CONSTRAINED CAPACITY OF MIMO RAYLEIGH FADING CHANNELS

A Thesis

by

WENYAN HE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2011

Major Subject: Electrical Engineering

CONSTRAINED CAPACITY OF MIMO

RAYLEIGH FADING CHANNELS

A Thesis

by

WENYAN HE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Costas N. Georghiades
Committee Members,	Scott L. Miller
	Krzysztof A. Michalski
	Jeffrey D. Hart
Head of Department,	Costas N. Georghiades

May 2011

Major Subject: Electrical Engineering

ABSTRACT

Constrained Capacity of MIMO Rayleigh Fading Channels. (May 2011)

Wenyan He, B.Eng., Zhejiang University;

M.Eng., Zhejiang University

Chair of Advisory Committee: Dr. Costas N. Georghiades

In this thesis channel capacity of a special type of multiple-input multiple-output (MIMO) Rayleigh fading channels is studied, where the transmitters are subject to a finite phase-shift keying (PSK) input alphabet. The constraint on the input alphabet makes an analytical solution for the capacity beyond reach. However we are able to simplify the final expression, which requires a single expectation and thus can be evaluated easily through simulation. To facilitate simulations, analytical expressions are derived for the eigenvalues and eigenvectors of a covariance matrix involved in the simplified capacity expression. The simplified expression is used to provide some good approximations to the capacity at low signal-to-noise ratios (SNRs). Involved in derivation of the capacity is the capacity-achieving input distribution. It is proved that a uniform prior distribution is capacity achieving. We also show that it is the only capacity-achieving distribution for our channel model. On top of that we generalize the uniqueness case for an input distribution to a broader range of channels.

To my parents

ACKNOWLEDGMENTS

First and foremost I want to express my sincere gratitude toward my academic advisor, Dr. Georghiades. Over the years he not only brought me into the interesting world of wireless communications, but also exemplified how to be a man of honorable character, the latter of which to me weighs far more than the former.

I would like to thank my other committee members: Dr. Miller, Dr. Michalski and Dr. Hart. The many discussions I had with Dr. Miller will benefit me for the rest of my life. Also, I learned a great deal in the classes taught by Dr. Michalski and Dr. Hart.

I made friends with some students who joined Wireless Communications Group at Texas A&M University before and after me. Their names are too long to be listed here, but their friendship will forever be treasured.

TABLE OF CONTENTS

CHAPTER — Page

- I INTRODUCTION 1
  - A. Previous Work on MIMO Capacity 2
  - B. Proposed Research 4
  - C. Organization of Thesis 5
- II MIMO CAPACITY 7
  - A. Ergodic Capacity 8
  - B. Outage Capacity 9
- III PSK CONSTRAINED MIMO CAPACITY 13
  - A. PSK Constrained MIMO Capacity 13
    - 1. The Covariance Matrix 15
    - 2. Eigenvalues and Eigenvectors $\mathbf{\Xi}$ 16
    - 3. Generation of a Correlated Gaussian Sequence 24
  - B. Some Approximations 25
  - C. Optimal Input Distribution 27
  - D. Simulations 29
- IV UNIQUENESS OF CAPACITY-ACHIEVING INPUT DISTRIBUTION 33
  - A. Mathematical Formulation 35
  - B. MIMO Channel 36
  - C. ISI Channel 41
  - D. Discussion 42
- V CONCLUSION 45

REFERENCES 46

VITA 52

LIST OF FIGURES

FIGURE Page

1 Block diagram of a MIMO system. 1

2 PSK signal space diagram. A signal point in the space is described as $s_k = \exp(j2k\pi/Q)$, $0 \leq k \leq Q-1$. In this particular example $Q = 8$. 5

3 Comparison between Monte Carlo simulation and numerical integral: $M = 2$, $N = 1$, $Q = 2$. 29

4 MIMO capacities with different types of input: $M = 2$, $N = 2$. 30

5 Capacity for two systems: $M = 4$ and $M = 8$ while $Q = N = 2$. . . . 31

6 Comparison between precise Monte Carlo simulation and the three approximations: $M = 4$, $N = 2$, $Q = 2$. 31

7 Comparison between precise Monte Carlo simulation and the three approximations: $M = 4$, $N = 2$, $Q = 16$. 32

CHAPTER I

INTRODUCTION

As IEEE 802.11n standards become more and more readily available in various electronics products, multi-input multi-output (MIMO) has never been more accepted by the general public. With 3G networks gradually becoming the main stream and in some markets even being replaced by 4G networks, consumers' appetite for more bandwidth has never been bigger. Under this backdrop, any research work related to MIMO is expected to have a significant impact on the society in the immediate future. The importance of MIMO capacity is that it provides a theoretical limit capping network throughput for reliable transmission of information over MIMO channels. A simple MIMO system is illustrated in Fig. 1.

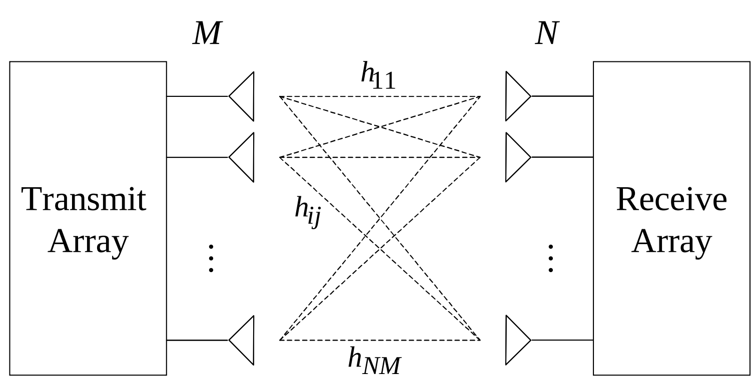

Fig. 1. Block diagram of a MIMO system.

This thesis follows the style of *IEEE Transactions on Information Theory.*

A. Previous Work on MIMO Capacity

After the pioneering work on MIMO systems was published by Telatar [1] and Foschini and Gans [2] in the late 1990's, MIMO capacity became a subject of significant research. Foschini [3] pointed out that the MIMO capacity could be substantially higher than that of a single-antenna system. He went further to establish the relationship between the M-dimensional architecture and M 1-dimensional architectures. The many currently available results on MIMO capacity are based on three exclusive assumptions: channel known at both the transmitter and receiver, channel known only at the receiver, and channel known at neither the transmitter nor the receiver. In the first category, the water-filling power allocation on the singular values of the channel matrix is shown to be optimal at the transmitter [1, 4]. In the second category, when the channel matrix entries are i.i.d., uniform power allocation is assumed at the transmitter since it has no knowledge of the channel state. Compared to the first two cases, the third, is mathematically more difficult to deal with and interest in it is relatively new. Marzetta and Hochwald [5] shed some light on this case by showing, for instance, that increasing the number of transmit antennas beyond the number of symbol periods in a coherence interval does not increase capacity. Zheng and Tse [6] also did some original work that falls into this category. They focused on the asymptotic capacity at high SNR and tried to give a geometric interpretation to the problem as sphere packing in the Grassmann manifold [7].

A brief overview of MIMO capacity can be found in [8]. Goldsmith *et al.* [9] gave a more detailed overview of recent results on single-user and multiuser MIMO capacity, in which multiple definitions of time-varying channel capacity are listed, e.g., outage capacity, ergodic capacity, and minimum rate capacity. Since a closed-form expression for the MIMO capacity is unavailable, the asymptotic behavior naturally becomes a primary topic of interest. When the number of transmit and receive antennas is large, the instantaneous MIMO capacity, as a random variable, can be well approximated by Gaussian distribution [10, 11]. Some nice analytical results are presented in [12] under the assumption that the channel is known at both transmitter and receiver. There are more asymptotic results available when the channel is assumed known at the receiver only. For example, Rapajic and Popescu derived a closed-form expression for the limiting capacity in [13], whose accuracy was corroborated in [11]. Sengupta and Mitra provided the limiting mean and variance of the conditional capacity given the channel gain matrix [14]. When the Rayleigh fading is correlated rather than independent, the corresponding results can be found in [12, 14], the former of which shows correlation reduces asymptotic capacity subject to uniform power allocation at the transmitter. Various approaches for calculating MIMO channel capacity have been used. Janaswamy obtained a series expression for the flat, uncorrelated Rayleigh fading case using Mellin transform [15]. Alfano *et al.* studied the capacity for a semi-correlated Rayleigh fading channel where correlation is present at only one side [16] and that for an uncorrelated Rician fading channel [17]. Chiani *et al.* applied Wishart matrix theory to derive a closed-form expression for the characteristic function of the instantaneous MIMO capacity [18], which in turn uniquely determines the distribution function of the capacity. However, one limitation in [18] is that it assumes correlation only at the end of the link with fewer antennas. A nice complement can be found in [19], where Smith *et al.* gave a closed-form expression for the characteristic function of a semi-correlated channel with correlation present at the end with more antennas. Kang *et al.* also calculated the moment generating function of MIMO capacity for different scenarios [20, 21].

The above being said, we need to point out that all aforementioned results studied a MIMO system with an unrestricted input, or equivalently, Gaussian input since it maximizes mutual information between MIMO channel input and output when there is no constellation constraint on the input alphabet. Actually, there has been little effort on input-constrained MIMO capacity [22]–[25], although a finite input alphabet apparently makes more practical sense when an actual MIMO system is to be established. Baccarelli [22] derived upper and lower bounds for the "symmetric capacity" employing two-dimensional data constellations. Hochwald and ten Brink [23] computed the MIMO mutual information of constrained constellations in their study of near-capacity performance of the LSD/APP detector/decoder. Lapidoth and Moser [24] took advantage of a dual expression for channel capacity to derive upper bounds on MIMO capacity. Finally, Müller [25] applied the replica method to deriving the MIMO channel capacity for a binary input alphabet.

B. Proposed Research

Noticing lack of results on MIMO capacity of constrained constellations, we decided to pursue this research direction. As will become obvious in later chapters, a MIMO system with a constrained input constellation usually becomes mathematically intractable , which only makes the capacity more elusive. We choose a relatively easy starting point, where the input signal is in a PSK constellation, whose diagram is presented in in Fig. 2. After the final expression for the capacity is obtained, which carries a double expectation, it is further simplified to a single expectation. Given this expectation, an analytical solution for the capacity is still beyond reach, which means Monte Carlo simulation is unavoidable. Yet we are able to find the analytical eigenvalues and eigenvectors for a covariance matrix included in the single-expectation capacity expression. This greatly boosts computational efficiency. After these results are ready, we also obtain some approximations to the capacity at low SNR, two of which are in closed form. The closed-form approximations can be used to find some insight into our MIMO system rather than solve a pure mathematical problem.

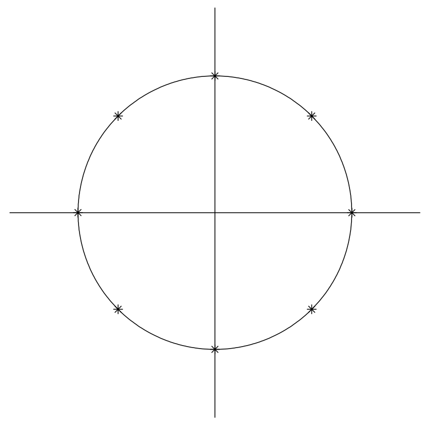

Fig. 2. PSK signal space diagram. A signal point in the space is described as $s_k = \exp(j2k\pi/Q)$, $0 \leq k \leq Q-1$. In this particular example $Q = 8$.

During the process of deriving the capacity, we have to establish that the uniform input distribution achieves capacity for our PSK signaling system. A by-product is, this capacity-achieving distribution is unique. Encouraged by this result on uniqueness, we want to put it in a somewhat unified framework. So we take a look at it for all MIMO channels with any discrete input alphabet, e.g., QAM. It will be shown that the uniqueness conclusion indeed works for any discrete input.

C. Organization of Thesis

The remaining part of this thesis is organized as follows. Chapter II summarizes available results on generic MIMO channels, such as the definition of outage capacity.

It eventually leads to PSK input constrained MIMO channel in Chapter III, where we not only study the capacity but also learn the capacity-achieving input distribution. Once the final simplified expression is obtained for the capacity, we try to find some approximations that work well within low SNR regime and that provide some insight. Even though uniqueness of capacity-achieving input distribution is employed in Chapter III, its proof is deferred to Chapter IV. In Chapter IV we try to prove uniqueness of capacity-achieving input distribution for a broad range of channels with a discrete input alphabet, of which PSK constellation is only a special example.

CHAPTER II

MIMO CAPACITY

MIMO systems have collected extensive attention in the new millennium due to its promise for a considerable increase in capacity, which is commonly considered to be a viable means for satisfying the ever-increasing demand for a higher data rate. When the concept of a multi-antenna system was originally introduced [1, 2, 3], no constraint was imposed on the input constellation. A Gaussian input was shown to achieve the capacity, or equivalently, to maximize the mutual information between input and output of a MIMO channel. However, it is impossible for a true Gaussian input to be realized in practice. The best one can do is approximate the Gaussian input with some sort of discrete input, which usually leads to a large and (maybe) irregular constellation. A more practically feasible way is to directly utilize a constrained input. Two popular options are phase-shift keying (PSK) and quadrature amplitude modulation (QAM).

The ground-breaking work by Telatar [1], Foschini and Gans [2] toward the end of last millennium not only provided some exciting results on capacity of multiple-antenna Gaussian channels but also stimulated a huge wave of enthusiasm toward various topics involving MIMO systems, including MIMO channel capacity, MIMO channel coding, space-time coding, etc. Even though the capacity topic has lost its original appeal to many researchers in the wireless communications area, there are still some interesting problems left to be solved.

This thesis focuses on one rather small topic in the MIMO capacity category and tries to dig deep into it. But before we make the jump, it is sensible to review others' accomplishments in the broader topic first.

Under most circumstances, a MIMO system can be very well characterized by the following simple model,

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{n} \tag{2.1}$$

where $\mathbf{H}$ is a complex Gaussian matrix, $\boldsymbol{n}$ represents complex Gaussian noise with a scaled identity matrix as its covariance matrix.

Among different notions of capacity [9], ergodic capacity and outage capacity are the two most often studied.

A. Ergodic Capacity

For ergodic capacity to be a legitimate characterization of a fading channel, the channel matrix $\mathbf{H}$ in (2.1) as a random process needs to be "ergodic" or changing fast enough. In other words, the fading has to be fast.

Early research on MIMO ergodic capacity assumed no constraint on input signals. Under this assumption or lack thereof, Telatar [1], Foschni and Gans [2] separately obtained a similar form of ergodic capacity for channel model (2.1),

$$C_{\mathrm{G}} = \mathrm{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{\gamma_s}{M}\mathbf{H}\mathbf{H}^{\dagger}\right)\right], \tag{2.2}$$

where the subscript G is used to emphasize the Gaussian characteristics of the capacity-achieving input signal.

More recently people started paying attention to capacity for a MIMO channel under discrete and finite input signaling, such as PSK [26]. In this case, the ergodic capacity becomes

$$\begin{aligned} C_{\text{erg}} &= M \log_2 Q - N \log_2 e \\ &\quad - \frac{1}{Q^M} \sum_{\boldsymbol{x}' \in \mathbb{X}} \mathrm{E}_{\mathbf{H}} \left[\mathrm{E}_{\hat{\boldsymbol{n}}} \left[\log_2 \sum_{\boldsymbol{x} \in \mathbb{X}} \exp \left\{ - \frac{\|\hat{\boldsymbol{n}} + \mathbf{H}(\boldsymbol{x}' - \boldsymbol{x})\|^2}{\sigma^2} \right\} \right] \right] \\ &= M \log_2 Q - N \log_2 e - \mathrm{E}_{\mathbf{H}} \left[\mathrm{E}_{\hat{\boldsymbol{n}}} \left[\log_2 \sum_{\boldsymbol{x} \in \mathbb{X}} \exp \left\{ - \frac{\|\hat{\boldsymbol{n}} + \mathbf{H}(\bar{\boldsymbol{x}} - \boldsymbol{x})\|^2}{\sigma^2} \right\} \right] \right] \end{aligned} \tag{2.3}$$

where M is the number of transmit antennas, N is the number of receive antennas, Q is the size of the PSK signal constellation, $\hat{\boldsymbol{n}} \sim \mathcal{N}_c(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, and

$$\bar{\boldsymbol{x}} = \left(\frac{1}{\sqrt{M}} \; \frac{1}{\sqrt{M}} \; \cdots \; \frac{1}{\sqrt{M}} \right)^{\mathrm{T}}.$$

Note that (2.3) is based on the conclusion that $p(\boldsymbol{x}) = 1/Q^M$.

B. Outage Capacity

The basic assumption for ergodic capacity is that the total transmission time is much longer than the coherence time of a fading channel. If this is not satisfied, as is the case in some real-time applications, e.g., speech transmission over wireless channels, the whole concept of ergodic capacity is no longer valid. In that case, we need to resort to a different definition of capacity, i.e., the outage capacity,

$$C_{\text{out}}(q) = \sup \left\{ R \geq 0 \colon \; \Pr\left[I(\boldsymbol{x}; \boldsymbol{y}) < R \right] \leq q \right\}, \tag{2.4}$$

where $q \in (0, 1)$ is the so-called outage probability.

Telatar also considered outage probability in [1],

$$P_{\text{out}}(R, P) = \inf_{\substack{\mathbf{Q} \geq 0 \\ \mathrm{tr}(\mathbf{Q}) \leq P}} \Pr\left[\log_2 \det(\mathbf{I}_N + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger) < R \right],$$

where P is the power constraint, R is the supportable rate, $\mathbf{Q} \geq 0$ means $\mathbf{Q}$ is a positive semi-definite matrix. This definition of outage probability is essentially the same (from a different angle though) as the definition of outage capacity in (2.4). Telatar conjectured that the optimal input covariance matrix $\mathbf{Q}$ is a diagonal matrix with the power equally shared among a subset of the transmit antennas [1]. Furthermore, the higher the rate, which inevitably leads to a higher outage probability, the fewer transmit antennas that should be put into service.

The motivation behind the pursuit of outage capacity is that the MIMO channel has zero Shannon capacity when the channel is non-ergodic, e.g., quasi-static or slow-fading. Due to the highly complicated probability density function of the instantaneous capacity, it was next to impossible to evaluate the cumulative distribution function of $I(\boldsymbol{x};\boldsymbol{y})$ as in (2.3), much less its pre-image R in (2.4). Therefore, let us look at the outage capacity with unconstrained input first, in which case the mutual information in (2.4) is

$$I(\boldsymbol{x};\boldsymbol{y}) = \log_2 \det\left(\mathbf{I}_N + \frac{\gamma_s}{M}\mathbf{H}\mathbf{H}^\dagger\right), \tag{2.5}$$

which is often called the instantaneous capacity. There was some nice effort at approximating the pdf of the above $I(\boldsymbol{x};\boldsymbol{y})$ by a Gaussian distribution [10, 11]. Although the approximation is only good under the assumption that the number of transmit and receive antennas is asymptotically large, their simulation results illustrate the surprising accuracy of the Gaussian approximation even for moderate-sized MIMO arrays. Basically, they showed that $I(\boldsymbol{x};\boldsymbol{y})$ in (2.5) converges in distribution to a Gaussian random variable, which enables one to attain an asymptotic formula for the capacity. Using singular value decomposition (SVD), Ge *et al.* [27] derived precise and good approximate statistical characteristics of $I(\boldsymbol{x};\boldsymbol{y})$ for different MIMO channels. Lately, some random matrix theories, more specifically, about Wishart matrix, have attracted attention and been applied to specification of outage capacity [18, 28, 29].

To compute outage capacity efficiently, Shi *et al.* [29] approximated the distribution of $I(\boldsymbol{x};\boldsymbol{y})$ by Gaussian and Gamma distributions and specified very simple and accurate formulas for outage probabilities.

Let

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^{\dagger}, & N \leq M \\ \mathbf{H}^{\dagger}\mathbf{H}, & N > M \end{cases} .$$

Then $\mathbf{W}$ is a Wishart matrix [30]. Let $N_{\min} = \min\{M, N\}$ and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_{N_{\min}}]^{\mathrm{T}}$ be the nonzero eigenvalues of the $N_{\min} \times N_{\min}$ matrix $\mathbf{W}$. Then the mutual information (2.5) can be rewritten as [18]

$$I = \sum_{i=1}^{N_{\min}} \log_2 \left(1 + \frac{\gamma_s}{M}\lambda_i\right) \tag{2.6}$$

where I is the short form of $I(\boldsymbol{x};\boldsymbol{y})$ (the same in the following). According to (2.6), once the joint distribution of $\boldsymbol{\lambda}$ is specified, the characteristic function of I, denoted by $\phi_I(z)$, is determined accordingly because

$$\phi_I(z) = \mathrm{E}\left[e^{j2\pi I z}\right] = \int_{0 \leq x_{N_{\min}} \leq \cdots \leq x_1} \cdots \int f_{\boldsymbol{\lambda}}(\boldsymbol{x}) \prod_{i=1}^{N_{\min}} \left(1 + \frac{\gamma_s}{M}\lambda_i\right)^{\frac{j2\pi z}{\ln 2}} \, d\boldsymbol{x}$$

where $\boldsymbol{x} = [x_1, \ldots, x_{N_{\min}}]^{\mathrm{T}}$ and $d\boldsymbol{x} = dx_1 \cdots dx_{N_{\min}}$. For the uncorrelated case, the joint pdf of ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{N_{\min}}$ of $\mathbf{W}$ is [18, 31]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{x}) = \frac{|\mathbf{V}(\boldsymbol{x})|^2}{\prod_{i=1}^{N_{\min}} [(N_{\max} - i)! \cdot (N_{\min} - i)!]} \prod_{i=1}^{N_{\min}} e^{-x_i} x_i^{N_{\max} - N_{\min}}$$

where $N_{\max} = \max\{M, N\}$, $\mathbf{V}(\boldsymbol{x})$ is the Vandermonde matrix with $x_1, \ldots, x_{N_{\min}}$ as its parameters. The final closed-form expressions for $\phi_I(z)$ can be found in [18] for both uncorrelated and correlated cases. Once the characteristic function is available, the pdf of the mutual information is simply

$$f_I(I) = \int_{-\infty}^{\infty} \phi_I(z) e^{-j2\pi I z} dz.$$

According to (2.4), the cdf of I is actually more helpful, which is easy to get from the preceding equation,

$$F_I(I) = \int_{-\infty}^{\infty} \phi_I(z) \left(\frac{1 - e^{-j2\pi I z}}{j2\pi z} \right) dz.$$

And the probability in (2.4) is nothing more than $F_I(R)$.

Up to this point, all accomplishments on outage capacity described are based on the assumption that the input is Gaussian. When the input is constrained to be, say, PSK, things will change. For example, (2.5) doesn't hold anymore. Outage capacity with a constrained input is still an open problem.

CHAPTER III

PSK CONSTRAINED MIMO CAPACITY

After the general discussion in Chapter II on capacity of a generic MIMO channel, we are now ready to focus our attention on the capacity of a MIMO fading channel under PSK signaling. We will also target the input distribution that achieves the capacity.

A. PSK Constrained MIMO Capacity

A Gaussian MIMO channel with Rayleigh fading and PSK input is described by

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{n} \tag{3.1}$$

where $\mathbf{H} \in \mathbb{C}^{N\times M}$, $\boldsymbol{x} \in \mathbb{X} \triangleq \Big\{(x_1\ x_2\ \cdots\ x_M)^{\mathrm{T}} \mid x_m = \exp\{j2\pi q/Q\}/\sqrt{M},\ 1 \le m \le M,\ 0 \le q \le Q-1\Big\}$ (superscript T denotes transpose), $\boldsymbol{y}, \boldsymbol{n} \in \mathbb{C}^{N\times 1}$, Q is the number of points in the PSK signal constellation, and M and N are the number of transmit and receive antennas, respectively. The definition of $\mathbb{X}$ essentially enforces the Q-ary PSK input constraint and also implies uniform power allocation at the transmitter. The values $\boldsymbol{x}$ can take guarantee that the average energy from each transmit antenna is $1/M$ [32]. Elements of $\mathbf{H}$ are i.i.d. complex Gaussian $\sim \mathcal{N}_c(0,1)$. Those of noise $\boldsymbol{n}$ are i.i.d. complex Gaussian $\sim \mathcal{N}_c(0,\sigma^2)$ where $\sigma^2 = 1/\gamma_s$ [32] and γ_s is the symbol signal-to-noise ratio E_s/N_0.

According to [1], under the assumption that the channel has no feedback, in which case the channel input $\boldsymbol{x}$ and the channel matrix $\mathbf{H}$ are independent, the mutual information between $\boldsymbol{x}$ and $\boldsymbol{y}$ with $\mathbf{H}$ known at the receiver is $I(\boldsymbol{x};(\boldsymbol{y},\mathbf{H})) =$

$I(\boldsymbol{x};\boldsymbol{y}\mid\mathbf{H})$, where

$$\begin{aligned} I(\boldsymbol{x};\boldsymbol{y}\mid\mathbf{H}) &= h(\boldsymbol{y}\mid\mathbf{H}) - h(\boldsymbol{y}\mid\boldsymbol{x},\mathbf{H}) = h(\boldsymbol{y}\mid\mathbf{H}) - h(\boldsymbol{n}) \\ &= -\int f(\mathbf{H})\,d\mathbf{H}\int f(\boldsymbol{y}\mid\mathbf{H})\log_2 f(\boldsymbol{y}\mid\mathbf{H})\,d\boldsymbol{y} - N\log_2(\pi e\sigma^2). \end{aligned} \tag{3.2}$$

The probability density function of the received vector $\boldsymbol{y}$ conditioned on $\mathbf{H}$ is

$$f(\boldsymbol{y}\mid\mathbf{H}) = \sum_{\boldsymbol{x}\in\mathbb{X}} p(\boldsymbol{x})f(\boldsymbol{y}\mid\boldsymbol{x},\mathbf{H}) = \sum_{\boldsymbol{x}\in\mathbb{X}} p(\boldsymbol{x})\left(\frac{1}{\pi\sigma^2}\right)^N \exp\left(-\frac{\|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|^2}{\sigma^2}\right), \tag{3.3}$$

where $\|\cdot\|$ denotes Euclidean norm. To maintain focus on PSK constrained MIMO capacity, we defer until Section C to show that the maximum of $I(\boldsymbol{x};(\boldsymbol{y},\mathbf{H}))$, which is by definition the channel capacity C, is achieved when the input is uniformly distributed, i.e., $p(\boldsymbol{x}) = 1/Q^M$ in (3.3). Hence it follows from (3.2) and (3.3) that the maximum achievable rate for the channel in (1) is

$$C = M\log_2 Q - N\log_2 e - \frac{1}{Q^M}\sum_{\boldsymbol{x}'\in\mathbb{X}} \mathrm{E}_{\mathbf{H}}\left\{\mathrm{E}_{\hat{\boldsymbol{n}}}\left[\log_2\sum_{\boldsymbol{x}\in\mathbb{X}}\exp\left(-\frac{\|\hat{\boldsymbol{n}}+\mathbf{H}(\boldsymbol{x}'-\boldsymbol{x})\|^2}{\sigma^2}\right)\right]\right\}, \tag{3.4}$$

where $\hat{\boldsymbol{n}} \sim \mathcal{N}_c(\mathbf{0},\sigma^2\mathbf{I})$. Furthermore, by noting that the distribution of $\mathbf{H}$ is rotationally invariant and that $\mathbf{U}\boldsymbol{x}$ is a point in the same PSK constellation as any PSK point $\boldsymbol{x}$ for $\mathbf{U} = \sqrt{M}\,\mathrm{diag}(\boldsymbol{x}') = \mathrm{diag}\left(e^{j2i'_1\pi/Q},\cdots,e^{j2i'_M\pi/Q}\right)$, (3.4) reduces to

$$C = M\log_2 Q - N\log_2 e - \mathrm{E}_{\mathbf{H}}\left\{\mathrm{E}_{\hat{\boldsymbol{n}}}\left[\log_2\sum_{\boldsymbol{x}\in\mathbb{X}}\exp\left(-\frac{\|\hat{\boldsymbol{n}}+\mathbf{H}(\bar{\boldsymbol{x}}-\boldsymbol{x})\|^2}{\sigma^2}\right)\right]\right\} \tag{3.5}$$

where $\bar{\boldsymbol{x}} = \left(\frac{1}{\sqrt{M}}\ \frac{1}{\sqrt{M}}\ \cdots\ \frac{1}{\sqrt{M}}\right)^{\mathrm{T}}$.

Numerical evaluation of (3.5) can be done using Monte Carlo averaging [33]. However, the two expectations in (3.5) make the evaluation a little inconvenient and give little insight into capacity computation. So next we would like to combine the two independent complex Gaussian random variables $\hat{\boldsymbol{n}}$ and $\mathbf{H}$ into one. The idea is to determine the covariance matrix for the concatenated random vector $\hat{\boldsymbol{n}}+\mathbf{H}(\bar{\boldsymbol{x}}-\boldsymbol{x})$ in (3.5) as $\boldsymbol{x}$ runs through all different values in $\mathbb{X}$. This concatenated vector is zero-mean, complex Gaussian. Thus, its covariance matrix uniquely determines its statistical characteristics. Fortunately, we can also analytically specify all nonzero eigenvalues and their associated eigenvectors for this covariance matrix rather than numerically, which greatly reduces complexity in generating an appropriately correlated Gaussian sequence. In addition, the availability of closed-form expressions for the eigenvalues and eigenvectors lays a foundation for some approximations presented in Section B.

1. The Covariance Matrix

For brevity, we introduce a new variable $\boldsymbol{z}_a \triangleq \bar{\boldsymbol{x}} - \boldsymbol{x}$, $1 \le a \le Q^M$. The subscript a varies with $\boldsymbol{x}$, i.e., as $\boldsymbol{x}$ traverses the Q^M different values in its domain, a changes from 1 through Q^M. The Q^M random vectors $\hat{\boldsymbol{n}}+\mathbf{H}\boldsymbol{z}_a$ ($1 \le a \le Q^M$) in the exponent in (3.5) can be concatenated to yield a $Q^M N \times 1$ complex random vector, denoted by $\tilde{\boldsymbol{q}}$, whose u^{th} (where $u = (a-1)N + b$, $1 \le a \le Q^M$, $1 \le b \le N$) component $\tilde{q}_u$ is a zero mean complex Gaussian scalar. The statistics of complex Gaussian vector $\tilde{\boldsymbol{q}}$ are completely specified by the covariance matrix of $\tilde{\boldsymbol{q}}$, denoted by $\widetilde{\boldsymbol{\Xi}}$. Since $\widetilde{\boldsymbol{\Xi}} = \mathrm{E}\left[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}^\dagger\right]$ (the symbol $\dagger$ denotes conjugate transpose), the $(u,v)^{\text{th}}$ element of $\widetilde{\boldsymbol{\Xi}}$ is

$$\tilde{\xi}_{uv} = \mathrm{E}\left[\tilde{q}_u \tilde{q}_v^*\right] = \left(\sigma^2 + \boldsymbol{z}_c^\dagger \boldsymbol{z}_a\right)\delta_{bd} \tag{3.6}$$

for $u = (a-1)N + b$, $v = (c-1)N + d$, $1 \le a, c \le Q^M$, $1 \le b, d \le N$, where $*$ denotes complex conjugate, and δ_{bd} is the Kronecker delta. Let $\xi_{ac} = \sigma^2 + \boldsymbol{z}_c^\dagger \boldsymbol{z}_a$. Equation (3.6) implies that $\widetilde{\boldsymbol{\Xi}}$ is a block matrix composed of $Q^M \times Q^M$ submatrices with the $(a,c)^{\text{th}}$ submatrix equal to an $N \times N$ identity matrix $\mathbf{I}_N$ multiplied by a scaling factor ξ_{ac}, where the pair (a,c) specifies the position of the corresponding submatrix in $\widetilde{\boldsymbol{\Xi}}$, i.e.,

$$\widetilde{\boldsymbol{\Xi}} = \begin{pmatrix} \xi_{1,1}\mathbf{I}_N & \xi_{1,2}\mathbf{I}_N & \cdots & \xi_{1,Q^M}\mathbf{I}_N \\ \xi_{2,1}\mathbf{I}_N & \xi_{2,2}\mathbf{I}_N & \cdots & \xi_{2,Q^M}\mathbf{I}_N \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{Q^M,1}\mathbf{I}_N & \xi_{Q^M,2}\mathbf{I}_N & \cdots & \xi_{Q^M,Q^M}\mathbf{I}_N \end{pmatrix} = \boldsymbol{\Xi} \otimes \mathbf{I}_N, \tag{3.7}$$

where $\boldsymbol{\Xi} \triangleq (\xi_{ac})_{Q^M \times Q^M}$ and the symbol $\otimes$ denotes Kronecker product. According to (3.7), $\widetilde{\boldsymbol{\Xi}}$ is a Hermitian matrix because $\boldsymbol{\Xi}$ is Hermitian.

Let $\boldsymbol{\Xi} = \boldsymbol{\Theta}\boldsymbol{\Lambda}\boldsymbol{\Theta}^\dagger$ be the eigen-decomposition of $\boldsymbol{\Xi}$. Then because of (3.7) the eigen-decomposition of $\widetilde{\boldsymbol{\Xi}}$, $\widetilde{\boldsymbol{\Xi}} = \widetilde{\boldsymbol{\Theta}}\widetilde{\boldsymbol{\Lambda}}\widetilde{\boldsymbol{\Theta}}^\dagger$, is readily obtained as $\widetilde{\boldsymbol{\Theta}} = \boldsymbol{\Theta} \otimes \mathbf{I}_N$, $\widetilde{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \otimes \mathbf{I}_N$ [34, Theorem 4.2.12, pp. 245]. This helps lower the complexity involved in the diagonalization procedure for $\widetilde{\boldsymbol{\Xi}}$ as the size of $\boldsymbol{\Xi}$ stays independent of N, the number of receive antennas.

2. Eigenvalues and Eigenvectors $\boldsymbol{\Xi}$

Given its size, $\boldsymbol{\Xi}$ has Q^M eigenvalues, counting multiplicity. Among them, $(M+1)$ are positive and the rest are zero. To substantiate this, let $\mathbf{Z}$ be the $M \times Q^M$ matrix whose a^{th} column is $\boldsymbol{z}_a$, $\hat{\boldsymbol{z}}_a = \left(\begin{smallmatrix} \boldsymbol{z}_a \\ \sigma \end{smallmatrix} \right)$ and $\widehat{\mathbf{Z}}$ be the matrix composed of columns $\{\hat{\boldsymbol{z}}_a\}_{a=1}^{Q^M}$. Then $\boldsymbol{\Xi} = \widehat{\mathbf{Z}}^\dagger \widehat{\mathbf{Z}}$. Since $\boldsymbol{z}_a = \bar{\boldsymbol{x}} - \boldsymbol{x}$, M columns of $\mathbf{Z}$ can be selected to form a scaled identity matrix $\left(1 - e^{j2\pi/Q}\right)/\sqrt{M}\mathbf{I}_M$, which implies $\text{rank}(\mathbf{Z}) = M$. Choose the corresponding M columns out of $\widehat{\mathbf{Z}}$ and also pick one particular column in $\widehat{\mathbf{Z}}$, $(0 \cdots 0 \ \sigma)^{\text{T}}$. These $(M+1)$ columns form a nonsingular matrix. Since this nonsingular matrix is part of $\widehat{\mathbf{Z}}$ and $\widehat{\mathbf{Z}}$ has a row number of $M+1$, $\text{rank}(\widehat{\mathbf{Z}}) = M+1$, which in turn leads to $\text{rank}(\boldsymbol{\Xi}) = M+1$ given the relationship between $\widehat{\mathbf{Z}}$ and $\boldsymbol{\Xi}$ [35]. Note that $\boldsymbol{\Xi}$ is positive semidefinite, which makes its eigenvalues nonnegative. In addition, the number of its positive eigenvalues equals its rank [35]. So $\boldsymbol{\Xi}$ has $(M+1)$ positive eigenvalues.

We now turn our attention to how to actually analytically determine the nonzero eigenvalues and their associated eigenvectors of $\mathbf{\Xi}$. Analytical evaluation will not only avoid numerical instabilities but will also make computation faster. So, our aim is to mathematically derive the matrices $\mathbf{\Theta}$ and $\mathbf{\Lambda}$. Note that we restrict our attention to the $(M+1)$ positive eigenvalues and their associated eigenvectors.

The ath column of $\mathbf{Z}$ is,

$$\begin{aligned}\boldsymbol{z}_a &= \bar{\boldsymbol{x}} - \boldsymbol{x}(a) \\ &= \frac{1}{\sqrt{M}} \begin{pmatrix} 1 - e^{j\frac{2i_{M-1}\pi}{Q}} \\ \vdots \\ 1 - e^{j\frac{2i_0\pi}{Q}} \end{pmatrix} = \frac{-2j}{\sqrt{M}} \begin{pmatrix} \sin\left(\frac{i_{M-1}\pi}{Q}\right) e^{j\frac{i_{M-1}\pi}{Q}} \\ \vdots \\ \sin\left(\frac{i_0\pi}{Q}\right) e^{j\frac{i_0\pi}{Q}} \end{pmatrix},\end{aligned}$$

where $a = 1 + i_{M-1} \cdots i_0(a-1, Q)$ and $i_{M-1} \cdots i_0(a-1,\, Q) = \sum_{k=0}^{M-1} i_k Q^k$ is the Q-ary representation of $a-1$. In the following, wherever no ambiguity arises, the two variables for a Q-ary number are omitted for brevity. As for the row structure of $\mathbf{Z}$, let its top and bottom rows be

$$\frac{-2j}{\sqrt{M}} \boldsymbol{z}^{\mathrm{r}}_{M-1} \quad \text{and} \quad \frac{-2j}{\sqrt{M}} \boldsymbol{z}^{\mathrm{r}}_0$$

respectively, where the superscript r stands for row. For an arbitrary row $\frac{-2j}{\sqrt{M}} \boldsymbol{z}^{\mathrm{r}}_k$ $(0 \le k < M)$,

$$\boldsymbol{z}^{\mathrm{r}}_k = \mathbf{1}^{\mathrm{T}}_{Q^{M-1-k}} \otimes \left\{ \left[0 \;\; \sin\left(\frac{\pi}{Q}\right) e^{j\frac{\pi}{Q}} \;\cdots\; \sin\left(\frac{(Q-1)\pi}{Q}\right) e^{j\frac{(Q-1)\pi}{Q}} \right] \otimes \mathbf{1}^{\mathrm{T}}_{Q^k} \right\}$$

where $\mathbf{1}$ denotes an all 1-entry column vector of proper size. Then

$$\mathbf{\Xi} = \widehat{\mathbf{Z}}^{\dagger} \widehat{\mathbf{Z}} = \mathbf{Z}^{\dagger} \mathbf{Z} + \sigma^2 \mathbf{1}^{\mathrm{T}}_{Q^M} \mathbf{1}_{Q^M} = \frac{4}{M} \left(\boldsymbol{g}^{\mathrm{r}}_{i_{M-1} \cdots i_0} \right)_{Q^M \times Q^M}$$

where

$$\boldsymbol{g}^{\mathrm{r}}_{i_{M-1} \cdots i_0} \triangleq \sum_{k=0}^{M-1} \sin\left(\frac{i_k \pi}{Q}\right) e^{-j\frac{i_k\pi}{Q}} \boldsymbol{z}^{\mathrm{r}}_k + \mathbf{\Sigma}^{\mathrm{r}}$$

is one out of Q^M rows, $i_{M-1}\cdots i_0$ is a Q-ary number used here as a subscript for the row vectors and $\mathbf{\Sigma}^{\mathrm{r}} = \frac{M}{4}\left(\sigma^2\ \cdots\ \sigma^2\right)$ is a $1 \times Q^M$ row vector.

To obtain the (nonzero) eigenvalues and the associated eigenvectors of $\mathbf{\Xi}$, we simply need to solve the vector equation $\lambda\boldsymbol{\phi} = \mathbf{\Xi}\boldsymbol{\phi}$ (where $\lambda \neq 0$ and $\boldsymbol{\phi} = (\phi_0\ \phi_1\ \cdots\ \phi_{Q^M-1})^{\mathrm{T}}$). Here $\boldsymbol{\phi}$ refers to a generic eigenvector of $\mathbf{\Xi}$. $\lambda\boldsymbol{\phi} = \mathbf{\Xi}\boldsymbol{\phi}$ actually consists of Q^M equations,

$$\lambda\phi_0 = \frac{4}{M}\boldsymbol{g}_0^{\mathrm{r}}\boldsymbol{\phi} = \frac{4}{M}\mathbf{\Sigma}^{\mathrm{r}}\boldsymbol{\phi}, \tag{3.8a}$$

$$\begin{aligned}
\lambda\phi_{p\cdot Q^k} &= \frac{4}{M}\boldsymbol{g}_{p\cdot Q^k}^{\mathrm{r}}\boldsymbol{\phi} = \frac{4}{M}\left[\sin\left(\frac{p\pi}{Q}\right)e^{-j\frac{p\pi}{Q}}\boldsymbol{z}_k^{\mathrm{r}} + \mathbf{\Sigma}^{\mathrm{r}}\right]\boldsymbol{\phi}\\
&= \frac{4}{M}\sin\left(\frac{p\pi}{Q}\right)e^{-j\frac{p\pi}{Q}}\boldsymbol{z}_k^{\mathrm{r}}\boldsymbol{\phi} + \lambda\phi_0\\
\Longrightarrow \sin\left(\frac{p\pi}{Q}\right)&e^{-j\frac{p\pi}{Q}}\boldsymbol{z}_k^{\mathrm{r}}\boldsymbol{\phi} = \frac{M}{4}\lambda(\phi_{p\cdot Q^k} - \phi_0),
\end{aligned} \tag{3.8b}$$

where $1 \le p \le Q-1$, $p \cdot Q^k$ in $\boldsymbol{g}_{p\cdot Q^k}^{\mathrm{r}}$ is a decimal instead of a Q-ary number, and finally,

$$\begin{aligned}
\lambda\phi_{i_{M-1}\cdots i_0} &= \frac{4}{M}\boldsymbol{g}_{i_{M-1}\cdots i_0}^{\mathrm{r}}\boldsymbol{\phi}\\
&= \frac{4}{M}\left[\sum_{k=0}^{M-1}\sin\left(\frac{i_k\pi}{Q}\right)e^{-j\frac{i_k\pi}{Q}}\boldsymbol{z}_k^{\mathrm{r}} + \mathbf{\Sigma}^{\mathrm{r}}\right]\boldsymbol{\phi}\\
&= \sum_{k=0}^{M-1}\lambda\left(\phi_{i_k\cdot Q^k} - \phi_0\right) + \lambda\phi_0\\
\Longrightarrow \phi_{i_{M-1}\cdots i_0} &= \sum_{k=0}^{M-1}(\phi_{i_k\cdot Q^k} - \phi_0) + \phi_0.
\end{aligned} \tag{3.8c}$$

From (3.8c) we know that each element of $\boldsymbol{\phi}$ can be expressed as a linear combination of ϕ_0 and $\phi_{p\cdot Q^k}$ ($1 \le p \le Q-1$, $0 \le k \le M-1$). So once we get ϕ_0 and $\phi_{p\cdot Q^k}$, the eigenvector $\boldsymbol{\phi}$ is determined. After some mathematical manipulation, (3.8a) becomes

$$\left[Q^{M-1}(M+Q-MQ)-\frac{\lambda}{\sigma^2}\right]\phi_0+Q^{M-1}\sum_{t=1}^{Q-1}\sum_{m=0}^{M-1}\phi_{t\cdot Q^m}=0. \tag{3.9}$$

On the other hand,

$$\begin{aligned}\boldsymbol{z}_k^{\mathrm{r}}\boldsymbol{\phi} &= \sum_{m=0}^{Q^M-1} z_{k,m}^{\mathrm{r}}\phi_m \\ &= \sum_{t=0}^{Q-1}\sin\left(\frac{t\pi}{Q}\right)e^{j\frac{t\pi}{Q}}\sum_{i_{M-1}\cdots i_0\backslash i_k=0\cdots 0,\ i_k=t}^{Q-1\cdots Q-1}\phi_{i_{M-1}\cdots i_0} \\ &= -\frac{j}{2}Q^{M-1}(M-1)(Q-1)\phi_0+Q^{M-1} \\ &\quad\cdot\sum_{t=1}^{Q-1}\left[\sin\left(\frac{t\pi}{Q}\right)e^{j\frac{t\pi}{Q}}\phi_{t\cdot Q^k}+\frac{j}{2}\sum_{\substack{m=0,\\ m\neq k}}^{M-1}\phi_{t\cdot Q^m}\right].\end{aligned} \tag{3.10}$$

Note that in the preceding equation the notation $i_{M-1}\cdots i_0\backslash i_k$ means $i_{M-1}\cdots i_{k+1}i_{k-1}\cdots i_0$ (i.e., the sequence excluding the kth term). It follows from (3.8b), (3.10), and (3.9) that

$$\begin{aligned}&\left\{\left[\frac{M}{4}\csc\left(\frac{p\pi}{Q}\right)e^{j\frac{p\pi}{Q}}+\frac{j}{2\sigma^2}\right]\lambda-\frac{j}{2}Q^{M-1}\right\}\phi_0-\frac{M}{4}\lambda\csc\left(\frac{p\pi}{Q}\right)e^{j\frac{p\pi}{Q}}\phi_{p\cdot Q^k} \\ &\quad -\frac{j}{2}Q^{M-1}\sum_{t=1}^{Q-1}e^{j\frac{2t\pi}{Q}}\phi_{t\cdot Q^k}=0,\end{aligned} \tag{3.11}$$

which represents $(Q-1)$ equations for a fixed k as $1\leq p\leq Q-1$. Subtracting the first from the remaining $(Q-2)$ equations, we have

$$\left[\csc\left(\frac{p\pi}{Q}\right)e^{j\frac{p\pi}{Q}}-\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}}\right]\phi_0+\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}}\phi_{Q^k}-\csc\left(\frac{p\pi}{Q}\right)e^{j\frac{p\pi}{Q}}\phi_{p\cdot Q^k}=0 \tag{3.12}$$

for $2\leq p\leq Q-1$. The first equation included in (3.11) corresponding to $p=1$ along with the $(Q-2)$ equations represented by (3.12) leads to

$$\left[\frac{M\lambda - Q^M}{4}\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} + j\frac{\lambda}{2\sigma^2}\right]\phi_0 + \frac{Q^M - M\lambda}{4}\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}}\phi_{Q^k} = 0. \quad (3.13)$$

Equations (3.9), (3.12) and (3.13) combined together give rise to the following coefficient matrix of size $[(Q-1)M+1] \times [(Q-1)M+1]$:

$$\mathbf{C} = \begin{pmatrix} Q^{M-1}(M+Q-MQ) - \frac{\lambda}{\sigma^2} & Q^{M-1}\mathbf{1}^{\mathrm{T}}_{(Q-1)M} \\ \mathbf{1}_M \otimes \mathbf{A} & \mathbf{I}_M \otimes \mathbf{B} \end{pmatrix} \quad (3.14)$$

for the unknown vector $\left(\phi_0\ \phi_{Q^0} \cdots \phi_{(Q-1)\cdot Q^0} \cdots \phi_{Q^{M-1}} \cdots \phi_{(Q-1)\cdot Q^{M-1}}\right)^{\mathrm{T}}$, where

$$\mathbf{A} = \begin{pmatrix} \frac{M\lambda - Q^M}{4}\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} + j\frac{\lambda}{2\sigma^2} \\ \csc\left(\frac{2\pi}{Q}\right)e^{j\frac{2\pi}{Q}} - \csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} \\ \vdots \\ \csc\left[\frac{(Q-1)\pi}{Q}\right]e^{j\frac{(Q-1)\pi}{Q}} - \csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} \end{pmatrix}$$

and

$$\mathbf{B} = \begin{pmatrix} \frac{Q^M - M\lambda}{4}\csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} & 0 & 0 & \cdots & 0 \\ \csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} & -\csc\left(\frac{2\pi}{Q}\right)e^{j\frac{2\pi}{Q}} & 0 & \cdots & 0 \\ \csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} & 0 & -\csc\left(\frac{3\pi}{Q}\right)e^{j\frac{3\pi}{Q}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \csc\left(\frac{\pi}{Q}\right)e^{j\frac{\pi}{Q}} & 0 & 0 & \cdots & -\csc\left[\frac{(Q-1)\pi}{Q}\right]e^{j\frac{(Q-1)\pi}{Q}} \end{pmatrix}.$$

Note that when $\lambda = Q^M/M$, the first row of $\mathbf{B}$ is all zero, which makes the $[(Q-1)k+2]$th row $(0 \le k \le M-1)$ of $\mathbf{C}$ equal to $\left(j\frac{Q^M}{2M\sigma^2}\ \mathbf{0}^{\mathrm{T}}_{(Q-1)M}\right)$. Hence $\det(\mathbf{C}) = 0$. Because of (3.8c) and the assumption that $\boldsymbol{\phi}$ is an eigenvector, ϕ_0, $\phi_{p\cdot Q^k}$ can not all be zero. Then a necessary and sufficient condition for a particular λ to be an eigenvalue of $\boldsymbol{\Xi}$ is $\det(\mathbf{C}) = 0$ as $\det(\lambda\mathbf{I} - \boldsymbol{\Xi})$ equals $\det(\mathbf{C})$ multiplied by a constant. Since $\lambda = Q^M/M$ makes $\det(\mathbf{C}) = 0$, it is a nonzero eigenvalue of $\boldsymbol{\Xi}$. Now let us assume $\lambda \neq Q^M/M$ so as to find other nonzero eigenvalues of $\boldsymbol{\Xi}$. Under this assumption, the matrix in (3.14) can be further simplified to

$$\mathbf{C}' = \begin{pmatrix} -M\lambda^2 + Q^M(M\sigma^2 + M + 1)\lambda - Q^{2M}\sigma^2 & \mathbf{0}^{\mathrm{T}}_{(Q-1)M} \\ \mathbf{1}_M \otimes \mathbf{A}' & -\mathbf{I}_{(Q-1)M} \end{pmatrix} \tag{3.15}$$

where

$$\mathbf{A}' = \begin{pmatrix} 1 + j\frac{2\lambda}{(M\lambda - Q^M)\sigma^2} \sin\left(\frac{\pi}{Q}\right) e^{-j\frac{\pi}{Q}} \\ \vdots \\ 1 + j\frac{2\lambda}{(M\lambda - Q^M)\sigma^2} \sin\left[\frac{(Q-1)\pi}{Q}\right] e^{-j\frac{(Q-1)\pi}{Q}} \end{pmatrix}.$$

Let determinant of $\mathbf{C}'$ be 0 and we obtain the following quadratic equation

$$M\lambda^2 - Q^M(M\sigma^2 + M + 1)\lambda + Q^{2M}\sigma^2 = 0, \tag{3.16}$$

whose two different roots are the simple (i.e., multiplicity 1) eigenvalues of $\mathbf{\Xi}$, denoted by

$$\lambda_M = \frac{Q^M}{2M}\left[(M\sigma^2 + M + 1) - \sqrt{(M\sigma^2 + M + 1)^2 - 4M\sigma^2}\right] \tag{3.17a}$$

and

$$\lambda_{M+1} = \frac{Q^M}{2M}\left[(M\sigma^2 + M + 1) + \sqrt{(M\sigma^2 + M + 1)^2 - 4M\sigma^2}\right] \tag{3.17b}$$

respectively. At the beginning of this subsection it was already established that $\mathbf{\Xi}$ has $M+1$ positive eigenvalues. Thus another nonzero eigenvalue of $\mathbf{\Xi}$, i.e., $\lambda = Q^M/M$ is of multiplicity-$(M-1)$. In other words, $\lambda_1 = \lambda_2 = \cdots = \lambda_{M-1} = Q^M/M$.

Now that we have all nonzero eigenvalues available, it is time to shift our attention to their associated eigenvectors. Those eigenvectors corresponding to the zero eigenvalue are of no use and hence are omitted.

When λ in (3.15) is either λ_M or λ_{M+1}, the elements of the associated eigenvector

$\boldsymbol{\phi}$ satisfy

$$\phi_{p\cdot Q^k} = \phi_0 \left[1 + j\frac{2\lambda}{(M\lambda - Q^M)\sigma^2} \sin\left(\frac{p\pi}{Q}\right) \exp\left(-j\frac{p\pi}{Q}\right)\right]. \tag{3.18}$$

Based on (3.8c) and (3.18), we know

$$\phi_{i_{M-1}\cdots i_0} = \phi_0 \left[1 + j\frac{2\lambda}{(M\lambda - Q^M)\sigma^2} \sum_{k=0}^{M-1} \sin\left(\frac{i_k\pi}{Q}\right) \exp\left(-j\frac{i_k\pi}{Q}\right)\right]. \tag{3.19}$$

To normalize $\boldsymbol{\phi}$, we simply make (after some manipulation)

$$\begin{aligned} 1 = \|\boldsymbol{\phi}\|^2 &= \sum_{i_{M-1}\cdots i_0 = 0\cdots 0}^{Q-1\cdots Q-1} |\phi_0|^2 \left|1 + j\frac{2\lambda}{(M\lambda - Q^M)\sigma^2} \sum_{k=0}^{M-1} \sin\left(\frac{i_k\pi}{Q}\right) \exp\left(-j\frac{i_k\pi}{Q}\right)\right|^2 \\ &= |\phi_0|^2 Q^M \left[1 + \frac{2M\lambda}{(M\lambda - Q^M)\sigma^2} + \frac{M(M+1)\lambda^2}{(M\lambda - Q^M)2\sigma^4}\right], \end{aligned}$$

which yields (choosing the positive real number)

$$\phi_0 = Q^{-\frac{M}{2}} \left[1 + \frac{2M\lambda}{(M\lambda - Q^M)\,\sigma^2} + \frac{M(M+1)\lambda^2}{(M\lambda - Q^M)^2\,\sigma^4}\right]^{-\frac{1}{2}}. \tag{3.20}$$

When $\lambda = Q^M/M$, the second row of $\mathbf{C}$ in (3.14) dictates

$$\phi_0 = 0. \tag{3.21}$$

Then the first row implies

$$\sum_{p=1}^{Q-1} \sum_{k=0}^{M-1} \phi_{p\cdot Q^k} = 0. \tag{3.22}$$

Finally, $\mathbf{B}$ yields

$$\phi_{p\cdot Q^k} = \sin\left(\frac{p\pi}{Q}\right) \csc\left(\frac{\pi}{Q}\right) e^{j\frac{(1-p)\pi}{Q}} \phi_{Q^k}. \tag{3.23}$$

Combining (3.22) and (3.23), it is obvious that

$$\sum_{k=0}^{M-1} \phi_{Q^k} = 0. \tag{3.24}$$

Because $\phi_0 = 0$, (3.8c) reduces to

$$\boldsymbol{\phi} = \frac{\sqrt{M}}{2j} \csc\left(\frac{\pi}{Q}\right) e^{j\frac{\pi}{Q}} \mathbf{Z}^{\dagger} \left[\phi_{Q^0} \ \phi_{Q^1} \ \cdots \ \phi_{Q^{M-1}}\right]^{\mathrm{T}}. \tag{3.25}$$

Next our task is to find $(M-1)$ orthonormal eigenvectors associated with $\lambda = Q^M/M$ whose entries satisfy (3.21), (3.24) and (3.25) simultaneously. Fix ϕ_{Q^0} in (3.24) to be $-\sin\left(\frac{\pi}{Q}\right) e^{-j\frac{\pi}{Q}}$ and let $\phi_{Q^1}, \cdots, \phi_{Q^{M-1}}$ be $\sin\left(\frac{\pi}{Q}\right) e^{-j\frac{\pi}{Q}}$ in turn while others are 0. For example,

$$\phi_{Q^0} = -\sin\left(\frac{\pi}{Q}\right) \exp\left(-j\frac{\pi}{Q}\right),$$
$$\phi_{Q^1} = \sin\left(\frac{\pi}{Q}\right) \exp\left(-j\frac{\pi}{Q}\right),$$

and $\phi_{Q^2} = \cdots = \phi_{Q^{M-1}} = 0$. A total of $(M-1)$ column vectors are constructed this way. Corresponding to these $(M-1)$ vectors, $\boldsymbol{\phi} = \left(\boldsymbol{z}_k^{\mathrm{r}} - \boldsymbol{z}_{M-1}^{\mathrm{r}}\right)^{\dagger}$ as $0 \le k \le M-2$. Since $\boldsymbol{z}_{M-1}^{\mathrm{r}}, \boldsymbol{z}_{M-2}^{\mathrm{r}}, \cdots, \boldsymbol{z}_0^{\mathrm{r}}$ are M linearly independent row vectors (recall that $\mathrm{rank}(\mathbf{Z}) = M$), $\left(\boldsymbol{z}_{M-2}^{\mathrm{r}} - \boldsymbol{z}_{M-1}^{\mathrm{r}}\right)^{\dagger}, \cdots, \left(\boldsymbol{z}_0^{\mathrm{r}} - \boldsymbol{z}_{M-1}^{\mathrm{r}}\right)^{\dagger}$ are linearly independent as well. Following standard Gram-Schmidt orthogonalization procedures, we obtain the $(M-1)$ orthonormal vectors $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{M-1}$, which can be described as follows.

Let $\boldsymbol{\zeta}_p$ $(1 \le p \le M-1)$ be a $Q^p \times 1$ vector whose $\left(\sum_{k=0}^{p-1} i_k Q^k\right)$th entry is

$$\frac{1}{p} \sum_{k=0}^{p-1} \sin\left(\frac{i_k \pi}{Q}\right) \exp\left[j\frac{(1-i_k)\pi}{Q}\right]$$

and $\boldsymbol{\eta}$ be a $Q-1 \times 1$ vector whose qth $(0 \le q \le Q)$ entry is

$$\eta_q = \sin\left(\frac{q\pi}{Q}\right) \exp\left[j\frac{(1-q)\pi}{Q}\right].$$

Then

$$\boldsymbol{\theta}_p = 2Q^{-M/2} \sqrt{\frac{p}{p+1}} \left(\mathbf{1}_{Q^p} \otimes \boldsymbol{\eta} - \boldsymbol{\zeta}_p \otimes \mathbf{1}_Q\right) \otimes \mathbf{1}_{Q^{M-p-1}}.$$

3. Generation of a Correlated Gaussian Sequence

To efficiently evaluate the capacity in (3.5), a zero-mean complex Gaussian sequence with covariance matrix $\widetilde{\boldsymbol{\Xi}} = \boldsymbol{\Xi} \otimes \mathbf{I}_N$ should be generated. Let $\boldsymbol{\Lambda}$ (the diagonal matrix obtained in eigen-decomposition of $\boldsymbol{\Xi}$) be arranged such that its nonzero eigenvalues are on the lower right corner. Then $\mathbf{R} = \boldsymbol{\Theta}\boldsymbol{\Lambda}^{1/2}$ ($\boldsymbol{\Lambda}^{1/2}$ is the component-wise square root of $\boldsymbol{\Lambda}$) is of the form $\mathbf{R} = [\mathbf{0}\ \mathbf{D}]$ where $\mathbf{0}$ is a zero matrix of proper size and

$$\mathbf{D} = \left[\sqrt{\lambda_1}\boldsymbol{\theta}_1 \cdots \sqrt{\lambda_{M+1}}\boldsymbol{\theta}_{M+1}\right],$$

λ_k $(1 \leq k \leq M+1)$ is the kth nonzero eigenvalue of $\boldsymbol{\Xi}$ and $\boldsymbol{\theta}_k$ is the eigenvector associated with λ_k. Recall now that a vector of Gaussian random variables with covariance matrix $\widetilde{\boldsymbol{\Xi}} = \widetilde{\mathbf{R}}\widetilde{\mathbf{R}}^\dagger$, where $\widetilde{\mathbf{R}} = \mathbf{R} \otimes \mathbf{I}_N$, can be generated as $\boldsymbol{v} = \widetilde{\mathbf{R}}\tilde{\boldsymbol{s}}$ where $\tilde{\boldsymbol{s}}$ is a vector of i.i.d. $\mathcal{N}_c(0,1)$ entries. It is clear that

$$\boldsymbol{v}_a = \sum_{m=1}^{M+1} [\mathbf{D}]_{am}\boldsymbol{s}_m, \qquad 1 \leq a \leq Q^M$$

where $\boldsymbol{v}_a$ is the ath length-N subvector in $\boldsymbol{v}$ and $\boldsymbol{s}_m$ is the $(Q^M - M - 1 + m)$th length-N subvector in $\tilde{\boldsymbol{s}}$. Then $r_a = ||\boldsymbol{v}_a||^2$ is the numerator in the argument of the exponents to be summed in (3.5). Alternatively, one can generate an $(M+1) \times N$ matrix $\mathbf{S}$ of i.i.d. $\mathcal{N}_c(0,1)$ entries and evaluate $\mathbf{Q} = \mathbf{DS}$. Clearly the rows of $\mathbf{Q}$ are equivalent to the vectors $\boldsymbol{v}_a$. Hence r_a equals the squared norm of the ath row of $\mathbf{Q}$. With the help of r_a, (3.5) can be rewritten as

$$C = M\log_2 Q - N\log_2 e - \mathrm{E}\left[\log_2 \sum_{a=1}^{Q^M} \exp\left(-\frac{r_a}{\sigma^2}\right)\right], \tag{3.26}$$

in which the expectation is over r_a.

B. Some Approximations

With $\mathbf{D}$ explicitly available now based on $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{M+1}$ and $\lambda_1, \cdots, \lambda_{M+1}$, i.e.,

$$\mathbf{D} = \left[\sqrt{\lambda_1}\boldsymbol{\theta}_1 \; \sqrt{\lambda_2}\boldsymbol{\theta}_2 \; \cdots \; \sqrt{\lambda_{M+1}}\boldsymbol{\theta}_{M+1}\right],$$

we can attempt to make some approximations regarding the input-constrained capacity given in (3.26).

When γ_s is small, or $\sigma^2 = 1/\gamma_s$ is large, λ_{M+1} is much greater than λ_1 through λ_M while the elements of $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{M+1}$ have about the same order of magnitude. Then, the contribution to r_a in (3.26) from all terms other than the one associated with $\sqrt{\lambda_{M+1}}\theta_{M+1,\,a-1}$ is negligible. This means, we only need to generate a $1 \times N$ i.i.d. Gaussian sequence $\boldsymbol{s} = [s_1 \cdots s_N]$ and $r_a \approx \lambda_{M+1}|\theta_{M+1,\,a-1}|^2 \sum_{b=1}^{N} |s_b|^2$, which is what we call Approximation 1.

Equation (3.19) applies to the eigenvectors associated with either λ_M or λ_{M+1}. For the eigenvector $\boldsymbol{\theta_{M+1}}$ associated with λ_{M+1}, when σ^2 is large, its entries

$$|\theta_{M+1,\,a-1}|^2 \approxeq \theta_{M+1,\,0}^2 \left[1 + \frac{4\lambda_{M+1}}{(M\lambda_{M+1} - Q^M)\,\sigma^2} \sum_{k=0}^{M-1} \sin^2\left(\frac{i_k \pi}{Q}\right)\right]$$

for $1 \leq a \leq Q^M$, where $\theta_{M+1,0}$ is given in (3.20) with λ replaced by λ_{M+1}. Then the key part in (3.26)

$$\begin{aligned}\sum_{a=1}^{Q^M} e^{-r_a/\sigma^2} &\approx \sum_{a=1}^{Q^M} \exp\left\{-\frac{\lambda_{M+1}\theta_{M+1,0}^2}{\sigma^2} \sum_{b=1}^{N} |s_b|^2 \right.\\ &\qquad \left. \cdot \left[1 + \frac{4\lambda_{M+1}}{(M\lambda_{M+1} - Q^M)\,\sigma^2} \sum_{k=1}^{M-1} \sin^2\left(\frac{i_k \pi}{Q}\right)\right]\right\} \\ &= \exp\left(-\frac{\lambda_{M+1}\theta_{M+1,0}^2}{\sigma^2} \sum_{b=1}^{N} |s_b|^2\right) \left\{\sum_{p=0}^{Q-1} \exp\left[-\frac{4\lambda_{M+1}^2\theta_{M+1,0}^2 \sum_{b=1}^{N} |s_b|^2}{(M\lambda_{M+1} - Q^M)\,\sigma^4} \sin^2\left(\frac{p\pi}{Q}\right)\right]\right\}^M,\end{aligned}$$

which leads to

$$\mathrm{E}\left[\log_2 \sum_{a=1}^{Q^M} \exp\left(-\frac{r_a}{\sigma^2}\right)\right] \approx -\frac{\lambda_{M+1}\theta_{M+1,0}^2 \log_2 e}{\sigma^2} \mathrm{E}\left[\sum_{b=1}^{N} |s_b|^2\right]$$

$$+ M\mathrm{E}\left\{\log_2 \sum_{p=0}^{Q-1} \exp\left[-\frac{4\lambda_{M+1}^2\theta_{M+1,0}^2 \sum_{b=1}^{N} |s_b|^2}{\left(M\lambda_{M+1} - Q^M\right)\sigma^4} \sin^2\left(\frac{p\pi}{Q}\right)\right]\right\} \tag{3.27}$$

$$\approx -\frac{\lambda_{M+1}\theta_{M+1,0}^2 \log_2 e}{\sigma^2} N + M\left\{\log_2 Q - \frac{2\lambda_{M+1}^2\theta_{M+1,0}^2 \log_2 e}{\left(M\lambda_{M+1} - Q^M\right)\sigma^4} \mathrm{E}\left[\sum_{b=1}^{N} |s_b|^2\right]\right\} \tag{3.28}$$

$$= M \log_2 Q - \left[\frac{\lambda_{M+1}\theta_{M+1,0}^2}{\sigma^2} + \frac{2M\lambda_{M+1}^2\theta_{M+1,0}^2}{\left(M\lambda_{M+1} - Q^M\right)\sigma^4}\right] N \log_2 e. \tag{3.29}$$

Note that from (3.27) to (3.28), an approximation based on the Taylor series expansion $\log_2\left\{\sum_{p=0}^{Q-1} \exp\left[x \sin^2\left(\frac{p\pi}{Q}\right)\right]\right\} \approx \log_2 Q + \frac{1}{2} x \log_2 e$ (for small x) is applied. With (3.29), (3.26) reduces to

$$C \approx \left(\lambda_{M+1}\theta_{M+1,0}^2\gamma_s + \frac{2M\lambda_{M+1}^2\theta_{M+1,0}^2\gamma_s^2}{M\lambda_{M+1} - Q^M} - 1\right) N \log_2 e, \tag{3.30}$$

which is our Approximation 2 and can be shown to be independent of Q. It is easily proved with the help of a little calculus that

$$\lim_{\gamma_s \to 0} \frac{1}{\gamma_s}\left(\lambda_{M+1}\theta_{M+1,0}^2\gamma_s + \frac{2M\lambda_{M+1}^2\theta_{M+1,0}^2\gamma_s^2}{M\lambda_{M+1} - Q^M} - 1\right) = 1.$$

This leads (3.30) to Approximation 3,

$$C \approx \gamma_s N \log_2 e. \tag{3.31}$$

The significance of (3.31) is that it indicates when SNR is very low, the PSK-input MIMO capacity is independent of both the number of transmit antennas M and the size of the signal constellation Q. Actually (3.31) is a special case of the results presented by Verdú [36, eqs. (16), (53), and (56)]. It also is in agreement with conclusions in [37], where Oyman *et al.* pointed out any extra antenna should be placed at the receiver end to increase capacity in the low-SNR regime. The difference is [37] assumed Gaussian input, while here we have shown a similar conclusion holds for PSK constrained input.

Suppose now that there is no constellation constraint on the input alphabet. Then the well-known formula for MIMO capacity due to Gaussian input is [1]

$$\begin{aligned} C &= \mathrm{E}\left[\log_2 \det\left(\mathbf{I}_N + \frac{\gamma_s}{M}\mathbf{H}\mathbf{H}^\dagger\right)\right] \\ &= \mathrm{E}\left[\log_2 \prod_{k=1}^{N}(1+\mu_k\gamma_s/M)\right] \approx \frac{\gamma_s}{M}\log_2 e \sum_{k=1}^{N}\mathrm{E}\left[\mu_k\right] \\ &= \frac{\gamma_s}{M}\log_2 e \mathrm{E}\left[\mathrm{tr}\left(\mathbf{H}\mathbf{H}^\dagger\right)\right] = \gamma_s N \log_2 e \end{aligned} \tag{3.32}$$

where μ_k $(1 \leq k \leq N)$ is an eigenvalue of $\mathbf{H}\mathbf{H}^\dagger$. The approximation holds when γ_s is small. Note that the final equation of (3.32) is identical to (3.31), which indicates PSK input-constrained MIMO capacity is approximately equal to the unconstrained capacity at low SNR.

C. Optimal Input Distribution

Earlier in Section A we omitted an indispensable part in deriving the capacity to focus on PSK constrained MIMO capacity itself. Since the final results were based on the assumption that the channel capacity is achieved when the input is uniformly distributed, it is necessary to prove the assumption. We achieve this by showing the uniform input distribution maximizes the mutual information $I(\boldsymbol{x};\boldsymbol{y} \mid \mathbf{H})$ in (3.2).

First $I(\boldsymbol{x};\boldsymbol{y} \mid \mathbf{H})$ is rewritten below following (3.2) and (3.3),

$$I(\boldsymbol{x};\boldsymbol{y} \mid \mathbf{H}) = -N\log_2 e - \beta \mathrm{E}_{\mathbf{H}}\left[\int \sum_{\boldsymbol{x}\in\mathbb{X}} p(\boldsymbol{x}) e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\boldsymbol{x}\|^2} \ln\left(\sum_{\boldsymbol{x}\in\mathbb{X}} p(\boldsymbol{x}) e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\boldsymbol{x}\|^2}\right) d\boldsymbol{z}\right], \tag{3.33}$$

where $\alpha = 1/\sigma$ and β is a positive constant.

Let $p_1 \geq \ldots \geq p_{Q^M}$, where $p_k = p(\boldsymbol{x}_k)$, be a non-uniform input distribution. Obviously $p_1 > p_{Q^M}$. Due to the circular symmetry in the PSK constellation, there is a certain diagonal and unitary matrix $\mathbf{U}$ such that $\boldsymbol{x}_{Q^M} = \mathbf{U}\boldsymbol{x}_1$. Construct the following mapping: $\tilde{\boldsymbol{x}}_k = \mathbf{U}\boldsymbol{x}_k$, $1 \leq k \leq Q^M$. The mapping is clearly invertible. Given the special properties of $\mathbf{U}$, $(\tilde{\boldsymbol{x}}_1 \ldots \tilde{\boldsymbol{x}}_{Q^M})$ is simply a permutation of $(\boldsymbol{x}_1 \ldots \boldsymbol{x}_{Q^M})$. Besides, $\tilde{p}_k = p(\tilde{\boldsymbol{x}}_k) = p(\boldsymbol{x}_k) = p_k$. Recall that $\boldsymbol{x}_{Q^M} = \mathbf{U}\boldsymbol{x}_1 = \tilde{\boldsymbol{x}}_1$. So in this new input distribution $\tilde{p}_k$, $\boldsymbol{x}_{Q^M}$ has the highest probability (which is greater than or equal to the probability of $\boldsymbol{x}_1$) while $p(\boldsymbol{x}_1) > p(\boldsymbol{x}_{Q^M})$ in the original input distribution p_k. Then these two input distributions are different. Now for input distribution $\tilde{p}_k$, the mutual information based on (3.33) is

$$
\begin{aligned}
I(\boldsymbol{x}; (\boldsymbol{y}, \mathbf{H})) = -\beta \int \mathrm{E}_{\mathbf{H}} \left[\sum_{k=1}^{Q^M} \tilde{p}_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\tilde{\boldsymbol{x}}_k\|^2} \ln \left(\sum_{k=1}^{Q^M} \tilde{p}_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\tilde{\boldsymbol{x}}_k\|^2} \right) \right] d\boldsymbol{z} \\
- N \log_2 e \\
= -N \log_2 e - \beta \int \mathrm{E}_{\mathbf{H}} \left[\sum_{k=1}^{Q^M} p_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\mathbf{U}\boldsymbol{x}_k\|^2} \ln \left(\sum_{k=1}^{Q^M} p_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\mathbf{U}\boldsymbol{x}_k\|^2} \right) \right] d\boldsymbol{z} \quad (3.34) \\
= -N \log_2 e - \beta \int \mathrm{E}_{\mathbf{H}} \left[\sum_{k=1}^{Q^M} p_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\boldsymbol{x}_k\|^2} \ln \left(\sum_{k=1}^{Q^M} p_k e^{-\|\boldsymbol{z}-\alpha\mathbf{H}\boldsymbol{x}_k\|^2} \right) \right] d\boldsymbol{z} \quad (3.35)
\end{aligned}
$$

as $I(\boldsymbol{x}; (\boldsymbol{y}, \mathbf{H})) = I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H})$. From (3.34) to (3.35), the rotational invariance of $\mathbf{H}$ and the fact that $\mathbf{U}$ is unitary are employed. Note that (3.35) is actually the mutual information corresponding to the original input distribution p_k. This means an arbitrary non-uniform input distribution always leads to a different input distribution that achieves the same mutual information.

In Chapter IV we will establish that mutual information is a *strictly* concave function of prior distribution, which implies $I(\boldsymbol{x}; (\boldsymbol{y}, \mathbf{H}))$ has a unique maximizer that achieves the capacity. So we have in effect shown that any non-uniform input distribution fails to achieve capacity because it violates the uniqueness. Hence the capacity-achieving input distribution has to be the uniform distribution.

D. Simulations

We present simulation results to verify our conclusions drawn earlier.

To substantiate (3.26), we compare the result based on the integral form of (3.5) to that of Monte Carlo simulations based on (3.26). Because of computational limitations in computing (3.5), we only look at a binary-input (i.e., $Q = 2$) MIMO system with $M = 2$, $N = 1$. Fig. 3 illustrates such a comparison. Fig. 4 includes

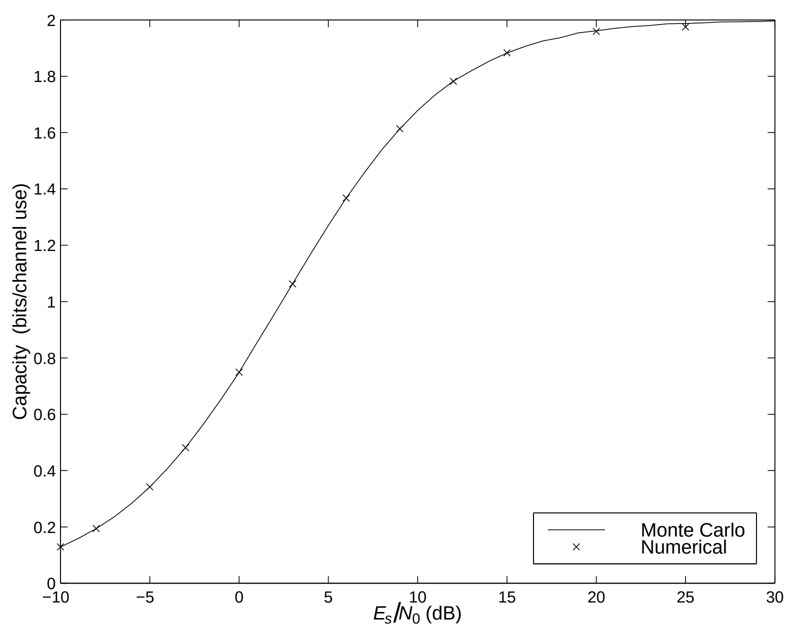

Fig. 3. Comparison between Monte Carlo simulation and numerical integral: $M = 2$, $N = 1$, $Q = 2$.

MIMO capacity subject to different types of input: Gaussian, BPSK, 4PSK, 8PSK, and 16PSK. The part at low SNR verifies the identicalness between (3.31) and (3.32).

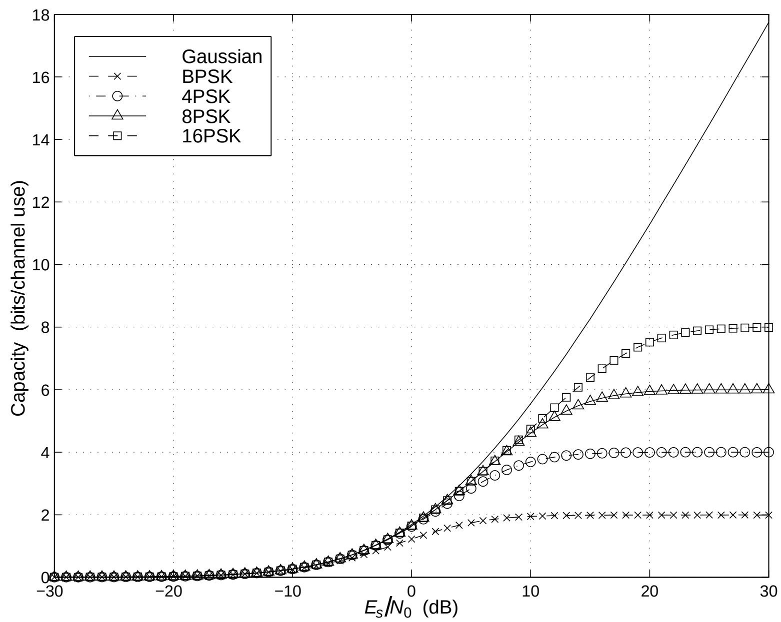

Fig. 4. MIMO capacities with different types of input: $M = 2$, $N = 2$.

Fig. 5 gives the capacity for the two systems with $M = 4$ and $M = 8$, respectively, while $Q = N = 2$ in both. It is clear that in low SNR regions, the MIMO capacity is almost independent of M. Figures 6 and 7 substantiate our Approximations 1, 2, and 3.

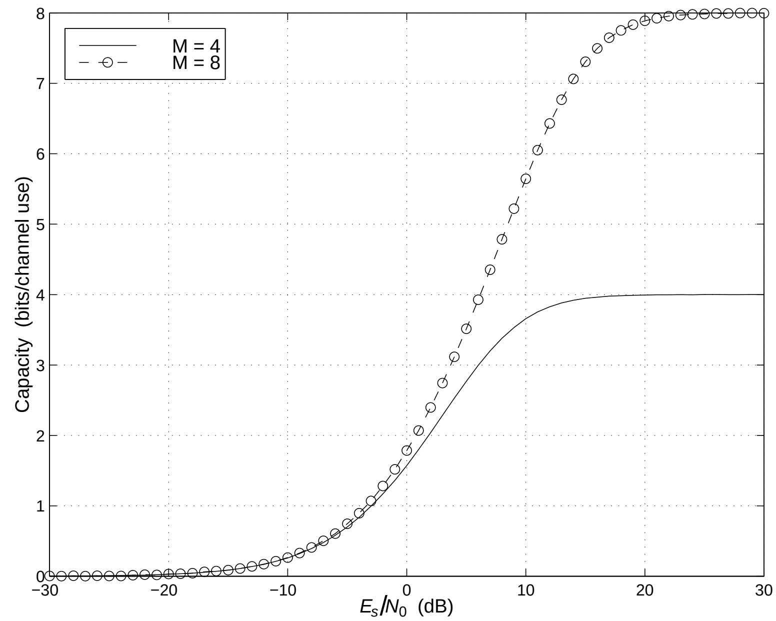

Fig. 5. Capacity for two systems: $M = 4$ and $M = 8$ while $Q = N = 2$.

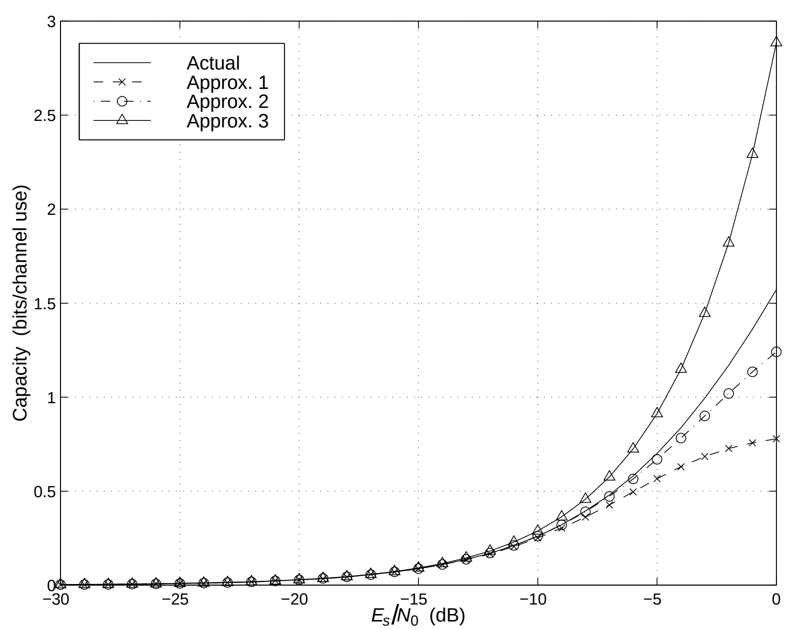

Fig. 6. Comparison between precise Monte Carlo simulation and the three approximations: $M = 4,\ N = 2,\ Q = 2$.

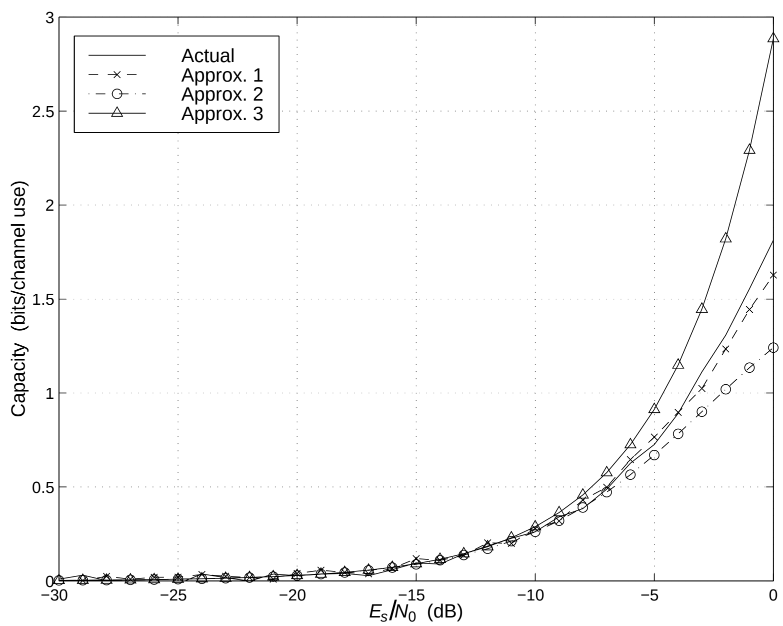

Fig. 7. Comparison between precise Monte Carlo simulation and the three approximations: $M = 4,\ N = 2,\ Q = 16$.

CHAPTER IV

UNIQUENESS OF CAPACITY-ACHIEVING INPUT DISTRIBUTION

Finding the capacity-achieving prior distribution has always been an integral part in determining channel capacity. In some sense, this optimum input distribution is the starting point of capacity computation and/or analysis. Sometimes it is rather straightforward to obtain the optimum input, e.g., for an additive white Gaussian noise (AWGN) channel with no constraint on its input alphabet [4]. Other times, the task is more involved but still doable. A case in point is the Gaussian vector for a Rayleigh MIMO channel [1]. The well-known Blahut-Arimoto algorithm [38, 39] is another good example, in which a numerical rather than analytical input distribution is obtained iteratively. Yet for many (if not most) other channels, it is mathematically intractable to identify the capacity-achieving input distribution, even for a seemingly simple channel, such as the ISI channel with a finite input alphabet [40]. In this case a lower bound on the capacity is sought since the optimum input distribution is unavailable.

Until recently, the capacity-achieving prior distribution has received relatively little attention compared to capacity itself. Yet, it has gradually been realized that the prior distribution deserves consideration as well because, to some extent, of the difficulty in determining the capacity of MIMO and ISI channels with a finite input alphabet. For the former, even though it is already known thanks to Telatar [1] that the ergodic capacity achieving input distribution for a Rayleigh fading MIMO channel is Gaussian, its counterpart for a discrete-alphabet MIMO channel is still very much in the air. The outage capacity achieving distribution is even more elusive. For example, Telatar simply used the same Gaussian distribution in his outage capacity analysis as in the ergodic capacity derivations (implicitly) based on the conjecture that the same Gaussian distribution achieves the outage capacity as well even though the hypothesis might be untrue. Baccarelli assumed a uniform input distribution in his effort to derive some lower bounds on capacity for a MIMO channel with PSK and QAM inputs [22]. As for the case of the ISI channel capacity, many researchers tried to get around the problem of finding the capacity-achieving input distribution by assuming a uniformly distributed input since the so-called i.u.d. (independent and uniformly distributed) capacity is expectedly close to the true capacity for practical channels [41, 42]. So, strictly speaking, the i.u.d. capacity is a tight lower bound for the ISI channel capacity.

Certain results were announced in some newly finished work. Chan *et al.* derived a necessary and sufficient condition for a capacity-achieving input probability distribution for a conditionally Gaussian channel (including AWGN and Rayleigh fading MIMO channels) subjected to a bounded-input constraint [43]. They further claimed that the probability measure was discrete under certain conditions. Huang and Meyn focused on the AWGN channel only and concluded that the optimal distribution, even if continuous, could be well approximated by a simple, discrete distribution that nearly achieves capacity [44]. Fozunbal *et al.* put capacity analysis for continuous-alphabet channels in a unified analytical framework and were able to derive equations that fit various scenarios [45].

In this chapter, we will study the optimal input distribution from a different perspective. More specifically, we want to show the uniqueness of the capacity-achieving prior distribution for a class of channels, including MIMO fading channels and ISI channels, with a discrete input alphabet. The work by Fozunbal *et al.*, as its title suggests, focuses on the continuous-alphabet case. Its results do not necessarily carry over to the discrete case, which is our target. At first glance, the problem might seem trivial as mutual information is proven to be a concave function of the prior distribution for a general channel [4]. However, the *strict* concavity of this function, to the authors' best knowledge, has never been established universally. Hence we believe it needs to be proven for any individual channel, which is the motivation behind our work.

A. Mathematical Formulation

Since the problem discussed in this chapter is more generalized than the one in Chapter III, we use a slightly different notation to avoid confusion. The channel model is described in matrix form with the following simple formula:

$$\boldsymbol{y} = \boldsymbol{\mathcal{H}}\boldsymbol{x} + \boldsymbol{n}, \tag{4.1}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the channel input and output, respectively, $\boldsymbol{n}$ is the channel noise assumed to be i.i.d. zero-mean Gaussian with components of variance σ^2, and independent of both $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{x}$; and $\boldsymbol{\mathcal{H}}$ characterizes the channel state. Depending on what channel (4.1) describes, MIMO or ISI, $\boldsymbol{\mathcal{H}}$ has different definitions, which will be given accordingly later. Note that, unlike the channel model assumed in Chapter III, here $\boldsymbol{x}$ is not necessarily a point in a PSK signal constellation as long as it belongs to a finite input alphabet. Let $\mathbb{X}$ be the input alphabet. Then since a discrete-alphabet constraint is imposed on $\boldsymbol{x}$, the cardinality of $\mathbb{X}$ is finite and the distribution of $\boldsymbol{x}$ can be characterized by a probability mass function (PMF) instead of a probability density function (PDF). The PMF of $\boldsymbol{x}$ will be interchangeably denoted by $p(\boldsymbol{x})$ and $\boldsymbol{p}$.

B. MIMO Channel

Here we focus on the MIMO fading channel with perfect channel state information at the receiver and the optimal input distribution that achieves its ergodic capacity. In this case, $\boldsymbol{\mathcal{H}}$ in (4.1) is a matrix whose entries are continuous complex random variables and independent of $\boldsymbol{x}$. Suppose the MIMO system has M transmit antennas and N receive antennas. Then the sizes of $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{\mathcal{H}}$, and $\boldsymbol{n}$ are $M \times 1$, $N \times 1$, $N \times M$, $N \times 1$, respectively. Let the input alphabet size (or cardinality of $\mathbb{X}$) be $\mathcal{S}$.

The task is to show that the mutual information $I(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{\mathcal{H}}))$ is a *strictly* concave function of prior distribution $p(\boldsymbol{x})$. The chain rule and independence between $\boldsymbol{x}$ and $\boldsymbol{\mathcal{H}}$ give [1]

$$I(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{\mathcal{H}})) = I(\boldsymbol{x}; \boldsymbol{y} \mid \boldsymbol{\mathcal{H}}) = \mathrm{E}_{\boldsymbol{\mathcal{H}}} \left[I(\boldsymbol{x}; \boldsymbol{y} \mid \boldsymbol{\mathcal{H}} = \mathbf{H}) \right].$$

In the following the condition $\boldsymbol{\mathcal{H}} = \mathbf{H}$ will be replaced by $\mathbf{H}$ for convenience where no ambiguity arises.

Based on all assumptions on $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{\mathcal{H}}$, we have

$$I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H}) = -N \log_2 e - \int_{\mathbb{C}^N} f(\boldsymbol{y} \mid \mathbf{H}) \log_2 \left[\sum_{\boldsymbol{x} \in \mathbb{X}} p(\boldsymbol{x}) \exp \left(-\frac{\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2}{\sigma^2} \right) \right] d\boldsymbol{y} \tag{4.2}$$

where $\mathbb{C}^N$ denotes the N-dimensional complex space and

$$f(\boldsymbol{y} \mid \mathbf{H}) = \sum_{\boldsymbol{x} \in \mathbb{X}} p(\boldsymbol{x}) \left(\frac{1}{\pi \sigma^2} \right)^2 \exp \left(-\frac{\|\boldsymbol{y} - \mathbf{H}\boldsymbol{x}\|^2}{\sigma^2} \right). \tag{4.3}$$

Substituting (4.3) into (4.2) and after some simple mathematical manipulation, we have

$$I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H}) = -N \log_2 e - \beta \int_{\mathbb{C}^N} \sum_{\boldsymbol{x} \in \mathbb{X}} p(\boldsymbol{x}) e^{-\|\boldsymbol{z} - \alpha \mathbf{H}\boldsymbol{x}\|^2} \ln \left(\sum_{\boldsymbol{x} \in \mathbb{X}} p(\boldsymbol{x}) e^{-\|\boldsymbol{z} - \alpha \mathbf{H}\boldsymbol{x}\|^2} \right) d\boldsymbol{z}, \tag{4.4}$$

where $\alpha = 1/\sigma$ and β is a positive constant. Let $\boldsymbol{p} = p(\boldsymbol{x}) = (p_1, \ldots, p_{\mathcal{S}})$ be the probability vector for $\boldsymbol{x}$. Then $\sum_{k=1}^{\mathcal{S}} p_k = 1$, $p_k \geq 0$, which implies the convexity of the feasible region of $\boldsymbol{p}$. We intend to show that, for a fixed but arbitrary $\mathbf{H}$, $I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H})$ is a strictly concave function of $\boldsymbol{p}$, or equivalently, $J(\boldsymbol{p}) \triangleq \int f_{\boldsymbol{z}}(\boldsymbol{p}) \ln f_{\boldsymbol{z}}(\boldsymbol{p})\, d\boldsymbol{z}$, where $f_{\boldsymbol{z}}(\boldsymbol{p}) = \sum_{k=1}^{\mathcal{S}} p_k \exp\left(-\|\boldsymbol{z} - \alpha \mathbf{H} \boldsymbol{x}_k\|^2\right)$, is a strictly convex function of $\boldsymbol{p}$.

The partial derivative of $J(\boldsymbol{p})$ w.r.t. p_i is

$$\begin{aligned}\frac{\partial J(\boldsymbol{p})}{\partial p_i} &= \int_{\mathbb{C}^N} \left[\frac{\partial f_{\boldsymbol{z}}(\boldsymbol{p})}{\partial p_i} \ln f_{\boldsymbol{z}}(\boldsymbol{p}) + \frac{\partial f_{\boldsymbol{z}}(\boldsymbol{p})}{\partial p_i}\right] d\boldsymbol{z} \\ &= \int_{\mathbb{C}^N} \exp(-\|\boldsymbol{z} - \alpha \mathbf{H} \boldsymbol{x}_i\|^2) \ln f_{\boldsymbol{z}}(\boldsymbol{p})\, d\boldsymbol{z} + \pi^N\end{aligned}$$

and the partial second derivative is

$$\frac{\partial^2 J(\boldsymbol{p})}{\partial p_i \partial p_j} = \int_{\mathbb{C}^N} \exp(-\|\boldsymbol{z} - \alpha \mathbf{H} \boldsymbol{x}_i\|^2) \exp(-\|\boldsymbol{z} - \alpha \mathbf{H} \boldsymbol{x}_j\|^2) \frac{1}{f_{\boldsymbol{z}}(\boldsymbol{p})}\, d\boldsymbol{z}. \tag{4.5}$$

Define

$$g_i(\boldsymbol{z}, \boldsymbol{p}) \triangleq \exp(-\|\boldsymbol{z} - \alpha \mathbf{H} \boldsymbol{x}_i\|^2) / \sqrt{f_{\boldsymbol{z}}(\boldsymbol{p})}, \tag{4.6}$$

which is clearly a smooth function. Then (4.5) becomes

$$\frac{\partial^2 J(\boldsymbol{p})}{\partial p_i \partial p_j} = \int_{\mathbb{C}^N} g_i(\boldsymbol{z}, \boldsymbol{p}) g_j(\boldsymbol{z}, \boldsymbol{p})\, d\boldsymbol{z} \triangleq g_{ij}. \tag{4.7}$$

Denote the Hessian matrix of $J(\boldsymbol{p})$ by $\nabla^2 J(\boldsymbol{p})$ whose (i, j)th $(1 \leq i, j \leq \mathcal{S})$ entry, defined as g_{ij} in (4.7), is specified in (4.5).

Next we try to substantiate the positive-definiteness of $\nabla^2 J(\boldsymbol{p})$ for any fixed but arbitrary $\boldsymbol{p}$, which in turn leads to the strict convexity of $J(\boldsymbol{p})$. It is recalled that one necessary and sufficient condition for positive-definiteness of an $\mathcal{S} \times \mathcal{S}$ matrix is that its $r \times r$ $(\forall 1 \leq r \leq \mathcal{S})$ principal submatrix has a positive determinant [35]. Let $\mathbf{G}_r = [g_{ij}]_{r \times r}$. According to (4.7), each entry of $\mathbf{G}_r$ is an inner product, $\langle g_i(\boldsymbol{z}, \boldsymbol{p}), g_j(\boldsymbol{z}, \boldsymbol{p}) \rangle = \int_{\mathbb{C}^N} g_i(\boldsymbol{z}, \boldsymbol{p}) g_j(\boldsymbol{z}, \boldsymbol{p})\, d\boldsymbol{z}$ (the validity of this inner product space is easy to verify). $\mathbf{G}_r$ is the Gram matrix of $g_1(\boldsymbol{z}, \boldsymbol{p}), \ldots, g_r(\boldsymbol{z}, \boldsymbol{p})$. Therefore, $\det(\mathbf{G}_r) \geq 0$ with equality if and only if $g_1(\boldsymbol{z}, \boldsymbol{p}), \ldots, g_r(\boldsymbol{z}, \boldsymbol{p})$ are linearly dependent [46, pp. 178, Theorem 8.7.2].

It should be noted that in the following $\boldsymbol{p}$ is fixed but arbitrary. To prove $g_1(\boldsymbol{z}, \boldsymbol{p}), \ldots, g_r(\boldsymbol{z}, \boldsymbol{p})$ are linearly independent, we start from the definition of linearly independent functions, i.e., $\sum_k \lambda_k g_k(\boldsymbol{z}, \boldsymbol{p}) \equiv 0$ for all different values of $\boldsymbol{z}$ implies $\lambda_1 = \ldots = \lambda_r = 0$. Under the assumption that $\sum_k \lambda_k g_k(\boldsymbol{z}, \boldsymbol{p}) \equiv 0$ for all $\boldsymbol{z}$'s, we have $\sum_k \lambda_k g_k(\boldsymbol{z}_m, \boldsymbol{p}) = 0$ for r specially picked values of $\boldsymbol{z}$, i.e., $\{\boldsymbol{z}_m \mid m = 1, \ldots, r\}$. From there we will try to prove $\lambda_1 = \ldots = \lambda_r = 0$, which in turn enables $\{g_k(\boldsymbol{z}, \boldsymbol{p}) \mid 1 \leq k \leq r\}$ to meet the definition of linearly independent functions. Toward this end, let

$$\boldsymbol{z}_k = \alpha \mathbf{H} \boldsymbol{x}_k, \quad k = 1, \ldots, r. \tag{4.8}$$

Construct r vectors by evaluating $g_1(\boldsymbol{z}, \boldsymbol{p}), \ldots, g_r(\boldsymbol{z}, \boldsymbol{p})$ at $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r$, respectively. The result is an $r \times r$ matrix,

$$[g_i(\boldsymbol{z}_k, \boldsymbol{p})]_{r \times r} = \begin{pmatrix} g_1(\boldsymbol{z}_1, \boldsymbol{p}) & g_1(\boldsymbol{z}_2, \boldsymbol{p}) & \cdots & g_1(\boldsymbol{z}_r, \boldsymbol{p}) \\ g_2(\boldsymbol{z}_1, \boldsymbol{p}) & g_2(\boldsymbol{z}_2, \boldsymbol{p}) & \cdots & g_2(\boldsymbol{z}_r, \boldsymbol{p}) \\ \vdots & \vdots & \ddots & \vdots \\ g_r(\boldsymbol{z}_1, \boldsymbol{p}) & g_r(\boldsymbol{z}_2, \boldsymbol{p}) & \cdots & g_r(\boldsymbol{z}_r, \boldsymbol{p}) \end{pmatrix}. \tag{4.9}$$

Combining (4.6), (4.8), and (4.9), we have

$$[g_i(\boldsymbol{z}_k,\boldsymbol{p})]_{r\times r} = \frac{1}{\sqrt{\prod_{k=1}^{r} f_{\boldsymbol{z}_k}(\boldsymbol{p})}} \begin{pmatrix} e^{-\|\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_1)\|^2} & \cdots & e^{-\|\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_1)\|^2} \\ \vdots & \ddots & \vdots \\ e^{-\|\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_r)\|^2} & \cdots & e^{-\|\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_r)\|^2} \end{pmatrix}$$
$$= \frac{1}{\sqrt{\prod_{k=1}^{r} f_{\boldsymbol{z}_k}(\boldsymbol{p})}} \begin{pmatrix} e^{-\|\Re\{\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_1)\}\|^2} & \cdots & e^{-\|\Re\{\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_1)\}\|^2} \\ \vdots & \ddots & \vdots \\ e^{-\|\Re\{\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_r)\}\|^2} & \cdots & e^{-\|\Re\{\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_r)\}\|^2} \end{pmatrix}$$
$$\circ \begin{pmatrix} e^{-\|\Im\{\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_1)\}\|^2} & \cdots & e^{-\|\Im\{\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_1)\}\|^2} \\ \vdots & \ddots & \vdots \\ e^{-\|\Im\{\alpha\mathbf{H}(\boldsymbol{x}_1-\boldsymbol{x}_r)\}\|^2} & \cdots & e^{-\|\Im\{\alpha\mathbf{H}(\boldsymbol{x}_r-\boldsymbol{x}_r)\}\|^2} \end{pmatrix}, \tag{4.10}$$

where the symbol $\circ$ denotes Hadamard product [35], i.e., componentwise multiplication of two matrices of the same size. Note that $\mathbf{H}$ in (4.10) is a realization of the channel matrix $\boldsymbol{\mathcal{H}}$. So there is no guarantee that $\Re\{\alpha\mathbf{H}(\boldsymbol{x}_i-\boldsymbol{x}_j)\} \neq 0$ and $\Im\{\alpha\mathbf{H}(\boldsymbol{x}_i-\boldsymbol{x}_j)\} \neq 0$ for $1 \leq i \neq j \leq r$. However, given the nature of $\boldsymbol{\mathcal{H}}$, a matrix whose entries are continuous random variables, we have

$$\Pr\left(\Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\}, \Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\} \mid 1 \leq i \neq j \leq r\right) = 1. \tag{4.11}$$

In other words, the $\boldsymbol{\mathcal{H}}$'s that contribute in the expectation over $\boldsymbol{\mathcal{H}}$ in $\mathrm{E}_{\boldsymbol{\mathcal{H}}}\left[I(\boldsymbol{x};\boldsymbol{y} \mid \boldsymbol{\mathcal{H}}=\mathbf{H})\right]$ are such that $\Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\}$ and $\Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\}$, $\forall 1 \leq i \neq j \leq r$, with probability one. Note that $\mathrm{E}_{\boldsymbol{\mathcal{H}}}\left[I(\boldsymbol{x};\boldsymbol{y} \mid \boldsymbol{\mathcal{H}}=\mathbf{H})\right]$ is an integral over the integration variable $\mathbf{H}$. Then in this integral we only need to keep $\mathbf{H}$ that satisfies $\Re\{\mathbf{H}\boldsymbol{x}_i\} \neq \Re\{\mathbf{H}\boldsymbol{x}_j\}$ and $\Im\{\mathbf{H}\boldsymbol{x}_i\} \neq \Im\{\mathbf{H}\boldsymbol{x}_j\}$ simultaneously, which also applies to (4.10).

Lemma 3.1 in [47] and the Schur product theorem [35, Theorem 7.5.3, pp. 458] are crucial in proving the nonsingularity of the matrix in (4.10). Hence they are repeated below with a little rephrasing.

Lemma 1. *[47] The* $n \times n$ *matrix whose elements have the values* $\exp(-\|\boldsymbol{w}_j - \boldsymbol{w}_k\|^2)$*,* $j, k = 1, 2, \ldots, n$*, is positive definite for all choices of different* $\boldsymbol{w}_k$*'s in* $\mathbb{R}^d$*, where* d *is any positive integer.*

Lemma 2. *[35] If* $\mathbf{A}$ *and* $\mathbf{B}$ *are* $n \times n$ *positive semidefinite matrices, then* $\mathbf{A} \circ \mathbf{B}$ *is also positive semidefinite. Further, if both* $\mathbf{A}$ *and* $\mathbf{B}$ *are positive definite, then so is* $\mathbf{A} \circ \mathbf{B}$*.*

With the help of Lemma 1 and Lemma 2, it is straightforward to show that the square matrix $[g_i(\boldsymbol{z}_k, \boldsymbol{p})]$ $(1 \leq i, k \leq r)$ in (4.10) is positive definite and thus invertible. The nonsingularity of matrix $[g_i(\boldsymbol{z}_k, \boldsymbol{p})]$ leads to the linear independence of $g_1(\boldsymbol{z}, \boldsymbol{p}), \ldots, g_r(\boldsymbol{z}, \boldsymbol{p})$, which in turn implies $\det(\mathbf{G}_r) > 0$. Then it follows that $\nabla^2 J(\boldsymbol{p})$ is positive definite and $J(\boldsymbol{p})$ is strictly convex over the feasible region of $\boldsymbol{p}$.

So far $I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H})$ has been shown to be a strictly concave function of $\boldsymbol{p}$ for any $\mathbf{H}$. For convenience let us introduce a new function $u(\boldsymbol{p}, \mathbf{H}) \triangleq I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H})$, which is continuous due to the definition of mutual information. Then

$$I(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{\mathcal{H}})) = \mathrm{E}_{\boldsymbol{\mathcal{H}}}\left[I(\boldsymbol{x}; \boldsymbol{y} \mid \mathbf{H})\right] = \int f(\mathbf{H}) u(\boldsymbol{p}, \mathbf{H})\, d\mathbf{H}.$$

Now, for any $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, and any pair of scalars λ_1, λ_2 such that $0 < \lambda_1, \lambda_2 < 1$ and $\lambda_1 + \lambda_2 = 1$, we have

$$\int f(\mathbf{H}) \left\{u(\lambda_1 \boldsymbol{p}_1 + \lambda_2 \boldsymbol{p}_2, \mathbf{H}) - [\lambda_1 u(\boldsymbol{p}_1, \mathbf{H}) + \lambda_2 u(\boldsymbol{p}_2, \mathbf{H})]\right\} d\mathbf{H} > 0, \tag{4.12}$$

because the integrand in (4.12) is positive and continuous w.r.t. $\mathbf{H}$. This demonstrates the strict concavity of $I(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{\mathcal{H}}))$ w.r.t. $\boldsymbol{p}$. Since the channel capacity is the maximum of $I(\boldsymbol{x}; (\boldsymbol{y}, \boldsymbol{\mathcal{H}}))$ over $\boldsymbol{p}$, the strict concavity of the mutual information implies the uniqueness of the capacity-achieving input distribution.

C. ISI Channel

Here we only focus on the type of ISI channels existing in magnetic recording, i.e., with real and deterministic channel coefficients, since the one for multipath fading channels is mathematically identical to the case discussed in Section B. For the ISI channel under discussion with a binary input alphabet, the original channel model in the matrix form introduced in Section A

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \boldsymbol{n} \tag{4.13}$$

(note that $\boldsymbol{\mathcal{H}}$ in (4.1) is replaced by $\mathbf{H}$ in (4.13) because the channel matrix is now assumed to be deterministic) can be reformulated as

$$y_k = \sum_{i=0}^{L} h_k x_{k-i} + n_k, \tag{4.14}$$

where $x_k \in \{\pm 1\}$ and $h_0,\ h_L \neq 0$. It should be kept in mind that even though (4.13) looks essentially the same as (4.1), the two are indeed different because of some underlying differences between ISI and MIMO channels. For example, in (4.13) all variables are real-valued, $\mathbf{H}$ is deterministic, and each component of $\boldsymbol{x}$ is binary.

The case for the ISI channel follows the same line as that for the MIMO channel. So we will skip the steps common to both and instead will only explain those specific to the ISI channel. Based on (4.14), we have, for $k > L$,

$$\begin{pmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ \vdots \end{pmatrix} = \begin{pmatrix} h_L & h_{L-1} & \cdots & h_1 & h_0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & h_L & h_{L-1} & \cdots & h_1 & h_0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & h_L & h_{L-1} & \cdots & h_1 & h_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} x_{k-L} \\ x_{k-L+1} \\ \vdots \\ x_{k-1} \\ x_k \\ \vdots \end{pmatrix} + \begin{pmatrix} n_k \\ n_{k+1} \\ n_{k+2} \\ \vdots \end{pmatrix}. \tag{4.15}$$

Comparing (4.15) to (4.14), we can easily obtain the definitions of $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{n}$, and particularly $\mathbf{H}$. From (4.15) the channel matrix $\mathbf{H}$ is clearly an upper triangular matrix with diagonal entries $h_L \neq 0$, which makes $\mathbf{H}$ full-column-rank.

Even though $\mathbf{H}$ is now deterministic, the mutual information $I(\boldsymbol{x}; \boldsymbol{y})$ is still (4.4) with some slight adjustments. For example, the constant term is different and of course the integration region is now over a multi-dimensional real space. The rest of the proof for the uniqueness is almost identical to those procedures presented in Section B. The minor differences are addressed below. In Section B the continuous random nature of $\boldsymbol{\mathcal{H}}$ helped prove $\Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Re\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\}$ and $\Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_i\} \neq \Im\{\boldsymbol{\mathcal{H}}\boldsymbol{x}_j\}$, $\forall 1 \leq i \neq j \leq r$, with probability one. Here we take advantage of the fact that $\mathbf{H}$ is of full column rank, which implies $\mathbf{H}(\boldsymbol{x}_i - \boldsymbol{x}_j) \neq \mathbf{0}$. Since now $\boldsymbol{x}$, $\boldsymbol{y}$, and $\mathbf{H}$ are all real, the part for imaginary parts in (4.10) should be discarded and only Lemma 1 is applicable.

D. Discussion

Though the channel noise $\boldsymbol{n}$ is assumed to be i.i.d. Gaussian in Section A, it could have other forms as well. For example, when $\boldsymbol{n}$ is colored Gaussian, the exponent in

(4.2) and (4.3) needs to include the covariance matrix of $\boldsymbol{n}$. Some proper adjustments should also be made accordingly in the rest of Section B. In this case Lemma 1 is no longer applicable. But Theorem 3.3 in [47], which implies positive-definiteness for a matrix whose entries assume a more generic functional form than the exponential quadratic function stated in Lemma 1, may very well be used to prove the positive definiteness of the matrix in (4.9). If the noise $\boldsymbol{n}$ is non-Gaussian, it becomes more difficult. We don't have a complete proof for the uniqueness in this case. But our conjecture is that it still holds. Here is the argument. The Gaussian exponential function in the integrand in (4.5) will be replaced by some other function, i.e., the definition of $g_i(\boldsymbol{z}, \boldsymbol{p})$ will be different. But the new $\{g_i(\boldsymbol{z}, \boldsymbol{p})\}$ may still be linearly independent because, intuitively, we can always find r different values of $\boldsymbol{z}$, $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r\}$, that make the new $[g_i(\boldsymbol{z}_k, \boldsymbol{p})]_{r \times r}$ invertible as $\boldsymbol{z}_k$ can be any point in the N-dimensional complex space.

Note that we never specified in Section B the statistics of the MIMO channel. It is because that was irrelevant in the proof of uniqueness. Only the continuous random nature of $\boldsymbol{\mathcal{H}}$ was utilized. And the entries of $\boldsymbol{\mathcal{H}}$ do not have to be i.i.d. The only constraint on $\boldsymbol{\mathcal{H}}$ is (4.11). Statistical independence among column vectors of $\boldsymbol{\mathcal{H}}$ is good enough for it. Yet even this is a sufficient but not necessary condition.

It needs to be pointed out that when studying ISI channel coding, many researchers only focus on input $\boldsymbol{x}$ with i.i.d. components. To them, it makes more sense to add an i.i.d. input constraint in any consideration of the ISI channel capacity. Unfortunately, this additional constraint destroys the convexity of the feasible region of the input distribution $\boldsymbol{p}$ and thus the proof provided in this paper becomes invalid even though the uniqueness might still hold.

Based on the above discussion, we are able to reach the following conclusion. Any channel that can be described by (4.1) with a complex discrete alphabet for input $\boldsymbol{x}$, additive Gaussian noise $\boldsymbol{n}$, and a continuous random channel matrix $\boldsymbol{\mathcal{H}}$ whose columns are statistically independent, has a unique capacity-achieving prior distribution. So does a channel that can be described by (4.13) with a real discrete input alphabet and a real deterministic channel matrix $\mathbf{H}$ of full column rank. Furthermore, for the same models except that the additive noise is non-Gaussian, we conjecture an identical conclusion.

CHAPTER V

CONCLUSION

In this thesis we studied MIMO channel capacity. Special attention was paid to the capacity of a MIMO fading channel with i.i.d. Rayleigh fading subject to PSK input constraint. We showed that the uniform input distribution achieves capacity and derived a capacity formula which can be easily computed through Monte Carlo simulation. To facilitate generating the correlated Gaussian variables needed we derived analytical expressions for the nonzero eigenvalues and their associated eigenvectors of the requisite covariance matrix. These analytical expressions lead to approximations to the MIMO capacity at low SNR, two of which, (3.30) and (3.31), are closed-form. Both are in agreement with results derived from the theory of capacity per unit cost [36, 48]. They also enable us to conclude there is no need to apply Gaussian input at low SNR because any PSK input can achieve about the same data rate. Another point worth mentioning is that Fig. (5) implies in order to enhance the information-theoretic performance of a system with a PSK constellation and medium to high SNR, more antennas should be used on the transmitting side than the receiving side when the total number of the two is fixed. This verifies a similar conclusion reached by Müller under the binary input constraint [25]. On the other hand, (3.31) suggests the contrary at low SNR, i.e., more antennas should be put on the receiving end.

During the process of reaching the final expression for the PSK constrained capacity, we learned that uniform input distribution is the unique capacity-achieving prior distribution. We then were able to extend this conclusion to a broad range of channels, including a generic MIMO channel with an arbitrary input alphabet and ISI channels.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[3] G. J. Foschini, "Layered space-time architecture for wireless communcation in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, pp. 41–59, 1996.

[4] T. Cover and J. Thomas, *Elements of Information Theory.* New York: John Wiley & Sons, Inc., 1991.

[5] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, 1999.

[6] L. Zheng and D. N. C. Tse, "Information theoretic limits for non-coherent multi-antenna communications," *Proc. IEEE Wireless Commun. Network. Conf.*, Chicago, IL, 2000, pp. 18–22.

[7] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multi-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, 2002.

[8] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281–302, 2003.

[9] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–702, 2003.

[10] M. A. Kamath, B. L. Hughes, and X. Yu, "Gaussian approximations for the capacity of MIMO Rayleigh fading channels," *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, 2002, pp. 614–618.

[11] P. J. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," *Proc. IEEE ICC'2002*, New York, 2002, pp. 406–410.

[12] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, 2002.

[13] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Commun.*, vol. 48, pp. 1245–1248, 2000.

[14] A. M. Sengupta and P. P. Mitra, "Capacity of multivariate channels with multiplicative noise: I. Random matrix techniques and large-N expansions for full transfer matrices," *LANL arXiv:physics/0010081*, vol. 1, 2000. Downloadable at *http://arxiv.org/pdf/physics/0010081v1*.

[15] R. Janaswamy, "Analytical expressions for the ergodic capacities of certain MIMO systems by the Mellin transform," *Proc. IEEE GLOBECOM'2003*, San Francisco, CA, 2003, pp. 287–291.

[16] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, "Capacity of MIMO channels with one-sided correlation," *Proc. IEEE ISSSTA'2004*, Sydney, Australia, 2004, pp. 515–519.

[17] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, “Mutual information and eigenvalue distribution of MIMO Ricean channels,” *Proc. ISITA’2004*, Parma, Italy, 2004.

[18] M. Chiani, M. Z. Win, and A. Zanella, “On the capacity of spatially correlated MIMO Rayleigh-fading channels,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2363–2371, 2003.

[19] P. J. Smith, S. Roy, and M. Shafi, “Capacity of MIMO systems with semicorrelated flat fading,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2781–2788, 2003.

[20] M. Kang, M.-S. Alouini, “Impact of correlation on the capacity of MIMO channels,” *Proc. IEEE ICC’2003*, Anchorage, AK, 2003, pp. 2623–2627.

[21] M. Kang, L. Yang, and M.-S. Alouini, “Capacity of MIMO Rician channels with multiple correlated Rayleigh co-channel interferers,” *Proc. IEEE GLOBECOM’2003*, San Francisco, CA, 2003, pp. 1119–1123.

[22] E. Baccarelli, “Evaluation of the reliable data rates supported by multiple-antenna coded wireless links for QAM transmission,” *IEEE J. Select. Areas Commun.*, vol. 19, pp. 295–304, 2001.

[23] B. M. Hochwald and S. ten Brink, “Achieving near-capacity on a multiple-antenna channel,” *IEEE Trans. Commun.*, vol. 51, pp. 389–399, 2003.

[24] A. Lapidoth and S. M. Moser, “Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2426–2467, 2003.

[25] R. R. Müller, “Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation,” *IEEE Trans. Signal Processing*,

vol. 51, pp. 2821–2828, 2003.

[26] W. He and C. N. Georghiades, "Computing the capacity of a MIMO fading channel under PSK signaling," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1794–1803, 2005.

[27] H. Ge, K. D. Wong, M. Barton, and J. C. Liberti, "Statistical characterization of multiple-input multiple-output (MIMO) channel capacity," *Proc. IEEE WCNC 2002*, Orlando, FL, 2002, pp. 789–793.

[28] R. K. Mallik, "The pseudo-Wishart distribution and its application to MIMO systems," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2761–2769, 2003.

[29] Z. Shi, C. Schlegel, and B. Farhang-Boroujeny, "Efficient capacity outage approximations for MIMO channels," preprint.

[30] A. Edelman, "Eigenvalues and condition numbers of random matrices," *SIAM J. Matrix Anal. Appl.*, vol. 9, pp. 543–560, 1988.

[31] M. Chiani, M. Z. Win, A. Zanella, R. K. Mallik, and J. H. Winters, "Bounds and approximations for optimum combining of signals in the presence of multiple cochannel interferers and thermal noise," *IEEE Trans. Commun.*, vol. 51, pp. 296–307, 2003.

[32] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, 1999.

[33] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, pp. 55–67, 1982.

[34] R. A. Horn and C. R. Johsnson, *Topics in Matrix Analysis.* Cambridge, UK: Cambridge University Press, 1991.

[35] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, UK: Cambridge University Press, 1985.

[36] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319–1343, 2002.

[37] Ö. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 2784–2795, 2003.

[38] R. E. Blahut, "Computation of channel capacity and rate-distortion fucntions," *IEEE Trans. Inform. Theory*, vol. 18, pp. 460–473, 1972.

[39] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 18, pp. 14–20, 1972.

[40] S. Shamai and R. Laroita, "The intersymbol interference channel: lower bounds on capacity and channel precoding loss," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1388–1404, 1996.

[41] X. Ma, N. Varnica, and A. Kavčić, "Matched information rate codes for binary ISI channels," in *Proc. Int. Symp. on Informmation Theory*, Lausanne, Switzerland, July 2002, pp. 269.

[42] A. Kavčić, X. Ma, and N. Varnica, "Matched information rate codes for partial response channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 973–989, 2005.

[43] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2073–2088, 2005.

[44] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2336–2351, 2005.

[45] M. Fozunbal, S. W. McLaughlin, and R. W. Schafer, "Capacity analysis for continuous-alphabet channels with side information," Part I & II, *IEEE Trans. Inform. Theory*, vol. 51, pp. 3075–3101, 2005.

[46] P. J. Davis, *Interpolation and Approximation*, New York: Blaisdell Publishing Company, 1963.

[47] M. J. D. Powell, "The theory of radial basis function approximation in 1990," *Advances in Numerical Analysis Volume II* (W. Light ed.), Clarendon Press, Oxford, 1992.

[48] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, 1990.

VITA

Wenyan He came from China. Prior to joining Texas A&M University, he obtained his B.Eng. and M.Eng. degrees in biomedical engineering from Zhejiang University, China, in 1997 and 2000, respectively. His permanent address is:

Department of Electrical & Computer Engineering
Texas A&M University
College Station, TX 77840
He can be reached via email at *tamwenyan@gmail.com*.

The typist for this thesis was Wenyan He.