Capacity and Performance Analysis of Advanced Multiple Antenna Communication Systems

A THESIS SUBMITTED TO THE UNIVERSITY COLLEGE LONDON IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

By

Caijun Zhong

Department of Electronic and Electrical Engineering

University College London

London, United Kingdom

March 2010

Declaration

I, Caijun Zhong, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgments

Foremost, I thank my supervisor, Dr. Kai-Kit Wong, for giving me the opportunity to be a part of his research group. I am greatly indebted to him for his full support, constant encouragement, and guidance in the past three years. I just hope I can live up to his expectations in my future career.

I am very grateful to Dr. Gan Zheng, who has helped me a lot in the beginning of my PhD study. Special thanks to Dr. Shi Jin for introducing random matrix theory as well as collaborating on various topics, working with Shi has been fantastic. My thanks also goes to Dr. McKay Matthew, for the insightful discussion and collaboration.

Finally, I would like to thank my parents, my sister and brother-in-law for their love and support.

Abstract

Multiple-input multiple-output (MIMO) antenna systems have been shown to be able to substantially increase date rate and improve reliability without extra spectrum and power resources. The increasing popularity and enormous prospect of MIMO technology calls for a better understanding of the performance of MIMO systems operating over practical environments. Motivated by this, this thesis provides an analytical characterization of the capacity and performance of advanced MIMO antenna systems.

First, the ergodic capacity of MIMO Nakagami-m fading channels is investigated. A unified way of deriving ergodic capacity bounds is developed under the majorization theory framework. The key idea is to study the ergodic capacity through the distribution of the diagonal elements of the quadratic channel $\mathbf{HH}^{\dagger}$ which is relatively easy to handle, avoiding the need of the eigenvalue distribution of the channel matrix which is extremely difficult to obtain. The proposed method is first applied on the conventional point-to-point MIMO systems under Nakagami-m fading, and later extended to the more general distributed MIMO systems.

Second, the ergodic capacity of MIMO multi-keyhole and MIMO amplify-and-forward (AF) dual-hop systems is studied. A set of new statistical properties involving product of random complex Gaussian matrix, i.e., probability density function (p.d.f.) of an unordered eigenvalue, p.d.f. of the maximum eigenvalue, expected determinant and log-determinant, is derived. Based on these, analytical closed-form expressions for the ergodic capacity of the systems are obtained and the connection between the product channels and conventional point-to-point MIMO channels is also revealed.

Finally, the effect of co-channel interference is investigated. First, the performance of optimum combining (OC) systems operating in Rayleigh-product channels is analyzed based on novel closed-form expression of the cumulative distribution function (c.d.f.) of the maximum eigenvalue of the resultant channel matrix. Then, for MIMO Rician channels and MIMO Rayleigh-product channels, the ergodic capacity at low signal-to-noise ratio (SNR) regime is studied, and the impact of various system parameters, such as transmit and receive antenna number, Rician factor, channel mean matrix and interference-to-noise-ratio, is examined.

Publications

The following is a list of journal and conference publications produced during my Ph.D. study.

Journal papers

1. **C. Zhong**, K. K. Wong, and S. Jin, "Capacity bounds for MIMO Nakagami-m fading channels", **IEEE Transactions on Signal Processing**, vol. 57, no. 9, pp. 3613–3623, Sep. 2009.

2. **C. Zhong**, S. Jin, K. K. Wong, and M. R. McKay, "Outage analysis for optimal beamforming MIMO systems in multi-keyhole channels," to appear in **IEEE Transactions on Signal Processing**.

3. S. Jin, M. R. McKay, **C. Zhong**, and K. K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," to appear in **IEEE Transactions on Information Theory**.

4. **C. Zhong**, S. Jin, and K. K. Wong, "MIMO Rayleigh-product channels with co-channel interference," **IEEE Transactions on Communications**, vol. 57, no. 6, pp. 1824–1835, June 2009.

5. **C. Zhong**, G. Zheng, and K. K. Wong,"Feasibility conditions of linear multiuser MIMO antenna systems in the asymptotic regime," **IEEE Communications Letters**, vol. 11, no. 12, pp. 979–981, Dec. 2007.

6. **C. Zhong**, S. Jin, and K. K. Wong, "Dual-hop systems with noisy relay and interference-limited destination," to appear in **IEEE Transactions on Communications**.

Conference papers

1. **C. Zhong**, K. K. Wong, and S. Jin, "On the Ergodic Capacity of MIMO Nakagami-fading channels," in proceedings of **IEEE International Symposium on Information Theory (ISIT 2008)**, Toronto, Canada, July 2008, pp.131–135.

2. **C. Zhong**, S. Jin, K. K. Wong, and M. R. McKay, "Performance analysis of transmit beamforming in multi-keyhole MIMO channels," **IEEE International Conference on Communications (ICC 2009)**, Dresden, Germany, June 2009, pp. 1–5.

3. S. Jin, M. R. McKay, **C. Zhong**, and K. K. Wong, "Ergodic capacity analysis of amplify-and-forward MIMO dual-hop systems," **IEEE International Symposium on Information Theory (ISIT 2008)**, Toronto, Canada, July 2008, pp. 1903-1907.

4. **C. Zhong**, S. Jin, and K. K. Wong, "Performance analysis of a Rayleigh-product MIMO channel with receiver correlation and cochannel interference," **IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008)**, Las Vegas, USA, March 2008, pp. 2877–2880.

5. **C. Zhong**, S. Jin, and K. K. Wong, "Low SNR capacity of MIMO channels with a single interferer," accepted in **IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC 2009)**, Perugia, Italy, June 2009, pp. 578–582.

6. **C. Zhong**, S. Jin, and K. K. Wong, "Outage probability of dual-hop relay channels in the presence of interference," **IEEE 69th Vehicular Technology Conference (VTC2009-Spring)**, Barcelona, Spain, Apr. 2009, pp. 1–5.

7. **C. Zhong**, S. Jin, K. K. Wong, "Outage Probability Analysis for Systems Using Amplify-and-Forward Relays with Direct Link," **London Communications Symposium 2009**, London, UK, Sep. 2009.

Notations

Here, we introduce the notations adopted in this thesis. Unless otherwise clearly indicated, boldface upper-case letters generally denote matrices, while boldface lower-case letters generally denote column vectors.

$\in$	Belongs to.
$\sim$	Follows certain distribution.
$\succ$	Majorization relationship.
$\otimes$	Kronecker product.
$\triangleq$	Defined as.
π	3.1415926
$\sum$	Summation symbol
$\prod$	Product symbol
∞	Infinite symbol
!	Factorial
$\rightarrow$	Approach symbol
$[\mathbf{H}]_{i,j}$ or $\mathbf{H}_{ij}$	(i,j)-th element of matrix $\mathbf{H}$.
$\mathbf{0}_{m\times n}$	$m \times n$ matrix with all elements being zero.
$\mathbf{I}_n$	$n \times n$ Identity matrix.
$\mathcal{R}^m$,$\mathcal{C}^m$	Real and complex $m \times 1$ vector.
$\mathcal{R}^{m\times n}$, $\mathcal{C}^{m\times n}$	Real and complex $M \times N$ matrix.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector with mean $\mathbf{m}$ and covariance $\mathbf{C}$.
$\mathtt{E}(\cdot)$	Expectation.
$\mathbf{H}^T$	Transpose of matrix $\mathbf{H}$.
$\mathbf{H}^*$	Complex conjugate of matrix $\mathbf{H}$.
$\mathbf{H}^\dagger$	Conjugate transpose of matrix $\mathbf{H}$.
$\mathtt{tr}(\mathbf{H})$	Trace of matrix $\mathbf{H}$.
$\lvert\mathbf{H}\rvert$ or $\det(\mathbf{H})$	Determinant of square matrix $\mathbf{H}$.
$\min(x, y)$	Minimum of x and y.
$\max(x, y)$	Maximum of x and y.

$I_n(\cdot)$	Bessel function of the first kind
$B(\cdot,\cdot)$	Beta function
$U(\cdot,\cdot,\cdot)$	Confluent hypergeometric function of the second kind
$\psi(\cdot)$	Digamma function
$\mathsf{exp}(\cdot)$	Exponential function
e^x	Exponential function
$E_n(\cdot)$	Exponential integral function of order n
$\Gamma(\cdot)$	Gamma function
$\log_2(\cdot)$	Logarithm in base 2.
$G_{m,n}^{s,t}(\cdot)$	Meijer-G function
$K_n(\cdot)$	Modified Bessel function of the second kind
$\ln(\cdot)$	Natural logarithm function
$\sqrt{\cdot}$	Square root function
$Q(\cdot)$	Standard Gaussian Q-function

Abbreviations

1G	1st Generation Mobile Communication Systems
2G	2nd Generation Mobile Communication Systems
3G	3rd Generation Mobile Communication Systems
4G	4th Generation Mobile Communication Systems
3GPP-LTE	Third Generation Partnership Project
AF	Amplify-and-Forward
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BFSK	Binary Frequency Shift-Keying
BPSK	Binary Phase Shift-Keying
CCI	Co-Channel Interference
CDF	Cumulative Distribution Function
C-MIMO	Co-located MIMO
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
D-MIMO	Distributed MIMO
DSL	Digital Subscriber Line
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communication
IEEE	Institute of Electrical and Electronics Engineers
INR	Interference to Noise Ratio
LAN	Local Area Network
LOS	Line of Sight
MAN	Metropolitan Area Network
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
OC	Optimum Combining
PAM	Phase Amplitude Modulation
PAN	Personal Area Network
PDF	Probability Density Function
RMT	Random Matrix Theory

SDMA	Space Division Multiple Access
SISO	Single Input Single Output
SIMO	Single Input Multiple Output
SM	Spatial Multiplexing
SER	Symbol Error Rate
SINR	Signal to Interference and Noise Ratio
STBC	Space Time Block Code
SNR	Signal Noise Rate
SVD	Singular Value Decomposition
UMTS	Universal Mobile Telecommunication System
VBLAST	Vertical Bell Laboratories Layered Space Time
WiMAX	Worldwide Interoperability for Microwave Access
WAN	Wide Area Network
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian
ZF	Zero Forcing

Contents

Chapter 1

Introduction

Wireless communication is, by any measure, the most vibrant area and fastest growing segment of the communication field today. The constantly evolving and developing wireless technologies are changing the way people live, work, and entertain. Indeed, wireless communication has now become an integral part of people's daily life and a critical business tool with the proliferation of cellular phones and laptop computers. Moreover, the popularity of wireless communication is set to increase with the development of various new wireless systems and applications, and such a trend is inevitable due to the advantages inherited from the nature of wireless communication.

1.1 Benefits of Wireless Communication

Compared with the wireline communication counterpart, wireless communication offers a number of significant benefits. First, probably the most prominent and important feature of wireless communication is the provision of convenient and reliable tetherless connectivity. This offers greater flexibility and mobility. Unlike with a wired connection, people are no longer tied to their dedicated place, instead, they will be able to move freely and access network resource from any location within the wireless coverage area.

Another direct consequence of the tetherless connectivity is that wireless communication presents an promising approach to bring network access to the areas which would be difficult to connect to a wired network. For instance, possible applications include remote monitoring of natural environments such as glaciers, volcanoes and bodies of water, monitoring the condition of historic buildings where wiring is difficult, dangerous, or undesirable.

In addition, wireless networks are generally easier to deploy and setup compared with the wired networks because they remove the need of extensive cabling and patching, which also implies that wireless networks are more cost-effective. This is an extremely desirable benefit for those applications that only employ temporary networks, for example, trade shows, exhibitions and construction sites.

Finally, the maintenance and management of wireless networks are relatively simple and low cost. Wire-

less networks allow great expandability, i.e., one can easily add users to the current wireless network with existing equipment without requiring additional wiring, as well as efficiently removing existing users from the current wireless network.

Because of these attractive advantages, wireless communication has captured the attention of the industry and the imagination of the public. Various wireless networks and applications have been developed to explore these benefits. In the next section, we briefly review some current wireless systems and networks in operation.

1.2 Current Wireless Networks

Depending on the service range, mobility and data transmission rate, wireless networks generally fall into four different categories: Wireless Personal Area Network (PAN), wireless Local Area Network (LAN), wireless Metropolitan Area Network (MAN), and wireless Wide Area Network (WAN).

1.2.1 Wireless PAN

A wireless PAN is a type of wireless network that interconnects personal devices within a relatively short range (typically up to 10m or so), e.g., from a laptop to a nearby printer or from a cell phone to a wireless headset. It can support both low-rate and high-rate applications with different technologies.

Wireless PAN is standardized under the IEEE 802.15 series [32] . Currently, the market for wireless PAN has been dominated by Bluetooth (IEEE 802.15.1) products, which provide low-rate services with low-power consumption, i.e. wireless control of and communication between a mobile phone and a hands-free headset, wireless mouse, keyboard, and wireless game consoles. Another technology under development for low-rate wireless PAN is defined by the ZigBee specification (IEEE 802.15.4) which is intended to be simpler and less expensive than Bluetooth. For high-rate applications, such as digital imaging and multimedia services, technologies are under development based on the WiMedia specification (IEEE 802.15.3).

Overall, the technology for wireless PANs is in its infancy and is undergoing rapid development and research, and it is expected that this technology will find its application in various new environments to provide simple, easy to use connection to other devices and networks.

1.2.2 Wireless LAN

A wireless LAN is a type of network that provides high-speed data to wireless devices which are generally stationary or moving at pedestrian speeds within a small region, for instance, residential house, office building, university campus, or airport. With the proliferation of laptops, wireless LAN has become increasingly popular due to its ease of installation, as well as the location freedom provided.

Wireless LAN is standardized under the IEEE 802.11 series [31]. At the moment, there are primely three

different wireless LAN standards which have been implemented in the marketplace. IEEE 802.11b is the first standard with wide commercial acceptance and success. It operates in the 2.4 GHz band with a maximum speed of 11Mbps. The second standard is IEEE 802.11a which operates at 5 GHz band and provides a maximum speed of 70Mbps by adopting Orthogonal Frequency Division Multiplexing (OFDM) modulation. Another wireless LAN standard is IEEE 802.11g, which combines the advantages of 802.11b (relatively large coverage) and 802.11a (higher throughput) by defining the application of the OFDM transmission scheme in the 2.4 GHz band. It can provide access speed of up to 54Mbps.

To address the increasing high demand for high-speed high-quality wireless services, IEEE 802.11n, a new wireless LAN standard has been proposed in 2006, which will significantly improve the network throughput over previous standards, i.e., it can provide a maximum speed of 540Mbps. The proposal is expected to be approved in Jan 2010.

1.2.3 Wireless MAN

A wireless MAN is a type of network which mainly aims at providing broadband wireless access in larger geographical area than a LAN, ranging from several blocks of buildings to an entire city. Its main advantage is fast deployment and relatively low cost, and it has been considered as an attractive alternative solution to the wired last mile access systems such as Digital Subscriber Line (DSL) and cable modem access, especially for very crowded geographic areas like big cities and rural areas where wired infrastructure is difficult to deploy.

Wireless MAN is standardized under the IEEE 802.16 series [33], and is also known as Broadband Wireless Access standard. Based on the IEEE 802.16 standard, Worldwide Interoperability for Microwave Access (WiMAX) technology has been put forward by the industry alliance called the WiMAX Forum. The initial standard IEEE 802.16d only supports fixed applications which are often referred to as “fixed WiMAX”. Later, another amendment IEEE 802.16e introduced support for mobility, which is known as “Mobile WiMAX”. WiMAX supports very robust data throughput. The technology could provide approximately 40Mbps per channel. However, services across this channel would be shared by multiple customers which means that the typical rate available to users will be around 3Mbps.

A new standard (IEEE 802.16m) intending to provide data rate of 100Mbps for mobile applications and 1 Gbps for fixed applications is currently under development. The proposed work plan is expected to complete by December 2009 and ready for approval by March 2010.

1.2.4 Wireless WAN

A wireless WAN is a form of network which uses mobile telecommunication cellular network technologies such as Universal Mobile Telecommunication System (UMTS), General Packet Radio Service (GPRS) or Global System for Mobile Communication (GSM) to offer regionally, nationwide, or even globally voice and date services.

Wireless WAN has gone through rapid development in the last three decades. In 1980s, the first generation (1G) mobile communication systems were deployed, while the second generation (2G) mobile systems started to operate since 1990s. Both the 1G and 2G systems focus primarily on voice communications, while the 2G system has enhanced voice quality and has better spectrum management over the 1G system. The 2G systems provide data rate in the range of 9.6 – 14.4 Kbps. Currently, the third generation (3G) systems have started to roll out at full pace, and it is expected that 3G systems will provide higher transmission rate: a minimum speed of 2Mpbs and maximum of 14.4Mbps for stationary users, and 348Kbps in a moving vehicle.

While the improvement on the quality of service by 3G systems is obvious and impressive, more emerging applications are calling for higher date rate wireless service. At the moment, the industry and standardization body have already started to work on the fourth generation (4G) systems, which is intended to be a complete replacement for the current networks and be able to provide voice, data, and streamed multimedia to users on an "anytime, anywhere" basis. It is expected that the 4G systems will be able to deliver data rate of 1Gbps for stationary applications and 100Mbps for mobile applications.

1.3 Motivation

In the light of the above description of the current wireless networks, one can conclude that despite significant improvement on the provision of wireless services, there is an underlying strong demand for higher date rate wireless services, mainly driven by wireless data applications, as well as users' expectation of wire-equivalent quality wireless service.

Providing such high-rate high-quality wireless services is extremely challenging due to the inherent harsh wireless propagation environment. Compared to wired communication, wireless communication faces two fundamental problems that make fast and reliable wireless connection difficult to achieve, namely, interference and fading (variation of the channel strength over time and frequency due to the small-scale effect of multipath fading, as well as larger-scale fading effects such as path loss via distance attenuation and shadowing by obstacles such as tall buildings and mountains). In addition, wireless communication is required to carefully address the resource management problem, i.e. how to efficiently allocate and utilize power and spectrum (two principle resources in wireless communication).

Responding to these challenges, multiple-input multiple-output (MIMO) antenna systems were proposed independently by Telatar [94] and Foschini and Gans [23]. By introducing multiple antennas at both sides of the communication link, MIMO systems are able to substantially increase date rate and improve reliability without extra spectrum and power resources. The remarkable prospect of MIMO systems has not only sparked huge research interests in the research community, but also attracted enormous attentions from the industry and has led to practical implementation in real communication systems. For instance, MIMO technology has already been incorporated into various industry standards, i.e., wireless LAN IEEE 802.11n standard, wireless MAN IEEE 802.16e, Third Generation Partnership Project Long Term

Evolution (3GPP-LTE) Release 8 [1]. In general, MIMO technology is likely to become a prominent feature of future wireless communication systems.

The huge potential of MIMO technology has sparked a surge of research activities, which greatly strengthen our understanding of the fundamental limits and performance of MIMO channels. However, most of these research works are based on a relatively simple channel model, for instance, the channel is assumed to be a single random matrix and is subjected to Rayleigh fading or Rician fading. On the other hand, the increasing popularity of MIMO technology calls for a better understanding of the performance of MIMO systems operating in more practical environments. Motivated by this, this thesis looks into several general and practical channel models, such as Nakagami-m MIMO fading channels, double-scattering MIMO channels, multi-keyhole MIMO channels, and AF dual-hop MIMO channels, and investigates the fundamental capacity limits of these channels, as well as the performance of certain popular signal processing schemes. The objective of the thesis is to enhance our understanding of MIMO systems operating in these general MIMO channels, and to derive a set of new analytical results for understanding the performance of these advanced MIMO systems.

1.4 Dissertation Contributions and Outline

The rest of the thesis is organized as follows. Chapter 2 provides some background on wireless communication systems. Chapter 3 introduces two key mathematical theories, i.e. majorization theory and random matrix theory (RMT), on which many results of this thesis are based. The following chapters present the major contributions of the thesis. From chapter 4 to chapter 6, we focus on single user point-to-point communication systems, while in chapter 7 and chapter 8, the impact of co-channel interference will be investigated.

Specifically, chapter 4 considers the ergodic capacity of MIMO Nakagami-m fading channels. In contrast to the RMT approach adopted in previous research works on the ergodic capacity analysis, a unified way of deriving ergodic capacity bounds is developed under the majorization theory framework. The key idea is to study the ergodic capacity through the distribution of the diagonal elements of the quadratic channel $\mathbf{HH}^\dagger$ which is relatively easy to handle, avoiding the need of the eigenvalue distribution of the channel matrix which is extremely difficult to obtain. We first apply this method on the conventional point-to-point MIMO systems under Nakagami-m fading, and later extend the analysis to the more general distributed MIMO systems.

Chapter 5 examines the performance of multi-keyhole MIMO channels in details. This chapter studies the ergodic capacity of multi-keyhole MIMO channels and also the performance of practical transmission scheme Maximum Ratio Combining (MRC) is also investigated. The analysis is based on a set of new statistical properties of multi-keyhole MIMO channels, which include closed-form expressions for the distributions of an unordered eigenvalue and maximum eigenvalue, as well as solutions for the expected log-determinant and expected characteristic polynomial. Finally, the capacity and performance in multi-

keyhole channels are compared with those of rich-scattering MIMO Rayleigh channels.

Chapter 6 analyzes the ergodic capacity of AF dual-hop MIMO systems. Expression for the exact ergodic capacity, simplified closed-form expressions for the high SNR regime, and tight closed-form upper and lower bounds are presented. These results are obtained from the new closed-form expressions for various statistical properties of the equivalent AF MIMO dual-hop relay channel, such as the distribution of an unordered eigenvalue and certain random determinant properties which are derived by employing recent tools from finite-dimensional RMT literatures. In contrast to prior results which deal with asymptotic large antenna number systems, our expressions apply for arbitrary numbers of antennas and arbitrary relay configurations. The impact of the system and channel characteristics, such as the antenna configuration and the relay power gain, are investigated, and a number of interesting relationships between the dual-hop AF MIMO relay channel and conventional point-to-point MIMO channels in various asymptotic regimes are revealed.

Chapter 7 investigates the impact of co-channel interference under Rayleigh-product fading. Specifically, we study the performance of the OC transmission scheme in an interference-limited scenario. The analysis is based on novel expressions of the c.d.f. and p.d.f. of the maximum eigenvalue of the resultant channel matrix. An important special case, i.e., keyhole channel, is investigated in detail, where the ergodic capacity, outage performance and symbol error rate (SER) are analyzed based on various closed-form expressions for exact and asymptotic measures derived.

Chapter 8 studies the ergodic capacity of general MIMO systems with a single interferer in the low SNR regime. In contrast to prior results which deal with the interference limited scenario, our results are general and include both the interference and additive noise. Moreover, in addition to the MIMO Rician channels, MIMO Rayleigh product channels are considered. Exact analytical expressions for the minimum energy per information bit and wideband slope are derived for both systems. Based on these, the impact of system parameters, such as transmit and receive antenna number, Rician factor, channel mean matrix and interference-to-noise-ratio, are examined.

Chapter 9 gives some concluding remarks and enumerates future lines of work.

Chapter 2

Wireless Background

2.1 Wireless Communication Systems

A typical wireless communication system consists of a transmitter and receiver, as well as a number of functional blocks which facilitate information transmission. Generally, before the signal is ready for transmission, it usually goes through the following steps: source coding (encoding the source message into binary bit stream and removing redundant information), encryption (providing security for the communication by preventing unauthorized users from understanding messages), channel coding (adding redundancy to improve the reliability of the communication system) and modulation (converting digital symbols to waveforms which are compatible with the transmission channel). Similarly, the signal received at the receiver end goes through a reverse processing order to recover the original message, i.e., demodulation, channel decoding, decryption, and source decoding. Figure 2.1 gives a simple illustration of a typical wireless communication system.

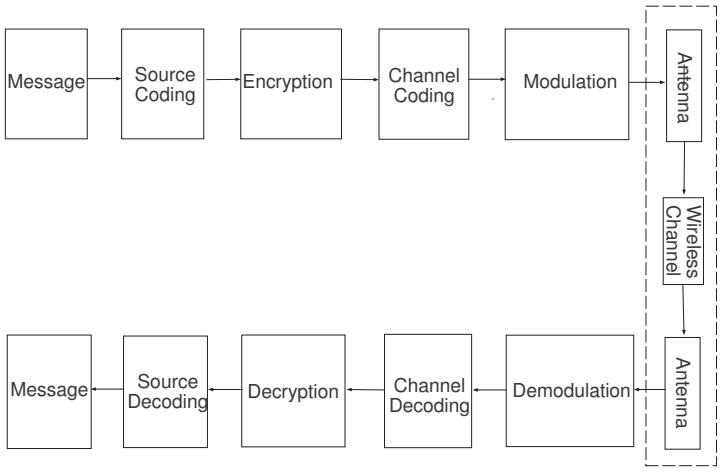

Figure 2.1. Schematic diagram of wireless communication systems

In this thesis, we mainly focus on understanding the impact of fading channels on the performance of communication systems (the dash line block). By doing so, the input and output relationship of a

communication system can be mathematically described by

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

where x is the transmitted symbol, h is the fading channel coefficient, n is the additive white Gaussian noise, and y is the received signal. In the following, we introduce the characteristic of the channel, and how it affects the performance of the system.

2.1.1 Wireless Fading Channels

A defining characteristic of wireless communication channels is the variation of the channel strength over time and over frequency, which is usually termed as "fading". The exact and precise mathematical description of this fading phenomena is either unknown or too complex for tractable analysis. Instead, a large amount of effort has been devoted to characterize the fading channel in a statistical approach. As a result, there exists a wide range of applicable statistical models corresponding to various physical propagation environments, which are relatively accurate and simple to analyze.

The fading effect is usually divided into two types, namely large-scale fading, mainly due to path loss as a function of distance and shadowing by large objects such as mountains and tall buildings, and small-scale fading, due to the constructive and destructive combination of randomly scattered, reflected, diffracted, and delayed multiple path signals. In the following, we give a mathematical description for several typical and important channel models which will be analyzed in this thesis.

1. Log-Normal Shadowing

   Empirical measurements reveal a general consensus that shadowing can be modeled by a log-normal distribution for various outdoor and indoor environments. The standard log-normal distribution can be expressed as

   $$p(r) = \frac{10}{\ln 10\sqrt{2\pi}\sigma r}\exp\left(-\frac{(10\log_{10} r - \mu)^2}{2\sigma^2}\right), \tag{2.2}$$

   where μ (dB) and σ (dB) are the mean and the standard deviation of $10\log_{10} r$, respectively.

2. Rayleigh Fading

   For small-scale fading, Rayleigh fading is probably one of the most frequently used models. It provides a good fit for multipath fading channels with no direct line-of-sight (LOS) path. The channel fading amplitude α is distributed according to

   $$p(\alpha) = \frac{2\alpha}{\Omega}\exp\left(-\frac{\alpha^2}{\Omega}\right), \alpha \geq 0, \tag{2.3}$$

   where $\Omega = \mathrm{E}\{\alpha^2\}$ is the mean value.

3. Rician Fading

In contrast to Rayleigh fading, Rician fading is often used to model propagation paths consisting of one strong direct LOS component and many random weaker components. The channel fading amplitude distribution can be expressed as

$$p(r) = \frac{2(1+n^2)e^{-n^2}r}{\Omega}\exp\left(-\frac{(1+n^2)r^2}{\Omega}\right) I_0\left(2nr\sqrt{\frac{1+n^2}{\Omega}}\right), r \geq 0, \tag{2.4}$$

where n is the fading parameter, which ranges from 0 to ∞, and is related to the Rician K factor by $K = n^2$ which corresponding to the ratio of the power of the LOS component to the average power of the scattered component. $I_0(\cdot)$ is the Bessel function of the first kind [26].

4. Nakagami-m Fading

   Nakagami-m fading is a more general fading distribution, which encompasses Rayleigh distribution as a special case, and can approximate well the Rician distribution. The channel fading amplitude distribution is given by

$$p(r) = \frac{2m^m r^{2m-1}}{\Omega^m \Gamma(m)}\exp\left(-\frac{mr^2}{\Omega}\right), r \geq 0 \tag{2.5}$$

   where m is the fading parameter, which ranges from $1/2$ to ∞. When $m = 1$, Nakagami-m distribution reduces to Rayleigh distribution. Moreover, the Rician distribution can be approximated by Nakagami-m distribution via a one-to-one mapping between the m parameter and the K parameter as follows [87]:

$$m = \frac{(1+K)^2}{1+2K}, K \geq 0, \tag{2.6}$$

   or

$$K = \frac{\sqrt{m^2-m}}{m-\sqrt{m^2-m}}, m \geq 1. \tag{2.7}$$

2.1.2 Performance Measures

An important aspect of communication research is to predict or evaluate the performance of various wireless communication systems. The elegant analytical tools developed by researchers not only offer system engineers a simple, yet accurate means for the performance evaluation, but also shed insight on the manner in which this performance depends on the key system parameters, thereby, providing guidance to the system engineers in the design of their systems.

There are several measures of performance related to practical wireless communication system design, i.e., channel capacity, outage probability, signal-to-noise ratio (SNR), signal-to-interference-and-noise ratio (SINR), symbol error rate (SER). This section gives brief introduction of these key measures that will be investigated through out the thesis.

2.1.2.1 Channel Capacity

Channel capacity is a term invented by Claude Shannon. In his landmark paper [86], he defined the channel capacity as the maximum rate of communication for which arbitrarily small error probability can be achieved. Mathematically, channel capacity is defined as the maximum of the mutual information between the transmitter and the receiver. For the channel model described by (2.1), the instantaneous channel capacity is given by [17]

$$C = \log_2 \left(1 + |h|^2 \frac{P}{\sigma^2}\right), \tag{2.8}$$

where P is the power of the transmit symbol, i.e., $\mathtt{E}\{xx^*\} = P$, and σ^2 is the noise level.

Depending on the underlying assumptions on the property of the fading channel h, several different notions of capacity emerged, i.e., ergodic capacity and outage capacity [3].

For ergodic capacity, the basic assumption here is that the transmission time is so long as to reveal the long-term ergodic properties of the fading process which is assumed to be an ergodic process in time. Mathematically, the ergodic capacity can be expressed as

$$C_e = \mathtt{E}_{|h|}\{C\}. \tag{2.9}$$

The ergodicity assumption is not necessarily satisfied in practical communication systems operating on fading channels. For the case where no significant channel variability occurs during the whole transmission, there may be a non-negligible probability that the value of the actual transmitted rate, no matter how small, exceeds the instantaneous channel capacity. In such case, $q\%$ outage capacity C_{out} should be considered, which is defined as the channel capacity C which is guaranteed to be supported by $(100 - q)\%$ of the channel realizations, required to provide a reliable service, i.e.,

$$\Pr\{C \leq C_{\mathsf{out}}\} \leq q\%. \tag{2.10}$$

2.1.2.2 SNR

The SNR, denoted as γ, is usually measured at the output of the communication systems, and is directly related to the data detection process. It is generally easy to evaluate, and more importantly, it often serves as an excellent indicator of the overall fidelity of the system. The output SNR is defined by

$$\gamma = \frac{\text{Power of signal component in the output}}{\text{Power in the noise component in the output}} = \frac{P|h|^2}{\sigma^2}. \tag{2.11}$$

In the context of fading channels, the average SNR $\bar{\gamma}$ is often taken as the performance measure, which is defined by

$$\bar{\gamma} = \mathtt{E}_{|h|}\{\gamma\}. \tag{2.12}$$

2.1.2.3 SINR

Wireless communication systems are generally subjected to co-channel interferences, for instance, in cellular systems, the received signals are often impaired by interference signals due to frequency reuse in the neighboring cells. Assuming only one strong interferer, mathematically, the input and output signal model for the desired user can be expressed as

$$y = hx + gs + n, \tag{2.13}$$

where g is the fading coefficient of the interferer-destination channel, and s is the interference signal satsfying $\mathtt{E}\{ss^*\} = P_s$.

When co-channel interference is taken into consideration, SINR, denoted as β, becomes a natural performance measure, which is defined by

$$\gamma = \frac{\text{Power of the desired-user's signal power in the output}}{\text{Sum of the power in the interference and noise components in the output}} = \frac{|h|^2 P}{|g|^2 P_s + \sigma^2}. \tag{2.14}$$

In the context of fading channels, the average SINR $\bar{\beta}$ is often taken as the performance measure, which is defined by

$$\bar{\gamma} = \mathtt{E}_{|h|,|g|}\{\gamma\}. \tag{2.15}$$

2.1.2.4 Outage Probability

Outage probability is another standard performance criterion denoted by $\mathsf{P_{out}}$ and defined as the probability that the instantaneous channel capacity below a specified value, or equivalently, the probability that the output SNR (or SINR) falls below a pre-defined acceptable threshold. Mathematically speaking, the outage probability is the c.d.f. of SNR evaluated at the specified threshold, i.e.,

$$P_{\mathsf{out}} = \int_0^{\gamma_{th}} p_\gamma(\gamma) d\gamma, \tag{2.16}$$

where γ_{th} is the predefined threshold, and $p_\gamma(\gamma)$ is the p.d.f. of SNR γ.

2.1.2.5 SER

The average SER, denoted by P_{SER}, is the one that is most revealing about the nature of the system behavior and is generally the most difficult performance criterion to compute. It is defined as the probability that a transmitted data symbol is detected in error at the receiver. The SER is typically modulation/detection scheme dependent, and is directly related to the instantaneous SNR (or SINR for multiuser systems). For many modulation schemes of interest, i.e, binary phase shift-keying (BPSK), binary-frequency shift-keying (BFSK) and M-ary phase amplitude modulation (PAM), the average SER

can be evaluated as [75]

$$P_{\mathsf{SER}} = \mathtt{E}_{\gamma}\left\{\alpha Q\left(\sqrt{2\beta\gamma}\right)\right\}, \tag{2.17}$$

where α, and β are modulation-specific constants, and $Q(\cdot)$ is the standard Gaussian Q-function.

2.2 MIMO Systems

In the previous section, we have introduced the conventional single-input single-output (SISO) communication system, several statistical channel fading models and various important performance measures. Now we turn our attention to the theme of this thesis, namely, MIMO systems. In this section, we brief discuss the MIMO fading channel model, benefits of MIMO systems, as well as some popular transmission schemes proposed to realize the benefits provided by MIMO systems.

2.2.1 MIMO Channels

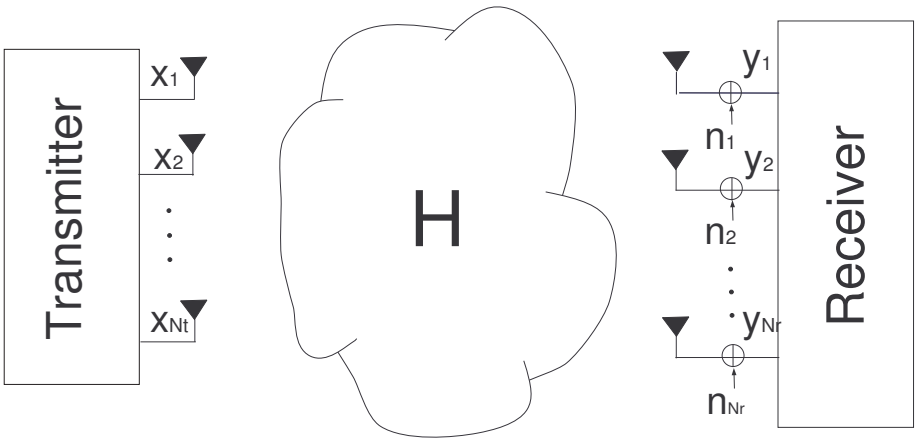

Figure 2.2. Diagram of a MIMO system with N_t transmit antennas and N_r receive antennas.

Figure 2.2 illustrates a MIMO system with N_t transmit antennas and N_r receive antennas. Mathematically, the complex baseband model is characterized by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.18}$$

where

$\mathbf{y} \in \mathcal{C}^{N_r \times 1}$ is the received signal at the receiver.

$\mathbf{x} \in \mathcal{C}^{N_t \times 1}$ is the transmit signal with sum power constraint $\mathtt{E}\{\mathbf{x}^{\dagger}\mathbf{x}\} = P$.

$\mathbf{n} \in \mathcal{C}^{N_r \times 1}$ is the noise vector with $\mathtt{E}\{\mathbf{n}\mathbf{n}^{\dagger}\} = \sigma^2\mathbf{I}$.

$\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$ is the channel matrix with (i, j)th element corresponding to the multiplicative fading parameter for the channel between the jth transmit antenna and ith receive antenna.

The characteristic of the MIMO channel is determined by the distribution of the elements of channel matrix $\mathbf{H}$, which in turn varies according to the underlying physical propagation environment. For instance, the elements of $\mathbf{H}$ can follow Rayleigh distribution, Rician distribution or Nakagami-m distribution as in the SISO channels. In this thesis, we will mainly focus on MIMO Nakagami-m, MIMO Rician and MIMO double-scattering fading channels, the mathematical description of these fading channels will be given in the corresponding chapters.

2.2.2 Benefits of MIMO Systems

The introduction of multiple antennas into communication systems has offered extra degree of freedom which can be exploited to provide various gains over conventional SISO systems, i.e., array gain (or power gain), diversity gain and multiplexing gain. In the following, we give a brief account of these gains.

2.2.2.1 Array Gain

Array gain is defined as the improvement of average SNR at the receiver by coherently combining the signals from multiple transmitters or receivers. For instance, in a single-input multiple-output (SIMO) system, the signal at the receiver is expressed as

$$\mathbf{y} = \mathbf{h}x + \mathbf{n}, \tag{2.19}$$

where $\mathbf{h}$ is the channel, and $\mathbf{n}$ is the noise $\mathtt{E}\{\mathbf{n}\mathbf{n}^\dagger\} = \sigma^2\mathbf{I}$. It is easy to show that the OC vector is given by $\frac{\mathbf{h}^\dagger}{||\mathbf{h}||}$, resulting the received SNR as $\frac{||\mathbf{h}||^2}{\sigma^2}$, which clearly indicates the advantage when compared with the SISO SNR $\frac{||h||^2}{\sigma^2}$. It is important to note that realization of array gain requires channel state information (CSI) at the transmitter or receiver.

2.2.2.2 Diversity Gain

Fading is the most prominent feature of a wireless communication channel, and diversity is an efficient means of combating channel fading. The general principle behind diversity is that the overall link reliability can be improved by observing multiple independent copies of the transmitted signal at the receiver. The diversity gain is usually measured in terms of how fast the bit error rate of a communication system decays with the increase of the SNR, i.e., the error exponent.

Diversity gain in SISO systems can be obtained in time or frequency domain, i.e. by repeating the same message several times, which however incurs a penalty in terms of date rate. In multiple antenna systems, another form of diversity is available, namely, spatial diversity, which includes receive diversity and transmit diversity. In contrast to the temporal and frequency diversity, the realization of spatial diversity does not incur any penalty in data rate, instead, it provides an array gain introduced earlier.

Receive diversity can be obtained in a system with multiple receive antennas by smartly combining the

multiple independent copies observed at the individual antenna. However, transmit diversity is generally much difficult to exploit since it requires sophisticated coding schemes. The most popular approach to realize transmit diversity is the so called Space-Time Coding (STC), which performs coding across space (transmit antennas) and time to extract diversity. We will give a simple introduction of the principle of STC later in this chapter.

2.2.2.3 Multiplexing Gain

Multiplexing gain is the most outstanding advantage of MIMO systems, and is defined as the linear increase in data rate without additional power or spectrum expenditure. Unlike array gain or diversity gain, which can be realized by either SIMO or multiple-input single-output (MISO) systems, multiplexing gain requires multiple antennas at both the transmitter and receiver ends.

The basic principle is to split a high-rate input data sequence into multiple lower-rate sequences, which are then modulated and independently sent in parallel via each of the transmit antennas, while the receiver employs appropriate signal processing technique to undo the mixing of the MIMO channel to detect the signals corresponding to each of the transmitted data streams.

Multiplexing gain and diversity gain can be achieved simultaneously by appropriate coding, in fact, there exists a fundamental tradeoff between the multiplexing gain and diversity gain for a given system as first discovered in [111], since which, the design of efficient and practical coding schemes achieving the optimal diversity-multiplexing tradeoff curve has been an extremely active area of research.

2.2.3 Transmission Schemes

As discussed in the previous section, MIMO antenna systems can be exploited to increase the spectral efficiency (multiplexing gain) or improve the link reliability (diversity gain). In the following, we introduce several popular transmission schemes proposed in the literature to realize these benefits. Before going into details, it is worth pointing out the critical value of CSI at the transmitter (CSIT)[1] in the design of practical transmission schemes. Generally, the availability of CSI limits the choice of transmission schemes, moreover, the more CSI, the better the performance of the system.

When there is no CSIT, popular design approaches include STC and layered architectures. The STC is a diversity oriented approach, which aims at improving the signal quality, reducing the SER, and providing better coverage. The idea of STC is to introduce intelligently controlled redundancy in the transmitted signal, both over space and time, which allows the receiver to recover the signal even in difficult propagation situations. There are mainly two types of STC techniques: space-time block coding (STBC) [2, 92] and space-time trellis coding (STTC) [91], both of which can achieve the full spatial diversity offered by the MIMO channel. STBC is relatively simpler than STTC. We now introduce a simple STBC scheme proposed by Alamouti in 1998 [2] for a system with two transmit antennas. For

[1] CSI at the receiver (CSIR) is relatively easy to obtain, i.e., via pilot training, hence, we assume that CSIR is always available.

the Alamouti code, the code matrix is given by

$$\mathbf{G} = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}. \tag{2.20}$$

At the first time slot, the first and second antenna transmit signals x_1 and x_2, respectively. At the second time slot, $-x_2^*$ and x_1^* are transmitted from the first and second antenna, respectively. Due to the orthogonal nature of the code matrix, the optimal diversity order can be obtained by only performing the MRC on the received the signal at the receiver.

Another scheme which does not require CSIT is known as layered architectures, also termed as layered STCs. In contrast to STC which is diversity based, layered architectures are capacity based which aim at realizing the linear capacity increase provided by MIMO systems. An important example is the famous Vertical Bell-Labs Layered Space Time (VBLAST) [22]. For this system, the transmitter splits a high-rate input data sequence into multiple lower-rate sequences, which are then modulated and independently sent in parallel via each of the transmit antennas, while the receiver employs appropriate signal processing technique to undo the mixing of the MIMO channel to detect the signals corresponding to each of the transmitted data streams.

When there is perfect CSIT, even superior performance can be achieved by beamforming strategy known as Maximum Ratio Transmission (MRT) [55]. The idea of MIMO-MRT is to steer a single transmitted symbol stream along the best eigenmode of the channel, by which, not only the system is robust against fading (achieve full diversity order), it also provides a boosted received SNR known as array gain. For instance, for a MIMO system described by (2.18), the transmitted signal vector is $\mathbf{x} = \mathbf{w}_{\mathsf{opt}} s$, with s representing the information symbol, and $\mathbf{w}_{\mathsf{opt}}$ denoting the optimal transmit weight vector. At the receiver, the signals on each receive antenna are linearly combined according to the MRC principle using the optimal receive weight vector $(\mathbf{H}\mathbf{w}_{\mathsf{opt}})^\dagger$ to give

$$z = \mathbf{w}_{\mathsf{opt}}{}^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{w}_{\mathsf{opt}} s + \mathbf{w}_{\mathsf{opt}}{}^\dagger \mathbf{H}^\dagger \mathbf{n}. \tag{2.21}$$

It is easy to show the optimal transmit weight vector is the dominant eigenvector of $\mathbf{H}^\dagger\mathbf{H}$ (i.e., the eigenvector corresponding to the maximum eigenvalue). The resultant SNR can be expressed as

$$\gamma = \frac{P}{\sigma^2} \lambda_{\mathsf{max}}, \tag{2.22}$$

where λ_{max} is the maximum eigenvalue of $\mathbf{H}\mathbf{H}^\dagger$.

Moreover, in the context of co-channel interference, which is inevitable in a cellular communication system, it is proved that OC [100] can effectively suppress the interference and achieve good performance. The idea of OC is similar to that of MIMO-MRC with the difference that the transmission direction is chosen to be the best eigenmode of the effective channel which takes into account of the interference.

When there is partial CSIT, such as channel mean matrix or correlation matrix, hybrid schemes can be employed [40, 96]. The basic idea is to enhance the performance of a STC, for which no CSIT is required, by combining it with some type of beamforming, for which the existing CSIT can be exploited.

2.2.4 Capacity of MIMO Channels

The enormous interests in MIMO systems are mainly inspired by the significant information-theoretical results reported in pioneering works by [23] and [94], independently, where the authors have proved that the capacity of MIMO system scales linearly with the minimum number of the transmit or receive antennas. In this section, we give a brief review of the results obtained in [23, 94].

For a system described by (2.18), the mutual information expression was derived in [23, 94] as

$$I = \log_2 \det \left(\mathbf{I} + \frac{1}{\sigma^2} \mathbf{H} \mathbf{R}_s \mathbf{H}^\dagger \right), \tag{2.23}$$

and therefore the channel capacity is given by

$$C = \max_{\mathrm{tr}\{\mathbf{R}_s\} \leq P} I, \tag{2.24}$$

where the optimization is taken on the signal covariance matrix $\mathbf{R}_s = \mathtt{E}\{\mathbf{x}\mathbf{x}^\dagger\}$ with P being the total transmit power. Therefore, the ergodic capacity can be expressed as

$$C_e = \mathtt{E}_{\mathbf{H}} \left\{ \max_{\mathrm{tr}\{\mathbf{R}_s\} \leq P} I \right\}. \tag{2.25}$$

The ergodic capacity depends heavily on the availability of CSIT. When perfect CSIT is available, the transmitter can adapt its power according to the so called waterfilling principle [17] to maximize the mutual information. In this case, we perform the singular value decomposition (SVD) on the channel matrix $\mathbf{H}$, which results in

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger \tag{2.26}$$

where $\mathbf{U} \in \mathcal{C}^{N_r \times N_r}$ and $\mathbf{V} \in \mathcal{C}^{N_t \times N_t}$ are the unitary matrices and $\mathbf{D} \in \mathcal{C}^{N_r \times N_r}$ is the diagonal matrix containing the singular values t_i, $i = 1, \cdots, r$ of $\mathbf{H}$ where $r = \min(N_t, N_r)$. The maximum mutual information is achieved when $\mathbf{R}_s = \mathbf{U}\mathbf{\Lambda}_s\mathbf{U}^\dagger$, where $\mathbf{\Lambda}_s$ is a diagonal matrix with elements given as

$$[\mathbf{\Lambda}_s]_{ii} = \left(\mu - \frac{\sigma^2}{t_i^2} \right)^+, \tag{2.27}$$

where $(a)^+$ denotes $\max(0, a)$, and μ is a constant to be decided to meet the power constraints $\sum_{i=1}^{r} (\mathbf{\Lambda}_s)_{ii} = P$. Therefore, the ergodic capacity is given by

$$C_e = \mathtt{E} \left\{ \sum_{i=1}^{r} \left(\log_2 \left(\frac{\mu t_i^2}{\sigma^2} \right) \right)^+ \right\}. \tag{2.28}$$

For the case where there is no CSIT, adaptation at the transmitter is not possible, and equal power allocation is the most reasonable strategy, i.e. $\mathbf{R}_s = \frac{P}{N_t}\mathbf{I}$[2]. Hence, the ergodic capacity can be expressed as

$$C_e = \mathtt{E}_{\mathbf{W}} \left\{ \log_2 \det \left(\mathbf{I}_r + \frac{P}{N_t} \mathbf{W} \right) \right\}, \tag{2.29}$$

where $r \times r$ matrix $\mathbf{W}$ is defined as

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger & N_r < N_t, \\ \mathbf{H}^\dagger\mathbf{H} & N_t \leq N_r. \end{cases} \tag{2.30}$$

By eigenvalue decomposition, (2.29) can be alternatively expressed as

$$C_e = \sum_{i=1}^{r} \mathtt{E}_{\lambda_i} \left\{ \log_2 \det \left(1 + \frac{P}{N_t} \lambda_i \right) \right\} = r\mathtt{E}_{\lambda} \left\{ \log_2 \det \left(1 + \frac{P}{N_t} \lambda \right) \right\}, \tag{2.31}$$

where $\{\lambda_i\}$, $i = 1, \cdots, r$ are the r eigenvalues of matrix $\mathbf{W}$ and λ is an unordered eigenvalue of matrix $\mathbf{W}$. From Equation (2.31), it is easy to observe that the ergodic capacity of MIMO systems scales linearly with the minimum number of transmit and receive antennas.

[2]It is worth to point out that equal power allocation is not necessarily optimal in all case [35].

Chapter 3

Mathematical Preliminaries

The research works conducted in this thesis heavily rely on two important mathematical theories. First of all, the majorization theory, which provides a powerful tool for establishing inequalities. Majorization theory has already been applied in wireless communications systems, for instance, in the design of optimal linear precoding scheme for MIMO systems [73], and in the analysis of impact of channel correlation on the ergodic capacity [42]. To make this thesis self-contained, we give necessary definitions and essential results on majorization theory in the following section.

In the second part of this chapter, we introduce the finite Random Matrix Theory (RMT). First, we give the basic definitions regarding multi-variates complex Gaussian distribution, followed by brief description of Wishart matrix. Then, we present a host of novel statistical results for a particular group of random matrices possessing a matrix product structure, including p.d.f. and c.d.f. of an unordered eigenvalue, p.d.f. and c.d.f. of the largest eigenvalue, expected determinant as well as the expected log-determinant of the random matrix of interest. These new results are applied in the performance analysis of various MIMO systems in the following chapters.

3.1 Majorization Theory

This section provides basic and necessary definitions on majorization theory as well as some essential results, which will be applied in Chapter 4 for the ergodic capacity analysis of MIMO Nakagami-m fading channels.

Definition 3.1 *[62] For any* $\mathbf{x} \in \mathbb{R}^n$*, let* $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$ *denote the components of vector* $\mathbf{x}$ *in decreasing order, and let*

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)} \tag{3.1}$$

denote the components of vector $\mathbf{x}$ *in increasing order.*

Definition 3.2 *[62, 1.A.1] For any vector $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}$ is majorized by $\mathbf{y}$ (or $\mathbf{y}$ majorizes $\mathbf{x}$) if*

$$\begin{cases} \sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, & 1 \leq k \leq n-1 \\ \sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}. \end{cases} \tag{3.2}$$

The notation $\mathbf{x} \prec \mathbf{y}$, or equivalently, by $\mathbf{y} \succ \mathbf{x}$, is used to denote the case where $\mathbf{y}$ majorizes $\mathbf{x}$. Alternatively, the previous conditions can be rewritten as

$$\begin{cases} \sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)}, & 1 \leq k \leq n-1 \\ \sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}. \end{cases} \tag{3.3}$$

Definition 3.3 *[62, 3.A.1] A real-valued function $\phi(\cdot)$ defined on a set $\mathcal{A} \subseteq R^n$ is said to be Schur-convex on $\mathcal{A}$ if*

$$\mathbf{x} \prec \mathbf{y} \quad on \quad \mathcal{A} \quad \Rightarrow \quad \phi(\mathbf{x}) \leq \phi(\mathbf{y}). \tag{3.4}$$

Similarly, $\phi(\cdot)$ is said to be Schur-concave on $\mathcal{A}$ if

$$\mathbf{x} \succ \mathbf{y} \quad on \quad \mathcal{A} \quad \Rightarrow \quad \phi(\mathbf{x}) \leq \phi(\mathbf{y}). \tag{3.5}$$

As a consequence, if $\phi(\cdot)$ is Schur-convex on $\mathcal{A}$ then $-\phi(\cdot)$ is Schur-concave on $\mathcal{A}$ and vice-versa.

Example 3.1 *[62, p.7] For any $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{1} \in \mathbb{R}^n$ denote the constant vector with the i-th element given by $\mathbf{1}_i \triangleq \frac{1}{n} \sum_{j=1}^{n} x_j$, then*

$$\mathbf{1} \prec \mathbf{x}. \tag{3.6}$$

This means that the vector of equal entries is majorized by any vector with the same sum-value.

Example 3.2 *For any $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{y} \in \mathbb{R}^n$ denote the vector with the first element being the only non-zero element $\sum_{i=1}^{n} x_i$, namely $\mathbf{y} = [\sum_{i=1}^{n} x_i, 0, \ldots, 0]$, then*

$$\mathbf{y} \succ \mathbf{x}. \tag{3.7}$$

In this example, it further states that the vector with only one non-zero element majorizes any vector with the same sum-value.

Lemma 3.1 *[62, 3.C.1] If $g : \mathbb{R} \to \mathbb{R}$ is convex, then the symmetric convex function*

$$\phi(\mathbf{x}) = \sum_{i=1}^{n} g(x_i) \tag{3.8}$$

is Schur-convex. Similarly, if g is concave, then $\phi(\mathbf{x}) = \sum_{i=1}^{n} g(x_i)$ is Schur-concave.

Lemma 3.2 *[62, 9.B.1] Let $\mathbf{Q}$ be an $n \times n$ Hermitian matrix with diagonal elements denoted by the vector $\mathbf{d}$ and the eigenvalues denoted by the vector $\boldsymbol{\lambda}$, then*

$$\boldsymbol{\lambda} \succ \mathbf{d}. \tag{3.9}$$

Lemma 3.3 *For fixed s, t, and $t \geq s$, let $\mathbf{x} = [x_1, \ldots, x_{st}]$ denote an st-dimensional vector with joint density $p(\mathbf{x})$, where $\{x_i\}$ are i.i.d. gamma random variables. Define the vector $\mathbf{y}^{(1)} = [y_1^{(1)}, \ldots, y_t^{(1)}] \in \mathbb{R}^t$ where $y_i^{(1)}$ is the summation of any s elements of $\mathbf{x}$ and for any $y_i^{(1)}, y_j^{(1)}$ such that $i \neq j$, they do not involve any common elements of $\mathbf{x}$. Similarly, $\mathbf{y}^{(2)} = [y_1^{(2)}, \ldots, y_s^{(2)}, 0, \ldots, 0] \in \mathbb{R}^t$ can be defined, such that $y_i^{(2)}$ is the summation of any t elements of $\mathbf{x}$ and that for any $y_i^{(2)}, y_j^{(2)}$ with $i \neq j$, they do not involve any common elements of $\mathbf{x}$. Then, we have*

$$\mathtt{E}\left[\tau(\mathbf{y}^{(1)})\right] \geq \mathtt{E}\left[\tau(\mathbf{y}^{(2)})\right], \tag{3.10}$$

where $\tau(u_1, \ldots, u_t) \triangleq \sum_{i=1}^{t} \log_2(1 + a u_i)$ and $a > 0$.

Proof: The above lemma is a special case of a more general result due to Boland *et al.* [9] . □

3.2 RMT

In this section, we give basic definitions and some preliminary results on complex multi-variate Gaussian random distribution. In addition, we derive a set of new RMT results, which will be applied in the capacity and performance analysis of various MIMO channels later on.

3.2.1 Definitions and Preliminary Results

Definition 3.4 *The complex multivariate gamma function $\tilde{\Gamma}_n(m)$ is defined as*

$$\tilde{\Gamma}_n(m) \triangleq \int_{\mathbf{A}=\mathbf{A}^\dagger>0} \mathtt{e}^{\mathtt{tr}(-\mathbf{A})} \det(\mathbf{A})^{m-n} (d\mathbf{A}) = \pi^{\frac{n(n-1)}{2}} \prod_{i=1}^{n} \Gamma(m-j+1). \tag{3.11}$$

Definition 3.5 *Let $\mathbf{X}$ be an $n \times n$ matrix with non-zero eigenvalues $x_1, \cdots, x_L$. The Vandermonde determinant is defined as*

$$V_L(\mathbf{X}) \triangleq \det\left(\{x_i^{L-j}\}_{i,j=1,\cdots,L}\right) = \prod_{i<j}^{L} (x_i - x_j). \tag{3.12}$$

Definition 3.6 *The n-variate complex Gaussian distribution with mean vector $\mathbf{v} \in \mathcal{C}^{n\times 1}$ and covariance matrix $\boldsymbol{\Omega} \in \mathcal{C}^{n\times n} > 0$ is denoted by $\mathcal{CN}_n(\mathbf{v}, \boldsymbol{\Omega})$.*

Definition 3.7 *[34] The random matrix $\mathbf{X} \in \mathcal{C}^{n\times m}$ is said to have a matrix-variate complex Gaussian distribution with mean matrix $\mathbf{M} \in \mathcal{C}^{n\times m}$ and covariance matrix $\mathbf{\Omega} \otimes \mathbf{\Sigma}$, where $\mathbf{\Omega} \in \mathcal{C}^{n\times n}$ and $\mathbf{\Sigma} \in \mathcal{C}^{m\times m}$ are positive definite matrix, if*

$$\mathsf{vec}(\mathbf{X}^\dagger) \sim \mathcal{CN}_{nm}\left(\mathsf{vec}(\mathbf{M}^\dagger), \mathbf{\Omega} \otimes \mathbf{\Sigma}\right). \tag{3.13}$$

And is denoted as $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{M}, \mathbf{\Omega} \otimes \mathbf{\Sigma})$.

Lemma 3.4 *[34] If the $n \times m$ matrix $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{M}, \mathbf{\Omega} \otimes \mathbf{\Sigma})$, where $\mathbf{\Omega} \in \mathcal{C}^{n\times n}$ and $\mathbf{\Sigma} \in \mathcal{C}^{m\times m}$ are positive definite matrix, then the p.d.f. of $\mathbf{X}$ is given by*

$$f(\mathbf{X}) = \frac{\mathsf{e}^{\mathsf{tr}\left(-\mathbf{\Omega}^{-1}(\mathbf{X}-\mathbf{M})\mathbf{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})^\dagger\right)}}{\pi^{nm}\det(\Omega)^m \det(\Sigma)^n}. \tag{3.14}$$

Lemma 3.5 *[34] Let $\mathbf{X} \sim \mathcal{CN}_{n,m}\left(\mathbf{0}_{n\times m}, \mathbf{\Omega} \otimes \mathbf{I}_m\right)$, with $n \leq m$. Then $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ has a complex central Wishart Distribution $W_n(m, \mathbf{\Omega})$ with p.d.f.*

$$f(\mathbf{W}) = \frac{\mathsf{e}^{\mathsf{tr}\left(-\mathbf{\Omega}^{-1}\mathbf{W}\right)}\det(\mathbf{W})^{m-n}}{\tilde{\Gamma}_n(m)\det(\mathbf{\Omega})^m}, \tag{3.15}$$

where $\tilde{\Gamma}_n(m)$ is the complex multivariate gamma function.

Lemma 3.6 *[21] Let $\mathbf{W} \sim W_n(m, \mathbf{I}_n)$. Then the joint p.d.f. of the ordered eigenvalues $\mathbf{\Lambda} = \mathsf{diag}(\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0)$ of $\mathbf{W}$ is given by*

$$f(\mathbf{\Lambda}) = \frac{\mathsf{e}^{\mathsf{tr}(-\mathbf{\Lambda})}\det(\mathbf{\Lambda})^{m-n}V_n(\mathbf{\Lambda})^2}{\prod_{i=1}^n \Gamma(n-i+1)\prod_{i=1}^n \Gamma(m-i+1)}. \tag{3.16}$$

3.2.2 New Random Eigenvalue Distribution Results

We now present some new results on the eigenvalue distribution of certain complex random matrices. These analytical expressions will be used to characterize various performance measures of certain MIMO channel in the following chapters.

Lemma 3.7 *Let $\mathbf{H} \sim \mathcal{CN}_{m,n}(\mathbf{0}_{m\times n}, \mathbf{I} \otimes \mathbf{I})$, and $\mathbf{\Omega}$ is an $m \times m$ positive definite matrix with eigenvalue $\omega_1 > \omega_2 > \cdots > \omega_m > 0$. Then, the marginal p.d.f. of an unordered eigenvalue λ of matrix $\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}$ is given by*

$$f(\lambda) = \frac{1}{s\prod_{i<j}^m(\omega_j - \omega_i)}\sum_{l=1}^m \sum_{k=m-s+1}^m \frac{\lambda^{n+k-m-1}e^{-\lambda/\omega_l}\omega_l^{m-n-1}}{\Gamma(n-m+k)}D_{l,k} \tag{3.17}$$

where $D_{l,k}$ is the (l,k)th cofactor of an $m \times m$ matrix $\mathbf{D}$ whose (i,j)th entry is

$$\{\mathbf{D}\}_{i,j} = \omega_i^{j-1}. \tag{3.18}$$

where $s = \min(m, n)$.

Proof: See Appendix A.1. □

This lemma presents a new expression for the unordered eigenvalue distribution of a complex semi-correlated central Wishart matrix. In prior work [5], two separate alternative expressions for this p.d.f. were obtained for the specific scenarios $n \leq m$ and $n > m$ respectively; the latter case[1] being a complicated expression in terms of determinants with entries depending on the inverse of a certain Vandermonde matrix. Here, Lemma 3.7 presents a simpler and more computationally-efficient unified expression, which applies for arbitrary m and n.

Lemma 3.8 *Let* $\mathbf{H} \sim \mathcal{CN}_{N_r,N_t}(\mathbf{0}_{N_r \times N_t}, \mathbf{I} \otimes \mathbf{I})$ *and* a *being a positive constant. Then the joint p.d.f. of the eigenvalues* $\{0 \leq \omega_1 < \cdots < \omega_q \leq 1/a\}$ *of random matrix* $\mathbf{H}^\dagger(\mathbf{I} + a\mathbf{H}\mathbf{H}^\dagger)^{-1}\mathbf{H}$ *is given by*

$$f(\omega_1, \ldots, \omega_q) = \mathcal{K} \prod_{i<j}^{q} (\omega_j - \omega_i)^2 \prod_{i=1}^{q} \frac{\omega_i^{p-q} e^{-\frac{\omega_i}{1-a\omega_i}}}{(1-a\omega_i)^{p+q}}, \tag{3.19}$$

where $q = \min(N_r, N_t)$, $p = \max(N_r, N_t)$ *and*

$$\mathcal{K} = \left(\prod_{i=1}^{q} \Gamma(q-i+1)\, \Gamma(p-i+1) \right)^{-1}. \tag{3.20}$$

The p.d.f. of an unordered eigenvalue $\omega \in \{\omega_1, \cdots, \omega_q\}$ *is given by*

$$f(\omega) = \frac{1}{q} \sum_{i=0}^{q-1} \sum_{j=0}^{i} \sum_{l=0}^{2j} \frac{\mathcal{A}(i,j,l,p,q)\, \omega^{p-q+l}}{(1-a\omega)^{p-q+l+2}} \exp\left(-\frac{\omega}{1-a\omega} \right), \tag{3.21}$$

where

$$\mathcal{A}(i,j,l,\kappa_1,\kappa_2) = \frac{(-1)^l \binom{2i-2j}{i-j} \binom{2j+2\kappa_1-2\kappa_2}{2j-l} (2j)!}{2^{2i-l}\, (\kappa_1 - \kappa_2 + j)!\, j!}. \tag{3.22}$$

Proof: See Appendix A.2. □

Now, armed with Lemma 3.7 and Lemma 3.8, we are ready to derive the following theorem, which will be used to evaluate the ergodic capacity of MIMO dual-hop systems.

Theorem 3.1 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_r,N_s}(\mathbf{0}_{N_r \times N_s}, \mathbf{I} \otimes \mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_d,N_r}(\mathbf{0}_{N_d \times N_r}, \mathbf{I} \otimes \mathbf{I})$, a *being a positive real number. Then the marginal p.d.f. of an unordered eigenvalue* λ *of* $\mathbf{H}_1^\dagger \mathbf{H}_2^\dagger (\mathbf{I} + a\mathbf{H}_2\mathbf{H}_2^\dagger)^{-1}\mathbf{H}_2\mathbf{H}_1$ *is*

[1] For this case ($n > m$), the random matrix $\mathbf{H}^\dagger \mathbf{\Omega} \mathbf{H}$ has reduced rank and the corresponding distribution, conditioned on $\mathbf{\Omega}$, is commonly referred to as *pseudo*-Wishart [58].

given by

$$f_{\lambda}\left(\lambda\right)=\frac{2e^{-\lambda a}\mathcal{K}}{s}\sum_{l=1}^{q}\sum_{k=q-s+1}^{q}\sum_{i=0}^{q+N_s-l}\frac{\binom{q+N_s-l}{i}a^{q+N_s-l-i}}{\Gamma\left(N_s-q+k\right)}\lambda^{(2N_s+2k+p-q-i-3)/2}K_{p+q-i-1}\left(2\sqrt{\lambda}\right)G_{l,k}, \tag{3.23}$$

where $q=\min(N_r,N_d)$*,* $p=\max(N_r,N_d)$*,* $s=\min(N_s,q)$*,* $K_v\left(\cdot\right)$ *is the modified Bessel function of the second kind and* $G_{l,k}$ *is the* (l,k)*th cofactor of a* $q\times q$ *matrix* $\mathbf{G}$ *whose* (m,n)*th entry is*

$$\left\{\mathbf{G}\right\}_{m,n}=a^{q-p-m-n+1}\Gamma\left(p-q+m+n-1\right)U\left(p-q+m+n-1,p+q,1/a\right) \tag{3.24}$$

with $U\left(\cdot,\cdot,\cdot\right)$ *denoting the confluent hypergeometric function of the second kind [26, (9.211.4)].*

Proof: See Appendix A.3. □

To this end, we present another theorem regarding the p.d.f. of an unordered eigenvalue of a matrix involving a product of two independent complex random matrices, which will be used for deriving the ergodic mutual information expression for MIMO multi-keyhole channels.

Theorem 3.2 *Let* $\mathbf{H}_1\sim\mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t\times N_k},\mathbf{I}\otimes\mathbf{I})$*,* $\mathbf{H}_2\in\mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r\times N_k},\mathbf{I}\otimes\mathbf{I})$*, and* $\mathbf{A}\in\mathcal{C}^{N_k\times N_k}$*. Then the marginal p.d.f. of an unordered eigenvalue of* $\mathbf{H}_1\mathbf{A}^{\dagger}\mathbf{H}_2^{\dagger}\mathbf{H}_2\mathbf{A}\mathbf{H}_1^{\dagger}$ *is given by*

$$f(\lambda)=\frac{1}{p\prod_{i<j}^{N_k}(b_j-b_i)}\sum_{i=1}^{N_k}\sum_{j=N_k-p+1}^{N_k}\frac{2b_i^{N_k-1-\frac{m+n}{2}}\lambda^{\frac{m+n}{2}-N_k+j-1}K_{n-m}\left(2\sqrt{\frac{\lambda}{b_i}}\right)}{\Gamma(n-N_k+j)\Gamma(m-N_k+j)}\mathcal{D}_{i,j}, \tag{3.25}$$

where $m=\max(N_t,N_r)$*,* $n=\min(N_t,N_r)$*,* $p=\min(n,N_k)$*,* $b_1\leq b_2\leq\cdots\leq b_{N_k}$ *denote the non-zero eigenvalues of* $\mathbf{B}\triangleq\mathbf{A}\mathbf{A}^{\dagger}$*, and* $\mathcal{D}_{i,j}$ *is the* (i,j)*th cofactor of the matrix* $\mathbf{\Xi}$ *whose* (l,k)*th entry equals*

$$\left[\mathbf{\Xi}\right]_{l,k}=b_l^{k-1},\ \textit{for}\ 1\leq l,k\leq N_k\ . \tag{3.26}$$

Proof: See Appendix A.4. □

The following theorem presents the c.d.f. of the maximum eigenvalue of a matrix involving a product of two independent complex random matrices, and it will be used for deriving the outage probability of transmit beamforming systems in MIMO multi-keyhole channels.

Theorem 3.3 *Let* $\mathbf{H}_1\sim\mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t\times N_k},\mathbf{I}\otimes\mathbf{I})$*,* $\mathbf{H}_2\sim\mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r\times N_k},\mathbf{I}\otimes\mathbf{I})$*, and* $\mathbf{A}\in\mathcal{C}^{N_k\times N_k}$*. Then the cumulative distribution function (c.d.f.) of the maximum eigenvalue of* $\mathbf{H}_1\mathbf{A}^{\dagger}\mathbf{H}_2^{\dagger}\mathbf{H}_2\mathbf{A}\mathbf{H}_1^{\dagger}$ *is given by*

$$F_{\lambda_{\max}}(x)=\frac{(-1)^{\frac{p(p-1)}{2}}\det(\mathbf{\Phi}(x))}{\prod_{i=1}^{p}\Gamma(n-i+1)\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{3.27}$$

where $m = \max(N_t, N_r)$, $n = \min(N_t, N_r)$, $p = \min(n, N_k)$, $b_1 \leq b_2 \leq \cdots \leq b_{N_k}$ *denote the non-zero eigenvalues of* $\mathbf{B} \triangleq \mathbf{A}\mathbf{A}^\dagger$, *and* $\mathbf{\Phi}(x)$ *is an* $N_k \times N_k$ *matrix whose* (l,k)*th entry is given by*

$$[\mathbf{\Phi}(x)]_{l,k} = \begin{cases} b_l^{k-1}, & k \leq N_k - p, \\ g(x)_{l,k}, & k > N_k - p, \end{cases} \tag{3.28}$$

with

$$g(x)_{l,k} = \Gamma(q-k+1) b_l^{2N_k-p-k} - b_l^{N_k-n-1} \sum_{t=0}^{m+q-n-k} \frac{x^t}{\Gamma(t+1)} 2(b_l x)^{\frac{q-t-k+1}{2}} K_{q-t-k+1}\left(2\sqrt{\frac{x}{b_l}}\right). \tag{3.29}$$

Proof: See Appendix A.5. □

We now present a theorem which gives first-order expansion for the c.d.f. given in Theorem 3.3 when $n = 1$, which will be used for deriving the diversity order, array gain, and asymptotic outage probability of transmit beamforming systems in MISO/SIMO multi-keyhole channels.

Theorem 3.4 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t \times N_k}, \mathbf{I} \otimes \mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r \times N_k}, \mathbf{I} \otimes \mathbf{I})$, *and* $\mathbf{A} \in \mathcal{C}^{N_k \times N_k}$, $m = \max(N_t, N_r)$, $n = \min(N_t, N_r)$, $p = \min(n, N_k)$, $d = \min(m, N_k)$. *When* $n = 1$, *the asymptotic expansion of the c.d.f. of the maximum eigenvalue* $\lambda_{\max}$ *of* $\mathbf{H}_1\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1^\dagger$ *is given by*

$$F_{\lambda_{\max}}(x) = \frac{a_1}{d} x^d + o(x^d) \tag{3.30}$$

where

$$a_1 = \begin{cases} \frac{\Gamma(m-N_k)}{\Gamma(m)\Gamma(N_k)\prod_{i=1}^{N_k} b_i}, & m > N_k, \\ \frac{1}{\Gamma(m)^2}\left(\frac{\psi(1)+\psi(m)-\ln x}{\prod_{i=1}^{N_k} b_i} + \frac{(-1)^{m-1}\det(\mathbf{\Phi}^3)}{\prod_{i<j}^{N_k}(b_j - b_i)}\right), & m = N_k, \\ \frac{(-1)^{m-1}}{\Gamma(m)^2}\frac{\det(\mathbf{\Phi}^4)}{\prod_{i<j}^{N_k}(b_j - b_i)}, & m < N_k, \end{cases} \tag{3.31}$$

where $\psi(\cdot)$ *is the digamma function [26], and* $\mathbf{\Phi}^3$ *and* $\mathbf{\Phi}^4$ *are* $N_k \times N_k$ *matrices with entries*

$$[\mathbf{\Phi}^3]_{l,k} = \begin{cases} b_l^{k-1}, & k = 1, \cdots, N_k - 1, \\ b_l^{-1} \ln b_l, & k = N_k, \end{cases} \tag{3.32}$$

and

$$[\mathbf{\Phi}^4]_{l,k} = \begin{cases} b_l^{k-1}, & k = 1, \cdots, N_k - 1, \\ b_l^{N_k-m-1} \ln b_l, & k = N_k, \end{cases} \tag{3.33}$$

respectively.

Proof: See Appendix A.6. □

The following theorem presents the exact c.d.f. expression of the maximum eigenvalue of a random matrix involving a product of three independent complex random matrices. This will be used to analyze

the outage probability of optimum combining system operating in interference-limited Rayleigh-product channels.

Theorem 3.5 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_s}(\mathbf{0}_{N_t\times N_s}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_s,N_t}(\mathbf{0}_{N_s\times N_t}, \mathbf{I}\otimes\mathbf{I})$ *and* $\mathbf{H}_3 \sim \mathcal{CN}_{N_r,N_I}(\mathbf{0}_{N_r\times N_I}, \mathbf{I}\otimes\mathbf{I})$, $N_I \geq N_r$. *Define* $m = \min(N_r, N_s)$, $n = \max(N_r, N_s)$, $p = \max(0, m - N_t)$, $q = \max(m, N_t)$. *Then the c.d.f. of the maximum eigenvalue of matrix* $\frac{1}{N_s}\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\left(\mathbf{H}_3\mathbf{H}_3^\dagger\right)^{-1}\mathbf{H}_1\mathbf{H}_2$ *is given by:*
1) When $N_t \leq N_r$ *or* $N_t \geq N_r \geq N_s$,

$$\mathcal{F}_{\lambda_{\max}}(x) = \frac{\prod_{i=1}^{m}(-1)^{pN_t}\Gamma(N_{\mathcal{I}}+N_s-i+1)\det(\mathbf{\Delta}(x))}{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}-N_r+m-i+1)\Gamma(m-i+1)\Gamma(n-i+1)}, \tag{3.34}$$

where $\mathbf{\Delta}(x)$ *is defined by*

$$[\mathbf{\Delta}(x)]_{i,j} = \begin{cases} (-1)^{m-N_t-i}B(n+i-j, N_{\mathcal{I}}-N_r+m-i+j), & i \leq p, \\ B(m+n+p-i-j+1, N_{\mathcal{I}}-q-N_r+i+j-1) - R(x), & i > p, \end{cases} \tag{3.35}$$

with

$$R(x) = \sum_{k=0}^{q-i}\frac{(xN_s)^k}{\Gamma(k+1)}\Gamma(N_{\mathcal{I}}-N_r-p+i+j+k-1)U(N_{\mathcal{I}}-N_r-p+i+j+k-1, i+j+k-p-n-m, xN_s), \tag{3.36}$$

2) $N_t \geq N_s \geq N_r$ *or* $N_s \geq N_t \geq N_r$,

$$\mathcal{F}_{\lambda_{\max}}(x) = \frac{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}+n-i+1)\det(\mathbf{\Theta}(x))}{\prod_{j=1}^{m}\Gamma(N_{\mathcal{I}}-j+1)\Gamma(n-j+1)\Gamma(m-j+1)\prod_{i=1}^{N_t}\Gamma(N_t-i+1)} \tag{3.37}$$

where $\mathbf{\Theta}(x)$ *is an* $N_r \times N_r$ *matrix whose entries are defined by*

$$[\mathbf{\Theta}(x)]_{i,j} = \Gamma(N_t-i+1)\left[B(N_s+N_r-i-j+1, N_{\mathcal{I}}-N_r+i+j-1) - \right.$$
$$\left.\sum_{k=0}^{N_t-i}\frac{(xN_s)^k}{\Gamma(k+1)}\Gamma(N_{\mathcal{I}}-N_r+i+j+k-1)U(N_{\mathcal{I}}-N_r+i+j+k-1, i+j+k-N_r-N_s, xN_s)\right], \tag{3.38}$$

Proof: See Appendix A.7. □

When $N_s = 1$ in Theorem 3.5, which corresponds to the interference limited keyhole scenario, the c.d.f. of the maximum eigenvalue can be further simplified as shown in the following corollary.

Corollary 3.1 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_s}(\mathbf{0}_{N_t\times N_s}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_s,N_t}(\mathbf{0}_{N_s\times N_t}, \mathbf{I}\otimes\mathbf{I})$ *and* $\mathbf{H}_3 \sim \mathcal{CN}_{N_r,N_I}(\mathbf{0}_{N_r\times N_I}, \mathbf{I}\otimes\mathbf{I})$, $N_I \geq N_r$. *Then the c.d.f. of the non-zero eigenvalue of*

$\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\left(\mathbf{H}_3\mathbf{H}_3^\dagger\right)^{-1}\mathbf{H}_1\mathbf{H}_2$ *is expressed as*

$$\mathcal{F}_{\lambda_{\max}}(x) = 1 - \frac{\Gamma(N_{\mathcal{I}}+1)}{\Gamma(N_r)\Gamma(N_{\mathcal{I}}-N_r+1)} \sum_{k=0}^{N_t-1}\frac{x^k}{\Gamma(k+1)}\Gamma(N_{\mathcal{I}}-N_r+k+1)U(N_{\mathcal{I}}-N_r+k+1,k-N_r+1,x). \tag{3.39}$$

The following theorem gives the p.d.f. of the maximum eigenvalue of of a random matrix involving a product of three independent random matrices. This will be used to investigate the ergodic capacity of optimum combining system operating in interference-limited Rayleigh-product channels.

Theorem 3.6 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_s}(\mathbf{0}_{N_t\times N_s}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_s,N_t}(\mathbf{0}_{N_s\times N_t}, \mathbf{I}\otimes\mathbf{I})$ *and* $\mathbf{H}_3 \sim \mathcal{CN}_{N_r,N_I}(\mathbf{0}_{N_r\times N_I}, \mathbf{I}\otimes\mathbf{I})$, $N_I \geq N_r$. *Define* $m = \min(N_r,N_s)$, $n = \max(N_r,N_s)$, $p = \max(0, m-N_t)$, $q = \max(m,N_t)$. *Then the p.d.f. of the maximum eigenvalue of matrix* $\frac{1}{N_s}\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\left(\mathbf{H}_3\mathbf{H}_3^\dagger\right)^{-1}\mathbf{H}_1\mathbf{H}_2$ *is given by:*
1) When $N_t \leq N_r$ *or* $N_t \geq N_r \geq N_s$,

$$f_{\lambda_{\max}}(x) = \frac{(-1)^{pN_t}\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}+N_s-i+1)\sum_{l=m-N_t+1}^{m}\det(\mathbf{\Delta}_l(x))}{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}-N_r+m-i+1)\Gamma(m-i+1)\Gamma(n-i+1)} \tag{3.40}$$

where $\mathbf{\Delta}_l(x)$ *is an* $m\times m$ *matrix defined by*

$$[\mathbf{\Delta}_l(x)]_{i,j} = \begin{cases} [\mathbf{\Delta}(\mathbf{x})]_{i,j}, & i\neq l, \\ \frac{N_s(N_sx)^{q-i}\Gamma(N_{\mathcal{I}}-N_r-p+j+q)}{\Gamma(q-i+1)}U(N_{\mathcal{I}}-N_r-p+j+q, j-p-n-m+q+1, xN_s), & i=l, \end{cases} \tag{3.41}$$

2) $N_t \geq N_s \geq N_r$ *or* $N_s \geq N_t \geq N_r$,

$$f_{\lambda_{\max}}(x) = \frac{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}+n-i+1)\sum_{l=1}^{N_r}\det(\mathbf{\Theta}_l(x))}{\prod_{j=1}^{m}\Gamma(N_{\mathcal{I}}-j+1)\Gamma(n-j+1)\Gamma(m-j+1)\prod_{i=1}^{N_t}\Gamma(N_t-i+1)}, \tag{3.42}$$

where, $\mathbf{\Theta}_l(x)$ *is an* $N_r\times N_r$ *matrix defined by*

$$[\mathbf{\Theta}_l(x)]_{i,j} = x^{N_t-i}N_s^{N_t-i+1}\Gamma(N_{\mathcal{I}}-N_r+N_t+j)U(N_{\mathcal{I}}-N_r+N_t+j, N_t-N_r-N_s+j+1, xN_s). \tag{3.43}$$

Proof: See Appendix A.8. □

3.2.3 New Random Determinant Results

We now turn our attention to the statistical properties of the determinant of certain random complex matrix. The derived expressions will be used to derive tight ergodic capacity (or ergodic mutual information) upper bounds or lower bounds in the following chapters.

Lemma 3.9 *Let* $\mathbf{H} \sim \mathcal{CN}_{m,n}(\mathbf{0}_{m\times n}, \mathbf{I}\otimes\mathbf{I})$*, and* $\mathbf{\Omega}$ *is an* $m\times m$ *positive definite matrix with eigenvalue* $\omega_1 > \omega_2 > \cdots > \omega_m > 0$*.* a *is a positive real number. Then, the expected determinant of* $\mathbf{I}_n + a\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}$ *is given by*

$$E\left\{\det\left(\mathbf{I}_n + a\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}\right)\right\} = \frac{\det\left(\mathbf{\Delta}\right)}{\prod_{i<j}^{q}\left(\omega_j - \omega_i\right)}, \tag{3.44}$$

where $\mathbf{\Delta}$ *is a* $m\times m$ *matrix with entries*[2]

$$\left\{\mathbf{\Delta}\right\}_{l,k} = \begin{cases} \omega_l^{k-1}, & k \le m-n, \\ \omega_l^{k-1}\left(1 + a\omega_l\left(n-m+k\right)\right), & k > m-n. \end{cases} \tag{3.45}$$

Proof: See Appendix A.9. □

This lemma presents a new expression for the expected characteristic polynomial of a complex semi-correlated central Wishart matrix. In prior work [82, 108], alternative expressions were obtained via a different approach (i.e., by exploiting a classical characteristic polynomial expansion for the determinant). Those results, however, involved summations over subsets of numbers, with each term involving determinants of partitioned matrices. In contrast, our result in Lemma 3.7 is more computationally-efficient, involving only a single determinant with simple entries. Moreover, it is more amenable to the further analysis in this paper, leading to the following two important theorems.

Theorem 3.7 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_r,N_s}(\mathbf{0}_{N_r\times N_s}, \mathbf{I}\otimes\mathbf{I})$*,* $\mathbf{H}_2 \sim \mathcal{CN}_{N_d,N_r}(\mathbf{0}_{N_d\times N_r}, \mathbf{I}\otimes\mathbf{I})$*,* a *and* b *being positive real numbers. Define* $q = \min(N_d, N_r)$ *and* $p = \max(N_d, N_r)$*. Then the expected determinant of* $\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger(\mathbf{I} + a\mathbf{H}_2\mathbf{H}_2^\dagger)^{-1}\mathbf{H}_2\mathbf{H}_1$ *is given by*

$$E\left\{\det\left(\mathbf{I} + b\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger(\mathbf{I} + a\mathbf{H}_2\mathbf{H}_2^\dagger)^{-1}\mathbf{H}_2\mathbf{H}_1\right)\right\} = \mathcal{K}\det\left(\bar{\mathbf{\Xi}}\right), \tag{3.46}$$

where $\bar{\mathbf{\Xi}}$ *is a* $q\times q$ *matrix with entries*

$$\left\{\bar{\mathbf{\Xi}}\right\}_{m,n} = \begin{cases} a^{1-\tau}\vartheta_{\tau-1}(a), & n \le q - N_s, \\ a^{1-\tau}\vartheta_{\tau-1}(a) + ba^{-\tau}\left(N_s - q + n\right)\vartheta_\tau(a), & n > q - N_s \end{cases} \tag{3.47}$$

with $\tau = p - q + m + n$*, and*

$$\vartheta_\tau(a) = \Gamma\left(\tau\right)U\left(\tau, p+q, 1/a\right)\ . \tag{3.48}$$

[2] When $m < n$, $\{\mathbf{\Delta}\}_{l,k} = \omega_l^{k-1}\left(1 + a\omega_l\left(n-m+l\right)\right)$.

Proof: Utilizing *Lemma 3.9*, [84, Lemma 2] and (10.29) yields the desired result. □

Theorem 3.8 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t\times N_k}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r\times N_k}, \mathbf{I}\otimes\mathbf{I})$, *and* $\mathbf{A} \in \mathcal{C}^{N_k\times N_k}$. *Define* $m = \max(N_t, N_r)$, $n = \min(N_t, N_r)$, $p = \min(n, N_k)$, *and let* $b_1 \leq b_2 \leq \cdots \leq b_{N_k}$ *denote the non-zero eigenvalues of* $\mathbf{B} \triangleq \mathbf{A}\mathbf{A}^\dagger$. *Then the expected determinant of the matrix* $\mathbf{I} + \frac{\gamma}{N_t}\mathbf{H}_1\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1^\dagger$ *(for some constant* γ*) is given by*

$$\mathtt{E}\left\{\det\left(\mathbf{I} + \frac{\gamma}{N_t}\mathbf{H}_1\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1^\dagger\right)\right\} = \frac{\det(\mathbf{\Delta})}{\prod_{i<j}^{N_k}(b_j - b_i)}, \tag{3.49}$$

where $\mathbf{\Delta}$ *is an* $N_k \times N_k$ *matrix with entries*

$$[\mathbf{\Delta}]_{l,k} = \begin{cases} b_l^{k-1}, & k \leq N_k - p, \\ b_l^{k-1}\left(1 + \frac{\gamma b_l}{N_t}(m - N_k + k)(n - N_k + k)\right), & k > N_k - p. \end{cases} \tag{3.50}$$

Proof: See Appendix A.10. □

Lemma 3.10 *Let* $\mathbf{H} \sim \mathcal{CN}_{m,n}(\mathbf{0}_{m\times n}, \mathbf{I}\otimes\mathbf{I})$, *and* $\mathbf{\Omega}$ *is an* $m\times m$ *positive definite matrix with eigenvalue* $\omega_1 > \omega_2 > \cdots > \omega_m > 0$. *Define* $s = \min(m, n)$, *and*

$$\mathbf{\Phi} = \begin{cases} \mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}, & m \geq n, \\ \mathbf{\Omega}\mathbf{H}_1\mathbf{H}_1^\dagger, & m < n. \end{cases} \tag{3.51}$$

The expected log-determinant of $\mathbf{\Phi}$ *is given by*

$$E\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} = \sum_{k=1}^{s}\psi\left(n - s + k\right) + \frac{\sum_{k=m-s+1}^{m}\det\left(\mathbf{Y}_k\right)}{\prod_{i<j}^{m}\left(\omega_j - \omega_i\right)}, \tag{3.52}$$

where $\psi\left(\cdot\right)$ *is the digamma function [26], and* $\mathbf{Y}_k$ *is an* $m \times m$ *matrix with entries*

$$\left\{\mathbf{Y}_k\right\}_{i,j} = \begin{cases} \omega_i^{j-1}, & j \neq k, \\ \omega_i^{j-1}\ln\omega_i, & j = k. \end{cases} \tag{3.53}$$

When $m = s$, *(3.52) reduces to*

$$E\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} = \sum_{k=1}^{s}\psi\left(n - s + k\right) + \ln\det\left(\mathbf{\Omega}\right)\ . \tag{3.54}$$

Proof: See Appendix A.11. □

We note that the above expected natural logarithm of the determinant for $m \geq n$ has been investigated in [57], where the derived expression is rather complicated, involving summations of determinants whose elements are in terms of the inverse of a certain Vandermonde matrix. We also note the $m < n$ and

$m = n = s$ cases have been considered in [28, 108]. Our result, in contrast, gives a simple unified expression which embodies all of these cases. Moreover, based on Lemma 3.10, we obtain the following two important theorems.

Theorem 3.9 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_r,N_s}(\mathbf{0}_{N_r\times N_s}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_d,N_r}(\mathbf{0}_{N_d\times N_r}, \mathbf{I}\otimes\mathbf{I})$, *a being positive real numbers. Define* $q = \min(N_d, N_r)$, $p = \max(N_d, N_r)$, $s = \min(N_s, q)$ *and*

$$\mathbf{\Phi} = \begin{cases} \mathbf{H}_1^\dagger \mathbf{Q} \mathbf{H}_1, & q \geq N_s, \\ \mathbf{Q}\mathbf{H}_1\mathbf{H}_1^\dagger, & q < N_s. \end{cases} \tag{3.55}$$

where $\mathbf{Q}$ *is a* $q \times q$ *matrix with the same non-zero eigenvalues as* $\mathbf{H}_2^\dagger(\mathbf{I} + a\mathbf{H}_2\mathbf{H}_2^\dagger)^{-1}\mathbf{H}_2$. *Then the expected log-determinant of* $\mathbf{\Phi}$ *is given by*

$$E\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} = \sum_{k=1}^{s}\psi\left(N_s - s + k\right) + \mathcal{K}\sum_{k=q-s+1}^{q}\det\left(\mathbf{W}_k,\right) \tag{3.56}$$

where $\mathbf{W}_k$ *is a* $q \times q$ *matrix with entries*

$$\left\{\mathbf{W}_k\right\}_{m,n} = \begin{cases} a^{1-\tau}\vartheta_{\tau-1}(a), & n \neq k, \\ \varsigma_{m+n}(a), & n = k, \end{cases} \tag{3.57}$$

where τ *and* $\vartheta_{\tau-1}(\cdot)$ *are defined as in (3.48), and*

$$\varsigma_t(a) = \sum_{i=0}^{2q-t} a^{2q-t-i}\Gamma\left(p+q-i-1\right)\binom{2q-t}{i}\left(\psi\left(p+q-i-1\right) - \sum_{l=0}^{p+q-i-2} g_l\left(\frac{1}{a}\right)\right), \tag{3.58}$$

where $g_l(\cdot)$ *denotes the auxiliary function*

$$g_l(x) = e^x E_{l+1}(x) \tag{3.59}$$

with $E_{l+1}(\cdot)$ *denoting the exponential integral function of order* $l+1$.

When $q = s$, *(3.56) reduces to*

$$\begin{aligned} E\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} = \sum_{k=1}^{s}\psi\left(N_s - s + k\right) + \sum_{i=0}^{q-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\sum_{k=0}^{2q-l-2}\binom{2q-l-2}{k}\mathcal{A}\left(i,j,l,p,q\right) \\ \times a^{2q-l-2-k}\Gamma\left(p+q-k-1\right)\left(\psi\left(p+q-k-1\right) - \sum_{m=0}^{p+q-k-2} g_m\left(1/a\right)\right). \end{aligned} \tag{3.60}$$

Proof: See Appendix A.12. □

Theorem 3.10 *Let* $\mathbf{H}_1 \sim \mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t\times N_k}, \mathbf{I}\otimes\mathbf{I})$, $\mathbf{H}_2 \sim \mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r\times N_k}, \mathbf{I}\otimes\mathbf{I})$, *and* $\mathbf{A} \in \mathcal{C}^{N_k\times N_k}$. *Define* $m = \max(N_t, N_r)$, $n = \min(N_t, N_r)$, $p = \min(n, N_k)$, *and let* $b_1 \leq b_2 \leq \cdots \leq b_{N_k}$

denote the non-zero eigenvalues of $\mathbf{B} \triangleq \mathbf{A}\mathbf{A}^\dagger$. *Define*

$$\mathbf{\Phi} \triangleq \begin{cases} \mathbf{H}_1^\dagger \mathbf{A}^\dagger \mathbf{H}_2^\dagger \mathbf{H}_2 \mathbf{A} \mathbf{H}_1, & p = N_t, \\ \mathbf{A}^\dagger \mathbf{H}_2^\dagger \mathbf{H}_2 \mathbf{A} \mathbf{H}_1 \mathbf{H}_1^\dagger, & p = N_k, \\ \mathbf{H}_2 \mathbf{A} \mathbf{H}_1 \mathbf{H}_1^\dagger \mathbf{A}^\dagger \mathbf{H}_2^\dagger, & p = N_r. \end{cases} \tag{3.61}$$

The expected log-determinant of $\mathbf{\Phi}$ *is given by*

$$\mathtt{E}\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} = \sum_{s=1}^{p} \psi(m-p+s) + \sum_{s=1}^{p} \psi(n-p+s) + \frac{\sum_{s=N_k-p+1}^{N_k} \det(\mathbf{Y}_s)}{\prod_{i<j}^{N_k}(b_j - b_i)}, \tag{3.62}$$

where $\psi(\cdot)$ *is the digamma function [26], and* $\mathbf{Y}_s$ *is an* $N_k \times N_k$ *matrix with entries*

$$[\mathbf{Y}_s]_{l,k} = \begin{cases} b_l^{k-1}, & k \neq s, \\ b_l^{k-1} \ln b_l, & k = s. \end{cases} \tag{3.63}$$

When $p = N_k$, *(3.62) reduces to*

$$\mathtt{E}\{\ln\det(\mathbf{\Phi})\} = \sum_{s=1}^{p} \psi(m-p+s) + \sum_{s=1}^{p} \psi(n-p+s) + \ln\det(\mathbf{B}). \tag{3.64}$$

Proof: The result can be obtained by applying Lemma 3.10 twice along with some algebraic manipulations. □

3.3 Conclusion

This chapter has introduced two key mathematical tools employed in the thesis, namely, majorization theory and finite RMT. The first section has given a brief introduction on majorization theory which will be primarily used in Chapter 4 for the ergodic capacity analysis of MIMO Nakagami-m fading channels.

The second part presented one of the major contributions of the thesis, namely, a set of new random matrix results, i.e., eigenvalue distribution, expectation of determinant and log-determinant properties, which find direct applications in the capacity and performance analysis of various MIMO systems.

Specifically, the first crucial result was presented in Lemma 3.7, which gives a unified expression for the p.d.f. of the unordered eigenvalue of a semi-correlated Wishart matrix. This convenient and nice expression plays a key role in the derivation of the p.d.f. of the unordered eigenvalue of certain product matrices shown in Theorem 3.1 and Theorem 3.2, which serve as the essential mathematical tools when studying the exact ergodic capacity performance of MIMO dual-hop AF systems in Chapter 6 and MIMO multi-keyhole systems in Chapter 5, respectively.

We then presented new results for the c.d.f. of the maximum eigenvalue of product matrices arising from the analysis of MIMO multi-keyhole channels in Theorem 3.3, as well as its asymptotic first-order ex-

pansion in Theorem 3.4. These expressions will be applied in the outage probability, diversity order and coding gain analysis of the optimal beamforming scheme operating over MIMO multi-keyhole channels in Chapter 5. In parallel with Theorem 3.3, Theorem 3.5 and Theorem 3.6 showed new expressions for the product matrices emerged from the analysis of interference-limited Rayleigh product channels. These results will be directly employed in Chapter 7 for studying the performance of optimum combining scheme, by deriving closed-form expressions for various important performance metrics of interest such as ergodic capacity, outage probability, diversity order and SER.

In addition, new random determinant properties have been considered. Similar to the approach dealing with the eigenvalue distribution, the first critical step was to derive simple and unified expressions for the expected determinant and log-determinant of a semi-correlated Wishart matrix shown in Lemma 3.9 and Lemma 3.10, respectively. These results were directly invoked in the derivation of the expected determinant and log-determinant of product matrices of interest. In particular, Theorem 3.7 and Theorem 3.9 provide new results for MIMO dual-hop AF systems, which will be utilized to investigate the ergodic capacity upper and lower bound in Chapter 6. The expressions exhibited in Theorem 3.8 and Theorem 3.10 will be used in Chapter 5 to study the ergodic capacity bounds of MIMO multi-keyhole channels.

Chapter 4

Capacity Bounds for MIMO Nakagami-m Fading Channels

4.1 Introduction

Understanding the fundamental limits of multiple antenna wireless channels has gained enormous attention from the research community since the invention of MIMO antenna systems. One important area of research is to derive exact capacity expression or tight capacity bound, which provides efficient means for evaluating the MIMO channel capacity. And this has been done for various statistical channel models of interest, e.g., MIMO Rayleigh fading channels [14, 46, 82, 84, 89] or MIMO Rician fading channels [6, 18, 36, 45, 63, 107].

Although Rayleigh and Rician fading channels are arguably the most popular statistical models for fading, a more powerful model, namely Nakagami-m fading, was proposed to capture a variety of physical channel environments [71]. The generality of Nakagami-m fading channel model not only allows to embrace both the Rayleigh and Rician fading scenarios, but more importantly, it has been found to be a very good fitting for the mobile radio channel [90]. However, despite its generality, there has been very limited works available on the capacity of multiple antenna Nakagami-m fading channels in the literature. For a SIMO or MISO Nakagami-m fading channel, exact capacity expressions were obtained in [110]. In the latest results of [24], Fraidenraich et. al derived exact capacity formulas for 2×2 and 2×3 Nakagami-m channels, with the fading parameter m being restricted to be an integer. This contemporary list of references indicates that despite the need to know the fundamental limits of Nakagami-m MIMO channels, little is understood.

In this chapter, we investigate the ergodic capacity of MIMO Nakagami-m fading channels with arbitrary real $m \geq 1/2$ and arbitrary finite number of antennas at both ends. Two models are considered, namely, conventional co-located MIMO (C-MIMO) and distributed MIMO (D-MIMO) systems. We derive tight upper and lower capacity bounds for both models. In addition, a simple and concise ergodic capacity upper bound is obtained in the high SNR regime, which enables the analysis of the impact of the channel

fading parameter m on the ergodic capacity. Moreover, we also look into the asymptotic behavior of the ergodic capacity in the large-system limit when the number of antennas at one or both side(s) goes to infinity.

4.2 System Model

In this section, we introduce the mathematical models for D-MIMO and C-MIMO antenna systems. The D-MIMO model reflects the distinctive large-scale fading effects for each antenna-pair, making it useful for analyzing a MIMO channel with the antennas distributed in a large area. On the other hand, the C-MIMO model will be used for the analysis of a traditional point-to-point MIMO channel where the antennas at either side are co-located, and have the same large-scale fading.

4.2.1 MIMO Systems

We consider a general D-MIMO system, where there are N_r receive antennas and L radio ports located far apart, each with N_t transmit antennas.[1] The antennas at a given port are assumed to go through the same large-scale fading, while the antennas at different ports undergo different large-scale fading (i.e., different path losses and shadowing effects). The received signal vector $\mathbf{y} \in \mathbb{C}^{N_r}$ can be related to the transmitted symbol vector $\mathbf{x} \in \mathbb{C}^{N_t}$ by

$$\mathbf{y} = \sqrt{\frac{P}{LN_t}}\mathbf{H}\mathbf{\Phi}^{\frac{1}{2}}\mathbf{x} + \boldsymbol{\zeta}, \tag{4.1}$$

where $\mathbf{x}$ has the covariance matrix of $\mathtt{E}[\mathbf{x}\mathbf{x}^\dagger] = \mathbf{Q}$, P denotes the total transmit power, $\boldsymbol{\zeta} \in \mathbb{C}^{N_r}$ is the complex AWGN vector with zero mean and the covariance matrix of $\mathtt{E}[\boldsymbol{\zeta}\boldsymbol{\zeta}^\dagger] = N_0\mathbf{I}$, and

$$\mathbf{\Phi}^{\frac{1}{2}} = \mathsf{diag}\left(\sqrt{\frac{l_1}{D_1^v}}, \ldots, \sqrt{\frac{l_1}{D_1^v}}, \ldots, \sqrt{\frac{l_L}{D_L^v}}, \ldots, \sqrt{\frac{l_L}{D_L^v}}\right) \in \mathbb{R}^{N_tL \times N_tL} \tag{4.2}$$

is a diagonal matrix accounting for the large-scale fading effect, in which the path loss is characterized by D_l^{-v} for some exponent v (typically from 4 to 6 depending on the environments), $\{l_i\}_{i=1}^L$ are independent and log-normal random variables (i.e., with the corresponding means $\{\mu_i\}$ and standard deviations $\{\sigma_i\}$), with p.d.f. given by

$$p(l) = \frac{\eta}{\sqrt{2\pi\sigma^2 l}} e^{-\frac{(\eta \ln l - \mu)^2}{2\sigma^2}}, \tag{4.3}$$

where $\eta = \frac{10}{\ln 10} \approx 4.3429$, and $\mathbf{H} \in \mathbb{C}^{N_r \times N_tL}$ is the channel matrix addressing the small-scale fading, and the elements of $\mathbf{H} = [h_{ij}]$ are assumed to be i.i.d. with uniformly distributed phase and the magnitude, $x = |h_{ij}|$, following a Nakagami-m p.d.f.

$$p(x) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m x^{2m-1} e^{-\left(\frac{m}{\Omega}\right)x^2}, \quad \text{for } x \geq 0 \text{ and } m \geq 0.5, \tag{4.4}$$

[1] Different number of antennas for each port can be easily accommodated in the formulation.

where $\Gamma(\cdot)$ denotes the gamma function, $m \triangleq \frac{\mathtt{E}^2[x^2]}{\mathtt{Var}[x^2]}$, and $\Omega \triangleq \mathtt{E}[x^2]$.

Note that in the above model, the overall channel between the transmitter and the receiver is expressed as a product of the small-scale fading and the large-scale fading, as in [74, 75].

When $L = 1$, this D-MIMO model degenerates to the conventional C-MIMO system. In this case, we focus only on the small scale fading effect and the large scale fading can be ignored, as it is identical for every antenna pair. Hence, (4.1) can be reduced to

$$\mathbf{y} = \sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{x} + \boldsymbol{\zeta}. \tag{4.5}$$

4.2.2 Ergodic Capacity

We assume that CSI is known perfectly at the receiver, and that an equal-power allocation across the transmit antennas is used, i.e., $\mathbf{Q} = \sqrt{\frac{P}{N_t}}\mathbf{I}$. Therefore, for C-MIMO systems, the ergodic capacity can be expressed as

$$\mathcal{C} = \mathtt{E}\left[\log_2 \det\left(\mathbf{I} + \frac{P}{N_0 N_t}\mathbf{H}\mathbf{H}^\dagger\right)\right]. \tag{4.6}$$

Similarly, for D-MIMO systems, we have

$$\mathcal{D} = \mathtt{E}\left[\log_2 \det\left(\mathbf{I} + \frac{P}{L N_t N_0}\mathbf{H}\boldsymbol{\Phi}\mathbf{H}^\dagger\right)\right]. \tag{4.7}$$

In the following sections, we first develop exact capacity bounds for C-MIMO systems based on majorization theory, and then extend the analysis to the more general D-MIMO systems.

4.3 Capacity Bounds of C-MIMO Nakagami-m Channels

In this section, we derive ergodic capacity upper and lower bounds for C-MIMO Nakagami-m fading channels, where only small-scale fading is considered. In addition, we study the high SNR regime, in which simpler results can be obtained to gain insight on the system performance. The analysis we carry out is mainly based on majorization theory. For convenience, we define

$$s \triangleq \min(N_t, N_r), \tag{4.8}$$

$$t \triangleq \max(N_t, N_r). \tag{4.9}$$

4.3.1 Ergodic Capacity Upper Bounds

Utilizing majorization theory, we derive several upper bounds of the ergodic capacity for Nakagami-m channels, which are now given in the following theorems.

Theorem 4.1 *The ergodic capacity of MIMO Nakagami-m fading channels is upper bounded by*

$$\mathcal{C} \leq \bar{\mathcal{C}}_1 = \frac{s}{\Gamma(tm)\ln 2} G_{3,2}^{1,3}\left(\frac{P}{N_t N_0}\frac{\Omega}{m}\Big|_{1,0}^{1-tm,1,1} \right). \tag{4.10}$$

Proof: See Appendix B.1. □

Similar upper bounds can be obtained using different majorization relationships. Nevertheless, among them, the upper bound $\bar{\mathcal{C}}_1$ is the tightest. This result is given in the following theorem.

Theorem 4.2 *Ergodic capacity upper bounds $\bar{\mathcal{C}}_1$, $\bar{\mathcal{C}}_2$, $\bar{\mathcal{C}}_3$ satisfy the following relationship:*

$$\bar{\mathcal{C}}_1 \leq \bar{\mathcal{C}}_2 \ \textit{and} \ \bar{\mathcal{C}}_1 \leq \bar{\mathcal{C}}_3, \tag{4.11}$$

where

$$\begin{cases} \bar{\mathcal{C}}_2 = \dfrac{t}{\Gamma(sm)\ln 2} G_{3,2}^{1,3}\left(\dfrac{P}{N_t N_0}\dfrac{\Omega}{m}\Big|_{1,0}^{1-sm,1,1} \right), \\ \bar{\mathcal{C}}_3 = \dfrac{s}{\Gamma(stm)\ln 2} G_{3,2}^{1,3}\left(\dfrac{P}{sN_t N_0}\dfrac{\Omega}{m}\Big|_{1,0}^{1-stm,1,1} \right). \end{cases} \tag{4.12}$$

Proof: See Appendix B.2. □

All the three bounds are expressed in closed form and can be evaluated very efficiently using standard softwares like Mathematica. Since the upper bounds involve the Meijer G-function, they do not offer much physical insight on the capacity performance. In the following, we consider the tightest capacity upper bound, $\bar{\mathcal{C}}_1$, in the high SNR regime to derive simpler expressions for more insights.

Corollary 4.1 *For MIMO Nakagami-m fading channels, in the high SNR regime, the ergodic capacity upper bound $\bar{\mathcal{C}}_1$ can be approximated as*

$$\bar{\mathcal{C}}_1 \approx \bar{\mathcal{C}}_{\mathsf{hsnr}} = s\log_2\left(\frac{P}{N_t N_0}\right) + \frac{s}{\ln 2}\left[\psi(tm) - \ln\left(\frac{m}{\Omega}\right)\right], \tag{4.13}$$

where $\psi(\cdot)$ is the digamma function [26, (8.365.4)].

Proof: At high SNRs, $\log_2\left(1+\frac{P}{N_t N_0}x\right)$ can be approximated by $\log_2\left(\frac{P}{N_t N_0}x\right)$. As such, we can have

$$\bar{\mathcal{C}}_{\mathsf{hsnr}} = \frac{s}{\ln 2}\int_0^\infty \ln\left(\frac{P}{N_t N_0}r\right)\frac{\left(\frac{m}{\Omega}\right)^{tm}}{\Gamma(tm)} r^{tm-1}e^{-\frac{m}{\Omega}r}dr \tag{4.14}$$

$$= s\log_2\left(\frac{P}{N_t N_0}\right) + \frac{s}{\ln 2}\frac{\left(\frac{m}{\Omega}\right)^{tm}}{\Gamma(tm)}\int_0^\infty \ln(r) r^{tm-1}e^{-\frac{m}{\Omega}r}dr \tag{4.15}$$

$$= s\log_2\left(\frac{P}{N_t N_0}\right) + \frac{s}{\ln 2}\left[\psi(tm) - \ln\left(\frac{m}{\Omega}\right)\right], \tag{4.16}$$

where the following integration formula has been used [26, (4.352.1)]

$$\int_0^\infty t^{v-1}e^{-at}\ln t dt = \frac{1}{a^v}\Gamma(v)[\psi(v) - \ln a]. \tag{4.17}$$

□

Corollary 4.2 *The ergodic capacity upper bound approximation, $\bar{\mathcal{C}}_{\mathsf{hsnr}}$, is a monotonic increasing function of the channel fading parameter m.*

Proof: We prove the corollary by showing the first derivative of $\bar{\mathcal{C}}_{\mathsf{hsnr}}$ with respect to m is strictly greater than zero regardless of s and t. This is done as follows.

$$\begin{aligned}
\frac{d\bar{\mathcal{C}}_{\mathsf{hsnr}}}{dm} &= \frac{s}{\ln 2}\left[\psi^{(1)}(tm) - \frac{1}{m}\right] && (4.18)\\
&= \frac{s}{\ln 2}\left[t\sum_{k=0}^{\infty}\frac{1}{(tm+k)^2} - \frac{1}{m}\right] && (4.19)\\
&> \frac{s}{\ln 2}\left[t\sum_{k=0}^{\infty}\frac{1}{(tm+k)(tm+k+1)} - \frac{1}{m}\right] && (4.20)\\
&= \frac{s}{\ln 2}\left[t\sum_{k=0}^{\infty}\left(\frac{1}{tm+k} - \frac{1}{tm+k+1}\right) - \frac{1}{m}\right] && (4.21)\\
&= 0, && (4.22)
\end{aligned}$$

where from (4.18) to (4.19), we have used the derivative property of digamma function [26, (8.363.8)]

$$\psi^{(n)}(x) = (-1)^{n+1}n!\sum_{k=0}^{\infty}\frac{1}{(x+k)^{n+1}}. \tag{4.23}$$

□

Corollary 4.2 is quite intuitive since a larger m corresponds to a less severe fading environment, and the ergodic capacity is anticipated to increase with m.

4.3.2 Upper Bound for Large Systems at High SNR

It was revealed in [15] that the ergodic capacity of a general MIMO fading channel grows linearly with the minimum number of antennas at both ends in the large-system limit where the numbers of antennas at both ends approach infinity. However, the asymptotic result when the number of antennas at only one side goes to infinity is not available. Here, we derive such results for Nakagami channels through the capacity upper bound approximation $\bar{\mathcal{C}}_{\mathsf{hsnr}}$.

Three cases are of interest: (i) $N_r \to \infty$ while N_t being fixed, (ii) $N_t \to \infty$ while N_r being fixed, and (iii) both $N_t, N_r \to \infty$ while keeping $\frac{N_t}{N_r} = \beta$ fixed. For convenience, we assume $\Omega = 1$. In the

analysis, the following approximation is used [4, (6.3.18)]

$$\psi(x) \approx \ln x, \;\; \text{if } x \to \infty. \tag{4.24}$$

For case (i),

$$\bar{\mathcal{C}}_{\mathsf{hsnr}} \overset{N_r \to \infty}{=} N_t \log_2 \left(\frac{P}{N_t N_0} \right) + N_t \log_2 N_r. \tag{4.25}$$

As such, asymptotically, the ergodic capacity increases logarithmically with the number of receive antennas. Considering case (ii), we then have

$$\bar{\mathcal{C}}_{\mathsf{hsnr}} \overset{N_t \to \infty}{=} N_r \log_2 \left(\frac{P}{N_t N_0} \right) + N_r \log_2 N_t \tag{4.26}$$

$$= N_r \log_2 \left(\frac{P}{N_0} \right). \tag{4.27}$$

The result indicates that the increase in the number of transmit antennas does not provide any capacity gain, which aligns with previous studies. Finally, for case (iii), we consider two separate cases, namely, $\beta \leq 1$ and $\beta > 1$. When $\beta \leq 1$, we have

$$\frac{\bar{\mathcal{C}}_{\mathsf{hsnr}}}{N_t} \overset{N_r \to \infty}{=} \log_2 \left(\frac{P}{N_0} \right) + \log_2 \beta^{-1}, \tag{4.28}$$

which shows that a linear increase in the ergodic capacity is achieved as long as N_t and N_r increase at the same rate. On the other hand, when $\beta > 1$, we have

$$\frac{\bar{\mathcal{C}}_{\mathsf{hsnr}}}{N_r} \overset{N_r \to \infty}{=} \log_2 \left(\frac{P}{N_0} \right), \tag{4.29}$$

which is independent of β, showing no capacity benefit from increasing N_t beyond N_r.

The above asymptotic results not only agree with that in [15] which indicates the linear capacity increase with the minimum number of antennas, but also provide additional insights on how the capacity grows with the greater number of antennas. Besides, the above scaling law for Nakagami-fading channels reveals the same asymptotic behavior as for Rayleigh-fading channels seen in [28]. From the scaling results, we notice that in all the three cases, the ergodic capacity is independent of m, which is intuitive as the increasing number of antennas helps to eliminate the effect of fading.

4.3.3 Ergodic Capacity Lower Bounds

In this subsection, our focus is on the derivation of ergodic capacity lower bound for the general Nakagami-m fading channels based on majorization theory.

Theorem 4.3 *The ergodic capacity of Nakagami-m fading channels is lower bounded by*

$$\mathcal{C} \geq \underline{\mathcal{C}}_1 = \frac{1}{\Gamma(stm) \ln 2} G_{3,2}^{1,3} \left(\frac{P}{N_t N_0} \frac{\Omega}{m} \Big|_{1,0}^{1-stm,1,1} \right). \tag{4.30}$$

Proof: See Appendix B.3. □

When $s = 1$, the channel degenerates to a SIMO or MISO system. The low bound $\underline{\mathcal{C}}_1$ becomes exact and is the same as the upper bound $\bar{\mathcal{C}}_2$. The lower bound is, however, not tight for a general MIMO channel (i.e., when $s > 1$) as will be shown in the following high SNR analysis.

Corollary 4.3 *In the high SNR regime, the lower bound $\underline{\mathcal{C}}_1$ can be approximated by*

$$\underline{\mathcal{C}}_1 \approx \underline{\mathcal{C}}_{\mathsf{hsnr}} = \log_2\left(\frac{P}{N_t N_0}\right) + \frac{1}{\ln 2}\left[\psi(stm) - \ln\left(\frac{m}{\Omega}\right)\right]. \tag{4.31}$$

Proof: The proof is similar to the proof for Corollary 4.1 and is omitted. □

Using the high SNR approximation of the lower bound, $\underline{\mathcal{C}}_{\mathsf{hsnr}}$, and further considering it in the asymptotic large-system limit where s or t (or both) approaches infinity, we get

$$\underline{\mathcal{C}}_{\mathsf{hsnr}} \overset{s,t\to\infty}{=} \log_2\left(\frac{N_r P}{N_0}\right). \tag{4.32}$$

The lower bound shows no linear capacity increase with the minimum number of the antennas, which does not align with the known results in the high SNR regime. Therefore, the bound, $\underline{\mathcal{C}}_1$, is not tight at least in the high SNR regime. Nevertheless, we shall show that this bound has an interpretation of a low SNR approximation and may give a reasonably tight bound at low SNRs.

To see this, we assume that $N_t \geq N_r$ [$N_t < N_r$ can be dealt with similarly by using (10.156)], so

$$\det\left(\mathbf{I} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{H}^\dagger\right) = \prod_{i=1}^{N_r}\left(1 + \frac{P}{N_t N_0}\lambda_k\right), \tag{4.33}$$

where λ_k, for $k = 1, \ldots, N_r$, are the eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$. Expanding the product at the right-hand-side of (4.33), and ignoring the second and higher order terms in $\frac{P}{N_0}$ (for low SNRs), we get

$$\mathcal{C} \geq \mathtt{E}\left[\log_2\left(1 + \frac{P}{N_t N_0}\|\mathbf{H}\|^2\right)\right], \tag{4.34}$$

where $\|\mathbf{H}\|$ is the Frobenius norm of $\mathbf{H}$. The right-hand-side of (4.34) is exactly the ergodic capacity lower bound $\underline{\mathcal{C}}_1$. Therefore, $\underline{\mathcal{C}}_1$ should be reasonably tight in the low SNR regime, while it degrades with the number of antennas, due to the significance of the higher order terms.

4.4 Capacity Bounds for D-MIMO Nakagami-m Channels

Here, we consider a D-MIMO channel which undergos composite Nakagami-m and log-normal fading, and attempt to derive similar capacity bounds for this channel.

4.4.1 Ergodic Capacity Upper Bounds

Theorem 4.4 *The ergodic capacity of a composite D-MIMO channel is upper bounded by*

$$\mathcal{D} \leq \bar{\mathcal{D}}_1 = \frac{N_t}{\Gamma(N_r m)\ln 2} \sum_{i=1}^{L} \frac{1}{\sqrt{\pi}} \sum_{j=1}^{N} w_j V_i(a_j), \tag{4.35}$$

where $V_i(t) = G_{3,2}^{1,3}\left(\frac{P}{LN_tN_0} \frac{e^{\frac{\sqrt{2}\sigma_i t+\mu_i}{\eta}}}{D_i^v} \frac{\Omega}{m} \middle| \begin{matrix} 1-mN_r, 1, 1 \\ 1, 0 \end{matrix} \right)$, *with* $\{a_j\}_{j=1}^N$ *corresponding to the zeros of the* N*-th order Hermite polynomial and* $\{w_j\}_{j=1}^N$ *are the weight factors tabulated in Table 25.10 of [4].*

Proof: See Appendix B.4. □

In the proof of Theorem 4.4, Gaussian-Hermite quadratic integration has been employed to approximate the infinite integral. While (4.35) can be used to compute the upper bound for the general composite Nakagami-m and log-normal fading channels, the computation of Meijer G-function can still be time-consuming at extreme low SNRs [e.g., < -15 (dB)]. A simpler expression is possible for the special case such as the Rayleigh and log-normal composite channel, and is given below.

Corollary 4.4 *The ergodic capacity of a composite Rayleigh and log-normal D-MIMO fading channel (i.e.,* $m = 1$ *and* $\Omega = 1$*) is upper bounded by*

$$\mathcal{D}_{\mathsf{Rayleigh}} \leq \bar{\mathcal{D}}_2 = \frac{N_t}{\ln 2} \sum_{i=1}^{L} \frac{1}{\sqrt{\pi}} \sum_{j=1}^{N} w_j T_i(a_j), \tag{4.36}$$

where $T_i(t) = e^{\frac{\frac{LN_tD_i^vN_0}{\sqrt{2}\sigma_i t+\mu_i}}{Pe^{\frac{\eta}{\eta}}}} \sum_{k=0}^{N_r-1} E_{k+1}\left(\frac{LN_tD_i^vN_0}{Pe^{\frac{\sqrt{2}\sigma_i t+\mu_i}{\eta}}} \right)$ *with* $E_n(x)$ *denoting the exponential integral of order* n *[26], and* $\{w_j\}$ *and* $\{a_j\}$ *are defined in (4.35).*

Proof: The outline of the proof is similar to that of the general Nakagami-m and log-normal composite channel. Specifically, the proof requires the capacity expression in [82], and the Gaussian-Hermite quadratic integration approximation. The accuracy of the Gaussian-Hermite approximation has been studied in [74], which has shown that the approximation is very accurate for $N \geq 4$. □

The above capacity bounds, though in closed form, are too complex to gain insights. It is thus of interest to consider the high SNR regime for simplification, which we do in the following.

Corollary 4.5 *For composite Nakagami-*m *and log-normal fading channels, in the high SNR regime, the ergodic capacity upper bound* $\bar{\mathcal{D}}_1$ *can be approximated as*

$$\bar{\mathcal{D}}_1 \approx \bar{\mathcal{D}}_{\mathsf{hsnr}} = LN_t \log\left(\frac{P}{LN_tN_0}\right) + \frac{LN_t}{\ln 2}\left[\psi(N_r m) - \ln\left(\frac{m}{\Omega}\right)\right] - N_t v \sum_{i=1}^{L} \log D_i + \frac{N_t}{\eta \ln 2} \sum_{i=1}^{L} \mu_i. \tag{4.37}$$

Proof: See Appendix B.5. □

The above result is quite informative. This clearly indicates the separate effects of small-scale and large-scale fading on the channel ergodic capacity. (4.37) decomposes the ergodic capacity into two parts: The first part accounts for the small-scale fading which is equivalent to a MIMO system with $N_t L$ transmit antennas and N_r receive antennas operating in Nakagami-m fading channels, while the second part explains the large-scale fading effect, plus the path loss effect on the ergodic capacity. The impact of log-normal fading can also be seen from the mean fading parameters $\{\mu_i\}$.

4.4.2 Ergodic Capacity Lower Bounds

In this subsection, we derive a lower bound for the ergodic capacity of the composite log-normal and Nakagami fading channels. To do so, however, the lower bound for the general D-MIMO system is not available due to the lack of analytical p.d.f. of the sum of weighted i.i.d. gamma random variables. We thus consider a special case when the number of ports is $L = 1$.

Theorem 4.5 *For the composite log-normal and Nakagami fading channels, when $L = 1$, the ergodic capacity is lower bounded by*

$$\mathcal{D} \geq \underline{\mathcal{D}}_1 = \frac{1}{\Gamma(stm)\ln 2}\frac{1}{\sqrt{\pi}}\sum_{i=1}^{N} w_i U(a_i), \tag{4.38}$$

where $U(t) = G_{3,2}^{1,3}\left(\frac{Pe^{\frac{\sqrt{2}\sigma t+\mu}{\eta}}}{N_t N_0 D^v}\frac{\Omega}{m}\middle|\begin{matrix}1-stm,1,1\\1,0\end{matrix}\right)$, $\{w_i\}$ and $\{a_i\}$ have been defined in (4.35).

Proof: See Appendix B.6. □

Corollary 4.6 *For the composite log-normal and Rayleigh fading channels (i.e., $m = 1$ and $\Omega = 1$), when $L = 1$, the ergodic capacity is lower bounded by*

$$\mathcal{D}_{\text{Rayleigh}} \geq \underline{\mathcal{D}}_2 = \frac{1}{\ln 2}\frac{1}{\sqrt{\pi}}\sum_{i=1}^{N} w_i Z(a_i), \tag{4.39}$$

where $Z(t) = e^{\frac{N_t N_0 D^v}{Pe^{\frac{\sqrt{2}\sigma t+\mu}{\eta}}}}\sum_{k=0}^{N_t N_r-1} E_{k+1}\left(\frac{N_t N_0 D^v}{Pe^{\frac{\sqrt{2}\sigma t+\mu}{\eta}}}\right)$, $\{w_i\}$ and $\{a_i\}$ are defined as in (4.35).

Proof: The proof is similar to that of Theorem 4.5, and is omitted. □

4.5 Numerical Results

In this section, we present some numerical results to examine the tightness of various capacity upper and lower bounds developed in the above sections.

For the Monte Carlo simulation method used in the section, as well as those in the remaining chapters, unless otherwise specified, it means that the simulation results are obtained by computing the desired function with repeated random sampling input which follows certain distribution function. Moreover, the desired number of samples and the distribution function are generally application dependent. For this particular case here, we generate 100,000 complex matrix $\mathbf{H}$ with the element of $\mathbf{H}$ follows the i.i.d. Nakagami-m distribution. For each sample, we compute the capacity according to Eq. (4.6) and the final simulation result is computed by averaging over 100,000 sample results. As for various parameters such as N_t, N_r, m and SNR range, they are randomly chosen according to two main principles: of practical interest and reasonable computation cost. However, in some cases, the parameters are carefully chosen to illustrate or verify certain properties.

Figure 4.1 plots three capacity upper bounds presented in Theorem 4.1 and Theorem 4.2 when $m = 0.5$ and $N_t = 3$, $N_r = 6$. As we can see, all the three bounds are quite tight at the low SNR regime. However, $\overline{\mathcal{C}}_2$ becomes loose when the SNR increases. Also, we observe that $\bar{\mathcal{C}}_1$ is the tightest upper bound, which agrees with the analytical result in Theorem 4.2.

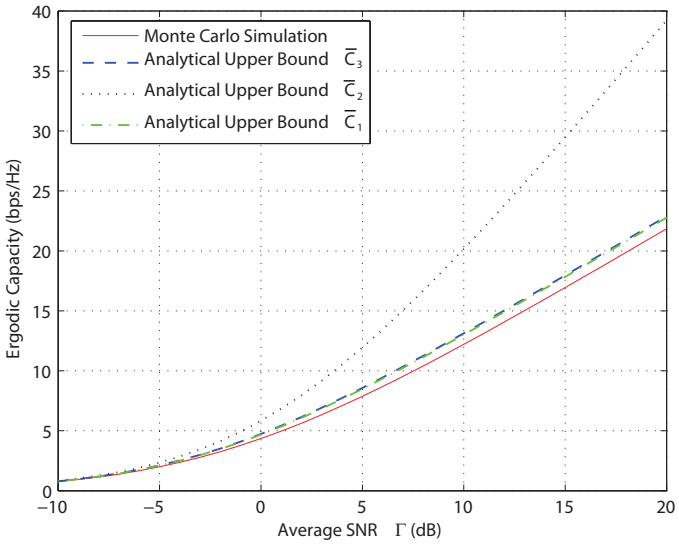

Figure 4.1. Analytical upper bound $\overline{\mathcal{C}}_1, \overline{\mathcal{C}}_2, \overline{\mathcal{C}}_3$ when $m = 0.5$, $N_t = 3$, $N_r = 6$.

Figure 4.2 plots capacity upper bound $\bar{\mathcal{C}}_1$ and the high SNR approximation, $\bar{\mathcal{C}}_{\mathsf{hsnr}}$ against Monte Carlo simulation results when $m = 0.5$, $N_t = 3$ and $N_r = 1, 3, 6, 12, 30$. As can be seen, the upper bound $\bar{\mathcal{C}}_1$ always overlaps with the exact capacity results when $s = 1$, which is expected because when $s = 1$, the overall MIMO channel reduces to a SIMO or MISO channel, and the upper bound (4.10) becomes exact. In other words, our results include those in [110] as a special case. In addition, it is observed that the upper bound is generally very tight, e.g., when $s = 3$ and $t = 6$ or $t = 12$, and almost overlaps with the exact results if $t = 30$, with the only exception occurs when $s = t$, or $N_t = N_r$.

Figure 4.3 examines impact of fading parameter m on the ergodic capacity of the system for $N_t = 2$ and

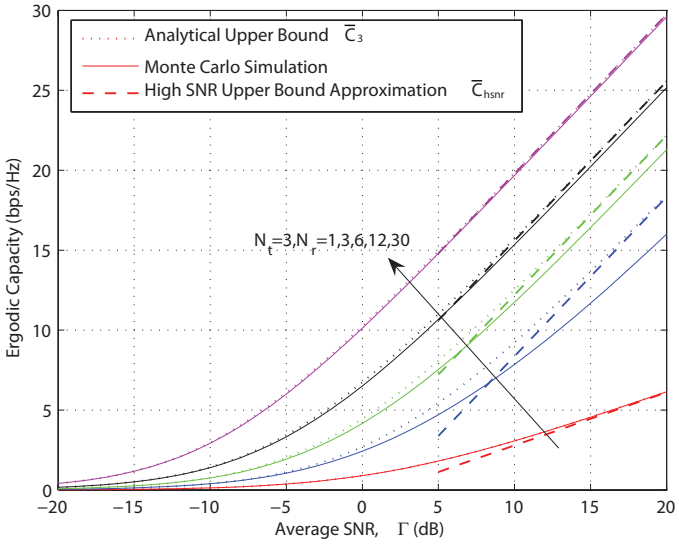

Figure 4.2. Ergodic capacity of Nakagami-m fading channel: Analytical upper bound versus simulation results when $m = 0.5$ (one-sided Gaussian distribution).

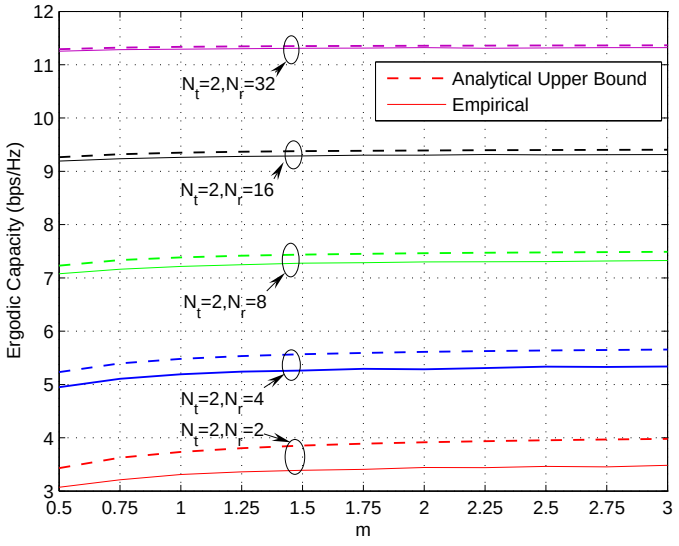

Figure 4.3. Ergodic capacity of Nakagami-m fading channel: Analytical upper bound versus simulation results with m and N_r as parameters.

$N_r = 2, 4, 8, 16, 32$ at average SNR of 5 (dB). We observe that, with the increase of N_r, the impact of channel fading parameter m on the ergodic capacity decreases gradually. For instance, when $N_r = 2$, the ergodic capacity increases considerably when m increases from 0.5 to 3. However, when $N_r = 30$, the difference is inappreciable.

Figure 4.4 shows the analytical lower bound curve $\underline{\mathcal{C}}_1$ for different antenna configurations. It can be observed that the lower bound is is reasonably tight in all case. In particularly, it performs very good at the extreme low SNR regime (i.e. < -15dB). In addition, the tightness of the lower bound degrades with the number of antennas, due to the significance of the higher order terms.

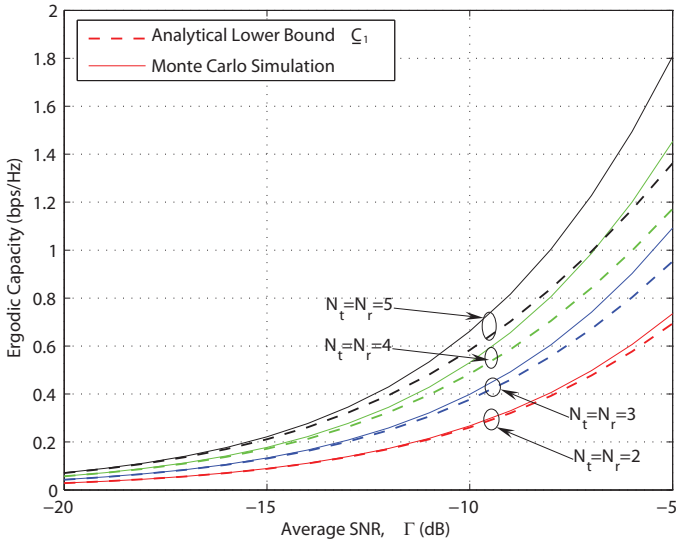

Figure 4.4. Ergodic capacity of Nakagami-m fading channel: Analytical lower bound versus simulation results when $m = 2$ for different antenna configurations.

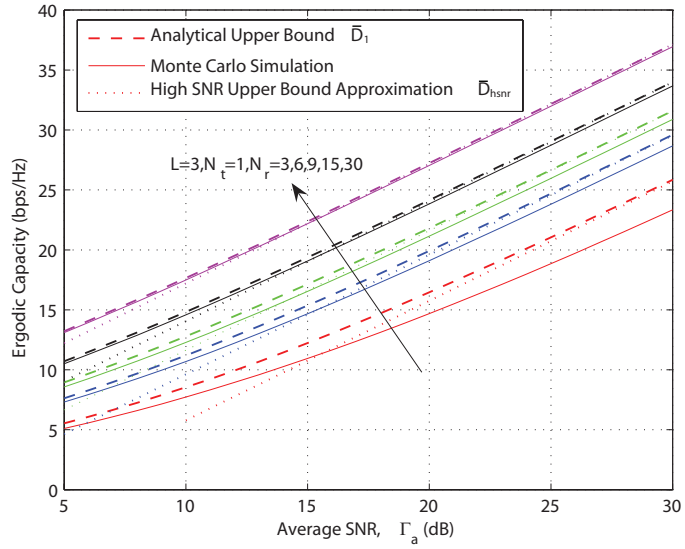

Figure 4.5. Ergodic capacity of composite Nakagami-m and log-normal fading channels with $\sigma = 8$ (dB), $D_1 = 1000$ (m), $D_2 = 1500$ (m), $D_3 = 2000$ (m), m=1/2, for different N_r.

Figure 4.5 plots the ergodic capacity of composite Nakagami-m and log-normal channels against the average received SNR per antenna, Γ_a[2]. Figure 4.6 plots the composite Rayleigh and log-normal case. In the simulation, we chose $\mu_i = 0$, $\sigma_i = \sigma$, for $i = 1, \ldots, L$ for simplicity. In both figures, we observe that the upper bound becomes tighter when N_r is greater which is consistent with the Nakagami-m only case.

Figure 4.7 investigates the performance of the lower bound. Similar to the lower bound of C-MIMO

[2] $\Gamma_a \triangleq \frac{PS}{N_0} e^{\frac{\sigma^2}{2\eta^2}}$, where S is the average path loss defined as $S = \frac{1}{L}\sum_{i=1}^{L} \frac{1}{D_i^v}$.

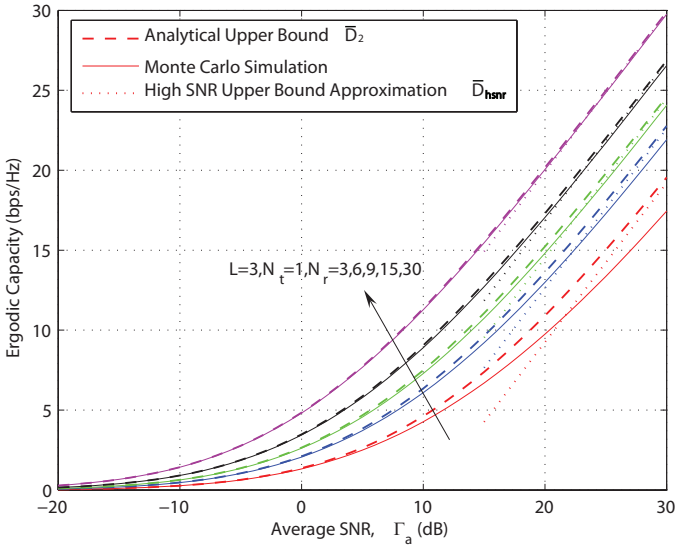

Figure 4.6. Ergodic capacity of composite Rayleigh and log-normal fading channels with $\sigma = 8$ (dB), $D_1 = 1000$ (m), $D_2 = 1500$ (m), $D_3 = 2000$ (m), for different N_r.

channels, results illustrate that the lower bound is tight at low SNRs while it becomes looser at high SNRs, particularly when the number of antennas increases.

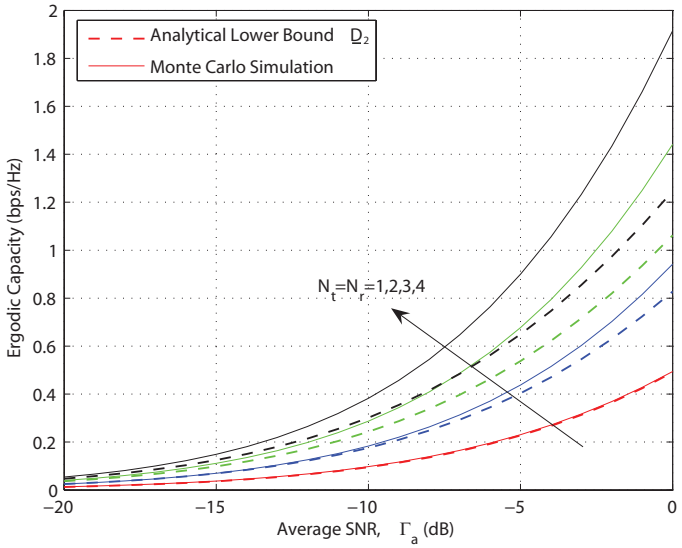

Figure 4.7. Ergodic capacity of composite Rayleigh and log-normal fading channels: Analytical lower bound versus simulation results when $L = 1$, $\sigma = 8$ (dB), $D = 1000$ (m), for different N_t, N_r.

4.6 Conclusion

In this chapter, by virtue of majorization theory, we have derived ergodic capacity bounds for C-MIMO and D-MIMO fading channels. In C-MIMO, the capacity for Nakagami-m fading channels was investigated in detail, where we derived several tight upper bounds in terms of Meijer G-function, and in the

high SNR regime, a simple closed-form upper bound was presented to gain insight on the impacts of the system parameters, such as fading severity m, the number of antennas N_t and N_r, etc. We also derived a tight capacity lower bound for the low SNR regime. The same capacity analysis was also performed for D-MIMO channels undergoing the composite (long-term) log-normal and (short-term) Nakagami fading where similar upper and lower bounds were derived.

Chapter 5

Mutual Information and Outage Analysis of Multi-Keyhole MIMO Channels

5.1 Introduction

The extraordinary gains of MIMO antenna systems have typically been demonstrated under the key assumption that the scattering is sufficiently rich to ensure that the MIMO channel matrix describing the channel gains between the transmitting and receiving antennas is of full-rank. However, it has also been shown that when the scattering environment is not-so-rich, the channel may exhibit reduced-rank behavior. In this case, the most commonly studied scenario is the single-keyhole (or pinhole) scenario, which describes the extreme scenario with the channel matrix having unit rank. This phenomenon has been validated theoretically [16] and experimentally [7, 8], and the performance of keyhole channels has been extensively studied for various settings [29, 51, 68, 69, 78, 83]. However, in practice the extremely rank-deficient behavior implied by the single-keyhole assumption may be too restrictive. This has motivated the so-called multi-keyhole channel model in [52–54], which generalizes and extends the applicability of the single-keyhole model. In fact, the multi-keyhole model provides a highly generalized channel description which embraces arbitrary rank behavior, and includes the conventional single-keyhole and rich scattering Rayleigh MIMO channel scenarios as special cases. The multi-keyhole channel is closely related to the double-scattering channel model proposed in [25], with the two models becoming equivalent when there is no correlation between the transmit and receive antennas.

In contrast to rich scattering MIMO channels, there are very few analytical results pertaining to the multi-keyhole channel model. In [53], it was revealed that the asymptotic instantaneous capacity of the multi-keyhole channel is described by summing the capacities of each individual keyhole. In [82], an upper bound for the capacity was presented. In [25], some approximations were provided for the p.d.f. of the eigenvalues of the channel correlation matrix, and these were used to study the performance of the multi-keyhole channel. In [104], the asymptotic diversity-multiplexing tradeoff (DMT) was considered, and [85] investigated the performance of orthogonal STBC (OSTBC) systems. Very recently, [37, 112]

studied the performance of MIMO multi-keyhole systems with CSIT, considering the special case of the multi-keyhole channel for which the power of each keyhole is unity.

In this chapter, we present a thorough investigation of multi-keyhole MIMO channels. Based on some newly derived statistical expressions for a product of complex random matrices presented in Chapter 3, we examine the mutual information and outage performance of MIMO systems operating over multi-keyhole channels. We consider two important scenarios. First, we derive new exact closed-form expressions and simplified upper and lower bounds for the ergodic mutual information, assuming that the transmitter has no access to the CSI but the receiver has perfect knowledge. We then present new performance results for optimal transmit beamforming scheme assuming both the transmitter and receiver have perfect CSI.

5.2 System Model

We consider a communication link with N_t transmit and N_r receive antennas operating in frequency non-selective channels. The received signals can be expressed in vector form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{5.1}$$

where $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$, and $\mathbf{x} = [x_1, x_2, \ldots, x_{N_t}]^T$ is the transmit symbol vector, with $\mathtt{E}_{\mathbf{x}}\{\|\mathbf{x}\|^2\} = P$. The matrix $\mathbf{H}$ represents the MIMO channel, which we model according to the multiple keyhole structure as follows [53, 54]

$$\mathbf{H} = \sum_{k=1}^{N_k} a_k \mathbf{h}_{r,k}\mathbf{h}_{t,k}^{\dagger} = \mathbf{H}_r\mathbf{A}\mathbf{H}_t^{\dagger}, \tag{5.2}$$

where $\mathbf{H}_r = [\mathbf{h}_{r,1}, \ldots, \mathbf{h}_{r,N_k}]$, $\mathbf{H}_t = [\mathbf{h}_{t,1}, \ldots, \mathbf{h}_{t,N_k}]$, and $\mathbf{A} = \mathsf{diag}(a_1, \ldots, a_{N_k})$, with N_k denoting the total number of independent keyholes, and a_k representing the complex gain for the kth keyhole. Moreover, $\mathbf{H}_r$ and $\mathbf{H}_t$ are mutually-independent matrices $\sim \mathcal{CN}_{N_r,N_k}(\mathbf{0}_{N_r\times N_k}, \mathbf{I}\otimes\mathbf{I})$ and $\sim \mathcal{CN}_{N_t,N_k}(\mathbf{0}_{N_t\times N_k}, \mathbf{I}\otimes\mathbf{I})$, respectively. Let $\mathbf{B} \triangleq \mathbf{A}\mathbf{A}^{\dagger}$, We assume that channel is normalized such that $\mathtt{E}\{\mathsf{trace}(\mathbf{H}\mathbf{H}^{\dagger})\} = N_rN_t$, and therefore $\mathsf{trace}(\mathbf{B}) = 1$.

For a general multi-keyhole MIMO channel, when the transmitter has no access to CSI while the receiver has perfect knowledge, the ergodic mutual information of a MIMO multi-keyhole channel is readily given by

$$\mathcal{I}(\gamma) = \mathtt{E}_{\mathbf{H}_r,\mathbf{H}_t}\left\{\log_2\det\left(\mathbf{I} + \frac{\gamma}{N_t}\mathbf{H}^{\dagger}\mathbf{H}\right)\right\}, \tag{5.3}$$

where $\gamma \triangleq \frac{P}{\sigma^2}$ is the SNR per transmit antenna.

For transmit beamforming system, the transmitted signal vector is $\mathbf{x} = \mathbf{w}_{\mathsf{opt}}x$, with x representing the information symbol, and $\mathbf{w}_{\mathsf{opt}}$ denoting the optimal transmit weight vector given by the dominant eigenvector of $\mathbf{H}^{\dagger}\mathbf{H}$ (i.e., the eigenvector corresponding to the maximum eigenvalue). At the receiver, the signals on each receive antenna are linearly combined according to the MRC principle using the

optimal receive weight vector $(\mathbf{H}\mathbf{w}_{\mathsf{opt}})^{\dagger}$ to give

$$z = \mathbf{w}_{\mathsf{opt}}{}^{\dagger}\mathbf{H}^{\dagger}\mathbf{H}\mathbf{w}_{\mathsf{opt}}x + \mathbf{w}_{\mathsf{opt}}{}^{\dagger}\mathbf{H}^{\dagger}\mathbf{n}\ . \tag{5.4}$$

This linear transmit-receive processing, commonly known as MIMO beamforming, maximizes the instantaneous SNR at the receiver, which is given by

$$\rho = \gamma \mathbf{w}_{\mathsf{opt}}^{\dagger}\mathbf{H}^{\dagger}\mathbf{H}\mathbf{w}_{\mathsf{opt}} = \gamma\lambda_{\max}, \tag{5.5}$$

where $\lambda_{\max}$ denotes the maximum eigenvalue of $\mathbf{H}^{\dagger}\mathbf{H}$.

5.3 Ergodic Mutual Information Analysis

In this section, we study the ergodic mutual information of a general multi-keyhole MIMO channel.

(5.3) can be alternatively expressed as

$$\mathcal{I}(\gamma) = \mathrm{E}_{\mathbf{H}_r,\mathbf{H}_t}\left\{\log_2\det\left(\mathbf{I} + \frac{\gamma}{N_t}\mathbf{H}_t^{\dagger}\mathbf{A}^{\dagger}\mathbf{H}_r^{\dagger}\mathbf{H}_r\mathbf{A}\mathbf{H}_t\right)\right\} \tag{5.6}$$

$$= p\mathrm{E}_{\lambda}\left\{\log_2\left(1 + \frac{\gamma}{N_t}\lambda\right)\right\}, \tag{5.7}$$

where λ is an unordered eigenvalue of random matrix $\mathbf{H}^{\dagger}\mathbf{H}$. Utilizing the p.d.f. expression presented in Theorem 3.3, we derive a new closed-form expression for the ergodic mutual information, as given by the following result.

Theorem 5.1 *When equal-power allocation is employed, the ergodic mutual information of a MIMO multi-keyhole channel is expressed as*

$$\mathcal{I}(\gamma) = \frac{\log_2 e}{\prod_{i<j}^{N_k}(b_j - b_i)}\sum_{i=1}^{N_k}\sum_{j=N_k-p+1}^{N_k}\frac{b_i^{j-1}\mathcal{D}_{i,j}G_{4,2}^{1,4}\left(\frac{\gamma b_i}{N_t}\middle|{}^{N_k+j-1-m,N_k+j-1-n,1,1}_{1,0}\right)}{\Gamma(n-N_k+j)\Gamma(m-N_k+j)}, \tag{5.8}$$

where $G_{m,n}^{p,q}(x)$ is the Meijer-G function [26], and and $\mathcal{D}_{i,j}$ is the (i,j)th cofactor of the matrix $\mathbf{\Xi}$ whose (l,k)th entry equals

$$[\mathbf{\Xi}]_{l,k} = b_l^{k-1},\ \textit{for}\ 1 \le l,k \le N_k\ . \tag{5.9}$$

Proof: The result can be proved by simply invoking Theorem 3.3 and performing a simple integral with the help of [26, (7.821.3)]. □

Whilst Theorem 5.1 presents an exact closed-form expression for the ergodic mutual information, for the extremely low SNR regime (e.g. $\gamma \le -15$dB) or when N_t grows large, the evaluation of the Meijer-G function can be computationally expensive. Hence, in those cases, it is better to use the bounds shown in the following theorems or some asymptotic results in [53].

Theorem 5.2 *The ergodic mutual information of a MIMO multi-keyhole channel is upper bounded by*

$$\mathcal{I}(\gamma) \leq \overline{\mathcal{I}}(\gamma) = \log_2 \left(\frac{\det(\mathbf{\Delta})}{\prod_{i<j}^{N_k} (b_j - b_i)} \right). \tag{5.10}$$

Proof: The result is obtained by applying Jensen's inequality and then invoking Theorem 3.8. □

To this end, we present the following lemma which will be used for simplifying the mutual information upper bound expression given in Theorem 5.2 for special case $n = 1$.

Lemma 5.1 *For distinct numbers $\{x_i\}$, $i = 1, \ldots, t$, let $\mathbf{V}_t$ denote the Vandermonde matrix of $\{x_i\}$ with entries*

$$[\mathbf{V}_t]_{l,k} = x_l^{k-1}, \text{ for } 1 \leq l, k \leq t. \tag{5.11}$$

Likewise, define the matrix $\mathbf{X}_{t,k}$, for $k = 0, 1, \ldots, t-2$, as

$$\mathbf{X}_{t,k} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^k & x_1^{k+2} & x_1^{k+3} & \cdots & x_1^t \\ 1 & x_2 & \cdots & x_2^k & x_2^{k+2} & x_2^{k+3} & \cdots & x_2^t \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_t & \cdots & x_t^k & x_t^{k+2} & x_t^{k+3} & \cdots & x_t^t \end{bmatrix}. \tag{5.12}$$

Then, the determinant of $\mathbf{X}_{t,k}$ can be computed as

$$\det(\mathbf{X}_{t,k}) = \det(\mathbf{V}_t) S_{t-1-k}(x_1, \ldots, x_t), \tag{5.13}$$

where $S_k(x_1, \cdots, x_t)$ is the elementary symmetric polynomial in t variables $x_1, \ldots, x_t$, defined by

$$S_k(x_1, \cdots, x_t) \triangleq \sum_{1 \leq i_1 < \cdots < i_k \leq t} x_{i_1} x_{i_2} \cdots x_{i_k}. \tag{5.14}$$

Proof: See Appendix C.1. □

Corollary 5.1 *When $n = 1$, the ergodic mutual information upper bound $\overline{\mathcal{I}}(\cdot)$ reduces to*

$$\overline{\mathcal{I}}(\gamma) = \log_2 (1 + \gamma m). \tag{5.15}$$

Proof: The result is obtained by using Lemma 5.1 and the assumption that $\sum_{i=1}^{N_k} b_i = 1$. □

Corollary 5.1 indicates that for SIMO or MISO multi-keyhole systems, the number of keyholes does not have a significant impact on the ergodic mutual information. This result is quite intuitive, since for SIMO or MISO, the multiplexing gain is limited to unity, regardless of the number of keyholes. This is in contrast to MIMO systems, in which case the number of keyholes may adversely effect the ergodic

mutual information. This effect is most significant when the number of keyholes is less than the number of transmit and receive antennas, in which case the number of keyholes limits the multiplexing gain of the system.

Theorem 5.3 *The ergodic mutual information of a MIMO multi-keyhole channel is lower bounded by*

$$\mathcal{I}(\gamma) \geq \underline{\mathcal{I}}(\gamma) = p \log_2 \left(1 + \frac{\gamma}{N_t} \mathsf{exp}\left(\frac{1}{p}\left[\sum_{s=1}^{p} \psi(m-p+s) + \sum_{s=1}^{p} \psi(n-p+s) + \frac{\sum_{s=N_k-p+1}^{N_k} \det(\mathbf{Y}_s)}{\prod_{i<j}^{N_k}(b_j - b_i)}\right]\right)\right), \tag{5.16}$$

where $\mathbf{Y}_s$ *has been defined in (3.63).*

Proof: Utilizing the result in [72, Theorem 1] and Theorem 3.10 yields the desired result. □

Corollary 5.2 *When* $p = 1$*, the ergodic mutual information lower bound* $\underline{\mathcal{I}}(\cdot)$ *reduces to*

$$\underline{\mathcal{I}}(\gamma) = \log_2 \left(1 + \frac{\gamma}{N_t} \mathsf{exp}\left(\psi(m) + \psi(n) + \frac{\det(\mathbf{Y}_{N_k})}{\prod_{i<j}^{N_k}(b_j - b_i)}\right)\right). \tag{5.17}$$

If $p = N_k = 1$*, then*

$$\underline{\mathcal{I}}(\gamma) = \log_2 \left(1 + \frac{\gamma}{N_t} \mathsf{exp}\left(\psi(m) + \psi(n)\right)\right). \tag{5.18}$$

Proof: The proof is straightforward and is omitted. □

5.4 Outage Analysis of MIMO-MRC system

In this section, we analyze the outage performance of MIMO-MRC system in multi-keyhole MIMO channels, which is defined as the probability that the received SNR drops below some predefined threshold, i.e.,

$$\mathrm{P}_{\mathrm{out}}(\gamma_{\mathsf{th}}) \triangleq \Pr(\rho < \gamma_{\mathsf{th}}). \tag{5.19}$$

From Equation (5.5), it is clear that the performance of this optimal MIMO beamforming system is determined by the statistics of $\lambda_{\max}$. Invoking Theorem 3.3, we obtain the following result

Theorem 5.4 *The outage probability of the optimal MIMO beamforming system in multi-keyhole channels can be expressed as*

$$\mathrm{P}_{\mathrm{out}}(\gamma_{\mathsf{th}}) = \frac{(-1)^{\frac{p(p-1)}{2}} \det\left(\mathbf{\Phi}\left(\frac{\gamma_{\mathsf{th}}}{\gamma}\right)\right)}{\prod_{i=1}^{p}\Gamma(n-i+1)\prod_{i<j}^{N_k}(b_j - b_i)}, \tag{5.20}$$

and $\mathbf{\Phi}(x)$ *is an* $N_k \times N_k$ *matrix whose* (l,k)*th entry is given by*

$$[\mathbf{\Phi}(x)]_{l,k} = \begin{cases} b_l^{k-1}, & k \leq N_k - p, \\ g(x)_{l,k}, & k > N_k - p, \end{cases} \tag{5.21}$$

with

$$g(x)_{l,k} = \Gamma(q-k+1) b_l^{2N_k-p-k} - b_l^{N_k-n-1} \sum_{t=0}^{m+q-n-k} \frac{x^t}{\Gamma(t+1)} 2(b_l x)^{\frac{q-t-k+1}{2}} K_{q-t-k+1}\left(2\sqrt{\frac{x}{b_l}}\right). \tag{5.22}$$

Proof: The proof follows directly from Thoerem 3.3. □

Theorem 5.4 is quite general and valid for arbitrary antenna and keyhole numbers. It only involves standard functions and can be efficiently evaluated by mathematical tools such as Matlab or Mathematica. However, the expression is too complicated to gain any insights. Therefore, we look into a special case, where simpler expressions can be obtained.

Corollary 5.3 *In the low outage regime, the outage probability of the optimal beamforming system in MISO/SIMO multi-keyhole channels can be approximated as*

$$\mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}}) \approx \begin{cases} \frac{\Gamma(m-N_k)}{\Gamma(m)\Gamma(N_k+1)\prod_{i=1}^{N_k} b_i} \left(\frac{\gamma_{\mathsf{th}}}{\gamma}\right)^{N_k}, & m > N_k, \\ \frac{1}{\Gamma(m)\Gamma(m+1)\prod_{i=1}^{N_k} b_i} \ln\left(\frac{\gamma}{\gamma_{\mathsf{th}}}\right) \left(\frac{\gamma_{\mathsf{th}}}{\gamma}\right)^{N_k}, & m = N_k, \\ \frac{(-1)^{m-1}}{\Gamma(m)\Gamma(m+1)} \frac{\det(\mathbf{\Phi}^4)}{\prod_{i<j}^{N_k}(b_j - b_i)} \left(\frac{\gamma_{\mathsf{th}}}{\gamma}\right)^{m}, & m < N_k. \end{cases} \tag{5.23}$$

Proof: This result is easily obtained by using Theorem 3.4, and noting that when $m = N_k$, the $\ln x$ term inside the brackets in (3.31) dominates the constant terms as $x \to 0$. □

Corollary 5.3 indicates that the diversity order of MISO/SIMO systems is given by $d = \min(m, N_k)$. Besides revealing the diversity order of MISO/SIMO multi-keyhole systems, Corollary 5.3 also shows that the power distribution of the keyholes (the b_i's) affects the array gain of the system. To gain insights into this effect, let us first consider the case $m \geq N_k$. To this end, it is convenient to apply tools from majorization theory [62], which leads to the following result.

Corollary 5.4 *When* $m \geq N_k$, $\mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}})$ *is a Schur-convex function with respect to* b_i, $i = 1, \cdots, N_k$.

Proof: The proof follows from the fact that $\prod_{i=1}^{N_k} b_i$ is a Schur-concave function [73]. □

The alternative case $m < N_k$ is more difficult to analyze due to the determinant expression involving the b_i's in (5.23). However, for the special case when $N_k = 2$ (or $m = 1$, i.e., SISO multi-keyhole channel), we have the following result.

Table 5.1. Power distribution among keyholes.

N_k	$(b_1, \ldots, b_{N_k})$
1	(1)
2	(0.4, 0.6)
3	(0.2, 0.3, 0.5)
4	(0.1, 0.2, 0.3, 0.4)
5	(0.05, 0.1, 0.2, 0.3, 0.35)
6	(0.05, 0.1, 0.12, 0.13, 0.25, 0.35)

Corollary 5.5 *When $N_k = 2$ and $m = 1$, $\mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}})$ is a Schur-convex function with respect to b_i, for $i = 1, \ldots, N_k$.*

Proof: See Appendix C.2. □

Both Corollary 5.4 and Corollary 5.5 indicate that as the power distribution amongst the keyholes (i.e., the b_i's) becomes "less spread", the outage probability improves.

5.5 Numerical Results

In this section, we provide some numerical results to verify the analytical expressions derived in the above sections. In all simulations, the power distribution among keyholes is given in Table 5.1, and the simulation results are obtained based on 100,000 independent channel realizations.

Figure 5.1 plots the exact curve according to (5.8) against Monte-Carlo simulation curves for $N_t = 5$, $N_r = 3$ with different N_k. As observed from Figure 5.1, the analytical curves match perfectly with the Monte-Carlo simulation curves for all cases, which confirms the correctness of the analytical results. It is also observed that the mutual information of multi-keyhole channel is always inferior to that of a standard MIMO channel with same N_t and N_r.

Figure 5.2 compares the ergodic mutual information upper (5.10) and lower bound (5.16) with the exact results in (5.8). We see that both bounds are very tight. Moreover, in the low SNR regime the upper bound and exact results coincide, whilst in the high SNR regime the lower bound and exact results coincide.

Figure 5.3 illustrates the outage probability for the optimal MIMO beamforming system in multi-keyhole channels. We see that when the number of keyholes increases, the performance approaches that of a Rayleigh MIMO channel. Surprisingly, we also observe that there is a crossover point, indicating that at high outage levels (equivalently, at sufficiently low SNR), the performance of a multi-keyhole channel can be superior than that of a Rayleigh MIMO channel. However, despite this cross-over point, for outage levels of practical interest (eg. < 0.1), achieving a given outage level requires lower SNR for a Rayleigh MIMO channel compared with multi-keyhole channels.

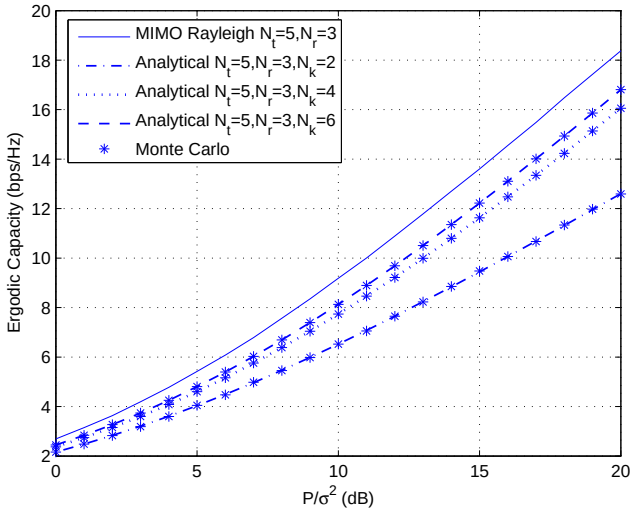

Figure 5.1. Ergodic mutual information of multi-keyhole MIMO channels with equal-power allocation.

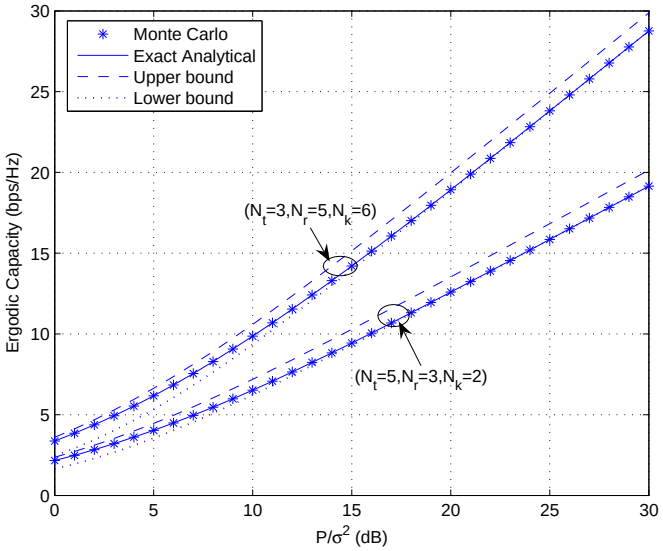

Figure 5.2. Ergodic mutual information bounds for MIMO multi-keyhole channels.

Figure 5.4 compares the exact outage probability curves based on Theorem 5.4, low outage approximation curves based on Corollary 5.3, and Monte-Carlo simulated curves. Results are shown for a system with $N_t = 4$, $N_r = 1$, and different numbers of keyholes. It can be observed that the approximation curves are very accurate in the low outage regime. We also see that the slope of the outage curves are determined by the minimum of N_k and m; thereby confirming our diversity analysis.

Figure 5.5 illustrates the impact of the power distribution on the outage curves when $N_t = 4$, $N_r = 1$, and $N_k = 2$. Three curves are plotted according to the power distributions (b_1, b_2) corresponding to $(0.05, 0.95)$, $(0.15, 0.85)$ and $(0.45, 0.55)$, respectively. Note that

$$(0.05, 0.95) \succ (0.15, 0.85) \succ (0.45, 0.55), \tag{5.24}$$

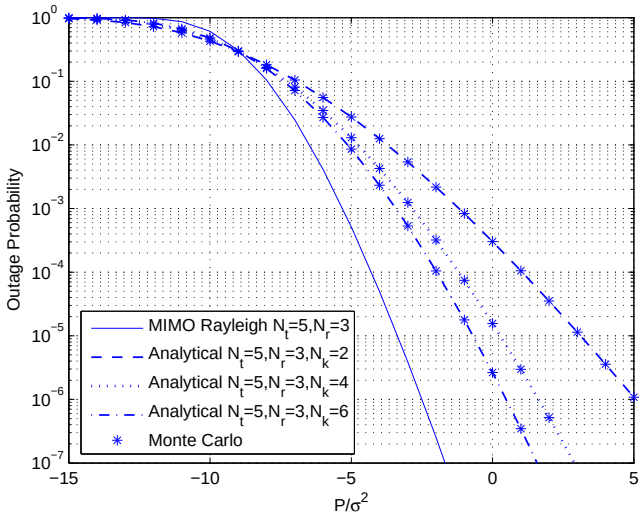

Figure 5.3. The outage probability of the optimal MIMO beamforming system in multi-keyhole channels.

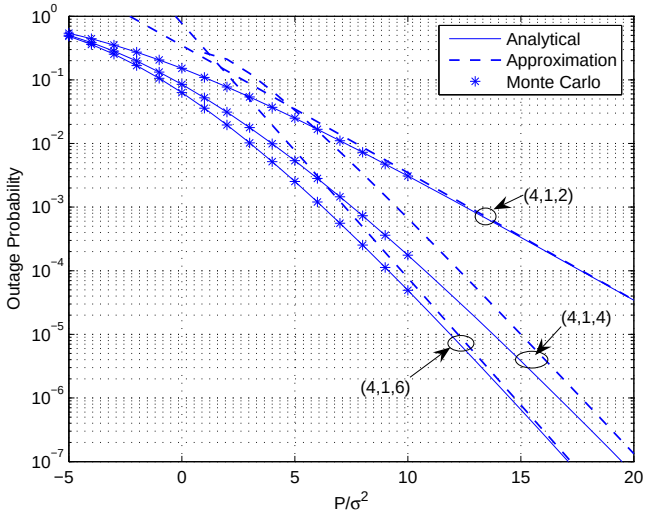

Figure 5.4. The outage probability of the optimal beamforming system in MISO multi-keyhole channels for different number of keyholes with $N_t = 4$, $N_r = 1$.

which, from Corollary 5.4, implies that

$$\mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}})_{(0.05,0.95)} \geq \mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}})_{(0.15,0.85)} \geq \mathrm{P}_{\mathrm{out}}^{n=1}(\gamma_{\mathsf{th}})_{(0.45,0.55)} \ . \tag{5.25}$$

The outage curves in Figure 5.5 confirm this analysis.

5.6 Conclusion

Multi-keyhole MIMO channels bridge the gap between single-keyhole and rich-scattering (full-rank) MIMO channels. In this chapter, we have provided an analytical characterization of the statistical prop-

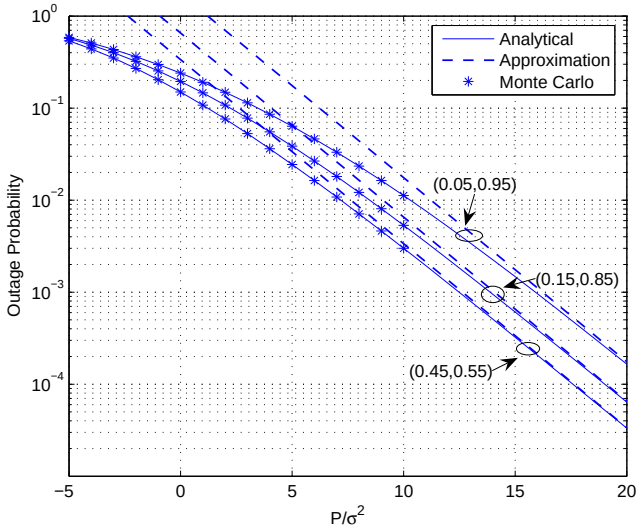

Figure 5.5. The outage probability of the optimal beamforming system in MISO multi-keyhole channels with different power distributions with $N_t = 4$, $N_r = 1$, $N_k = 2$.

erties of multi-keyhole MIMO channel matrices. In particular, we derived exact expressions for the p.d.f. of an unordered eigenvalue, exact and asymptotic expressions for the distribution of the maximum eigenvalue, as well as closed-form expressions for the expected log-determinant and expected characteristic polynomial. These results were applied to investigate the ergodic MIMO mutual information, and the outage probability of optimal transmit beamforming in multi-keyhole MIMO channels. The findings suggest that performance in multi-keyhole MIMO channels is generally inferior to that of rich-scattering MIMO channels.

Chapter 6

Capacity of AF MIMO Dual-Hop Systems

6.1 Introduction

Point-to-point MIMO communication systems have been receiving considerable attention in the last decade due to their potential for providing linear capacity growth and significant performance improvements over conventional SISO systems [23, 94]. Recently, the application of MIMO techniques in a cooperative communication setting [49, 50, 79, 80] has become a topic of increasing interest as a means of achieving further performance improvements in wireless networks [12, 13, 70, 99, 101].

A great deal of research works has been conducted to gain fundamental understanding of the capacity of this class of systems [11, 65, 66, 97, 98, 105]. In [11], the ergodic capacity of AF MIMO dual-hop systems was examined for a large numbers of relay antennas K, and was shown to scale with $\log K$. Asymptotic ergodic capacity results were also obtained in [97] by means of the replica method from statistical physics. In [65, 66], the asymptotic network capacity was examined as the number of source/desination antennas M and relay antennas K grew large with a fixed-ratio $K/M \rightarrow \beta$ using tools from large-dimensional RMT. It was demonstrated that for $\beta \rightarrow \infty$, the relay network behaved equivalently to a point-to-point MIMO link. The results of [65, 66] were further elaborated in [105] where a general asymptotic ergodic capacity formula was presented for multi-level AF relay networks. Recently, the asymptotic mean and variance of the mutual information in correlated Rayleigh fading was studied in [98].

All of these prior capacity results, however, were derived by employing *asymptotic methods* (i.e., by letting the system dimensions grow to infinity). There appear to be no analytical ergodic capacity results which apply for AF MIMO dual hop systems with arbitrary finite antenna and relaying configurations. In this chapter we derive new exact analytical results, simple closed-form high SNR expressions, and tight closed-form upper and lower bounds on the ergodic capacity of AF MIMO dual-hop systems. In contrast to previous results, our expressions apply for any finite number of MIMO antennas and for arbitrary number of relay antennas.

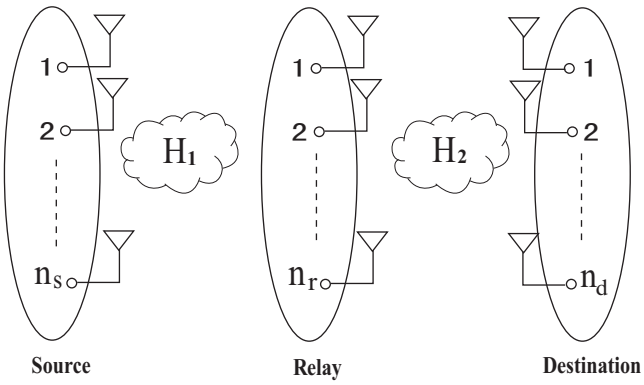

Figure 6.1. Schematic diagram of a MIMO dual-hop system, where there is no direct link between source and destination.

6.2 System Model

We employ the same AF MIMO dual-hop system model as in [65, 66]. In particular, suppose that there are n_s source antennas, n_r relay antennas and n_d destination antennas, which we represent by the 3-tuple (n_s, n_r, n_d). All terminals operate in half-duplex mode, and as such communication occurs from source to relay and from relay to destination in two separate time slots. It is assumed that there is no direct communication link between the source and destination, as sketched in Figure 6.1. The end-to-end input-output relation of this channel is then given by

$$\mathbf{y} = \mathbf{H}_2\mathbf{F}\mathbf{H}_1\mathbf{s} + \mathbf{H}_2\mathbf{F}\mathbf{n}_{n_r} + \mathbf{n}_{n_d} \tag{6.1}$$

where $\mathbf{s}$ is the transmit symbol vector, $\mathbf{n}_{n_r}$ and $\mathbf{n}_{n_d}$ are the relay and destination noise vectors respectively, $\mathbf{F} = \sqrt{\alpha/\left(n_r\left(1+\rho\right)\right)}\mathbf{I}_{n_r}$ (α corresponds to the overall power gain of the relay terminal) is the forwarding matrix at the relay terminal which simply forwards scaled versions of its received signals, and $\mathbf{H}_1 \sim \mathcal{CN}_{n_r,n_s}(\mathbf{0}_{n_r\times n_s}, \mathbf{I}\otimes\mathbf{I})$ and $\mathbf{H}_2 \sim \mathcal{CN}_{n_d,n_r}(\mathbf{0}_{n_d\times n_r}, \mathbf{I}\otimes\mathbf{I})$ denote the channel matrices of the first hop and the second hop respectively. The input symbols are chosen to be i.i.d. zero-mean circulant symmetric complex Gaussian (ZMCSCG) random variables and the per antenna power is assumed to be ρ/n_s, i.e., $E\left\{\mathbf{s}\mathbf{s}^\dagger\right\} = \left(\rho/n_s\right)\mathbf{I}_{n_s}$. The additive noise at the relay and destination are assumed to be white in both space and time and are modeled as ZMCSCG with unit variance, i.e., $E\left\{\mathbf{n}_{n_r}\mathbf{n}_{n_r}^\dagger\right\} = \mathbf{I}_{n_r}$ and $E\left\{\mathbf{n}_{n_d}\mathbf{n}_{n_d}^\dagger\right\} = \mathbf{I}_{n_d}$. We assume that the source and relay have no CSI, and that the destination has perfect knowledge of both $\mathbf{H}_2$ and $\mathbf{H}_2\mathbf{H}_1$.

The ergodic capacity (in b/s/Hz) of the AF MIMO dual-hop system described above can be written as [65, 66, 97]

$$C = \frac{1}{2}E\left\{\log_2\det\left(\mathbf{I}+\mathbf{R}_s\mathbf{R}_n^{-1}\right)\right\} \tag{6.2}$$

where $\mathbf{R}_s$ and $\mathbf{R}_n$ are $n_d \times n_d$ matrices given by

$$\mathbf{R}_s = \frac{\rho a}{n_s}\mathbf{H}_2\mathbf{H}_1\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger \tag{6.3}$$

and

$$\mathbf{R}_n = \mathbf{I}_{n_d} + a\mathbf{H}_2\mathbf{H}_2^\dagger \tag{6.4}$$

respectively, with

$$a = \frac{\alpha}{n_r\left(1+\rho\right)}. \tag{6.5}$$

Using the identity

$$\det\left(\mathbf{I}+\mathbf{AB}\right) = \det\left(\mathbf{I}+\mathbf{BA}\right), \tag{6.6}$$

(6.2) can be alternatively expressed as

$$C\left(\rho\right) = \frac{1}{2}E\left\{\log_2\det\left(\mathbf{I}_{n_s} + \frac{\rho a}{n_s}\mathbf{H}_1^\dagger\mathbf{H}_2^\dagger\mathbf{R}_n^{-1}\mathbf{H}_2\mathbf{H}_1\right)\right\}. \tag{6.7}$$

Next, we utilize the singular value decomposition to write $\mathbf{H}_2 = \mathbf{U}_2\mathbf{D}_2\mathbf{V}_2^\dagger$, where

$$\mathbf{D}_2 = \mathrm{diag}\left\{\lambda_1,\ldots,\lambda_{\min(n_d,n_r)}\right\} \tag{6.8}$$

is an $n_d \times n_r$ diagonal matrix, with diagonal elements pertaining to the increasing ordered singular values, and $\mathbf{U}_2 \in \mathcal{C}^{n_d\times n_d}$ and $\mathbf{V}_2 \in \mathcal{C}^{n_r\times n_r}$ are unitary matrices containing the respective eigenvectors. Since $\mathbf{H}_1$ is invariant under left and right unitary transformation, the ergodic capacity in (6.7) can be further simplified as

$$C\left(\rho\right) = \frac{1}{2}E\left\{\log_2\det\left(\mathbf{I}_{n_r} + \frac{\rho a}{n_s}\mathbf{H}_1^\dagger\mathbf{\Psi}\mathbf{H}_1\right)\right\}, \tag{6.9}$$

where

$$\mathbf{\Psi} = \begin{cases} \mathrm{diag}\left\{\frac{\lambda_1^2}{1+a\lambda_1^2},\ldots,\frac{\lambda_{n_r}^2}{1+a\lambda_{n_r}^2}\right\}, & n_r \le n_d, \\ \mathrm{diag}\left\{\frac{\lambda_1^2}{1+a\lambda_1^2},\ldots,\frac{\lambda_{n_d}^2}{1+a\lambda_{n_d}^2},\underbrace{0,\ldots,0}_{n_r-n_d}\right\}, & n_r > n_d. \end{cases} \tag{6.10}$$

It is then easily established that

$$C\left(\rho\right) = \frac{1}{2}E\left\{\log_2\det\left(\mathbf{I}_{n_s} + \frac{\rho a}{n_s}\tilde{\mathbf{H}}_1^\dagger\mathbf{L}\tilde{\mathbf{H}}_1\right)\right\}, \tag{6.11}$$

where $\tilde{\mathbf{H}}_1^{\dagger} \sim \mathcal{CN}_{n_s,q}\left(\mathbf{0}, \mathbf{I}_{n_s} \otimes \mathbf{I}_q\right)$, with $q = \min\left(n_d, n_r\right)$, and

$$\mathbf{L} = \operatorname{diag}\left\{\lambda_i^2 / \left(1 + a\lambda_i^2\right)\right\}_{i=1}^{q}. \tag{6.12}$$

Equivalently, we can now write

$$C\left(\rho\right) = \frac{s}{2}\int_0^{\infty} \log_2\left(1 + \frac{\rho a}{n_s}\lambda\right) f_{\lambda}\left(\lambda\right) d\lambda, \tag{6.13}$$

where $s = \min\left(n_s, q\right)$, λ denotes an unordered eigenvalue of the random matrix $\tilde{\mathbf{H}}_1^{\dagger}\mathbf{L}\tilde{\mathbf{H}}_1$, and $f_{\lambda}\left(\cdot\right)$ denotes the corresponding p.d.f..

6.3 Exact Ergodic Capacity Analysis

In this section, we present new analytical expressions for the ergodic capacity of AF MIMO dual-hop systems.

Theorem 6.1 *The exact ergodic capacity of AF MIMO dual-hop systems can be expressed as*

$$C\left(\rho\right) = \mathcal{K}\sum_{l=1}^{q}\sum_{k=q-s+1}^{q}\sum_{i=0}^{q+n_s-l}\frac{\binom{q+n_s-l}{i}a^{q+n_s-l-i}}{\Gamma\left(n_s-q+k\right)}G_{l,k}\mathcal{J}_{i,k}, \tag{6.14}$$

where

$$\mathcal{K} = \left(\prod_{i=1}^{q}\Gamma\left(q-i+1\right)\Gamma\left(p-i+1\right)\right)^{-1}, \tag{6.15}$$

and

$$\mathcal{J}_{i,k} = \int_0^{\infty}\log_2\left(1+\frac{\rho a}{n_s}\lambda\right)e^{-\lambda a}\lambda^{(2n_s+2k+p-q-i-3)/2}K_{p+q-i-1}\left(2\sqrt{\lambda}\right)d\lambda\,. \tag{6.16}$$

Proof: The above result can be obtained by substituting the p.d.f. expression of an unordered eigenvalue derived in Theorem 3.1 in (6.13). □

The integral in (6.16) can be evaluated either numerically, or can be expressed as an infinite series involving Meijer-G functions. To this end, we examine the ergodic capacity relationship of AF MIMO dual-hop systems and single-hop MIMO systems in the following subsection.

6.3.1 Analogies with Single-Hop MIMO Ergodic Capacity

Let $C^{\mathrm{SH-MIMO}}(n_s, n_d, \rho)$ denote the ergodic capacity of a conventional single-hop i.i.d. Rayleigh fading MIMO channel matrix $\mathbf{H} \sim \mathcal{CN}_{n_d,n_s}(\mathbf{0}_{n_d\times n_s}, \mathbf{I}\otimes\mathbf{I})$, with n_s transmit and n_d receive antennas, and

average SNR ρ, i.e..

$$C^{\text{SH}-\text{MIMO}}(n_s, n_d, \rho) = E\left\{\log_2 \det\left(\mathbf{I}_{n_d} + \frac{\rho}{n_s}\mathbf{H}\mathbf{H}^\dagger\right)\right\}. \tag{6.17}$$

Here, we demonstrate four particular cases for which the AF MIMO dual-hop channel relates directly to single-hop MIMO channels, in terms of ergodic capacity.

- As the number of relay antennas grows large, i.e., $n_r \to \infty$, the ergodic capacity of AF MIMO dual-hop systems becomes

$$\lim_{n_r\to\infty} C(\rho) = \frac{1}{2}C^{\text{SH}-\text{MIMO}}\left(n_s, n_d, \frac{\rho\alpha}{1+\rho+\alpha}\right). \tag{6.18}$$

  A proof is presented in Appendix D.1. Note that a similar phenomenon has been derived in [11], for the special case $n_s = n_d$. Here, (6.18) generalizes that result for arbitrary source and destination antenna configurations.

- As the number of source antennas grows large, i.e., $n_s \to \infty$, the ergodic capacity of AF MIMO dual-hop systems becomes

$$\lim_{n_s\to\infty} C(\rho) = \frac{1}{2}C^{\text{SH}-\text{MIMO}}(n_r, n_d, \alpha) - \frac{1}{2}C^{\text{SH}-\text{MIMO}}\left(n_r, n_d, \frac{\alpha}{1+\rho}\right). \tag{6.19}$$

  A proof is presented in Appendix D.2. Interestingly, we see that as ρ grows large, the right-most term in (6.19) disappears, and the AF MIMO dual-hop capacity becomes equivalent to one half of the ergodic capacity of a single-hop MIMO channel with n_r transmit antennas, n_d receive antennas, and average SNR α.

- As the number of destination antennas grows large, i.e., $n_d \to \infty$, the ergodic capacity of AF MIMO dual-hop systems becomes

$$\lim_{n_d\to\infty} C(\rho) = \frac{1}{2}C^{\text{SH}-\text{MIMO}}(n_s, n_r, \rho). \tag{6.20}$$

  The result is trivially obtained by directly taking $\lambda_i^2 \to \infty$ in (6.11). We see that the AF MIMO dual-hop capacity becomes equivalent to one half of the ergodic capacity of a single-hop MIMO channel with n_s transmit antennas, n_r receive antennas, and average SNR ρ.

- As the power gain of the relay grows large, i.e., $\alpha \to \infty$, the ergodic capacity of AF MIMO dual-hop systems becomes

$$\lim_{\alpha\to\infty} C(\rho) = \frac{1}{2}C^{\text{SH}-\text{MIMO}}(n_s, q, \rho). \tag{6.21}$$

  The result is trivially obtained by directly taking $\alpha \to \infty$ in (6.11). Thus we see the interesting result that even as the relay power gain becomes very large, the capacity of AF MIMO dual-hop

channels remains bounded, and in fact becomes equivalent to one half of the ergodic capacity of a single-hop MIMO channel with n_s transmit antennas, $q = \min(n_r, n_d)$ receive antennas, and average SNR ρ.

We note that for each of the cases (6.18)–(6.21), closed-form expressions can be obtained by directly invoking known results from the single-hop MIMO capacity literature (eg. see [82]).

In order to obtain further simplified closed-form results, it is useful to investigate the ergodic capacity in the high SNR regime. This is presented in the subsection below.

6.3.2 High SNR Capacity Analysis

For the high SNR regime, we consider one important scenario where the source and relay powers grow large proportionately.

Here we have $\alpha \to \infty$, $\rho \to \infty$, with $\alpha/\rho = \beta$, for some fixed β. Then $\rho a \to \frac{\alpha}{n_r}$ and $a \to \beta/n_r$, and the ergodic capacity at high SNR reduces to

$$C(\rho)|_{\alpha,\rho\to\infty,\alpha/\rho=\beta} = \frac{1}{2} E\left\{\log_2 \det\left(\mathbf{I}_{n_s} + \frac{\rho\beta}{n_s n_r}\tilde{\mathbf{H}}_1^{\dagger}\bar{\mathbf{L}}\tilde{\mathbf{H}}_1\right)\right\}, \tag{6.22}$$

where $\bar{\mathbf{L}} = \operatorname{diag}\left\{\lambda_i^2/\left(1+(\beta/n_r)\lambda_i^2\right)\right\}_{i=1}^{q}$. We can express (6.22) in the general form [57]

$$C(\rho)|_{\alpha,\rho\to\infty,,\alpha/\rho=\beta} = S_\infty\left(\frac{\rho|_{\mathrm{dB}}}{3\mathrm{dB}} - \mathcal{L}_\infty\right) + o(1), \tag{6.23}$$

where $3\mathrm{dB} = 10\log_{10}(2)$. Here, the two key parameters are S_∞, which denotes the high-SNR slope in bits/s/Hz/(3 dB) given by

$$S_\infty = \lim_{\alpha,\rho\to\infty} \frac{C(\rho)|_{\alpha,\rho\to\infty,\alpha/\rho=\beta}}{\log_2(\rho)} \tag{6.24}$$

and $\mathcal{L}_\infty$, which represents the high-SNR power offset in 3 dB units given by

$$\mathcal{L}_\infty = \lim_{\alpha,\rho\to\infty}\left(\log_2(\rho) - \frac{C(\rho)|_{\alpha,\rho\to\infty,\alpha/\rho=\beta}}{S_\infty}\right). \tag{6.25}$$

From (6.22), we can evaluate $\mathcal{S}_\infty$ and $\mathcal{L}_\infty$ in closed-form as follows.

Theorem 6.2 *For the case $\alpha \to \infty$, $\rho \to \infty$, with $\alpha/\rho = \beta$, the high-SNR slope and high-SNR power offset of AF MIMO dual-hop systems are given by*

$$S_\infty = \frac{s}{2} \text{ bit/s/Hz/(3dB)} \tag{6.26}$$

and[1]

$$\mathcal{L}_\infty(n_s,n_r,n_d)=\log_2\left(\frac{n_sn_r}{\beta}\right)-\frac{1}{s\ln 2}\left[\sum_{k=1}^{s}\psi\left(n_s+k-s\right)+\mathcal{K}\sum_{k=q-s+1}^{q}\det\left(\bar{\mathbf{W}}_k\right)\right] \tag{6.27}$$

respectively, where $\bar{\mathbf{W}}_k$ *is a* $q\times q$ *matrix with entries*

$$\left\{\bar{\mathbf{W}}_k\right\}_{m,n}=\begin{cases}\left(\frac{\beta}{n_r}\right)^{1-\tau}\vartheta_{\tau-1}\left(\frac{\beta}{n_r}\right), & n\neq k,\\ \varsigma_{m+n}\left(\frac{\beta}{n_r}\right), & n=k.\end{cases} \tag{6.28}$$

For the case $q=s$ *(i.e. corresponding to* $\min(n_s,n_r,n_d)=n_d$ *or* $\min(n_s,n_r,n_d)=n_r$*), the high SNR power offset (6.27) admits the alternative form*

$$\mathcal{L}_\infty(n_s,n_r,n_d)=\log_2\left(\frac{n_sn_r}{\beta}\right)-\frac{1}{s\ln 2}\left[\sum_{k=1}^{s}\psi\left(n_s-s+k\right)+\sum_{i=0}^{q-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\sum_{k=0}^{2q-l-2}\binom{2q-l-2}{k}\right.$$
$$\left.\times\,\mathcal{A}\left(i,j,l,p,q\right)\left(\frac{\beta}{n_r}\right)^{2q-l-2-k}\Gamma\left(p+q-k-1\right)\left(\psi\left(p+q-k-1\right)-\sum_{m=0}^{p+q-k-2}g_m\left(\frac{n_r}{\beta}\right)\right)\right]. \tag{6.29}$$

Proof: See Appendix D.3. □

Interestingly, we see that the high SNR slope depends only on the minimum system dimension, i.e. $s=\min(n_s,n_r,n_d)$, whereas the high SNR power offset is a much more intricate function of n_s, n_r, and n_d.

It is important to note that Theorem 6.2 presents an exact characterization of the key high SNR ergodic capacity parameters, $\mathcal{S}_\infty$ and $\mathcal{L}_\infty(\cdot)$, for arbitrary numbers of antennas at the source, relay, and destination terminals. We now examine some particular cases of Theorem 6.2, in which these expressions reduce to simple forms.

Corollary 6.1 *Let* $n_r=1$*. Then* $\mathcal{S}_\infty=1/2$*, and* $\mathcal{L}_\infty(\cdot)$ *reduces to*

$$\mathcal{L}_\infty(n_s,1,n_d)=\log_2\left(\frac{n_s}{\beta}\right)-\frac{1}{\ln 2}\left[\psi\left(n_s\right)+\psi\left(n_d\right)-\sum_{m=0}^{n_d-1}g_m\left(\frac{1}{\beta}\right)\right]. \tag{6.30}$$

Note that, as n_s grows large, $\psi\left(n_s\right)=\ln n_s+o(1)$ [4, (6.3.18)], where the $o(1)$ term disappears as $n_s\to\infty$, and as such we have

$$\lim_{n_s\to\infty}\mathcal{L}_\infty(n_s,1,n_d)=\log_2\left(\frac{1}{\beta}\right)-\frac{1}{\ln 2}\left[\psi\left(n_d\right)-\sum_{m=0}^{n_d-1}g_m\left(\frac{1}{\beta}\right)\right]. \tag{6.31}$$

[1] Note that here we explicitly indicate the dependence of the high SNR power offset on n_s, n_r, and n_d.

Corollary 6.2 *Let $n_d = 1$. Then $\mathcal{S}_\infty = 1/2$, and $\mathcal{L}_\infty(\cdot)$ reduces to*

$$\mathcal{L}_\infty(n_s, n_r, 1) = \log_2\left(\frac{n_s n_r}{\beta}\right) - \frac{1}{\ln 2}\left[\psi(n_s) + \psi(n_r) - \sum_{m=0}^{n_r-1} g_m\left(\frac{n_r}{\beta}\right)\right]. \tag{6.32}$$

In this case, as n_s grows large we have

$$\lim_{n_s\to\infty} \mathcal{L}_\infty(n_s, n_r, 1) = \log_2\left(\frac{n_r}{\beta}\right) - \frac{1}{\ln 2}\left[\psi(n_r) - \sum_{m=0}^{n_r-1} g_m\left(\frac{n_r}{\beta}\right)\right]. \tag{6.33}$$

Based on these results, we can easily examine the effect of the relative power gain factor β on the ergodic capacity. In particular, noting that $g_l(x)$ in (3.59) is a monotonically decreasing function of x in the interval[2] $[0, \infty)$, we see that increasing β, whilst having no effect on the high SNR capacity slope $\mathcal{S}_\infty$, results in decreasing the high SNR power offset $\mathcal{L}_\infty(\cdot)$, and therefore increasing the ergodic capacity in the high SNR regime.

Corollary 6.3 *Let $n_s = n_r = 1$. Adding k destination antennas, while not altering S_∞, would reduce the high SNR power offset as*

$$\begin{aligned}\delta(n_d, k) &\triangleq \mathcal{L}_\infty(1, 1, n_d + k) - \mathcal{L}_\infty(1, 1, n_d) \\ &= -\frac{1}{\ln 2}\sum_{l=n_d}^{n_d+k-1}\left(\frac{1}{\ell} + g_l\left(\frac{1}{\beta}\right)\right).\end{aligned} \tag{6.34}$$

Note that, to obtain this result, we have invoked the definition of the digamma function [26]. Since $g_l(x) > 0$ for $x \in [0, \infty)$, it is clear that the high SNR power offset $\mathcal{L}_\infty(\cdot)$ in (6.34) is a decreasing function of k, thereby confirming the intuitive notion that adding more antennas to the destination terminal has the effect of improving the ergodic capacity.

6.4 Ergodic Capacity Upper Bound

The following theorem presents a new tight upper bound on the ergodic capacity of AF MIMO dual-hop systems.

Theorem 6.3 *The ergodic capacity of AF MIMO dual-hop systems is upper bounded by*

$$C(\rho) \le C_U(\rho) = \frac{1}{2}\log_2\left(\mathcal{K}\det(\bar{\boldsymbol{\Xi}})\right), \tag{6.35}$$

where $\bar{\boldsymbol{\Xi}}$ is defined in (3.47).

[2]This conclusion is easily established by noting that $\mathrm{d}/\mathrm{d}x\left(g_l(x)\right) = e^x\left[E_{l+1}(x) - E_l(x)\right]$, and using [4, Eq. 5.1.17].

Proof: Application of Jensen's inequality gives[3]

$$C(\rho) \leqslant \frac{1}{2}\log_2 E\left\{\det\left(\mathbf{I}_{n_s} + \frac{\rho a}{n_s}\tilde{\mathbf{H}}_1^{\dagger}\mathbf{L}\tilde{\mathbf{H}}_1\right)\right\}. \tag{6.36}$$

The result now follows by using Theorem 3.7. □

The following corollaries present some example scenarios for which the upper bound (6.35) reduces to simplified forms.

Corollary 6.4 *For the case $n_s \to \infty$, $C_U(\rho)$ becomes*

$$\lim_{n_s\to\infty} C_U(\rho) = \frac{1}{2}\log_2\left(\mathcal{K}\det(\bar{\mathbf{\Xi}}_1)\right), \tag{6.37}$$

where $\bar{\mathbf{\Xi}}_1$ is a $q \times q$ matrix with entries

$$\left\{\bar{\mathbf{\Xi}}_1\right\}_{m,n} = a^{1-\tau}\vartheta_{\tau-1}(a) + \rho a^{1-\tau}\vartheta_{\tau}(a). \tag{6.38}$$

Proof: The proof is straightforward and is omitted. □

This result shows that in AF MIMO dual-hop systems, when the numbers of antennas at both the relay and destination remain fixed, the ergodic capacity remains bounded as the number of source antennas grows large. This is in agreement with the results in Section 6.3.1.

Note that for the scenarios $n_r \to \infty$ and $n_d \to \infty$, simplified closed-form results can also be obtained by taking the corresponding limits in (6.37) or, alternatively, by using the equivalent single-hop MIMO capacity relations in (6.18) and (6.20), and applying known upper bounds for single-hop MIMO channels in [72]. We omit these expressions here for the sake of brevity.

Corollary 6.5 *Let $n_r = 1$. Then, $C_U(\rho)$ reduces to*

$$C_U^{n_r=1}(\rho) = \frac{1}{2}\log_2\left(1 + \rho n_d e^{\frac{1+\rho}{\alpha}} E_{n_d+1}\left(\frac{1+\rho}{\alpha}\right)\right). \tag{6.39}$$

When $n_d \to \infty$, $C_U^{n_r=1}(\rho)$ becomes

$$\lim_{n_d\to\infty} C_U^{n_r=1}(\rho) = \frac{1}{2}\log_2(1+\rho). \tag{6.40}$$

When $\alpha \to \infty$, $C_U^{n_r=1}(\rho)$ becomes

$$\lim_{\alpha\to\infty} C_U^{n_r=1}(\rho) = \frac{1}{2}\log_2(1+\rho). \tag{6.41}$$

[3]Note that this inequality has also been applied in the ergodic capacity analysis of single-user single-hop MIMO systems (see eg. [36, 63, 108]).

Proof: See Appendix D.4. □

This shows the interesting result that if a single relay antenna is employed, then when either the number of destination antennas n_d or the relay gain α grows large, the ergodic capacity is upper bounded by the capacity of an AWGN SISO channel.

Corollary 6.6 *In the high SNR regime, (i.e., as $\rho \to \infty$) for fixed relay gain α, $C_U(\rho)$ becomes*

$$\lim_{\rho\to\infty} C_U(\rho) = \frac{1}{2}\log_2\left(\mathcal{K}\det(\tilde{\mathbf{\Xi}})\right), \tag{6.42}$$

where $\tilde{\mathbf{\Xi}}$ is a $q \times q$ matrix with entries

$$\left\{\tilde{\mathbf{\Xi}}\right\}_{m,n} = \begin{cases} \Gamma(\tau-1), & n \le q-n_s, \\ \Gamma(\tau-1)\left(1+\frac{\alpha}{n_s n_r}(n_s-q+n)(\tau-1)\right), & n > q-n_s. \end{cases} \tag{6.43}$$

Proof: The proof follows from the observation that when $\rho \to \infty$, then $a \to 0$, and that asymptotic first-order expansion for confluent hypergeometric function U [4] can be expressed as

$$U(c,b,z) = z^{-c} + o(1), \quad z \to \infty. \tag{6.44}$$

□

6.5 Ergodic Capacity Lower Bound

The following theorem presents a new tight lower bound on the ergodic capacity of AF MIMO dual-hop systems.

Theorem 6.4 *The ergodic capacity of AF MIMO dual-hop systems is lower bounded by*

$$C(\rho) \ge C_L(\rho) = \frac{s}{2}\log_2\left(1+\frac{\rho a}{n_s}\exp\left(\frac{1}{s}\left[\sum_{k=1}^{s}\psi(n_s-s+k)+\mathcal{K}\sum_{k=q-s+1}^{q}\det(\mathbf{W}_k)\right]\right)\right), \tag{6.45}$$

where $\mathbf{W}_k$ is defined as in (3.57).

Proof: See Appendix D.5. □

The following corollaries present some example scenarios for which the lower bound (6.45) reduces to simplified forms.

Corollary 6.7 *For the case $n_s \to \infty$, $C_L(\rho)$ reduces to*

$$\lim_{n_s \to \infty} C_L(\rho) = \frac{s}{2} \log_2 \left(1 + \rho a \exp \left(\frac{\mathcal{K}}{s} \sum_{k=1}^{q} \det(\mathbf{W}_k) \right) \right). \tag{6.46}$$

Proof: When $n_s \to \infty$, $\psi(n_s - q + k)$ can be approximated as [4, (6.3.18)]

$$\begin{aligned} \psi(n_s - q + k)|_{n_s \to \infty} &\approx \ln(n_s - q + k) \\ &\approx \ln n_s. \end{aligned} \tag{6.47}$$

Substituting (6.47) into (6.45) yields the desired result. □

Again, we note that for the scenarios $n_r \to \infty$ and $n_d \to \infty$, simplified closed-form results can also be obtained by taking the corresponding limits in (6.37) or, alternatively, by using (6.18) and (6.20), and applying known lower bounds for single-hop MIMO channels in [72].

Corollary 6.8 *For the case $n_r = 1$, $C_L(\rho)$ reduces to*

$$C_L^{n_r=1}(\rho) = \frac{1}{2} \log_2 \left(1 + \frac{\rho\alpha}{n_s(1+\rho)} \exp \left(\psi(n_s) + \psi(n_d) - e^{(1+\rho)/\alpha} \sum_{l=0}^{n_d-1} E_{l+1}\left(\frac{1+\rho}{\alpha} \right) \right) \right). \tag{6.48}$$

When $n_s \to \infty$, $C_L^{n_r=1}(\rho)$ becomes

$$\lim_{n_s \to \infty} C_L^{n_r=1}(\rho) = \frac{1}{2} \log_2 \left(1 + \frac{\rho\alpha}{1+\rho} \exp \left(\psi(n_d) - e^{(1+\rho)/\alpha} \sum_{l=0}^{n_d-1} E_{l+1}\left(\frac{1+\rho}{\alpha} \right) \right) \right). \tag{6.49}$$

When $n_d \to \infty$, $C_L^{n_r=1}(\rho)$ becomes

$$\lim_{n_d \to \infty} C_L^{n_r=1}(\rho) = \frac{1}{2} \log_2 \left(1 + \frac{\rho\alpha}{n_s(1+\rho)} \exp \left(\psi(n_s) + \psi\left(\frac{1+\rho}{\alpha} \right) \right) \right). \tag{6.50}$$

When $\alpha \to \infty$, $C_L(\rho)$ becomes

$$\lim_{\alpha \to \infty} C_L^{n_r=1}(\rho) = \frac{1}{2} \log_2 \left(1 + \frac{\rho}{n_s} \exp(\psi(n_s)) \right). \tag{6.51}$$

Proof: See Appendix D.6. □

As also observed from the upper bound in Corollary 6.5, this result shows that for a system with a single relay antenna, when the relay gain α grows large, the ergodic capacity of an AF MIMO dual-hop channel is lower bounded by the capacity of an AWGN SISO channel (with scaled average SNR).

Corollary 6.9 *In the high SNR regime, (i.e., as $\rho \to \infty$) for fixed relay gain α, $C_L(\rho)$ becomes*

$$\lim_{\rho\to\infty} C_L(\rho) = \frac{s}{2}\log_2\left(1+\frac{\alpha}{n_r n_s}\exp\left(\frac{\mathcal{K}}{s}\sum_{k=q-s+1}^{q}\det\left(\tilde{\mathbf{W}}_k\right)\right)\right), \tag{6.52}$$

where $\tilde{\mathbf{W}}_k$ is a $q \times q$ matrix with entries

$$\left\{\tilde{\mathbf{W}}_k\right\}_{m,n} = \begin{cases} \Gamma(\tau-1), & n \neq k, \\ \Gamma(\tau-1)\left[\psi(n_s-q+n)+\psi(\tau-1)\right], & n = k. \end{cases} \tag{6.53}$$

Proof: Using the following approximation [4]

$$E_v(z) \approx \tfrac{1}{z}e^{-z}\left(1+o\left(\tfrac{1}{z}\right)\right) \quad |z| \to \infty \quad , \tag{6.54}$$

$\varsigma_{m+n}(a)$ can be approximated as

$$\varsigma_{m+n}(a)\big|_{\rho\to\infty} \approx \Gamma(\tau-1)\,\psi(\tau-1) \ , \tag{6.55}$$

which leads to the final result. □

6.6 Numerical Results

In this section, we verify our analytical expressions and examine the tightness of various upper and lower bounds proposed in this chapter through Monte-Carlo simulations. The simulation results are computed by averaging over 100,000 independent channel realizations.

Figure 6.2 compares the exact analytical capacity of AF MIMO dual-hop systems, based on (6.14) and (6.16), with Monte-Carlo simulated curves for two different antenna and relay configurations. In both cases, there is an exact agreement between the analysis and simulations, as expected.

Figure 6.3 illustrates the relationship in Corollary 6.3, where the high SNR power offset shift $\delta(n_d, k)$ is plotted against n_d, for $k = 1$, $k = 2$, and $k = 4$. As expected, for a fixed value of k, $\delta(n_d, k)$ is an increasing function of n_d, approaching a limit of 0 dB as $n_d \to \infty$.

Figure 6.4 compares the closed-form upper bound (6.35) with the exact analytical ergodic capacity based on (6.14) and (6.16), for two different AF MIMO dual-hop system configurations. The results are shown as a function of SNR ρ, with $\alpha = 2\rho$. We see that the closed-form upper bound is very tight for all SNRs, for both system configurations considered. Moreover, we see that in the low SNR regime (e.g., $\rho \approx 5$ dB), the upper bound and exact capacity curves coincide.

Figure 6.5 plots the closed-form upper bound (6.39), closed-form lower bound (6.48), and the exact analytical ergodic capacity based on (6.14) and (6.16), for an AF MIMO dual-hop system with $n_r = 1$.

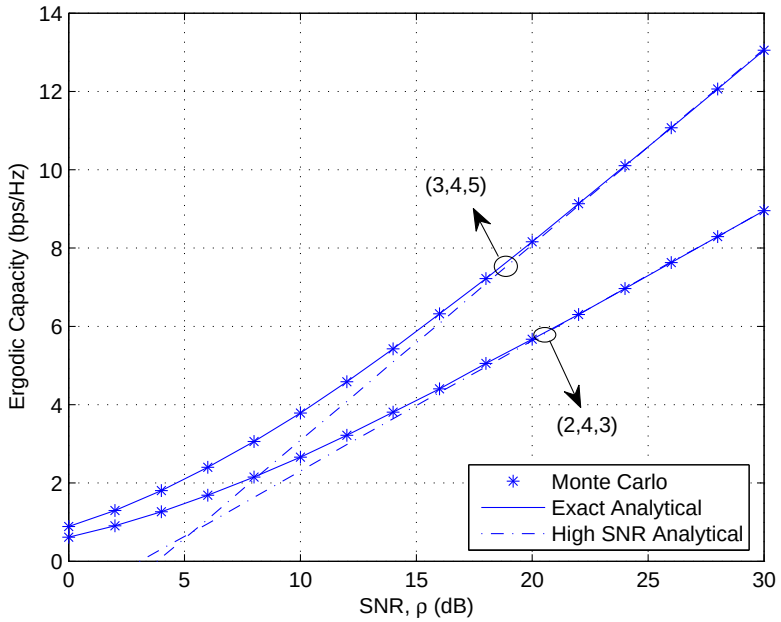

Figure 6.2. Comparison of exact analytical, high SNR analytical, and Monte Carlo simulation results for ergodic capacity of AF MIMO dual-hop systems with different antenna configurations. Results are shown for $\alpha/\rho = 2$.

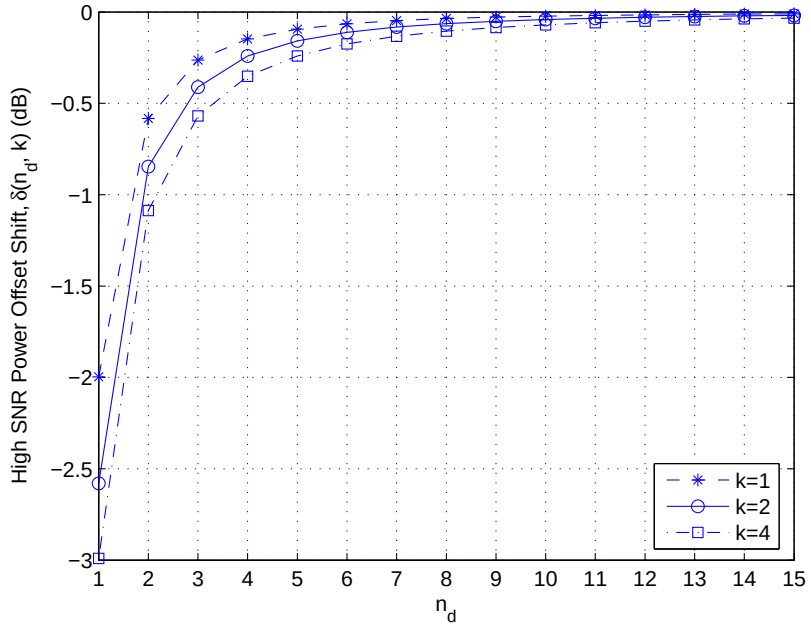

Figure 6.3. High SNR power offset shift, in decibels, obtaining by adding either (a) one antenna to the destination, (b) two antennas to the destination, or (c) four antennas to the destination. Results are shown for $n_s = n_r = 1$ and $\alpha/\rho = 2$.

The results are presented as a function of the relay gain α. We see that both the upper and lower bounds are quite tight for the entire range of α considered. The asymptotic approximations for the upper and lower bounds, based on (6.41) and (6.51) respectively, are also shown for further comparison, and are

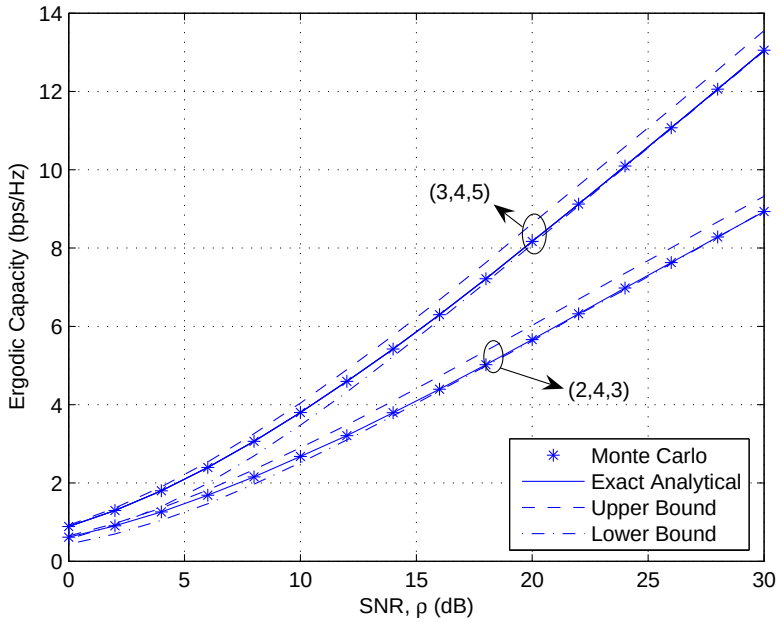

Figure 6.4. Comparison of bounds, exact analytical, high SNR analytical, and Monte Carlo simulation results for ergodic capacity of AF MIMO dual-hop systems with different antenna configurations. Results are shown for $\alpha/\rho = 2$.

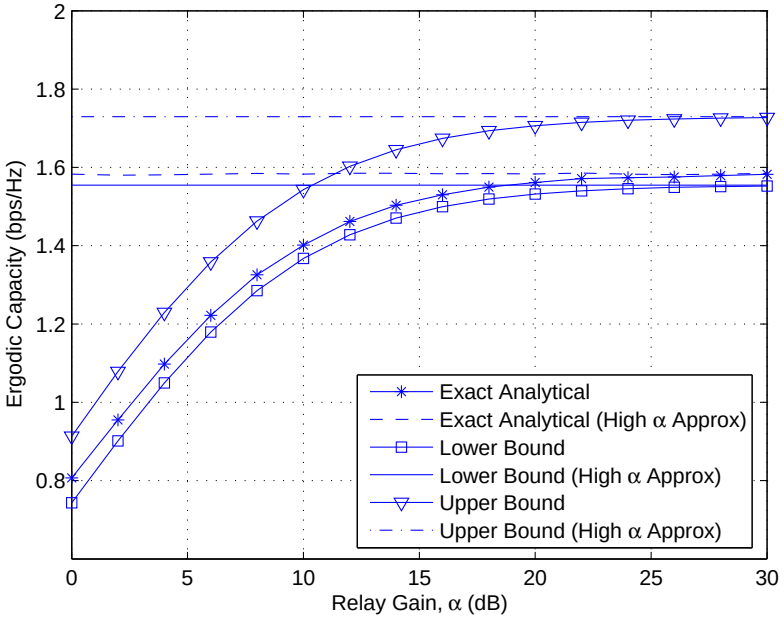

Figure 6.5. Comparison of capacity bounds, high α approximation, and exact analytical results for different relay gains. Results are shown for $n_r = 1$, $n_s = 2$, $n_d = 4$ and $\rho = 10\text{dB}$.

seen to converge for moderate values of α (e.g. within $\alpha \approx 20$ dB).

Figure 6.6 depicts the closed-form high SNR approximations for the exact ergodic capacity, as well as the respective upper and lower bounds, based on (6.42), and (6.52) respectively. For comparison, curves are

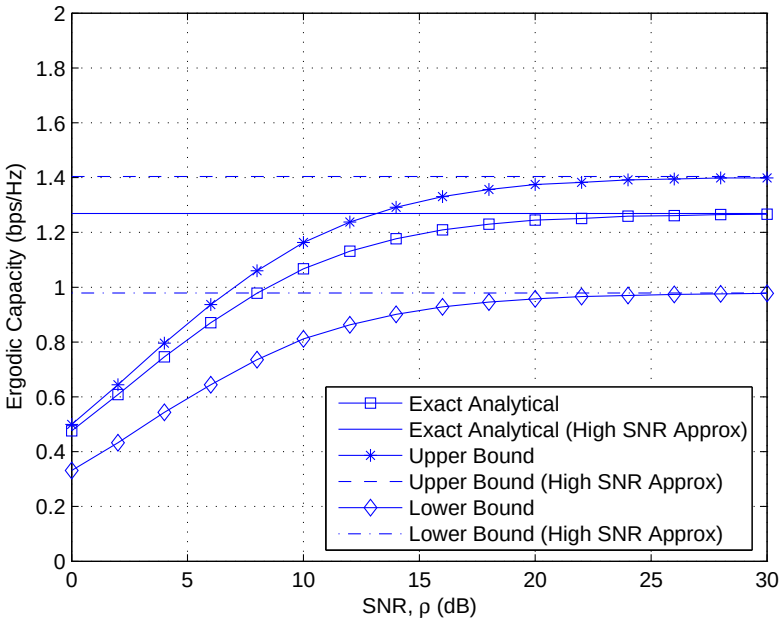

Figure 6.6. Comparison of capacity bounds, high SNR approximations, and exact analytical results. Results are shown for a system configuration $(3, 4, 2)$ and $\alpha = 2$.

also presented for the upper bound (6.35), lower bound (6.45), and the exact analytical ergodic capacity based in (6.14) and (6.16). Results are shown for an AF MIMO dual-hop system with configuration $(3, 4, 2)$. Clearly, the analytical high SNR approximations are seen to be very accurate for even moderate SNR levels (e.g., $\rho \approx 20$ dB).

6.7 Conclusion

This chapter presented an analytical characterization of the ergodic capacity of AF MIMO dual-hop relay channels under the common assumption that CSI is available at the destination terminal, but not at the relay or the source terminal. A new exact expression for the ergodic capacity, as well as simplified and insightful closed-form expressions for the high SNR regime were derived. Simplified closed-form upper and lower bounds were also presented, which were shown to be tight for all SNRs. The analytical results were made possible by first employing RMT techniques to derive new expressions for the p.d.f. of an unordered eigenvalue, as well as random determinant results for the equivalent AF MIMO dual-hop relay channel, described by a certain product of finite-dimensional complex random matrices. The analytical results were validated through comparison with numerical simulations.

Chapter 7

Performance Analysis of OC in Rayleigh-Product Channels

7.1 Introduction

Wireless communications systems are generally subjected to co-channel interferences. For multiple antenna systems, OC scheme which maximizes the received SINR by exploiting the CSI in a SIMO antenna system has been proposed in [100] to combat the effect of co-channel interference. Later in [103], the concept of OC was extended to a MIMO antenna system in which the transmit and receive antennas are jointly optimized to maximize the SINR by projecting the transmitted signal onto the strongest eigenspace of the interference-inverted channel matrix in quadratic form. The performance of OC systems have been extensively analyzed in the literature, see e.g., [27, 44, 47, 59–61, 64, 81, 106].

While these prior research works are fundamental in nature and profoundly important in understanding the performance of the MIMO OC systems, most adopted the assumption of a perfectly rich-scattering environment that renders a full-rank MIMO channel matrix, Hence, these results tend to be overly optimistic, and may fail to address practical environments such as keyholes [8] or the more general double-scattering channels [25].

Motivated by this, in this chapter, we intend to provide an accurate account of the real performance of OC systems operating in double-scattering channels. To allow useful results to be derived, we shall assume that the transmit and receive antennas are uncorrelated and the scattering matrix in the double-scattering model is identity, giving rise to a Rayleigh-product MIMO channel with co-channel interference. Furthermore, as in [44], we shall adopt the interference-limited assumption so that noise can be neglected and also the assumption that the co-channel interferers are of equal power and Rayleigh-faded.

In this chapter, we present the exact closed-form expression for the outage probability of the OP systems in interference-limited Rayleigh product channel based on the new statistical results derived in Chapter 3. To gain more insight, we apply these findings to develop further analytical results for the keyhole channel, an important special case of double scattering channel. In particular, the expressions in closed form

for the p.d.f. and c.d.f. (and their asymptotic expansions), the ergodic capacity, the outage probability and the SER of the optimally-combined keyhole channel with co-channel interference are derived.

7.2 System Model

Consider a MIMO system equipped with N_t antennas at the transmitter and N_r antennas at the receiver, and assume that there exist $N_{\mathcal{I}}$ co-channel interferers with $N_{\mathcal{I}} \geq N_r$. The received signals in vector form can be modeled as

$$\mathbf{y} = \sqrt{P_0}\mathbf{H}\mathbf{t}s_0 + \sum_{n=1}^{N_{\mathcal{I}}} \sqrt{P_n}\mathbf{h}_n s_n + \boldsymbol{\eta} \tag{7.1}$$

where s_0 is the transmitted signal of the desired user, and s_n $(n \geq 1)$ denotes the signals transmitted from the nth interferer with $\mathrm{E}[|s_n|^2] = 1\ \forall n$ so that $\{P_n\}$ are the transmitted power of the users. Additionally, $\boldsymbol{\eta}$ is the complex noise vector with independent elements following $\mathcal{CN}(0, \sigma^2)$, $\mathbf{t} \in \mathbb{C}^{N_t}$ denotes the transmit beamforming vector of the desired user with $\|\mathbf{t}\| = 1$, $\mathbf{h}_n \in \mathbb{C}^{N_r}$ is the complex channel vector of the nth interferer with independent elements following $\mathcal{CN}(0, 1)$, and

$$\mathbf{H} = \frac{1}{\sqrt{N_s}}\mathbf{H}_1\mathbf{H}_2 \tag{7.2}$$

in which $\mathbf{H}_1 \sim \mathcal{CN}_{N_r,N_s}(\mathbf{0}_{N_r\times N_s}, \mathbf{I}\otimes\mathbf{I})$ and $\mathbf{H}_2 \sim \mathcal{CN}_{N_s,N_t}(\mathbf{0}_{N_s\times N_t}, \mathbf{I}\otimes\mathbf{I})$ are random matrices, is the Rayleigh-product MIMO channel[1] between the transmitter and the desired receiver with N_s being the number of scatterers in the environment [104]. For ease of exposition, we define $\mathbf{H}_{\mathcal{I}} \triangleq [\mathbf{h}_1 \cdots \mathbf{h}_n]$. As a consequence, (7.1) can be re-expressed as

$$\mathbf{y} = \sqrt{P_0}\mathbf{H}\mathbf{t}s_0 + \mathbf{H}_{\mathcal{I}}\mathbf{P}_{\mathcal{I}}^{\frac{1}{2}}\mathbf{s}_{\mathcal{I}} + \boldsymbol{\eta} \tag{7.4}$$

where $\mathbf{P}_{\mathcal{I}} = \mathrm{diag}(P_1, \ldots, P_{N_{\mathcal{I}}})$, and $\mathbf{s}_{\mathcal{I}} = [s_1\ s_2 \cdots s_{N_{\mathcal{I}}}]^T$.

In this model, the desired user's channel is assumed to undergo double-scattering while the interferences do not. This setting is particularly useful for an uplink space-division multiple-access (SDMA) system in which the same spectrum is shared by a number of users within a cell. As a result, it can represent the scenario where the desired user is at the boundary of the cell and the co-channel interferences (other users) are much closer to the base station receiver. The double-scattering desired user link with single-scattering interferers is therefore an important benchmark for the performance of MIMO-SDMA systems using OC in the uplink.

To allow further analysis of the system, henceforth, we assume that the system is interference-limited, meaning that the noise can be neglected (though the results of this paper will be examined numerically

[1] In the double-scattering model [25], the channel matrix in (7.2) would have been written as

$$\mathbf{H} = \frac{1}{\sqrt{N_s}}\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{H}_1\boldsymbol{\Phi}^{\frac{1}{2}}\mathbf{H}_2\boldsymbol{\Xi}^{\frac{1}{2}} \tag{7.3}$$

where $\boldsymbol{\Phi}$ denotes the scatterer correlation matrix, and $\boldsymbol{\Sigma}$ and $\boldsymbol{\Xi}$ denote, respectively, the spatial correlation matrices at the receiver and the transmitter. Hence, the Rayleigh product model is the special case of the double-scattering model when $\boldsymbol{\Sigma} = \boldsymbol{\Phi} = \boldsymbol{\Xi} = \mathbf{I}$.

in the presence of noise in Section 7.4). This assumption is particularly reasonable when the interference to noise ratio (INR) is high. To make the analysis tractable, we also assume that $P_{\mathcal{I}} \triangleq P_1 = P_2 = \cdots = P_{N_{\mathcal{I}}}$.

The OC for (7.4) with CSI has been well known [103]. In particular, the optimum receiver combining, which maximizes the output signal-to-interference ratio (SIR) (with noise being ignored) by left-multiplying $\mathbf{y}$ with a vector $\mathbf{r}^\dagger$, is achieved by having

$$\mathbf{r} = \left(P_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}^\dagger\right)^{-1}\mathbf{H}\mathbf{t}, \tag{7.5}$$

which gives the SIR, ρ, as

$$\begin{aligned}\rho &= \mathbf{t}^\dagger\mathbf{H}^\dagger\left(P_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}^\dagger\right)^{-1}\mathbf{H}\mathbf{t} \\ &= \frac{1}{N_s P_{\mathcal{I}}}\mathbf{t}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\left(\mathbf{H}_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}^\dagger\right)^{-1}\mathbf{H}_1\mathbf{H}_2\mathbf{t}.\end{aligned} \tag{7.6}$$

According to the Rayleigh-Ritz theorem [30], γ is maximized by choosing $\mathbf{t} = \mathbf{u}_{\mathsf{max}}$, where $\mathbf{u}_{\mathsf{max}}$ denotes the eigenvector corresponding to the largest eigenvalue of the matrix

$$\mathbf{F} \triangleq \frac{1}{N_s}\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\left(\mathbf{H}_{\mathcal{I}}\mathbf{H}_{\mathcal{I}}^\dagger\right)^{-1}\mathbf{H}_1\mathbf{H}_2. \tag{7.7}$$

The corresponding maximum SIR is given by

$$\rho_{\mathsf{max}} = \frac{P_0}{P_{\mathcal{I}}}\lambda_{\mathsf{max}}, \tag{7.8}$$

where λ_{max} is the largest eigenvalue of $\mathbf{F}$. Apparently, the performance of (7.4) depends directly upon the statistical properties of λ_{max}.

7.3 Performance Analysis of OC Systems in Rayleigh-Product Channels

In this section, we study the performance of the OC systems in Rayleigh-product channels based on a set of newly derived closed-form expressions of the c.d.f. and p.d.f. the maximum eigenvalue λ_{max} presented in Chapter 3.

7.3.1 Outage Analysis of OC Systems in Rayleigh-Product Channels

Outage probability is an important performance metric in communication systems, which is defined as the probability that the system fails to achieve an acceptable SIR threshold, say, ρ_{th}. In this subsection, we present the closed-form outage expressions of the OC systems in Rayleigh-product channels.

Theorem 7.1 *For interference-limited Rayleigh-product channels, the outage probability of the OC systems can be computed as*

1) When $N_t \leq N_r$ or $N_t \geq N_r \geq N_s$,

$$\mathcal{P}_{\mathsf{out}} = \frac{\prod_{i=1}^{m}(-1)^{pN_t}\Gamma(N_{\mathcal{I}}+N_s-i+1)\det\left(\mathbf{\Delta}\left(\frac{P_I\rho_{\mathsf{th}}}{P_0}\right)\right)}{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}-N_r+m-i+1)\Gamma(m-i+1)\Gamma(n-i+1)}, \tag{7.9}$$

2) When $N_t \geq N_s \geq N_r$ or $N_s \geq N_t \geq N_r$,

$$\mathcal{P}_{\mathsf{out}} = \frac{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}+n-i+1)\det\left(\mathbf{\Theta}\left(\frac{P_I\rho_{\mathsf{th}}}{P_0}\right)\right)}{\prod_{j=1}^{m}\Gamma(N_{\mathcal{I}}-j+1)\Gamma(n-j+1)\Gamma(m-j+1)\prod_{i=1}^{N_t}\Gamma(N_t-i+1)}, \tag{7.10}$$

where $\mathbf{\Delta}(x)$ and $\mathbf{\Theta}(x)$ is an $N_r \times N_r$ matrix whose entries are defined in Theorem 3.5.

Proof: Following the definition of outage probability, the desired results can be obtained by directly invoking Theorem 3.5. □

The above theorem gives a complete characterization of the outage behavior of OC systems in interference-limited Rayleigh-product channels. Although these expressions can be efficiently evaluated by standard softwares, such as Mathematica, the expressions themselves are too complicated to gain physical insights. In the following section, we consider a special case of Rayleigh-product channels, namely keyhole channels, for which, we give a detailed performance investigation.

7.3.2 Performance Analysis of OC Systems in Keyhole Channels

In this subsection, we examine, in detail, the keyhole channel which is a special case of double-scattering or Rayleigh-product channels discussed in Section 7.2. We first give the exact closed-form expressions for the c.d.f. and p.d.f. when $N_s = 1$, then derive the asymptotic expressions for the c.d.f. and p.d.f., which will enable us to reveal some insightful properties. Using these new statistical results, we also derive the ergodic capacity, the outage probability and the SER of the keyhole channels.

When $N_s = 1$, the number of scatterers in the channel is one and this channel is usually referred as the keyhole, or pinhole, channel.

Corollary 7.1 *When $N_s = 1$, the c.d.f. of the non-zero eigenvalue of $\mathbf{F}$ is expressed as*

$$\mathcal{F}_{\lambda_{\mathsf{max}}}(x) = 1 - \frac{\Gamma(N_{\mathcal{I}}+1)}{\Gamma(N_r)\Gamma(N_{\mathcal{I}}-N_r+1)}\sum_{k=0}^{N_t-1}\frac{x^k\Gamma(N_{\mathcal{I}}-N_r+k+1)}{\Gamma(k+1)} U(N_{\mathcal{I}}-N_r+k+1, k-N_r+1, x). \tag{7.11}$$

Corollary 7.2 *When $N_s = 1$, the p.d.f. of the non-zero eigenvalue of $\mathbf{F}$ is given by*

$$f_{\lambda_{\mathsf{max}}}(x) = \frac{\Gamma(N_{\mathcal{I}}+1)\Gamma(N_{\mathcal{I}}+N_t-N_r+1)}{\Gamma(N_{\mathcal{I}}-N_r+1)\Gamma(N_t)\Gamma(N_r)}x^{N_t-1}U(N_{\mathcal{I}}+N_t-N_r+1, N_t-N_r+1, x). \tag{7.12}$$

The proofs of the above two corollaries are straightforward and thus omitted.

To this end, we derive the asymptotic expansions for the c.d.f. and p.d.f. of the non-zero eigenvalue of $\mathbf{F}$. The simple expressions obtained enable us to investigate the asymptotic outage probability later.

Theorem 7.2 *The asymptotic expansions for the c.d.f. and p.d.f. of the non-zero eigenvalue of* $\mathbf{F}$ *are given, respectively, by*

$$\mathcal{F}_{\lambda_{\max}}(x)|_{N_s=1} = ax^s + \mathcal{O}(x^{s+1}), \tag{7.13}$$

$$f_{\lambda_{\max}}(x)|_{N_s=1} = asx^{s-1} + \mathcal{O}(x^s), \tag{7.14}$$

where $s = \min(N_t, N_r)$ *and*

$$a = \begin{cases} \dfrac{\Gamma(N_{\mathcal{I}} + N_t - N_r + 1)\Gamma(N_r - N_t)}{\Gamma(N_{\mathcal{I}} - N_r + 1)\Gamma(N_t + 1)\Gamma(N_r)}, N_t \leq N_r - 1, \\ -\dfrac{\Gamma(N_{\mathcal{I}} + 1)[\ln x + \psi(N_{\mathcal{I}} + 1)]}{\Gamma(N_{\mathcal{I}} - N_r + 1)\Gamma(N_t + 1)\Gamma(N_r)}, N_t = N_r, \\ \dfrac{\Gamma(N_{\mathcal{I}} + 1)\Gamma(N_t - N_r)}{\Gamma(N_{\mathcal{I}} - N_r + 1)\Gamma(N_t)\Gamma(N_r + 1)}, N_t \geq N_r + 1, \end{cases} \tag{7.15}$$

where $\psi(\cdot)$ *is the digamma function which, for integer* n*, can be expressed as [26]*

$$\psi(n) = -K + \sum_{p=1}^{n-1} \frac{1}{p}, \tag{7.16}$$

where $K \approx 0.57721566$ *denotes the Euler's constant.*

Proof: The asymptotic expansion of $U(a, b, x)$ can be obtained from [4, (13.5.6)–(13.5.12)]. Utilizing these results together with Corollary 7.2, we can get the first-order expansion for the p.d.f. of the keyhole channel. The c.d.f. is obtained by an additional integration. □

7.3.2.1 Ergodic Capacity

Now, we present the ergodic capacity (b/s/Hz) of the MIMO OC system in keyhole channels with co-channel interference. This is obtained by evaluating

$$C = \mathtt{E}_{\lambda_{\max}}\left[\log_2\left(1 + \frac{P_0}{P_{\mathcal{I}}}\lambda_{\max}\right)\right]. \tag{7.17}$$

Applying the result in Corollary 7.2 and after some mathematical manipulations, (7.17) can be expressed in closed-form as presented in the following theorem.

Theorem 7.3 *The ergodic capacity of the interference-limited keyhole channel is given by*

$$C = \frac{\log_2 e G_{4,3}^{2,4}\left(\frac{P_0}{P_{\mathcal{I}}}\middle| \begin{matrix} 1,1,1-N_t,1-N_r \\ 1,1+N_{\mathcal{I}}-N_r,0 \end{matrix}\right)}{\Gamma(N_{\mathcal{I}}-N_r+1)\Gamma(N_t)\Gamma(N_r)}, \tag{7.18}$$

where $G_{p,q}^{m,n}\left(x\left|\begin{smallmatrix} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{smallmatrix}\right.\right)$ *is the Meijer-G function defined in [26, (9.301)].*

Proof: We start by expressing Hypergeometric function $U(\cdot,\cdot,\cdot)$ and $\ln(1+x)$ [76, (8.4.6.5)] in terms of Meijer-G function, i.e.,

$$\begin{cases} U(a,b,x) = \dfrac{1}{\Gamma(a)\Gamma(a-b+1)} G_{1,2}^{2,1}\left(x\left|\begin{smallmatrix} 1-a \\ 0,1-b \end{smallmatrix}\right.\right), \\ \ln(1+ax) = G_{2,2}^{1,2}\left(ax\left|\begin{smallmatrix} 1,1 \\ 1,0 \end{smallmatrix}\right.\right). \end{cases} \tag{7.19}$$

The desired result can be obtained with the help of the integration formula [26, (7.821.3)]. □

Although the explicit equation given above can be used to compute the ergodic capacity efficiently, it does not offer much insight about the system. Therefore, it is of interest to analyze the special case which gives simplified expressions. The following two corollaries characterize the ergodic capacity of the keyhole channel in the high and low SIR regimes.

Corollary 7.3 *In the low SIR regime, for* $N_{\mathcal{I}} > N_r$*, the ergodic capacity in (7.18) can be expressed as*

$$C \approx K\left(\frac{P_0}{P_{\mathcal{I}}}\right), \tag{7.20}$$

where

$$K = \frac{N_t N_r}{N_{\mathcal{I}} - N_r}. \tag{7.21}$$

Proof: At low SIRs, we can approximate the capacity formula (7.17) as $C \approx \frac{P_0}{P_{\mathcal{I}}}\mathtt{E}[\lambda_{\mathsf{max}}]$. Then, utilizing the integral formula in [26, (7.612.2)] yields

$$\begin{aligned} C &\approx \frac{\Gamma(N_{\mathcal{I}}+1)\Gamma(N_{\mathcal{I}}+N_t-N_r+1)}{\Gamma(N_{\mathcal{I}}-N_r+1)\Gamma(N_t)\Gamma(N_r)} \frac{P_0}{P_{\mathcal{I}}} \int_0^\infty x^{N_t} U(N_{\mathcal{I}}+N_t-N_r+1, N_t-N_r+1, x)dx \\ &= K\left(\frac{P_0}{P_{\mathcal{I}}}\right), \end{aligned} \tag{7.22}$$

where K has been defined in (7.21). For $N_{\mathcal{I}} = N_r$, nevertheless, due to the capacity approximation at low SIRs, the integral in (7.22) diverges and such an approximation is not available. □

Corollary 7.4 *In the high SIR regime, the ergodic capacity in (7.18) allows the following expression*

$$C \approx \log_2 \frac{P_0}{P_{\mathcal{I}}} + (\log_2 e) \times [\psi(N_r) + \psi(N_t) - \psi(N_{\mathcal{I}} - N_r + 1)]. \tag{7.23}$$

Proof: At high SIRs, we approximate the capacity in (7.17) by $C \approx \mathtt{E}\left[\log_2 \frac{P_0}{P_{\mathcal{I}}}\lambda_{\mathsf{max}}\right]$ so that

$$C \approx \log_2 \frac{P_0}{P_{\mathcal{I}}} + (\log_2 e) \times \frac{d}{d\tau}\mathtt{E}_{\lambda_{\mathsf{max}}}\left[\lambda_{\mathsf{max}}^{\tau}\right]\Big|_{\tau=0}$$
$$= \log_2 \frac{P_0}{P_{\mathcal{I}}} + (\log_2 e) \times \frac{\Gamma(N_{\mathcal{I}}+1)\Gamma(N_{\mathcal{I}}+N_t-N_r+1)}{\Gamma(N_{\mathcal{I}}-N_r+1)\Gamma(N_t)\Gamma(N_r)} \frac{d}{d\tau}\int_0^{\infty} x^{N_t+\tau-1}\theta(x)dx\Big|_{\tau=0}, \quad (7.24)$$

which leads to the result of (7.23). Note $\theta(x)$ is defined in (7.32), and in (7.24), the identity that $\frac{\Gamma(x)}{dx} = \psi(x)\Gamma(x)$ has been used. □

The results show that the ergodic capacity is affected by three important parameters besides the SIR, namely, $N_t, N_r, N_{\mathcal{I}}$. First of all, when $N_{\mathcal{I}} = N_r$, the ergodic capacity is a symmetric function of N_t and N_r and hence they have the same impact on the capacity. However, when $N_{\mathcal{I}}$ is strictly greater than N_r, we start with a system with $N_t = N_r$ and study the effect on the capacity by adding one more antenna at either transmitter side or receiver side. In the low SIR regime, as we can see, increasing the number of transmit antennas by one contributes to the increase of ergodic capacity by a factor of $\frac{N_r}{N_{\mathcal{I}}-N_r}$. In contrast, if the antenna is added at the receiver side, the ergodic capacity will be increased by $\frac{N_{\mathcal{I}}}{N_{\mathcal{I}}-N_r-1}\frac{N_r}{N_{\mathcal{I}}-N_r}$, which is obviously greater than $\frac{N_r}{N_{\mathcal{I}}-N_r}$. In the high SIR regime, it can also be observed that it is better to deploy the additional antennas at the receiver than at the transmitter side in terms of the ergodic capacity.

7.3.2.2 Outage Probability

According to Corollary 7.1, the exact outage probability of the keyhole channel can be found as

$$\mathcal{P}_{\mathsf{out}} = \mathrm{Prob}(\rho_{\mathsf{max}} < \rho_{\mathsf{th}}) = 1 - \frac{\Gamma(N_{\mathcal{I}}+1)}{\Gamma(N_r)\Gamma(N_{\mathcal{I}}-N_r+1)}\sum_{k=0}^{N_t-1}\frac{\left(\frac{P_{\mathcal{I}}\rho_{\mathsf{th}}}{P_0}\right)^k}{\Gamma(k+1)}\Gamma(N_{\mathcal{I}}-N_r+k+1)$$
$$U\left(N_{\mathcal{I}}-N_r+k+1, k-N_r+1, \frac{P_{\mathcal{I}}\rho_{\mathsf{th}}}{P_0}\right),$$

which, using the results in Theorem 7.2, can further be approximated, at high SIRs, as

$$\mathcal{P}_{\mathsf{out}} \approx \mathrm{Prob}(\rho_{\mathsf{max}} < \rho_{\mathsf{th}}) = a\left(\frac{P_0}{P_{\mathcal{I}}\rho_{\mathsf{th}}}\right)^{-s}. \quad (7.25)$$

In [104], it has been shown that the diversity order of a double-scattering MIMO channel is upper-bounded by $\frac{N_tN_sN_r}{\mathsf{max}(N_t,N_s,N_r)}$ which is achievable only if the following condition holds

$$2\max(N_t, N_s, N_r) + 1 \geq N_t + N_s + N_r. \quad (7.26)$$

For the keyhole channel with co-channel interference that we consider here, it can be easily shown that

this condition is satisfied and therefore

$$\text{Diversity Gain} = \frac{N_t N_r}{\max(N_t, 1, N_r)} = \min(N_t, N_r), \tag{7.27}$$

which coincides with the asymptotic result in (7.13). Intriguingly, note also that the diversity order does not depend on the number of co-channel interferers $N_{\mathcal{I}}$ which turns out to affect only the coding (or array) gain of the system. To exemplify this, let us focus on (7.15) for the case $N_t \leq N_r - 1$. When the number of interferers increases from, say, $N_{\mathcal{I}}$ to $\bar{N}_{\mathcal{I}} = N_{\mathcal{I}} + 1$, then we have

$$\begin{aligned} & a(\bar{N}_{\mathcal{I}}) \\ &= \frac{\Gamma(\bar{N}_{\mathcal{I}} + N_t - N_r + 1)\Gamma(N_r - N_t)}{\Gamma(\bar{N}_{\mathcal{I}} - N_r + 1)\Gamma(N_t + 1)\Gamma(N_r)} \\ &= \frac{N_{\mathcal{I}} + N_t - N_r + 1}{N_{\mathcal{I}} - N_r + 1} \frac{\Gamma(N_{\mathcal{I}} + N_t - N_r + 1)\Gamma(N_r - N_t)}{\Gamma(N_{\mathcal{I}} - N_r + 1)\Gamma(N_t + 1)\Gamma(N_r)} \\ &> a(N_{\mathcal{I}}). \end{aligned} \tag{7.28}$$

As a result, a increases with the number of interferers, which in turn decreases the coding gain because

$$\text{CODING GAIN} = \left(\frac{1}{a}\right)^{\frac{1}{s}}. \tag{7.29}$$

In addition, (7.15) demonstrates that both the c.d.f. and the p.d.f. of λ_{max} decay to zero more slowly for $N_t = N_r$ than for $N_t \neq N_r$, due to the term $\ln x$ in the leading factor a.

7.3.2.3 SER

In addition to the outage probability, SER is also a common metric used to characterize the performance of a communication system. For most modulation formats, the average SER can be evaluated as [75]

$$\text{SER}\left(\frac{P_0}{P_{\mathcal{I}}}\right) = \mathtt{E}_{\rho_{\mathsf{max}}}\left[\alpha Q\left(\sqrt{2\beta\rho_{\mathsf{max}}}\right)\right], \tag{7.30}$$

where $Q(\cdot)$ is the Gaussian Q-function, and α, β are modulation-specific constants. For example, BPSK requires $\alpha = 1, \beta = 1$; BFSK has $\alpha = 1, \beta = 0.5$ with orthogonal signaling or $\alpha = 1, \beta = 0.715$ with minimum correlation while for M-ary PAM, $\alpha = 2(M-1)/M, \beta = 3/(M^2 - 1)$.

Theorem 7.4 *The SER of the interference-limited keyhole channel is given by*

$$\textit{SER}\left(\frac{P_0}{P_{\mathcal{I}}}\right) = \frac{\alpha G_{3,3}^{3,2}\left(\frac{\beta P_0}{P_{\mathcal{I}}} \middle| \begin{matrix} 1 - N_t, 1 - N_r, 1 \\ 0, 1/2, 1 + N_{\mathcal{I}} - N_r \end{matrix}\right)}{2\sqrt{\pi}\Gamma(N_{\mathcal{I}} - N_r + 1)\Gamma(N_t)\Gamma(N_r)}. \tag{7.31}$$

Proof: Define

$$\theta(x) \triangleq U(N_{\mathcal{I}} + N_t - N_r + 1, N_t - N_r + 1, x). \tag{7.32}$$

Making use of (7.8) and Corollary 7.2, the SER of keyhole channel can be evaluated as

$$\begin{aligned}\text{SER}\left(\frac{P_0}{P_{\mathcal{I}}}\right) &= \int_0^{\infty}\left[\alpha Q\left(\sqrt{2\beta\frac{P_0}{P_{\mathcal{I}}}x}\right)\right] f_{\lambda_{\max}}(x)dx \\ &= \frac{\alpha\Gamma(N_{\mathcal{I}}+1)\Gamma(N_{\mathcal{I}}+N_t-N_r+1)}{\Gamma(N_{\mathcal{I}}-N_r+1)\Gamma(N_t)\Gamma(N_r)}\int_0^{\infty}\left[Q\left(\sqrt{2\beta\frac{P_0}{P_{\mathcal{I}}}x}\right)x^{N_t-1}\theta(x)\right]dx.\end{aligned} \tag{7.33}$$

Utilizing the relation that $Q(x) = \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right)$, where $\text{erfc}(\cdot)$ denotes the complementary error function [26], we express $\text{erfc}(x)$ in terms of Meijer-G function as [76]

$$\text{erfc}(\sqrt{x}) = \frac{1}{\sqrt{\pi}}G_{1,2}^{2,0}\left(x\left|{}^{1}_{0,1/2}\right.\right). \tag{7.34}$$

Now, applying (7.19), (7.34), and integrating using [26, (7.821.3)] yields the desired result. □

7.4 Numerical Results

In this section, we provide numerical results to confirm the correctness of the analytical results we have derived. In addition to that, various examples are also given to demonstrate how various system parameters impact on the performance of a Rayleigh-product MIMO channel using OC. In particular, we validate our two assumptions, namely, "equal power interference", and "interference-limited", we show that the assumed model can provide very good performance reference to the real system. All the simulation results are obtained based on 1,000,000 independent channel realizations.

Figure 7.1 plots the outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR (i.e., the SIR normalized by the threshold) $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ for various number of scatterers N_s when $N_t = 3$, $N_r = 5$, and $N_{\mathcal{I}} = 6$. Results in this figure indicate that the number of scatterers has a significant impact on the outage performance, which agrees with the expectation that the number of scatterers should somehow link with the diversity order of the channel which affects the outage probability. In addition, as can be seen in Figure 7.1, the results for both the Monte Carlo simulations and the analytical formulae agree perfectly with each other.

Figure 7.2 provides similar results as in Figure 7.1 but for various number of co-channel interferers $N_{\mathcal{I}}$ for two settings when $(N_t, N_r, N_s) = (3, 4, 1)$, and $(N_t, N_r, N_s) = (3, 4, 5)$. It shows that a larger number of co-channel interferers degrades the system performance and leads to an increase in the outage probability. Also, intriguingly, it is observed that $N_{\mathcal{I}}$ affects only on the coding gain but not the diversity order.

Now, numerical results in Figure 7.3 and Figure 7.4 are provided for the outage probability performance when $(N_t, N_r, N_s) = (3, 4, 7)$ under the case when the interference-limited system and equal-power interferers assumptions are no longer true. In Figure 7.3, we compare the analytical results to Monte Carlo results for the system with equal power interference $\frac{P_{\mathcal{I}}}{\sigma^2} = 3$ (dB) plus white Gaussian noise. While in Figure 7.4, we plot the analytical results against the Monte Carlo results for the system with unequal power inteference $P_i = \frac{2i-1}{N_{\mathcal{I}}}P_{\mathcal{I}}$ and white Gaussian noise. A close observation from the results

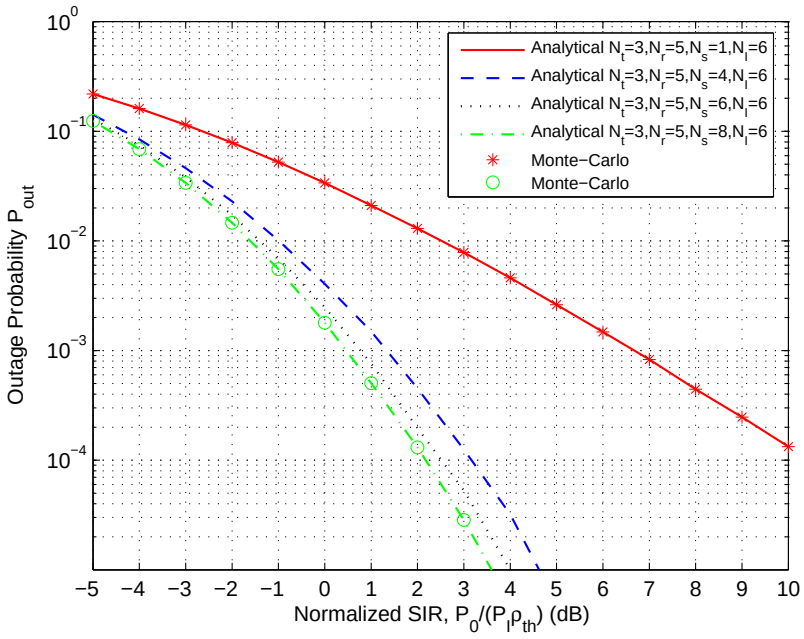

Figure 7.1. Outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ in decibels in Rayleigh-product channels with co-channel interference for various number of scatterers N_s.

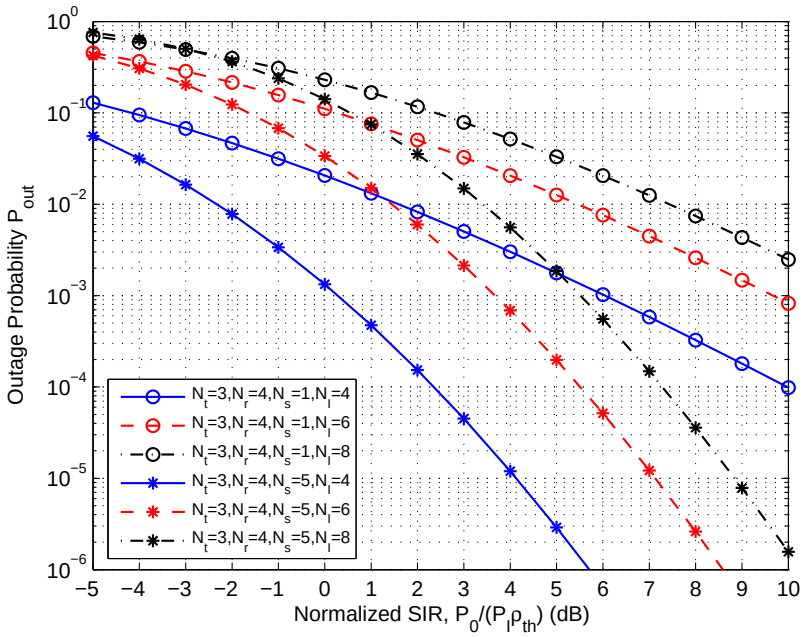

Figure 7.2. Outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ in decibels in Rayleigh-product channels for various number of co-channel interferers $N_{\mathcal{I}}$ with $N_t = 3, N_r = 4$ and for both the cases $N_s = 1$ and $N_s = 5$.

in the figures reveals that the gap between the analytical results and the Monte Carlo simulations closes down if the number of interferers increases. In particular, when $N_{\mathcal{I}} = 16$, the difference is inappreciable. Very interestingly, it is also observed that the analytical results are more accurate for the case with unequal-power interferers than the case with equal-power interferers for a given total interference power.

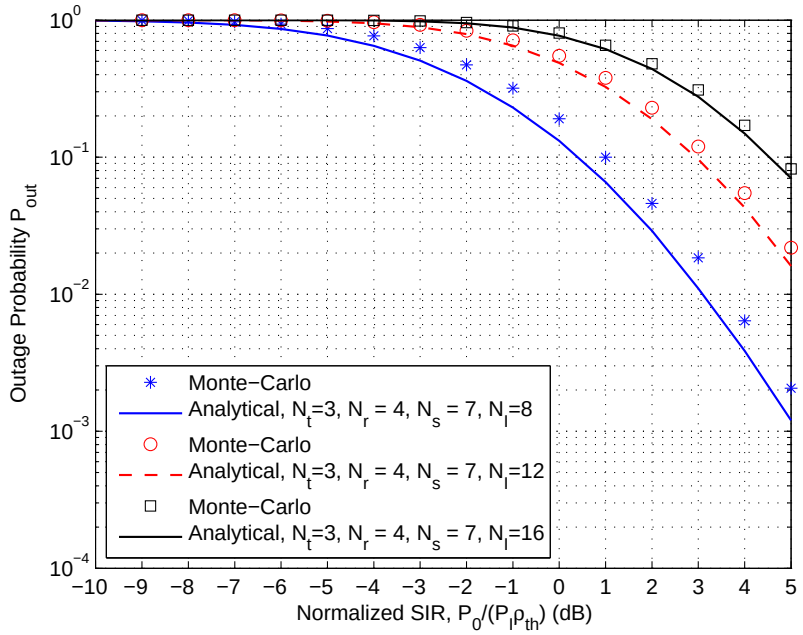

Figure 7.3. Outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ in decibels with equal-power co-channel interferers and white noise.

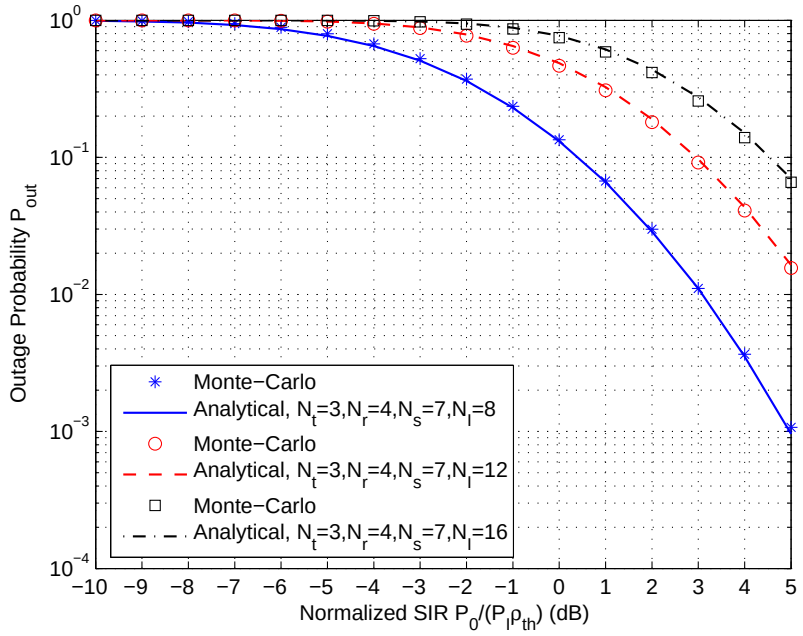

Figure 7.4. Outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ in decibels with unequal-power co-channel interferers $P_i = \frac{2i-1}{N_{\mathcal{I}}}P_{\mathcal{I}}$, for $i = 1, \ldots, N_{\mathcal{I}}$ and white noise.

Results in Figs. 7.5–7.7 are provided for keyhole MIMO channels (i.e., $N_s = 1$). Figure 7.5 illustrates the outage probability results for various number of transmit antennas N_t when $N_r = 5$ and $N_{\mathcal{I}} = 6$. The results for the exact analytical expression (7.11), the asymptotic expression (7.13) and the Monte-Carlo simulations are shown and compared. As we can see, the Monte Carlo and the exact analytical

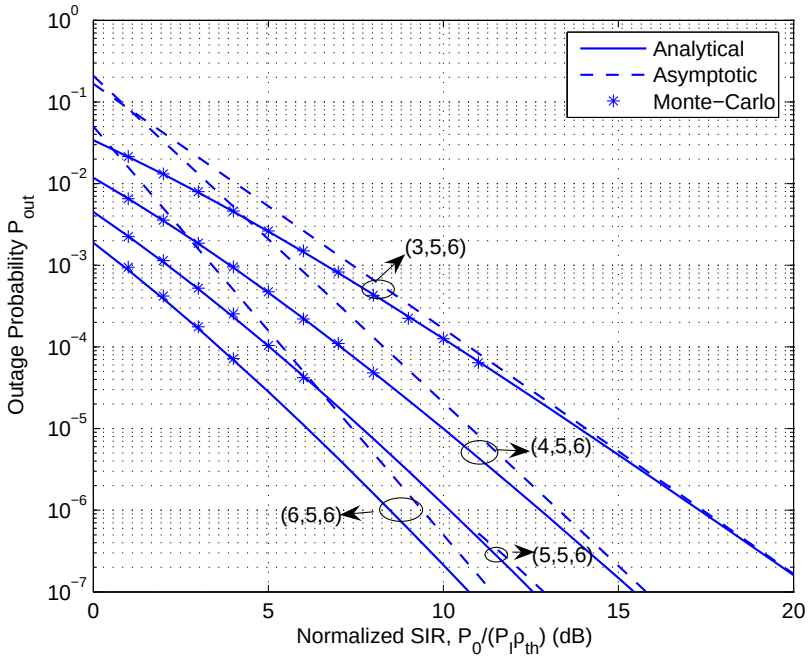

Figure 7.5. Outage probability $\mathcal{P}_{\text{out}}$ versus the normalized SIR $P_0/(P_{\mathcal{I}}\rho_{\text{th}})$ in decibels for keyhole MIMO channels for various number of transmit antennas N_t for 1) the exact analytical expressions, 2) the asymptotic expansions, and 3) the Monte-Carlo simulation results. In this figure, the notation $(N_t, N_r, N_{\mathcal{I}})$ has been used.

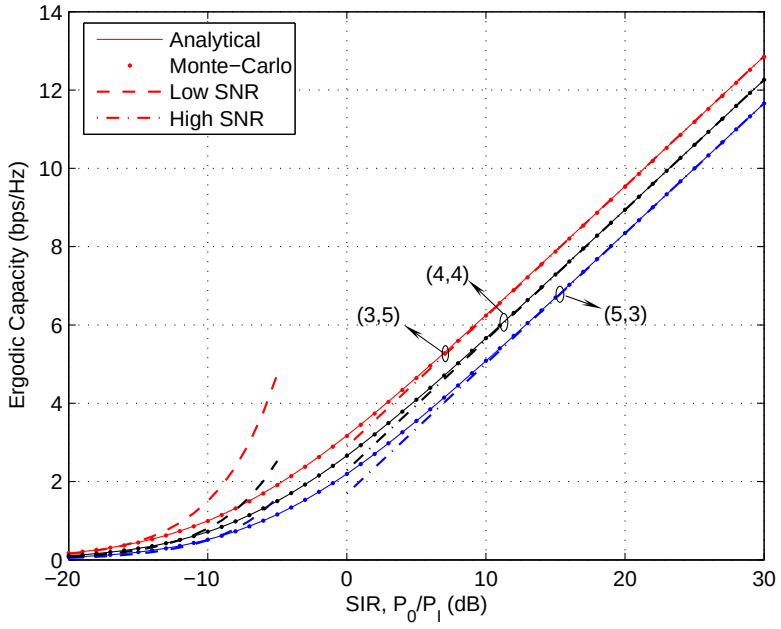

Figure 7.6. Ergodic capacity C versus SIR $P_0/P_{\mathcal{I}}$ in decibels for keyhole MIMO channels for different numbers of transmit and receive antennas with the same total number of antennas. In this figure, the notation (N_t, N_r) has been used and $N_{\mathcal{I}} = 6$ is assumed.

results match perfectly together while the exact results converge to the asymptotic results at high SIR values, which permits the asymptotic results to be used for the derivation of the channel diversity order

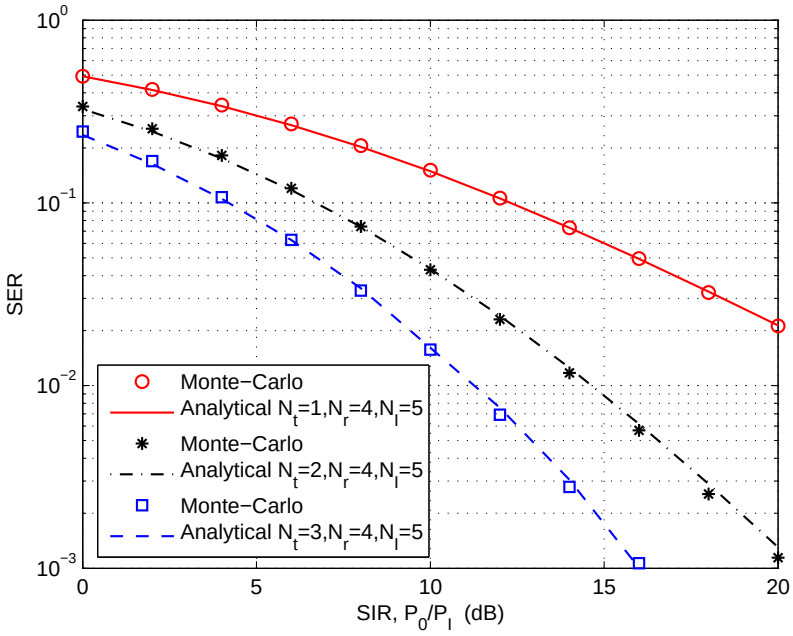

Figure 7.7. SER versus SIR $P_0/P_{\mathcal{I}}$ in decibels for 8-PSK keyhole MIMO channels for various number of the transmit antennas N_t.

[which equals $s = \min(N_t, N_r)$].

Figure 7.6 plots the ergodic capacity of keyhole channels for various number of transmit and receive antennas while the total number of antennas at both ends is kept to be 8 when $N_{\mathcal{I}} = 6$. Again, the results confirm that the analytical results are correct and match precisely with the Monte-Carlo results. On the other hand, results demonstrate that it is preferable to deploy more antennas at the receiver side than the transmitter side for maximizing the channel ergodic capacity.

Finally, Figure 7.7 plots the SER of keyhole channels for different antenna configurations when $N_{\mathcal{I}} = 5$ and coherent 8-PSK modulation ($\alpha = 2, \beta = 0.146$) is assumed. Results show a perfect agreement with the analytical and the Monte Carlo simulation results.

7.5 Conclusion

In this chapter, an analytical characterization of the performance of Rayleigh-product MIMO channels (a special case of double-scattering) using OC with co-channel interference was presented. With the interference-limited assumption of equal-power interferers, we have derived new closed-form expression for the outage probability of the OC systems operating in interference-limited Rayleigh-product channels. We have also developed a set of new results for an interference-limited keyhole MIMO channel, which includes the ergodic capacity, the outage probability and the SER, all in closed form. The findings suggest that co-channel interference does not affect the diversity order of the system, but instead, it degrades the outage performance by introducing a loss in the array gain.

Chapter 8

Low SNR Capacity Analysis of General MIMO Channels with Single Interferer

8.1 Introduction

A wide variety of digital communication systems (e.g., wireless sensor networks) operate in the low-power region where both spectral efficiency and the energy-per-bit are relatively low. In [95], Verdú proposed that the spectral efficiency in the low SNR (or wideband) regime can be analyzed through two parameters; namely, the minimum E_b/N_0 (with E_b denoting the average energy per information bit and N_0 being the noise spectral density) required for reliable communications, and the wideband slope (denoted by S_0). These low SNR metrics provide a useful reference in understanding the achievable rate at low SNRs, and have subsequently been elaborated in [39, 41, 56, 77, 93, 109], where the impacts of Rician K factor, spatial correlation, transmit and receiver CSI were investigated.

On the other hand, due to spectrum scarcity, communication systems are anticipated to be corrupted by interference. Therefore, it is of practical interest to investigate the low SNR capacity of MIMO systems in the presence of co-channel interference. Prior works on this topic were very limited in that explicit expressions for these two low SNR metrics were only derived for the interference-limited Rayleigh fading channels, the corresponding results for MIMO Rician fading channels appear to be limited [93].

Motivated by this, in this chapter, we first investigate the MIMO Rician fading channels with arbitrary mean matrix and K factor, in which we derive exact expressions for $E_b/N_{0_{\min}}$ and S_0 in the presence of both interference and noise. Based on these, we further study the special cases, namely, the Rician MISO channels, the rank-one deterministic channels and the MIMO Rayleigh channels, in which simple expressions can be obtained to illustrate the impacts of the number of transmit and receive antennas, the Rician K factor, the channel mean matrix, and the INR on the capacity. Also, asymptotic results in the large-system limit and high INR are developed. In addition, we provide the low SNR capacity analysis for Rayleigh-product MIMO fading channels [25] with interference.

8.2 System Model

Consider a communication link with N_t transmit and N_r receive antennas, corrupted by interference and AWGN. The received signals, $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$, can be expressed as[1]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{h}s + \mathbf{n}, \tag{8.1}$$

where $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the transmitted symbol vector satisfying $\mathtt{E}\{\|\mathbf{x}\|^2\} = P$, s is the interference symbol such that $\mathtt{E}\{|s|^2\} = P_I$, $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is the noise vector with i.i.d. entries following $\mathcal{CN}(0, N_0)$, and $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the MIMO channel between the transmitter and receiver while $\mathbf{h} \in \mathbb{C}^{N_r \times 1}$ denotes the channel vector between the interferer and the desired receiver.

In this paper, we investigate the low SNR capacity properties of two important MIMO channel models, namely: 1) Rician fading and 2) Rayleigh-product fading. They are described as follows:

- **Rician MIMO channels**—In this case, the channel matrix has the structure [20]

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\mathbf{H}_0 + \sqrt{\frac{1}{K+1}}\mathbf{H}_w, \tag{8.2}$$

  where K denotes the Rician K-factor, and $\mathbf{H}_w \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix containing i.i.d. zero-mean unit-variance complex Gaussian entries. On the other hand, $\mathbf{H}_0 \in \mathbb{C}^{N_r \times N_t}$ denotes the channel mean matrix, which is normalized to satisfy

$$\mathsf{tr}\left\{\mathbf{H}_0\mathbf{H}_0^\dagger\right\} = N_r N_t. \tag{8.3}$$

- **Rayleigh-product MIMO channels**—The channel matrix $\mathbf{H}$ can be expressed as [25]

$$\mathbf{H} = \frac{1}{\sqrt{N_s}}\mathbf{H}_1\mathbf{H}_2, \tag{8.4}$$

  where $\mathbf{H}_1 \in \mathbb{C}^{N_r \times N_s}$ and $\mathbf{H}_2 \in \mathbb{C}^{N_s \times N_t}$ are statistically independent matrices containing i.i.d. zero-mean unit-variance complex Gaussian entries, with N_s denoting the number of effective scatterers. By varying N_s, this model can describe various rank-deficient effects of a MIMO channel, e.g., it degenerates to Rayleigh fading if $N_s \to \infty$, and a keyhole channel if $N_s = 1$.

We assume that CSI is not known at the transmitter side but perfectly known at the receiver. Thus, an equal-power allocation policy is employed and the ergodic capacity is then expressed as [56]

$$C = \mathtt{E}\left\{\log_2 \mathsf{det}\left(\mathbf{I} + \frac{P}{N_0 N_t}\mathbf{H}^\dagger\left(\frac{P_I}{N_0}\mathbf{h}\mathbf{h}^\dagger + \mathbf{I}\right)^{-1}\mathbf{H}\right)\right\}. \tag{8.5}$$

[1] For mathematic tractability, we assume a single co-channel interferer.

As pointed out in [56], with co-channel interference, it is more suitable to define the SNR as

$$\rho \triangleq \left(\frac{P}{N_0}\right) \frac{\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{H}^\dagger(\rho_I \mathbf{h}\mathbf{h}^\dagger + \mathbf{I})^{-1}\mathbf{H}\right\}\right\}}{N_t N_r}, \tag{8.6}$$

where $\rho_I \triangleq \frac{P_I}{N_0}$ is regarded as the INR. This definition is different from the conventional one which defines the SNR as the average signal power divided by the noise power. The reason is that in the presence of interference, the noise is generally colored and SNR should be defined by averaging the SNRs along each of the principal directions of the noise space.

Based on the above definitions, the ergodic capacity expression in (8.5) can be rewriten as

$$C(\rho) = \mathtt{E}\left\{\log_2 \det\left(\mathbf{I} + \frac{\rho \mathbf{H}^\dagger \left(\rho_I \mathbf{h}\mathbf{h}^\dagger + \mathbf{I}\right)^{-1}\mathbf{H}}{\frac{1}{N_r}\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{H}^\dagger(\rho_I \mathbf{h}\mathbf{h}^\dagger + \mathbf{I})^{-1}\mathbf{H}\right\}\right\}}\right)\right\}. \tag{8.7}$$

At low SNRs, it has proved useful to investigate the capacity in terms of the normalized transmit energy per information bit, E_b/N_0, rather than the per-symbol SNR, ρ. This capacity can be well approximated for low E_b/N_0 levels by the following expression [95]

$$\mathtt{C}\left(\frac{E_b}{N_0}\right) \approx S_0 \log_2\left(\frac{\frac{E_b}{N_0}}{\frac{E_b}{N_0}_{\min}}\right), \tag{8.8}$$

in which $\frac{E_b}{N_0}_{\min}$ denotes the minimum energy per information bit required to convey any positive rate reliably and S_0 is the wideband slope [56, 95]. These are the two key parameters that dictate the capacity behavior in the low SNR regime, and can be obtained from $C(\rho)$ via [56][2]

$$\frac{E_b}{N_0}_{\min} = \frac{N_t N_r}{\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{H}^\dagger(\rho_I \mathbf{h}\mathbf{h}^\dagger + \mathbf{I})^{-1}\mathbf{H}\right\}\right\}} \frac{1}{\dot{C}(0)}, \tag{8.9}$$

$$S_0 = -\frac{2\left[\dot{C}(0)\right]^2}{\ddot{C}(0)} \ln 2, \tag{8.10}$$

where $\dot{C}(\cdot)$ and $\ddot{C}(\cdot)$ represent, respectively, the first- and second-order derivatives taken with respect to ρ. Note that $\mathtt{C}\left(\frac{E_b}{N_0}\right)$ implicity captures the second-order behavior of $C(\rho)$ as $\rho \to 0$.

[2]To facilitate the comparison to the interference free results, we adopt a slightly different definition of $\frac{E_b}{N_0}_{\min}$ from that in [56]. Specifically, in [56], E_b is normalized by the interference energy plus the noise energy while here E_b is normalized by the noise energy only. Therefore, the final result of $\frac{E_b}{N_0}_{\min}$ differs by a factor of $\rho_I + 1$.

8.3 Preliminaries

Here, we present some statistical results that will be used frequently in the following sections. For notational convenience, we define

$$D_1(m,t) \triangleq t^{-m}\Psi\left(m,m,t^{-1}\right), \tag{8.11}$$

$$D_2(m,t) \triangleq t^{-m}\Psi\left(m,m-1,t^{-1}\right), \tag{8.12}$$

where $\Psi(\cdot,\cdot,\cdot)$ is the confluent hypergeometric function defined in [26].

Lemma 8.1 *For any $m \times 1$ vector $\mathbf{h} \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$, and positive number t, let $\mathbf{\Lambda} \triangleq (t\mathbf{h}\mathbf{h}^\dagger + \mathbf{I})^{-1}$. Then, we have*

$$\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{\Lambda}\right\}\right\} = m - 1 + D_1(m,t), \tag{8.13}$$

$$\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{\Lambda}^2\right\}\right\} = m - 1 + D_2(m,t), \tag{8.14}$$

$$\mathtt{E}\left\{\mathtt{tr}^2\left\{\mathbf{\Lambda}\right\}\right\} = (m-1)^2 + 2(m-1)D_1(m,t) + D_2(m,t). \tag{8.15}$$

Proof: The result can be obtained by noting the unitary invariant of vector $\mathbf{h}$, and using the integration formula [26, (3.385.5)]. □

Lemma 8.2 *For any $m \times n$ matrix $\mathbf{H} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}\otimes\mathbf{I})$, $m \times 1$ vector $\mathbf{h} \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$, and positive constant t, we have*

$$\mathtt{E}\{\mathtt{tr}\{\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H}\}\} = n(m-1) + nD_1(m,t), \tag{8.16}$$

$$\mathtt{E}\{\mathtt{tr}\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\}\} = n(m-1)(n+m-1) + (n^2+n)D_2(m,t) + 2n(m-1)D_1(m,t), \tag{8.17}$$

$$\mathtt{E}\{\mathtt{tr}^2\{\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H}\}\} = n(m-1)(mn-n+1) + (n^2+n)D_2(m,t) + 2(m-1)n^2D_1(m,t), \tag{8.18}$$

where $\mathbf{\Lambda}$ has been defined in Lemma 8.1.

Proof: See Appendix E.1. □

Lemma 8.3 *When $m \to \infty$ or $t \to \infty$, we have $D_1(m,t) = D_2(m,t) = 0$. On the other hand, if $t \to 0$, $D_1(m,t) = D_2(m,t) = 1$.*

Proof: See Appendix E.2. □

8.4 Low SNR Capacity Analysis of Rician MIMO Channels with Single Interferer

In this section, we derive analytical expressions for $E_b/N_{0_{\min}}$ and the wideband slope, S_0, for Rician MIMO fading channels. The main result is given in the following theorem.

Theorem 8.1 *For MIMO Rician fading channels with a single interferer, we have*

$$\frac{E_b}{N_{0\,\min}} = \frac{\ln 2}{N_r - 1 + A_0}, \tag{8.19}$$

$$S_0 = \frac{2N_t(K+1)^2}{2K + B_0}, \tag{8.20}$$

where A_0 and B_0 are, respectively, given by

$$A_0 = D_1(N_r, \rho_I), \tag{8.21}$$

and

$$\begin{aligned} B_0 = \frac{1}{(N_r - 1 + D_1(N_r,\rho_I))^2} \Bigg[& \left(\frac{2K^2 \left(\mathsf{tr}\left\{ (\mathbf{H}_0\mathbf{H}_0^\dagger)^2 \right\} - N_t^2 N_r \right)}{N_t(N_r+1)} + 2(N_r - 1) \right) D_1(N_r,\rho_I) \\ & + \left(1 + (1+2K)N_t + \frac{K^2 \left(\mathsf{tr}\left\{ (\mathbf{H}_0\mathbf{H}_0^\dagger)^2 \right\} + N_t^2 N_r^2 \right)}{N_t N_r (N_r+1)} \right) D_2(N_r,\rho_I) \\ & + (N_r - 1) \left(N_r - 1 + (1+2K)N_t + \frac{K^2 \left(\mathsf{tr}\left\{ (\mathbf{H}_0\mathbf{H}_0^\dagger)^2 \right\} (N_r^2 - N_r - 1) + N_t^2 N_r^2 \right)}{N_t N_r (N_r^2 - 1)} \right) \Bigg]. \end{aligned} \tag{8.22}$$

Proof: See Appendix E.3. □

Theorem 8.1 is general and valid for mean matrix of arbitrary rank, $\mathbf{H}_0$, and any possible N_t, N_r, K and ρ_I. From (8.19), we observe that the Rician factor K and the structure of channel mean $\mathbf{H}_0$ (as long as $\mathsf{tr}\left\{\mathbf{H}_0\mathbf{H}_0^\dagger\right\} = N_t N_r$) do not affect $E_b/N_{0_{\min}}$, while the values of N_r and ρ_I have a direct impact. Also in (8.20), we see that all the parameters will affect the wideband slope S_0.

Based on (8.19), we can further investigate the impact of N_r and ρ_I on $E_b/N_{0_{\min}}$ as follows.

Corollary 8.1 *The $E_b/N_{0_{\min}}$ is a decreasing function of N_r (i.e., when N_r increases, $E_b/N_{0_{\min}}$ decreases) and is an increasing function of ρ_I (i.e., when ρ_I increases, $E_b/N_{0_{\min}}$ increases). Moreover, the increase in $E_b/N_{0_{\min}}$ due to interference is upper bounded by $\frac{\ln 2}{N_r(N_r-1)}$ for $N_r \geq 2$.*

Proof: See Appendix E.4. □

Corollary 8.2 *When $\rho_I \to 0$, Theorem 8.1 reduces to*

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r}, \tag{8.23}$$

$$S_0 = \frac{2(K+1)^2}{\frac{K^2 \mathsf{tr}\left\{(\mathbf{H}_0\mathbf{H}_0^\dagger)^2\right\}}{N_t^2 N_r^2} + (1+2K)\frac{N_t+N_r}{N_t N_r}}. \tag{8.24}$$

In particular, if the channel mean matrix, $\mathbf{H}_0$, is of rank-1, then S_0 can be reduced to

$$S_0 = \frac{2(K+1)^2}{K^2 + (1+2K)\frac{N_t+N_r}{N_t N_r}}. \tag{8.25}$$

Proof: The results can be obtained with the help of Lemma 3, together with the fact that $\mathsf{tr}\left\{(\mathbf{H}_0\mathbf{H}_0^\dagger)^2\right\} = N_t^2 N_r^2$ when $\mathbf{H}_0$ is of rank-1. □

Corollary 8.2 corresponds to the results for Rician MIMO fading channels in an interference-free environment, which generalizes the results in [56] where a rank-1 channel mean was considered.

To gain further insight, in the following, we look at three special cases: 1) Rician MISO fading channels, i.e., $N_r = 1$, 2) Rician MIMO channels of rank-1 mean in the large K regime, i.e., $K \to \infty$ and $\mathbf{H}_0 = \boldsymbol{\alpha}\boldsymbol{\beta}^\dagger$ (with complex column vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$), and 3) Rayleigh MIMO channels, i.e., $K = 0$.

8.4.1 Rician MISO Channels

Corollary 8.3 *For Rician MISO channels, i.e., $N_r = 1$, we have*

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{D_1(1,\rho_I)}, \tag{8.26}$$

$$S_0 = \frac{2N_t(K+1)^2 D_1(1,\rho_I)^2}{2K + [1+N_t(1+K)^2]D_2(1,\rho_I)}. \tag{8.27}$$

Proof: The result can be obtained by substituting $N_r = 1$ in Theorem 8.1. □

Corollary 8.4 *When $N_r = 1$, S_0 is an increasing function of N_t. When $0 \le K < 1 - D_2(1,\rho_I)$, S_0 is a decreasing function of K, while for $K \ge 1 - D_2(1,\rho_I)$, S_0 is an increasing function of K.*

Proof: See Appendix E.5. □

In contrast to the interference-free case, where the increase of Rician factor K always improves the wideband slope S_0 when $N_r = 1$, Corollary 8.4 reveals that the impact of K on S_0 depends on the interference level. Moreover, when $\rho_I \to \infty$, $S_0 = 0$ which aligns with the observations in [56] for interference-limited Rayleigh fading scenarios. However, the general impact of ρ_I on S_0 is more difficult to characterize, though simulation results indicate that S_0 decreases when ρ_I increases.

8.4.2 Rank-1 Mean Rician MIMO Channels for Large K

Corollary 8.5 *In the large K regime, for rank-1 mean Rician MIMO channels with a single interferer, it can be derived that*

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1 + A_0}, \tag{8.28}$$

$$S_0 = \frac{2(N_r+1)(N_r - 1 + D_1(N_r, \rho_I))^2}{N_r[N_r(N_r-1) + 2(N_r-1)D_1(N_r,\rho_I) + 2D_2(N_r,\rho_I)]}, \tag{8.29}$$

Proof: The desired results can be obtained by taking the limit $K \to \infty$ in Theorem 8.1. □

Corollary 8.5 indicates that in the large K regime, for rank-1 mean Rician MIMO fading channels, multiple transmit antennas are irrelevant in terms of $E_b/N_{0_{\min}}$ and S_0. This is actually an intuitive result. The reason is that the large K regime corresponds to the non-fading channel scenarios, and thus, varying the number of transmit antennas for a fixed total transmit power will not increase the receive signal energy and will not contribute to the capacity. In addition, N_r affects both $E_b/N_{0_{\min}}$ and S_0 in contrast to the interference-free case where N_r is only relevant in terms of $E_b/N_{0_{\min}}$ [56].

With the help of Lemma 8.3, we can further obtain the results in various asymptotic regimes:

- When $\rho_I \to 0$, Corollary 8.5 reduces to

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r}, \tag{8.30}$$

$$S_0 = 2. \tag{8.31}$$

  The above results correspond to the interference-free scenario, and conforms to those in [56].

- When $N_r \to \infty$, Corollary 8.5 reduces to

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1}, \tag{8.32}$$

$$S_0 = 2\left(1 - \frac{1}{N_r^2}\right) \approx 2. \tag{8.33}$$

  Compared with the interference-free case, the above results suggest that interference degrade the capacity performance by increasing $E_b/N_{0_{\min}}$ and decreasing S_0. The loss in $E_b/N_{0_{\min}}$ can be explained by the fact that one receive antenna is dedicated to suppress the single-antenna interference, while the rest, $N_r - 1$ antennas, contribute to normal communication.

- When $\rho_I \to \infty$ and $N_r \geq 2$, Corollary 8.5 reduces to

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1}, \tag{8.34}$$

$$S_0 = 2\left(1 - \frac{1}{N_r^2}\right). \tag{8.35}$$

Intriguingly, these results coincide with the case $N_r \to \infty$. Nevertheless, it is worth mentioning that the situations in application are very different. One is applicable for large N_r but arbitrary interference power ρ_I, while the other is valid for large ρ_I but arbitrary N_r.

8.4.3 Rayleigh MIMO Channels

Corollary 8.6 *For Rayleigh MIMO channels with a single interferer, we have*

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1 + A_0}, \tag{8.36}$$

$$S_0 = \frac{2N_t}{1 + B_1}, \tag{8.37}$$

where B_1 is defined as

$$B_1 \triangleq \frac{N_t(N_r - 1) + (N_t + 1)D_2(N_r, \rho_I) - D_1(N_r, \rho_I)^2}{[N_r - 1 + D_1(N_r, \rho_I)]^2}. \tag{8.38}$$

Proof: The results follow immediately by substituting $K = 0$ into Theorem 8.1. □

Corollary 8.6 shows that the number of transmit antennas affects the capacity performance through S_0. More insights can be gained by looking into the asymptotic regimes as follows.

- When $\rho_I \to 0$, we have

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r}, \tag{8.39}$$

$$S_0 = \frac{2N_tN_r}{N_t + N_r}. \tag{8.40}$$

  This scenario corresponds to the case for Rayleigh MIMO channels without interference, and the results are consistent with those derived in [56].

- When $N_r \to \infty$, we have

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1}, \tag{8.41}$$

$$S_0 = \frac{2N_t(N_r - 1)}{N_t + N_r - 1}. \tag{8.42}$$

- When $\rho_I \to \infty$ and $N_r \geq 2$, we have

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1}, \tag{8.43}$$

$$S_0 = \frac{2N_t(N_r - 1)}{N_t + N_r - 1}. \tag{8.44}$$

  Similar to the case of rank-1 mean Rician MIMO channels with a large K, it is observed that the results for $N_r \to \infty$ and $\rho_I \to \infty$ coincide. In addition, by comparing the above results

to the interference-free results, we see that in Rayleigh fading, $E_b/N_{0_{\min}}$ and S_0 for a MIMO channel with a single interferer behaves like a channel with one less receive antenna operating in an interference-free environment, which is different from the large K rank-1 mean Rician MIMO channel case where S_0 does not have this interpretation.

8.5 Low SNR Capacity Analysis of Rayleigh-Product MIMO Channels with Single Interferer

In this section, we develop the low SNR capacity results for Rayleigh-product MIMO channels.

Theorem 8.2 *For Rayleigh-product MIMO channels with a single interferer, we have*

$$\frac{E_b}{N_{0\,\min}} = \frac{\ln 2}{N_r - 1 + A_0}, \tag{8.45}$$

$$S_0 = \frac{2N_tN_s}{N_t + N_s + B_2}, \tag{8.46}$$

where A_0 has been defined in Theorem 8.1 and B_2 is given by

$$B_2 \triangleq \frac{(N_r-1)(N_tN_s+1) + (N_t+1)(N_s+1)D_2(N_r,\rho_I) - (N_s+N_t)D_1(N_r,\rho_I)^2}{[N_r - 1 + D_1(N_r,\rho_I)]^2}. \tag{8.47}$$

Proof: See Appendix E.6. □

Theorem 8.2 shows that the $E_b/N_{0_{\min}}$ for Rayleigh-product MIMO channels is the same as that for Rician MIMO fading channels, although the two channels have very different information-carrying capabilities. As such, the results of Corollary 8.1 also apply for Rayleigh-product channels. Nonetheless, this is not surprising as has already been reported in [95], and this is the consequence of the noise being additive Gaussian. This explains that $E_b/N_{0_{\min}}$ is not sufficient to indicate the capacity performance and motivates the need for higher order approximation of the capacity such as the wideband slope, S_0, which is generally different for different channels. In addition, it is observed that N_t and N_s affect the capacity performance through the wideband slope S_0 but not $E_b/N_{0_{\min}}$.

Corollary 8.7 *When $N_s \to \infty$, the wideband slope for Rayleigh-product fading with a single interferer becomes the same as that for Rayleigh fading scenarios.*

Proof: The corollary can be proved by noting that $\frac{B_2}{N_s}|_{N_s\to\infty} = B_1$. □

The above corollary indicates that the Rayleigh-product channels converges to a Rayleigh fading channel when $N_s \to \infty$. This result is quite intuitive since the large N_s correspondes to a rich scattering enviornment which is the scenario that fits well with the Rayleigh fading model.

The following asymptotic cases are looked at to gain further understanding.

- When $\rho_I \to 0$, we have

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r}, \tag{8.48}$$

$$S_0 = \frac{2N_tN_sN_r}{N_tN_s + N_rN_s + N_tN_r + 1}. \tag{8.49}$$

This scenario corresponds to the interference-free case for Rayleigh-product channels whose results have been derived in [85]. In addition, when $N_s = 1$, we further have

$$S_0 = \frac{2N_tN_r}{(N_t + 1)(N_r + 1)} \tag{8.50}$$

which provides the wideband slope for keyhole channels.

- When $N_r \to \infty$, we have

$$\frac{E_b}{N_0}_{\min} = \frac{\ln 2}{N_r - 1}, \tag{8.51}$$

$$S_0 = \frac{2N_tN_s(N_r - 1)}{N_tN_s + (N_r - 1)(N_s + N_t) + 1}. \tag{8.52}$$

- When $\rho_I \to \infty$ and $N_r \geq 2$, it can be easily shown that $E_b/N_{0_{\min}}$ and S_0 are, respectively, given by (8.51) and (8.52). In other words, the results for $N_r \to \infty$ and $\rho \to \infty$ coincide. Additionally, similar to Rayleigh MIMO channels, the penalty of having an interferer is illustrated through a reduction on the number of effective receive antennas by 1.

8.6 Numerical Results

In this section, we perform various simulations to further examine the derived analytical expressions. All the Monte-Carlo simulation results were obtained by averaging over 10^5 independent channel realizations. For MIMO Rician channels, the mean matrix is generated according to [10]

$$\mathbf{H}_0 = \sum_{l=1}^{L} \beta_l \boldsymbol{\alpha}(\theta_{r,l})\boldsymbol{\alpha}(\theta_{t,l})^T, \tag{8.53}$$

where β_l is the complex amplitude of the lth path, and $\boldsymbol{\alpha}(\theta_{t,l})$ and $\boldsymbol{\alpha}(\theta_{r,l})$ are the specular array responses corresponding to the lth dominant path at the transmitter and receiver, respectively. The array response is defined as $[1, e^{j2\pi d\cos(\theta)}, \cdots, e^{j2\pi d(N-1)\cos(\theta)}]^T$ where d is the antenna spacing in wavelengths. In all simulations, we assume that $d = 0.5$ at both the transmit and receive sides.

For 3×2 MIMO Rician channels, the mean matrix is constructed by assuming that there are two dominant paths (i.e., $L = 2$), with the arriving and departure angles given by $\theta_{r,1} = \theta_{t,1} = \frac{\pi}{2} + \frac{\pi}{8}$, $\theta_{r,2} = \theta_{t,2} = \frac{\pi}{2} - \frac{\pi}{8}$,[3] respectively. The complex coefficient β_l is chosen such that $\mathtt{tr}\{\mathbf{H}_0\mathbf{H}_0^\dagger\} = N_tN_r$. For rank-1 mean Rician fading MIMO channels, we assume $L = 1$, $\beta_1 = 1$ and $\theta_{r,1} = \theta_{t,1} = \frac{\pi}{2}$.

[3]These angles are randomly chosen for simulation purpose, and our results are applicable to arbitrary angles.

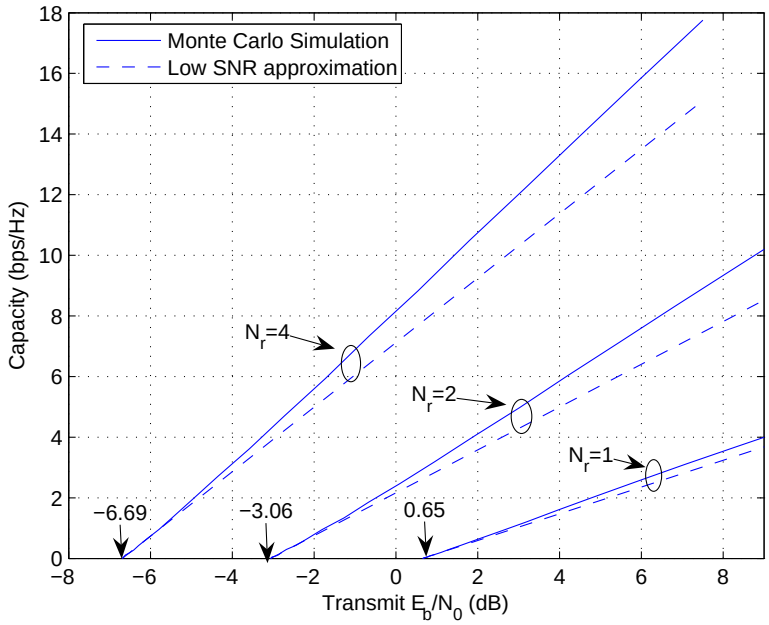

Figure 8.1. Low SNR capacity versus transmit E_b/N_0 for various N_r when $N_t = 3$ and $\rho_I = 0$ dB.

Fig. 8.1 investigates the impact of N_r on $E_b/N_{0\min}$. To isolate the effect of N_r, we have set $K = 0$ to eliminate the possible impact from channel mean matrix $\mathbf{H}_0$. From the results of Fig. 8.1, it can be seen that the increase of N_r helps to reduce the required $E_b/N_{0\min}$, which confirms the analysis of Corollary 8.1. Moreover, we observe that when N_r increases, so does the wideband slope S_0, which indicates the double benefits of increasing N_r. In addition, when compared with the Monte-Carlo simulation results, the analytical results show very high accuracy in terms of $E_b/N_{0\min}$, and also the wideband slope S_0 if the SNR of interest is sufficiently low, i.e., below 2 bps/Hz of capacity.

In Fig. 8.2, results for the low SNR capacity approximation are plotted for 3×2 Rician MIMO channels with $K = 1$. Results reveal a good agreement between the analysis and the simulations. We also see that the increase in the interference power degrades the capacity performance by increasing the required $E_b/N_{0\min}$, while the impact on S_0 is not so pronounced. Furthermore, the increase in $E_b/N_{0\min}$ from a channel without interference to that with a 10 dB of INR is about 0.1, which appears to be very close to the upper bound we obtained in Corollary 8.1 $(\ln 2)/(N_r(N_r - 1)) = 0.115$.

Results in Fig. 8.3 are provided for the capacity of 3×2 MIMO Rician channels for different Rician-K factors in the low SNR regime according to Theorem 8.1. The curves indicate the accuracy of our analytical expression and that the range for a good approximation improves if K increases. In particular, the approximation is very good for the capacity range from 0 to 10, when $K = 100$. Also, results demonstrate that the Rician K factor affects the capacity performance through the wideband slope S_0 but not the $E_b/N_{0\min}$, and more specifically, the wideband slope S_0 increases when K becomes larger. However, the increase is not very substantial. On the other hand, Fig. 8.4 plots the results for rank-1 mean 3×2 MIMO Rician channels in the large K regime both with and without interference. Results

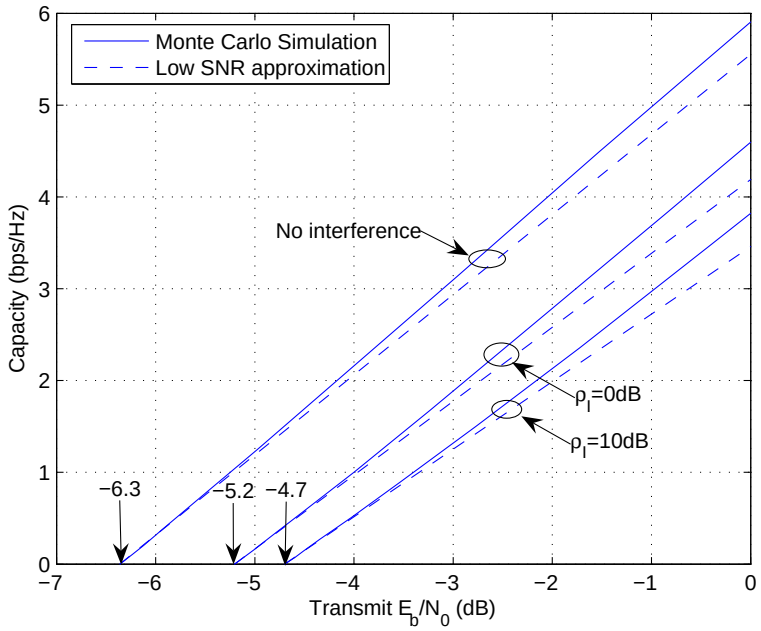

Figure 8.2. Low SNR capacity versus transmit E_b/N_0 for various ρ_I when $N_t = 2$ and $N_r = 3$.

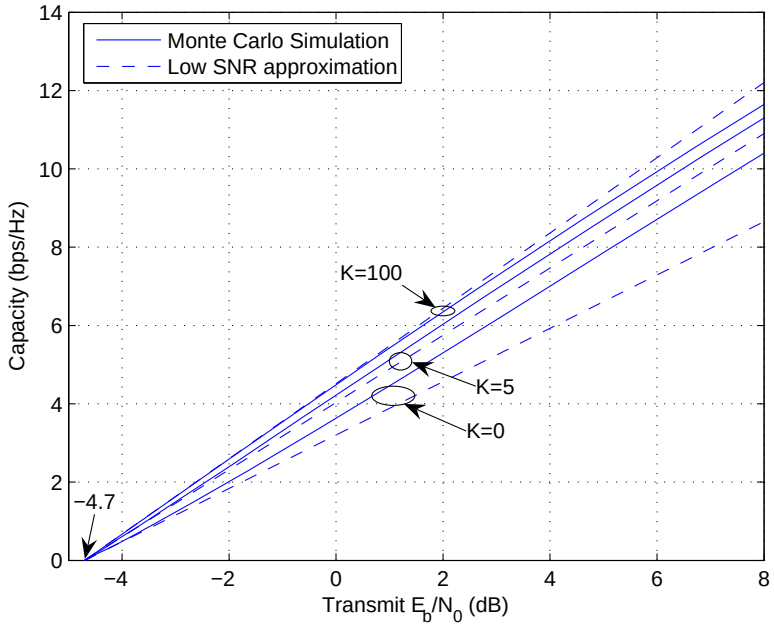

Figure 8.3. Low SNR capacity versus transmit E_b/N_0 for various Rician factor K when $N_t = 2$, $N_r = 3$ and $\rho_I = 10$ dB.

confirm the correctness of the analytical results in Corollary 8.5.

Results in Fig. 8.3 are provided for the capacity of 3×2 Rician fading MIMO channels for different Rician-K factors in the low SNR regime according to Theorem 8.1. The curves indicate the accuracy of our analytical expression and that the range for a good approximation improves if K increases. In particular, the approximation is very good for the capacity range from 0 to 10, when $K = 100$. Also,

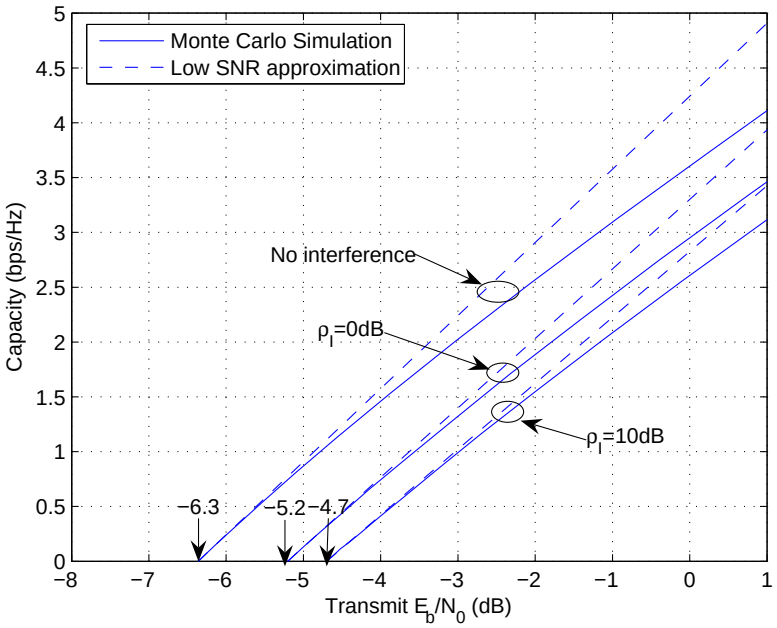

Figure 8.4. Low SNR capacity versus transmit E_b/N_0 for rank-1 mean Rician fading MIMO channels when $N_t = 2$, $N_r = 3$ and $K = 100$.

results demonstrate that the Rician K factor affects the capacity performance through the wideband slope S_0 but not the $E_b/N_{0\min}$, and more specifically, the wideband slope S_0 increases when K becomes larger. However, the increase is not very substantial. On the other hand, Fig. 8.4 plots the results for rank-1 mean 3×2 Rician MIMO channels in the large K regime both with and without interference. Results confirm the correctness of the analytical results in Corollary 8.5.

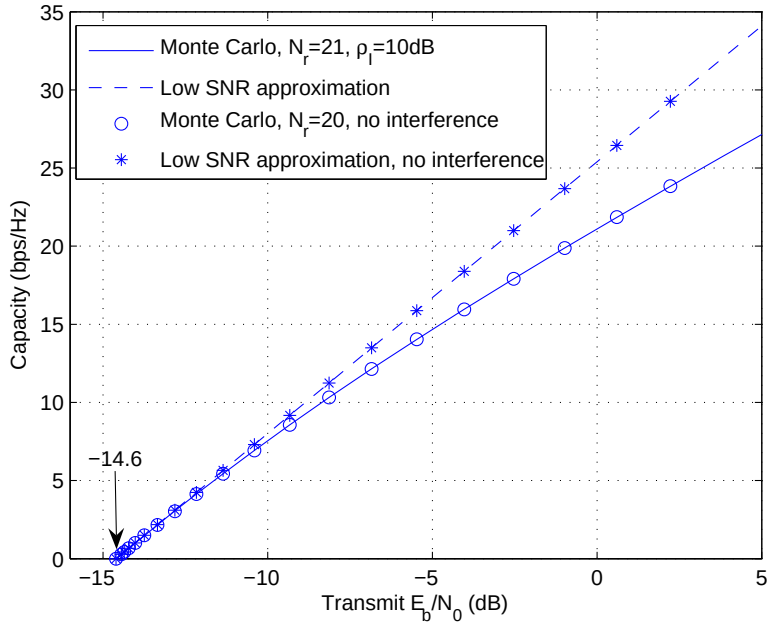

Figure 8.5. Low SNR capacity versus transmit E_b/N_0 when $N_t = 3$ and different N_r and ρ_I.

In Fig. 8.5, we provide the results for Rayleigh MIMO fading channels. Two system configurations are investigated: one for 3×21 channels with a single interferer of $\rho_I = 10$ dB, and the other for 3×20 channels without interference. As we can see, the results of the two systems almost overlap with inappreciable difference in the low SNR regime, which aligns with our analysis.

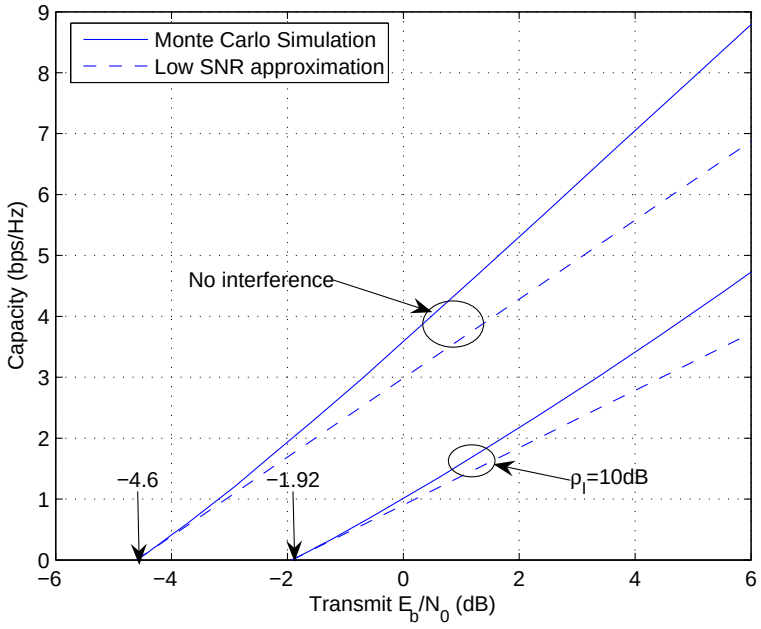

Figure 8.6. Low SNR capacity versus transmit E_b/N_0 for Rayleigh-product channel when $N_t = 3$, $N_s = 6$ and $N_r = 2$.

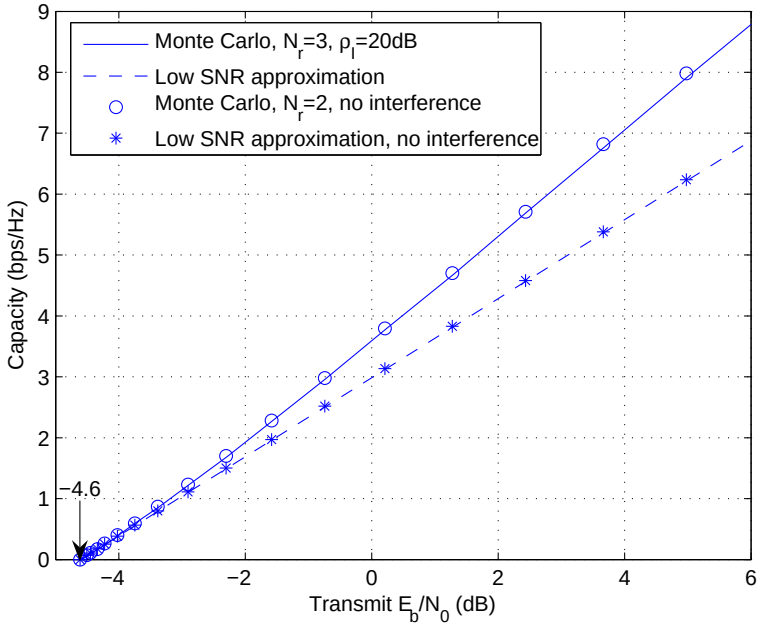

Figure 8.7. Low SNR capacity versus transmit E_b/N_0 for Rayleigh-product channel when $N_t = 2$, $N_s = 6$ and different N_r and ρ_I.

Results in Figs. 8.6 and 8.7 are provided for Rayleigh-product MIMO channels. Results demonstrate a good agreement between the analysis and simulations. Additionally, it is observed that the level of interference increases the required $E_b/N_{0\min}$ and reduces the wideband slope S_0. On the other hand, Fig. 8.7 plots the results for two systems both with $N_t = 2$ and $N_s = 6$: one with a single strong interferer of $\rho_I = 20$ dB and $N_r = 3$, and the other with $N_t = 2$ in an interference-free environment. Results for both systems overlap in the low SNR regime, which confirms our analysis.

8.7 Conclusion

This chapter has studied the capacity of Rician fading and Rayleigh-product MIMO channels with a single interferer in the low SNR regime. Exact expressions for the minimum energy per information bit, $E_b/N_{0\min}$, and the wideband slope, S_0, were derived for both channels. Also, we showed that interference degrades the capacity performance by increasing $E_b/N_{0\min}$ and reducing S_0. Moreover, the impact of other system parameters, such as the number of transmit and receive antennas, Rician factor K, the channel mean matrix $\mathbf{H}_0$, were investigated.

Chapter 9

Conclusions and Future Works

The theme of this thesis is on the capacity and performance analysis of MIMO antenna systems operating over general and practical propagation channels. There are primary two key aspects in the thesis: the first is to give a thorough investigation on the fundamental capacity limits of several important MIMO channels, while the second is to analyze in details for some practical transmission methods over matrix product channels. In the following, we summarize the main contributions and insights of this thesis, and discuss some possible lines for future works.

9.1 Summary of Contributions and Insights

In Chapter 3, a host of new finite RMT results were derived, which provided the essential mathematical tools for the performance analysis conducted in later chapters.

Chapter 4 presented a unified framework of developing capacity bounds for general MIMO Nakagami-m fading channels based on majorization theory. By exploiting the majorization relationships between the eigenvalues and diagonal elements of the random matrix of interest, we derived several capacity upper bounds and lower bounds for both C-MIMO and D-MIMO systems. Based on these analytical expressions, a number of insights were obtained:

- The ergodic capacity is a monotonic increasing function of the fading parameter m, i.e. if m becomes greater, the ergodic capacity increases. This finding is quite intuitive since a greater m corresponds to less severe fading, and the ergodic capacity is anticipated to increase with m.
- In the large system regime, the ergodic capacity of the system scales linearly with the minimum of the antenna numbers, and is independent of the fading parameter m. This again is an intuitive finding. Since whatever the value of m, the effect of channel fading could be completely eliminated with the increasing antenna number.
- Both the path loss and shadowing effects affect the ergodic capacity of the system. The path loss effect decreases the ergodic capacity as a function of the distance, while the shadowing effect

impacts the ergodic capacity through its mean fading parameter.

Chapter 5 considered the mutual information and outage performance of MIMO multi-keyhole channels. The results for mutual information were derived by directly invoking the new marginal p.d.f. expression of a product of random matrices (Thoerem 3.2), as well as certain random determinant (Theorem 3.8) and log-determinant results (Theorem 3.10) from Chapter 3, while the exact and approximate analytical outage expressions for MIMO MRC systems were obtained based on those maximum eigenvalue distributions, i.e., Theorem 3.3, and Theorem 3.4. From these analytical results, we obtain the following key insights:

- The mutual information of MIMO multi-keyhole channels is generally inferior to that of MIMO Rayleigh channels, and the number of keyholes will significantly affect the mutual information. For instance, when the keyhole number increases, the mutual information improves, and it gradually approaches the MIMO Rayleigh mutual information bound when the keyhole number is large enough.

- For MIMO MRC systems, the outage probability of multi-keyhole MIMO channels can be superior than that of MIMO Rayleigh channels at high outage regime (or equivalently, at sufficiently low SNR regime). However, for outage level of practical interest, (e.g., ¡ 0.1), achieving a given outage level requires lower SNR for MIMO Rayleigh channels compared with multi-keyhole channels.

- The keyhole power distribution does not affect the diversity order of the system, while the number of keyhole does.

- The impact of keyhole power distribution on the outage performance is characterized through Schur-concavity, and the findings suggest that the more equally distributed keyhole power, the better is the outage performance.

- By interpreting the keyhole power matrix as the correlation matrix, the results also apply to one-sided correlated Rayleigh-product channels. And for MIMO MRC system, this indicates that the higher the correlation of the channel, the worse the outage performance.

Chapter 6 investigated the capacity of MIMO dual-hop systems employing AF relay node. In contrast to prior results which primary focus on the asymptotic large antenna regime, we aimed at the finite antenna system. Therefore, the results are applicable for arbitrary number of transmit, relay and receive antennas. Exact capacity expression was derived by directly invoking Theorem 3.1 along with some basic mathematical manipulations, and tight capacity upper bound and lower bound were obtained based on Theorem 3.7 and Theorem 3.9, respectively. Moreover, several special cases were studied in great detail. From these analytical results, we see that the ergodic capacity of AF MIMO dual-hop systems is intimately related to that of single hop MIMO systems.

- The multiplexing gain of the system is one-half of the minimum of the numbers of transmit, relay and destination antennas, where the one-half comes from the fact that two time slots are consumed for the entire communication.
- When the number of relay antennas grows large, the ergodic capacity of AF MIMO dual-hop systems becomes one-half of that single hop MIMO systems with same transmit and receive antenna numbers but a properly adjusted SNR. This equivalent SNR is always smaller than the first hop SNR, which indicates that no matter how many relay antennas are employed and how much power is used, the capacity is always smaller than one-half of that single hop MIMO systems with the same transmit and receive antenna numbers with SNR being the first hop SNR.
- When the number of destination antennas grows large, the ergodic capacity of AF MIMO dual-hop systems becomes one-half of that of the first hop MIMO systems.
- When the relay power grows large, the ergodic capacity of AF MIMO dual-hop systems becomes one-half of that of a single hop MIMO systems with the same number of transmit antenna, but with the receive antenna number being the minimum of the relay and destination antenna numbers.

Chapter 7 examined the impact of co-channel interference in Rayleigh-product channels. The exact outage probability expressions for optimum combining systems were derived based on Theorem 3.5, and a detailed performance investigation on the special keyhole channels was carried out. From these analytical results, we have the following findings:

- The interference does not affect diversity order of the system, but it reduces the outage performance by contributing in the loss of array gain.
- The number of scatterers will significantly affect the outage performance, and the outage performance improves when the number of scatterers increases.
- For the special keyhole channels, the diversity order equals to the minimum of the transmit and receive antenna numbers.

Chapter 8 studied the ergodic capacity of general MIMO channels with a single interferer in the low SNR regime. For both Rician channels and Rayleigh-product channels, exact analytical expressions for the minimum energy per information bit, $E_b/N_{0_{\min}}$, and wideband slope, S_0, were derived. Several special cases were investigated in detail and the impact of transmit, receive antenna number, Rician K factor, channel mean matrix and INR on the ergodic capacity were analyzed. We gained the following insights:

- The minimum energy per information bit is the same for Rician channels and Rayleigh-product channels, while their wideband slopes differ significantly.
- Increasing the number of receive antennas helps to reduce the minimum energy per information bit required for reliable communication, while increasing the interference power requires more power to ensure reliable communication, and this extra power can be upper bounded.

- The structure of the channel mean matrix does not affect the ergodic capacity.
- For MISO Rician channels, the wideband slope is an increasing function of the transmit antenna number for small K, but is a decreasing function for large K.
- For large K and rank-1 mean Rician MIMO channels, the transmit antenna number does not affect the ergodic capacity.
- For Rayleigh-product channels, the number of scatters affects the ergodic capacity through the wideband slope but not the minimum energy per information bit.
- For both scenarios where either the number of receive antenna is large, or the interference power is large, the resulting capacity performance with one co-channel interference is equivalent to a system with the same number of transmit antennas but one less receive antenna in an interference free environment.

9.2 Future Works

In this section, we discuss several possible extensions of the problems investigated in this thesis.

A substantial portion of the thesis is devoted to the analysis of the newly emerged MIMO channels with product matrix structure. While the contributions made in the thesis have enhanced our knowledge of the fundamental limitations of the matrix product channels, nevertheless, much more theoretical works are needed to gain a thorough and deep understanding of the nature of the MIMO matrix product channels. We believe that matrix product channels will remain a fruitful area, and more important discoveries are expected to be made. In the following, we discuss some interesting directions.

The first extension is to consider more general settings. In particular, the MIMO multi-keyhole channel $\mathbf{H} = \mathbf{H}_1\mathbf{A}\mathbf{H}_2^\dagger$ studied in Chapter 5 is a special case of the more general model, referred to as the MIMO double scattering channel where the channel matrix is given by $\mathbf{H} = \mathbf{\Phi}_r^{1/2}\mathbf{H}_1\mathbf{\Phi}_s^{1/2}\mathbf{H}_2\mathbf{\Phi}_t^{1/2}$. It is of great interest to investigate the joint effect of transmit and receive correlation and rank-deficient phenomenon on the fundamental capacity of the system, as well as on the performance of various practical transmission schemes, i.e., OSTBC and transmit beamforming receiver combining systems.

Secondly, it is interesting to investigate the performance of practically appealing linear receiver system (i.e., ZF or MMSE receiver), as well as certain low complexity non-linear receiver system(i.e., ZF-DF or MMSE-DF receiver). While the performance of these simple receivers in a single MIMO Rayleigh fading channels has been extensively studied and well understood, there has been no available results in the literature for MIMO product channels. Therefore, it is important to look into this particular system, analytically characterize its performance, and compare it with the single MIMO Rayleigh channel case to gain more insights.

In addition, the case where partial CSI (i.e., channel mean, or channel correlation matrix) is available at the transmitter is also an important research topic. The partial CSI case has been well studied for the single full rank MIMO channels, where optimum precoding matrix and optimal power allocation schemes based on various performance metrics such as maximizing ergodic capacity, minimizing the outage probability, or minimizing the mean square error, have been derived for various scenarios. Making extension to the matrix product channels is therefore of great interest and importance, and it is expected that the rank deficiency phenomenon will play a key role in designing the system.

While the partial CSI is a more realistic assumption in most of the cases, limited feedback schemes are more attractive and practical, and it has received enormous attentions from the research community. In such systems, instead of feedback the exact channel matrix, only limited number of bits is feeded back to the transmitter, which greatly eases the demand and requirement on the feedback link. One particular popular limited feedback schemes is the antenna selection schemes, and its performance has been studied for two extreme cases, i.e., single full rank MIMO channels, degenerated single keyhole channels. To bridge the gap, it is therefore important to investigate the case with multiple keyholes. Some interesting questions arise naturally, i.e., how to select the transmit and receive antennas, whether it achieves the maximum diversity offered by the multi-keyhole channels.

Another interesting area to look into is the multi-hop communication systems, which has attracted enormous attentions from the research community due to its power saving and coverage extension advantages. As the MIMO technology becomes mature, it is expected that MIMO will be incorporated into the multi-hop system to form a MIMO multi-hop system. Therefore, analytical characterization of the performance of MIMO multi-hop system is a very important topic. For MIMO multi-hop channels, in essence, the effective channel of interest consists a product of multiple independent random matrices which is a generalization of the double-scattering channel model studied in this thesis. So far, only very limited asymptotic results are available in literature. It has been demonstrated that the asymptotic results fail to accurately capture the finite cases, which provides great motivation to investigate the performance of the system in the finite regime. The key challenge is to obtain the statistical properties of the resultant random product matrices, and it remains to be seen whether the conditional approach developed in the thesis can be applied in the general case.

In addition to the aforementioned point-to-point MIMO systems, another exciting area is the multi-user scenarios. Some preliminary results have been developed in Chapter 7 and Chapter 8, where the impact of co-channel interference is investigated. However, a number of questions remains to be addressed. To name a few, what is the optimal transmit precoding strategy for each individual user, what is the optimal power allocation scheme, what the the sum capacity of the system, and what is the optimality condition for the beamforming transmission scheme, etc.

To sum up, the newly emerged MIMO product channel is an important class of channel model, and considerable efforts are required to help improve our fundamental understanding of its performance.

Appendices

In this Appendix, we provide all the proofs for the main theorems and corollaries appeared in the thesis in detail. Specifically, Section A gives the proofs for the new statistical results of certain random matrices presented in Chapter 3, while the remaining sections give the proofs for the key theorems and corollaries appeared in Chapter 4, Chapter 5, Chapter 6, and Chapter 8, respectively.

A. Proofs for Chapter 3

A.1. Proof for Lemma 3.7

To prove this lemma, it is convenient to give a separate treatment for the two cases, $m < n$ and $m \geq n$.

(i) The $m < n$ Case

For this case, an expression for the p.d.f. $f(\lambda)$ has been given previously as [5]

$$f(\lambda) = \frac{\sum_{l=1}^{m}\sum_{k=1}^{m} \lambda^{n-m+k-1} e^{-\lambda/\beta_l} \tilde{D}_{l,k}}{m \det(\mathbf{L})^{n-m+1} \prod_{i=1}^{m} \Gamma(n-i+1) \prod_{i<j}^{m} (\omega_j - \omega_i)}, \tag{10.1}$$

where $\tilde{D}_{l,k}$ is the (l,k)th cofactor of a $m \times m$ matrix with entries

$$\left\{\tilde{\mathbf{D}}\right\}_{i,j} = \Gamma(n-m+j)\, \omega_i^{n-m+j}. \tag{10.2}$$

After some basic manipulations, we can express this cofactor as

$$\tilde{D}_{l,k} = \frac{\prod_{j=1}^{m} \Gamma(n-j+1)}{\Gamma(n-m+k)} \frac{\det(\mathbf{\Omega})^{n-m+1}}{\omega_l^{n-m+1}} D_{l,k}. \tag{10.3}$$

Substituting (10.3) into (10.1) yields the desired result.

(ii) The $m \geq n$ Case

For this case, we start by employing a result from [88, (11)] to express the joint p.d.f. of the unordered eigenvalues $\gamma_1, \ldots, \gamma_n$ of $\mathbf{H}^\dagger \mathbf{\Omega} \mathbf{H}$, as follows

$$f(\gamma_1, \ldots, \gamma_n) = \frac{\det(\mathbf{\Delta}_1) \prod_{i<j}^{n} (\gamma_j - \gamma_i)}{\prod_{i=1}^{n} \Gamma(n-i+1) \prod_{i<j}^{m} (\omega_j - \omega_i)}, \tag{10.4}$$

where $\mathbf{\Delta}_1$ is the $m \times m$ matrix

$$\mathbf{\Delta}_1 = \begin{bmatrix} 1 & \omega_1 & \cdots & \omega_1^{m-n-1} & \omega_1^{m-n-1} e^{-\frac{\gamma_1}{\omega_1}} & \cdots & \omega_1^{m-n-1} e^{-\frac{\gamma_n}{\omega_1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_m & \cdots & \omega_m^{m-n-1} & \omega_m^{m-n-1} e^{-\frac{\gamma_1}{\omega_m}} & \cdots & \omega_m^{m-n-1} e^{-\frac{\gamma_n}{\omega_m}} \end{bmatrix}. \tag{10.5}$$

The p.d.f. of a single unordered eigenvalue λ is found from (10.4) via

$$\begin{aligned} f(\lambda) &= \int_0^\infty \cdots \int_0^\infty f(\gamma_1, \ldots, \gamma_n)\, d\gamma_1 \cdots d\gamma_{n-1} \Big|_{\gamma_n = \lambda} \\ &= \frac{1}{\prod_{i=1}^n \Gamma(n-i+1) \prod_{i<j}^m (\omega_j - \omega_i)} \int_0^\infty \cdots \int_0^\infty \det(\mathbf{\Delta}_1) \det\left(\gamma_i^{j-1}\right) d\gamma_1 \cdots d\gamma_{n-1} \Big|_{\gamma_n = \lambda}, \end{aligned} \tag{10.6}$$

where we have used $\prod_{i<j}^n (\gamma_j - \gamma_i) = \det\left(\gamma_i^{j-1}\right)$. To evaluate the $n-1$ integrals, we expand $\det(\mathbf{\Delta}_1)$ along its last column and $\det\left(\gamma_i^{j-1}\right)$ along its last row, and then integrate term-by-term by virtue of [84, Lemma 2]. This yields

$$f(\lambda) = \frac{\sum_{l=1}^m \sum_{k=m-n+1}^m \beta_l^{m-n-1} e^{-y/\omega_l} \lambda^{m-n+k-1} \bar{D}_{l,k}}{n \prod_{i=1}^n \Gamma(n-i+1) \prod_{i<j}^m (\omega_j - \omega_i)}, \tag{10.7}$$

where $\bar{D}_{l,k}$ is the (l,k)th cofactor of a $m \times m$ matrix $\mathbf{\Xi} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \end{bmatrix}$, with entries

$$\{\mathbf{A}\}_{i,j} = \omega_i^{j-1} \quad i = 1, \ldots, m, \quad j = 1, \ldots, m-n \tag{10.8}$$

and

$$\{\mathbf{C}\}_{i,j} = \Gamma(j)\, \omega_i^{m-n+j-1} \quad i = 1, \ldots, m, \quad j = 1, \ldots, n\ . \tag{10.9}$$

Then, it can be shown that

$$\sum_{l=1}^m \sum_{k=m-n+1}^n \omega_l^{m-n-1} e^{-y/\omega_l} \lambda^{m-n+k-1} \bar{D}_{l,k} = \sum_{k=m-n+1}^m \det(\mathbf{D}_k), \tag{10.10}$$

where $\mathbf{D}_k$ is a $m \times m$ matrix with entries

$$\{\mathbf{D}_k\}_{i,j} = \begin{cases} \omega_i^{j-1}, & i = 1, \ldots, m, \quad j = 1, \ldots, m-n, \\ \Gamma(j)\, \omega_i^{m-n+j-1}, & i = 1, \ldots, m, \quad j = m-n+1, \ldots, m, \quad j \neq k, \\ \omega_i^{m-n-1} e^{-\lambda/\omega_i} \lambda^{j-m+n-1}, & i = 1, \ldots, m, \quad j = k, \end{cases} \tag{10.11}$$

Hence, we can rewrite (10.7) as

$$f(\lambda) = \frac{\sum_{k=m-n+1}^{m} \det(\mathbf{D}_k)}{n \prod_{i=1}^{n} \Gamma(n-i+1) \prod_{i<j}^{m} (\omega_j - \omega_i)}. \tag{10.12}$$

After some basic manipulations, (10.12) can be further simplified as

$$f(\lambda) = \frac{1}{n \prod_{i<j}^{m} (\omega_j - \omega_i)} \sum_{k=m-n+1}^{m} \frac{\lambda^{n-m+k-1}}{\Gamma(n-m+k)} \det\left(\bar{\mathbf{D}}_k\right), \tag{10.13}$$

where $\bar{\mathbf{D}}_k$ is a $m \times m$ matrix with entries

$$\left\{\bar{\mathbf{D}}_k\right\}_{i,j} = \begin{cases} \omega_i^{j-1}, & j \neq k, \\ e^{-\lambda/\omega_i} \omega_i^{n-m+1}, & j = k. \end{cases} \tag{10.14}$$

Finally, we apply Laplace's expansion to (10.13) to yield the desired result.

A.2. Proof for Lemma 3.8

The joint p.d.f. of the non-zero eigenvalues of matrix $\mathbf{H}\mathbf{H}^{\dagger}$ $\alpha_1, \ldots, \alpha_q$ is given in [34] as

$$f(\alpha_1, \cdots, \alpha_q) = \mathcal{K} e^{-\sum_{i=1}^{q} \alpha_i} \prod_{i=1}^{q} \alpha_i^{p-q} \prod_{i<j}^{q} (\alpha_j - \alpha_i)^2. \tag{10.15}$$

Recalling that

$$\alpha_i = \frac{\omega_i}{1 - a\omega_i}, \tag{10.16}$$

we derive the joint p.d.f. of $\omega_1, \ldots, \omega_q$ from (10.15) by applying a vector transformation [67]

$$f(\omega_1, \cdots, \omega_q) = f\left(\frac{\omega_1}{1-a\omega_1}, \cdots, \frac{\omega_q}{1-a\omega_q}\right) \left|\mathbf{J}\left((\alpha_1, \ldots, \alpha_q) \to (\omega_1, \ldots, \omega_q)\right)\right|, \tag{10.17}$$

where

$$\mathbf{J}\left((\alpha_1, \ldots, \alpha_q) \to (\omega_1, \ldots, \omega_q)\right) = \det \begin{bmatrix} \frac{\partial \alpha_1}{\partial \omega_1} & \cdots & \frac{\partial \alpha_1}{\partial \omega_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial \alpha_q}{\partial \omega_1} & \cdots & \frac{\partial \alpha_q}{\partial \omega_q} \end{bmatrix}. \tag{10.18}$$

From (10.16), we have

$$\frac{\partial \alpha_i}{\partial \omega_i} = \frac{1}{(1-a\omega_i)^2}. \tag{10.19}$$

Therefore the Jacobian transformation in (10.18) is evaluated as

$$\mathbf{J}\left((\alpha_1,\ldots,\alpha_q)\rightarrow(\omega_1,\ldots,\omega_q)\right)=\prod_{i=1}^{q}\frac{1}{(1-a\omega_i)^2}. \tag{10.20}$$

Substituting (10.15) and (10.20) into (10.17) yields

$$f\left(\omega_1,\cdots,\omega_q\right)=\mathcal{K}\prod_{i=1}^{q}\frac{\omega_i^{p-q}e^{-\frac{\omega_i}{1-a\omega_i}}}{(1-a\omega_i)^{p-q+2}}\prod_{i<j}^{q}\left(\frac{\omega_j}{1-a\omega_j}-\frac{\omega_i}{1-a\omega_i}\right)^2. \tag{10.21}$$

Finally, simplifying using

$$\prod_{i<j}^{q}\left(\frac{\omega_j}{1-a\omega_j}-\frac{\omega_i}{1-a\omega_i}\right)^2=\prod_{i<j}^{q}\left(\frac{\omega_j-\omega_i}{(1-a\omega_j)(1-a\omega_i)}\right)^2=\frac{\prod_{i<j}^{q}(\omega_j-\omega_i)^2}{\prod_{i=1}^{q}(1-a\omega_i)^{2(q-1)}} \tag{10.22}$$

yields the desired result.

We now derive the p.d.f. of an unordered eigenvalue ω. According to [82, (42)], the unordered eigenvalue p.d.f. of $\mathbf{HH}^\dagger$ is given by

$$f\left(\alpha\right)=\frac{1}{q}\sum_{i=0}^{q-1}\sum_{j=0}^{i}\sum_{l=0}^{2j}\mathcal{A}\left(i,j,l,p,q\right)\alpha^{p-q+l}e^{-\alpha}\,. \tag{10.23}$$

Recalling that $\omega=\alpha/\left(1+a\alpha\right)$, the result follows after applying a simple transformation.

A.3. Proof for Theorem 3.1

We start by re-expressing the unordered eigenvalue p.d.f. of a semi-correlated Wishart matrix $f(\lambda)$ in Lemma 3.7 as

$$f\left(\lambda\right)=\frac{1}{s\prod_{i<j}^{q}\left(\omega_j-\omega_i\right)}\sum_{k=q-s+1}^{q}\frac{\lambda^{N_s-q+j-1}}{\Gamma\left(N_s-q+j\right)}\det\left(\tilde{\mathbf{D}}_k\right), \tag{10.24}$$

where $\tilde{\mathbf{D}}_k$ is a $q\times q$ matrix with entries

$$\left\{\tilde{\mathbf{D}}_k\right\}_{m,n}=\begin{cases}\omega_m^{n-1}, & n\neq k,\\ e^{-\lambda/\omega_m}\omega_m^{q-N_s-1}, & n=k.\end{cases} \tag{10.25}$$

Now, utilizing Lemma 3.8, we can evaluate the unconditional p.d.f. as

$$\begin{aligned}f(\lambda)&=E_{\mathbf{\Omega}}\left[f(\lambda)\right]\\&=\frac{1}{s\prod_{i<j}^{q}(\omega_j-\omega_i)^2}\prod_{i=1}^{q}\left(\Gamma\left(q-i+1\right)\Gamma\left(p-i+1\right)\right)\sum_{k=q-s+1}^{q}\frac{\lambda^{N_s-q+k-1}}{\Gamma\left(N_s-q+k\right)}\bar{\mathcal{I}}_k\end{aligned} \tag{10.26}$$

where

$$\bar{\mathcal{I}}_k = \int_{0\le\omega_1<\cdots<\omega_q\le 1/a} \det(\tilde{\mathbf{D}}_k)\prod_{i<j}^{q}(\omega_j-\omega_i)\prod_{l=1}^{q}\frac{\omega_l^{p-q}e^{-\frac{\omega_l}{1-a\omega_l}}}{(1-a\omega_l)^{p+q}}d\omega_1\cdots d\omega_q$$
$$= \det(\tilde{\mathbf{Y}}_k), \tag{10.27}$$

where $\tilde{\mathbf{Y}}_k$ is a $q\times q$ matrix with entries

$$\{\tilde{\mathbf{Y}}_k\}_{m,n} = \begin{cases} \int_0^{1/a}\frac{x^{p-q+m+n-2}}{(1-ax)^{p+q}}e^{-\frac{x}{1-ax}}dx, & n\neq k,\\ \int_0^{1/a}\frac{x^{p-N_s+m-2}}{(1-ax)^{p+q}}e^{-\frac{x}{1-ax}}e^{-\lambda/x}dx, & n=k. \end{cases} \tag{10.28}$$

Let $t = x/(1-ax)$. Utilizing [26, (3.383.5)] and [26, (3.471.9)], the integrals in (10.28) can be evaluated, respectively, as[1]

$$\int_0^{1/a}\frac{x^{p-q+m+n-2}}{(1-ax)^{p+q}}e^{-\frac{x}{1-ax}}dx = \int_0^{\infty}t^{p-q+m+n-2}(1+at)^{2q-m-n}e^{-t}dt$$
$$= a^{q-p-m-n+1}\Gamma(p-q+m+n-1)\,U(p-q+m+n-1,p+q,1/a) \tag{10.29}$$

and

$$\int_0^{1/a}\frac{x^{p-N_s+m-2}}{(1-ax)^{p+q}}e^{-\frac{x}{1-ax}}e^{-\lambda/x}dx$$
$$= e^{-\lambda a}\int_0^{\infty}t^{p-N_s+m-2}(1+at)^{q+N_s-m}e^{-t-\lambda/t}dt$$
$$= e^{-\lambda a}\sum_{i=0}^{q+N_s-m}\binom{q+N_s-m}{i}a^{q+N_s-m-i}\int_0^{\infty}t^{p+q-i-2}e^{-t-\lambda/t}dt$$
$$= 2e^{-\lambda a}\sum_{i=0}^{q+N_s-m}\binom{q+N_s-m}{i}a^{q+N_s-m-i}\lambda^{(p+q-i-1)/2}K_{p+q-i-1}\left(2\sqrt{\lambda}\right). \tag{10.30}$$

Combining (10.26)–(10.30) and then applying Laplace's expansion yields the desired result.

A.4. Proof of Theorem 3.2

Due to the symmetry of the channel, we only deal with the case when $N_t \ge N_r$. The case for $N_r > N_t$ can be dealt with by simply exchanging N_t and N_r. We find it useful to give a separate treatment for the two cases: $N_r > N_k$ and $N_r \le N_k$. For convenience, we define $\mathbf{Q} = \mathbf{A}^{\dagger}\mathbf{H}_2^{\dagger}\mathbf{H}_2\mathbf{A}$.

(i) The $N_r > N_k$ Case

Since $N_t \ge N_r > N_k$, we observe that $\mathbf{F}$ has N_k distinct eigenvalues $0 < f_1 < \cdots < f_{N_k} < \infty$ and $\mathbf{Q}$

[1] Note that, by using the Binomial expansion, (10.29) can be alternatively expressed as

$$\int_0^{\infty}t^{p-q+m+n-2}(1+at)^{2q-m-n}e^{-t}dt = \sum_{i=0}^{2q-m-n}a^i\Gamma(p-q+m+n+i-1).$$

has N_k distinct eigenvalues $0 < q_1 < \cdots < q_{N_k} < \infty$. We assume that $0 < b_1 < \cdots < b_{N_k} < \infty$. First note that the non-zero eigenvalues of $\mathbf{Q}$ are the same as that of $\mathbf{W} \triangleq \mathbf{H}_2\mathbf{A}\mathbf{A}^\dagger\mathbf{H}_2^\dagger$. Therefore, the joint p.d.f. of the ordered eigenvalues of $\mathbf{Q}$ is given by [46]

$$f(q_1, \ldots, q_{N_k}) = \frac{1}{\prod_{i=1}^{N_k} \Gamma(N_r - i + 1) \det(\mathbf{B})^{N_r}} \prod_{i=1}^{N_k} q_i^{N_r - N_k} \frac{\det\left(q_i^{N_k - j}\right) \det\left(e^{-\frac{q_i}{b_j}}\right)}{\prod_{l<k}^{N_k} \left(\frac{1}{b_k} - \frac{1}{b_l}\right)}. \quad (10.31)$$

Utilizing Lemma 3.7 and (10.31), the marginal p.d.f. of an unordered eigenvalue λ of $\mathbf{H}_1^\dagger\mathbf{Q}\mathbf{H}_1$ can be obtained as

$$f(\lambda) = C_1 \sum_{s=1}^{N_k} \frac{\lambda^{N_t - N_k + s - 1}}{\Gamma(N_t - N_k + s)} \det(\mathbf{D}_s), \quad (10.32)$$

where

$$C_1 = \frac{1}{N_k \prod_{i=1}^{N_k} \Gamma(N_r - i + 1) \det(\mathbf{B})^{N_r - N_k + 1} \prod_{i<j}^{N_k} (b_j - b_i)}, \quad (10.33)$$

and $\mathbf{D}_s$ is an $N_k \times N_k$ matrix whose entries are defined as

$$[\mathbf{D}_s]_{l,k} = \begin{cases} b_l^{N_r - N_k + k} \Gamma(N_r - N_k + k), & k \neq s, \\ 2(\lambda b_l)^{\frac{N_r - N_t}{2}} K_{N_r - N_t}\left(2\sqrt{\frac{\lambda}{b_l}}\right), & k = s. \end{cases} \quad (10.34)$$

After some manipulation, we can then compute the determinant of $\mathbf{D}_s$ as

$$\det(\mathbf{D}_s) = \prod_{\substack{i=1 \\ i \neq s}}^{N_k} \Gamma(N_r - N_k + i) \det(\mathbf{B})^{N_r - N_k + 1} \det(\bar{\mathbf{D}}_s), \quad (10.35)$$

where $\bar{\mathbf{D}}_s$ is defined as

$$[\bar{\mathbf{D}}_s]_{l,k} = \begin{cases} b_l^{k-1}, & k \neq s, \\ 2\lambda^{\frac{N_r - N_t}{2}} b_l^{N_k - 1 - \frac{N_r + N_t}{2}} K_{N_r - N_t}\left(2\sqrt{\frac{\lambda}{b_l}}\right), & k = s. \end{cases} \quad (10.36)$$

Substituting (10.35) into (10.32) and applying the Laplace's expansion along the sth column of $\bar{\mathbf{D}}_s$, we have

$$f(\lambda) = \frac{1}{N_k \prod_{i<j}^{N_k} (b_j - b_i)} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \frac{2 b_i^{N_k - 1 - \frac{N_r + N_t}{2}} \lambda^{\frac{N_r + N_t}{2} - N_k + j - 1} K_{N_r - N_t}\left(2\sqrt{\frac{\lambda}{b_i}}\right)}{\Gamma(N_r - N_k + j)\Gamma(N_t - N_k + j)} \mathcal{D}_{i,j}, \quad (10.37)$$

where $\mathcal{D}_{i,j}$ is the (i,j)th cofactor of an $N_k \times N_k$ matrix $\mathbf{\Xi}$ defined in (3.26).

(ii) The $N_r \leq N_k$ Case

When $N_t \geq N_k \geq N_r$ or $N_k \geq N_t \geq N_r$, the joint p.d.f. of the ordered N_r eigenvalues $0 < q_1 < \cdots <$

$q_{N_r} < \infty$ of $\mathbf{Q}$ is in (10.4). Similar to the proof in the first part, we derive the p.d.f. as

$$f(\lambda) = C_2 \sum_{s=N_k-N_r+1}^{N_k} \frac{\lambda^{N_t+s-N_k-1}}{\Gamma(N_t-N_k+s)} \det(\mathbf{E}_s), \tag{10.38}$$

where

$$C_2 = \frac{1}{N_r \prod_{i=1}^{N_r} \Gamma(N_r-i+1) \prod_{i<j}^{N_k}(b_j-b_i)} \tag{10.39}$$

and $\mathbf{E}_s$ is an $N_k \times N_k$ matrix defined as

$$[\mathbf{E}_s]_{l,k} = \begin{cases} b_l^{k-1}, & k \le N_k - N_r, \\ b_l^{k-1}\Gamma(N_r-N_k+k), & k > N_k - N_r and k \ne s, \\ 2b_l^{N_k-N_r-1}(\lambda b_l)^{\frac{N_r-N_t}{2}} K_{N_r-N_t}\left(2\sqrt{\frac{\lambda}{b_l}}\right), & k = s. \end{cases} \tag{10.40}$$

In the above, the integration technique for the product of the determinant of two matrices of different dimensions given in [82] was used. After some manipulations, (10.38) can be further simplified as

$$f(\lambda) = \frac{1}{N_r \prod_{i<j}^{N_k}(b_j-b_i)} \sum_{s=N_k-N_r+1}^{N_k} \frac{\lambda^{N_t+s-N_k-1}}{\Gamma(N_t-N_k+s)\Gamma(N_r-N_k+s)} \det(\bar{\mathbf{E}}_s), \tag{10.41}$$

where $\bar{\mathbf{E}}_s$ is an $N_k \times N_k$ matrix defined by

$$[\bar{\mathbf{E}}_s]_{l,k} = \begin{cases} b_l^{k-1}, & k \ne s, \\ 2b_l^{N_k-N_r-1}(\lambda b_l)^{\frac{N_r-N_t}{2}} K_{N_r-N_t}\left(2\sqrt{\frac{\lambda}{b_l}}\right), & k = s. \end{cases} \tag{10.42}$$

Finally, we apply Laplace's expansion in (10.42) to yield

$$f(\lambda) = \frac{1}{N_r \prod_{i<j}^{N_k}(b_j-b_i)} \sum_{i=1}^{N_k} \sum_{j=N_k-N_r+1}^{N_k} \frac{2b_i^{N_k-1-\frac{N_r+N_t}{2}} \lambda^{\frac{N_t+N_r}{2}+j-N_k-1} K_{N_r-N_t}\left(2\sqrt{\frac{\lambda}{b_i}}\right)}{\Gamma(N_t-N_k+j)\Gamma(N_r-N_k+j)} \mathcal{D}_{i,j}, \tag{10.43}$$

where $\mathcal{D}_{i,j}$ is the (i,j)th cofactor of an $N_k \times N_k$ matrix $\mathbf{\Xi}$ defined in (3.26).

A.5. Proof of Theorem 3.3

Similar to the proof of Theorem 3.2, we only consider the case when $N_t \ge N_r$ and give a separate treatment for the two cases: $N_r > N_k$ and $N_r \le N_k$.

(i) The $N_r > N_k$ Case

When $N_t \ge N_r > N_k$, the c.d.f. conditioned on $\mathbf{A}^\dagger \mathbf{H}_2^\dagger \mathbf{H}_2 \mathbf{A}$ is given in [43]

$$F_{\lambda_{\max}|\mathbf{Q}}(x) = \frac{1}{\prod_{i=1}^{N_k}\Gamma(N_t-i+1)} \frac{\det\left(q_j^{N_t-i+1}\gamma\left(N_t-i+1, \frac{x}{q_j}\right)\right)}{\prod_{i=1}^{N_k} q_i^{N_t} \prod_{l<k}^{N_k}\left(\frac{1}{q_k}-\frac{1}{q_l}\right)}. \tag{10.44}$$

To obtain the unconditional c.d.f., we need to average over the joint p.d.f. of the ordered eigenvalues of

$\mathbf{Q}$, i.e.,

$$F_{\lambda_{\max}}(x) = \int_{\mathcal{D}_{ord}} F_{\lambda_{\max}|\mathbf{Q}}(x) f(\mathbf{Q}) dq_1 \cdots q_{N_k}, \tag{10.45}$$

where the integrals are taking over the region $\mathcal{D}_{ord} = \{0 \leq q_1 \cdots \leq q_{N_k} \leq \infty\}$.

The integral in (10.45) can be simplified as

$$F_{\lambda_{\max}}(x) = \frac{(-1)^{\frac{N_k(N_k-1)}{2}} \det\left(\bar{\mathbf{\Psi}}_1(x)\right)}{\prod_{i=1}^{N_k} \Gamma(N_r - i + 1) \prod_{i=1}^{N_k} \Gamma(N_t - i + 1) \det(\mathbf{B})^{N_r - N_k + 1} \prod_{i<j}^{N_k} (b_j - b_i)}, \tag{10.46}$$

where $\bar{\mathbf{\Psi}}_1(x)$ is an $N_k \times N_k$ matrix with entries

$$[\bar{\mathbf{\Psi}}_1(x)]_{l,k} = \Gamma(N_t - l + 1) b_k^{N_r - l + 1} \Gamma(N_r - l + 1) - \Gamma(N_t - l + 1) \sum_{t=0}^{N_t - l} \frac{x^t}{\Gamma(t+1)} 2(b_k x)^{\frac{N_r - t - l + 1}{2}} K_{N_r - t - l + 1}\left(2\sqrt{\frac{x}{b_k}}\right). \tag{10.47}$$

Further algebraic manipulations gives

$$F_{\lambda_{\max}}(x) = \frac{(-1)^{\frac{N_k(N_k-1)}{2}} \det(\mathbf{\Psi}_1(x))}{\prod_{i=1}^{N_k} \Gamma(N_r - i + 1) \prod_{i<j}^{N_k} (b_j - b_i)}, \tag{10.48}$$

where the entries of $\mathbf{\Psi}_1(x)$ are defined as

$$[\mathbf{\Psi}_1(x)]_{l,k} = b_l^{N_k - k} \Gamma(N_r - k + 1) - b_l^{N_k - N_r - 1} \sum_{t=0}^{N_t - k} \frac{x^t}{\Gamma(t+1)} 2(b_l x)^{\frac{N_r - t - k + 1}{2}} K_{N_r - t - k + 1}\left(2\sqrt{\frac{x}{b_l}}\right). \tag{10.49}$$

(ii) The $N_r \leq N_k$ Case

When $N_t \geq N_k \geq N_r$ or $N_k \geq N_t \geq N_r$, based on [112, Lemma 1] and (10.44), the c.d.f. of the maximum eigenvalue of $\mathbf{F}$ conditioned on $\mathbf{Q}$ is given by

$$F_{\lambda_{\max}|\mathbf{Q}}(x) = \frac{1}{\prod_{i=1}^{N_r} \Gamma(N_t - i + 1)} \frac{\det\left(q_j^{N_t - i + 1} \gamma\left(N_t - i + 1, \frac{x}{q_j}\right)\right)}{\prod_{i=1}^{N_r} q_i^{N_t} \prod_{l<k}^{N_r} \left(\frac{1}{q_k} - \frac{1}{q_l}\right)}. \tag{10.50}$$

Utilizing the joint ordered p.d.f. of the N_r eigenvalues $0 < q_1 < \cdots < q_{N_r} < \infty$ of $\mathbf{Q}$ given in [88], the unconditional c.d.f. can be obtained as

$$F_{\lambda_{\max}}(x) = \frac{(-1)^{\frac{N_r(N_r-1)}{2}} \det(\bar{\mathbf{\Psi}}_2(x))}{\prod_{i=1}^{N_r} \Gamma(N_t - i + 1) \prod_{i=1}^{N_r} \Gamma(N_r - i + 1) \prod_{i<j}^{N_k} (b_j - b_i)}, \tag{10.51}$$

where $\bar{\mathbf{\Psi}}_2(x)$ is an $N_k \times N_k$ matrix with entries

$$[\bar{\mathbf{\Psi}}_2(x)]_{l,k} = \begin{cases} b_l^{k-1}, & k \leq N_k - N_r, \\ [\mathbf{R}(x)]_{l,k}, & k > N_k - N_r, \end{cases} \tag{10.52}$$

in which $[\mathbf{R}(x)]_{l,k}$ is defined as

$$[\mathbf{R}(x)]_{l,k} = b_l^{2N_k-N_r-k}\Gamma(N_t+N_k-N_r-k+1)\Gamma(N_k-k+1)- \\ b_l^{N_k-N_r-1}\Gamma(N_t+N_k-N_r-k+1)\sum_{t=0}^{N_t+N_k-N_r-k}\frac{2x^t}{\Gamma(t+1)}(xb_l)^{\frac{N_k-k-t+1}{2}}K_{N_k-k-t+1}\left(2\sqrt{\frac{x}{b_l}}\right). \tag{10.53}$$

We can then further simplify (10.51) as

$$F_{\lambda_{\max}}(x) = \frac{(-1)^{\frac{N_r(N_r-1)}{2}}\det(\mathbf{\Psi}_2(x))}{\prod_{i=1}^{N_r}\Gamma(N_r-i+1)\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.54}$$

where the entries of $\mathbf{\Psi}_4(x)$ are given as

$$[\mathbf{\Psi}_2(x)]_{l,k} = \begin{cases} b_l^{k-1}, & k \le N_k - N_r, \\ u(x)_{l,k}, & k > N_k - N_r, \end{cases} \tag{10.55}$$

where

$$u(x)_{l,k} = b_l^{2N_k-N_r-k}\Gamma(N_k-k+1)- \\ b_l^{N_k-N_r-1}\sum_{t=0}^{N_t+N_k-N_r-k}\frac{2x^t}{\Gamma(t+1)}(xb_l)^{\frac{N_k-k-t+1}{2}}K_{N_k-k-t+1}\left(2\sqrt{\frac{x}{b_l}}\right). \tag{10.56}$$

A.6. Proof of Theorem 3.4

We focus on deriving the first order expansion for the p.d.f. of $\lambda_{\max}$. (The corresponding first order expansion for the c.d.f. can be obtained by simple integration.) For the SIMO/MISO multi-keyhole channel, we have $n = p = 1$, the p.d.f. of the maximum eigenvalue can be expressed as

$$f_{\lambda_{\max}}(x) = \frac{\det(\bar{\mathbf{\Phi}}(x))}{\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.57}$$

where $\bar{\mathbf{\Phi}}(x)$ is an $N_k \times N_k$ matrix with the (l,k)th entry equal to

$$[\bar{\mathbf{\Phi}}(x)]_{l,k} = \begin{cases} b_l^{k-1}, & k = 1,\cdots,N_k-1, \\ \frac{2x^{\frac{m-1}{2}}b_l^{N_k-\frac{3+m}{2}}K_{m-1}\left(2\sqrt{\frac{x}{b_l}}\right)}{\Gamma(m)}, & k = N_k. \end{cases} \tag{10.58}$$

To proceed, utilizing series representations of Bessel function $K_v(x)$ and $I_v(x)$ [4], we express the elements of the last column in a series form as given in (10.59). Therefore, the elements of the last

column can be expressed as

$$[\bar{\mathbf{\Phi}}(x)]_{l,N_k} = \frac{1}{\Gamma(m)} \Bigg(\underbrace{\sum_{k=0}^{m-2} \frac{\Gamma(m-1-k)}{\Gamma(k+1)} (-1)^k b_l^{N_k-k-2} x^k}_{Part\ I} + \underbrace{(-1)^m \sum_{k=0}^{\infty} \left(\ln\left(\frac{x}{b_l}\right) - \psi(k+1) - \psi(m+k) \right) \frac{b_l^{N_k-m-k-1} x^{m+k-1}}{\Gamma(k+1)\Gamma(m+k)}}_{Part\ II} \Bigg). \tag{10.59}$$

In order to get the first order expansion, we need to find the minimum exponent of x in (10.59) such that $\det(\bar{\mathbf{\Phi}}(x)) \neq 0$. To do this, we consider three separate cases: (i) $m > N_k$, (ii) $m = N_k$ and (iii) $m < N_k$.

(i) The $m > N_k$ case

Due to the multi-linear property of the determinant, we observe that, in $Part\ I$ in (10.59), the minimum exponent of x satisfying $\det(\bar{\mathbf{\Phi}}(x)) \neq 0$ is $N_k - 1$, i.e., for $k = N_k - 1$. Note also that since the minimum exponent of x in $Part\ II$ is $m - 1$, so it can be omitted. Hence, for small x, we compute $\det(\bar{\mathbf{\Phi}}(x))$ as

$$\det(\bar{\mathbf{\Phi}}(x)) = \frac{\Gamma(m-N_k) x^{N_k-1}}{\Gamma(m)\Gamma(N_k)} \det(\mathbf{\Phi}^1), \tag{10.60}$$

where

$$[\mathbf{\Phi}^1]_{l,k} = \begin{cases} b_l^{k-1}, & k = 1, \cdots, N_k - 1, \\ (-1)^{N_k-1} b_l^{-1}, & k = N_k. \end{cases} \tag{10.61}$$

After some mathematical manipulation, we can compute $\det(\mathbf{\Phi}^1)$ as

$$\det(\mathbf{\Phi}^1) = \prod_{i=1}^{N_k} b_i^{-1} \prod_{i<j}^{N_k} (b_j - b_i). \tag{10.62}$$

Pulling (10.60), (10.62) and (10.57) together, we have the first order expansion for the p.d.f. of $\lambda_{\max}$

$$f_{\lambda_{\max}}(x) = \frac{\Gamma(m-N_k)}{\Gamma(m)\Gamma(N_k) \prod_{i=1}^{N_k} b_i} x^{N_k-1} + o(x^{N_k-1}). \tag{10.63}$$

(ii) The $m = N_k$ case

In this case, the elements in $Part\ I$ in (10.59) do not contribute to the determinant computation, and the minimum exponent of x such that $\det(\bar{\mathbf{\Phi}}(x)) \neq 0$ comes from $Part\ II$ when $k = 0$.

Hence, for small x, we compute $\det(\bar{\mathbf{\Phi}}(x))$ as

$$\det(\bar{\mathbf{\Phi}}(x)) = \frac{(-1)^m x^{m-1}}{\Gamma(m)^2} \det(\mathbf{\Phi}^2), \tag{10.64}$$

where

$$[\mathbf{\Phi}^2]_{l,k} = \begin{cases} b_l^{k-1}, & k = 1, \cdots, N_k - 1, \\ \left(\ln\left(\frac{x}{b_l}\right) - \psi(1) - \psi(m)\right) b_l^{-1}, & k = N_k. \end{cases} \tag{10.65}$$

$\det(\mathbf{\Phi}^2)$ can be further simplified as

$$\det(\mathbf{\Phi}^2) = (-1)^{N_k-1}(\ln x - \psi(1) - \psi(m)) \prod_{i=1}^{N_k} {b_i}^{-1} \prod_{i<j}^{N_k} (b_j - b_i) + \det(\mathbf{\Phi}^3), \tag{10.66}$$

where matrix $\mathbf{\Phi}^3$ is defined in Theorem 3.4. To this end, the first order expansion for the p.d.f. of $\lambda_{\max}$ can be expressed as

$$f_{\lambda_{\max}}(x) = \frac{(-1)^m}{\Gamma(m)^2} \left((-1)^{N_k-1}(\ln x - \psi(1) - \psi(m)) \prod_{i=1}^{N_k} b_i^{-1} + \frac{\det(\mathbf{\Phi}^3)}{\prod_{i<j}^{N_k}(b_j - b_i)} \right) x^{m-1} + o(x^{m-1}). \tag{10.67}$$

(iii) The $m < N_k$ case

Similar to the case (ii), the minimum exponent of x such that $\det(\bar{\mathbf{\Phi}}(x)) \neq 0$ comes from $Part\ II$ when $k = 0$. Hence, for small x, we compute $\det(\bar{\mathbf{\Phi}}(x))$ as

$$\det(\bar{\mathbf{\Phi}}(x)) = \frac{(-1)^{m-1}x^{m-1}}{\Gamma(m)^2} \det(\mathbf{\Phi}^4), \tag{10.68}$$

where matrix $\mathbf{\Phi}^4$ is defined in Theorem 3.4. To this end, the first order expansion for the p.d.f. of $\lambda_{\max}$ can be expressed as

$$f_{\lambda_{\max}}(x) = \frac{(-1)^{m-1}}{\Gamma(m)^2} \frac{\det(\mathbf{\Phi}^4)}{\prod_{i<j}^{N_k}(b_j - b_i)} x^{m-1} + o(x^{m-1}). \tag{10.69}$$

A.7. Proof of Theorem 3.5

We first prove the first part of the theorem, while the second part follows similarly. For convenience, we consider three separate cases.

(i) $N_s \leq N_t \leq N_r$

Define $\mathbf{W} \triangleq \mathbf{H}_1^\dagger(\mathbf{H}_3\mathbf{H}_3^\dagger)^{-1}\mathbf{H}_1$. It is easy to observe that both $\mathbf{F} = \mathbf{H}_2^\dagger\mathbf{H}_1^\dagger(\mathbf{H}_3\mathbf{H}_3^\dagger)^{-1}\mathbf{H}_1\mathbf{H}_2$ and $\mathbf{W}$ have N_s non-zero eigenvalues $0 < \lambda_1 < \cdots < \lambda_{N_s} < \infty$ and $0 < \phi_1 < \cdots < \phi_{N_s} < \infty$, respectively. Utilizing the results in [43], the maximum eigenvalue of $\mathbf{F}$ conditioned on $\mathbf{W}$ is given by

$$\mathcal{F}_{\lambda_{\max,1}}(x|\mathbf{W}) = \frac{\det(\mathbf{\Psi}_1(x))}{\det(\mathbf{V}_1)\prod_{i=1}^{N_s}\Gamma(N_t - i + 1)}, \tag{10.70}$$

where $\mathbf{V}_1$ is an $N_s \times N_s$ matrix, with determinant of

$$\det(\mathbf{V}_1) = \left(\prod_{i=1}^{N_s} \phi_i^{N_t}\right) \prod_{1 \leq l \leq k \leq N_s} \left(\frac{1}{\phi_k} - \frac{1}{\phi_l}\right). \tag{10.71}$$

Also, $\mathbf{\Psi}_1(x)$ is an $N_s \times N_s$ matrix with entries given by

$$[\mathbf{\Psi}_1(x)]_{i,j} = \phi_j^{N_t-i+1}\gamma\left(N_t - i + 1, \frac{xN_s}{\phi_j}\right), \tag{10.72}$$

where

$$\gamma(p,x) = \int_0^x t^{p-1}e^{-t}dt = (p-1)!\left(1 - e^{-x}\sum_{k=0}^{p-1}\frac{x^k}{k!}\right),\ p = 1, 2, \ldots, \tag{10.73}$$

is the lower incomplete gamma function. To obtain the unconditional c.d.f. of $\lambda_{\mathsf{max},1}$, we must further average (10.70) over the joint p.d.f of $\phi_1, \ldots, \phi_{N_s}$ which is given by [48]

$$g_1(\mathbf{W}) = C_1\prod_{j=1}^{N_s}\phi_j^{N_r-N_s}(1+\phi_j)^{-N_{\mathcal{I}}-N_s}\prod_{1\le l\le k\le N_s}(\phi_l - \phi_k)^2, \tag{10.74}$$

where

$$C_1 = \frac{\prod_{j=1}^{N_s}\Gamma(N_{\mathcal{I}}+N_s-j+1)}{\prod_{j=1}^{N_s}\Gamma(N_{\mathcal{I}}+N_s-N_r-j+1)\Gamma(N_s-j+1)\Gamma(N_r-j+1)}. \tag{10.75}$$

The unconditional c.d.f. of $\lambda_{\mathsf{max},1}$ can be obtained by

$$\mathcal{F}_{\lambda_{\mathsf{max},1}}(x) = \int_{\mathbf{W}}\mathcal{F}_{\lambda_{\mathsf{max},1}}(x|\mathbf{W})g_1(\mathbf{W})d\mathbf{W}. \tag{10.76}$$

Substituting (10.74) into (10.76), we then have

$$\mathcal{F}_{\lambda_{\mathsf{max},1}}(x) = \frac{C_1}{\prod_{j=1}^{N_s}\Gamma(N_t-i+1)}L_1(x), \tag{10.77}$$

where

$$L_1(x) = \int_{\mathbf{W}}\det(\mathbf{\Psi}_1(x))\det\left(\left[\phi_j^{N_s-i}\right]_{i,j}\right)\prod_{j=1}^{N_s}\phi_j^{N_r-N_t-1}(1+\phi_j)^{-N_{\mathcal{I}}-N_s}d\mathbf{W}. \tag{10.78}$$

Now using the method proposed in [14], and applying (10.73) and [26, (3.383.5)], and after some mathematical manipulation, we have

$$\mathcal{F}_{\lambda_{\mathsf{max},1}} = C_1\det(\mathbf{\Psi}_1'(x)), \tag{10.79}$$

where

$$[\mathbf{\Psi}_1'(x)]_{i,j} = B(N_s+N_r+1-i-j, N_{\mathcal{I}}-N_r+i+j-1) - U_1(x) \tag{10.80}$$

in which

$$U_1(x) = \sum_{k=0}^{N_t-i}\frac{(xN_s)^k}{\Gamma(k+1)}$$
$$\Gamma(N_{\mathcal{I}}-N_r+i+j+k-1)U(N_{\mathcal{I}}-N_r+i+j+k-1, i+j+k-N_r-N_s, xN_s), \tag{10.81}$$

where $B(\cdot,\cdot)$ is the beta function [26, (8.380.1)].

(ii) $N_t \leq N_s \leq N_r$

From [43], we can obtain the c.d.f. of the maximum eigenvalue of $\mathbf{F}$ conditioned on $\mathbf{W}$ as

$$\mathcal{F}_{\lambda_{\mathsf{max},2}}(x|\mathbf{W}) = \frac{(-1)^{N_t(N_s-N_t)}\det(\mathbf{\Psi}_2(x))}{\det(\mathbf{V}_2)\prod_{i=1}^{N_t}\Gamma(N_t-i+1)}, \tag{10.82}$$

where $\mathbf{V}_2$ is an $N_s \times N_s$ matrix with determinant of

$$\det(\mathbf{V}_2) = \left(\prod_{i=1}^{N_s}\phi_i^{N_t}\right)\prod_{1\leq l\leq k\leq N_s}\left(\frac{1}{\phi_k}-\frac{1}{\phi_l}\right) \tag{10.83}$$

and $\mathbf{\Psi}_2(x)$ is an $N_s \times N_s$ matrix with entries given by

$$\mathbf{\Psi}_2(x)t_{i,j} = \begin{cases} \left(-\frac{1}{\phi_j}\right)^{N_s-N_t-i}, & i \leq N_s-N_t, \\ \phi_j^{N_s-i+1}\gamma\left(N_s-i+1, \frac{xN_s}{\phi_j}\right), & i > N_s-N_t. \end{cases} \tag{10.84}$$

In this case, $\mathbf{W}$ has only N_s non-zero eigenvalues $0 < \phi_1 < \cdots < \phi_{N_s} < \infty$, with the joint p.d.f. given by (10.74). The unconditional c.d.f. of $\lambda_{\mathsf{max},2}$ can be obtained by

$$\mathcal{F}_{\lambda_{\mathsf{max},2}}(x) = \int_{\mathbf{W}}\mathcal{F}_{\lambda_{\mathsf{max},2}}(x|\mathbf{W})g_1(\mathbf{W})d\mathbf{W}. \tag{10.85}$$

Substituting (10.74) and (10.82) into (10.85), we obtain

$$\mathcal{F}_{\lambda_{\mathsf{max},2}}(x) = \frac{C_1(-1)^{N_t(N_s-N_t)}}{\prod_{j=1}^{N_t}\Gamma(N_t-j+1)}L_2(x), \tag{10.86}$$

where $L_2(x) = \det(\mathbf{L}_2(x))$ and the entries of matrix $\mathbf{L}_2(x)$ are defined as

$$[\mathbf{L}_2(x)]_{i,j} = \begin{cases} (-1)^{N_s-N_t-i}B(N_r+i-j, N_{\mathcal{I}}+N_s-N_r-i+j), & i \leq N_s-N_t, \\ \Gamma(N_s-i+1)D(x), & i > N_s-N_t, \end{cases} \tag{10.87}$$

where

$$\begin{aligned} D(x) &= B(2N_s+N_r-N_t-i-j+1, N_{\mathcal{I}}+N_t-N_s-N_r+i+j-1) \\ &- \sum_{k=0}^{N_s-i}\frac{(xN_s)^k}{\Gamma(k+1)}\Gamma(i+j+k+N_{\mathcal{I}}+N_t-N_s-N_r-1) \\ &U(i+j+k+N_{\mathcal{I}}+N_t-N_s-Nr-1, i+j+k-2N_s-N_r+N_t, xN_s). \end{aligned} \tag{10.88}$$

(iii) $N_t \leq N_r \leq N_s$

In this case, $\mathbf{W}$ has only N_r non-zero eigenvalue $0 < \phi_1 < \cdots < \phi_{N_r} < \infty$, Utilizing the result

in [104], the c.d.f. of the maximum eigenvalue of $\mathbf{F}$ conditioned on $\mathbf{W}$ is given by

$$\mathcal{F}_{\lambda_{\max,3}}(x|\mathbf{W}) = \frac{(-1)^{N_t(N_r-N_t)}\det(\mathbf{\Psi}_3(x))}{\det(\mathbf{V}_3)\prod_{i=1}^{N_t}\Gamma(N_t-i+1)}, \tag{10.89}$$

where $\mathbf{V}_3$ is an $N_r \times N_r$ matrix with determinant of

$$\det(\mathbf{V}_3) = \left(\prod_{i=1}^{N_r}\phi_i^{N_t}\right)\prod_{1\le l\le k\le N_r}\left(\frac{1}{\phi_k}-\frac{1}{\phi_l}\right) \tag{10.90}$$

and $\mathbf{\Psi}_3(x)$ is an $N_r \times N_r$ matrix with entries

$$\mathbf{\Psi}_3(x)t_{i,j} = \begin{cases} \left(-\frac{1}{\phi_j}\right)^{N_r-N_t-i}, & i \le N_r-N_t, \\ \phi_j^{N_r-i+1}\gamma\left(N_r-i+1, \frac{xN_s}{\phi_j}\right), & i > N_r-N_t. \end{cases} \tag{10.91}$$

The joint p.d.f. for the N_r non-zero eigenvalues of $\mathbf{W}$ is given by [48]

$$g_2(\mathbf{W}) = C_2\prod_{j=1}^{N_r}\phi_j^{N_s-N_r}(1+\phi_j)^{-N_s-N_\mathcal{I}}\prod_{1\le l\le k\le N_s}(\phi_l-\phi_k)^2, \tag{10.92}$$

where

$$C_2 = \frac{\prod_{j=1}^{N_r}\Gamma(N_\mathcal{I}+N_s-j+1)}{\prod_{j=1}^{N_r}\Gamma(N_\mathcal{I}-j+1)\Gamma(N_s-j+1)\Gamma(N_r-j+1)}. \tag{10.93}$$

The unconditioned c.d.f. of $\lambda_{\max,3}$ can be obtained by

$$\mathcal{F}_{\lambda_{\max,3}}(x) = \int_{\mathbf{W}}\mathcal{F}_{\lambda_{\max,3}}(x|\mathbf{W})g_2(\mathbf{W})d\mathbf{W}. \tag{10.94}$$

Substituting (10.89) and (10.92) into (10.94), we get

$$\mathcal{F}_{\lambda_{\max,3}}(x) = \frac{C_2(-1)^{N_t(N_r-N_t)}}{\prod_{i=1}^{N_t}\Gamma(N_t-i+1)}L_3(x), \tag{10.95}$$

where $L_3(x)$ can be written in determinant form as $L_3(x) = \det(\mathbf{L}_3(x))$ which is defined as

$$[\mathbf{L}_3(x)]_{i,j} = \begin{cases} (-1)^{N_r-N_t-i}B(N_s+i-j, N_\mathcal{I}-i+j), & i \le N_r-N_t, \\ \Gamma(N_r-i+1)E(x), & i > N_r-N_t, \end{cases} \tag{10.96}$$

where

$$\begin{aligned} E(x) &= B(N_s-N_t+2N_r-i-j+1, N_\mathcal{I}+N_t-2N_r+i+j-1) \\ &- \sum_{k=0}^{N_r-i}\frac{(xN_s)^k}{\Gamma(k+1)}\Gamma(N_\mathcal{I}+N_t-2N_r+i+j+k-1) \\ &U(N_\mathcal{I}+N_t-2N_r+i+j+k-1, i+j+k+N_t-2N_r-N_s, xN_s). \end{aligned} \tag{10.97}$$

A.8. Proof of Theorem 3.6

Using the fact that $f_{\lambda_{\max}}(x) = d\mathcal{F}_{\lambda_{\max}}(x)/dx$, and a classical formula for the derivative of a determinant, the p.d.f. can be written as

$$f_{\lambda_{\max}}(x) = \frac{(-1)^{pN_t}\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}+N_s-i+1)\sum_{l=m-N_t+1}^{m}\det(\mathbf{\Delta}_l(x))}{\prod_{i=1}^{m}\Gamma(N_{\mathcal{I}}-N_r+m-i+1)\Gamma(m-i+1)\Gamma(n-i+1)}, \tag{10.98}$$

where $\mathbf{\Delta}_l(x)$ is an $m \times m$ matrix of x with the (i,j)th entries

$$[\mathbf{\Delta}_l(x)]_{i,j} = \begin{cases} [\mathbf{\Delta}(x)]_{i,j}, & i \neq l, \\ -\frac{dR(x)}{dx}, & i = l, \end{cases} \tag{10.99}$$

where $[\mathbf{\Delta}(x)]_{i,j}$ and $R(x)$ are given in Theorem 3.5. Now, to make the notation simpler, we define $A \triangleq N_{\mathcal{I}} - N_r - p + i + j$ and $B \triangleq i + j - p - n - m$. Also, we find the following differential property of $U(\cdot,\cdot,\cdot)$ useful [4, (13.4.20)]

$$U'(a,b,x) = -aU(a+1,b+1,x). \tag{10.100}$$

Using this result, we can express $\frac{dR(x)}{dx}$ as

$$\begin{aligned}
\frac{dR(x)}{dx} &= \frac{d}{dx}\left[\Gamma(A-1)U(A-1,B,xN_s) + \sum_{k=1}^{q-i}\frac{(xN_s)^k}{\Gamma(k+1)}\Gamma(A+k-1)U(A+k-1,B+k,xN_s)\right] \\
&= -N_s\Gamma(A)U(A,B+1,xN_s) + \sum_{k=1}^{q-i}\left[\frac{x^{k-1}N_s^k}{\Gamma(k)}\Gamma(A+k-1)U(A+k-1,B+k,xN_s)\right. \\
&\left. - \frac{x^kN_s^{k+1}}{\Gamma(k+1)}\Gamma(A+k)U(A+k,B+k+1,xN_s)\right] \\
&= -\frac{N_s(xN_s)^{q-i}}{\Gamma(q-i+1)}\Gamma(N_{\mathcal{I}}-N_r-p+j+q)U(N_{\mathcal{I}}-N_r-p+j+q,j-p-n-m+q+1,xN_s).
\end{aligned} \tag{10.101}$$

Substituting (10.101) into (10.99) yields the desired result.

A.9. Proof of Lemma 3.9

We will prove the lemma by giving a separate treatment for the two cases, $m < n$ and $m \geq n$.

(i) $m < n$ Case

In this case, we start by writing

$$\begin{aligned}
\mathtt{E}\left\{\det\left(\mathbf{I}_n + a\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}\right)\right\} &= \mathtt{E}\left\{\det\left(\mathbf{I}_m + a\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)\right\} \\
&= \mathtt{E}\left\{\prod_{i=1}^{m}(1+a\gamma_i)\right\},
\end{aligned} \tag{10.102}$$

where $\gamma_1, \ldots, \gamma_m$ are the ordered eigenvalues of $\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger$ with joint p.d.f. given in [14]. Using this result,

we can express (10.102) as

$$\begin{aligned}&\mathtt{E}\left\{\det\left(\mathbf{I}_m+a\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)\right\}\\&=\frac{\int_{\mathcal{D}_{\text{ord}}}\det\left(e^{-\gamma_j/\omega_i}\right)\prod_{i=1}^m(1+a\gamma_i)\,\omega_i^{m-n-1}\gamma_i^{n_s-m}\det(\gamma_i^{j-1})d\gamma_1\cdots d\gamma_m}{\prod_{i=1}^m\Gamma(n-i+1)\prod_{i<j}^m(\omega_j-\omega_i)},\end{aligned}\tag{10.103}$$

where the integrals are taken over the region $\mathcal{D}_{\text{ord}}=\{\infty\geq\gamma_1\geq\cdots\gamma_m\geq 0\}$. Applying [14, Corollary 2], (10.103) can be evaluated in closed form as

$$\mathtt{E}\left\{\det\left(\mathbf{I}_m+a\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)\right\}=\frac{\prod_{i=1}^m\omega_i^{m-n-1}\det\left(\mathbf{\Xi}_1\right)}{\prod_{i=1}^m\Gamma(n-i+1)\prod_{i<j}^m(\omega_j-\omega_i)},\tag{10.104}$$

where $\mathbf{\Xi}_1$ is an $m\times m$ matrix with entries

$$\{\mathbf{\Xi}_1\}_{l,k}=\omega_l^{n-m+k}\left(\Gamma(n-m+k)+a\omega_l\Gamma(n-m+k+1)\right).\tag{10.105}$$

Extracting common factors from the determinant in (10.104) and simplifying yields the desired result.

(ii) $m\geq n$ Case

In this case, we use the joint eigenvalue p.d.f. (10.4) to obtain

$$\begin{aligned}\mathtt{E}\left\{\det\left(\mathbf{I}_n+a\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}\right)\right\}&=\mathtt{E}\left\{\prod_{i=1}^n(1+a\gamma_i)\right\}\\&=\frac{\int_{\mathcal{D}_{\text{ord}}}\prod_{i=1}^n(1+a\gamma_i)\det\left(\mathbf{\Delta}_1\right)\det(\gamma_i^{j-1})d\gamma_1\cdots d\gamma_n}{\prod_{i=1}^n\Gamma(n-i+1)\prod_{i<j}^m(\omega_j-\omega_i)},\end{aligned}\tag{10.106}$$

where $\gamma_1,\ldots,\gamma_n$ are the ordered eigenvalues of $\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}$, $\mathbf{\Delta}_1$ is defined in (10.5), and the integration region is $\mathcal{D}_{\text{ord}}=\{\infty\geq\gamma_1\geq\cdots\gamma_n\geq 0\}$. Applying [84, Lemma 2], (10.106) can be evaluated in closed form as

$$\mathtt{E}\left\{\det\left(\mathbf{I}_n+a\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}\right)\right\}=\frac{\det\left(\mathbf{\Xi}_2\right)}{\prod_{i=1}^n\Gamma(n-i+1)\prod_{i<j}^m(\omega_j-\omega_i)},\tag{10.107}$$

where $\mathbf{\Xi}_2=\begin{bmatrix}\mathbf{A}_1 & \mathbf{C}_1\end{bmatrix}$ is an $m\times m$ matrix with entries

$$\{\mathbf{A}_1\}_{i,j}=\omega_i^{j-1},\quad j=1,\ldots,m-n\tag{10.108}$$

and

$$\{\mathbf{C}_1\}_{i,j}=\omega_i^{j+m-n-1}\left(\Gamma(j)+a\omega_i\Gamma(j+1)\right),\qquad j=1,\ldots,n.\tag{10.109}$$

Extracting common factors from $\det\left(\mathbf{\Xi}_2\right)$ and simplifying yields the desired result.

A.10. Proof of Theorem 3.8

Due to the symmetry of the channel, we only deal with the case when $N_t \geq N_r$. The case for $N_r > N_t$ can be dealt with by simply exchanging N_t and N_r. The expectation can be computed in the following way.

$$\mathtt{E}_{\mathbf{H}_1,\mathbf{H}_2}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1\right)\right\} = \mathtt{E}_{\mathbf{W}}\left\{\mathtt{E}_{\mathbf{H}_1|\mathbf{W}}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{W}\mathbf{H}_1\right)\right\}\right\}, \tag{10.110}$$

where $\mathbf{W} = \mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}$. The inner expectation is available in several forms from the literature, e.g., [82, 107]. However, the final expressions are rather complex, making further manipulations difficult. Recently, a simple and unified expression was derived in [38], which we use to get

$$\mathtt{E}_{\mathbf{H}_1|\mathbf{W}}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{W}\mathbf{H}_1\right)\right\} = \frac{\det(\mathbf{\Delta}_1)}{\prod_{i<j}^{v}(q_j-q_i)}, \tag{10.111}$$

where $\mathbf{\Delta}_1$ is a $v \times v$ matrix ($v \triangleq \min(N_r, N_k)$) with entries

$$[\mathbf{\Delta}_1]_{l,k} = q_l^{k-1}\left(1+\frac{\gamma}{N_t}q_l(N_t-v+k)\right), \tag{10.112}$$

where $q_1, \ldots, q_v$ are the v non-zero eigenvalues of $\mathbf{W}$.

To proceed, it is convenient to consider two separate cases: $N_r \geq N_k$ and $N_r < N_k$.

(i) The $N_r \geq N_k$ Case

In this case, the joint p.d.f. of the N_k ordered eigenvalues of $\mathbf{W}$ is given by (10.31). Hence, we take the expectation over $\mathbf{W}$ which gives

$$\mathtt{E}_{\mathbf{H}_1,\mathbf{H}_2}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1\right)\right\} = C_1\det(\mathbf{\Delta}_2), \tag{10.113}$$

where

$$C_1 = \frac{1}{\prod_{i=1}^{N_k}\Gamma(N_r-i+1)\det(\mathbf{B})^{N_r-N_k+1}\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.114}$$

and $\mathbf{\Delta}_2$ is an $N_k \times N_k$ matrix with entries

$$[\mathbf{\Delta}_2]_{l,k} = b_l^{N_r-N_k+k}\Gamma(N_r-N_k+k)\left(1+\frac{\gamma b_l}{N_t}(N_r-N_k+k)(N_t-N_k+k)\right). \tag{10.115}$$

After some manipulation, we can simplify (10.113) as

$$\mathtt{E}_{\mathbf{H}_1,\mathbf{H}_2}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1\right)\right\} = \frac{\det(\bar{\mathbf{\Delta}}_2)}{\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.116}$$

where $\bar{\mathbf{\Delta}}_2$ is defined as

$$[\bar{\mathbf{\Delta}}_2]_{l,k} = b_l^{k-1}\left(1+\frac{\gamma b_l}{N_t}(N_r-N_k+k)(N_t-N_k+k)\right). \tag{10.117}$$

(ii) The $N_r < N_k$ Case

In this case, the joint p.d.f. of the N_r ordered eigenvalues of $\mathbf{W}$ is given in [88]. Hence, we take the expectation over $\mathbf{W}$ which gives

$$\mathtt{E}_{\mathbf{H}_1,\mathbf{H}_2}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1\right)\right\}=C_2\det(\mathbf{\Delta}_3), \tag{10.118}$$

where

$$C_2=\frac{1}{\prod_{i=1}^{N_r}\Gamma(N_r-i+1)\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.119}$$

and

$$[\mathbf{\Delta}_3]_{l,k}=\begin{cases} b_l^{k-1}, & k\le N_k-N_r,\\ b_l^{k-1}\Gamma(N_r-N_k+k)\left(1+\frac{\gamma b_l}{N_t}(N_t-N_k+k)(N_r-N_k+k)\right), & k>N_k-N_r.\end{cases} \tag{10.120}$$

After some manipulations, we simplify (10.118) as

$$\mathtt{E}_{\mathbf{H}_1,\mathbf{H}_2}\left\{\det\left(\mathbf{I}+\frac{\gamma}{N_t}\mathbf{H}_1^\dagger\mathbf{A}^\dagger\mathbf{H}_2^\dagger\mathbf{H}_2\mathbf{A}\mathbf{H}_1\right)\right\}=\frac{\det(\bar{\mathbf{\Delta}}_3)}{\prod_{i<j}^{N_k}(b_j-b_i)}, \tag{10.121}$$

where $\bar{\mathbf{\Delta}}_3$ is defined as

$$[\bar{\mathbf{\Delta}}_3]_{l,k}=\begin{cases} b_l^{k-1}, & k\le N_k-N_r,\\ b_l^{k-1}\left(1+\frac{\gamma b_l}{N_t}(N_t-N_k+k)(N_r-N_k+k)\right), & k>N_k-N_r.\end{cases} \tag{10.122}$$

A.11. Proof of Lemma 3.10

To prove this lemma, it is convenient to give a separate treatment for the two cases, $m<n$ and $m\ge n$.

(i) $m<n$ Case

Now we need to calculate the expectation $\mathtt{E}\left\{\ln\det\left(\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)\right\}$. The moment generating function (m.g.f.) of $\ln\det\left(\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)$ is given by

$$\mathcal{M}_1(t)=\mathtt{E}\left\{\det\left(\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)^t\right\}. \tag{10.123}$$

Utilizing the joint p.d.f. of the eigenvalues $\gamma_1,\cdots,\gamma_m$ of $\mathbf{\Omega}\mathbf{H}\mathbf{H}^\dagger$, presented in [5, 14], we get

$$\mathcal{M}_1(t)=\frac{\int_{\mathcal{F}_{\rm ord}}\det\left(e^{-\gamma_j/\omega_i}\right)\prod_{i=1}^m\gamma_i^{n-m+t}\omega_i^{m-n-1}\prod_{i<j}^m(\gamma_j-\gamma_i)d\gamma_1\cdots d\gamma_m}{\prod_{i=1}^m\Gamma(n-i+1)\prod_{i<j}^m(\omega_j-\omega_i)}, \tag{10.124}$$

where the integrals are taken over the region $\mathcal{F}_{\rm ord}=\{\infty\ge\gamma_1\ge\cdots\gamma_m\ge0\}$. Applying [14, Corollary

2], (10.124) can be further simplified as

$$\mathcal{M}_1(t) = \frac{\det(\boldsymbol{\Xi}_3)}{\prod_{i=1}^{m}\Gamma(n-i+1)\prod_{i<j}^{m}(\omega_j-\omega_i)}, \tag{10.125}$$

where $\boldsymbol{\Xi}_3$ is an $m \times m$ matrix with entries

$$\{\boldsymbol{\Xi}_3\}_{i,j} = \omega_i^{m-n-1}\int_0^\infty e^{-y/\omega_i} y^{n-m+t+j-1} dy = \omega_i^{t+j-1}\Gamma(n-m+t+j). \tag{10.126}$$

From $\mathcal{M}_1(t)$, we get

$$\begin{aligned}\mathtt{E}\left\{\ln\det\left(\boldsymbol{\Omega}\mathbf{H}\mathbf{H}^\dagger\right)\right\} &= \frac{d}{dt}\mathcal{M}_1(t)\Big|_{t=0} \\ &= \frac{\sum_{k=1}^{m}\det(\boldsymbol{\Sigma}_k)}{\prod_{i=1}^{m}\Gamma(n-i+1)\prod_{i<j}^{m}(\omega_j-\omega_i)},\end{aligned} \tag{10.127}$$

where $\boldsymbol{\Sigma}_k$ is an $m \times m$ matrix whose entries are

$$\{\boldsymbol{\Sigma}_k\}_{i,j} = \begin{cases} \omega_i^{j-1}\Gamma(n-m+j), & j \neq k, \\ \omega_i^{j-1}\Gamma(n-m+j)\left[\psi(n-m+j)+\ln\omega_i\right], & j = k. \end{cases} \tag{10.128}$$

where $\psi(\cdot)$ is the digamma function. Now, $\det(\boldsymbol{\Sigma}_k)$ can be further simplified as

$$\det(\boldsymbol{\Sigma}_k) = \det\left(\tilde{\boldsymbol{\Sigma}}_k\right)\prod_{k=1}^{m}\Gamma(n-m+k) \tag{10.129}$$

where $\tilde{\boldsymbol{\Sigma}}_k$ is an $m \times m$ matrix with entries

$$\left\{\tilde{\boldsymbol{\Sigma}}_k\right\}_{i,j} = \begin{cases} \omega_i^{j-1}, & j \neq k, \\ \beta_i^{j-1}\left[\psi(n-m+j)+\ln\omega_i\right], & j = k. \end{cases} \tag{10.130}$$

By using the multi-linear property of determinants, along with some basic manipulations, we can write

$$\det\left(\tilde{\boldsymbol{\Sigma}}_k\right) = \psi(n-m+k)\det\left(\omega_i^{j-1}\right) + \det(\mathbf{Y}_k). \tag{10.131}$$

Substituting (10.129) and (10.131) into (10.127) and simplifying yields the desired result.

(ii) $m \geq n$ Case

We now evaluate the m.g.f. of $\ln\det\left(\mathbf{H}^\dagger\boldsymbol{\Omega}\mathbf{H}\right)$, which is given by

$$\mathcal{M}_2(t) = \mathtt{E}\left\{\det\left(\mathbf{H}^\dagger\boldsymbol{\Omega}\mathbf{H}\right)^t\right\}. \tag{10.132}$$

Utilizing (10.4), (10.132) can be expressed as

$$\mathcal{M}_2(t) = \frac{1}{\prod_{i=1}^{n}\Gamma(n-i+1)\prod_{i<j}^{m}(\omega_j-\omega_i)}\int_{\mathcal{D}_{\text{ord}}}\prod_{i=1}^{n}\gamma_i^t\det(\mathbf{\Delta}_2)\det(\gamma_i^{j-1})d\gamma_1,\ldots,d\gamma_n, \tag{10.133}$$

where $\mathcal{D}_{\text{ord}} = \{\infty \geq \gamma_1 \geq \cdots \gamma_n \geq 0\}$. Applying [84, Lemma 2] yields

$$\mathcal{M}_2(t) = \frac{\det(\mathbf{\Xi}_4)}{\prod_{i=1}^{n}\Gamma(n-i+1)\prod_{i<j}^{m}(\omega_j-\omega_i)}, \tag{10.134}$$

where $\mathbf{\Xi}_4 = \begin{bmatrix}\mathbf{A}_2 & \mathbf{C}_2\end{bmatrix}$ is an $m \times m$ matrix with entries

$$\{\mathbf{A}_2\}_{i,j} = \omega_i^{j-1}, \quad j = 1,\ldots,m-n \tag{10.135}$$

and

$$\{\mathbf{C}_2\}_{i,j} = \Gamma(t+j)\,\omega_i^{m-n+t+j-1}, \quad j = 1,\ldots,n. \tag{10.136}$$

From the m.g.f. (10.134), we can then obtain

$$\begin{aligned}\mathtt{E}\left\{\ln\det\left(\mathbf{H}^\dagger\mathbf{\Omega}\mathbf{H}\right)\right\} &= \frac{d}{dt}\mathcal{M}_2(t)\Big|_{t=0} \\ &= \frac{\sum_{k=m-n+1}^{m}\det(\mathbf{\Omega}_k)}{\prod_{i=1}^{n}\Gamma(n-i+1)\prod_{i<j}^{m}(\omega_j-\omega_i)},\end{aligned} \tag{10.137}$$

where $\mathbf{\Omega}_k$ is an $m \times m$ matrix with entries

$$\{\mathbf{\Omega}_k\}_{i,j} = \begin{cases}\omega_i^{j-1}, & j \neq k, \quad j = 1,\ldots,m-n, \\ \Gamma(n-m+j)\,\omega_i^{j-1}, & j \neq k, \quad j = m-n+1,\ldots,m, \\ \omega_i^{j-1}\Gamma(n-m+j)\left[\psi(n-m+j)+\ln\omega_i\right], & j = k.\end{cases} \tag{10.138}$$

By using the multi-linear property of determinants, along with some basic manipulations, we can obtain the desired result.

(iii) $m = s$ Case

In this case, starting with (3.52), we can write the determinant summation over k as

$$\sum_{k=1}^{m}\det(\mathbf{Y}_k) = \sum_{k=1}^{m}\sum_{\{\alpha\}}\text{sgn}(\alpha)\left[\prod_{i=1}^{m}\omega_{\alpha(i)}^{i-1}\right]\ln\omega_{\alpha(k)} \tag{10.139}$$

where the second summation is over all permutations $\alpha = \{\alpha(1),\ldots,\alpha(m)\}$ of the set $\{1,\ldots,m\}$,

with $\operatorname{sgn}(\alpha)$ denoting the sign of the permutation. We can further write

$$\begin{aligned}\sum_{k=1}^{m}\det\left(\mathbf{Y}_k\right) &= \sum_{\{\alpha\}}\operatorname{sgn}(\alpha)\left[\prod_{i=1}^{m}\omega_{\alpha(i)}^{i-1}\right]\sum_{k=1}^{m}\ln\omega_{\alpha(k)} \\ &= \ln\det\left(\operatorname{diag}\left\{\omega_i\right\}_{i=1}^{m}\right)\prod_{i<j}^{m}\left(\omega_j-\omega_i\right) \\ &= \ln\det\left(\mathbf{L}\right)\prod_{i<j}^{m}\left(\omega_j-\omega_i\right). \end{aligned} \tag{10.140}$$

Substituting (10.140) into (3.52) yields the final result.

A.12. Proof of Theorem 3.9

We start with Lemma 3.10 and further take expectation on $\mathbf{W}$ by using Lemma 3.8 as

$$\begin{aligned}\mathtt{E}\left\{\ln\det\left(\mathbf{\Phi}\right)\right\} &= \sum_{k=1}^{s}\psi\left(N_s-s+k\right) \\ &+\mathcal{K}\int_{0<\omega_1<\cdots<\omega_q\leq 1/a}\det\left(\beta_i^{j-1}\right)\prod_{i=1}^{q}g\left(\omega_i\right)\sum_{k=q-N_s+1}^{q}\det\left(\mathbf{Y}_k\right)d\omega_1\cdots d\omega_q,\end{aligned} \tag{10.141}$$

where

$$g\left(u\right)=\frac{u^{p-q}e^{-u/(1-au)}}{\left(1-au\right)^{p+q}}. \tag{10.142}$$

Using [84, Lemma 2], these integrals can be simplified to give

$$\mathtt{E}\left\{\ln\det\left(\mathbf{\Phi}\right)\right\}\sum_{k=1}^{s}\psi\left(N_s-s+k\right)+\mathcal{K}\sum_{k=q-N_s+1}^{q}\det\left(\tilde{\mathbf{W}}_k\right), \tag{10.143}$$

where $\tilde{\mathbf{W}}_k$ is a $q\times q$ matrix with entries

$$\left\{\tilde{\mathbf{W}}_k\right\}_{m,n}=\begin{cases}\int_0^{1/a}\frac{u^{p-q+m+n-2}}{(1-au)^{p+q}}e^{-\frac{u}{1-au}}du, & n\neq k,\\ \int_0^{1/a}\frac{u^{p-q+m+n-2}}{(1-au)^{p+q}}e^{-\frac{u}{1-au}}\ln u du, & n=k.\end{cases} \tag{10.144}$$

For the case $n \neq k$, a closed-form expression is given in (10.29). For the case $n = k$, we utilize [26, (4.358.5)] and [82, (47)], to obtain

$$\begin{aligned}
&\int_0^{1/a} \frac{u^{p-q+m+n-2}}{(1-au)^{p+q}} e^{-\frac{u}{1-au}} \ln u du \\
&\quad = \int_0^{\infty} t^{p-q+m+n-2} (1+at)^{2q-m-n} e^{-t} [\ln t - \ln(1+at)] dt \\
&\quad = \sum_{i=0}^{2q-m-n} a^{2q-m-n-i} \binom{2q-m-n}{i} \int_0^{\infty} t^{p+q-i-2} e^{-t} [\ln t - \ln(1+at)] dt \\
&\quad = \sum_{i=0}^{2q-m-n} a^{2q-m-n-i} \binom{2q-m-n}{i} \Gamma(p+q-i-1) \\
&\qquad\qquad \times \left[\psi(p+q-i-1) - e^{1/a} \sum_{l=0}^{p+q-i-2} E_{l+1}\left(\frac{1}{a}\right) \right]. \qquad (10.145)
\end{aligned}$$

Substituting (10.29) and (10.145) into (10.144) and (10.143) yields (3.54).

When $q = s$, we start with (3.54) and remove the conditioning on $\mathbf{L}$ to give

$$\mathtt{E}\{\ln\det(\mathbf{\Phi})\} = \sum_{k=1}^{q} \psi(N_s - q + k) + q \int_0^{\infty} f(\bar{\omega}) \ln \bar{\omega} d\bar{\omega} \qquad (10.146)$$

where $f(\bar{\omega})$ denotes the unordered eigenvalue p.d.f. of $\mathbf{\Omega}$ (i.e., p.d.f. of a randomly-selected $\bar{\omega} \in \{\omega_1, \cdots, \omega_q\}$). Substituting this p.d.f. from (3.21) and integrating using (10.145), we obtain the desired result.

B. Proofs for Chapter 4

B.1. Proof of Theorem 4.1

Let $\{\lambda_i\}_{i=1}^{s}$ be the s eigenvalues of the matrix $\mathbf{H}_{s\times t}\mathbf{H}_{s\times t}^{\dagger}$. Now, we consider the function $g(x) = \log_2(1+ax)$, for $a > 0$. The second derivative of $g(x)$ with respect to x is given by

$$\frac{d^2 g(x)}{dx^2} = \frac{-a^2 \ln 2}{(1+ax)^2} < 0. \qquad (10.147)$$

Hence, $g(x)$ is concave. Based on Lemma 3.1, we have the Schur-concave symmetric function

$$\phi(\boldsymbol{\lambda}) \triangleq \sum_{i=1}^{s} \log_2\left(1 + \frac{P}{N_t N_0}\lambda_i\right). \qquad (10.148)$$

Define the vector $\boldsymbol{\lambda} \triangleq [\lambda_{[1]}, \ldots, \lambda_{[s]}]$ and $\mathbf{d}^{(s)} \triangleq \left[d_{[1]}^{(s)}, \ldots, d_{[s]}^{(s)}\right]$, where $\{d_{[i]}^{(s)}\}_{i=1}^{s}$ are the diagonal elements of $\mathbf{H}_{s\times t}\mathbf{H}_{s\times t}^{\dagger}$. From Lemma 3.2, we have $\boldsymbol{\lambda} \succ \mathbf{d}^{(\mathbf{s})}$. As a result, $\mathcal{C} = \mathtt{E}[\phi(\boldsymbol{\lambda})] \leq \mathtt{E}[\phi(\mathbf{d}^{(s)})]$. To evaluate this, the p.d.f.'s for $\{d_{[i]}^{(s)}\}_{i=1}^{s}$ are required, which we obtain by first noting that $d_{[i]}^{(s)}$ is actually the sum of t squared i.i.d. Nakagami-m distributed random variables. Then, it is also known that $y = r^2$

with r being a Nakagami-m distributed random variable, has the following p.d.f.

$$p(y) = \frac{1}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m y^{m-1} e^{-\frac{m}{\Omega} y} \text{ for } y \geq 0, \tag{10.149}$$

which is a gamma distributed random variable, $x \sim \gamma(b, c)$, with the scale parameter $b = \frac{\Omega}{m} > 0$ and the shape parameter $c = m > 0$. Additionally, it is known in [19] that the sum of n statistically independent gamma variables with the shape parameters $\{c_i\}_{i=1}^n$ and a common scale parameter b is also a gamma variate with the parameters $\sum_{i=1}^n c_i$ and b. Thus, the p.d.f. of $d_{[i]}^{(s)}$ is

$$p(r) = \frac{1}{\Gamma(tm)} \left(\frac{m}{\Omega}\right)^{tm} r^{tm-1} e^{-\frac{m}{\Omega} r} \text{ for } r \geq 0. \tag{10.150}$$

As a result, the capacity bound can be evaluated as

$$\mathcal{C} \leq \bar{\mathcal{C}}_1 = \frac{s}{\ln 2} \int_0^\infty \ln\left(1 + \frac{P}{N_t N_0} r\right) \frac{1}{\Gamma(tm)} \left(\frac{m}{\Omega}\right)^{tm} r^{tm-1} e^{-\frac{m}{\Omega} r} dr \tag{10.151}$$

$$= \frac{s}{\ln 2} \int_0^\infty G_{2,2}^{1,2}\left(\frac{P}{N_t N_0} r \Big|_{1,0}^{1,1}\right) \frac{1}{\Gamma(tm)} \left(\frac{m}{\Omega}\right)^{tm} r^{tm-1} e^{-\frac{m}{\Omega} r} dr \tag{10.152}$$

$$= \frac{s}{\Gamma(tm) \ln 2} G_{3,2}^{1,3}\left(\frac{P}{N_t N_0} \frac{\Omega}{m} \Big|_{1,0}^{1-tm,1,1}\right), \tag{10.153}$$

where in (10.152), we have expressed $\ln(1 + ax)$ in terms of Meijer G-function [76, (8.4.6.5)] and in (10.153), we have used the integration formula [26, (7.813.1)]

$$\int_0^\infty x^{-\rho} e^{-\beta x} G_{p,q}^{m,n}\left(\alpha x|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) dx = \beta^{\rho-1} G_{p+1,q}^{m,n+1}\left(\frac{\alpha}{\beta}\Big|_{b_1,\ldots,b_q}^{\rho,a_1,\ldots,a_p}\right) \tag{10.154}$$

if $p + q < 2(m + n)$, $|\angle\alpha| < \left(m + n - \frac{1}{2}p - \frac{1}{2}q\right)\pi$, $|\angle\beta| < \frac{\pi}{2}$, $\mathtt{Re}(b_j - \rho) > -1$, for $j = 1, \ldots, m$.

B.2. Proof of Theorem 4.2

We first derive another two ergodic capacity upper bounds, and then compare them with $\bar{\mathcal{C}}_1$. Now, define $R \triangleq \sum_{i=1}^s \lambda_i$, and vector $\mathbf{1}_s \triangleq \left[\frac{R}{s}, \ldots, \frac{R}{s}\right]$. Noting that R is the trace of the resultant channel matrix, R is the sum of st squared i.i.d. Nakagami-m random variables with the p.d.f.

$$p(r) = \frac{1}{\Gamma(stm)} \left(\frac{m}{\Omega}\right)^{stm} r^{stm-1} e^{-\frac{m}{\Omega} r} \text{ for } r \geq 0. \tag{10.155}$$

With the help of Example 3.1 and following similar steps in the proof of Theorem 4.1, the bound $\bar{\mathcal{C}}_2$ can be easily derived and is given in Theorem 4.2.

To derive the capacity upper bound $\bar{\mathcal{C}}_3$, we use the determinant property [30]

$$\det(\mathbf{I}_{s\times s} + \mathbf{A}_{s\times t}\mathbf{B}_{t\times s}) = \det(\mathbf{I}_{t\times t} + \mathbf{B}_{t\times s}\mathbf{A}_{s\times t}), \tag{10.156}$$

and then the ergodic capacity expression (4.6) can be rewritten as

$$\mathcal{C} = \mathtt{E}\left[\log_2 \det\left(\mathbf{I}_{t\times t} + \frac{P}{N_0 N_t}\mathbf{H}_{s\times t}^{\dagger}\mathbf{H}_{s\times t}\right)\right]. \tag{10.157}$$

Define the vector $\mathbf{d}^{(t)} \triangleq \left[d_{[1]}^{(t)}, \ldots, d_{[t]}^{(t)}\right]$, where $\{d_{[i]}^{(t)}\}_{i=1}^{t}$ are the diagonal elements of $\mathbf{H}_{s\times t}^{\dagger}\mathbf{H}_{s\times t}$, which are the sums of s i.i.d. gamma random variables. Also define

$$\boldsymbol{\lambda}^{(t)} \triangleq \left[\lambda_{[1]}, \ldots, \lambda_{[s]}, \underbrace{0, \ldots, 0}_{t-s}\right], \tag{10.158}$$

where $\{\lambda_i\}_{i=1}^{s}$ are the non-zero eigenvalues of $\mathbf{H}_{s\times t}^{\dagger}\mathbf{H}_{s\times t}$. From Lemma 3.2, we then have $\boldsymbol{\lambda}^{(t)} \succ \mathbf{d}^{(t)}$. Following the similar steps as in the proof of Theorem 4.1, we get $\bar{\mathcal{C}}_3$ in Theorem 4.2.

To show the relative tightness of the capacity upper bounds, we note from Example 3.1 that $\mathbf{d}^{(s)} \succ \mathbf{1}_s$. Due to the Schur-concavity of $\phi(\cdot)$, we have $\phi(\mathbf{d}^{(s)}) \leq \phi(\mathbf{1}_s)$, which leads to $\bar{\mathcal{C}}_1 \leq \bar{\mathcal{C}}_2$. On the other hand, with $\mathbf{d}^{(s)}$ and $\mathbf{d}^{(t)}$ defined earlier, they constitute two different divisions of st gamma random variables according to the rules in Theorem 3.3. Applying Theorem 3.3 and after some simple integrations yields the result $\bar{\mathcal{C}}_1 \leq \bar{\mathcal{C}}_3$, which completes the proof.

B.3. Proof of Theorem 4.3

In order to utilize the majorization theory result in Theorem 4.1, we need to find a vector which majorizes the eigenvalue vector of the channel. From Example 3.2, we know that

$$\left[\sum_{i=1}^{s}\lambda_i, 0, \ldots, 0\right] \succ [\lambda_1, \ldots, \lambda_s]. \tag{10.159}$$

Therefore, the ergodic capacity is lower bounded by

$$\begin{aligned}
\mathcal{C} &\geq \underline{\mathcal{C}}_1 && (10.160)\\
&= \mathtt{E}\left[\log_2\left(1 + \frac{P}{N_t N_0}\sum_{i=1}^{s}\lambda_i\right)\right] && (10.161)\\
&= \frac{1}{\Gamma(stm)\ln 2} G_{3,2}^{1,3}\left(\frac{P}{N_t N_0}\frac{\Omega}{m}\middle|_{1,0}^{1-stm,1,1}\right). && (10.162)
\end{aligned}$$

B.4. Proof of Theorem 4.4

As we know, the p.d.f. of the diagonal elements is required when majorization theory is applied to derive the capacity upper bound. However, the diagonal elements of $\mathbf{H}\boldsymbol{\Phi}\mathbf{H}^{\dagger}$ are weighted sums of i.i.d. gamma random variables, of which the p.d.f. expression in closed form is unavailable. To circumvent this, we

use the determinant identity (10.156) to rewrite (4.7) as

$$\mathcal{D} = \log_2 \det \left(\mathbf{I} + \frac{P}{LN_tN_0} \mathbf{\Phi}^{\frac{1}{2}} \mathbf{H}^\dagger \mathbf{H} \mathbf{\Phi}^{\frac{1}{2}} \right). \tag{10.163}$$

Define $\mathbf{W} \triangleq \mathbf{\Phi}^{1/2} \mathbf{H}^\dagger \mathbf{H} \mathbf{\Phi}^{1/2}$, and let $\{w_k \equiv w_i^{(j)}\}$, for $i = 1, \ldots, L$, $j = 1, \ldots, N_t$ and $k = 1, \ldots, LN_t$, be the diagonal entries of $\mathbf{W}$. It is easy to see that $w_i^{(j)} = \frac{l_i}{D_i^v} x_i^{(j)}$ where $x_i^{(j)}$ is the sum of N_r i.i.d. gamma random variables and therefore $w_i^{(j)}$ is gamma distributed $\gamma(\frac{\Omega}{m}, mN_r)$.

According to Lemma 3.2, the capacity is upper bounded by

$$\begin{aligned} \mathcal{D} &\le \bar{\mathcal{D}}_1 \\ &= \sum_{i=1}^{L} N_t \mathrm{E} \left[\log_2 \left(1 + \frac{P}{LN_tN_0} \frac{l_i}{D_i^v} x_i \right) \right], \end{aligned} \tag{10.164}$$

where the superscript (j) for $x_i^{(j)}$ is not needed anymore. The expectation in (10.164) is taken over both l_i and x_i. Utilizing the results in (4.10), (10.164) can be expressed as

$$\bar{\mathcal{D}}_1 = \frac{N_t}{\Gamma(N_r m) \ln 2} \sum_{i=1}^{L} \int_0^\infty G_{3,2}^{1,3} \left(\frac{P}{LN_tN_0} \frac{l_i}{D_i^v} \frac{\Omega}{m} \middle|_{1,0}^{1-mN_r,1,1} \right) f(l_i) dl_i. \tag{10.165}$$

Now, substituting (4.3) into (10.165), and changing of variables, namely, $t_i = \frac{\eta \ln l_i - \mu_i}{\sqrt{2}\sigma_i}$, gives

$$\bar{\mathcal{D}}_1 = \frac{N_t}{\Gamma(N_r m) \ln 2} \sum_{i=1}^{L} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} V_i(t) \, e^{-t^2} dt \tag{10.166}$$

where $V_i(t) = G_{3,2}^{1,3} \left(\frac{P}{LN_tN_0} \frac{e^{\frac{\sqrt{2}\sigma_i t + \mu_i}{\eta}}}{D_i^v} \frac{\Omega}{m} \middle|_{1,0}^{1-mN_r,1,1} \right)$. In general, the integration in (10.166) cannot be expressed in closed form but can be efficiently evaluated by Gauss-Hermite quadratic integration [87]. To conclude, we have the capacity upper bound

$$\mathcal{D} \le \bar{\mathcal{D}}_1 = \frac{N_t}{\Gamma(N_r m) \ln 2} \sum_{i=1}^{L} \frac{1}{\sqrt{\pi}} \sum_{j=1}^{N} w_j V_i(a_j) \tag{10.167}$$

where $\{a_j\}_{j=1}^N$ are the zeros of the N-th order Hermite polynomial and $\{w_j\}_{j=1}^N$ are the weight factors tabulated in Table 25.10 of [4].

B.5. Proof of Corollary 4.5

At high SNRs, we approximate $\log_2(1 + ax) \approx \log_2(ax)$ to evaluate the capacity upper bound as

$$\bar{\mathcal{D}}_{\mathsf{hsnr}} = \sum_{i=1}^{L} N_t \mathrm{E} \left[\log_2 \left(\frac{P}{LN_tN_0} \frac{l_i}{D_i^v} x_i \right) \right] = \mathcal{I}_1 + \mathcal{I}_2, \tag{10.168}$$

where

$$\mathcal{I}_1 = LN_t \log_2\left(\frac{P}{LN_tN_0}\right) + \frac{LN_t}{\ln 2}\left[\psi(N_r m) - \ln\left(\frac{m}{\Omega}\right)\right] - N_t v \sum_{i=1}^{L} \log_2 D_i \tag{10.169}$$

and

$$\begin{aligned}\mathcal{I}_2 &= N_t \sum_{i=1}^{L} \int_0^{\infty} \log_2(l_i) f(l_i) dl_i \\ &= \frac{N_t}{\ln 2} \sum_{i=1}^{L} \frac{\eta}{\sqrt{2\pi}\sigma_s} \int_0^{\infty} \ln(l_i) \frac{1}{l_i} e^{-\frac{(\eta \ln l_i - \mu_i)^2}{2\sigma_i^2}} dl_i \\ &= \frac{N_t}{\ln 2} \sum_{i=1}^{L} \frac{1}{\sqrt{\pi}\eta} \int_{-\infty}^{\infty} \left(\sqrt{2}\sigma_i t + \mu_i\right) e^{-t^2} dt = \frac{N_t}{\eta \ln 2} \sum_{i=1}^{L} \mu_i.\end{aligned} \tag{10.170}$$

In (10.170), we have used the following integration results

$$\begin{cases} \int_{-\infty}^{\infty} x e^{-x^2} dx = 0, \\ \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}. \end{cases} \tag{10.171}$$

B.6. Proof of Theorem 4.5

When $L = 1$, the ergodic capacity formula (4.7) reduces to

$$\mathcal{D} = \log_2 \det\left(\mathbf{I} + \frac{Pl}{N_tN_0D^v}\mathbf{H}\mathbf{H}^{\dagger}\right). \tag{10.172}$$

Conditioned on the random variable l, we can then use the result of Theorem 4.3 to get

$$\mathcal{D} \geq \underline{\mathcal{D}}_1 = \mathtt{E}_l\left[\frac{1}{\Gamma(stm)\ln 2} G_{3,2}^{1,3}\left(\frac{Pl}{N_tN_0D^v}\frac{\Omega}{m}\Big|_{1,0}^{1-stm,1,1}\right)\right]. \tag{10.173}$$

We then use the Gauss-Hermite quadratic integration technique to evaluate (10.173) so that

$$\underline{\mathcal{D}}_1 = \frac{1}{\Gamma(stm)\ln 2}\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} U\left(e^{\frac{\sqrt{2}\sigma t + \mu}{\eta}}\right) e^{-t^2} dt = \frac{1}{\Gamma(stm)\ln 2}\frac{1}{\sqrt{\pi}} \sum_{i=1}^{N} w_i U(a_i), \tag{10.174}$$

where $U(t) = G_{3,2}^{1,3}\left(\frac{Pe^{\frac{\sqrt{2}\sigma t+\mu}{\eta}}}{N_tN_0D^v}\frac{\Omega}{m}\Big|_{1,0}^{1-stm,1,1}\right)$, and $\{w_i\}$ and $\{a_i\}$ have been defined in (4.35).

C. Proofs for Chapter 5

C.1. Proof of Lemma 5.1

We prove the lemma by induction. First of all, consider $t = 2$, then we have $k = 0$. It is easy to verify that

$$\det(\mathbf{X}_{2,0}) = \det(\mathbf{V}_2) S_1(x_1, x_2). \tag{10.175}$$

Now, assume that the following is true

$$\det(\mathbf{X}_{t,k}) = \det(\mathbf{V}_t) S_{t-1-k}(x_1, \ldots, x_t) \tag{10.176}$$

and consider

$$\mathbf{X}_{t+1,k} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^k & x_1^{k+2} & x_1^{k+3} & \cdots & x_1^t & x_1^{t+1} \\ 1 & x_2 & \cdots & x_2^k & x_2^{k+2} & x_2^{k+3} & \cdots & x_2^t & x_2^{t+1} \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 1 & x_t & \cdots & x_t^k & x_t^{k+2} & x_t^{k+3} & \cdots & x_t^t & x_t^{t+1} \\ 1 & x_{t+1} & \cdots & x_{t+1}^k & x_{t+1}^{k+2} & x_{t+1}^{k+3} & \cdots & x_{t+1}^t & x_{t+1}^{t+1} \end{bmatrix}. \tag{10.177}$$

To compute $\det(\mathbf{X}_{t+1,k})$, we can write

$$\det(\mathbf{X}_{t+1,k}) = (-1)^{t+2} \prod_{i<t+1} (x_i - x_{t+1}) \det(\mathbf{Z}_t), \tag{10.178}$$

where $\mathbf{Z}_t$ is a $t \times t$ matrix defined as

$$\mathbf{Z}_t = \begin{bmatrix} 1 & x_1 & \cdots & x_1^{k-1} & x_1^k(x_1 + x_{t+1}) & x_1^{k+2} & \cdots & x_1^t \\ 1 & x_2 & \cdots & x_2^{k-1} & x_2^k(x_2 + x_{t+1}) & x_2^{k+2} & \cdots & x_2^t \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_t & \cdots & x_t^{k-1} & x_t^k(x_t + x_{t+1}) & x_t^{k+2} & \cdots & x_t^t \end{bmatrix}. \tag{10.179}$$

Due to the multi-linear property of determinants, and also the assumption (10.176), we have

$$\det(\mathbf{Z}_t) = \det(\mathbf{V}_t)(x_{t+1} S_{t-1-k}(x_1, \ldots, x_t) + S_{t-1-(k-1)}(x_1, \ldots, x_t)). \tag{10.180}$$

With

$$\binom{t}{k} + \binom{t}{k-1} = \binom{t+1}{k}, \tag{10.181}$$

where

$$\binom{n}{k} \triangleq \frac{n!}{k!(n-k)!}, \tag{10.182}$$

we can verify that

$$S_{(t+1)-(1+k)}(x_1, \ldots, x_{t+1}) = x_{t+1} S_{t-(1+k)}(x_1, \ldots, x_t) + S_{(t+1)-(1+k)}(x_1, \ldots, x_t). \tag{10.183}$$

Finally, combining (10.178), (10.180) and (10.183) yields

$$\mathbf{X}_{t+1,k} = \det(\mathbf{V}_{t+1}) S_{(t+1)-(1+k)}(x_1, \ldots, x_{t+1}), \tag{10.184}$$

which by induction completes the proof.

C.2. Proof of Corollary 5.5

When $N_k = 2$ and $m = 1$, the determinant of matrix $\mathbf{\Phi}^4$ can be computed as

$$\det(\mathbf{\Phi}^4) = \ln b_2 - \ln b_1. \tag{10.185}$$

Therefore, to prove Corollary 5.5, we need to show that function $f(x, y)$ defined as

$$f(x, y) = \frac{\ln x - \ln y}{x - y}, \tag{10.186}$$

is a Schur convex function. It is easy to observe that $f(x, y)$ is a symmetric function. Therefore, from Schur's condition [62], we only need to show that

$$g(x, y) \triangleq (x - y)\left(\frac{\partial f(x, y)}{\partial x} - \frac{\partial f(x, y)}{\partial y}\right) \geq 0. \tag{10.187}$$

To this end, $g(x, y)$ can be computed as

$$g(x, y) = \frac{1}{x} + \frac{1}{y} - \frac{2}{x - y} \ln \frac{x}{y}, \tag{10.188}$$

which is symmetric. Hence, without loss of generality, we assume $x > y$ and let $t = \frac{x}{y}$, and

$$f(t) \triangleq \frac{1}{2}\left(t - \frac{1}{t}\right) - \ln t. \tag{10.189}$$

Then, we have

$$g(x, y) = \frac{2}{x - y} f(t). \tag{10.190}$$

The first derivative of $f(t)$ with respect to t can be computed as

$$\frac{df(t)}{dt} = \frac{1}{2}\left(1 + \frac{1}{t^2}\right) - \frac{1}{t} \geq 0. \tag{10.191}$$

In addition, we have $f(t)|_{t \to 1^+} = 0$. Hence, we can conclude that $f(t) \geq 0$ for $t > 1$. Therefore, we have $g(x, y) \geq 0$, which completes the proof.

D. Proofs for Chapter 6

D.1. Proof of (6.18)

When $n_r \to \infty$, the ergodic capacity expression (6.11) can be expressed as

$$\lim_{n_r \to \infty} C(\rho) = \frac{1}{2} \mathtt{E}\left\{\log_2 \det\left(\mathbf{I}_{n_s} + \frac{\rho\alpha}{n_s(1 + \rho)} \tilde{\mathbf{H}}_1^{\dagger} \tilde{\mathbf{L}}_1 \tilde{\mathbf{H}}_1\right)\right\}, \tag{10.192}$$

where $\tilde{\mathbf{L}}_1 = \mathrm{diag}\left\{\lambda_i^2/\left(n_r\left(1+a\lambda_i^2\right)\right)\right\}$. Noting that $q = n_d$, by the Law of Large Numbers we have

$$\lim_{n_r\to\infty} \frac{\mathbf{H}_2\mathbf{H}_2^\dagger}{n_r} = \mathbf{I}_{n_d} \tag{10.193}$$

which implies that

$$\lim_{n_r\to\infty} \frac{\lambda_i^2}{n_r} = 1\ , \qquad i = 1,\ldots,n_d\ . \tag{10.194}$$

Recalling (6.5), application of (10.194) in (10.192) yields

$$\lim_{n_r\to\infty} C\left(\rho\right) = \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\mathbf{I}_{n_s} + \frac{\rho\alpha}{n_s(1+\rho+\alpha)}\mathbf{H}^\dagger\mathbf{H}\right)\right\}, \tag{10.195}$$

where $\mathbf{H}$ is an $n_d \times n_s$ i.i.d. Rayleigh fading MIMO channel matrix. Applying the identity (6.6) to (10.195) yields the desired result.

D.2. Proof of (6.19)

Using (6.6), the ergodic capacity expression (6.11) can be alternatively written as

$$C\left(\rho\right) = \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\mathbf{I}_q + \frac{\rho a}{n_s}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^\dagger\mathbf{L}\right)\right\}\ . \tag{10.196}$$

By the Law of Large Numbers we have

$$\lim_{n_s\to\infty} \frac{\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^\dagger}{n_s} \to \mathbf{I}_q \tag{10.197}$$

and hence (10.196) reduces to

$$\lim_{n_s\to\infty} C\left(\rho\right) = \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\mathbf{I}_q + \rho a\mathbf{L}\right)\right\}\ . \tag{10.198}$$

Substituting (6.12) into (10.198), after some simple manipulations we easily obtain

$$\lim_{n_s\to\infty} C\left(\rho\right) = \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\mathbf{I}_q + \left(\rho+1\right)a\mathbf{H}_2^\dagger\mathbf{H}_2\right)\right\} - \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\mathbf{I}_q + a\mathbf{H}_2^\dagger\mathbf{H}_2\right)\right\}\ . \tag{10.199}$$

Substituting (6.5) into (10.199) and applying the identity (6.6) yields the desired result.

D.3. Proof of Theorem 6.2

We will consider the following cases separately; namely, $q < n_s$ and $q \geq n_s$.

(i) $q < n_s$ Case

We start by applying the identity (6.6) to obtain the ergodic capacity, in the high SNR regime, as

$$C(\rho)|_{\alpha,\rho\to\infty,\alpha/\rho=\beta} = \frac{1}{2}\left[q\log_2\rho - q\log_2\left(\frac{\beta}{n_s n_r}\right) + \mathtt{E}\left\{\log_2\det\left(\bar{\mathbf{L}}\mathbf{H}_1\tilde{\mathbf{H}}_1^\dagger\right)\right\}\right] . \tag{10.200}$$

The high SNR slope can be calculated as

$$S_\infty = \tfrac{q}{2} \quad \text{bit/s/Hz (3dB)} . \tag{10.201}$$

Applying (6.25), the high SNR power offset is given by

$$\mathcal{L}_\infty = \frac{q}{2}\log_2\left(\frac{\beta}{n_s n_r}\right) - \frac{1}{2}\mathtt{E}\left\{\log_2\det\left(\bar{\mathbf{L}}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^\dagger\right)\right\} . \tag{10.202}$$

Invoking Theorem 3.9 and simplifying yields the high SNR power offset for case $q < n_s$.

(ii) $q \geq n_s$ Case

In the high SNR regime, the ergodic capacity can be approximated as

$$C(\rho)|_{\alpha,\rho\to\infty,\alpha/\rho=\beta} = \frac{1}{2}\left[n_s\log_2(\rho) - n_s\log_2\left(\frac{\beta}{n_s n_r}\right) + \mathtt{E}\left\{\log_2\det\left(\tilde{\mathbf{H}}_1^\dagger\bar{\mathbf{L}}\tilde{\mathbf{H}}_1\right)\right\}\right] . \tag{10.203}$$

In this case, the high SNR slope is

$$S_\infty = \frac{n_s}{2} \quad \text{bits/s/Hz (3dB)} \tag{10.204}$$

and the high SNR power offset can be obtained as

$$\mathcal{L}_\infty = \frac{n_s}{2}\log_2\left(\frac{\beta}{n_s n_r}\right) - \frac{1}{2}E\left\{\log_2\det\left(\tilde{\mathbf{H}}_1\bar{\mathbf{L}}\tilde{\mathbf{H}}_1^\dagger\right)\right\} . \tag{10.205}$$

The result follows by applying Theorem 3.9.

D.4. Proof of Corollary 6.5

Substituting $n_r = 1$ into (6.35) yields

$$C_U^{n_r=1}(\rho) = \frac{1}{2}\log_2\left(a^{-n_d}\left[U\left(n_d, n_d+1, \frac{1+\rho}{\alpha}\right) + \rho n_d U\left(n_d+1, n_d+1, \frac{1+\rho}{\alpha}\right)\right]\right) . \tag{10.206}$$

Using the following properties of the confluent hypergeometric function of the second kind [26]:

$$U(a,a,z) = e^z z^{1-a}E_a(z) \tag{10.207}$$

and

$$U\left(a, a+1, z\right) = z^{-a}, \tag{10.208}$$

we get the final expression for $C_U^{n_r=1}(\rho)$ in (6.39). Note that $C_U^{n_r=1}(\rho)$ can be lower and upper bounded as

$$C_{U,1}^{n_r=1}(\rho) < C_U^{n_r=1}(\rho) \leq C_{U,2}^{n_r=1}(\rho), \tag{10.209}$$

with

$$C_{U,1}^{n_r=1}(\rho) = \frac{1}{2}\log_2\left(1 + \rho n_d \frac{1}{\frac{1+\rho}{\alpha} + n_d + 1}\right) \tag{10.210}$$

and

$$C_{U,2}^{n_r=1}(\rho) = \frac{1}{2}\log_2\left(1 + \rho n_d \frac{1}{\frac{1+\rho}{\alpha} + n_d}\right), \tag{10.211}$$

where we have used the inequality [4, (5.1.19)]. Taking $n_d \to \infty$, we see that both (10.210) and (10.211) converge to the same limit in (6.40). Taking $\alpha \to \infty$ and ultilizing [4, (5.1.23)], we obtain (6.41).

D.5. Proof of Theorem 6.4

We will use the lower bound derived in [72, Theorem 1] and consider the following cases separately; namely, $q < n_s$ and $q \geq n_s$.

(i) $q < n_s$ Case

Applying the (6.6) and [72, Theorem 1] to (6.11), we lower bound the ergodic capacity, conditioned on $\mathbf{L}$, as

$$C\left(\rho\right) \geq q\log_2\left(1 + \frac{\rho\alpha}{n_s n_r}\exp\left(\frac{1}{q}\mathtt{E}\left\{\ln\det\left(\mathbf{L}\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^{\dagger}\right)\right\}\right)\right). \tag{10.212}$$

Now, using *Theorem 3.9* yields the desired result.

(ii) $q \geq n_s$ Case

In this case, the lower bound can be written as

$$C\left(\rho\right) \geq n_s\log_2\left(1 + \frac{\rho\alpha}{n_s n_r}\exp\left(\frac{1}{n_s}\mathtt{E}\left\{\ln\det\left(\tilde{\mathbf{H}}_1\mathbf{L}\tilde{\mathbf{H}}_1^{\dagger}\right)\right\}\right)\right). \tag{10.213}$$

Again, we use *Theorem 3.9* to obtain the desired result.

D.6. Proof of Corollary 6.8

Taking $n_s \to \infty$ and using [4, (6.3.18)], we get (6.49).

For the case $n_d \to \infty$, we first apply [4, (5.1.19)] and [26, (8.365.3)] to obtain the following approximation

$$\exp\left(\frac{1+\rho}{\alpha}\right)\sum_{l=1}^{n_d-1} E_{l+1}\left(\frac{1+\rho}{\alpha}\right) \approx \psi\left(n_d+\frac{1+\rho}{\alpha}\right)-\psi\left(\frac{1+\rho}{\alpha}\right) . \tag{10.214}$$

Furthermore, substituting (10.214) into (6.48) and using [26, (8.365.5)] and [4, (6.3.18)] yields (6.50).

Now consider the case $\alpha \to \infty$. Utilizing the recurrence relation for the exponential integral [4, (5.1.14)], the summation in (6.48) can be alternatively written as

$$\begin{aligned}
&\exp\left(\frac{1+\rho}{\alpha}\right)\sum_{l=1}^{n_d-1} E_{l+1}\left(\frac{1+\rho}{\alpha}\right) \\
&\quad= \exp\left(\frac{1+\rho}{\alpha}\right)E_1\left(\frac{1+\rho}{\alpha}\right)+\sum_{l=1}^{n_d-1}\frac{1}{l}\left[1-\frac{1+\rho}{\alpha}\exp\left(\frac{1+\rho}{\alpha}\right)E_l\left(\frac{1+\rho}{\alpha}\right)\right] \\
&\quad= \exp\left(\frac{1+\rho}{\alpha}\right)\left[E_1\left(\frac{1+\rho}{\alpha}\right)-\sum_{l=1}^{n_d-1}\left(\frac{1+\rho}{\alpha l}\right)E_l\left(\frac{1+\rho}{\alpha}\right)\right]+\psi\left(n_d\right)+\gamma
\end{aligned} \tag{10.215}$$

where $\gamma = 0.577215\ldots$ is the Euler's constant. Note that, in deriving (10.215), we have applied the definition of the digamma function [26, (8.365.4)]. Using the series expansion given in [4, (5.1.11)], when $\alpha \to \infty$, we get

$$E_1\left(\frac{1+\rho}{\alpha}\right)\Big|_{\alpha\to\infty} \to -\gamma-\ln\left(\frac{1+\rho}{\alpha}\right) \tag{10.216}$$

and therefore

$$\sum_{l=1}^{n_d-1}\left(\frac{1+\rho}{\alpha l}\right)E_l\left(\frac{1+\rho}{\alpha}\right)\Big|_{\alpha\to\infty} \to 0 . \tag{10.217}$$

Applying (10.215)–(10.217) in (6.48) yields the desired result.

E. Proofs for Chapter 8

E.1. Proof for Lemma 8.2

Utilizing the unitary invariant property of the distributions of $\mathbf{H}$ and $\mathbf{h}$, conditioned on $\mathbf{\Lambda}$, we have

$$\mathtt{E}\left\{\mathsf{tr}\{\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H}\}|\mathbf{\Lambda}\right\} = \mathtt{E}\left\{\mathsf{tr}\{\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H}\}|\mathbf{\Lambda}\right\} \tag{10.218}$$

$$= \mathtt{E}\left\{(\mathsf{vec}(\mathbf{H}))^\dagger(\mathbf{I}\otimes\mathbf{\Lambda})\mathsf{vec}(\mathbf{H})\}|\mathbf{\Lambda}\right\} \tag{10.219}$$

$$= n\mathsf{tr}(\mathbf{\Lambda}), \tag{10.220}$$

in which (10.219) comes from [67, Lemma 2.2.3]. Similarly, conditioned on $\mathbf{\Lambda}$, $\mathtt{E}\{\mathtt{tr}\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\}\}$ can be expressed as

$$\mathtt{E}\left\{\mathtt{tr}\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\}|\mathbf{\Lambda}\right\} = \mathtt{E}\{\mathtt{tr}\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\}|\mathbf{\Lambda}\} = \mathtt{tr}^2\{\mathbf{I}\}\mathtt{tr}\{\mathbf{\Lambda}^2\} + \mathtt{tr}^2\{\mathbf{\Lambda}\}\mathtt{tr}\{\mathbf{I}\}, \tag{10.221}$$

where (10.221) comes from [85, Lemma 6]. Finally, conditioned on $\mathbf{\Lambda}$, and with the help of [85, Lemma 5], we have

$$\mathtt{E}\left\{\mathtt{tr}^2\{\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H}\}|\mathbf{\Lambda}\right\} = \mathtt{tr}\{\mathbf{I^2}\}\mathtt{tr}\{\mathbf{\Lambda}^2\} + \mathtt{tr}^2\{\mathbf{I}\}\mathtt{tr}^2\{\mathbf{\Lambda}\}. \tag{10.222}$$

The desired results can be obtained by further taking expectation on $\mathbf{\Lambda}$ with the help of Lemma 8.1.

E.2. Proof of Lemma 8.3

First, we note that the confluent hypergeometric function can be expressed in terms of exponential integral function $E_n(\cdot)$ [102], so that

$$\Psi\left(m, m, \frac{1}{t}\right) = t^{m-1}e^{\frac{1}{t}}E_m\left(\frac{1}{t}\right), \tag{10.223}$$

$$\Psi\left(m, m-1, \frac{1}{t}\right) = \frac{t^{m-2}}{m-1}\left[e^{\frac{1}{t}}E_{m-1}\left(\frac{1}{t}\right)\left(m-1+\frac{1}{t}\right)-1\right]. \tag{10.224}$$

Moreover, the exponential integral function satisfies the following inequality [4, (5.1.19)]

$$\frac{1}{x+n} < e^x E_n(x) \leq \frac{1}{x+n-1}, x > 0. \tag{10.225}$$

Then, we can establish the following two inequalities:

$$\frac{1}{1+tm} < D_1(m,t) \leq \frac{1}{1+t(m-1)}, \tag{10.226}$$

$$0 < D_2(m,t) \leq \frac{1}{t(m-1)[t(m-2)-1]}. \tag{10.227}$$

For the cases $m \to \infty$ and $t \to \infty$, it is easy to see that both sides of (10.226) and (10.227) approach 0 and therefore, we have $D_1(m,t) = D_2(m,t) = 0$. Now, consider the case $t \to 0$. It is easily observed that both sides of (10.226) will approach 1 and hence, $D_1(m,t) = 1$. However, since $\frac{1}{t(m-1)[t(m-2)-1]} \to \infty$ when $t \to 0$, the two sides of (10.227) diverge. To obtain the limit of $D_2(m,t)$, define $a \triangleq \frac{1}{t}$. Utilizing the property of confluent hypergeometric function [4, (13.4.24)] and [4, (13.4.21)], we have

$$a^m\Psi\left(m, m-1, a\right) = (1-m)a^m\Psi\left(m, m, a\right)) - a^{m+1}\Psi^{'}\left(m, m, a\right) \tag{10.228}$$

$$= (1-m)a^m\Psi\left(m, m, a\right) + ma^{m+1}\Psi\left(m+1, m+1, a\right). \tag{10.229}$$

Then, from (10.226), we have

$$a^m\Psi\left(m, m, a\right) = 1, \text{ as } a \to \infty. \tag{10.230}$$

Therefore, (10.229) reduces to

$$a^m \Psi\left(m, m-1, a\right) = (1-m) + m = 1, \text{ as } a \to \infty, \tag{10.231}$$

which completes the proof.

E.3. Proof of Theorem 8.1

With the help of the following determinant property,

$$\frac{d}{dx} \ln \det(\mathbf{I} + x\mathbf{A}) \mid_{x=0} = \mathsf{tr}\{\mathbf{A}\}, \tag{10.232}$$

$$\frac{d^2}{d^2x} \ln \det(\mathbf{I} + x\mathbf{A}) \mid_{x=0} = -\mathsf{tr}\{\mathbf{A}^2\}, \tag{10.233}$$

we compute the first and second derivatives of $C(\rho)$ at $\rho = 0$ as

$$\dot{C}(0) = N_r \log_2 e, \tag{10.234}$$

$$\ddot{C}(0) = -\frac{N_r^2 \mathtt{E}\left\{\mathsf{tr}\left\{\left(\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right)^2\right\}\right\} \log_2 e}{\mathtt{E}^2\left\{\mathsf{tr}\left\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right\}\right\}}. \tag{10.235}$$

As a result, $\frac{E_b}{N_0}_{\min}$ and S_0 can be computed according to (8.9) as

$$\frac{E_b}{N_0}_{\min} = \frac{N_t \ln 2}{\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right\}\right\}}, \tag{10.236}$$

$$S_0 = \frac{2\mathtt{E}^2\left\{\mathsf{tr}\left\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right\}\right\}}{\mathtt{E}\left\{\mathsf{tr}\left\{\left(\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right)^2\right\}\right\}}. \tag{10.237}$$

To proceed, we need to compute $\mathtt{E}\left\{\mathsf{tr}\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\}\right\}$ and $\mathtt{E}\{\mathsf{tr}\{(\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H})^2\}\}$.

Following similar steps as in [56], we have

$$\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right\}\right\} = \frac{K}{K+1} \mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0 \mathbf{H}_0^\dagger \mathbf{\Lambda}\right\}\right\} + \frac{1}{K+1} \mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w \mathbf{H}_w^\dagger \mathbf{\Lambda}\right\}\right\}. \tag{10.238}$$

The first term of (10.238) can be computed with the help of [56, Lemma 3] as

$$\frac{K}{K+1} \mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0 \mathbf{H}_0^\dagger \mathbf{\Lambda}\right\}\right\} = \frac{K}{K+1} \frac{1}{N_r} \mathsf{tr}\left\{\mathbf{H}_0 \mathbf{H}_0^\dagger\right\} \mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}\right\}\right\} = \frac{K N_t}{K+1} \left[D_1(N_r, \rho_I) + N_r - 1\right], \tag{10.239}$$

where in (10.239), we have used the fact that $\mathsf{tr}\{\mathbf{H}_0 \mathbf{H}_0^\dagger\} = N_t N_r$ and the result of Lemma 8.1. On the other hand, the second term of (10.238) can be obtained directly from Lemma 8.2. As such,

$$\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H}\right\}\right\} = N_t \left[D_1(N_r, \rho_I) + N_r - 1\right]. \tag{10.240}$$

Now, it remains to derive the expression for $\mathtt{E}\{\mathsf{tr}\{(\mathbf{H}^\dagger \mathbf{\Lambda} \mathbf{H})^2\}\}$. Utilizing the zero mean property of $\mathbf{H}_w$

and after some basic algebraic manipulations, we have

$$\mathtt{E}\left\{\mathsf{tr}\left\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\right\}\right\} = \frac{1}{(K+1)^2}\mathtt{E}\left\{\mathsf{tr}\left\{\left(\mathbf{H}_w\mathbf{H}_w^\dagger\mathbf{\Lambda}\right)^2\right\}\right\} + \frac{K^2}{(K+1)^2}\mathtt{E}\left\{\mathsf{tr}\left\{\left(\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}\right)^2\right\}\right\} + \frac{2K}{(K+1)^2}\left(\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w\mathbf{H}_w^\dagger\mathbf{\Lambda}\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}\right\}\right\} + \mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w^\dagger\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_0^\dagger\mathbf{\Lambda}\mathbf{H}_0\right\}\right\}\right). \quad (10.241)$$

The first term can be easily obtained directly from Lemma 8.2. Therefore, here, we focus on the last three terms. With the help of [56, Lemma 3], we compute the second term as

$$\mathtt{E}\left\{\mathsf{tr}\left\{\left(\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}\right)^2\right\}\right\} = \frac{\mathsf{tr}\left\{(\mathbf{H}_0\mathbf{H}_0^\dagger)^2\right\}}{N_r^2-1}\left(\mathtt{E}\left\{\mathsf{tr}^2\left\{\mathbf{\Lambda}\right\}\right\} - \frac{1}{N_r}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}^2\right\}\right\}\right) + \frac{N_t^2N_r^2}{N_r^2-1}\left(\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}^2\right\}\right\} - \frac{1}{N_r}\mathtt{E}\left\{\mathsf{tr}^2\left\{\mathbf{\Lambda}\right\}\right\}\right). \quad (10.242)$$

Similarly, the third and fourth terms can be obtained as

$$\begin{aligned}
\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w\mathbf{H}_w^\dagger\mathbf{\Lambda}\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}\right\}\right\} &= \frac{1}{N_r}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w\mathbf{H}_w^\dagger\right\}\right\}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}^2\right\}\right\} && (10.243)\\
&= N_t\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0\mathbf{H}_0^\dagger\mathbf{\Lambda}^2\right\}\right\} && (10.244)\\
&= \frac{N_t}{N_r}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0\mathbf{H}_0^\dagger\right\}\right\}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}^2\right\}\right\} && (10.245)\\
&= N_t^2\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}^2\right\}\right\}, && (10.246)
\end{aligned}$$

and

$$\begin{aligned}
\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w^\dagger\mathbf{\Lambda}\mathbf{H}_w\mathbf{H}_0^\dagger\mathbf{\Lambda}\mathbf{H}_0\right\}\right\} &= \frac{1}{N_r}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}\right\}\right\}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w\mathbf{H}_0^\dagger\mathbf{\Lambda}\mathbf{H}_0\mathbf{H}_w^\dagger\right\}\right\} && (10.247)\\
&= \frac{1}{N_tN_r}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}\right\}\right\}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_w\mathbf{H}_w^\dagger\right\}\right\}\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{H}_0^\dagger\mathbf{\Lambda}\mathbf{H}_0\right\}\right\} && (10.248)\\
&= N_t\mathtt{E}^2\left\{\mathsf{tr}\left\{\mathbf{\Lambda}\right\}\right\}. && (10.249)
\end{aligned}$$

As a result,

$$\mathtt{E}\left\{\mathsf{tr}\left\{(\mathbf{H}^\dagger\mathbf{\Lambda}\mathbf{H})^2\right\}\right\} = \frac{1}{(K+1)^2}\left[\frac{K^2\left(\mathsf{tr}\left\{(\mathbf{H}_0\mathbf{H}_0^\dagger)^2\right\} - N_t^2N_r\right)}{N_r^2-1} + N_t\right]\mathtt{E}\left\{\mathsf{tr}^2\left\{\mathbf{\Lambda}\right\}\right\} + \frac{1}{(K+1)^2}\left[\frac{K^2\left(N_t^2N_r^3 - \mathsf{tr}\left\{(\mathbf{H}_0\mathbf{H}_0^\dagger)^2\right\}\right)}{N_r(N_r^2-1)} + (1+2K)N_t^2\right]\mathtt{E}\left\{\mathsf{tr}\left\{\mathbf{\Lambda}^2\right\}\right\} + \frac{2KN_t}{(K+1)^2}\mathtt{E}^2\left\{\mathsf{tr}\left\{\mathbf{\Lambda}\right\}\right\}. \quad (10.250)$$

Finally, applying Lemma 8.1 yields the desired result.

E.4. Proof of Corollary 8.1

Define the function $f(N_r) \triangleq N_r - 1 + A_0$. Hence, we are required to prove that $f(N_r)$ is an increasing function of N_r which we do by considering

$$f(N_r+1) - f(N_r) = 1 + D_1(N_r+1, \rho_I) - D_1(N_r, \rho_I). \tag{10.251}$$

With the help of (10.226), we can bound $f(N_r+1) - f(N_r) \geq 0$ by

$$\begin{aligned} f(N_r+1) - f(N_r) &> 1 + \frac{1}{1+\rho_I(N_r+1)} - \frac{1}{1+\rho_I(N_r-1)} && (10.252)\\ &= 1 - \frac{2\rho_I}{(1+\rho_I N_r)^2 - \rho_I^2} && (10.253)\\ &\geq 1 - \frac{2\rho_I}{1+2\rho_I} > 0, && (10.254) \end{aligned}$$

which completes the first half of the proof.

To prove the corresponding part for ρ_I, we define another function $g(\rho_I)$ as $g(\rho_I) = D_1(N_r, \rho_I)$ and then show that $g(\rho_I)$ is monotonically decreasing. To do so, we compute the first derivative of $g(\rho_I)$ with the help of the derivative formula of a confluent hypergeometric function [4, (13.4.20)]

$$\begin{aligned} g'(\rho_I) &= N_r \rho_I^{-N_r-2} \Psi(N_r+1, N_r+1, \rho_I^{-1}) - N_r \rho_I^{-N_r-1} \Psi(N_r, N_r, \rho_I^{-1}) && (10.255)\\ &< N_r \rho_I^{-1} \frac{1}{1+\rho_I N_r} - N_r \rho_I^{-1} \frac{1}{1+\rho_I N_r} = 0. && (10.256) \end{aligned}$$

Hence, $g(\rho_I)$ is a monotonic decreasing function. Therefore, we have

$$g(\rho_I \to \infty) \leq g(\rho_I) \leq g(\rho_I \to 0). \tag{10.257}$$

With the help of Lemma 8.3, we have $g(\rho_I \to 0) = 1$ and $g(\rho_I \to \infty) = 0$. As a consequence, the increase in $E_b/N_{0\min}$ is bounded by

$$\frac{\ln 2}{N_r - 1 + g(\rho_I \to \infty)} - \frac{\ln 2}{N_r - 1 + g(\rho_I \to 0)} = \ln 2 \left(\frac{1}{N_r} - \frac{1}{N_r - 1} \right) = \frac{\ln 2}{N_r(N_r-1)}, \tag{10.258}$$

which completes the proof.

E.5. Proof of Corollary 8.4

The first derivative of S_0 with respect to N_t can be obtained as

$$S_0^{'}(N_t) = \frac{2(K+1)^2 D_1(1,\rho_I)^2 (2K + D_2(1,\rho_I))}{(2K + (1+N_t(K+1)^2 D_2(1,\rho_I)))^2} \geq 0, \tag{10.259}$$

which has proved the first claim. Similarly, the first derivative of S_0 with respect to K is given by

$$S_0^{'}(K) = \frac{4N_t(K+1)D_1(1,\rho_I)^2(K+D_2(1,\rho_I)-1)}{(2K+(1+N_t(K+1)^2D_2(1,\rho_I)))^2}. \tag{10.260}$$

To complete the proof, we further have

$$D_2(1,\rho_I) = \frac{1}{\rho_I}\int_0^\infty e^{-\frac{1}{\rho_I}x}(1+x)^{-2}dx = \int_0^\infty e^{-x}(1+\rho_I x)^{-2}dx \leq \int_0^\infty e^{-x}dx = 1. \tag{10.261}$$

Because of the fact that $0 \leq D_2(1,\rho_I) \leq 1$, we have $S_0'(K) > 0$ if $0 \leq K < 1 - D_2(1,\rho_I)$ and $S_0'(K) \leq 0$ if $K \geq 1 - D_2(1,\rho_I)$, which has proved the second claim.

E.6. Proof of Theorem 8.2

Following the same steps as in the proof of Theorem 8.1, for Rayleigh-product MIMO channels, it can be derived that

$$\frac{E_b}{N_0}_{\min} = \frac{N_t \ln 2}{\mathtt{E}\left\{\frac{1}{N_s}\mathtt{tr}\left\{\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1\mathbf{H}_2\right\}\right\}}, \tag{10.262}$$

$$S_0 = \frac{2\mathtt{E}^2\left\{\frac{1}{N_s}\mathtt{tr}\left\{\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1\mathbf{H}_2\right\}\right\}}{\mathtt{E}\left\{\frac{1}{N_s^2}\mathtt{tr}\left\{\left(\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1\mathbf{H}_2\right)^2\right\}\right\}}. \tag{10.263}$$

First define $\mathbf{W} \triangleq \mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1$, conditioned on $\mathbf{W}$, we get

$$\mathtt{E}\left\{\frac{1}{N_s}\mathtt{tr}\left\{\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1\mathbf{H}_2\right\}\middle|\,\mathbf{W}\right\} = \frac{1}{N_s}\mathtt{E}\left\{\mathtt{tr}\left\{\mathbf{H}_2^\dagger\mathbf{W}\mathbf{H}_2\right\}\middle|\,\mathbf{W}\right\} = \frac{N_t}{N_s}\mathtt{tr}\{\mathbf{W}\}. \tag{10.264}$$

Similarly, we have

$$\mathtt{E}\left\{\frac{1}{N_s^2}\mathtt{tr}\left\{\left(\mathbf{H}_2^\dagger\mathbf{H}_1^\dagger\mathbf{\Lambda}\mathbf{H}_1\mathbf{H}_2\right)^2\right\}\middle|\,\mathbf{W}\right\} = \frac{1}{N_s^2}\mathtt{E}\left\{\mathtt{tr}\{(\mathbf{H}_2^\dagger\mathbf{W}\mathbf{H}_2)^2\}\middle|\,\mathbf{W}\right\} \tag{10.265}$$

$$= \frac{1}{N_s^2}\left(\mathtt{tr}\{\mathbf{I}\}^2\mathtt{tr}\{\mathbf{W}^2\} + \mathtt{tr}^2\{\mathbf{W}\}\mathtt{tr}\{\mathbf{I}\}\right) \tag{10.266}$$

$$= \frac{1}{N_s^2}\left(N_t^2\mathtt{tr}\{\mathbf{W}^2\} + N_t\mathtt{tr}^2\{\mathbf{W}\}\right) \tag{10.267}$$

With the help of Lemma 8.2 and further taking expectation on $\mathbf{W}$ in (10.264) and (10.267), the desired result can be obtained after some basic algebraic manipulations.

References

[1] 3GPP technical specification release 8, http://www.3gpp.org/Release-8

[2] Alamouti, S. M., "A simple transmit diversity scheme for wireless communications," *IEEE Journal Selected Areas in Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.

[3] Biglieri, E., J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[4] Abramowitz, M. and I. A. Stegun, *Handbook of mathematical functions*, New York: Dover Publication Inc., 1974.

[5] Alfano, G., A. Lozano, A. M. Tulino, and S. Verdú, "Capacity of MIMO channels with one-sided correlation," in Proc. *IEEE Int. Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, pp. 515–519, Sydney, Australia, Aug. 2004.

[6] Alfano, G., A. Lozano, A. M. Tulino, and S. Verdú, "Mutual information and eigenvalue distribution of MIMO Ricean channels," in Proc. *ISITA 04*, pp. 1040–1045, Parma, Italy, Oct. 2004.

[7] Almers, P., F. Tufvesson and A. F.Molisch, "Measurements of keyhole effect in a wireless multiple-input multiple-output (MIMO) channel," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 373–375, Aug. 2003.

[8] Almers, P., F. Tufvesson and A. F. Molisch, "Keyhole effects in MIMO wireless channels - measurements and theory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3596–3604, Dec. 2006.

[9] Boland, P. J., F. Proschan, and Y. L. Tong, *Some majorization inequalities for functions of exchangeable random variables*, National Technical Information Service, Access Number: ADA188207, Oct. 1987.

[10] Bolcskei, H., M. Borgmann and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 427–439, Apr. 2003.

[11] Bolcskei, H., R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. on Wireless Commun.*, vol. 5, no. 6, pp. 1433–1444, Jun. 2006.

[12] Borade, S. P., L. Zheng, and R. G. Gallager, "Maximizing degrees of freedom in wireless networks," in Proc. *Allerton Annual Conference on Communication, Control and Computing*, pp. 561–570, Illinois, USA, 2003.

[13] Borade, S. P., L. Zheng, and R. G. Gallager, "Amplify-and-forward in wireless relay networks: Rate, diversity, and network size," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3302–3318, Oct. 2007.

[14] Chiani, M., M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[15] Chuah, C., 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, no. 3, pp. 637–650, Mar. 2002.

[16] Chizhik, D., G. J. Foschini and R. A. Valenzuela, "Capacity of multi-element transmit and receive antennas: Correlations and keyholes," *Elect. Letters*, vol. 36, no. 13, pp. 1099–1100, Jun. 2000.

[17] Cover, M. T. and J. A. Thomas, *Elements of Information Theory*, Wiley-interscience, New York, 1991.

[18] Cui, X. W., Q. T. Zhang and Z. M. Feng, "Generic procedure for tightly bounding the capacity of MIMO correlated rician fading channels," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 890–898, May 2005

[19] Evans, M., N. Hastings, and B. Peacock, *Statistical distributions*, 3rd ed. New York: Wiley, 2000.

[20] Farrokhi, F. R., G. J. Foschini, A. Lozano, and R. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Letters*, vol. 5, no. 3, pp. 85–87, Mar. 2001.

[21] Fisher, R. A. "The sampling distribution of some statistics obtained from non-linear equations." *Annals of Eugenics*, vol. 9, pp. 238–249, 1939.

[22] Foschini, G. J., "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Lab. Tech. Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[23] Foschini, G. J. and M. J. Gans "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Per. Commun.*, Vol. 6, pp. 311–335, Mar. 1998.

[24] Fraidenraich, G., O. Leveque and J. M. Cioffi, "On the MIMO channel capacity for Nakagami-m channel," *IEEE Trans. Inform. Theory*, vol.54, no.8, pp. 3745–3751, Aug. 2008.

[25] Gesbert, D., H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[26] Gradshteyn, I. S. and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Orlando, FL: Academic Press, 5th ed., 1994.

[27] Gao, H., P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh fading additive interference channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 666–672, May 1998.

[28] Grant, A. "Rayleigh fading multi-antenna channels," *EURASIP J. Applied Signal Processing*, vol. 3, pp. 316–329, 2002.

[29] Gong, Y. and K. B. Letaief, "Space-time block codes in keyhole fading channels: error rate analysis and performance results," in Proc. *IEEE Veh. Tech. Conf.(VTC) Spring*, vol. 4, pp. 1903–1907, Melbourne, Australia, 2006.

[30] Horn, R. A. and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1985.

[31] IEEE 802.11 standard, http://www.ieee802.org/11.

[32] IEEE 802.15 standard, http://www.ieee802.org/15.

[33] IEEE 802.16 standard, http://www.ieee802.org/16.

[34] James, A. T., "Distribution of matrix variates and latent roots derived from normal samples." *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.

[35] Jayaweera, S. K. and V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.

[36] Jin, S., X. Gao and X. You, "On the ergodic capacity of rank-1 Rician-fading MIMO channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 502–517, Feb. 2007.

[37] Jin, S., M. R. McKay, K. K. Wong, and X. Gao, "Transmit beamforming in Rayleigh product MIMO channels: Capacity and performance analysis," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 5204–5221, Oct. 2008.

[38] Jin, S., M. R. McKay, C. Zhong, and K. K. Wong, "Ergodic capacity analysis of Amplify-and-Forward MIMO dual-hop systems," accepted in *IEEE Trans. Inform. Theory*, 2009.

[39] Jin, S., M. R. McKay, K. K. Wong and X. Li, "Low SNR capacity of double-scattering MIMO channels with transmitter channel Knowledge," in Proc. *IEEE Int. Conf. Commun.(ICC)*, Dresden, German, 2009.

[40] Jongren, G., M. Skoglund and B. Ottersen, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. on Inform. Theory*, vol, 48, no. 3, pp. 611–627, Mar. 2002.

[41] Jorswieck, E. A. and H. Boche, "Multiple-antenna capacity in the low-power regime:channel knowledge and correlation," in Proc. *IEEE Int. Conf. Acoustic, Speech, and Sig. Proc.(ICASSP)*, pp. 385–388, Philadelphia, USA, 2005.

[42] Jorswieck, E. A. and H. Boche, "Optimal transmission strategies and impact of correlation in multi-antenna systems with different types of channel state information", *IEEE Trans. on Signal Processing*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[43] Kang, M. and M. -S. Alouini, "Impact of correlation on the capacity of MIMO channels," in Proc. *IEEE Int. Conf. on Commun.(ICC)*, pp. 2623–2627, Paris, France, Jun. 2003.

[44] Kang, M. and M. S. Alouini, "Quadratic forms in complex Gaussian matrices and performance analysis of MIMO systems with co-channel interference," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 418–431, Mar. 2004.

[45] Kang, M. and M. S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[46] Kang, M. and M. -S. Alouini, "Capacity of correlated MIMO Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 143–155, Jan. 2006.

[47] Kang, M., L. Yang, and M. S. Alouini, "Outage probability of MIMO optimum combining in the presence of unbalanced co-channel interferers and noise," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1661–1668, Jul. 2006.

[48] Khatri, C. G. "Classical statistical analysis based on a certain multivariate complex Gaussian distribution," *Ann. Math. Statist.*, vol. 36, no. 1. pp. 98–114, Feb. 1965.

[49] Laneman, J. N. and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless nertworks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[50] Laneman, J. N., D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[51] Levin, G. and S. Loyka, "Capacity Distribution of a correlated keyhole channel," in Proc. *Canadian workshop on inform. theory(CWIT)*, pp. 319-322, Montreal, QC, Jun. 2005.

[52] Levin, G. and S. Loyka, "Multi-keyhole and measure of correlation in MIMO channels," in Proc. *23rd Biennial Symp. on Commun.*, Kingston, May 2006.

[53] Levin, G. and S. Loyka, "Multi-keyhole MIMO channels: Asymptotic analysis of outage capacity," in Proc. *Int. Symp. on Inf. Theory(ISIT)*, pp.1305-1309, Seattle, WA, Jul. 2006.

[54] Levin, G. and S. Loyka, "On the outage capacity distribution of correlated keyhole MIMO channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 3232–3245, Jul. 2008.

[55] Lo, T. K. Y., "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.

[56] Lozano, A., A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[57] Lozano, A., A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[58] Mallik, R. K., "The pseudo-wishart distribution and its application to mimo systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2761–2769, Oct. 2003.

[59] Mallik, R. K., M. Z. Win, M. Chiani, and A. Zanella, "Bit-error probability for optimum combining of binary signals in the presence of interference and noise," *IEEE Trans. on Wireless Commun.*, vol. 3, no. 2, pp. 395–407, Mar. 2004.

[60] Mallik, R. K. and Q. T. Zhang, "Optimum combining with correlated interference," *IEEE Trans. on Wireless Commun.*, vol. 4, no. 5, pp. 2340–2348, Sep. 2005.

[61] Mallik, R. K., "Optimized diversity combining with imperfect channel estimation," *IEEE Trans. on Inform. Theory*, vol. 52, no. 3, pp. 1176–1184, Mar. 2006.

[62] Marshall, A. W. and I. Olkin, *Inequalities: Theory of majorization and its applications*, Academic Press, New York, 1979.

[63] McKay, M. R. and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.

[64] McKay, M. R., A. Zanella, I. B. Collings, and M. Chiani, "Error probability and SINR analysis of optimum combining in Rician fading,"*IEEE Trans. Commun.*, vol. 57, no. 3, pp. 676–687, Mar. 2009.

[65] Morgenshtern, V. I. and H. Bolcskei, "Random matrix analysis of large relay networks," in Proc. *Allerton Conf. Comm.*, pp. 106–112, Illinois, USA, Sep. 2006.

[66] Morgenshtern, V. I. and H. Bolcskei, "Crystallization in large wireless networks," *IEEE Trans. on Inform. Theory*, vol. 53, no. 10, pp. 3319–3349, Oct. 2007.

[67] Muirhead, R. J. *Aspects of Multivariate Statistical Theory*. New York: John Wiley and Sons, 1982.

[68] Muller, A. and J. Speidel, "Bit error rate analysis of orthogonal space-time block codes in Nakagami-m keyhole channels," in Proc. *IEEE Intern. Conf. on Commun. and Networking in China (CHINACOM)*, pp. 1–5, Beijing, China, Oct. 2006.

[69] Muller, A. and J. Speidel, "Characterization of mutual information of spatially correlated MIMO channels with keyhole," in Proc. *Intern. Conf. Commun.(ICC)*, pp. 750–755, Glasgow, Scotland, Jun. 2007.

[70] Nabar, R. U., H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE J. Selected Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[71] Nakagami, M. “The m-distribution–A general formula for intensity distribution of rapid fading,” in *Statistical methods in Radio Wave Propagation*, W. G. Hoffman, Ed. Oxford, U.K.: Pergamon, 1960.

[72] Oyman, O., R. U. Nabar, H. Bolcskei, and A. J. Paulraj, “Characterizing the statistical properties of mutual information in MIMO channels,” *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.

[73] Palomar, D., J. M. Cioffi and M. A. Lagunas, “Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization,” *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381–2401, 2003.

[74] Park, M., C. Chae and R. W. Heath Jr. “Ergodic capacity of spatial multiplexing MIMO systems with ZF receivers for log-normal shadowing and Rayleigh fading channels,” in Proc. *IEEE Int. Sym. Personal, Indoor and Mobile Radio Commun.(PIMRC)*. Athens, 2007.

[75] Proakis, J. G. *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[76] Prudnikov, A. P., Y. A. Brychkov and O. I. Marichev, *Integrals and series, Volume 3: More special functions.* New York: Gordon and Breach Science Publishers, 1990.

[77] Rao, C. and B. Hassibi, “Analysis of multiple antenna wireless links at low SNR,” *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2123–2130, Sep. 2004.

[78] Sanayei, S., A. Hedayat and A. Nosratinia, “Space-time codes in keyhole channels: analysis and design,” in Proc. *IEEE Globecom Conf.(GLOBECOM)*, vol. 6, pp. 3768-3772, Dallas, TX, Dec. 2004.

[79] Sendonaris, A., E. Erkip, and B.Aazhang, “User cooperation diversity-Part I: system description,” *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[80] Sendonaris, A., E. Erkip, and B.Aazhang, “User cooperation diversity-Part II: implementation aspects and performance analysis,” *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[81] Shah, A. and A. M. Haimovich, “Performance analysis of optimum combining in wireless communication with Rayleigh fading and cochannel interference,” *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 473–479, Apr. 1998.

[82] Shin, H. and J. H. Lee, “Capacity of multi-antenna fading channels: Spatial fading correlation, double scattering, and keyhole,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[83] Shin, H. and J. H. Lee, “Effect of keyholes on the symbol error rate of space-time block codes,” *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 27–29, Jan. 2003.

[84] Shin, H., M. Z. Win, J. H. Lee and M. Chiani, “On the capacity of doubly correlated MIMO Channels,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2265, Aug. 2006.

[85] Shin, H., and M. Z. Win, "MIMO diversity in presence of double scattering," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008.

[86] Shannon, C. E. and W. Weaver, "A Mathematical Theory of Communication," Urbana, IL: Univ. of Illinois Press, 1949.

[87] Simon, M. G. and M. S. Alouini, *Digital communications over fading channels*, New York: John Wiley & Sons, 2000.

[88] Smith, P. J., L. M. Garth and S. Loyka, "Exact capacity distribution for MIMO systems with small numbers of antenna," *IEEE Commun. Lett.*, vol. 7, no. 10, pp. 481–483, Oct. 2003.

[89] Smith, P. J., S. Roy, and M. Shafi, "Capacity of MIMO systems with semi-correlated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[90] Suzuki, H. "A statistical model for urban radio propagation," *IEEE Trans. Commun.*, vol. 25, no. 7, pp. 673–679, Jul. 1977.

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

[92] Tarokh, V., H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. on Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[93] Tulino, A. M., A. Lozano, and S. Verdu, "The impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[94] Telatar, I. E. "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecommun.*, vol 10, pp. 585–596. Nov.-Dec. 1999.

[95] Verdú, S. "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[96] Vu, M. and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with nonzero mean in space-time coded systems," *IEEE Trans. Signal Processing*, vo. 54, no. 6, pp. 2318–2336, Jun. 2006.

[97] Wagner, J., B. Rankov, and A. Wittneben, "On the asymptotic capacity of the Rayleigh fading amplify-and-forward MIMO relay channel," in Proc. *IEEE Int. Symp. on Information Theory (ISIT)*, pp. 2711–2715, Nice, France, Jun. 2007.

[98] Wagner, J., B. Rankov, and A. Wittneben, "Large n analysis of amplify-and-forward MIMO relay channels with correlated rayleigh fading," to appear in *IEEE Trans. on Inform. Theory*, 2008.

[99] Wang, B., J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[100] Winters, J. H., "Optimum combining in digital mobile radio with cochannel interference," *IEEE Journal Selected Areas in Commun.*, vol. SAC-2, no. 4, pp. 528–539, Jul. 1984.

[101] Wittneben, A. and B. Rankov, "Impact of cooperative relays on the capacity of rank-deficient MIMO channels," in Proc. *Mobile and Wireless Communications Summit (IST)*, pp. 421–425, Aveiro, Potugal, 2003.

[102] Wolfram function site, http://functions.wolfram.com.

[103] Wong, K. K., R. S. K. Cheng, K. B. Letaief, and R. D. Murch, "Adaptive antennas at the mobile and base stations in an OFDM/TDMA system," *IEEE Trans. Commun*, vol. 49, no. 1, pp. 195–206, Jan. 2001.

[104] Yang, S. and J. C. Belflore, "Diversity-multiplexing tradeoff of double scattering MIMO channels," submitted to *IEEE Trans. Inform. Theory*, Mar. 2006. Available at http://arxiv.org/abs/cs/0603124.

[105] Yeh, S. and O. Leveque, "Asymptotic capacity of multi-level amplify-and-forward relay networks," in Proc. *IEEE Int. Symp. on Information Theory (ISIT)*, pp. 1436–1440, Nice, France, Jun. 2007.

[106] Yue, D. and Q. Zhang, "Outage Performance analysis of correlated MIMO optimum combining systems with and without co-channel interference," in *Proc. IEEE Info. Theory Workshop (ITW)*, pp. 269–273, Chendu, China, Oct. 2006.

[107] Zhang, Q. T. and D. P. Liu, "A simple capacity formula for correlated diversity Rician fading channles," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 708–711, Nov. 2002.

[108] Zhang, Q. T., X. W. Cui, and X. M. Li, "Very tight capacity bounds for MIMO-correlated Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 681–688, Mar. 2005.

[109] Zhang, W. and J. N. Laneman, "Benefits of spatial correlation for multi-antenna non-coherent communication over fading channels at low SNR," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 887–896, Mar. 2007.

[110] Zheng, F. and T. Kaiser, "On the channel capacity of multiantenna systems with Nakagami fading," *EURASIP J. Applied Sig. Proc.*, vol. 2006, article id 39436, pp. 1–11, 2006.

[111] Zheng, L. and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Thoery*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[112] Zhong, C., S. Jin and K. K. Wong, "MIMO rayleigh-product channels with co-channel interference," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1824–1835, Jun. 2009.