac{2^{b_1^{(1)}} - 1}{x} \right) (x-y)^2 e^{-x} e^{-y} dy dx}_{I_{1,1}} \\
&+ \Gamma \underbrace{\int_0^{\infty} \int_0^{Kx} \left(\frac{2^{b_1^{(2)}} - 1}{y} \right) (x-y)^2 e^{-x} e^{-y} dy dx}_{I_{1,2}} \\
&+ \Gamma \underbrace{\int_0^{\infty} \int_{Kx}^{x} \left(\frac{2^{b_2^{(1)}-1} - 1}{x} \right) (x-y)^2 e^{-x} e^{-y} dy dx}_{I_{2,1}} \\
&+ \Gamma \underbrace{\int_0^{\infty} \int_{Kx}^{x} \left(\frac{2^{b_2^{(2)}-1} - 1}{y} \right) (x-y)^2 e^{-x} e^{-y} dy dx}_{I_{2,2}} . \qquad (245)
\end{aligned}
$$

However $I_{1,2} = 0$ unless $b_1^{(2)} = 0$ since

$$\int_0^{Kx} \frac{e^{-y}}{y} dy \to \infty \qquad (246)$$

if $K > 0$. Therefore, we always set $b_1^{(2)} = 0$ on a 2×2 Rayleigh-fading channel. Other integrals can be evaluated as follows

$$I_{1,1} = (b_1^{(1)} - 1) \left\{ 2\log(K+1) - \left(\frac{2K}{K+1} \right)^2 \right\}, \tag{247}$$

$$I_{2,1} = (b_2^{(1)} - 1) \left\{ \left(\frac{K-1}{K+1} \right)^2 + 2\left(\frac{K-1}{K+1} \right) + 2\log\left(\frac{2}{K+1} \right) \right\}, \tag{248}$$

and

$$I_{2,2} = (b_2^{(2)} - 1) \left\{ 1 + \frac{K-3}{(K+1)^2} - \frac{1}{K+1} + 2\log\left(\frac{K+1}{2K} \right) \right\}, \tag{249}$$

where we use

$$\int_0^\infty \frac{e^{-x}}{x} dx - \int_0^\infty \frac{e^{-ax}}{x} = \log(a) \tag{250}$$

and

$$\int_0^\infty x^2 e^{-x} \underbrace{\left(\int_{Kx}^{x} \frac{e^{-t}}{t} dt \right)}_{u} dx = \frac{5}{4} - \frac{1}{(K+1)^2} - \frac{2}{K+1} + 2\log\left(\frac{K+1}{2K} \right), \tag{251}$$

where

$$\frac{du}{dx} = \frac{e^{-x} - e^{-Kx}}{x}. \tag{252}$$

CHAPTER 8

BIT-ERROR RATE PERFORMANCE

So far, we have examined the information-theoretical aspects of MIMO fading channels with an emphasis on the case where the transmitter knows CSI (a closed-loop system). In this chapter, we turn our attention to a realistic communication system with discrete and finite constellations and practical channel codes. We use the Binary Search and Fixed Allocation strategies, proposed in Chapter 7, to distribute rate and power to measure the performance of a closed-loop system in terms of bit-error rate (BER). Recall that there are two types of energy constraint: the short-term constraint in (51) and the long-term constraint in (50). First, we begin with the short-term constraint, for which we compare the BER performance with and without CSI at the transmitter. Then, we move to the long-term constraint to see the advantage of controlling power at the transmitter with the long-term constraint.

8.1 Short-Term Energy Constraint

With the short-term constraint, zero outage probability is unattainable even when the transmitter has CSI. Thus, the performance is limited by the outage probability.

8.1.1 Uncoded Performance of Flat-Fading MIMO

This section investigates the BER performance on a flat-fading MIMO channel without an outer code. With CSI known at the transmitter, recall that eigenbeamforming creates a bank of M scalar channels:

$$\begin{aligned} z^{(1)} &= \sqrt{s^{(1)}}a^{(1)} + w^{(1)} \\ &\vdots \quad \vdots \qquad\qquad . \\ z^{(M)} &= \sqrt{s^{(M)}}a^{(M)} + w^{(M)} \end{aligned} \tag{253}$$

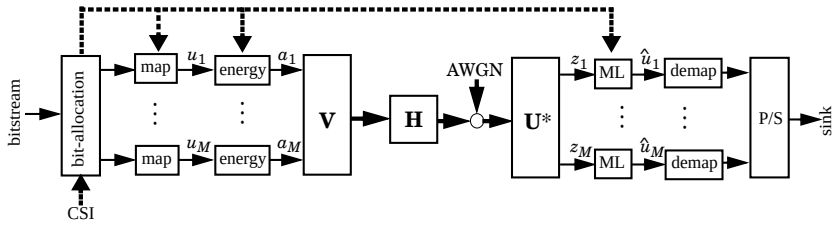

Figure 49: Block diagram for flat-fading MIMO.

Let $e^{(m)} = \mathbb{E}\left[|a^{(m)}|^2 \left| \{s^{(m)}\} \right.\right]$ denote the energy for the mth scalar channel. Then, by the short-term constraint, $\sum_{m=1}^{M} e^{(m)} = \bar{E}$ is always satisfied. If N_0 is the variance of $w^{(m)}$, SNR can be written as $\rho = \bar{E}/N_0$. From Chapter 4 and Chapter 5, the diversity and spatial multiplexing orders of (253) are $M_T M_R$ and M, respectively.

The block diagram for the uncoded-modulation system is illustrated in Figure 49. At the transmitter, the bit-allocation block takes R message bits as an input and divides them into a partition $\{r^{(1)}, r^{(2)}, \ldots, r^{(M)}\}$ satisfying $\sum_{m=1}^{M} r^{(m)} = R$, where $r^{(m)}$ is the rate for the mth scalar channel. We use uncoded QAM constellations, so that $r^{(m)}$ must be chosen from $\{0, 2, \ldots, 6\}$, that is, $B_{\text{MAX}} = 6$ and $\beta = 1$.[1] In this experiment, we use Full-Search Allocation in Definition 7.2 and Fixed Allocation in Section 7.2 to determine $\{r^{(m)}\}$. For each m, $r^{(m)}$ bits are mapped into $a^{(m)}$ from QAM constellations. The energy for each $a^{(m)}$ is initially determined by $e^{(m)} = (2^{r^{(m)}} - 1)/s^{(m)}$. Then, we normalize $\{e^{(m)}\}$ such that $\sum_{m=1}^{M} e^{(m)} = \bar{E}$ is satisfied [17].

In Figure 49, the detection of $\{z^{(m)}\}$ is straightforward since eigenbeamforming removes the spatial interference of a MIMO channel. Since $\{r^{(m)}\}$ and $\{e^{(m)}\}$ are known, the receiver independently detects $a^{(m)}$ from $z^{(m)}$ in the maximum-likelihood sense. The detection complexity is linearly proportional to M.

To compare with the case where the transmitter is ignorant of CSI, we consider

[1] When $r^{(m)} = 3$, hexagonal constellation instead of 8-QAM is used for better performance. Also we avoid BPSK ($r^{(m)} = 1$) since QPSK (4-QAM) conveys more information while exhibits the same performance at a given SNR per bit.

orthogonal STBC [57]. We mean by orthogonal that symbols in a code can be detected independently, so that the detection complexity linearly increases with M. Equivalently, the input/output relation can be written as

$$z = \frac{||\mathbf{H}||_F}{\sqrt{M}} a + w, \tag{254}$$

where $\mathbf{H}$ is the channel matrix, a is the transmitted symbol, and w is additive white Gaussian noise with variance N_0. From (254), we know that (1) the maximum diversity order ($M_T M_R$) is achieved; (2) the spatial multiplexing order is unity, less than M of the maximum order. According to [57], an orthogonal STBC with the spatial multiplexing order of unity exists only for $M_T = 2$, known as Alamouti's code [1]. For $M_T > 2$, the spatial multiplexing order must be sacrificed to maintain orthogonality. However, for simplicity, we assume that (254) holds for any M_T.

We investigate the BER performance of eigenbeamforming in (253) with the Full-Search allocation based on the full-search set in Definition 7.2 for $M \times M$ Rayleigh fading. We also examine the performance of orthogonal STBC in (254) as a benchmark. For spectral efficiencies in the range $R \in \{2, 3, 4, 5, 6\}$, Figure 50 and Figure 51 show BER results for $M = 2$ and $M = 4$, respectively. For both $M = 2$ and $M = 4$, we can confirm that both eigenbeamforming and STBC have the same diversity order (the slope of BER curves). However, for a given BER, eigenbeamforming requires less SNR per bit than STBC. This advantage is more conspicuous at higher R and/or larger M due to the difference in spatial multiplexing orders.

To emphasize the SNR advantage of eigenbeamforming, we estimate an *SNR per bit requirement*, which we define as ρ/R, so as to achieve a BER of 10^{-4}. The results are summarized in Figure 52 for $M = 2$ and $M = 4$. We also plot the SNR per bit requirement for eigenbeamforming with Fixed Allocation in Section 7.2 with $\{[2,0],[3,0],[4,0],[5,0],[6,0]\}$ for $M = 2$ and $\{[2,0,0,0],\ [3,0,0,0],\ [2,2,0,0],\ [3,2,0,0],\ [4,2,0,0]\}$ for $M = 4$.

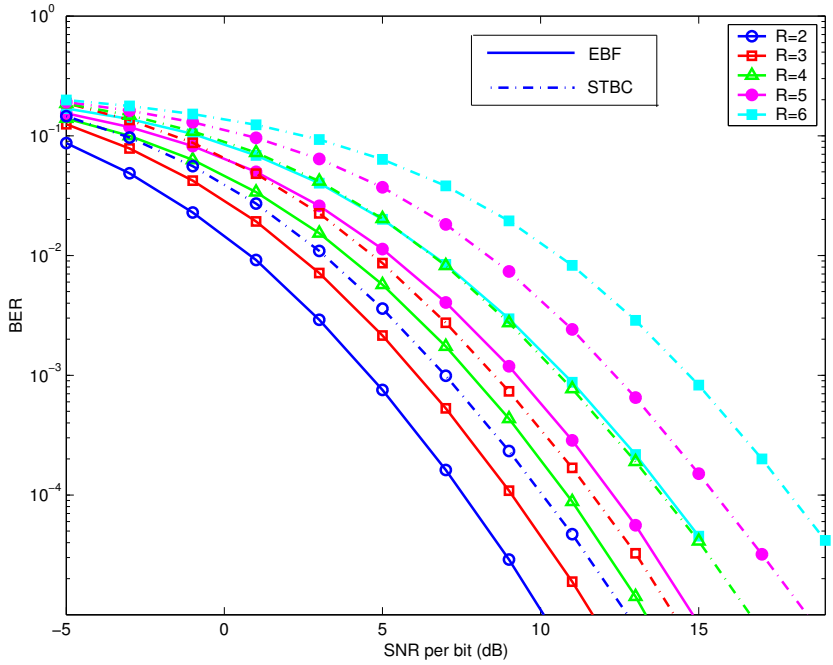

Figure 50: Uncoded BER performance of eigenbeamforming and orthogonal STBC on 2×2 Rayleigh fading.

For $M = 2$, eigenbeamforming with Fixed Allocation has an advantage of approximately 2.1 dB over STBC. Note that this advantage is uniform for all R. This is because Fixed Allocation uses the first scalar channel only and thus it has the same diversity order ($M_T M_R$) and the spatial multiplexing order (unity) as STBC. When Full-Search Allocation is used, which achieves a spatial multiplexing order of M, the advantage of eigenbeamforming becomes larger than STBC. In this case, the advantage grows with R, from 2.2 dB at $R = 2$ to 3.6 dB at $R = 6$.

For $M = 4$, the advantage of eigenbeamforming is more obvious. The SNR (vertical) gap between eigenbeamforming and STBC apparently increases with R, as much as 8.4 dB at $R = 6$. An interesting point in Figure 52 is that Fixed Allocation performs very closely to Full-Search Allocation when $M = 4$, while their difference is more obvious when $M = 2$, as expected from Figure 46, where Fixed Allocation is nearly optimal when M is four or more.

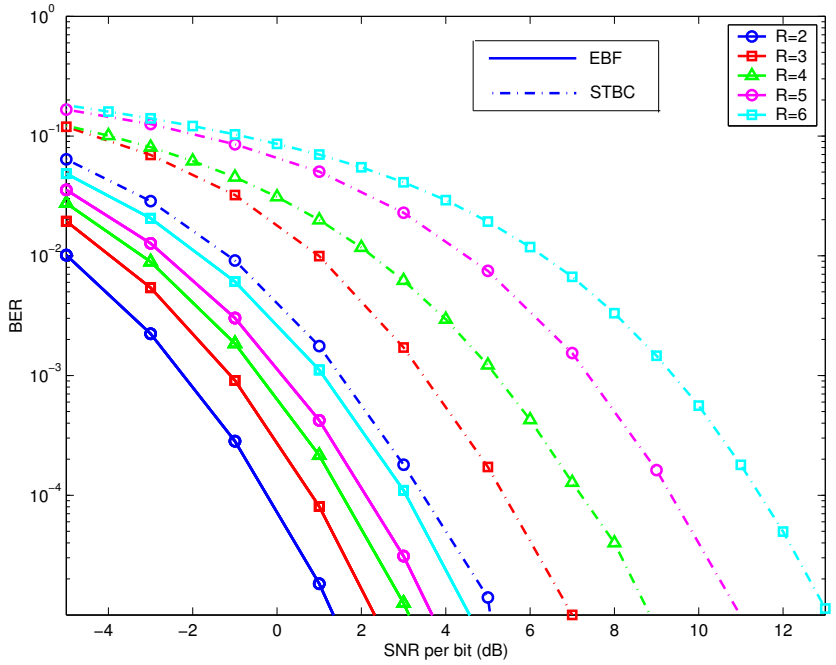

Figure 51: Uncoded BER performance of eigenbeamforming and orthogonal STBC on 4×4 Rayleigh fading.

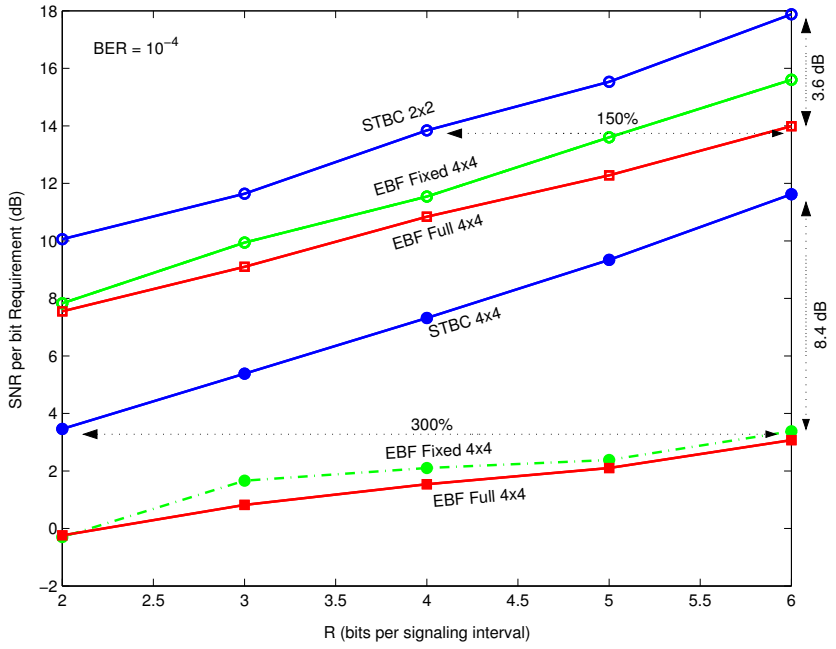

Figure 52: SNR per bit requirement with the short-term energy constraint on 2×2 and 4×4 Rayleigh fading.

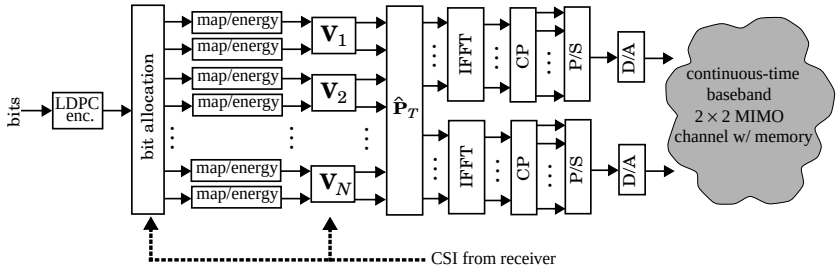

Figure 53: Block diagram for a MIMO-OFDM system with an LDPC outer code at the transmitter.

8.1.2 MIMO-OFDM with an Outer Code

We extend the flat-fading study in the previous section to MIMO-OFDM with a low-density parity-check (LDPC) code [6]. Here, we consider more realistic situations. Channel generation follows the indoor channel model [61], where six uncorrelated Rayleigh faded taps are obtained form the modified Jakes simulator at a Doppler frequency of 48.33 Hz [53], which generates a continuous-time baseband 2×2 channel with memory. The 2×2 channel with memory is transformed into a set of memoryless 2×2 channels by OFDM, as described in Figure 1.

The block diagrams for the transmitter and receiver are illustrated in Figure 53 and Figure 54, respectively, for a closed-loop 2×2 MIMO system. We also consider Alamouti's code, for which the bit-allocation block is replaced by serial-to-parallel (S/P) converter, and Alamouti's encoders are used instead of the transmit filter bank $\{\mathbf{V}_n\}$ in the transmitter. At the receiver, Alamouti's combiners replace the receive filter banks $\{\mathbf{U}_n^*\}$. Note that Alamouti's code requires two OFDM blocks for encoding and combining.

An OFDM block consists of $N = 1024$ tones with 64 intervals for cyclic prefix. Two training blocks are sent for every 10 data blocks for channel estimation [40]. We use regular LDPC codes, whose length covers an OFDM block, which are applied across space and frequency to provide coding gain as well as frequency diversity.

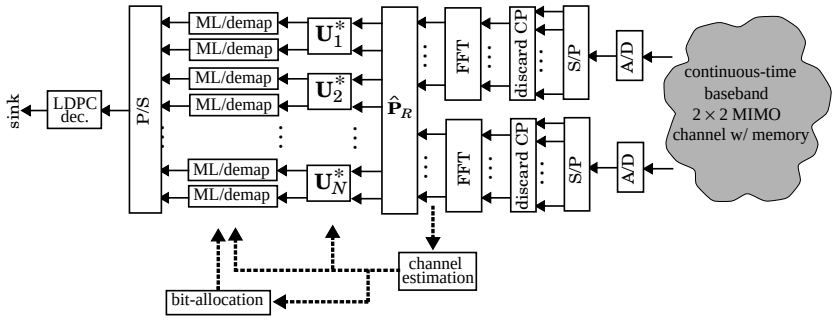

Figure 54: Block diagram for a MIMO-OFDM system with an LDPC outer code at the receiver.

For eigenbeamforming, Fixed Allocation ($\{[1,0],[4,0],[6,0]\}$) is used for BPSK, 16-QAM, 64-QAM constellations. BER results are shown in Figure 55 for code rates of 1/2, 3/4, and 7/8 [31]. For each code rate, the advantage of eigenbeamforming (Fixed Allocation) is clear over STBC (Alamouti's code).

8.2 Long-Term Energy Constraint

This section examines the BER performance of eigenbeamforming with the long-term energy constraint in (50). We use the discrete-time baseband model in Figure 1 with the uniform power profile. Each $\mathbf{H}_n$ is transformed into a bank of scalar channels in Figure 3 by the transmit and receive filter banks $\{\mathbf{V}_n\}$ and $\{\mathbf{U}_n^*\}$, respectively.

The overall diagram is illustrated in Figure 56. The LDPC block takes 768 message bits and produce 1024 coded bits ($\beta = 3/4$ rate). Bit-allocation block determines $\{r_n^{(m)}\}$ satisfying $\sum_{n=1}^{N}\sum_{m=1}^{M} r_n^{(m)} = 768$. We only use square QAM constellations (4-QAM, 16-QAM, and 64-QAM), so that $r_n^{(m)}$ must be chosen from $\{0, 2\beta, 4\beta, 6\beta\}$. For each m and n, $r_n^{(m)}/\beta$ coded bits are mapped into a complex symbol $a_n^{(m)}$, chosen from corresponding constellation. The energy for each symbol is given by

$$e_n^{(m)} = \mathbb{E}\left[|a_n^{(m)}|^2 \middle| \{s_n^{(m)}\}\right] = \Gamma \frac{2^{r_n^{(m)}} - 1}{s_n^{(m)}/N_0}. \tag{255}$$

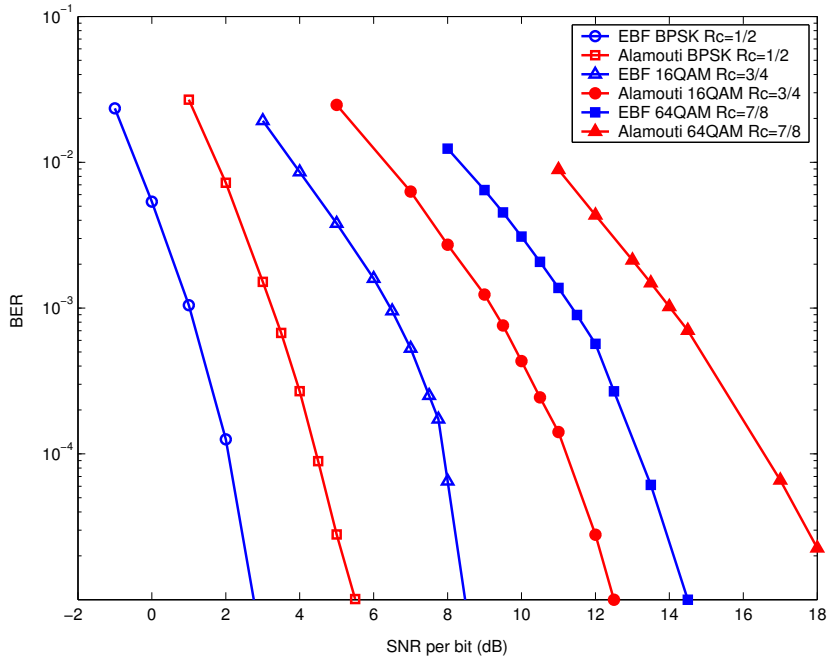

Figure 55: BER results for eigenbeamforming with Fixed Allocation and STBC in 2×2 MIMO-OFDM with LDPC codes.

Then, SNR is

$$\rho = \Gamma \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{2^{r_n^{(m)}}-1}{s_n^{(m)}}\right], \tag{256}$$

where Γ is the SNR gap defined in Section 7.1. If the channel code is ideal, $\Gamma = 1$ achieves zero BER. In practice, the transmitter can control SNR via Γ.

At the receiver, the log-likelihood ratio (LLR) is calculated from a collection of $z_n^{(m)}$, where $z_n^{(m)} = \sqrt{s_n^{(m)}}a_n^{(m)} + w_n^{(m)}$ from (39), and is fed into LDPC decoder.

Example 8.1. Consider 4×4 $N = 128$ MIMO-OFDM with the long-term energy constraint in (50) at $R = 6$ bits per signaling interval. We generate $100,000$ independent $\{s_n^{(m)}\}$ to obtain BER results of eigenbeamforming with (1) Campello bit allocation [14] over MN scalar channels; (2) Binary-Search Allocation with the FUSE constraint (FUSE-BINARY); (3) Fixed Allocation with the FUSE constraint (FUSE-FIXED). From Chapter 6, we expect that the performance of Campello is limited by C_0, corresponding to water-filling over MN scalar channels. Also, FUSE-BINARY

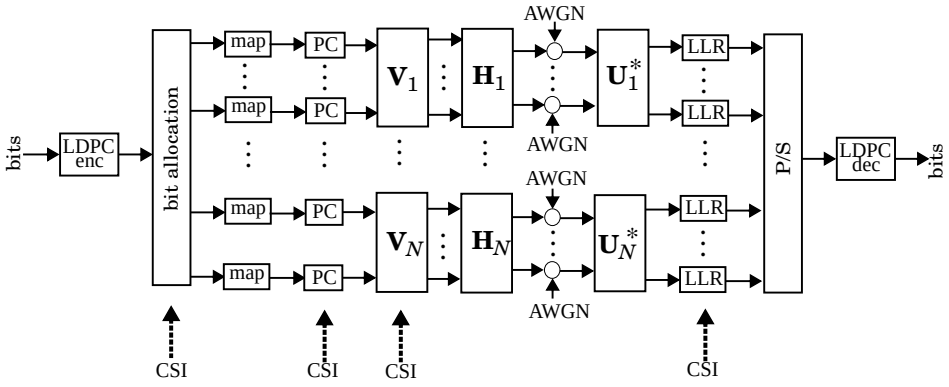

Figure 56: Block diagram for MIMO-OFDM with power control and an LDPC outer code.

and FUSE-FIXED have $C_{0,\mathrm{FUSE}}$ and $C_{0,\mathrm{FIX}}$ as their limits.

Figure 57 illustrates BER results when channel has $L = 0$ memory (flat fading). As a benchmark, Figure 57 also shows BER results when the short-term energy constraint in (51) is used. The difference between long-term and short-term energy constraint is obvious. At a BER of 10^{-4}, the advantage of long-term constraint is more than 2.6 dB in SNR. We remark that the diversity order for long-term constraint is infinite since it achieves zero outage probability as discussed in Chapter 4. Thus, the slope of BER curve is governed only by LDPC codes. If we use a longer code, we expect a sharper drop in BER curves. On the other hand, the slope for short-term constraint is limited by the diversity order, which is 16 on a 4×4 spatially uncorrelated MIMO channel with $L = 0$ memory.

Among three bit-allocation strategies, there is little difference between Campello and FUSE-BINARY since $C_{0,\mathrm{FUSE}} = C_{0,\mathrm{OPT}}$ when $L = 0$. The 0.1 dB difference between Campello and FUSE-BINARY comes from the fact Binary-Search Allocation uses two candidates, while Campello allocation considers all candidates. Theoretically, to achieve 6 bits per signaling interval ($C_{0,\mathrm{OPT}} = C_{0,\mathrm{FUSE}} = C_{0,\mathrm{FIX}} = 6$), the SNR requirements for $C_{0,\mathrm{OPT}}$, $C_{0,\mathrm{FUSE}}$, and $C_{0,\mathrm{FIX}}$ are $\rho_{0,\mathrm{OPT}} = 3.65$ dB, $\rho_{0,\mathrm{FUSE}} = 3.75$ dB, and $\rho_{0,\mathrm{FUSE}} = 4.1$ dB. Approximately, in Figure 57, the gap between Campello

and FUSE-BINARY matches $\rho_{0,\mathrm{OPT}} - \rho_{0,\mathrm{FUSE}} = 0.1$ dB and the gap between FUSE-BINARY and FUSE-FIXED matches $\rho_{0,\mathrm{FUSE}} - \rho_{0,\mathrm{FIX}} = 0.35$ dB, which match the separations of BER curves in Figure 57. ♦

Example 8.2. In Figure 58, we consider the same environment as in the previous example excepts that the channel has $L = 4$ memory. A distinct difference from Figure 57 is that the performance gap between long-term and short-term constraints is reduced. At a BER of 10^{-4}, the SNR gap is approximately 0.9 dB. Another distinct change is the slope of BER curves (diversity order) for short-term constraint. Owing to $L = 4$ memory, the frequency diversity helps improve the diversity order significantly.

Another observation from Figure 58 is that the BER curve of Campello allocation is shifted to the left, while the BER curves of FUSE-BINARY and FUSE-FIXED remain unmoved compared to Figure 57. The performance improvement of Campello allocation is also from the increase in channel memory. Thus, the gap between Campello and FUSE-BINARY increases from 0.1 dB to 0.38 dB according to the analysis in Chapter 6, which also match the separation in Figure 58 while the gap between FUSE-BINARY and FUSE-FIXED remains the same. ♦

The implications of the above two examples are

- The proposed bit-allocation strategies perform so well. With the long-term constraint, we can approach the zero-outage capacities $C_{0,\mathrm{OPT}}$, $C_{0,\mathrm{FUSE}}$, and $C_{0,\mathrm{FIX}}$ within about 3 dB at a BER of 10^{-4} with a powerful outer channel code.

- The SNR differences in the theoretical results of Figure 44 well match the SNR differences of BER curves in Figure 57 and Figure 58.

- As the diversity order grows from 16 in Figure 57 to 64 in Figure 58, the performance with the short-term constraint approaches the performance with

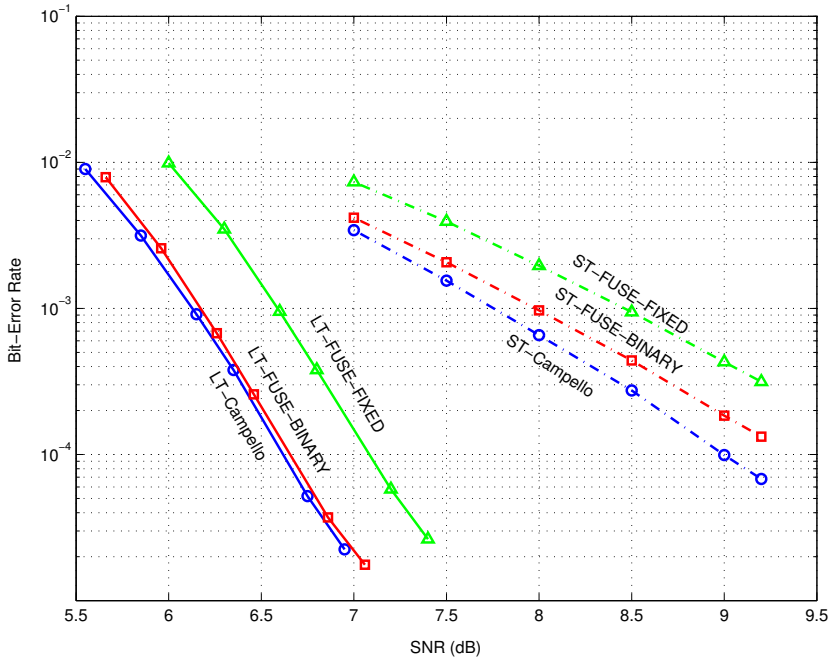

Figure 57: BER results of 4×4 MIMO-OFDM with $N = 128$ and $L = 0$ at $R = 6$ bits per signaling interval, averaged over 10^5 independent channel realizations.

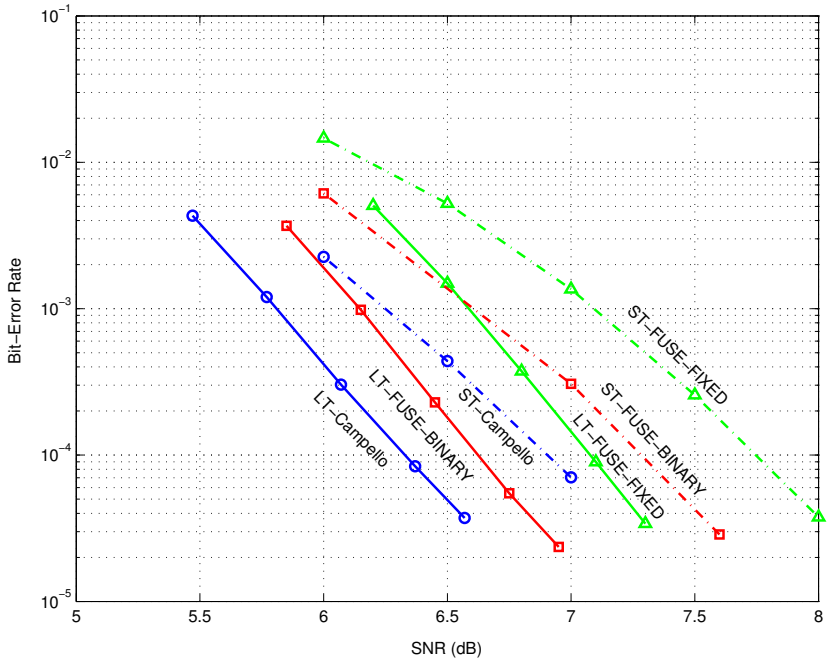

Figure 58: BER results of 4×4 MIMO-OFDM with $N = 128$ and $L = 4$ at $R = 6$ bits per signaling interval, averaged over 10^5 independent channel realizations.

the long-term constraint. In the end, we guess that the performance advantage of the long-term constraint vanishes as the diversity order with the short-term constraint tends to infinity.

CHAPTER 9

CONCLUSIONS

9.1 Summary

In this thesis, we analyzed power allocation with simplifying constraints and proposed simple bit-allocation strategies for a closed-loop MIMO-OFDM system. The conclusion is that the transmitter need not perform adaptive allocation in order to approach the capacity in MIMO-OFDM with eigenbeamforming. Instead, a combination of eigenbeamforming and fixed allocation is sufficient. We proved that fixed allocation is asymptotically optimal as the number of antennas tends to infinity. The convergence is fast, and a nonadaptive MIMO-OFDM system performs well for a moderate number of antennas.

Main contributions include

- First, we examine asymptotic behaviors of outage and ergodic capacity at high SNR in Chapter in terms of geometric means of eigenvalues 4. From the asymptotic results, we confirmed that knowing CSI has little impact on average capacity from Proposition 5.6 and Proposition 5.6. If $M_T \leq M_R$, the advantage completely disappears at high SNR. We also saw that zero-outage capacity grows with the number of channel memory (L) up to the average capacity at high SNR in Proposition 5.7.

- In Chapter 4, we introduce the outage-region capacity, which outperforms the zero-outage capacity by introducing an outage region when the channel is bad. Procedures for the optimal choice of outage region were presented in Section 4.4. From simulations, we saw that the advantage of outage-region Capacity becomes

less dominant as M grows. However, the outage-region capacity still shows better peak-power performance.

- Chapter 6 introduced FUSE and Fixed constraints. We proved that the penalty by each constraint converges to zero when compared to the optimal allocation of Theorem 4.4 in Proposition 6.2 and Proposition 6.4, respectively. We also showed that the penalties are small at a finite number of antennas by investigating asymptotic behaviors at high SNR in Corollary 6.2 and Proposition 6.5.

- Based on the theoretical results in Chapter 6, we proposed simple bit-allocation strategies: Binary-Search and Fixed allocation in Section 7.2. The proposed schemes are very low in complexity, but approaches the optimal performance of Full-Search allocation in Section 7.1. We extended both strategies to MIMO-OFDM combined with FUSE constraint in Chapter 6. Proposition 6.4 guarantees that a nonadaptive allocation based on Fixed allocation and FUSE constraint performs near optimally.

We also list relatively minor contributions:

- We discussed the diversity order of frequency-selective channels when CSI is unknown to the transmitter. Proposition 5.1 showed that the diversity order of a MIMO channel with memory L is bounded by $M_T M_R(L+1)$. Numerical results in Chapter 8 showed that the bound is actually achieved.

- We developed adaptive methods for updating the receive filter of eigenbeamforming in Section 3.3. Both blind and data-aided methods are discussed.

- We proved that eigenbeamforming with the short-term energy constraint in (51) achieves full diversity and spatial multiplexing orders in Proposition 5.2. We also showed the advantage of knowing CSI at the transmitter with the short-term constraint over CSI-ignorant transmitter.

9.2 Directions for Future Research

One of the underlying assumptions in this work is that the MIMO channel taps $\{\mathbf{G}_l\}$ are uncorrelated with each other. This assumption results in Lemma 2.1, namely memoryless MIMO channels $\{\mathbf{H}_n\}$ at each OFDM tone is statistically identical to each other, which is important in the analysis of the FUSE allocation in Chapter 6 and the proposed bit-allocation strategies in Chapter 7. In reality, this assumption might not always hold, as considered in [50]. The correlation between channel taps would result in a larger penalty than we expect for uncorrelated taps when we use the FUSE allocation. Moreover, it might be necessary to introduce a new strategy that considers the correlation. Also, we can analytically investigate the performance in the presence of the correlation, as in [50].

In the development of robust bit allocation in Section 7.3, we only consider the change in the line-of-sight (LOS) component. Another factor that affects the performance of the proposed bit allocation is the spatial correlation. When the spatial correlation becomes high, the deficiency in the channel matrix rank limits the number of available spatial channels. Thus, a change in spatial correlation leads to a severe performance degradation. As in the robust strategy for the LOS change, we expect that a few additions in the search set would be sufficient to cover the wide range of correlation changes. Also, a clever choice will compensate for the change in both LOS and spatial correlation.

We briefly mention the adaptive implementation of eigenbeamforming filter in Section 3.3. In practice, especially in a TDD system, the adaptive eigenbeamforming would be critical in a closed-loop system. Eventually, we develop more efficient adaptive algorithms and combine them with the proposed bit-allocation strategies and outer channel codes to obtain BER results. BER results will be valuable for practical implementation of closed-loop MIMO systems.

For some minor future work, we will do the following:

- We will prove or disprove that the uniform power profile minimizes the outage probability when L is given.

- We will analytically show that the CDF of arithmetic mean of eigenvalues in (146) converges to a step function as $L \to \infty$, as conjectured in Section 5.2.1.2.

- We will compare the peak-power performance between the zero-outage capacity C_0 and the outage-region capacity C_{OR}, which will show the advantage of using outage region in terms of peak-power probability.

REFERENCES

[1] Alamouti, S. M., "A simple transmitter diversity scheme for wireless communications," *IEEE J. Selected Areas Commun.*, vol. 16, pp. 1451–1458, October 1998.

[2] Alouni, M.-S. and Goldsmith, A. J., "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Vehic. Tech.*, vol. 48, pp. 1165–1181, July 1999.

[3] Asztély, D., "On antenna arrays in mobile communication systems: fast fading and GSM base station receiver algorithms," (Royal Institute of Technology, Stockholm, Sweden, IR-S3-SB-9611), 1996.

[4] Bai, Z. D., Silverstein, J. W., and Yin, Y. Q., "A note on the largest eigenvalue of a large dimensional sample covariance matrix," *J. Multivariate Anal.*, vol. 26, pp. 166–168, 1988.

[5] Bai, Z. D. and Yin, Y. Q., "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *Ann. Probab.*, vol. 21, pp. 1275–1294, 1993.

[6] Barry, J. R., Lee, E. A., and Messerschmitt, D. G., *Digital Communication*. Kluwer Academic Publishers, 3rd ed., 2003.

[7] Biglieri, E., Caire, G., and Taricco, G., "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1273–1289, May 2001.

[8] Biglieri, E., Proakis, J., and Shamai (Shitz), S., "Fading channels: information-theoretic and communication aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, October 1998.

[9] Bingham, J. A. C., "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Commun. Magazine*, vol. 28, April 1990.

[10] Bölcskei, H., Gesbert, D., and Paulraj, A. J., "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, February 2002.

[11] Brandenburg, L. H. and Wyner, A. D., "Capacity of the Gaussian channel with memory: the multivariate case," *Bell Syst. Tech. J.*, pp. 745–778, May/June 1974.

[12] Brunner, C., Utschick, W., and Nossek, J. A., "Exploiting the short-term and long-term channel properties in space and time: eigenbeamforming concepts for the BS in WCDMA," *European Transactions on Telecommunication*, vol. 12, no. 5, pp. 365–378, 2001.

[13] Caire, G., Taricco, G., and Biglieri, E., "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.

[14] Campello, J., "Practical bit-loading for DMT," in *Proc. IEEE ICC 1999*, (Vancourver, Canada), pp. 801–805, 1999.

[15] Cardoso, J. F. and Laheld, B. H., "Equivariant adaptive source separtion," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, December 1996.

[16] Causey, R. T., "Blind multiuser detection based on second-order statistics," *Ph.D. Dissertation, Georgia Institute of Technology, USA*, 1999.

[17] Chow, P. S., Cioffi, J. M., and Bingham, J. A. C., "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773–775, February/March/April 1995.

[18] Chuah, C., Tse, D., Kahn, J., and Valenzuela, R. A., "Capacity scaling of MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, March 2002.

[19] Cioffi, J. M., *Lecture Notes for Advanced Digital Communications*. Stanford University, 2002.

[20] Cover, T. M. and Thomas, J. A., *Elements of Inform. Theory*. John Wiley and Sons, 1991.

[21] Dahl, T., Christophersen, N., , and Gesbert, D., "Blind MIMO estimation based on the power method," *IEEE Trans. Signal Processing (to appear)*, 2004.

[22] Dahl, T., Silva, S., Christophersen, N., and Gesbert, D., "Blind MIMO communication based on subspace estimation," *IEEE Trans. Signal Processing (submitted)*, 2004.

[23] Edelman, A., "Eigenvalues and condition numbers of random matrices," *Ph.D. Dissertation, Massachusetts Institute of Technology, USA*, 1989.

[24] Foschini Jr., G. J. and Gans, M. J., "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, pp. 36–54, March 1998.

[25] Gesbert, D., Bölcskei, H., Gore, D. A., and Paulraj, A. J., "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Trans. Commun.*, vol. 50, pp. 1926–1934, December 2002.

[26] Giannakis, G., Hua, Y., Stoica, P., and Tong, L., *Signal Processing Advances in Communications*. Prentice Hall, New Jersey, 2001.

[27] Golden, G., Foschini, G., Venezuela, R., and Wolniansky, P., "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *IEEE Electronics Letters*, vol. 35, pp. 14–16, January 1999.

[28] Goldsmith, A. J. and Varaiya, P. P., "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, November 1997.

[29] Gradshteyn, I. S. and Ryzhik, I. M., *Table of Integrals, Series, and Products*. Academic Press, 6th ed., 2000.

[30] Grant, A., "Rayleigh fading multi-antenna channels," *EURASIP JASP*, vol. 2002, pp. 316–329, March 2002.

[31] Ha, J., Mody, A., Sung, J., Barry, J., McLaughin, S., and Stüber, G., "LDPC Coded OFDM with Alamouti/SVD Diversity Techniques," *Wireless Personal Communications*, vol. 23, pp. 183–194, 2002.

[32] Hanly, S. and Tse, D., "Multi-access fading channels—part II: delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2816–2831, November 1998.

[33] Horn, R. A. and Johnson, C. R., *Matrix Analysis*. Cambridge University Press, 1985.

[34] Jafar, S. A., Vishwanath, S., , and Goldsmith, A., "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Internation Conf. Commun. (ICC) 2001*, (St.Petersburg, Russia), pp. 2266–2270, 2001.

[35] Jonsson, D., "Some limit theorems for the eigenvalues of a sample covariance matrix," *J. Multivariate Anal.*, vol. 12, pp. 1–38, 1982.

[36] Karhunen, J., "Adaptive algorithms for estimaing eigenvectors of correlation type matrices," in *Proc. of IEEE Conference on Acoustics, Speech, and Signal Processing*, (San Diego, CA, USA), pp. 14.6.1–14.6.4, 1984.

[37] Knopp, R. and Caire, G., "Power control and beamforming for systems with multiple transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 638–648, October 2002.

[38] Kurpjuhn, T. P., Joham, M., Utschick, W., and Nossek, J. A., "Experimental studies about eigenbeamforming in standardization MIMO channels," in *Proc. IEEE Vehic. Tech. Conf. (VTC) 2002 Fall*, (Vancouver, Canada), pp. 185–189, 2002.

[39] Leon-Garcia, A., *Probability and Random Processes for Electrical Engineering*. Addison Wesley, 2nd ed., 1994.

[40] Mody, A. N. and Stüber, G. L., "Parameter estimation for OFDM with transmit receive diversity," in *Proc. of IEEE Veh. Tech. Conf. 2001*, (Rhodes, Greece), 2001.

[41] Mogensen, P., Frederiksen, F., Dam, H., Olesen, K., and Larsen, S., "A hardware testbed for evaluation of adaptive antennas in GSM/UMTS," in *Proc. of IEEE PIMRC 1996*, (Taipei, Taiwan), pp. 540–544, 1996.

[42] Müller, R. R., "A random matrix model of communication via antenna arrays," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2495–2506, September 2002.

[43] Naguib, A. F., "On the matched filter bound of transmit diversity techniques," in *Proc. IEEE Commun. Theory Workshop 2001*, (La Casa Del Zorro Desert Resort, CA), pp. 596–603, 2001.

[44] Oyman, Ö., Nabar, R. U., Bölcskei, H., and Paulraj, A. J., "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 2784–2795, November 2003.

[45] Palomar, D. P., Cioffi, J. M., , and Lagunas, M. A., "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Processing*, vol. 51, pp. 2381–2401, September 2003.

[46] Proakis, J. G., *Digital Communications*. McGraw-Hill, 3rd ed., 1995.

[47] Raleigh, G. G. and Cioffi, J. M., "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, March 1998.

[48] Reial, A. and Wilson, S. G., "Capacity-maximizing transmitter processing for fading matrix channel," in *Proc. Commun. Theory Mini Conf.*, pp. 6–10, 1999.

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

[50] Scaglione, A. and Salhotra, A., "The statistics of the MIMO frequency selective fading AWGN channel capacity," *submitted to IEEE Trans. Inform. Thoery*.

[51] Scaglione, A., Stoica, P., Barbarossa, S., Giannakis, G. B., , and Sampath, H., "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[52] Shin, H. and Lee, J. H., "Closed-form formulas for ergodic capacity of MIMO Rayleigh fading channels," in *Proc. of IEEE ICC 2003*, (Anchorage, Alaska, USA), pp. 2996–3000, 2003.

[53] Stüber, G. L., *Principles of Mobile Communication*. Kluwer Academic Publishers, 1996.

[54] Sung, J. H. and Barry, J. R., "Space-time processing with channel knowledge at the transmitter," in *Proceedings of EUROCON'01*, (Bratislave, Slovakia), pp. 26–29, 2001.

[55] Sung, J. H. and Barry, J. R., "Bit-Allocation Strategies for Closed-Loop MIMO OFDM," in *Proceedings of Vehic. Tech. Conf. VTC 2003*, (Orlando, FL, USA), 2003.

[56] Tarokh, V., Jafarkhani, H., and Calderbank, A. R., "Space-time block codes for wireless communications: performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–460, March 1999.

[57] Tarokh, V., Jafarkhani, H., and Calderbank, A. R., "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[58] Tarokh, V., Naguib, A., Seshadri, N., and Calderbank, A. R., "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol. 47, pp. 199–207, February 1999.

[59] Tarokh, V., Seshadri, N., and Calderbank, A. R., "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, March 1998.

[60] Telatar, I. E., "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–595, November/December 1999.

[61] Valenzuela, A. S. R., "A Statistical Model for Indoor Multipath Propagation," *IEEE JSAC*, vol. 5, pp. 128–137, February 1987.

[62] Varadarajan, B. and Barry, J. R., "The rate-diversity trade-off for linear space-time codes," in *Proc. of IEEE Vehicular Technology Conference 2002*, (Vancouver, Canada), 2002.

[63] Varadarajan, B. and Barry, J. R., "Optimization of full-rate full-diveristy linear space-time codes using the union bound," in *Proc. of 2003 Information Theory Workshop*, (Paris, France), pp. 210–213, 2003.

[64] Viterbi, A. J., "Very low rate convolution codes for maximum theoretical performance of spread-spectrum multiple-access channels," *IEEE JSAC*, vol. 8, pp. 641–649, May 1990.

[65] Wennström, M., "On MIMO systems and adaptive arrays for wireless communication: analysis and practical issues," *Ph.D. Dissertation, Uppsala University, Sweden*, 2002.

[66] Wennström, M., Helin, M., Rydberg, A., and Öberg, T., "On the optimality and performance of transmit and receive space diversity in MIMO channels," in *Proc. of IEE Technical Seminar on MIMO Communication Systems: From Concept to Implementation*, (London, U.K.), 2001.

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

VITA

Joon Hyun Sung was born in Seoul, Korea. He joined Seoul National University, Korea in 1992 and received his Bachelor of Science degree in Electrical Engineering in 1996. He received his Master of Science degree in Electrical Engineering from Seoul National University in 1998. He joined Georgia Institute of Technology and completed the requirements for the degree of Doctor of Philosophy in Electrical Engineering in July 2004. He has been a student member of IEEE sine 1999.