University of Alberta

Unitary Integrations for Unified MIMO Capacity and Performance Analysis

by

Alireza Ghaderipoor

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering

Fall 2009
Edmonton, Alberta

Examining Committee

Chintha Tellambura, Electrical and Computer Engineering

Masoud Ardakani, Electrical and Computer Engineering

Hai Jiang, Electrical and Computer Engineering

Peter Hooper, Mathematical and Statistical Sciences

Teng Joon Lim, Electrical and Computer Engineering, University of Toronto

To my wife, mother and father ...

.

" Those who turn in a glance
dust into gold
Shall they ever our sight behold? "

— Hafez Shirazi

Abstract

Integrations over the unitary group are required in many applications including the joint eigenvalue distributions of the Wishart matrices. In this thesis, a universal integration framework is proposed to use the character expansions for any unitary integral with general rectangular complex matrices in the integrand. The proposed method is applied to solve some of the well–known but not solved in general form unitary integrals in their general forms, such as the generalized Harish–Chandra–Itzykson–Zuber integral. These integrals have applications in quantum chromodynamics and color–flavor transformations in physics. The unitary integral results are used to obtain new expressions for the joint eigenvalue distributions of the semi–correlated and full–correlated central Wishart matrices, as well as the i.i.d. and uncorrelated noncentral Wishart matrices, in a unified approach. Compared to the previous expressions in the literature, these new expressions are much easier to compute and also to apply for further analysis. In addition, the joint eigenvalue distribution of the full–correlated case is a new result in random matrix theory. The new distribution results are employed to obtain the individual eigenvalue densities of Wishart matrices, as well as the capacity of multiple–input multiple–output (MIMO) wireless channels. The joint eigenvalue distribution of the i.i.d. case is used to obtain the largest eigenvalue density and the bit error rate (BER) of the optimal beamforming in finite–series expressions. When complete channel state information is not available at the transmitter, a codebook of beamformers is used by the transmitter and the receiver. In this thesis, a codebook design method using the genetic algorithm is proposed, which reduces the design complexity and achieves large minimum–distance codebooks. Exploiting the specific structure of these beamformers, an order and bound algorithm is proposed to reduce the beamformer selection complexity at the receiver side. By employing a geometrical approach, an approximate BER for limited feedback beamforming is derived in finite–series expressions.

Acknowledgements

This is a great opportunity to express my respect and gratitude to my supervisor, Dr. Chintha Tellambura. The first thing that his name brings to mind is his kind and caring personality. I would like to thank him for inspiring me all along the way, as well as for his supervision and support during my PhD.

The financial support for my PhD has been provided mainly by the Alberta Ingenuity Fund and *i*CORE and partly by NSERC.

I would like to thank Dr. Norman C. Beaulieu and Dr. Masoud Ardakani for their kind support. As a member of the *i*CORE Wireless Communications Lab (*i*WCL) at the University of Alberta, I would like to thank the students, postdoctoral fellows and administrative staff for making *i*WCL a friendly environment for research, discussions, seminars and occasional conference trips. Here, I take the opportunity to thank Sandra Abello for her caring assistance.

For this thesis, I was able to discuss with Dr. Plamen Koev on random matrix theory, and Dr. A. B. Balantekin on character expansions of groups. I thank them for their helpful comments.

I would like to thank Moslem Noori, Mahdi Hajiaghayi and Payam Dehghani for their collaboration and occasional discussions, which provided helpful comments and feedback for my research.

I would like to thank my master degree supervisor, Dr. Hamidreza Jamali, for his kind support, as well as all my friends, with whom I had memorable times, in particular, Mohsen Eslami, Amirmasoud Rabiei, Alireza Farhangfar, Seyed Ali Arefifar, Mahdi Shahbakhti and Mahdi Ramezani.

And finally, my last and the biggest thank yous go to my wife for her invaluable support and patience, and my family for their endless support from thousands of miles away. This thesis is dedicated to my wife, mother and father.

Table of Contents

List of Figures

List of Tables

List of Symbols

$\mathbf{r}_N$	Representation of $\mathrm{GL}(N, \mathcal{C})$	14
$d_{\mathbf{r}_N}$	Dimension of the representation $\mathbf{r}_N$	14
$\det_M[f(i,j)]$	Determinant of a $M \times M$ matrix with the (i,j)-th element given by $f(i,j)$	14
$\mathbf{X}^{(\mathbf{r}_N)}$	Representation matrix of $\mathbf{X}$ under representation $\mathbf{r}_N$	15
$\chi_{\mathbf{r}_N}$	Character function of representation $\mathbf{r}_N$	15
$\Delta_N(\mathbf{x})$	Vandermonde determinant of the N–dimensional vector $\mathbf{x}$	15
$\mathrm{D}\mathbf{U}$	Standard Haar measure of $\mathcal{U}(N)$	15
$\mathrm{etr}\{\mathbf{X}\}$	$\exp(\mathrm{tr}\{\mathbf{X}\})$	17
$\alpha_{\mathbf{r}_N}$	Character expansion coefficient	17
$I_n(\cdot)$	Modified Bessel function	31
$u(\cdot)$	Step function	41
$\gamma(\alpha, x)$	The lower incomplete Gamma function	42
$\Gamma(\alpha)$	Gamma function	45
$\Phi(a, b; x)$	The confluent hypergeometric function	48
$\Gamma(\alpha, x)$	The upper incomplete Gamma function	54
$\mathcal{I}$	The mutual information between the transmitter and receiver	59
$g(z)$	Moment generating function of $\mathcal{I}$	59
$\lvert x \rvert$	Absolute value of x	73
$\boldsymbol{\mathcal{F}}$	The codebook of beamformers	74
N_b	Number of feedback bits	74
L	The size of $\boldsymbol{\mathcal{F}}$	74
$\mathbb{I}$	The index of the best beamformer	74
$\mathrm{S}(\mathbf{a})$	Sign of the permutation vector $\mathbf{a}$	108

List of Abbreviations

Chapter 1

Introduction

In 1895, Guglielmo Marconi opened the way for modern wireless communications by transmitting the three-dot Morse code for the letter 'S' over a distance of three kilometers by using electromagnetic waves. From this beginning, wireless communications has been a rapidly growing segment of the communications industry, particularly in the last decade, and has been developed into a key element of modern society. From satellite transmission, radio and television broadcasting to the now ubiquitous multimedia internet-enabled mobile cell phones and laptops, wireless communications has revolutionized the way societies function. However, wireless systems face many technical challenges in overcoming resource limitations.

Frequency bandwidth has been always a constraint in wireless communications. Particularly after the realization of the first–generation digital cellular systems and due to the rapid growth in digital cell phone users, as well as the demand for high–speed wireless internet networks, bandwidth efficiency (bits/second/Hertz) became a primary challenge in wireless systems. Similarly, reducing the power consumption, or equivalently, increasing the battery life of mobile devices, such as cell phones, laptops, sensors, and others is another challenge in wireless communications. On the other hand, in his influential paper in 1948 [1], Shannon showed that in a single–input single–output (SISO) channel with white Gaussian noise, when the channel bandwidth is limited to B (Hertz), and the signal–to–noise power ratio (SNR) is limited to S/N, the channel capacity (C), i.e., the maximum achievable throughput, is limited to

$$C = B \log\left(1 + \frac{S}{N}\right) \tag{1.1}$$

where C is measured in bits per second if the logarithm is taken in base 2, or nats per second if the natural logarithm is used.

Since 1948, the state–of–the–art coding and modulation techniques have been developed to reach Shannon's fundamental limit [2]– [6]. However, due to the limited power in mobile devices and the related delay introduced by signal processing modules, high–performance coding techniques may not always be feasible in practice. Moreover, the signal power S in (1.1) is decayed in wireless channels due to fading and shadowing effects [7]– [9]. Hence, during the 1970s for satellite communications and, later, in the 1990s for digital cellular mobile systems, single–input multiple–output (SIMO) systems were developed to use multiple receiver antennas for diversity gains [10] and receiver beamforming [11]. However, even by increasing the number of receiver antennas to N_r, the improvement in the capacity in high SNR is approximately equal to $B\log(N_r)$, which is not adequate for high–speed wireless networks. These facts led researchers to use multiple antennas at the transmitter side [12].

1.1 MIMO Systems

Multiple–input multiple–output (MIMO) wireless systems deploy multiple antennas at both the transmitter and receiver [13]. Signals are transmitted simultaneously from transmit antennas using the same bandwidth. A MIMO wireless system with N_t transmit antennas and N_r receive antennas, denoted by MIMO(N_t, N_r), is illustrated in Figure 1.1, where the linear transformation between the transmit and receive antennas can be modeled as

$$\mathbf{Y} = \sqrt{\rho}\,\mathbf{H}\mathbf{S} + \mathbf{N} \tag{1.2}$$

where $\mathbf{Y} \in \mathcal{C}^{N_r \times T}$ is the complex received matrix, $\mathbf{S} \in \mathcal{C}^{N_t \times T}$ is the complex transmitted matrix, $\mathbf{N} \in \mathcal{C}^{N_r \times T}$ is the additive white noise matrix, and $\mathbf{H} \in \mathcal{C}^{N_r \times N_t}$ is the complex channel matrix. Here, T is the number of signaling intervals used for transmitting the matrix $\mathbf{S}$. The elements of matrix $\mathbf{S}$ satisfy the power constraint condition: $\mathrm{E}\{\mathbf{S}^*\mathbf{S}\} = \mathbf{I}_T$, where $(\cdot)^*$, $\mathrm{E}\{\cdot\}$, and $\mathbf{I}_T$ denote the Hermitian (transpose conjugate), expectation operator, and T dimensional identity matrix, respectively. The coefficient $\sqrt{\rho}$ ensures that the total transmit power in each channel use is ρ, independent of the number of transmit antennas. The elements of matrix $\mathbf{N}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, denoted by $\mathcal{CN}(0, 1)$. The elements of matrix $\mathbf{H}$ are the channel gains between each pair of transmit and receive antennas, and are

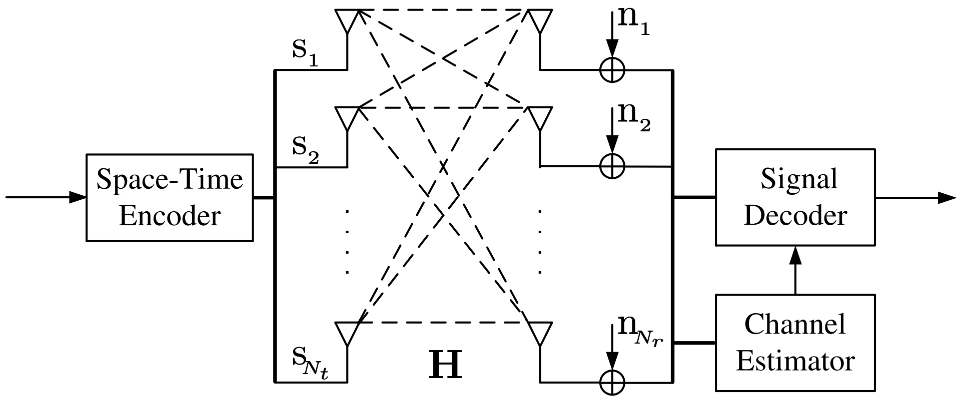

Figure 1.1: The block diagram of MIMO(N_t, N_r) wireless system.

modeled by complex Gaussian random variables.

When the elements of matrix **S** are selected independently according to the input data stream, the system is called spatial multiplexing. The first spatial multiplexing system, called Bell Laboratories Layered Space-Time (BLAST), was implemented at Bell Labs in 1996 by Foschini [14]. On the other hand, when the elements of matrix **S** are selected based on a coding strategy, space-time codes are obtained [15]– [19]. In 1995, Telatar [20] analyzed the capacity of MIMO systems by assuming that the channel gains are i.i.d. $\mathcal{CN}(0,1)$ random variables. He found that the capacity of a MIMO system increases linearly with the minimum number of transmit and receive antennas $\big(\min(N_t, N_r)\big)$. This intriguing discovery along with Foschini's laboratory prototype sparked the massive international research effort on MIMO wireless systems that continues today.

The system model in (1.2) shows the fundamental role of the channel matrix **H** in MIMO systems. In other words, the channel capacity directly depends on **H**. Moreover, the space–time code designs and the decoding methods for such codes depend on the availability of **H** for the transmitter and receiver. In general, the MIMO literature falls into three main categories based on the availability of channel state information (CSI) for the transmitter and receiver.

1.1.1 CSI available for neither transmitter nor receiver

In this scenario, **H** is not available for the transmitter and receiver. Data transmission is possible in this case if the transmit matrix **S** is a unitary matrix [21], [22]; i.e., $\mathbf{SS}^* = \mathbf{I}_{N_t}$. Some of the unitary space–time coding and decoding techniques proposed for MIMO systems with unknown $\mathbf{H}$ can be found in [23]– [29]. The capacity analysis of MIMO systems with unknown $\mathbf{H}$ can be found in [30]. This analysis is only for the case in which the channel gains are i.i.d. $\mathcal{CN}(0,1)$ random variables. This paper along with the analysis in [31] reveals that MIMO capacity analysis with unknown $\mathbf{H}$ requires unitary integrations for obtaining the joint pdf (probability density function) of the received matrix elements. Such unitary integrals cannot be solved by using the previous available mathematical tools. As a result, there are no results in the literature on the capacity of MIMO systems with unknown $\mathbf{H}$ other than the i.i.d. case.

1.1.2 CSI available for receiver only

The common practical assumption is that the channel matrix $\mathbf{H}$ is known to the receiver, an assumption that is realized by using training signals [32]. In the literature, when the availability of $\mathbf{H}$ is not explicitly stated, it is commonly assumed that the channel matrix in known to the receiver only. Most of the MIMO literature deals with this case. Such literature can be divided into three main categories. In the first category, the space–time code designs are developed, e.g., orthogonal codes [15], [17], quasi–orthogonal codes [33]– [36], and algebraic codes [37]– [41]. In the second category, the decoding techniques for MIMO systems are considered [42]– [47], where the focus is on solving the following NP–hard optimization problem by using the exact or approximate solutions:

$$\widetilde{\mathbf{S}} = \arg\min_{\forall \mathbf{S}} \|\mathbf{Y} - \mathbf{H}\mathbf{S}\|^2 \tag{1.3}$$

Here, $\|\cdot\|$ is the Frobenius norm [48]. The main challenge in space–time code design is to reach the promising system capacity. This challenge usually leads to a tradeoff between diversity and multiplexing [49]. In decoding techniques, however, the main challenge is a tradeoff between computation complexity and performance [50].

In the third category, the capacity and performance of MIMO systems are investigated. The joint distribution of channel gains comes into account for MIMO performance analysis (probability of detection error) [51]– [54]; e.g., the pairwise error probability of sending the transmit matrix $\mathbf{S}$ in (1.2) and detecting the matrix $\widetilde{\mathbf{S}}$ in (1.3) (maximum likelihood detection) is equal to

$$\mathrm{P}\Big(\mathbf{S} \to \widetilde{\mathbf{S}} \,\big|\mathbf{H}\Big) = Q\left(\sqrt{\frac{\rho}{2}\|\mathbf{H}(\mathbf{S} - \widetilde{\mathbf{S}})\|^2}\right) \tag{1.4}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} \, du \tag{1.5}$$

is the Q–function.

The joint eigenvalue distribution of matrix $\mathbf{HH}^*$ comes into account for MIMO capacity analysis [55] since the capacity of the system (1.2) is equal to

$$C = \mathrm{E}_{\boldsymbol{\lambda}}\Big\{ B \sum_{i=1}^{M} \log\big(1 + \rho\lambda_i\big) \Big\} \tag{1.6}$$

where $M = \min\{N_t, N_r\}$, and $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_M\}$ are the M nonzero eigenvalues of the matrix $\mathbf{HH}^*$.

Due to the complexity of MIMO performance analysis, most of the results in the literature are not exact solutions, but rather are based on the assumption of a large number of antennas [56] and/or a high SNR regime [16]. Similarly, most of the literature on MIMO capacity analysis relies on the large dimension (asymptotic) random matrix theory [57] and/or high SNR regime [58]. The exact capacity results are limited to some special cases, and almost all of them are based on the contributions by James [59].

1.1.3 CSI available for receiver and transmitter

In some wireless communication systems, CSI can be available at the transmitter side as well. This information can be full CSI [60]–[62] if the transmission is bidirectional (duplex) using the same channel, or, most commonly, partial CSI when a limited feedback channel is available. Feedback information is usually used to select a subset of transmit antennas [63]– [66] or to select a suitable precoder [67]– [72]. A MIMO system with precoding is modeled by

$$\mathbf{Y} = \sqrt{\rho}\,\mathbf{HFS} + \mathbf{N} \tag{1.7}$$

where $\mathbf{F} \in \mathcal{C}^{N_t \times M_t}$ is the complex precoder matrix, $\mathbf{S} \in \mathcal{C}^{M_t \times T}$ is the transmit matrix, and M_t is the number of transmit antennas of the virtual channel $\mathbf{HF}$. The optimum precoder can be obtained based on the eigen–structure of the channel matrix [60], [62]. As a result, the capacity and performance analysis of MIMO systems with precoding requires the JDE of the matrix $\mathbf{HH}^*$ as well.

The basic concepts of MIMO wireless systems were reviewed in this section, where it was shown that the eigenvalue distribution of matrix $\mathbf{HH}^*$ is a prerequisite for the capacity and performance analysis of MIMO systems.

1.2 Random Matrix Theory

Without a doubt, random matrix theory is the most influential part of advanced multivariate statistical analysis, which has been deeply affected by its applications in physics, mathematics, engineering, and other fields. Random matrix theory appears in fields as diverse as the Riemann hypothesis, stochastic differential equations, condensed matter physics, statistical physics, chaotic systems, numerical linear algebra, neural networks, multivariate statistics, information theory, signal processing, and small–world networks [57].

The history of random matrices goes back to the late 1920s when John Wishart, a Scottish agricultural statistician, computed the joint distribution of the elements of matrix $\mathbf{A}\mathbf{A}^T$ where the elements of matrix $\mathbf{A} \in \mathcal{R}^{N\times M}$ are i.i.d. Gaussian random variables with zero mean [73]. In honor of Wishart, matrices of the form $\mathbf{H}\mathbf{H}^*$ where $\mathbf{H}$ is a (real or complex) Gaussian random matrix are called (real or complex) Wishart matrices in the literature. When $\mathbf{H}$ is a zero mean Gaussian matrix, $\mathbf{H}\mathbf{H}^*$ is called a central Wishart matrix; otherwise, it is called noncentral. Similarly, when other attributes are mentioned for Wishart matrices, they reflect the original attributes of the Gaussian matrix $\mathbf{H}$.

Following Wishart's paper [73], the joint eigenvalue distribution of the central i.i.d. Wishart matrix was derived simultaneously and independently by Fisher [74], Hsu [75], Roy [76] and Girshick [77] in 1939. James [59] unified and developed all previous works on the joint eigenvalue distributions of Wishart matrices by defining the hypergeometric functions with matrix arguments. Most of the results in the literature on Wishart matrices, e.g., the distribution of the smallest eigenvalue, largest eigenvalue, ordered eigenvalues, a subset of eigenvalues, are based on James's paper.

Due to the difficulty of deriving and using the exact results, a large body of work has been published on the asymptotic analysis of random matrices. The first asymptotic results for large random matrices were obtained by Wigner in the 1950s in a series of papers [78]– [81] motivated by nuclear physics. To study heavy atoms, he proposed replacing the Schrödinger operator by a large random Hermitian matrix ($\mathbf{A}\mathbf{A}^* = \mathbf{I}$), hoping that its eigenvalues would correspond to the nucleus energy levels. Since then, research on the analysis of large dimensional random matrices has continued to attract interest in probability, statistics, physics and engineering.

Ref. [57] reports recent developments in the asymptotic theory of random matrices, particulary for the analysis of MIMO wireless systems.

As mentioned earlier, James [59] defined hypergeometric functions with matrix arguments to derive the joint eigenvalue distributions of some Wishart matrices. These hypergeometric functions, however, have some drawbacks:

1. Hypergeometric functions with matrix arguments are defined in terms of zonal polynomials, which are symmetric multivariate polynomials of the eigenvalues of a symmetric matrix. Unfortunately, "the zonal polynomials are notoriously difficult to compute" [82].

2. A close look at the papers using James' results for other applications reveals that hypergeometric functions with matrix arguments are not easy to use and manipulate. In addition, the results are again obtained in terms of hypergeometric functions with matrix arguments, which are difficult to compute [83].

3. By using hypergeometric functions with matrix arguments, James could derive the joint eigenvalue distributions of noncentral and semi–correlated Wishart matrices. A Wishart matrix is called semi–correlated if only the columns or only the rows of matrix $\mathbf{H}$ are correlated. To our best knowledge, the joint eigenvalue distribution of the full–correlated Wishart matrix has not been reported in the literature for a rectangular matrix $\mathbf{H}$ [57].

4. Deriving the joint eigenvalue distribution of a random matrix requires integrations over the unitary group [84]. Hence, the work by James can, in fact, be interpreted as a method for solving integrals over the unitary group. However, his approach to use hypergeometric functions with matrix arguments requires the integrand (argument) matrices to be positive definite [59].

5. Since hypergeometric functions with matrix arguments are not easy to manipulate, the final expressions of the unitary integrations with James' method may not be convenient for understanding phenomena in physics and engineering [85].

These facts motivated researchers, particularly in physics, to find other methods for unitary integrations.

1.3 Character Expansions for Unitary Integrations

Integrals over the unitary group have many applications in physics, e.g., quantum chromodynamics (QCD) [86], [87], flux line pinning in superconductors [88], scattering in quantum chaos [89], color-flavor transformation [90], as well as in mathematics [91] and in the theory of random matrices [92].

The character expansion method, introduced by A. B. Balantekin [86], is a powerful tool for solving unitary integrals. Although the expansion of an invariant function of a group into its characters (traces of the representation matrices) is intricate in general, Balantekin first presented a combinatorial formula to write the character expansions for the $\mathcal{U}(N)$ group[1] [93], and later in [86], he extended the results to more general situations than those covered in [93]. He applied character expansions to simplify the integrations over the unitary group and, in particular, to derive the well-known Harish-Chandra-Itzykson-Zuber integral [91], [94].

In [86], the coefficient matrices appearing in the integrand are nonzero–determinant square matrices. However, in some applications, such as the derivation of joint eigenvalue distribution of Wishart matrices, when $\mathbf{H}$ is not a square matrix, the resulting matrix in the integrand is not full rank and, therefore, is not a group member, so that character expansions cannot be applied.

In 2003, Schlittgen and Wettig [85] generalized two useful unitary integrals for color–flavor transformations. For the cases in which the coefficient matrices in the integrand are rectangular, these researchers considered the possibility of taking a limit of the final answer for the square coefficient matrices to obtain the result for the rectangular case. However, they realized that this approach is valid only for single unitary integrals.

Similarly, Simon, Moustakas and Marinelli [95] applied the character expansion method to obtain the joint eigenvalue distributions of Wishart matrices. These researchers first assumed that the matrix $\mathbf{H}$ is a square matrix, then followed Balantekin's approach exactly and, in the end, made some of eigenvalues to approach zero. However, they failed to obtain the correct joint eigenvalue distributions whenever double unitary integrals were involved.

[1] $\mathcal{U}(N)$ denotes the group of unitary matrices with dimension N.

1.4 Thesis Outline and Contributions

This thesis is organized as follows.

In Chapter 2, we briefly review the basic principles of the representation theory and characters of groups, and state the character expansions and the integration framework used by Balantekin [86]. Next, we generalize and develop essential tools for unitary integrations. We propose a universal integration framework to use character expansions for any unitary integral with general rectangular complex matrices in the integrand, without any specific assumptions about the matrices, e.g., to be diagonal, real, or positive definite.

In Chapter 3, we apply the proposed integration framework to solve some of the *well–known but not solved in general form* unitary integrals in their general forms, such as the generalization of the well–known Harish–Chandra–Itzykson–Zuber integral. These integrals have applications in quantum chromodynamics and color–flavor transformations in physics. The proposed integration method can be used to solve other unitary integrals accordingly.

We use the results of the unitary integrals to obtain new expressions for the joint eigenvalue distributions of the semi–correlated and full–correlated central Wishart matrices, as well as the i.i.d. and uncorrelated noncentral Wishart matrices, all in a unified approach. Compared to the previous expressions in the literature, these new expressions are much easier to compute and also to apply for further analysis. In addition, the joint eigenvalue distribution of the full–correlated central Wishart matrix is a new result in random matrix theory [57].

Due to the difficulty of manipulating the joint eigenvalue distribution results derived by James [59], and also due to the nature of most applications requiring the eigenvalue densities of large random matrices, the eigenvalue densities of Wishart matrices have been derived asymptotically in the literature [57]. However, because of the convenient mathematical form of the joint eigenvalue distributions derived in Chapter 3, we employ an innovative procedure to obtain the exact marginal eigenvalue densities of Wishart matrices in Chapter 4. Our results are in the form of finite summations of determinants, which can be easily computed and also employed for further analysis.

In Chapter 5, we use the results of Chapter 3 and 4 to obtain new expressions for the capacity of MIMO channels. Although the MIMO capacity results have been reported in the literature (using considerably more complicated methods), the approach presented in this thesis is significantly more simple and straightforward, and the new expressions can be easily computed.

In Chapter 6, we use the joint eigenvalue distribution of the i.i.d. central Wishart matrix to obtain the largest eigenvalue density and the bit error rate (BER) of the optimal transmit beamforming in finite–series expressions. MIMO systems achieve significant diversity and array gains by using transmit beamforming. When complete channel information is not available at the transmitter, a codebook of beamformers is used by both the transmitter and the receiver. For each channel realization, the best beamformer is selected at the receiver, and its index is sent back to the transmitter via a limited feedback channel. We propose a codebook design method by using the genetic algorithm, which reduces the design complexity, and achieves large minimum–distance codebooks. Exploiting the specific structure of these beamformers, we develop an order and bound algorithm to reduce the beamformer selection complexity at the receiver side. By employing a geometrical approach, we derive an approximate BER of limited feedback beamforming in finite–series expressions. Since limited feedback beamforming is a practical scenario in wireless communications, the approximate BERs are useful in the design of such systems.

Finally, in Chapter 7, we discuss the possible future work based on the contributions of this thesis.

Chapter 2

Preliminaries

In this chapter, the basic definitions from group theory and representation theory are briefly reviewed. The properties of characters and the character expansions for a few useful matrix expressions are presented. The original integration framework, presented by Balantekin [86] for unitary integrals with nonzero–determinant square coefficient matrices in the integrand, is stated and demonstrated by solving the Harish-Chandra-Itzykson-Zuber integral. Next, we propose a universal integration framework to use the character expansions for any unitary integral with general rectangular complex matrices in the integrand. No specific assumptions about the matrices, e.g., to be diagonal, real or positive definite, are made. To handle such unitary integrals, the essential algebraic machinery is developed.

2.1 Definitions

This section will briefly introduce a few relevant definitions and important results from group and representation theory. The interested reader may refer to several textbooks, e.g., [96]– [98], for more details.

Definition 2.1 : Assume $\mathcal{F}$ is the set of complex numbers ($\mathcal{C}$) or real numbers ($\mathcal{R}$). The set of all invertible $N \times N$ matrices with entries in $\mathcal{F}$, under matrix multiplication, forms a group. This group is called the general linear group of degree (or dimension) N over $\mathcal{F}$, and is denoted by $\mathrm{GL}(N, \mathcal{F})$. It is an infinite group with the identity matrix $\mathbf{I}_N$ as the identity element.

Definition 2.2 : The set of all $N \times N$ unitary matrices, under matrix multiplication, forms a group. This group is called the unitary group with degree (or dimension) N and is denoted by $\mathcal{U}(N)$. Clearly, the unitary group is a subgroup of

$\mathrm{GL}(N, \mathcal{C})$.

Definition 2.3 : If G and H are groups, those functions from G to H which "preserve the group structure" are called homomorphisms. Equivalently, a homomorphism from G to H is a function $\phi \colon G \to H$ which satisfies

$$\phi(g_1 g_2) = \phi(g_1)\, \phi(g_2) \quad \forall\, g_1, g_2 \in G.$$

2.1.1 Group Representations

A representation of a group G provides a way of visualizing G as a group of matrices. To be accurate, a representation is a homomorphism from G into a group of invertible matrices.

Definition 2.4 : Let G be a group and $\mathcal{F}$ be $\mathcal{R}$ or $\mathcal{C}$. A representation of G over $\mathcal{F}$ is a homomorphism ϕ from G to $\mathrm{GL}(N, \mathcal{F})$, for some N. N is called the degree or dimension of ϕ.

Therefore, if ϕ is a function from G to $\mathrm{GL}(N, \mathcal{F})$, then ϕ is a representation if and only if

$$\phi(gh) = \phi(g)\, \phi(h) \quad \forall\, g, h \in G.$$

Since a representation is a homomorphism, it follows that for every representation $\phi \colon G \to \mathrm{GL}(N, \mathcal{F})$, we have

$$\begin{aligned} \phi(1) &= \mathbf{I}_N\,, \quad \text{and} \\ \phi(g^{-1}) &= \phi(g)^{-1} \quad \forall\, g \in G. \end{aligned}$$

Example 2.1 : Let C_4 denote the cyclic group of order 4, with the following members:

$$e = 1 \quad , \quad a = \mathrm{e}^{\frac{j\pi}{2}} \quad , \quad b = \mathrm{e}^{j\pi} \quad , \quad c = \mathrm{e}^{\frac{j3\pi}{2}}$$

where we have $a^{-1} = c$, $c^{-1} = a$ and $b^{-1} = b$.

Let H denote a subgroup of $\mathrm{GL}(2, \mathcal{R})$, with the following members:

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

where $\mathbf{AC} = \mathbf{I}_2$, $\mathbf{CA} = \mathbf{I}_2$ and $\mathbf{BB} = \mathbf{I}_2$.

Now, if ϕ is defined as a function from C_4 to H such that

$$\begin{aligned}
\phi\colon C_4 &\to H \subset \mathrm{GL}(2, \mathcal{R}) \\
\phi(1) &= \mathbf{I}_2\,, \\
\phi(a) &= \mathbf{A}\,, \\
\phi(b) &= \mathbf{B}\,, \\
\phi(c) &= \mathbf{C}\,,
\end{aligned}$$

it is easy to verify that ϕ is a representation with dimension 2.

$\Diamond$

Let $\phi : G \to \mathrm{GL}(N, \mathcal{F})$ be a representation, and let $\mathbf{T} \in \mathrm{GL}(N, \mathcal{F})$. For all $\mathbf{A}, \mathbf{B} \in \mathrm{GL}(N, \mathcal{F})$, we have

$$\big(\mathbf{T}^{-1}\mathbf{A}\mathbf{T}\big)\big(\mathbf{T}^{-1}\mathbf{B}\mathbf{T}\big) = \mathbf{T}^{-1}\big(\mathbf{A}\mathbf{B}\big)\mathbf{T}\,.$$

We can use this observation to produce a new representation ψ from ϕ, by simply defining

$$\psi(g) = \mathbf{T}^{-1}\phi(g)\,\mathbf{T} \quad \forall\, g \in G.$$

Then for all $g, h \in G$,

$$\begin{aligned}
\psi(gh) &= \mathbf{T}^{-1}\phi(gh)\,\mathbf{T} \\
&= \mathbf{T}^{-1}\big(\phi(g)\,\phi(h)\big)\mathbf{T} \\
&= \big(\mathbf{T}^{-1}\phi(g)\,\mathbf{T}\big)\big(\mathbf{T}^{-1}\phi(h)\,\mathbf{T}\big) \\
&= \psi(g)\,\psi(h)
\end{aligned}$$

and, thus, ψ is indeed a representation.

Definition 2.5 : If $\phi\colon G \to \mathrm{GL}(N, \mathcal{F})$ and $\psi\colon G \to \mathrm{GL}(M, \mathcal{F})$ are two representations of G over $\mathcal{F}$, then ϕ is said to be equivalent to ψ if $N = M$, and there exists an invertible matrix $\mathbf{T} \in \mathrm{GL}(N, \mathcal{F})$ such that

$$\psi(g) = \mathbf{T}^{-1}\phi(g)\,\mathbf{T} \quad \forall\, g \in G.$$

Definition 2.6 : If $\phi : G \to \mathrm{GL}(N, \mathcal{F})$ is a representation of G over $\mathcal{F}$, then ϕ is called an irreducible representation if there exists no equivalent representation $\psi : G \to H \subset \mathrm{GL}(N, \mathcal{F})$ for ϕ such that all members in H are block diagonal matrices.

Definition 2.7 : The trace of the representation matrices are called the characters of the representation. If $\phi\colon G \to \mathrm{GL}(N, \mathcal{F})$ is a representation of G over $\mathcal{F}$, then the character function $\chi_\phi\colon G \to \mathcal{F}$ is defined as

$$\chi_\phi(g) = \mathrm{tr}\{\phi(g)\} \quad \forall\, g \in G.$$

2.1.2 Representations of $\mathrm{GL}(N, \mathcal{C})$, $\mathcal{U}(N)$ and character properties

A homomorphism $\phi\colon \mathrm{GL}(N, \mathcal{C}) \to \mathrm{GL}(d, \mathcal{C})$ is called a d–dimensional representation of $\mathrm{GL}(N, \mathcal{C})$.

Lemma 2.1 : The irreducible representations of $\mathrm{GL}(N, \mathcal{C})$ can be labeled by the N–dimensional ordered sets, denoted by $\mathbf{r}_N = \{r_1, r_2, \ldots, r_N\}$, where $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_N \geqslant 0$ are integers [97]. ▼

Lemma 2.2 : The dimension $d_{\mathbf{r}_N}$ of the irreducible representation $\mathbf{r}_N$ of $\mathrm{GL}(N, \mathcal{C})$ is given by [99]

$$d_{\mathbf{r}_N} = \left[\prod_{i=1}^{N} \frac{(r_i + N - i)!}{(N-i)!}\right] \det_N\left[\frac{1}{(r_i - i + j)!}\right] \tag{2.1}$$

where the matrix elements inside the determinant with $r_i - i + j < 0$ are zero. ▼

Throughout this thesis, $\det\limits_M[\, f(i,j)\,]$ denotes the determinant of a $M \times M$ matrix with the (i,j)-th element given by $f(i,j)$.

Another equivalent expression for $d_{\mathbf{r}_N}$ is presented in Appendix C.1, which gives

$$d_{\mathbf{r}_N} = \frac{1}{\prod_{i=1}^{N}(N-i)!} \Delta_N(\mathbf{k}) \tag{2.2}$$

where $k_i \triangleq r_i + N - i$. This form of $d_{\mathbf{r}_N}$ is more convenient for certain unitary integrations.

Lemma 2.3 : The character of a group element $\mathbf{X} \in \mathrm{GL}(N, \mathcal{C})$ in its representation $\mathbf{r}_N$ is obtained by Weyl's character formula [100]:

$$\chi_{\mathbf{r}_N}(\mathbf{X}) = \mathrm{tr}\Big\{\mathbf{X}^{(\mathbf{r}_N)}\Big\} = \frac{\det\limits_N\left[\, x_i^{r_j + N - j}\,\right]}{\Delta\limits_N(\mathbf{x})} \tag{2.3}$$

where $\mathbf{X}^{(\mathbf{r}_N)} \in \mathrm{GL}(d_{\mathbf{r}_N}, \mathcal{C})$ denotes the representation matrix of $\mathbf{X}$, $\mathbf{x} = (x_1, \ldots, x_N)^T$ are the eigenvalues of matrix $\mathbf{X}$, and

$$\Delta_N(\mathbf{x}) = \Delta(x_1, x_2, ..., x_N) = \det_N\left[\, x_i^{N-j}\,\right] = \prod_{i<j}(x_i - x_j) \tag{2.4}$$

is the Vandermonde determinant of vector $\mathbf{x}$. ▼

Lemma 2.4 : Let $\mathbf{r}_N$ and $\mathbf{r}'_N$ be two representations of $\mathcal{U}(N)$. The orthogonality relation between the unitary group matrix elements implies that [97]

$$\int \mathrm{D}\mathbf{U}\, U_{ij}^{(\mathbf{r}_N)} U_{kl}^{(\mathbf{r}'_N)*} = \frac{1}{d_{\mathbf{r}_N}}\, \delta_{\mathbf{r}_N \mathbf{r}'_N}\, \delta_{ik}\, \delta_{jl} \tag{2.5}$$

where $U_{ij}^{(\mathbf{r}_N)}$ denotes the (i,j)-th element of the representation matrix of $\mathbf{U}$, and $d_{\mathbf{r}_N}$ is the dimension of the representation. Here, the integral is over all unitary matrices $\mathbf{U} \in \mathcal{U}(N)$, and $\mathrm{D}\mathbf{U}$ denotes the standard Haar measure of $\mathcal{U}(N)$ [97]. ▼

Proposition 2.1 : Assuming $\mathbf{A}, \mathbf{B} \in \mathrm{GL}(N, \mathcal{C})$, $\mathbf{U} \in \mathcal{U}(N)$, and $\mathbf{r}_N$ and $\mathbf{r}'_N$ are two representations of $\mathrm{GL}(N, \mathcal{C})$, then

$$\int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A})\, \chi_{\mathbf{r}'_N}(\mathbf{U}^*\mathbf{B}) = \frac{1}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{A}\mathbf{B})\, \delta_{\mathbf{r}_N \mathbf{r}'_N}\,.$$

▼

Proof: From (2.3), we have $\chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A}) = \mathrm{tr}\{(\mathbf{U}\mathbf{A})^{(\mathbf{r}_N)}\}$. Since a representation is a homomorphism, i.e., $(\mathbf{U}\mathbf{A})^{(\mathbf{r}_N)} = \mathbf{U}^{(\mathbf{r}_N)}\mathbf{A}^{(\mathbf{r}_N)}$, we obtain

$$\chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A}) = \mathrm{tr}\left\{\mathbf{U}^{(\mathbf{r}_N)}\mathbf{A}^{(\mathbf{r}_N)}\right\} = \sum_{k_2=1}^{N}\sum_{k_1=1}^{N} U_{k_2k_1}^{(\mathbf{r}_N)} A_{k_1k_2}^{(\mathbf{r}_N)}\,.$$

Therefore, we have

$$\begin{aligned}
\int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A})\chi_{\mathbf{r}'_N}(\mathbf{U}^*\mathbf{B}) &= \sum_{k_4=1}^{N}\sum_{k_3=1}^{N}\sum_{k_2=1}^{N}\sum_{k_1=1}^{N} A_{k_1k_2}^{(\mathbf{r}_N)} B_{k_3k_4}^{(\mathbf{r}'_N)} \int \mathrm{D}\mathbf{U}\, U_{k_2k_1}^{(\mathbf{r}_N)} U_{k_3k_4}^{(\mathbf{r}'_N)*} \\
&= \frac{1}{d_{\mathbf{r}_N}}\, \delta_{\mathbf{r}_N \mathbf{r}'_N} \sum_{k_4=1}^{N}\sum_{k_3=1}^{N} A_{k_4k_3}^{(\mathbf{r}_N)} B_{k_3k_4}^{(\mathbf{r}_N)} \\
&= \frac{1}{d_{\mathbf{r}_N}}\, \delta_{\mathbf{r}_N \mathbf{r}'_N}\, \mathrm{tr}\left\{\mathbf{A}^{(\mathbf{r}_N)}\mathbf{B}^{(\mathbf{r}_N)}\right\} \\
&= \frac{1}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{A}\mathbf{B})\, \delta_{\mathbf{r}_N \mathbf{r}'_N}
\end{aligned}$$

where the second equality comes from Lemma 2.4. ■

Proposition 2.2 : Assuming $\mathbf{A}, \mathbf{B} \in \mathrm{GL}(N, \mathcal{C})$, $\mathbf{U} \in \mathcal{U}(N)$, and $\mathbf{r}_N$ is a representation of $\mathrm{GL}(N, \mathcal{C})$, then

$$\int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A}\mathbf{U}^*\mathbf{B}) = \frac{1}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{A})\, \chi_{\mathbf{r}_N}(\mathbf{B})\,.$$

▼

Proof: From (2.3), we have $\chi_{\mathbf{r}_N}(\mathbf{UAU}^*\mathbf{B}) = \mathrm{tr}\{(\mathbf{UAU}^*\mathbf{B})^{(\mathbf{r}_N)}\}$. Since a representation is a homomorphism, i.e., $(\mathbf{UAU}^*\mathbf{B})^{(\mathbf{r}_N)} = \mathbf{U}^{(\mathbf{r}_N)}\mathbf{A}^{(\mathbf{r}_N)}\mathbf{U}^{*(\mathbf{r}_N)}\mathbf{B}^{(\mathbf{r}_N)}$, we obtain

$$\begin{aligned}\chi_{\mathbf{r}_N}(\mathbf{UAU}^*\mathbf{B}) &= \mathrm{tr}\Big\{\mathbf{U}^{(\mathbf{r}_N)}\mathbf{A}^{(\mathbf{r}_N)}\mathbf{U}^{*(\mathbf{r}_N)}\mathbf{B}^{(\mathbf{r}_N)}\Big\}\\ &= \sum_{k_4=1}^{N}\sum_{k_3=1}^{N}\sum_{k_2=1}^{N}\sum_{k_1=1}^{N} U^{(\mathbf{r}_N)}_{k_4k_3}A^{(\mathbf{r}_N)}_{k_3k_2}U^{*(\mathbf{r}_N)}_{k_1k_2}B^{(\mathbf{r}_N)}_{k_1k_4}\,.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\int \mathrm{D}\mathbf{U}\,\chi_{\mathbf{r}_N}(\mathbf{UAU}^*\mathbf{B}) &= \sum_{k_4=1}^{N}\sum_{k_3=1}^{N}\sum_{k_2=1}^{N}\sum_{k_1=1}^{N} A^{(\mathbf{r}_N)}_{k_3k_2}B^{(\mathbf{r}_N)}_{k_1k_4}\int \mathrm{D}\mathbf{U}\,U^{(\mathbf{r}_N)}_{k_4k_3}U^{*(\mathbf{r}_N)}_{k_1k_2}\\ &= \frac{1}{d_{\mathbf{r}_N}}\sum_{k_2=1}^{N}\sum_{k_1=1}^{N} A^{(\mathbf{r}_N)}_{k_2k_2}B^{(\mathbf{r}_N)}_{k_1k_1}\\ &= \frac{1}{d_{\mathbf{r}_N}}\,\mathrm{tr}\Big\{\mathbf{A}^{(\mathbf{r}_N)}\Big\}\,\mathrm{tr}\Big\{\mathbf{B}^{(\mathbf{r}_N)}\Big\}\\ &= \frac{1}{d_{\mathbf{r}_N}}\,\chi_{\mathbf{r}_N}(\mathbf{A})\chi_{\mathbf{r}_N}(\mathbf{B})\end{aligned}$$

where the second equality comes from Lemma 2.4.

■

Remark 2.1 : Propositions 2.1 and 2.2 are the fundamental tools for applying the character expansions for unitary integrations. In both proofs, we used the fact that

$$\begin{aligned}\chi_{\mathbf{r}_N}(\mathbf{ABC}) &= \mathrm{tr}\Big\{(\mathbf{ABC})^{(\mathbf{r}_N)}\Big\}\\ &= \mathrm{tr}\Big\{\mathbf{A}^{(\mathbf{r}_N)}\mathbf{B}^{(\mathbf{r}_N)}\mathbf{C}^{(\mathbf{r}_N)}\Big\}\end{aligned}\tag{2.6}$$

where the second equality occurs because a representation is a homomorphism. Apparently, all matrices should be square with dimension N and nonzero eigenvalues.

Now, if $\mathbf{B} \in \mathcal{C}^{N\times M}$ and $\mathbf{C} \in \mathcal{C}^{M\times N}$, $N > M$, we can take $\mathbf{BC}$ as a matrix with nonzero eigenvalues to obtain

$$\begin{aligned}\chi_{\mathbf{r}_N}(\mathbf{ABC}) &= \mathrm{tr}\Big\{(\mathbf{ABC})^{(\mathbf{r}_N)}\Big\}\\ &= \mathrm{tr}\Big\{\mathbf{A}^{(\mathbf{r}_N)}(\mathbf{BC})^{(\mathbf{r}_N)}\Big\}\,,\end{aligned}$$

and later, we can make $N-M$ eigenvalues of the matrix $\mathbf{BC}$ approach zero, without violating the definition of homomorphism. However, in [95], the authors assume that $N = M$, and that the eigenvalues of the matrix $\mathbf{BC}$ are nonzero. Consequently, the authors use the same equation as (2.6) in order to solve the integrals. Hence, the results are mathematically legitimate only for $N = M$, and for $N > M$, the definition of homomorphism is violated, even when $N - M$ eigenvalues of the matrix **BC** approach zero.

▲

2.2 Character Expansions

In [86], Balantekin proves the following theorem:

Theorem 2.1 : Consider the power series expansion of function $G(x)$ as

$$G(x) = \sum_{n=0}^{\infty} A_n \, x^n \, .$$

Given this series is convergent for N different values of x: $\{x_1, x_2, \ldots, x_N\}$, then we have

$$\prod_{i=1}^{N} G(x_i) = \sum_{\mathbf{n}_N} \det_N \left[\, A_{n_i - i + j} \,\right] \chi_{\mathbf{n}_N}(\mathbf{X})$$

where $\mathbf{X}$ is a member of $\mathrm{GL}(N, \mathcal{C})$ with eigenvalues $\{x_1, x_2, \ldots, x_N\}$, $\mathbf{n}_N$ is an irreducible representation of $\mathrm{GL}(N, \mathcal{C})$, and the summation is over all irreducible representations of $\mathrm{GL}(N, \mathcal{C})$.

▼

In the following, we present a few useful character expansions obtained by using Theorem 2.1:

1. If $\mathbf{X} \in \mathrm{GL}(N, \mathcal{C})$, then the following equation holds for $\mathbf{X}$:

$$\mathrm{etr}\{\mathbf{X}\} = \sum_{\mathbf{r}_N} \alpha_{\mathbf{r}_N} \chi_{\mathbf{r}_N}(\mathbf{X}) \tag{2.7}$$

where $\mathrm{etr}\{\mathbf{X}\} = \exp(\mathrm{tr}\{\mathbf{X}\})$, the summation is over all irreducible representations of $\mathrm{GL}(N, \mathcal{C})$, and the expansion coefficient $\alpha_{\mathbf{r}_N}$ is defined as

$$\alpha_{\mathbf{r}_N} = \det_N \left[\frac{1}{(r_i - i + j)!} \right] = \left[\prod_{i=1}^{N} \frac{(N-i)!}{(r_i + N - i)!} \right] d_{\mathbf{r}_N} \tag{2.8}$$

where $d_{\mathbf{r}_N}$ is the dimension of the representation $\mathbf{r}_N$. The matrix elements inside the determinant with $r_i - i + j < 0$ are zero.

To gain an insight into the character expansions, note that for $N = 1$, (2.7) is reduced to the Taylor series of e^x; i.e.,

$$\mathrm{etr}\{x\} = \sum_{r=0}^{\infty} \frac{1}{r!}\, x^r\, .$$

2. If $\mathbf{X} \in \mathrm{GL}(N, \mathcal{C})$ and $a \in \mathcal{C}$, then we have

$$\mathrm{etr}\{a\mathbf{X}\} = \sum_{\mathbf{r}_N} \beta_{\mathbf{r}_N} \chi_{\mathbf{r}_N}(\mathbf{X})$$

   where the expansion coefficient $\beta_{\mathbf{r}_N}$ is defined as $\beta_{\mathbf{r}_N} = a^{(r_1+r_2+\cdots+r_N)}\alpha_{\mathbf{r}_N}$.

3. Assume $\mathbf{X} \in \mathrm{GL}(N, \mathcal{C})$, $a \in \mathcal{C}$ and $\nu \in \mathcal{Z}$, then

$$\det_N[\,\mathbf{X}\,]^{\nu}\, \mathrm{etr}\{a\,\mathbf{X}\} = \sum_{\mathbf{r}_N} \beta_{\mathbf{r}_N}^{(\nu)} \chi_{\mathbf{r}_N}(\mathbf{X})$$

   where the expansion coefficient $\beta_{\mathbf{r}_N}^{(\nu)}$ is defined as

$$\beta_{\mathbf{r}_N}^{(\nu)} = a^{(r_1+r_2+\cdots+r_N)-N\nu} \det_N\left[\frac{1}{(r_i - i + j - \nu)!}\right].$$

Interested readers can refer to [86] for the derivation procedure and more examples of character expansions.

2.3 Unitary Integration by Character Expansions

To solve unitary integrals, Balantekin [86] applied the following integration steps:

1. Expansion of the integrand by using the character expansion method.
2. Integration over unitary matrices by using the available results on the unitary group, i.e., Propositions 2.1 and 2.2.
3. Re–summation of the expansion by using the Cauchy–Binet formula, i.e., Lemma B.1 in Appendix B.

He used character expansions to solve the well–known Harish-Chandra-Itzykson-Zuber integral [94] as follows.

Example 2.2 : The Harish-Chandra-Itzykson-Zuber integral is

$$\mathbb{I}_1^N = \int \mathrm{D}\mathbf{U}\,\mathrm{etr}\{\beta \mathbf{U}\mathbf{A}\mathbf{U}^*\mathbf{B}\} \tag{2.9}$$

where $\mathbf{U} \in \mathcal{U}(N)$, $\mathbf{A}, \mathbf{B} \in \mathrm{GL}(N, \mathcal{C})$, and $\beta \in \mathcal{C}$ is a scalar. Since the matrix $\mathbf{UAU}^*\mathbf{B}$ is a member of $\mathrm{GL}(N, \mathcal{C})$, we can freely apply the character expansions to (2.9). By absorbing β into $\mathbf{A}$ and using the expansion formula in (2.7), we obtain

$$
\begin{aligned}
\mathbb{J}_1^N &= \int \mathrm{D}\mathbf{U} \operatorname{etr}\{\mathbf{UAU}^*\mathbf{B}\} \\
&= \sum_{\mathbf{r}_N} \alpha_{\mathbf{r}_N} \int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{UAU}^*\mathbf{B}) \\
&= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}} \chi_{\mathbf{r}_N}(\mathbf{A})\, \chi_{\mathbf{r}_N}(\mathbf{B}) && (2.10) \\
&= \sum_{\mathbf{r}_N} \left[\prod_{i=1}^{N} \frac{(N-i)!}{(r_i + N - i)!} \right] \frac{\det_N \left[a_i^{r_j + N - j} \right]}{\Delta_N(\mathbf{a})} \times \frac{\det_N \left[b_i^{r_j + N - j} \right]}{\Delta_N(\mathbf{b})} && (2.11)
\end{aligned}
$$

where (2.10) is obtained by applying Proposition 2.2, and (2.11) is the result of substituting $\alpha_{\mathbf{r}_N}$ from (2.8) and the characters from (2.3), where the vectors $\mathbf{a} = (a_1, a_2, \ldots, a_N)^T$ and $\mathbf{b} = (b_1, b_2, \ldots, b_N)^T$ represent the eigenvalues of the matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. Thus,

$$
\mathbb{J}_1^N = \frac{\prod_{i=1}^{N} (N-i)!}{\Delta_N(\mathbf{a})\, \Delta_N(\mathbf{b})} \sum_{\mathbf{k}_N} \det_N \left[a_i^{k_j} \right] \det_N \left[b_i^{k_j} \right] \prod_{i=1}^{N} \frac{1}{k_i!} \qquad (2.12)
$$

where $k_j = r_j + N - j$.

By considering the power series expansion of e^z, and applying the Cauchy–Binet formula (Lemma B.1 in Appendix B), we have

$$
\mathbb{J}_1^N = \frac{\prod_{i=1}^{N} (N-i)!}{\Delta_N(\mathbf{a})\, \Delta_N(\mathbf{b})} \det_N \left[\exp(a_i b_j) \right] . \qquad (2.13)
$$

By replacing $\mathbf{A}$ in $\mathbb{J}_1^N$ by $\beta\mathbf{A}$ or, equivalently, replacing a_i by βa_i in (2.13) and some factoring of β, we conclude that

$$
\mathbb{I}_1^N = \frac{\beta^{-\frac{N(N-1)}{2}} \prod_{i=1}^{N} (N-i)!}{\Delta_N(\mathbf{a})\, \Delta_N(\mathbf{b})} \det_N \left[\exp(\beta a_i b_j) \right] . \qquad (2.14)
$$

$\Diamond$

In Balantekin's approach [86], the coefficient matrices appearing in the integrand are nonzero–determinant square matrices. However, in some applications, including the capacity analysis of MIMO wireless systems when correlations exist between antenna array gains in one side of the system, we have to solve the following generalized integral over $\mathcal{U}(N)$:

$$
\mathbb{I}_1^{N,M} = \int \mathrm{D}\mathbf{U} \operatorname{etr}\{\beta \mathbf{UABU}^*\mathbf{C}\} \qquad (2.15)
$$

where $\mathbf{U} \in \mathcal{U}(N)$, and $\mathbf{A} \in \mathcal{C}^{N\times M}$, $\mathbf{B} \in \mathcal{C}^{M\times N}$ $(N \geqslant M)$ and $\mathbf{C} \in \mathcal{C}^{N\times N}$ are general complex matrices. As a result, the matrix $\mathbf{AB}$ has $N - M$ zero eigenvalues [48]. Consequently, the resulting matrix in the integrand (i.e., $\mathbf{UABU}^*\mathbf{C}$) is not invertible and, therefore, is not a member of $\mathrm{GL}(N, \mathcal{C})$. Thus, we cannot directly apply character expansions to calculate (2.15).

To handle this problem, the authors in [95] propose the following integration steps:

1. Assume momentarily $N = M$ so that the matrix integrand is a group member.
2. Apply Balantekin's three–step method to solve the unitary integral.
3. Find the limit of the final result when $N - M$ eigenvalues approach zero.

Unfortunately, this integration approach is not mathematically legitimate, as we explained in Remark 2.1. To be more specific, whenever applying the above approach violates Remark 2.1, the result of the integral is wrong.

In this thesis, we propose the following universal integration framework:

1. Take the unitary integrals in order from the largest dimension to the smallest.
2. Assume that the matrix integrand has distinct nonzero eigenvalues so that the matrix integrand is a group member.
3. Apply the character expansion method.
4. Integrate over the corresponding unitary matrix by using the available results on the unitary group; i.e., Propositions 2.1 and 2.2.
5. Take the limit of the result when some of the eigenvalues should be zero.
6. Repeat steps (2), (4) and (5) in the case of multiple unitary integrals.
7. Re–sum the final result by using the generalized Cauchy–Binet formula (Appendix B) where applicable.

Unlike the approach in [95], we take the limits on the eigenvalues exactly after each integration. This approach guarantees that the unitary integrations are performed in the original dimensions of the unitary matrices and prevents the possible errors due to the misuse of Propositions 2.1 and 2.2, as we explained in Remark 2.1.

To obtain the limits according to the above framework, we need additional algebraic results, which will be presented in the next section. The Cauchy–Binet formula is generalized in Appendix B.

2.4 Essential Limits

Proposition 2.3 : Assume $\mathbf{A} \in \mathcal{C}^{N\times M}$ and $\mathbf{B} \in \mathcal{C}^{M\times N}$ are of the rank M ($M \leqslant N$), and x_i's, $i = 1, \ldots, N$, are the eigenvalues of the matrix $\mathbf{AB}$. Then

$$\lim_{\{x_{M+1},\ldots,x_N\}\to 0} \chi_{\mathbf{r}_N}(\mathbf{AB}) = \begin{cases} \chi_{\mathbf{r}_M}(\mathbf{BA}) & \text{, if} \quad r_{M+1} = \cdots = r_N = 0\,; \\ 0 & \text{, otherwise}\,; \end{cases}$$

where $\{x_{M+1}, \ldots, x_N\}$ represent the $N-M$ zero eigenvalues of the matrix $\mathbf{AB}$, and $\mathbf{r}_N$ and $\mathbf{r}_M$ are irreducible representations of $\mathrm{GL}(N, \mathcal{C})$ and $\mathrm{GL}(M, \mathcal{C})$, respectively.

▼

Proof: From the definition of the character of a matrix (2.3), and noting that $\det_N\left[\, x_i^{r_j+N-j} \,\right] = \det_N\left[\, x_j^{r_i+N-i} \,\right]$, we can define $f_i(x_j) = x_j^{r_i+N-i}$ and apply Lemma A.1 in Appendix A to (2.3) to obtain

$$\begin{aligned} \lim_{\{x_{M+1},\ldots,x_N\}\to x_0} \chi_{\mathbf{r}_N}(\mathbf{AB}) &= \lim_{\{x_{M+1},\ldots,x_N\}\to x_0} \frac{\det_N\left[\, x_j^{r_i+N-i} \,\right]}{\Delta(x_1, \ldots, x_N)} \\ &= \frac{\det_N[\,\mathbf{X}\,]}{\Delta(x_1, \ldots, x_M) \prod_{i=1}^{M} (x_i - x_0)^{N-M} \prod_{j=1}^{N-M-1} j!} \end{aligned}$$

where

$$\mathbf{X} = \left(x_j^{r_i+N-i} \Big|_{j=1}^{M} \,,\; \frac{(r_i + N - i)!}{(r_i - i + j)!}\, x_0^{r_i-i+j} \Big|_{j=M+1}^{N} \right)$$

and $i = 1, \ldots, N$ generates all rows of $\mathbf{X}$. Note that all entries of $\mathbf{X}$ with $r_i - i + j < 0$ are zero.

As observed, all diagonal entries ($i = j$) of $\mathbf{X}$ from $i, j = M+1, \ldots, N$ are in the form of $\frac{(r_i+N-i)!}{r_i!}\, x_0^{r_i}$. Therefore, if $x_0 \to 0$, then $\det_N[\,\mathbf{X}\,] = 0$ unless $r_i = 0$, $i = M+1, \ldots, N$. In this case,

$$\lim_{\{x_{M+1},\ldots,x_N\}\to 0}\; \lim_{\{r_{M+1},\ldots,r_N\}\to 0} \chi_{\mathbf{r}_N}(\mathbf{AB}) = \frac{\det_N \begin{bmatrix} \mathbf{Q}_{M\times M} & \mathbf{0}_{M\times(N-M)} \\ \mathbf{W}_{(N-M)\times M} & \mathbf{P}_{(N-M)\times(N-M)} \end{bmatrix}}{\Delta(x_1, \ldots, x_M) \prod_{i=1}^{M} x_i^{N-M} \prod_{j=1}^{N-M-1} j!}$$

where

$$\mathbf{Q} = \begin{pmatrix} x_1^{r_1+N-1} & \cdots & x_M^{r_1+N-1} \\ \vdots & \ddots & \vdots \\ x_1^{r_M+N-M} & \cdots & x_M^{r_M+N-M} \end{pmatrix}$$

and

$$\mathbf{P} = \begin{pmatrix} (N-M-1)! & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 2! & 0 & 0 \\ 0 & \cdots & 0 & 1! & 0 \\ 0 & \cdots & 0 & 0 & 0! \end{pmatrix}.$$

By column factoring of $\mathbf{Q}$, we obtain

$$\begin{aligned} \det_N \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{W} & \mathbf{P} \end{bmatrix} &= \det_M[\,\mathbf{Q}\,] \det_{N-M}[\,\mathbf{P}\,] \\ &= \left[\det_M \left[x_j^{r_i+M-i} \right] \prod_{i=1}^{M} x_i^{N-M} \right] \prod_{j=1}^{N-M-1} j! \end{aligned}$$

independent of $\mathbf{W}$. Therefore,

$$\begin{aligned} \lim_{\{x_{M+1},\ldots,x_N\}\to 0} \lim_{\{r_{M+1},\ldots,r_N\}\to 0} \chi_{\mathbf{r}_N}(\mathbf{AB}) &= \frac{\det_M \left[x_j^{r_i+M-i} \right]}{\Delta(x_1,\ldots,x_M)} \\ &= \chi_{\mathbf{r}_M}(\mathbf{BA}) \end{aligned}$$

where the last equality occurs because the nonzero eigenvalues of the matrices $\mathbf{AB}$ and $\mathbf{BA}$ are equal [48].

■

Proposition 2.4 : Assume $\mathbf{r}_N$ is an irreducible representation of $\mathrm{GL}(N,\mathcal{C})$ with dimension $d_{\mathbf{r}_N}$, and $\alpha_{\mathbf{r}_N}$ is the corresponding expansion coefficient defined in (2.8). Then

$$\begin{aligned} \lim_{\{r_{M+1},\ldots,r_N\}\to 0} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}} &= \lim_{\{r_{M+1},\ldots,r_N\}\to 0} \left[\prod_{i=1}^{N} \frac{(N-i)!}{(r_i+N-i)!} \right] \\ &= \frac{\prod_{i=1}^{N}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)! \prod_{i=M+1}^{N}(N-i)!} \\ &= \frac{\prod_{i=1}^{M}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)!}. \end{aligned}$$

■

Proposition 2.5 : Assume $\mathbf{r}_N$ and $\mathbf{r}_M$ are irreducible representations of $\mathrm{GL}(N,\mathcal{C})$ and $\mathrm{GL}(M,\mathcal{C})$ $(N \geqslant M)$, respectively, and $\alpha_{\mathbf{r}_N}$ and $\alpha_{\mathbf{r}_M}$ are the corresponding expansion coefficients defined in (2.8). Then

$$\lim_{\{r_{M+1},\ldots,r_N\}\to 0} \alpha_{\mathbf{r}_N} = \alpha_{\mathbf{r}_M} .$$

▼

Proof: From the definition of $\alpha_{\mathbf{r}_N}$ in (2.8), and noting that the matrix elements inside the determinant with $r_i - i + j < 0$ are zero, we have

$$\begin{aligned}\lim_{\{r_{M+1},\ldots,r_N\}\to 0} \alpha_{\mathbf{r}_N} &= \lim_{\{r_{M+1},\ldots,r_N\}\to 0} \det_N \left[\frac{1}{(r_i - i + j)!} \right] \\ &= \det_N \begin{bmatrix} \mathbf{Q}_{M\times M} & \mathbf{T}_{M\times(N-M)} \\ \mathbf{0}_{(N-M)\times M} & \mathbf{R}_{(N-M)\times(N-M)} \end{bmatrix} \\ &= \det_M [\,\mathbf{Q}\,] \det_{N-M} [\,\mathbf{R}\,]\end{aligned}$$

where $Q_{ij} = \Big[(r_i - i + j)! \Big]^{-1}$ for $i, j = 1, \ldots, M$, and

$$\mathbf{R} = \begin{pmatrix} \frac{1}{0!} & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{(N-M-2)!} & \frac{1}{(N-M-1)!} \\ 0 & \frac{1}{0!} & \frac{1}{1!} & \cdots & \frac{1}{(N-M-3)!} & \frac{1}{(N-M-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{0!} & \frac{1}{1!} \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{0!} \end{pmatrix}$$

so that $\det_{N-M} [\,\mathbf{R}\,] = 1$. Thus,

$$\begin{aligned}\lim_{\{r_{M+1},\ldots,r_N\}\to 0} \alpha_{\mathbf{r}_N} &= \det_M [\,\mathbf{Q}\,] \\ &= \alpha_{\mathbf{r}_M} .\end{aligned}$$

■

2.5 Summary

In this chapter, after a short review of basic definitions from group and representation theory, the characters of $\mathrm{GL}(N,\mathcal{C})$ and $\mathcal{U}(N)$ and their properties were presented. In particular, we derived Propositions 2.1 and 2.2 as the major tools for using character expansions for unitary integrations. The key observation in Remark 2.1 shows that Propositions 2.1 and 2.2 should be carefully applied to unitary integrals. Therefore, we proposed a new integration framework to avoid the possible errors due to the misuse of Propositions 2.1 and 2.2. In addition, we derived the essential algebraic tools for our proposed integration method in Section 2.4, as well as in Appendices A and B, to be able to solve unitary integrals with general rectangular complex matrices in the integrand. Some of such unitary integrals are investigated in the next chapter.

Chapter 3

Generalization of Some Unitary Integrals and Applications

In this chapter, our proposed integration framework in Section 2.3 is applied to solve three well–known but not solved in general form unitary integrals in their general forms, such as the generalization of the Harish–Chandra–Itzykson–Zuber integral [94]. Such generalized integrals have applications in physics such as quantum chromodynamics (QCD) [86], [87], flux line pinning in superconductors [88], scattering in quantum chaos [89], and color–flavor transformation [90]. In addition, the joint eigenvalue distributions of Wishart matrices require unitary integrals that are special cases of these generalized integrals.

We use the results of these integrals to derive new expressions for the joint eigenvalue distributions of the semi–correlated and full–correlated central Wishart matrices, as well as the i.i.d. and uncorrelated noncentral Wishart matrices, in a unified approach. The new expressions can be easily computed and also used for further analysis, such as in the applications presented in the next chapters. In addition, the joint eigenvalue distribution of the full–correlated central Wishart matrix is a new result in random matrix theory [57].

3.1 Generalized Unitary Integrals

The unitary integrals that we consider in this thesis are as follows:

1. The first integral is the generalization of the Harish–Chandra–Itzykson–Zuber integral [94]:

$$\mathbb{I}_1^{N,M} = \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\beta \mathbf{U}\mathbf{A}\mathbf{B}\mathbf{U}^*\mathbf{C}\} \tag{3.1}$$

where $\mathbf{U} \in \mathcal{U}(N)$, $\mathbf{C} \in \mathcal{C}^{N\times N}$, $\beta \in \mathcal{C}$ is a complex scalar, and $\mathbf{A} \in \mathcal{C}^{N\times M}$ and $\mathbf{B} \in \mathcal{C}^{M\times N}$ are general complex matrices with rank M ($M \leqslant N$). This integral has been solved before only for special cases, where $N = M$ and/or $\mathbf{A} = \mathbf{B}^*$, and $\mathbf{C}$ is a positive definite Hermitian matrix [85].

2. The second integral is

$$\mathbb{I}_2^{N,M} = \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\beta\,(\mathbf{U}\mathbf{A}\mathbf{V}^*\mathbf{B} + \mathbf{C}\mathbf{V}\mathbf{D}\mathbf{U}^*)\} \tag{3.2}$$

where $\mathbf{U} \in \mathcal{U}(N)$, $\mathbf{V} \in \mathcal{U}(M)$, $\mathbf{A}, \mathbf{C} \in \mathcal{C}^{N\times M}$ and $\mathbf{B}, \mathbf{D} \in \mathcal{C}^{M\times N}$ are general rectangular complex matrices, and β is a complex scalar. To our best knowledge, this integral was previously known only for the case of $\mathbf{D} = \mathbf{A}^*$ and $\mathbf{B} = \mathbf{C}^*$ [59], [85], [101].

3. The third integral is

$$\mathbb{I}_3^{N,M} = \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\beta \mathbf{U}\mathbf{A}\mathbf{V}^*\mathbf{B}\mathbf{V}\mathbf{C}\mathbf{U}^*\mathbf{D}\} \tag{3.3}$$

where $\mathbf{U} \in \mathcal{U}(N)$, $\mathbf{V} \in \mathcal{U}(M)$, $\mathbf{B} \in \mathcal{C}^{M\times M}$, $\mathbf{D} \in \mathcal{C}^{N\times N}$, $\mathbf{A} \in \mathcal{C}^{N\times M}$ and $\mathbf{C} \in \mathcal{C}^{M\times N}$ are general complex matrices, and β is a complex scalar. To our best knowledge, this integral has been solved in the literature only for the case that $N = M$, $\mathbf{A} = \mathbf{C}^*$, and $\mathbf{B}$ and $\mathbf{D}$ are positive definite Hermitian matrices [59], [95], [102].

3.1.1 Calculation of $\mathbb{I}_1^{N,M}$

To simplify the calculation of $\mathbb{I}_1^{N,M}$, we first absorb β into $\mathbf{C}$ and solve the following integral:

$$\mathbb{J}_1^{N,M} = \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\mathbf{U}\mathbf{A}\mathbf{B}\mathbf{U}^*\mathbf{C}\}\ . \tag{3.4}$$

When $M < N$, the matrix $\mathbf{AB}$ is not a full rank matrix; i.e., $\mathbf{AB} \notin \mathrm{GL}(N, \mathcal{C})$. Therefore, we cannot apply the character expansion method directly. To handle this problem, we assume momentarily that the matrix $\mathbf{AB}$ has a full rank of N. Thus, by using the expansion formula in (2.7), we obtain

$$\begin{aligned} \widetilde{\mathbb{J}}_1^{N,M} &= \sum_{\mathbf{r}_N} \alpha_{\mathbf{r}_N} \int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{U}\mathbf{A}\mathbf{B}\mathbf{U}^*\mathbf{C}) \\ &= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{A}\mathbf{B})\, \chi_{\mathbf{r}_N}(\mathbf{C}) \end{aligned} \tag{3.5}$$

where the second equality is obtained by using Proposition 2.2. The full–rank assumption allowed us to take the integration over $\mathcal{U}(N)$ by grouping the coefficient matrices according to Remark 2.1. Thus, to obtain $\mathbb{J}_1^{N,M}$, we should take the limit of $\widetilde{\mathbb{J}}_1^{N,M}$ when $N-M$ eigenvalues of the matrix $\mathbf{AB}$ approach zero. That is,

$$\mathbb{J}_1^{N,M} = \lim_{\{x_{M+1},\dots,x_N\}\to 0} \widetilde{\mathbb{J}}_1^{N,M} \tag{3.6}$$

where x_i, $i=1,\dots,N$, are the eigenvalues of matrix $\mathbf{AB}$.

By defining

$$\overline{\chi}_{\mathbf{r}_M}(\mathbf{C}) = \frac{\det\limits_{N}\left[y_i^{r_j+N-j}\Big|_{j=1}^{M},\, y_i^{N-j}\Big|_{j=M+1}^{N} \right]}{\Delta\limits_{N}(\mathbf{y})} \tag{3.7}$$

where the vector $\mathbf{y}=(y_1,\dots,y_N)^T$ represents the eigenvalues of matrix $\mathbf{C}$, and by applying Propositions 2.3 and 2.4 to (3.5), we obtain

$$\begin{aligned}
\mathbb{J}_1^{N,M} &= \lim_{\{x_{M+1},\dots,x_N\}\to 0} \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{AB})\, \chi_{\mathbf{r}_N}(\mathbf{C}) \\
&= \prod_{i=1}^{M}(N-i)! \sum_{\mathbf{r}_M} \frac{1}{\prod_{i=1}^{M}(r_i+N-i)!}\, \chi_{\mathbf{r}_M}(\mathbf{BA})\, \overline{\chi}_{\mathbf{r}_M}(\mathbf{C}) \\
&= \frac{\prod_{i=1}^{M}(N-i)!}{\Delta\limits_{M}(\mathbf{x})\, \Delta\limits_{N}(\mathbf{y})} \times \\
&\qquad \sum_{\mathbf{r}_M} \det_{M}\left[x_i^{r_j+M-j} \right] \det_{N}\left[y_i^{r_j+N-j}\Big|_{j=1}^{M},\, y_i^{N-j}\Big|_{j=M+1}^{N} \right] \prod_{i=1}^{M} \frac{1}{(r_i+N-i)!} \\
&= \frac{\prod_{i=1}^{M}(N-i)!}{\Delta\limits_{M}(\mathbf{x})\, \Delta\limits_{N}(\mathbf{y})} \times \\
&\qquad \sum_{\mathbf{k}_M} \det_{M}\left[x_i^{k_j} \right] \det_{N}\left[y_i^{k_j+N-M}\Big|_{j=1}^{M},\, y_i^{N-j}\Big|_{j=M+1}^{N} \right] \prod_{i=1}^{M} \frac{1}{(k_i+N-M)!}
\end{aligned} \tag{3.8}$$

where the vector $\mathbf{x}=(x_1,\dots,x_M)^T$ represents the eigenvalues of the matrix $\mathbf{BA}$, and $k_j = r_j + M - j$.

By considering the power series expansion of e^z, and applying the generalized Cauchy–Binet formula (Lemma B.2 in Appendix B), we have

$$\mathbb{J}_1^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!}{\Delta\limits_{M}(\mathbf{x})\, \Delta\limits_{N}(\mathbf{y}) \prod_{i=1}^{M} x_i^{N-M}} \det_{N} \begin{bmatrix} \exp(x_i y_j)\Big|_{M}^{i=1} \\ y_j^{N-i}\Big|_{N}^{i=M+1} \end{bmatrix}. \tag{3.9}$$

Finally, by replacing $\mathbf{C}$ in $\mathbb{J}_1^{N,M}$ by $\beta\mathbf{C}$ or, equivalently, replacing y_j by βy_j in (3.9) and some factoring of β, we conclude that

$$\mathbb{I}_1^{N,M} = \frac{\beta^{M(\frac{M+1}{2}-N)} \prod_{i=1}^{M}(N-i)!}{\underset{M}{\Delta}(\mathbf{x}) \underset{N}{\Delta}(\mathbf{y}) \prod_{i=1}^{M} x_i^{N-M}} \det_N \begin{bmatrix} \exp(\beta x_i y_j) \Big|_{M}^{i=1} \\ \\ y_j^{N-i} \quad \Big|_{N}^{i=M+1} \end{bmatrix}. \tag{3.10}$$

■

In the case in which $N \leqslant M$, the matrix $\mathbf{AB}$ has a full rank of N. Therefore, one can use (3.10) to calculate $\mathbb{I}_1^{N,M}$ by setting $N = M$. In other words, $\mathbb{I}_1^{N,M} = \mathbb{I}_1^{N,N}$ when $N \leqslant M$.

Example 3.1 : Although we solved $\mathbb{I}_1^{N,M}$ without assuming any specific structure for the matrix $\mathbf{C}$, it is of interest to find $\mathbb{I}_1^{N,M}$ when all eigenvalues of matrix $\mathbf{C}$ are identical, or, equivalently, when $y_j = y\,, j = 1, \ldots, N$. In this scenario, by applying Lemma A.1 in Appendix A, we can write

$$\lim_{\{y_1,\cdots,y_N\}\to y} \frac{\det_N \begin{bmatrix} \exp(\beta x_i y_j) \Big|_{M}^{i=1} \\ \\ y_j^{N-i} \quad \Big|_{N}^{i=M+1} \end{bmatrix}}{\underset{N}{\Delta}(\mathbf{y})} = \frac{\det_N \begin{bmatrix} \mathbf{Q}_{M\times M} & \mathbf{T}_{M\times(N-M)} \\ \mathbf{0}_{(N-M)\times M} & \mathbf{P}_{(N-M)\times(N-M)} \end{bmatrix}}{\prod_{j=1}^{N-1} j!} \tag{3.11}$$

where

$$\mathbf{Q} = \begin{pmatrix} x_1^{N-1}\exp(x_1 y) & \cdots & x_1^{N-M}\exp(x_1 y) \\ \vdots & \ddots & \vdots \\ x_M^{N-1}\exp(x_M y) & \cdots & x_M^{N-M}\exp(x_M y) \end{pmatrix}$$

and

$$\mathbf{P} = \begin{pmatrix} (N-M-1)! & \cdots & (N-M-1)(N-M-2)y^{N-M-3} & (N-M-1)y^{N-M-2} & y^{N-M-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 2! & 2\,y & y^2 \\ 0 & \cdots & 0 & 1! & y \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

By column factoring of $\mathbf{Q}$, we obtain

$$\begin{aligned}\det_N\begin{bmatrix}\mathbf{Q} & \mathbf{T}\\ \mathbf{0} & \mathbf{P}\end{bmatrix} &= \det_M[\,\mathbf{Q}\,]\det_{N-M}[\,\mathbf{P}\,]\\ &= \det_M\Big[\,x_i^{M-j}\,\Big]\prod_{i=1}^{M} x_i^{N-M}\exp(x_i y)\prod_{j=0}^{N-M-1} j!\\ &= \Delta_M(\mathbf{x})\prod_{i=1}^{M} x_i^{N-M}\exp(x_i y)\prod_{j=M+1}^{N}(N-j)! \end{aligned} \tag{3.12}$$

independent of $\mathbf{T}$. By substituting (3.12) into (3.11) and (3.9), we can write

$$\mathbb{J}_1^{N,M}(y) \triangleq \lim_{\{y_1,\dots,y_N\}\to y}\mathbb{J}_1^{N,M} = \exp\left(y\sum_{i=1}^{M}x_i\right).$$

Finally, by replacing y with βy, we obtain

$$\mathbb{I}_1^{N,M}(y) \triangleq \lim_{\{y_1,\dots,y_N\}\to y}\mathbb{I}_1^{N,M} = \exp\left(\beta y\sum_{i=1}^{M}x_i\right).$$

This specific result justifies the calculation of $\mathbb{I}_1^{N,M}$ presented in this section because if we take the eigenvalue decomposition of the positive definite matrix $\mathbf{C}$ as $\mathbf{C}=\mathbf{V}\mathbf{D}\mathbf{V}^*$ where $\mathbf{D}=\operatorname{diag}(y_1,...,y_N)$, for the case of equal eigenvalues for $\mathbf{C}$, directly from (3.1) we have

$$\begin{aligned}\mathbb{I}_1^{N,M}(y) &= \int \mathrm{D}\mathbf{U}\,\operatorname{etr}\{\beta y\mathbf{U}\mathbf{A}\mathbf{B}\mathbf{U}^*\mathbf{V}\mathbf{I}_N\mathbf{V}^*\}\\ &= \exp\big(\beta y\operatorname{tr}\{\mathbf{A}\mathbf{B}\}\big)\int \mathrm{D}\mathbf{U}\\ &= \exp\left(\beta y\sum_{i=1}^{M}x_i\right)\end{aligned}$$

where $\mathbf{I}_N$ is the identity matrix. $\Diamond$

3.1.2 Calculation of $\mathbb{I}_2^{N,M}$

To simplify the calculation of $\mathbb{I}_2^{N,M}$, we absorb β into $\mathbf{B}$ and $\mathbf{C}$ to solve the following integral:

$$\mathbb{J}_2^{N,M} = \int \mathrm{D}\mathbf{V}\int \mathrm{D}\mathbf{U}\,\operatorname{etr}\{\mathbf{U}\mathbf{A}\mathbf{V}^*\mathbf{B}+\mathbf{C}\mathbf{V}\mathbf{D}\mathbf{U}^*\}\,. \tag{3.13}$$

We assume that $N \geqslant M$ and that all rectangular matrices $\mathbf{A},\mathbf{B},\mathbf{C}$ and $\mathbf{D}$ have a full rank of M. Note that since $\operatorname{tr}\{\mathbf{U}\mathbf{A}\mathbf{V}^*\mathbf{B}\}=\operatorname{tr}\{\mathbf{B}\mathbf{U}\mathbf{A}\mathbf{V}^*\}$ and $\operatorname{tr}\{\mathbf{C}\mathbf{V}\mathbf{D}\mathbf{U}^*\}=\operatorname{tr}\{\mathbf{V}\mathbf{D}\mathbf{U}^*\mathbf{C}\}$ [48], we can write

$$\mathbb{J}_2^{M,N} = \int \mathrm{D}\mathbf{U}\int \mathrm{D}\mathbf{V}\operatorname{etr}\{\mathbf{V}\mathbf{D}\mathbf{U}^*\mathbf{C}+\mathbf{B}\mathbf{U}\mathbf{A}\mathbf{V}^*\}\,.$$

Thus, for the case of $N < M$, one may use the result of (3.13) by replacing $M, N, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ with $N, M, \mathbf{D}, \mathbf{C}, \mathbf{B}, \mathbf{A}$, respectively.

When $N \geqslant M$, both N–dimensional matrices $\mathbf{E} = \mathbf{AV}^*\mathbf{B}$ and $\mathbf{F} = \mathbf{CVD}$ in (3.13) are of the rank M and are not a member of $\mathrm{GL}(N, \mathcal{C})$. However, we assume momentarily that both matrices $\mathbf{E}$ and $\mathbf{F}$ have a full rank of N. Thus, we can use N–dimensional representations and character expansions to obtain

$$\begin{aligned}
\widetilde{\mathbb{J}}_2^{N,M} &= \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\mathbf{UE}\}\ \mathrm{etr}\{\mathbf{U}^*\mathbf{F}\} \\
&= \sum_{\mathbf{r}_N} \sum_{\mathbf{r}'_N} \alpha_{\mathbf{r}_N} \alpha_{\mathbf{r}'_N} \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{UE})\, \chi_{\mathbf{r}'_N}(\mathbf{U}^*\mathbf{F}) \\
&= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}^2}{d_{\mathbf{r}_N}} \int \mathrm{D}\mathbf{V}\, \chi_{\mathbf{r}_N}(\mathbf{EF}) \\
&= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}^2}{d_{\mathbf{r}_N}} \int \mathrm{D}\mathbf{V}\, \chi_{\mathbf{r}_N}(\mathbf{AV}^*\mathbf{BCVD}) \qquad (3.14)
\end{aligned}$$

where the second equality comes from (2.7), and the third equality is obtained by using Proposition 2.1. Note that the full–rank assumption allowed us to take the integration over $\mathcal{U}(N)$ by grouping the coefficient matrices according to Remark 2.1. However, since $\mathbf{A}$ is a $N \times M$ matrix and $\mathbf{D}$ is a $M \times N$ matrix, and the character function is the character of representation $\mathbf{r}_N$, the coefficient matrices in (3.14) cannot be properly grouped according to Remark 2.1 to apply Proposition 2.2. Thus, we cannot apply Proposition 2.2 to (3.14) unless we take the limit of $\widetilde{\mathbb{J}}_2^{N,M}$ when $N - M$ eigenvalues of the $N \times N$ matrix $\mathbf{AV}^*\mathbf{BCVD}$ approach zero; i.e.,

$$\begin{aligned}
\mathbb{J}_2^{N,M} &= \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \widetilde{\mathbb{J}}_2^{N,M} \\
&= \int \mathrm{D}\mathbf{V} \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \sum_{\mathbf{r}_N} \left[\frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}}\right] \alpha_{\mathbf{r}_N}\, \chi_{\mathbf{r}_N}(\mathbf{AV}^*\mathbf{BCVD}) \qquad (3.15)
\end{aligned}$$

where η_i's, $i = 1, 2, \ldots, N$, are the eigenvalues of the matrix $\mathbf{AV}^*\mathbf{BCVD}$.

By applying Propositions 2.3, 2.4 and 2.5 to (3.15), we obtain

$$\begin{aligned}
\mathbb{J}_2^{N,M} &= \sum_{\mathbf{r}_M} \frac{\prod_{i=1}^{M}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)!}\, \alpha_{\mathbf{r}_M} \int \mathrm{D}\mathbf{V} \chi_{\mathbf{r}_M}(\mathbf{VDAV}^*\mathbf{BC}) \\
&= \sum_{\mathbf{r}_M} \frac{\prod_{i=1}^{M}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)!} \left[\frac{\alpha_{\mathbf{r}_M}}{d_{\mathbf{r}_M}}\right] \chi_{\mathbf{r}_M}(\mathbf{DA})\, \chi_{\mathbf{r}_M}(\mathbf{BC})
\end{aligned}$$

where the second equality is obtained by using Proposition 2.2. By applying Weyl's character formula (2.3), and expansion coefficient (2.8), we have

$$\mathbb{J}_2^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{a})\,\underset{M}{\Delta}(\mathbf{b})} \sum_{\mathbf{k}_M} \det_M\left[\, a_i^{k_j} \,\right] \det_M\left[\, b_i^{k_j} \,\right] \prod_{i=1}^{M} \frac{1}{k_i!\,(k_i+N-M)!} \tag{3.16}$$

where $k_i \triangleq r_i + M - i$, and M–dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ are the eigenvalues of the matrices $\mathbf{DA}$ and $\mathbf{BC}$, respectively.

By considering the power series expansion of the modified Bessel function as

$$\frac{I_n(2\,z)}{z^n} = \sum_{k=0}^{\infty} \frac{z^{2k}}{k!\,(k+n)!}\,, \tag{3.17}$$

and by applying the Cauchy–Binet formula (Lemma B.1 in Appendix B), we obtain

$$\mathbb{J}_2^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{a})\,\underset{M}{\Delta}(\mathbf{b})} \det_M\left[\frac{I_{N-M}\left(2\sqrt{a_i b_j}\right)}{\left(\sqrt{a_i b_j}\right)^{N-M}} \right]. \tag{3.18}$$

Finally, by replacing $\mathbf{B}$ and $\mathbf{C}$ in $\mathbb{J}_2^{N,M}$ (3.13) by $\beta\mathbf{B}$ and $\beta\mathbf{C}$ or, equivalently, replacing b_j in (3.18) by $\beta^2 b_j$, and making some simplifications, we conclude that

$$\mathbb{I}_2^{N,M} = \frac{\beta^{M(1-N)}\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{a})\,\underset{M}{\Delta}(\mathbf{b})\prod_{i=1}^{M}(a_i b_i)^{\frac{N-M}{2}}} \det_M\left[\, I_{N-M}\left(2\beta\sqrt{a_i b_j}\right) \right]. \tag{3.19}$$

■

3.1.3 Calculation of $\mathbb{I}_3^{N,M}$

To simplify the calculation of $\mathbb{I}_3^{N,M}$, we absorb β into $\mathbf{B}$ to solve the following integral:

$$\mathbb{J}_3^{N,M} = \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \mathrm{etr}\{\mathbf{UAV}^*\mathbf{BVCU}^*\mathbf{D}\}\,. \tag{3.20}$$

Without loss of generality, we assume that $N \geqslant M$, the matrices $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ have a full rank of M, and the matrix $\mathbf{D}$ has a full rank of N. Since $\mathrm{tr}\{\mathbf{UAV}^*\mathbf{BVCU}^*\mathbf{D}\} = \mathrm{tr}\{\mathbf{VCU}^*\mathbf{DUAV}^*\mathbf{B}\}$, we can write

$$\mathbb{J}_3^{M,N} = \int \mathrm{D}\mathbf{U} \int \mathrm{D}\mathbf{V} \mathrm{etr}\{\mathbf{VCU}^*\mathbf{DUAV}^*\mathbf{B}\}\,.$$

Hence, for the case in which $N < M$, one may use the result of (3.20) by replacing $M, N, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ with $N, M, \mathbf{C}, \mathbf{D}, \mathbf{A}, \mathbf{B}$, respectively.

When $N \geqslant M$, the N–dimensional matrix $\mathbf{E} = \mathbf{AV}^*\mathbf{BVC}$ in (3.20) is not a full rank matrix; i.e., $\mathbf{E} \notin \mathrm{GL}(N, \mathcal{C})$. However, we assume momentarily that the matrix $\mathbf{E}$ has a full rank of N. Thus, we can use N–dimensional representations and character expansion to obtain

$$
\begin{aligned}
\widetilde{\mathbb{J}}_3^{N,M} &= \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U} \,\mathrm{etr}\{\mathbf{UEU}^*\mathbf{D}\} \\
&= \sum_{\mathbf{r}_N} \alpha_{\mathbf{r}_N} \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, \chi_{\mathbf{r}_N}(\mathbf{UEU}^*\mathbf{D}) \\
&= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}} \int \mathrm{D}\mathbf{V}\, \chi_{\mathbf{r}_N}(\mathbf{E})\, \chi_{\mathbf{r}_N}(\mathbf{D}) \\
&= \sum_{\mathbf{r}_N} \frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}}\, \chi_{\mathbf{r}_N}(\mathbf{D}) \int \mathrm{D}\mathbf{V}\, \chi_{\mathbf{r}_N}(\mathbf{AV}^*\mathbf{BVC})
\end{aligned}
\tag{3.21}
$$

where the second equality comes from (2.7), and the third equality is obtained by using Proposition 2.2. Note that the full–rank assumption allowed us to take the integration over $\mathcal{U}(N)$ by grouping the coefficient matrices according to Remark 2.1. However, since $\mathbf{A}$ is a $N \times M$ matrix and $\mathbf{C}$ is a $M \times N$ matrix, and the character function is the character of representation $\mathbf{r}_N$, the coefficient matrices in (3.21) cannot be properly grouped according to Remark 2.1 to apply Proposition 2.2. Thus, we cannot apply Proposition 2.2 to (3.21) unless we take the limit of $\widetilde{\mathbb{J}}_3^{N,M}$ when $N - M$ eigenvalues of the $N \times N$ matrix $\mathbf{AV}^*\mathbf{BVC}$ approach zero; i.e.,

$$
\begin{aligned}
\mathbb{J}_3^{N,M} &= \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \widetilde{\mathbb{J}}_3^{N,M} \\
&= \int \mathrm{D}\mathbf{V} \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \sum_{\mathbf{r}_N} \left[\frac{\alpha_{\mathbf{r}_N}}{d_{\mathbf{r}_N}}\right] \chi_{\mathbf{r}_N}(\mathbf{D})\, \chi_{\mathbf{r}_N}(\mathbf{AV}^*\mathbf{BVC})
\end{aligned}
\tag{3.22}
$$

where η_i's, $i = 1, 2, \ldots, N$, are the eigenvalues of the matrix $\mathbf{AV}^*\mathbf{BVC}$.

By applying Propositions 2.3 and 2.4 to (3.22), we obtain

$$
\begin{aligned}
\mathbb{J}_3^{N,M} &= \sum_{\mathbf{r}_M} \frac{\prod_{i=1}^{M}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)!}\, \overline{\chi}_{\mathbf{r}_M}(\mathbf{D}) \int \mathrm{D}\mathbf{V} \chi_{\mathbf{r}_M}(\mathbf{VCAV}^*\mathbf{B}) \\
&= \sum_{\mathbf{r}_M} \frac{\prod_{i=1}^{M}(N-i)!}{\prod_{i=1}^{M}(r_i+N-i)!} \left[\frac{1}{d_{\mathbf{r}_M}}\right] \overline{\chi}_{\mathbf{r}_M}(\mathbf{D})\, \chi_{\mathbf{r}_M}(\mathbf{CA})\, \chi_{\mathbf{r}_M}(\mathbf{B})
\end{aligned}
\tag{3.23}
$$

where $\overline{\chi}_{\mathbf{r}_M}(\mathbf{D})$ is defined in (3.7), and the second equality is obtained by employing Proposition 2.2. Substituting the characters from (2.3), $\overline{\chi}_{\mathbf{r}_M}(\mathbf{D})$ from (3.7), and $d_{\mathbf{r}_M}$

from (2.2) into (3.23), we obtain

$$\mathbb{J}_3^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{a})\,\underset{M}{\Delta}(\mathbf{b})\,\underset{N}{\Delta}(\mathbf{c})} \times \sum_{\mathbf{k}_M} \frac{\underset{M}{\det}\left[\, a_i^{k_j} \,\right] \underset{M}{\det}\left[\, b_i^{k_j} \,\right] \underset{N}{\det}\left[\, c_i^{k_j+N-M}\Big|_{j=1}^{M},\; c_i^{N-j}\Big|_{j=M+1}^{N} \,\right]}{\underset{M}{\Delta}(\mathbf{k}) \prod_{i=1}^{M}(k_i+N-M)!} \tag{3.24}$$

where $k_i = r_i + M - i$, the M–dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ represent the eigenvalues of the matrices $\mathbf{CA}$ and $\mathbf{B}$, respectively, and the N–dimensional vector $\mathbf{c}$ represents the eigenvalues of the matrix $\mathbf{D}$. Although we were able to generalize the Cauchy–Binet formula (Lemma B.2 in Appendix B) to the case of three or more determinants in the summation, since $\underset{M}{\Delta}(\mathbf{k})$ in (3.24) cannot be represented as a multiplication of any function of k_i's, it seems (3.24) cannot be further simplified.

By replacing $\mathbf{B}$ in $\mathbb{J}_3^{N,M}$ (3.20) by $\beta\mathbf{B}$ or, equivalently, replacing b_i in (3.24) by βb_i and making some simplifications, we conclude that

$$\mathbb{I}_3^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{a})\,\underset{M}{\Delta}(\mathbf{b})\,\underset{N}{\Delta}(\mathbf{c})\,\beta^{\frac{M(M-1)}{2}}} \times \sum_{\mathbf{k}_M} \frac{\underset{M}{\det}\left[\, a_i^{k_j} \,\right] \underset{M}{\det}\left[\, b_i^{k_j} \,\right] \underset{N}{\det}\left[\, c_i^{k_j+N-M}\Big|_{j=1}^{M},\; c_i^{N-j}\Big|_{j=M+1}^{N} \,\right]}{\underset{M}{\Delta}(\mathbf{k}) \prod_{i=1}^{M}\beta^{-k_i}(k_i+N-M)!} . \tag{3.25}$$

■

Remark 3.1 : In this section, we assumed the matrix integrand is a full–rank matrix to apply the character expansions, and took the limit of the result after the first integration step. This approach can be interpreted as

$$\int \mathrm{D}\mathbf{U} \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \operatorname{etr}\{\mathbf{X}\} = \lim_{\{\eta_{M+1},\ldots,\eta_N\}\to 0} \int \mathrm{D}\mathbf{U}\, \operatorname{etr}\{\mathbf{X}\} \tag{3.26}$$

where η_i's, $i = 1, 2, \ldots, N$, are the eigenvalues of matrix $\mathbf{X}$. The function $\operatorname{etr}\{\mathbf{X}\}$ is a bounded function; i.e., $\big|\operatorname{etr}\{\mathbf{X}\}\big| < \infty$, even when some of the eigenvalues of matrix $\mathbf{X}$ approach zero. Therefore, equation (3.26) holds for our analysis according to the Dominated Convergence Theorem [103].

▲

3.2 Joint Eigenvalue Distributions of Wishart Matrices

As we explained in Chapter 1, matrices of the form $\mathbf{H}\mathbf{H}^*$ where $\mathbf{H}$ is a Gaussian random matrix are known as Wishart matrices. Although such matrices were originally used by Wishart [73] for multivariate data analysis, they also have important applications in physics, mathematics, and information theory and communications for capacity and performance analysis of MIMO wireless systems.

In this section, we use the unitary integral results of the previous section to derive the joint eigenvalue distributions of Wishart matrices for common statistical assumptions.

Definition 3.1 : We define the $N \times M$, $N \geqslant M$, complex matrix $\mathbf{G}$ as the standard Gaussian random matrix when the elements of $\mathbf{G}$ are i.i.d. complex Gaussian random variables with zero mean and unit variance; i.e., $\mathcal{CN}(0,1)$.

3.2.1 Synopsis

Without loss of generality, we assume $\mathbf{H}$ is a $N \times M$, $N \geqslant M$, complex random matrix with Gaussian distribution. By assuming the singular value decomposition of $\mathbf{H}$ as $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ where $\mathbf{U} \in \mathcal{U}(N)$, $\mathbf{V} \in \mathcal{U}(M)$ and $\mathbf{\Sigma} = \mathrm{diag}(\{\sqrt{\lambda_i}\}) \in \mathcal{R}_+^{N\times M}$, it has been shown that [92]

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N,M}}{M!} \underset{M}{\Delta}(\boldsymbol{\lambda})^2 \prod_{i=1}^{M} \lambda_i^{N-M} \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U}\, p(\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*) \tag{3.27}$$

where the integrals are over all unitary matrices $\mathbf{U}$ and $\mathbf{V}$, $\mathrm{D}\mathbf{U}$ denotes the standard Haar measure of $\mathcal{U}(N)$ [97], $P(\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_M\})$ is the joint pdf of the nonzero eigenvalues of matrix $\mathbf{H}\mathbf{H}^*$, $p(\mathbf{H})$ is the joint pdf of the elements of $\mathbf{H}$, and

$$\mathcal{K}_{N,M}^{-1} = \prod_{j=1}^{M} (N-j)!(M-j)! \tag{3.28}$$

is a constant.

According to the joint distribution of the elements of $\mathbf{H}$, we substitute the corresponding $p(\mathbf{H})$ into (3.27) and evaluate the unitary integrals to obtain $P(\boldsymbol{\lambda})$.

3.2.2 $P(\boldsymbol{\lambda})$ for i.i.d. Central Wishart Matrix

We start with the most common case, where $\mathbf{H}$ is the standard complex Gaussian matrix. In this case, the joint pdf of the elements in $\mathbf{H}$ is given by

$$p(\mathbf{H}) = \mathrm{etr}\{-\mathbf{H}\mathbf{H}^*\}\,. \tag{3.29}$$

The pdf normalization factor in (3.29) has been absorbed into D$\mathbf{V}$D$\mathbf{U}$ in (3.27). Using the singular value decomposition of $\mathbf{H}$ as $\mathbf{H} = \mathbf{U\Sigma V}^*$, we have

$$p(\mathbf{H}) = \mathrm{etr}\{-\mathbf{\Sigma}^*\mathbf{\Sigma}\} = \exp\left(-\sum_{i=1}^{M}\lambda_i\right). \tag{3.30}$$

Since (3.30) is independent of both unitary matrices $\mathbf{U}$ and $\mathbf{V}$, from (3.27) we obtain

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N,M}}{M!}\underset{M}{\Delta}(\boldsymbol{\lambda})^2\prod_{i=1}^{M}e^{-\lambda_i}\lambda_i^{N-M} \tag{3.31}$$

without any need for complicated integrations over $\mathbf{U}$ and $\mathbf{V}$ [82].

3.2.3 $P(\boldsymbol{\lambda})$ for Semi–Correlated Central Wishart Matrix

When only the columns or the rows of matrix $\mathbf{H}$ are correlated, the corresponding Wishart matrix is called the semi–correlated Wishart matrix. Without loss of generality, we assume $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}$ where $\mathbf{R} \in \mathrm{GL}(N, \mathcal{C})$ is the correlation matrix, and $\mathbf{G}$ is the standard complex Gaussian matrix. In this case,

$$\begin{aligned} p(\mathbf{H}) &= \mathcal{N}_{\mathbf{R}}\mathrm{etr}\{-\mathbf{H}\mathbf{H}^*\mathbf{R}^{-1}\} \\ &= \mathcal{N}_{\mathbf{R}}\mathrm{etr}\{-\mathbf{U\Sigma\Sigma}^*\mathbf{U}^*\mathbf{R}^{-1}\} \end{aligned} \tag{3.32}$$

where $\mathcal{N}_{\mathbf{R}}^{-1} = \underset{N}{\det}[\,\mathbf{R}\,]^M$, and the second equality comes from $\mathbf{H} = \mathbf{U\Sigma V}^*$. Thus, from (3.27), we have

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N,M}}{M!}\mathcal{N}_{\mathbf{R}}\underset{M}{\Delta}(\boldsymbol{\lambda})^2\prod_{i=1}^{M}\lambda_i^{N-M}\int \mathrm{D}\mathbf{U}\,\mathrm{etr}\{-\mathbf{U\Sigma\Sigma}^*\mathbf{U}^*\mathbf{R}^{-1}\}. \tag{3.33}$$

If $N \leqslant M$, then $\mathbf{\Sigma\Sigma}^*$ and $\mathbf{R}^{-1}$ are both $N \times N$ matrices with full rank. However, when $N > M$, then the diagonal matrix $\mathbf{\Sigma\Sigma}^* = \mathrm{diag}\{\lambda_1, \ldots, \lambda_M, 0, \ldots, 0\}$ has $N - M$ zero eigenvalues. Therefore, the former case is a special case of the later case when $N = M$.

By employing (3.10), the joint pdf of the eigenvalues for the semi–correlated central Wishart is calculated as follows:

$$P(\boldsymbol{\lambda}) = \frac{(-1)^{M(\frac{M+1}{2}-N)}\underset{M}{\Delta}(\boldsymbol{\lambda})\prod_{j=1}^{N}y_j^M}{\underset{N}{\Delta}(\mathbf{y})\prod_{i=1}^{M}i!}\underset{N}{\det}\begin{bmatrix} \mathrm{e}^{-y_j\lambda_i}\Big|_{M}^{i=1} \\ y_j^{N-i}\Big|_{N}^{i=M+1} \end{bmatrix} \tag{3.34}$$

where the vector $\mathbf{y} = (y_1, \ldots, y_N)^T$ represents the eigenvalues of matrix $\mathbf{R}^{-1}$.

Accordingly, if we take $\mathbf{H} = \mathbf{G}\mathbf{T}^{\frac{1}{2}}$ where $\mathbf{T} \in \mathrm{GL}(M, \mathcal{C})$ is the correlation matrix, we can still use (3.34) to obtain the joint eigenvalue distribution by setting $M = N$.

The result in (3.34) is identical to the result in [95] because we had only one unitary integration, and the limit on the zero eigenvalues could be performed before or after the re–summation of the expansion. However, in the following sections where we have double unitary integrals, the results are not similar.

3.2.4 $P(\boldsymbol{\lambda})$ for Uncorrelated Noncentral Wishart Matrix

In this scenario, the matrix $\mathbf{H}$ can be modeled as $\mathbf{H} = \mathbf{G} + \mathbf{G}_0$ where $\mathbf{G}_0 \in \mathcal{C}^{N\times M}$ denotes the complex mean matrix, and $\mathbf{G}$ is the standard complex Gaussian matrix. In this case,

$$\begin{aligned} p(\mathbf{H}) &= \mathrm{etr}\{-(\mathbf{H}-\mathbf{G}_0)(\mathbf{H}-\mathbf{G}_0)^*\} \\ &= \exp\left(-\sum_{i=1}^{M}(\lambda_i+\gamma_i)\right)\mathrm{etr}\{\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{G}_0^* + \mathbf{G}_0\mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^*\} \end{aligned} \tag{3.35}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)^T$ are the M nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$, and the second equality comes from $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$. Thus, from (3.27) we have

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N,M}}{M!}\underset{M}{\Delta}(\boldsymbol{\lambda})^2\prod_{i=1}^{M}e^{-(\lambda_i+\gamma_i)}\lambda_i^{N-M}\int \mathrm{D}\mathbf{V}\int \mathrm{D}\mathbf{U}\,\mathrm{etr}\{\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{G}_0^* + \mathbf{G}_0\mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^*\}\,. \tag{3.36}$$

By employing (3.19), the joint pdf of the eigenvalues for the uncorrelated noncentral Wishart matrix is obtained as follows:

$$P(\boldsymbol{\lambda}) = \frac{1}{M!}\frac{\underset{M}{\Delta}(\boldsymbol{\lambda})}{\underset{M}{\Delta}(\boldsymbol{\gamma})}\prod_{i=1}^{M}\left[e^{-(\lambda_i+\gamma_i)}\left(\frac{\lambda_i}{\gamma_i}\right)^{\frac{N-M}{2}}\right]\det_M\left[I_{N-M}\left(2\sqrt{\lambda_i\gamma_j}\right)\right] \tag{3.37}$$

where the vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_M)^T$ represents the nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$.

Interested readers can refer to James' paper [59] to realize the simplicity and neatness of the character expansion method for unitary integrations compared to James' approach. The result in [59] has to be considerably manipulated to become handy like our result in (3.37). Such manipulations can be found in [104].

Although Simon et al. [95] try to obtain (3.37) by character expansions, due to the violation of Remark 2.1, they fail to obtain the correct joint pdf of the eigenvalues. The result in [95, Eq. (52)] is in the form of $I_0\left(2\sqrt{\lambda_i\gamma_j}\right)$ rather than

$I_{N-M}\left(2\sqrt{\lambda_i\gamma_j}\right)$ and, therefore, is incorrect. Interested readers can calculate [95, Eq. (52)] for a MIMO system with $N_t = 1$ and $N_r = 2$. This calculation results in a non–pdf function.

In the case where some of M nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$ are equal, one can use Lemma A.1 or Lemma A.2 in Appendix A to obtain the appropriate joint pdf.

3.2.5 $P(\boldsymbol{\lambda})$ for i.i.d. Noncentral Wishart Matrix

The i.i.d. noncentral Wishart matrix is a special case of the uncorrelated noncentral Wishart matrix when the mean matrix $\mathbf{G}_0$ is

$$\mathbf{G}_0 = g\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}_{N\times M}$$

where g is a complex scalar. In other words, all elements of the matrix $\mathbf{H}$ are i.i.d. $\mathcal{CN}(g,1)$ random variables. Therefore, $\gamma = MN\,|g|^2$ is the *only* nonzero eigenvalue of the matrix $\mathbf{G}_0\mathbf{G}_0^*$.

By denoting the eigenvalues of the matrices $\mathbf{DA}$ and $\mathbf{BC}$ in (3.18) by $\mathbf{a} = \mathbf{x}^2 = (x_1^2,\ldots,x_M^2)^T$ and $\mathbf{b} = \mathbf{y}^2 = (y_1^2,\ldots,y_M^2)^T$, respectively, we have

$$\mathbb{J}_2^{N,M} = \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{\underset{M}{\Delta}(\mathbf{y}^2)\prod_{i=1}^{M} y_i^{N-M}} \times \frac{\underset{M}{\det}\left[\, x_j^{-(N-M)} I_{N-M}\left(2y_i\, x_j\right)\right]}{\underset{M}{\Delta}(\mathbf{x}^2)}\,. \tag{3.38}$$

If we assume $(M-1)$ zero eigenvalues for the matrix $\mathbf{DA}$ in (3.38) or, equivalently, $x_1 = x$ and $x_i = 0\,,\ i = 2,\ldots,M$, then by applying Lemma A.2 in Appendix A to (3.38), we obtain

$$\begin{aligned}\mathbb{J}_2^{N,M}(x) &\triangleq \lim_{\{x_2,\ldots,x_M\}\to 0} \mathbb{J}_2^{N,M} \\ &= \frac{\prod_{i=1}^{M}(N-i)!(M-i)!}{x^{N+M-2}\underset{M}{\Delta}(\mathbf{y}^2)\prod_{i=1}^{M} y_i^{N-M}\prod_{i=0}^{M-2} i!(N-M+i)!} \\ &\qquad \times \underset{M}{\det}\left[\, I_{N-M}(2\,y_i x)\,,\ y_i^{N+M-2j}\Big|_{j=2}^{M}\right]. \end{aligned} \tag{3.39}$$

■

By employing (3.39), the joint pdf of eigenvalues for the i.i.d. noncentral Wishart matrix is obtained from (3.36) as follows:

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N-1,M-1}}{M!} \left[\frac{\mathrm{e}^{-\gamma}}{\gamma^{\frac{N+M}{2}-1}} \right] \underset{M}{\Delta}(\boldsymbol{\lambda}) \prod_{i=1}^{M} \left[\lambda_i^{\frac{N-M}{2}} \mathrm{e}^{-\lambda_i} \right]$$
$$\times \underset{M}{\det} \left[I_{N-M}\left(2\sqrt{\gamma\lambda_i}\right), \lambda_i^{\frac{N+M}{2}-j} \Big|_{j=2}^{M} \right]. \quad (3.40)$$

3.2.6 $P(\boldsymbol{\lambda})$ for Full–Correlated Central Wishart Matrix

When both the columns and the rows of the matrix $\mathbf{H}$ are correlated, the corresponding Wishart matrix is called the full–correlated Wishart matrix. Hence, $\mathbf{H}$ can be formulated as $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}\mathbf{T}^{\frac{1}{2}}$ where $\mathbf{R} \in \mathrm{GL}(N, \mathcal{C})$ and $\mathbf{T} \in \mathrm{GL}(M, \mathcal{C})$ are the correlation matrices, and $\mathbf{G}$ is the standard complex Gaussian matrix. In this case,

$$\begin{aligned} p(\mathbf{H}) &= \mathcal{N}_{\mathbf{R},\mathbf{T}} \,\mathrm{etr}\{-\mathbf{H}\mathbf{T}^{-1}\mathbf{H}^*\mathbf{R}^{-1}\} \\ &= \mathcal{N}_{\mathbf{R},\mathbf{T}} \,\mathrm{etr}\{-\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{T}^{-1}\mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^*\mathbf{R}^{-1}\} \end{aligned} \quad (3.41)$$

where $\mathcal{N}_{\mathbf{R},\mathbf{T}}^{-1} = \underset{N}{\det}[\,\mathbf{R}\,]^M \underset{M}{\det}[\,\mathbf{T}\,]^N$, and the second equality comes from $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$. Thus, from (3.27) we have

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{K}_{N,M}}{M!} \mathcal{N}_{\mathbf{R},\mathbf{T}} \underset{M}{\Delta}(\boldsymbol{\lambda})^2 \prod_{i=1}^{M} \lambda_i^{N-M} \int \mathrm{D}\mathbf{V} \int \mathrm{D}\mathbf{U} \,\mathrm{etr}\{-\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*\mathbf{T}^{-1}\mathbf{V}\boldsymbol{\Sigma}^*\mathbf{U}^*\mathbf{R}^{-1}\}. \quad (3.42)$$

By employing (3.25), the joint pdf of the eigenvalues for the full–correlated central Wishart matrix is calculated as follows:

$$P(\boldsymbol{\lambda}) = \frac{(-1)^{\frac{M(M-1)}{2}} \prod_{i=1}^{M} x_i^N \prod_{j=1}^{N} y_j^M}{M! \,\underset{M}{\Delta}(\mathbf{x}) \,\underset{N}{\Delta}(\mathbf{y})}$$
$$\times \sum_{\mathbf{k}_M} \frac{\underset{M}{\Delta}(\boldsymbol{\lambda}) \underset{M}{\det}\left[\lambda_i^{k_j+N-M}\right] \underset{M}{\det}\left[x_i^{k_j}\right] \underset{N}{\det}\left[y_i^{k_j+N-M}\Big|_{j=1}^{M}, y_i^{N-j}\Big|_{j=M+1}^{N} \right]}{\underset{M}{\Delta}(\mathbf{k}) \prod_{i=1}^{M} (-1)^{k_i} (k_i+N-M)!} \quad (3.43)$$

where the M–dimensional vector $\mathbf{x}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension M, and the N–dimensional vector $\mathbf{y}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension N.

The joint pdf of the eigenvalues for the full–correlated central Wishart matrix is obtained in [102] and [95], *only for* $M = N$. In addition, (3.43) cannot be obtained from the results in [102] or [95] by taking the limit on the zero eigenvalues. The reason is explained in Remark 2.1.

In the cases where some of the eigenvalues of the matrices $\mathbf{R}^{-1}$ and/or $\mathbf{T}^{-1}$ are equal, one can use Lemma A.1 in Appendix A to obtain the appropriate joint pdf.

Example 3.2 : One such example occurs when all eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$ are equal to one. This case is equivalent to the semi–correlated central Wishart case.

If we define

$$\mathcal{R}(\mathbf{k}_M) \triangleq \lim_{\{x_1,\ldots,x_M\}\to 1} \frac{\det_M\left[\, x_i^{k_j} \,\right]}{\Delta_M(\mathbf{x})},$$

then, by employing Lemma A.1 in Appendix A, we have

$$\begin{aligned}\mathcal{R}(\mathbf{k}_M) &= \frac{1}{\prod_{j=1}^{M-1} j!} \det_M\left[\frac{k_i!}{(k_i - M + j)!}\right] \\ &= \frac{1}{\prod_{j=1}^{M-1} j!} \Delta_M(\mathbf{k}) \end{aligned} \tag{3.44}$$

where the second equality is obtained by applying Lemma C.1 in Appendix C. Interested readers can apply (3.44) and the generalized Cauchy–Binet formula (Lemma B.2 in Appendix B) to (3.43), to obtain the joint pdf of the eigenvalues for the semi–correlated central Wishart matrix (3.34).

◊

3.3 Summary

In this chapter, we employed the integration framework proposed in Section 2.3 to solve three generalized unitary integrals. These integrals have applications in physics [85], [86] and the derivation of the joint eigenvalue distributions of Wishart matrices. Although the joint eigenvalue distributions of the semi–correlated central Wishart matrix and also noncentral cases were previously derived in the literature (by using considerably more complicated methods), this chapter demonstrates that the character expansion method can be used to obtain those results in a few simple steps. In addition, we were able to derive the joint eigenvalue distribution of the full–correlated non–square central Wishart matrix, which was not reported in the previous literature [57]. The approach presented in this chapter may be used to solve other unitary integrals accordingly.

Chapter 4

Eigenvalue Density of Wishart Matrices

In some applications, instead of joint eigenvalue distribution, the general eigenvalue density of a Wishart matrix (i.e., the distribution of a single eigenvalue without ordering) is of interest. Such applications in physics may include mesoscopics [105], high–energy physics [106], and econophysics [107]. Other applications may be found in the capacity and performance analysis of MIMO wireless systems, particularly when the channel state information is available for both the receiver and the transmitter [62]. It is shown that the optimum precoder matrix $\mathbf{F}$ in (1.7) should at the least diagonalize the channel matrix $\mathbf{H}$ [60]. Hence, the transmit signals are virtually transmitted over parallel channels configured in the eigenvalues of the channel matrix.

Due to the nature of most applications requiring the eigenvalue densities of large dimension random matrices, the eigenvalue density of Wishart matrices has been derived asymptotically in the literature [57]. However, for the capacity and performance analysis of MIMO systems with precoding, the exact eigenvalue densities are required since the number of transmit and/or receive antennas are relatively small.

The only paper that we found on the exact (not asymptotic) eigenvalue density of Wishart matrices is [108]. However, the unitary integration approach in this paper is identical to the approach in [95], which is not valid [85]. Nonetheless, the step–by–step approach in [108], from the joint eigenvalue density to the individual eigenvalue density, is valid, and results in neat expressions that are convenient for further analysis, particularly for the performance analysis of MIMO wireless channels with feedback. This fact motivated us to follow the same unified approach to obtain the eigenvalue densities of Wishart matrices.

4.1 Synopsis

Without loss of generality, we assume $\mathbf{H}$ is a $N \times M$, $N \geqslant M$, complex random matrix with Gaussian distribution. Given $P(\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_M\})$ is the joint pdf of the M nonzero eigenvalues of the matrix $\mathbf{HH}^*$, we define and derive the following quantities from $P(\boldsymbol{\lambda})$:

$$G_\nu(z) \triangleq \mathrm{E}\left\{\prod_{i=1}^{M}(\lambda_i - z)^\nu\right\} = \prod_{i=1}^{M}\int_0^\infty d\lambda_i\,(\lambda_i - z)^\nu\, P(\boldsymbol{\lambda}) \tag{4.1}$$

$$D(z) \triangleq \frac{\partial G_\nu(z)}{\partial \nu}\bigg|_{\nu=0} = \mathrm{E}\left\{\sum_{i=1}^{M}\log(\lambda_i - z)\right\} \tag{4.2}$$

$$P(\lambda) = \frac{1}{M}\lim_{\epsilon\to 0}\frac{D(\lambda - i\epsilon) - D(\lambda + i\epsilon)}{2\pi i} = \frac{1}{M}\,\mathrm{E}\left\{\sum_{i=1}^{M}u(\lambda - \lambda_i)\right\} \tag{4.3}$$

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{1}{M}\,\mathrm{E}\left\{\sum_{i=1}^{M}\delta(\lambda - \lambda_i)\right\} \tag{4.4}$$

where $u(\cdot)$ is the step function, and the multiplication before the integral in (4.1) represents the M–tuple integration. The partial differentiation in (4.2) produces a $\log(\lambda_i - z)$ function which becomes a step function after taking the limit in (4.3) since

$$\lim_{\epsilon\to 0}\mathrm{Im}\Big\{\log(-y + i\epsilon)\Big\} = \pi\, u(y)$$

for real y. The factor $\dfrac{1}{M}$ in (4.3) and (4.4) are introduced to make $P(\lambda)$ a cdf (cumulative density function) and $p(\lambda)$ a pdf.

In the following sections, we examine the above expressions for different $P(\boldsymbol{\lambda})$.

4.2 i.i.d. Central Wishart

We start with the popular case, where $\mathbf{H}$ is the standard complex Gaussian matrix. Substituting $P(\boldsymbol{\lambda})$ for this case from (3.31) into (4.1) results in

$$G_\nu(z) = \frac{\mathcal{K}_{N,M}}{M!}\prod_{i=1}^{M}\int_0^\infty d\lambda_i\,(\lambda_i - z)^\nu\,\mathrm{e}^{-\lambda_i}\lambda_i^{N-M}\underset{M}{\Delta}(\boldsymbol{\lambda})^2\,. \tag{4.5}$$

By using the Leibniz formula (C.4) in Appendix C.2 to expand the Vandermonde determinant (2.4), we can write

$$\underset{M}{\Delta}(\boldsymbol{\lambda})^2 = \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a})\,\mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \lambda_i^{2M-a_i-b_i}.$$

Thus, from (4.5), we obtain

$$\begin{aligned} G_\nu(z) &= \frac{\mathcal{K}_{N,M}}{M!} \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a})\,\mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty d\lambda_i \,(\lambda_i - z)^\nu \,\mathrm{e}^{-\lambda_i}\, \lambda_i^{N+M-a_i-b_i} \\ &= \mathcal{K}_{N,M} \times \frac{1}{M!} \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a})\,\mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty dt\,(t-z)^\nu \,\mathrm{e}^{-t}\, t^{N+M-a_i-b_i} \\ &= \mathcal{K}_{N,M} \underset{M}{\det}\left[\int_0^\infty dt\,(t-z)^\nu \,\mathrm{e}^{-t}\, t^{N+M-i-j} \right] \end{aligned} \tag{4.6}$$

where we use (C.5) in Appendix C.2 to obtain the determinant in (4.6). By differentiation with respect to ν (Appendix C.2), we have

$$D(z) = \frac{\partial G_\nu(z)}{\partial \nu}\bigg|_{\nu=0} = \mathcal{K}_{N,M} \sum_{m=1}^{M} \underset{M}{\det}[\,\mathbf{L}_m\,] \tag{4.7}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt\, \log(t-z)\,\mathrm{e}^{-t}\, t^{N+M-i-j} & ,\text{ if } \quad i = m\,; \\ (N+M-i-j)! & ,\text{ if } \quad i \neq m\,. \end{cases} \tag{4.8}$$

Considering that each determinant in (4.7) can be expanded by using the Laplace expansion formula [48] over the $(i=m)$-th row, the limit in (4.3) is absorbed into all $L_{m,mj}$ terms in (4.8). Therefore, the eigenvalue cdf becomes

$$P(\lambda) = \frac{\mathcal{K}_{N,M}}{M} \sum_{m=1}^{M} \underset{M}{\det}[\,\mathbf{L}'_m\,] \tag{4.9}$$

where

$$L'_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt\, u(\lambda - t)\,\mathrm{e}^{-t}\, t^{N+M-i-j} = \gamma\Big(N+M-i-j+1, \lambda\Big) & ,\text{ if } \quad i = m\,; \\ (N+M-i-j)! & ,\text{ if } \quad i \neq m\,; \end{cases}$$

and $\gamma\big(\alpha, x\big)$ is the lower incomplete Gamma function [109].

Hence, the eigenvalue density of the i.i.d. central Wishart matrix is achieved as

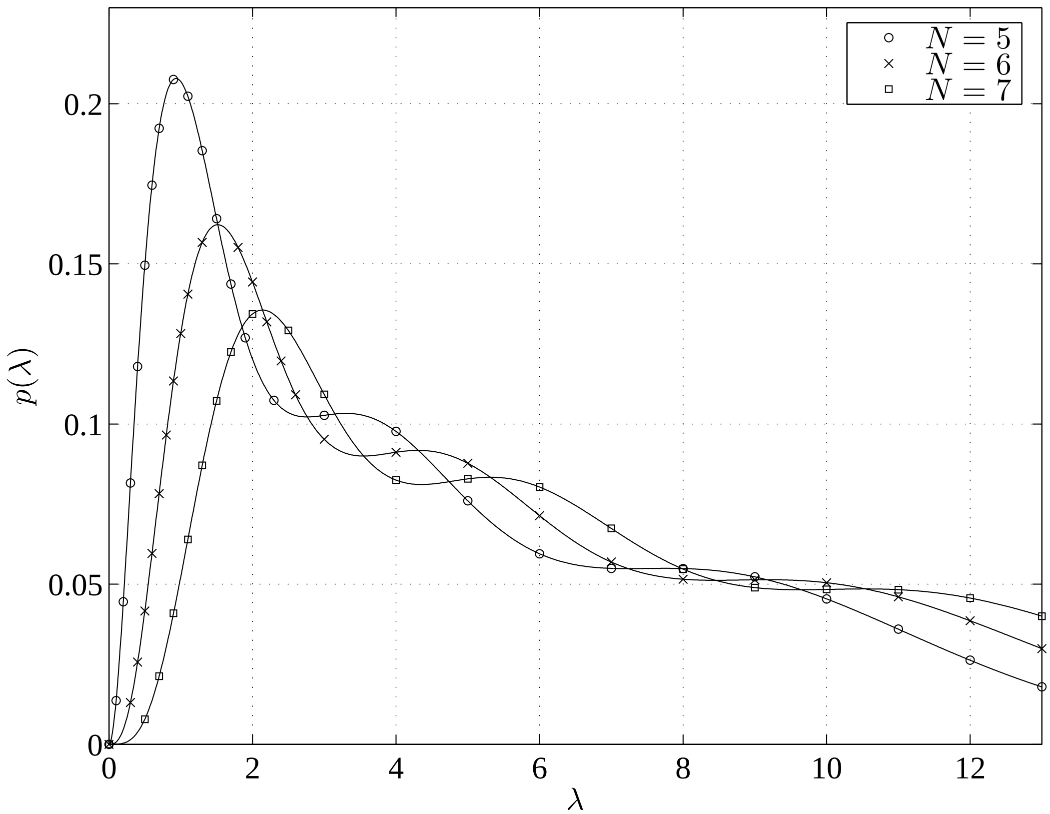

Figure 4.1: The eigenvalue density of i.i.d. central Wishart matrices with $M=3$ and $N=5,6$ and 7.

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{\mathcal{K}_{N,M}}{M} \sum_{m=1}^{M} \det_{M}[\,\mathbf{Q}_m\,] \tag{4.10}$$

where

$$Q_{m,ij} = \begin{cases} e^{-\lambda}\,\lambda^{N+M-i-j} & ,\text{ if } \quad i=m\,; \\ (N+M-i-j)! & ,\text{ if } \quad i\neq m\,. \end{cases}$$

■

Example 4.1 : Consider i.i.d. central Wishart matrices with $M=3$ and $N=5,6$ and 7. We use computer simulations to verify our analysis in this section. Figure 4.1 depicts the $p(\lambda)$ for these three cases where the solid curves are obtained from (4.10), and the symbols are obtained from computer simulations.

◊

4.3 Semi–Correlated Central Wishart

Similar to the assumption in Section 3.2.3, we assume $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}$ where $\mathbf{R} \in \mathrm{GL}(N, \mathcal{C})$ is the correlation matrix, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N\times M}$ is the standard complex Gaussian matrix. The joint eigenvalue distribution for the semi–correlated central Wishart matrix is given in (3.34), which is equal to

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{A}}{M!} \underset{M}{\Delta}(\boldsymbol{\lambda}) \underset{N}{\det} \begin{bmatrix} \mathrm{e}^{-y_j \lambda_i} \Big|_{M}^{i=1} \\ y_j^{N-i} \Big|_{N}^{i=M+1} \end{bmatrix} \tag{4.11}$$

where we define

$$\mathcal{A} \triangleq \frac{(-1)^{M(\frac{M+1}{2}-N)} \prod_{j=1}^{N} y_j^M}{\underset{N}{\Delta}(\mathbf{y}) \prod_{i=1}^{M-1} i!}, \tag{4.12}$$

and the vector $\mathbf{y} = (y_1, \ldots, y_N)^T$ represents the eigenvalues of matrix $\mathbf{R}^{-1}$. Substituting $P(\boldsymbol{\lambda})$ from (4.11) into (4.1) results in

$$G_\nu(z) = \frac{\mathcal{A}}{M!} \prod_{i=1}^{M} \int_0^\infty d\lambda_i \, (\lambda_i - z)^\nu \underset{M}{\Delta}(\boldsymbol{\lambda}) \underset{N}{\det} \begin{bmatrix} \mathrm{e}^{-y_j \lambda_i} \Big|_{M}^{i=1} \\ y_j^{N-i} \Big|_{N}^{i=M+1} \end{bmatrix}. \tag{4.13}$$

Using the Leibniz formula (C.4) in Appendix C.2 to expand both determinants in (4.13) gives

$$\begin{aligned} G_\nu(z) &= \frac{\mathcal{A}}{M!} \sum_{\mathbf{a}_M} \sum_{\mathbf{b}_N} \mathrm{S}(\mathbf{a}) \, \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty d\lambda_i \, (\lambda_i - z)^\nu \, \lambda_i^{M-a_i} \, \mathrm{e}^{-y_{b_i} \lambda_i} \prod_{i=M+1}^{N} y_{b_i}^{N-i} \\ &= \frac{\mathcal{A}}{M!} \sum_{\mathbf{a}_M} \sum_{\mathbf{b}_N} \mathrm{S}(\mathbf{a}) \, \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty dt \, (t - z)^\nu \, t^{M-a_i} \, \mathrm{e}^{-y_{b_i} t} \prod_{i=M+1}^{N} y_{b_i}^{N-i} \\ &= \mathcal{A} \sum_{\mathbf{b}_N} \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty dt \, (t - z)^\nu \, t^{M-i} \, \mathrm{e}^{-y_{b_i} t} \prod_{i=M+1}^{N} y_{b_i}^{N-i} \\ &= \mathcal{A} \underset{N}{\det} \begin{bmatrix} \int_0^\infty dt \, (t - z)^\nu \, t^{M-i} \, \mathrm{e}^{-y_j t} & \Big|_{M}^{i=1} \\ y_j^{N-i} & \Big|_{N}^{i=M+1} \end{bmatrix}. \end{aligned} \tag{4.14}$$

By differentiation with respect to ν (Appendix C.2), we have

$$D(z) = \frac{\partial G_\nu(z)}{\partial \nu}\Big|_{\nu=0} = \mathcal{A} \sum_{m=1}^{M} \underset{N}{\det}[\, \mathbf{L}_m \,] \tag{4.15}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt\, \log(t-z)\, \mathrm{e}^{-y_j t}\, t^{M-i} & \text{, if} \quad i = m \leqslant M\,; \\ y_j^{-(M-i+1)}\, \Gamma\big(M-i+1\big) & \text{, if} \quad i \neq m \leqslant M\,; \\ y_j^{N-i} & \text{, if} \quad i > M\,; \end{cases} \tag{4.16}$$

and $\Gamma\big(\alpha\big)$ is the Gamma function [109]. The index m in (4.15) goes to M only since the rest of the rows in (4.14) are independent of ν and result in zero determinants after differentiation.

Considering that each determinant in (4.15) can be expanded by using the Laplace expansion formula [48] over the $(i = m)$-th row, the limit in (4.3) is absorbed into all $L_{m,mj}$ terms in (4.16). Therefore,

$$P(\lambda) = \frac{\mathcal{A}}{M} \sum_{m=1}^{M} \det_{N}\big[\, \mathbf{L}'_m \,\big] \tag{4.17}$$

where

$$L'_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt\, u(\lambda - t)\, \mathrm{e}^{-y_j t}\, t^{M-i} & \text{, if} \quad i = m \leqslant M\,; \\ L_{m,ij} & \text{, otherwise}\,; \end{cases}$$

$$= \begin{cases} y_j^{-(M-i+1)}\, \gamma(M-i+1, \lambda y_j) & \text{, if} \quad i = m \leqslant M\,; \\ L_{m,ij} & \text{, otherwise.} \end{cases}$$

Hence, the eigenvalue density of the semi–correlated central Wishart matrix is achieved as

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{\mathcal{A}}{M} \sum_{m=1}^{M} \det_{N}\big[\, \mathbf{Q}_m \,\big] \tag{4.18}$$

where

$$Q_{m,ij} = \begin{cases} \mathrm{e}^{-\lambda y_j}\, \lambda^{M-i} & \text{, if} \quad i = m \leqslant M\,; \\ y_j^{-(M-i+1)}\, \Gamma(M-i+1) & \text{, if} \quad i \neq m \leqslant M\,; \\ y_j^{N-i} & \text{, if} \quad i > M\,. \end{cases}$$

$\blacksquare$

Accordingly, if we take $\mathbf{H} = \mathbf{G}\mathbf{T}^{\frac{1}{2}}$ where $\mathbf{T} \in \mathcal{C}^{M \times M}$ is a correlation matrix, we can still use (4.18) to obtain the eigenvalue density by setting $M = N$.

Example 4.2 : Consider a MIMO system where the receive and transmit antennas are deployed in a linear antenna array configuration. When the distance between antennas is small in comparison to the signal wavelength (λ_s), the channel gains become correlated. The elements of the correlation matrix $\mathbf{R}$ for this configuration are obtained from the following expression [95]:

$$R_{ab} = \int_{-180}^{180} \frac{d\phi}{\sqrt{2\pi\delta^2}} \exp\left(2\pi i(a-b)d_\lambda \sin\left(\frac{\phi\pi}{180}\right) - \frac{\phi^2}{2\delta^2}\right) \tag{4.19}$$

where δ in degrees is the angle spread measured from the vertical to the linear antenna array, and $d_\lambda = \dfrac{d_{\min}}{\lambda_s}$ is the normalized minimum distance between antennas.

Figure 4.2 illustrates the $p(\lambda)$ for the semi–correlated central Wishart matrices with $N = 5$, $M = 3$, $\delta = 10°$ and $d_\lambda = 0.8$, 1 and 2. The curves are obtained from (4.18), and the symbols are obtained from computer simulations. As we expected, when d_λ is small, the channel gains are more correlated, so that $p(\lambda)$ is sharper (denser) around zero. However, when d_λ is large, the channel gains are uncorrelated and, therefore, $p(\lambda)$ approaches the eigenvalue density of the 5×3 i.i.d. central Wishart matrix in Figure 4.1.

$\Diamond$

4.4 Uncorrelated Noncentral Wishart

Similar to the assumption in Section 3.2.4, assume the matrix $\mathbf{H}$ is modeled as $\mathbf{H} = \mathbf{G} + \mathbf{G}_0$ where $\mathbf{G}_0 \in \mathcal{C}^{N\times M}$ denotes the complex mean matrix, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N\times M}$ is the standard complex Gaussian matrix. The joint eigenvalue distribution for the uncorrelated noncentral Wishart matrix is given in (3.37), which is equal to

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{G}}{M!} \underset{M}{\Delta}(\boldsymbol{\lambda}) \det_{M}\left[\, f_j(\lambda_i)\,\right] \tag{4.20}$$

where we define

$$\mathcal{G} \triangleq \frac{1}{\underset{M}{\Delta}(\boldsymbol{\gamma})} \prod_{i=1}^{M} \frac{\mathrm{e}^{-\gamma_i}}{\gamma_i^{\frac{N-M}{2}}} \tag{4.21}$$

and

$$f_j(\lambda_i) \triangleq \mathrm{e}^{-\lambda_i}\, \lambda_i^{\frac{N-M}{2}}\, I_{N-M}\left(2\sqrt{\lambda_i\gamma_j}\right), \tag{4.22}$$

and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_M)^T$ are the M nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$.

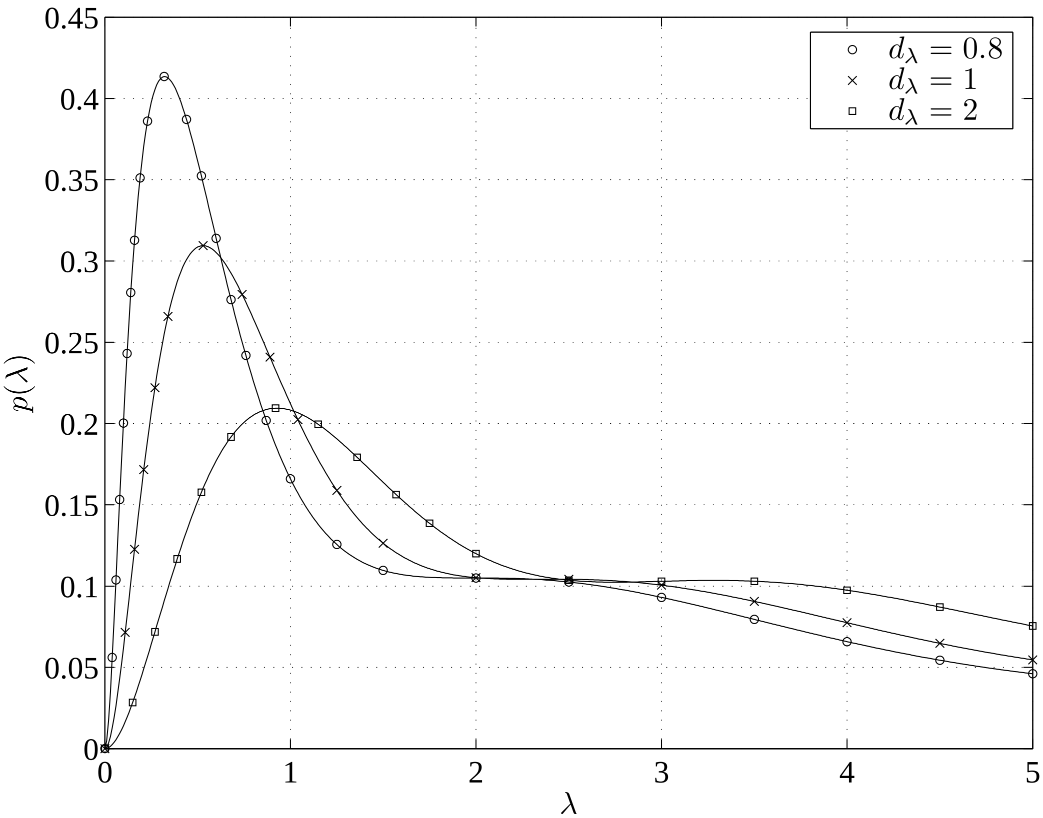

Figure 4.2: The eigenvalue density of semi–correlated central Wishart matrices with $N = 5$, $M = 3$, $\delta = 10^\circ$ and $d_\lambda = 0.8$, 1 and 2. The correlation factors are obtained from (4.19).

By substituting (4.20) into (4.1) and expanding the determinants by using the Leibniz formula (C.4) in Appendix C.2, we can write

$$
\begin{aligned}
G_\nu(z) &= \frac{\mathcal{G}}{M!} \prod_{i=1}^{M} \int_0^\infty d\lambda_i \, (\lambda_i - z)^\nu \, \underset{M}{\Delta}(\boldsymbol{\lambda}) \, \underset{M}{\det}\big[\, f_j(\lambda_i) \,\big] \\
&= \frac{\mathcal{G}}{M!} \prod_{i=1}^{M} \int_0^\infty d\lambda_i \, (\lambda_i - z)^\nu \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \prod_{i=1}^{M} \lambda_i^{M-a_i} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} f_{b_i}(\lambda_i) \\
&= \mathcal{G} \times \frac{1}{M!} \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a}) \, \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty dt \, (t - z)^\nu \, t^{M-a_i} f_{b_i}(t) \\
&= \mathcal{G} \, \underset{M}{\det}\left[\int_0^\infty dt \, (t - z)^\nu \, t^{M-i} f_j(t) \right] \\
&= \mathcal{G} \, \underset{M}{\det}\left[\int_0^\infty dt \, (t - z)^\nu \, \mathrm{e}^{-t} \, t^{\frac{N+M}{2} - i} \, I_{N-M}\left(2\sqrt{\gamma_j \, t}\,\right) \right]
\end{aligned}
\tag{4.23}
$$

where we use (C.5) in Appendix C.2 to obtain the determinant.

Differentiating (4.23) with respect to ν (Appendix C.2) results in

$$
D(z) = \frac{\partial G_\nu(z)}{\partial \nu}\bigg|_{\nu=0} = \mathcal{G} \sum_{m=1}^{M} \underset{M}{\det}[\, \mathbf{L}_m \,] \tag{4.24}
$$

where

$$
L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt \log(t - z) \, \mathrm{e}^{-t} \, t^{\frac{N+M}{2} - i} \, I_{N-M}\left(2\sqrt{\gamma_j \, t}\,\right) & , \text{if} \quad i = m\,; \\[2ex] \displaystyle\frac{(N-i)!}{(N-M)!} \, \gamma_j^{\frac{N-M}{2}} \, \Phi\Big(N - i + 1, N - M + 1; \gamma_j\Big) & , \text{if} \quad i \neq m\,; \end{cases}
$$

and $\Phi\big(a, b; x\big)$ is the confluent hypergeometric function [109]. Taking the limit in (4.3) by using (4.24) gives

$$
P(\lambda) = \frac{\mathcal{G}}{M} \sum_{m=1}^{M} \underset{M}{\det}\big[\, \mathbf{L}'_m \,\big] \tag{4.25}
$$

where

$$
L'_{m,ij} = \begin{cases} \displaystyle\int_0^\lambda dt \, \mathrm{e}^{-t} \, t^{\frac{N+M}{2} - i} \, I_{N-M}\left(2\sqrt{\gamma_j \, t}\,\right) & , \text{if} \quad i = m\,; \\[2ex] L_{m,ij} & , \text{otherwise.} \end{cases}
$$

Therefore, the eigenvalue density of the uncorrelated noncentral Wishart matrix is derived as

$$
p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{\mathcal{G}}{M} \sum_{m=1}^{M} \underset{M}{\det}[\, \mathbf{Q}_m \,] \tag{4.26}
$$

where

$$Q_{m,ij} = \begin{cases} \mathrm{e}^{-\lambda}\,\lambda^{\frac{N+M}{2}-i}\, I_{N-M}\Big(2\sqrt{\lambda\gamma_j}\Big) & , \text{ if } \quad i = m\,; \\ \dfrac{(N-i)!}{(N-M)!}\,\gamma_j^{\frac{N-M}{2}}\;\Phi\Big(N-i+1, N-M+1; \gamma_j\Big) & , \text{ if } \quad i \neq m\,; \end{cases}$$

■

Example 4.3 : By assuming that the three nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$ are $\boldsymbol{\gamma} = (2,5,7)^T$, the eigenvalue densities of the uncorrelated noncentral Wishart matrices with $M = 3$ and $N = 5, 6$ and 7 can be illustrated as in Figure 4.3, where the curves are obtained from (4.26), and the symbols are computer simulation results. As observed in Figure 4.3, a perfect agreement exists between the analytical results and the simulation results.

◊

4.5 i.i.d. Noncentral Wishart

The i.i.d. noncentral Wishart matrix is a special case of the uncorrelated noncentral Wishart matrix when all elements of the matrix $\mathbf{H}$ are i.i.d. $\mathcal{CN}(g,1)$ random variables; i.e., $\mathbf{H} \in \mathcal{CN}(g,1)^{N\times M}$. Therefore, $\gamma = MN\,|g|^2$ is the *only* nonzero eigenvalue of the matrix $\mathbf{G}_0\mathbf{G}_0^*$. The joint eigenvalue distribution for the i.i.d. noncentral Wishart matrix is given in (3.40), which is equal to

$$P(\boldsymbol{\lambda}) = \frac{\mathcal{G}_\gamma}{M!}\,\underset{M}{\Delta}(\boldsymbol{\lambda})\,\underset{M}{\det}\big[\, f_j(\lambda_i)\,\big] \tag{4.27}$$

where we define

$$\mathcal{G}_\gamma \triangleq \mathcal{K}_{N-1,M-1}\left[\frac{\mathrm{e}^{-\gamma}}{\gamma^{\frac{N+M}{2}-1}}\right] \tag{4.28}$$

and

$$f_j(\lambda_i) \triangleq \begin{cases} \mathrm{e}^{-\lambda_i}\,\lambda_i^{\frac{N-M}{2}}\; I_{N-M}\Big(2\sqrt{\gamma\lambda_i}\Big) & , \text{ if } \quad j = 1\,; \\ \mathrm{e}^{-\lambda_i}\,\lambda_i^{N-j} & , \text{ if } \quad j = 2, \ldots, N\,. \end{cases} \tag{4.29}$$

By substituting (4.27) into (4.1) and expanding the determinants by using the Leibniz formula (C.4) in Appendix C.2, we follow the same procedure presented in (4.23) to obtain

$$G_\nu(z) = \mathcal{G}_\gamma\,\underset{M}{\det}\left[\int_0^\infty dt\,(t-z)^\nu\, t^{M-i} f_j(t)\right]. \tag{4.30}$$

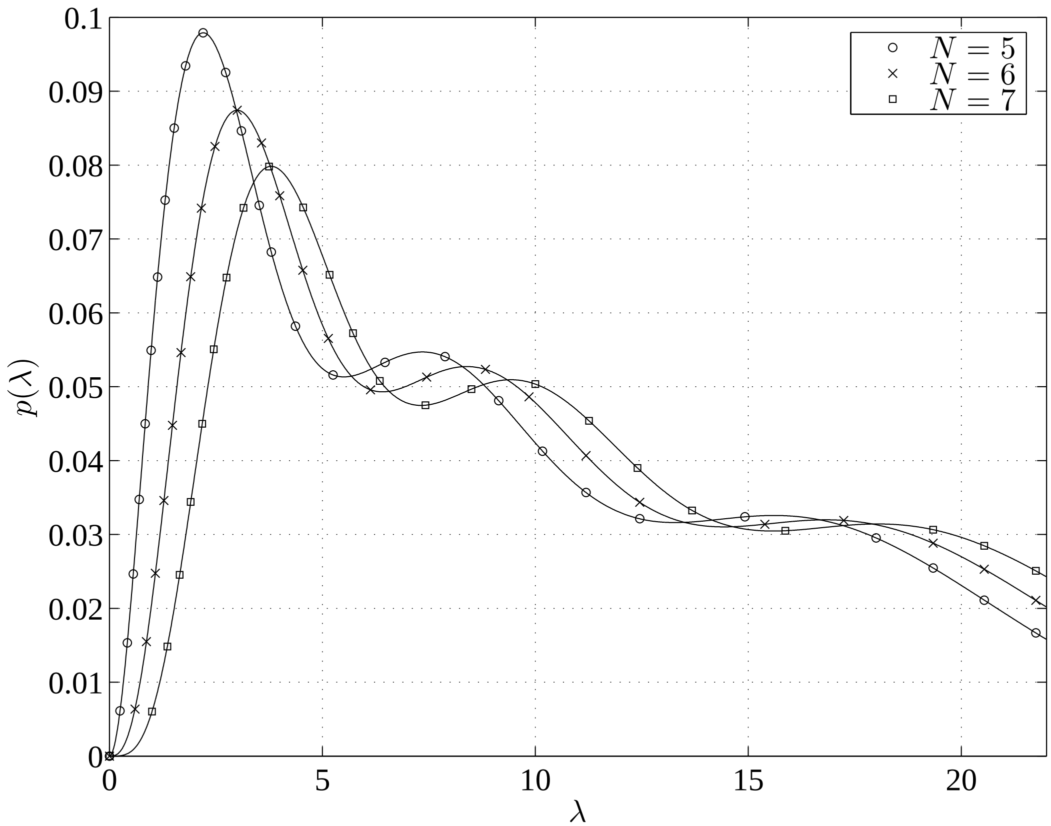

Figure 4.3: The eigenvalue density of uncorrelated noncentral Wishart matrices with $M = 3$, $\boldsymbol{\gamma} = (2, 5, 7)^T$ and $N = 5, 6$ and 7.

Differentiating (4.30) with respect to ν (Appendix C.2) results in

$$D(z) = \frac{\partial G_\nu(z)}{\partial \nu}\Big|_{\nu=0} = \mathcal{G}_\gamma \sum_{m=1}^{M} \det_M[\,\mathbf{L}_m\,] \tag{4.31}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty dt \log(t-z)\, \mathrm{e}^{-t}\, t^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma\, t}\right) & \text{, if } \quad i=m\,,\ j=1\,; \\ \displaystyle\int_0^\infty dt \log(t-z)\, \mathrm{e}^{-t}\, t^{N+M-i-j} & \text{, if } \quad i=m\,,\ j>1\,; \\ \dfrac{(N-i)!}{(N-M)!}\, \gamma^{\frac{N-M}{2}}\, \Phi\Big(N-i+1, N-M+1; \gamma\Big) & \text{, if } \quad i\neq m\,,\ j=1\,; \\ (N+M-i-j)! & \text{, if } \quad i\neq m\,,\ j>1\,. \end{cases}$$

Taking the limit in (4.3) by using (4.31) gives

$$P(\lambda) = \frac{\mathcal{G}_\gamma}{M} \sum_{m=1}^{M} \det_M\big[\,\mathbf{L}'_m\,\big] \tag{4.32}$$

where

$$L'_{m,ij} = \begin{cases} \displaystyle\int_0^\lambda dt\, \mathrm{e}^{-t}\, t^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma\, t}\right) & \text{, if } \quad i=m\,,\ j=1\,; \\ \gamma\Big(N+M-i-j+1, \lambda\Big) & \text{, if } \quad i=m\,,\ j>1\,; \\ L_{m,ij} & \text{, otherwise}\,; \end{cases}$$

and $\gamma\big(\alpha, x\big)$ is the lower incomplete Gamma function [109].

Therefore, the eigenvalue density of the i.i.d. noncentral Wishart matrix is obtained as

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{\mathcal{G}_\gamma}{M} \sum_{m=1}^{M} \det_M[\,\mathbf{Q}_m\,] \tag{4.33}$$

where

$$Q_{m,ij} = \begin{cases} \mathrm{e}^{-\lambda}\, \lambda^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma\lambda}\right) & \text{, if } \quad i=m\,,\ j=1\,; \\ \mathrm{e}^{-\lambda}\, \lambda^{N+M-i-j} & \text{, if } \quad i=m\,,\ j>1\,; \\ \dfrac{(N-i)!}{(N-M)!}\, \gamma^{\frac{N-M}{2}}\, \Phi\Big(N-i+1, N-M+1; \gamma\Big) & \text{, if } \quad i\neq m\,,\ j=1\,; \\ (N+M-i-j)! & \text{, if } \quad i\neq m\,,\ j>1\,. \end{cases}$$

■

4.6 Full–Correlated Central Wishart

Similar to the assumption in Section 3.2.6, we assume $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}\mathbf{T}^{\frac{1}{2}}$ where $\mathbf{R} \in \mathrm{GL}(N, \mathcal{C})$ and $\mathbf{T} \in \mathrm{GL}(M, \mathcal{C})$ are the correlation matrices, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N\times M}$ is the standard complex Gaussian matrix. The joint eigenvalue distribution for the full–correlated central Wishart matrix is given in (3.43), which is equal to

$$P(\boldsymbol{\lambda}) = \sum_{\mathbf{k}_M} \frac{\mathcal{B}}{M!\, \underset{M}{\Delta}(\mathbf{k})}\, \underset{M}{\Delta}(\boldsymbol{\lambda})\, \underset{M}{\det}\left[\, \lambda_i^{k_j+N-M} \,\right] \tag{4.34}$$

where $\mathbf{k}_M$ is a representation of $\mathrm{GL}(M, \mathcal{C})$, and the auxiliary parameter $\mathcal{B}$ represents the excluded terms of (3.43) in (4.34), which are independent of $\boldsymbol{\lambda}$ as well.

Since $\underset{M}{\Delta}(\boldsymbol{\lambda})$ can be written as (Appendix C.3)

$$\underset{M}{\Delta}(\boldsymbol{\lambda}) = \frac{1}{(-z)^{\frac{M(M-1)}{2}}}\, \underset{M}{\det}\left[\left(\frac{\lambda_i}{\lambda_i - z} \right)^{M-j} \right] \prod_{i=1}^{M} (\lambda_i - z)^{M-1}, \tag{4.35}$$

and by applying the Leibniz formula (C.4) in Appendix C.2 to expand the determinants in (4.34), we can write

$$\begin{aligned} G_\nu(z) =& \sum_{\mathbf{k}_M} \frac{(-z)^{-\frac{M(M-1)}{2}}\, \mathcal{B}}{M!\, \underset{M}{\Delta}(\mathbf{k})} \prod_{i=1}^{M} \int_0^\infty d\lambda_i\, (\lambda_i - z)^{\nu+M-1} \\ & \times \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \prod_{i=1}^{M} \left(\frac{\lambda_i}{\lambda_i - z} \right)^{M-a_i} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \lambda_i^{k_{b_i}+N-M} \\ =& \sum_{\mathbf{k}_M} \frac{(-z)^{-\frac{M(M-1)}{2}}\, \mathcal{B}}{\underset{M}{\Delta}(\mathbf{k})} \times \frac{1}{M!} \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a})\, \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} \int_0^\infty dt\, (t-z)^{\nu+a_i-1}\, t^{k_{b_i}+N-a_i} \\ =& \sum_{\mathbf{k}_M} \frac{(-z)^{-\frac{M(M-1)}{2}}\, \mathcal{B}}{\underset{M}{\Delta}(\mathbf{k})}\, \underset{M}{\det}\left[\int_0^\infty dt\, (t-z)^{\nu+i-1}\, t^{k_j+N-i} \right] \end{aligned} \tag{4.36}$$

where the last equality comes from (C.5) in Appendix C.2. Performing $(M-i)$ times integrations by part for the integral in (4.36) results in

$$\begin{aligned} G_\nu(z) =& \sum_{\mathbf{k}_M} \frac{(-z)^{-\frac{M(M-1)}{2}}\, \mathcal{B}}{\underset{M}{\Delta}(\mathbf{k})} \\ & \times \underset{M}{\det}\left[\frac{(-1)^{M-i}}{\prod_{l=1}^{M-i}(\nu+i+l-1)} \times \frac{(\alpha_j+M-i)!}{\alpha_j!} \int_0^\infty dt\, (t-z)^{\nu+M-1}\, t^{\alpha_j} \right] \end{aligned} \tag{4.37}$$

where $\alpha_j \triangleq k_j + N - M$. The row and column factoring from the determinant in (4.37) gives

$$G_\nu(z) = \sum_{\mathbf{k}_M} \frac{z^{-\frac{M(M-1)}{2}}\,\mathcal{B}}{\underset{M}{\Delta}(\mathbf{k})} \det_M\left[\frac{(\alpha_j + M - i)!}{\alpha_j!}\right] \prod_{j=1}^{M} \frac{\int_0^\infty dt\,(t-z)^{\nu+M-1}\,t^{\alpha_j}}{(\nu+j-1)^{j-1}}. \tag{4.38}$$

By following the same approach as shown in Lemma C.1 in Appendix C.1, it is easy to show that

$$\det_M\left[\frac{(\alpha_j + M - i)!}{\alpha_j!}\right] = \underset{M}{\Delta}(\boldsymbol{\alpha}) = \underset{M}{\Delta}(\mathbf{k} + N - M) = \underset{M}{\Delta}(\mathbf{k}) \tag{4.39}$$

where the last equality comes from Appendix C.3. The last equality in (4.39) allows us to cancel the term $\underset{M}{\Delta}(\mathbf{k})$ in (4.38) and use the generalized Cauchy–Binet formula as follows. Otherwise, the final result cannot be shown as a determinant.

Substituting $\mathcal{B}$ from (3.43), and (4.39) into (4.38) results in

$$\begin{aligned}
G_\nu(z) &= \sum_{\mathbf{k}_M} z^{-\frac{M(M-1)}{2}}\,\mathcal{B} \prod_{j=1}^{M} \frac{\int_0^\infty dt\,(t-z)^{\nu+M-1}\,t^{k_j+N-M}}{(\nu+j-1)^{j-1}} \\
&= \frac{(-z)^{-\frac{M(M-1)}{2}} \prod_{i=1}^{M} x_i^N \prod_{j=1}^{N} y_j^M}{\underset{M}{\Delta}(\mathbf{x})\,\underset{N}{\Delta}(\mathbf{y}) \prod_{i=1}^{M-1} (\nu+i)^i} \\
&\qquad \times \sum_{\mathbf{k}_M} \det_M\left[\,(-x_i)^{k_j}\,\right] \det_N\left[\, y_i^{k_j+N-M}\Big|_{j=1}^{M},\ y_i^{N-j}\Big|_{j=M+1}^{N}\right] \\
&\qquad\qquad \times \prod_{i=1}^{M} \int_0^\infty dt\,(t-z)^{\nu+M-1}\,\frac{t^{k_i+N-M}}{(k_i+N-M)!}
\end{aligned}$$

where the M–dimensional vector $\mathbf{x}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension M, and the N–dimensional vector $\mathbf{y}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension N ($N \geqslant M$).

Considering the fact that

$$\int_0^\infty dt\,(t-z)^{\nu+M-1}\,\mathrm{e}^{tx} = \sum_{k=0}^{\infty} \left[\int_0^\infty dt\,(t-z)^{\nu+M-1}\,\frac{t^k}{k!}\right] x^k,$$

and by applying the generalized Cauchy–Binet formula (Lemma B.2 in Appendix

B), we obtain

$$G_\nu(z) = \frac{(-1)^{M(N-\frac{M+1}{2})} \prod_{i=1}^M x_i^M \prod_{j=1}^N y_j^M}{z^{\frac{M(M-1)}{2}} \underset{M}{\Delta}(\mathbf{x}) \underset{N}{\Delta}(\mathbf{y}) \prod_{i=1}^{M-1} (\nu+i)^i} \det_N \begin{bmatrix} \int_0^\infty dt\, (t-z)^{\nu+M-1} \, \mathrm{e}^{-x_i y_j t} \Big|_M^{i=1} \\ y_j^{N-i} \qquad\qquad \Big|_N^{i=M+1} \end{bmatrix}$$

$$= \frac{(-1)^{M(N-\frac{M+1}{2})} z^{-\frac{M(M-1)}{2}}}{\underset{M}{\Delta}(\mathbf{x}) \underset{N}{\Delta}(\mathbf{y}) \prod_{i=1}^{M-1} (\nu+i)^i} \det_N \begin{bmatrix} (x_i y_j)^M \int_0^\infty dt\, (t-z)^{\nu+M-1} \, \mathrm{e}^{-x_i y_j t} \Big|_M^{i=1} \\ y_j^{N+M-i} \qquad\qquad \Big|_N^{i=M+1} \end{bmatrix}$$

$$= \frac{\mathcal{X}\, z^{-\frac{M(M-1)}{2}}}{\prod_{i=1}^{M-1} (\nu+i)^i} \det_N \begin{bmatrix} F\big(x_i y_j\,, \nu+M, z\big) \Big|_M^{i=1} \\ y_j^{N+M-i} \qquad \Big|_N^{i=M+1} \end{bmatrix} \tag{4.40}$$

where we define

$$\begin{aligned} F\big(x, \alpha, z\big) &\triangleq x^M \int_0^\infty dt\, (t-z)^{\alpha-1} \, \mathrm{e}^{-xt} \\ &= x^{M-\alpha} \, \mathrm{e}^{-xz} \, \Gamma\big(\alpha, -xz\big) \end{aligned} \tag{4.41}$$

and

$$\mathcal{X} \triangleq \frac{(-1)^{M(N-\frac{M+1}{2})}}{\underset{M}{\Delta}(\mathbf{x}) \underset{N}{\Delta}(\mathbf{y})} \tag{4.42}$$

and $\Gamma\big(\alpha, x\big)$ is the upper incomplete Gamma function [109]. For integer m,

$$F\big(x, m, z\big) = x^{M-m} \, (m-1)! \sum_{k=0}^{m-1} \frac{(-xz)^k}{k!} \,. \tag{4.43}$$

By differentiating (4.40) with respect to ν (Appendix C.2), we have

$$D(z) = \frac{\partial G_\nu(z)}{\partial \nu}\Big|_{\nu=0} = 1 - M + \frac{\mathcal{X}\, z^{-\frac{M(M-1)}{2}}}{\prod_{i=1}^{M-1} i^i} \sum_{m=1}^{M} \det_N [\, \mathbf{L}_m \,] \tag{4.44}$$

where

$$L_{m,ij} = \begin{cases} (x_i y_j)^M \int_0^\infty dt \log(t-z)\, (t-z)^{M-1} \, \mathrm{e}^{-x_i y_j t} & \text{, if} \quad i = m \leqslant M \,; \\ F\big(x_i y_j, M, z\big) & \text{, if} \quad i \neq m \leqslant M \,; \\ y_j^{N+M-i} & \text{, if} \quad i > M \,. \end{cases}$$

The index m in (4.44) goes to M only since the rest of the rows in (4.40) are independent of ν and result in zero determinants after differentiation.

Taking the limit in (4.3) by using (4.44) results in

$$P(\lambda) = \frac{\mathcal{X}\lambda^{-\frac{M(M-1)}{2}}}{M\prod_{i=1}^{M-1} i^i} \sum_{m=1}^{M} \det_{N}\left[\,\mathbf{L}'_m\,\right] \tag{4.45}$$

where

$$L'_{m,ij} = \begin{cases} E\big(x_i y_j \lambda\,, M\big) & ,\text{ if } \quad i = m \leqslant M\,; \\ F\big(x_i y_j\,, M, \lambda\big) & ,\text{ if } \quad i \neq m \leqslant M\,; \\ y_j^{N+M-i} & ,\text{ if } \quad i > M\,; \end{cases}$$

and we define

$$\begin{aligned} E\big(x, M\big) &\triangleq \int_0^x dt\,(-t)^{M-1}\,\mathrm{e}^{t-x} \\ &= (M-1)!\left[-\mathrm{e}^{-x} + \sum_{m=0}^{M-1} \frac{(-x)^m}{m!}\right]. \end{aligned} \tag{4.46}$$

Therefore, the eigenvalue density of the full–correlated central Wishart matrix is obtained as

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} = \frac{\mathcal{X}\lambda^{-\frac{M(M-1)}{2}-1}}{M\prod_{i=1}^{M-1} i^i} \sum_{m=1}^{M}\sum_{n=1}^{M} \det_{N}\left[\,\mathbf{Q}_{mn}\,\right] \tag{4.47}$$

where

$$Q_{mn,ij} = \begin{cases} -\Big[\Big(\frac{M(M-1)}{2} + x_i y_j \lambda\Big) E\big(x_i y_j \lambda, M\big) + (-x_i y_j \lambda)^M\Big] & ,\text{ if } \quad i = n = m \leqslant M\,; \\ -(M-1)\,F\big(x_i y_j, M-1, \lambda\big) & ,\text{ if } \quad i = n \neq m \leqslant M\,; \\ \lambda\,E\big(x_i y_j \lambda, M\big) & ,\text{ if } \quad i = m \neq n \leqslant M\,; \\ F\big(x_i y_j, M, \lambda\big) & ,\text{ if } \quad i \neq m, n \leqslant M\,; \\ y_j^{N+M-i} & ,\text{ if } \quad i > M\,; \end{cases}$$

and the functions $E\big(x, M\big)$ and $F\big(x, m, z\big)$ are defined in (4.46) and (4.43), respectively. As before, the M–dimensional vector $\mathbf{x}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension M, and the N–dimensional vector $\mathbf{y}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension N $(N \geqslant M)$.

■

Example 4.4 : To verify (4.47), we consider the MIMO wireless system presented in Example 4.2 where we now assume both the transmitter and the receiver antennas

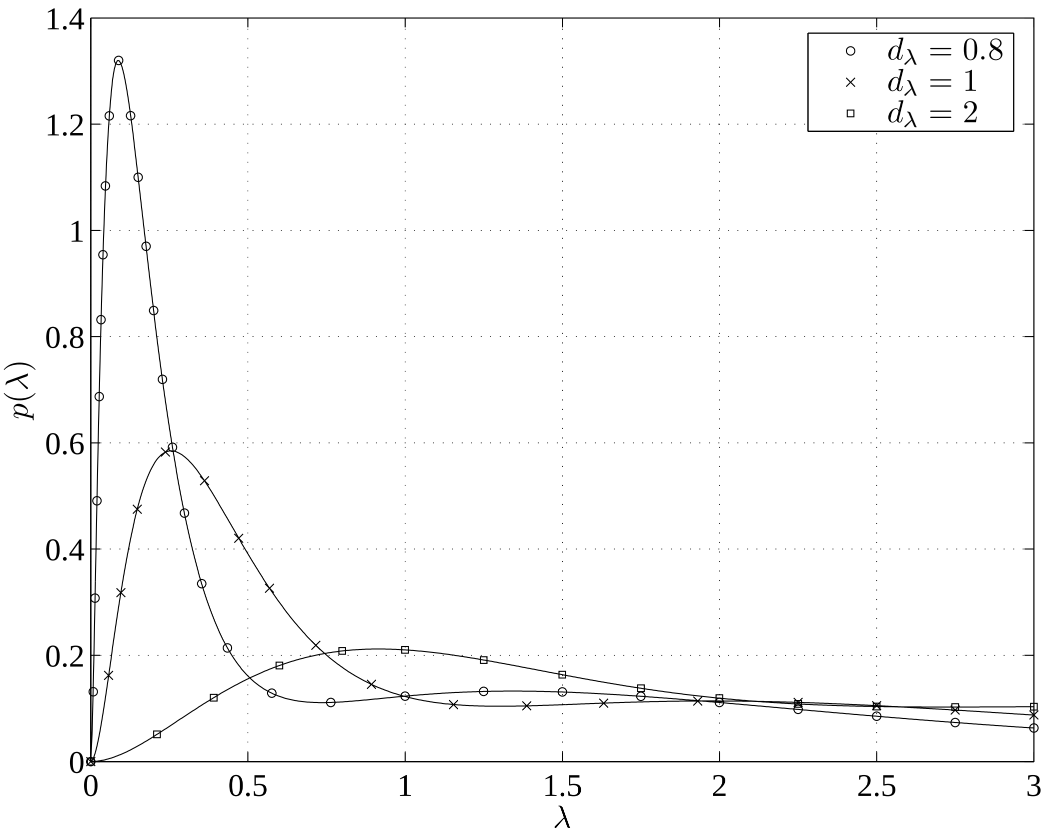

Figure 4.4: The eigenvalue density of full–correlated central Wishart matrices with $N = 5$, $M = 3$, $\delta = 10^\circ$ and $d_\lambda = 0.8$, 1 and 2. The correlation factors are obtained from (4.19).

are correlated in a linear antenna array deployment. Therefore, the correlation matrices $\mathbf{R}$ and $\mathbf{T}$ can be obtained from (4.19).

The $p(\lambda)$ for full–correlated central Wishart matrices with $N = 5$, $M = 3$, $\delta = 10^\circ$ and $d_\lambda = 0.8$, 1 and 2 are presented in Figure 4.4. The curves are obtained from (4.47), and the symbols are obtained from computer simulations. As expected, when the channel gains are more correlated or, equivalently, when d_λ is small, the $p(\lambda)$ is sharper (denser) around zero. However, when d_λ is large, the $p(\lambda)$ approaches the eigenvalue density of the 5×3 i.i.d. central Wishart matrix in Figure 4.1. $\Diamond$

4.7 Summary

Due to the nature of most applications, the eigenvalue densities of Wishart matrices have been derived asymptotically in the literature [57]. However, for the capacity and performance analysis of MIMO systems with precoding, the exact eigenvalue densities are required since the number of transmit and/or receive antennas are relatively small.

Because of the convenient mathematical form of the joint eigenvalue distributions derived in Chapter 3, we were able to derive the exact eigenvalue densities of Wishart matrices straightforwardly. Our results are in the form of finite summations of determinants, which can be easily computed and also employed for further analysis.

Chapter 5

Capacity Analysis of MIMO Systems

The capacity of MIMO systems is commonly analyzed by using the moment generating function of the mutual information between the transmitter and receiver, for various assumptions about the statistics of the channel matrix. The first derivative of the MGF then yields the capacity. Moreover, the probability of outage can be derived through a simple numerical integral [110]. The outage mutual information for the Gaussian uncorrelated channels, obtained by using the MGF, is presented in [111], and the capacity of MIMO systems when the channel is Ricean can be found in [104], and in [112] and [113] when the channel is semi–correlated. Specifically, [112] deals with the case where the number of correlated antennas is less than or equal to the number of uncorrelated antennas, and [113] with the opposite case. All these works are based on the results derived by James [59].

Recently, the character expansion method has been used in [102] for the capacity analysis of full–correlated MIMO channel when $N_r = N_t$, and in [95] for the capacity analysis of semi–correlated, Ricean and full–correlated MIMO channels with arbitrary numbers of transmit and receive antennas. However, we showed in the previous chapters that the unitary integration approach presented in [95] is mathematically invalid when $N_r \neq N_t$.

In this chapter, we use the joint eigenvalue distributions derived in Chapter 3 to obtain the capacity of corresponding MIMO channels. Although the expressions are not similar to our results, these capacity results have been derived before with much more effort. However, the approach presented in this thesis is significantly simple, compact, straightforward and unified.

5.1 Synopsis

By recalling the MIMO wireless system model from (1.2) in section 1.1, and assuming that the channel matrix $\mathbf{H}$ is known to the receiver only, the mutual information between the transmitter and receiver is given by [20]

$$\mathcal{I} = \log\left(\det_{N_r}[\,\mathbf{I}_{N_r} + \rho\,\mathbf{H}\mathbf{H}^*\,]\right)$$

where $\log(\cdot)$ is the natural logarithm. By defining the MGF of $\mathcal{I}$ as

$$\begin{aligned} g(z) &\triangleq \mathrm{E}\{\mathrm{e}^{z\mathcal{I}}\} \\ &= \mathrm{E}\left\{\det_{N_r}[\,\mathbf{I}_{N_r} + \rho\,\mathbf{H}\mathbf{H}^*\,]^z\right\}, \end{aligned} \tag{5.1}$$

the capacity of the system is obtained by direct differentiation of $g(z)$; i.e.,

$$C = \mathrm{E}\{\mathcal{I}\} = g'(0). \tag{5.2}$$

Unlike (1.6), the bandwidth B is set to unity here. Thus, the capacity in (5.2) is measured in nats per second per Hertz (nats/s/Hz).

From (5.1), it is clear that the generating function can be written simply in terms of the eigenvalues $\{\lambda_i\}$ of the matrix $\mathbf{H}\mathbf{H}^*$ as

$$\begin{aligned} g(z) &= \mathrm{E}\left\{\prod_{i=1}^{M}(1+\rho\lambda_i)^z\right\} \\ &= \prod_{i=1}^{M}\int_0^\infty d\lambda_i\,(1+\rho\lambda_i)^z\,P(\boldsymbol{\lambda}) \end{aligned} \tag{5.3}$$

where $M = \min\{N_t, N_r\}$, and $P(\boldsymbol{\lambda}=\{\lambda_1,\ldots,\lambda_M\})$ is the joint pdf of the M nonzero eigenvalues of matrix $\mathbf{H}\mathbf{H}^*$.

Considering the fact that

$$(1+\rho\lambda)^z = \rho^z\left[\lambda - \left(\frac{-1}{\rho}\right)\right]^z,$$

and comparing $g(z)$ in (5.3) with $G_\nu(z)$ in (4.1), we observe that

$$g(z) = \rho^{zM}\,G_z\left(\frac{-1}{\rho}\right). \tag{5.4}$$

Hence, instead of substituting the corresponding $P(\boldsymbol{\lambda})$ for different channel statistics into (5.3), we use the appropriate $G_\nu(z)$ from Chapter 4 to obtain the MGF of the mutual information, $g(z)$ and, consequently, the corresponding capacity.

5.2 i.i.d. Rayleigh MIMO Channel

When all channel gains in $\mathbf{H}$ are i.i.d. $\mathcal{CN}(0,1)$ random variables, the channel is called the standard Rayleigh fading channel. Therefore, the matrix $\mathbf{HH}^*$ becomes the i.i.d. central Wishart matrix whose joint eigenvalue distribution is given in (3.31). Substituting $G_\nu(z)$ for this case from (4.6) into (5.4) results in

$$\begin{aligned} g(z) &= \rho^{zM}\, \mathcal{K}_{N,M} \det_M \left[\int_0^\infty d\lambda \left(\lambda - \left(\frac{-1}{\rho} \right) \right)^z \mathrm{e}^{-\lambda}\, \lambda^{N+M-i-j} \right] \\ &= \mathcal{K}_{N,M} \det_M \left[\int_0^\infty d\lambda\, (1+\rho\lambda)^z\, \mathrm{e}^{-\lambda}\, \lambda^{N+M-i-j} \right] \end{aligned} \tag{5.5}$$

where $N = \max\{N_t, N_r\}$, $M = \min\{N_t, N_r\}$ and $\mathcal{K}_{N,M}$ is defined in (3.28).

Differentiating (5.5) with respect to z (Appendix C.2) gives

$$C = g'(0) = \mathcal{K}_{N,M} \sum_{m=1}^{M} \det_M [\, \mathbf{L}_m\,]$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty d\lambda\, \log(1+\rho\lambda)\, \mathrm{e}^{-\lambda}\, \lambda^{N+M-i-j} & \text{, if} \quad i = m\,; \\ (N+M-i-j)! & \text{, if} \quad i \neq m\,. \end{cases}$$

■

5.3 Semi–Correlated Rayleigh MIMO Channel

In this scenario, the channel matrix is assumed to be correlated at one side of the transmission only. Without loss of generality, we assume that the channel matrix is correlated at the receiver side. Thus, it can be modeled as $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}$ where $\mathbf{R} \in \mathrm{GL}(N_r, \mathcal{C})$ denotes the receiver correlation matrix, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N_r \times N_t}$ is a standard Rayleigh matrix. Therefore, the matrix $\mathbf{HH}^*$ becomes the semi–correlated central Wishart matrix whose joint eigenvalue distribution is given in (3.34). Substituting $G_\nu(z)$ for this case from (4.14) into (5.4) results in

$$\begin{aligned} g(z) &= \rho^{zM}\, \mathcal{A} \det_N \left[\begin{array}{cc} \displaystyle\int_0^\infty d\lambda \left(\lambda - \left(\frac{-1}{\rho} \right) \right)^z \lambda^{M-i}\, \mathrm{e}^{-y_j \lambda} & \Big|_{M}^{i=1} \\ y_j^{N-i} & \Big|_{N}^{i=M+1} \end{array} \right] \\ &= \mathcal{A} \det_N \left[\begin{array}{cc} \displaystyle\int_0^\infty d\lambda\, (1+\rho\lambda)^z\, \lambda^{M-i}\, \mathrm{e}^{-y_j \lambda} & \Big|_{M}^{i=1} \\ y_j^{N-i} & \Big|_{N}^{i=M+1} \end{array} \right] \end{aligned} \tag{5.6}$$

where $M = \min(N_t, N_r)$, $N = \max(N_t, N_r)$, $\mathcal{A}$ is defined in (4.12), and the vector $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ represents the eigenvalues of matrix $\mathbf{R}^{-1}$.

By differentiating (5.6) with respect to z (Appendix C.2), we have

$$C = g'(0) = \mathcal{A} \sum_{m=1}^{M} \det_{N}[\, \mathbf{L}_m \,] \tag{5.7}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty d\lambda \, \log(1+\rho\lambda)\, \lambda^{M-i}\, \mathrm{e}^{-y_j \lambda} & , \text{if} \quad i = m \leqslant M\,; \\ y_j^{-(M-i+1)}\, \Gamma\big(M-i+1\big) & , \text{if} \quad i \neq m \leqslant M\,; \\ y_j^{N-i} & , \text{if} \quad i > M\,. \end{cases}$$

The index m in (5.7) goes to M only since the rest of the rows in (5.6) are independent of z and result in zero determinants after differentiation.

■

5.4 Uncorrelated Ricean MIMO Channel

When the channel gains in $\mathbf{H}$ are complex Gaussian random variables with a nonzero mean, the channel is called a Ricean fading channel. In the uncorrelated scenario, the matrix $\mathbf{H}$ is modeled as $\mathbf{H} = \mathbf{G} + \mathbf{G}_0$ where $\mathbf{G}_0 \in \mathcal{C}^{N_r \times N_t}$ denotes the complex mean matrix, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N_r \times N_t}$ is a standard Rayleigh matrix. Therefore, the matrix $\mathbf{HH}^*$ becomes the uncorrelated noncentral Wishart matrix whose joint eigenvalue distribution is given in (3.37). Substituting $G_\nu(z)$ for this case from (4.23) into (5.4) results in

$$\begin{aligned} g(z) &= \rho^{zM}\, \mathcal{G} \det_{M}\left[\int_0^\infty d\lambda \left(\lambda - \left(\frac{-1}{\rho}\right)\right)^z \mathrm{e}^{-\lambda}\, \lambda^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma_j\,\lambda}\right) \right] \\ &= \mathcal{G} \det_{M}\left[\int_0^\infty d\lambda \, (1+\rho\lambda)^z\, \mathrm{e}^{-\lambda}\, \lambda^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma_j\,\lambda}\right) \right] \end{aligned} \tag{5.8}$$

where $N = \max(N_t, N_r)$, $M = \min(N_t, N_r)$, the parameter $\mathcal{G}$ is defined in (4.21), and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_M)^T$ are the M nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$.

Differentiating (5.8) with respect to z (Appendix C.2) gives

$$C = g'(0) = \mathcal{G} \sum_{m=1}^{M} \det_{M}[\, \mathbf{L}_m \,] \tag{5.9}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty d\lambda \log(1+\rho\lambda)\, \mathrm{e}^{-\lambda}\, \lambda^{\frac{N+M}{2}-i}\, I_{N-M}\left(2\sqrt{\gamma_j\,\lambda}\right) & \text{, if} \quad i = m\,; \\ \displaystyle\frac{(N-i)!}{(N-M)!}\, \gamma_j^{\frac{N-M}{2}}\, \Phi\Big(N-i+1, N-M+1; \gamma_j\Big) & \text{, if} \quad i \neq m\,; \end{cases}$$

and $\Phi\big(a, b; x\big)$ is the confluent hypergeometric function [109].

■

Example 5.1 : Consider MIMO wireless systems with $N_t = 3$ transmit antennas and $N_r = 3$, 4 and 5 receive antennas. Assuming $\boldsymbol{\gamma} = (0.1, 0.3, 0.7)^T$ are the three nonzero eigenvalues of the matrix $\mathbf{G}_0\mathbf{G}_0^*$ for given systems, Figure 5.1 illustrates the capacity of MIMO systems versus SNR (ρ) when the channel is uncorrelated Ricean fading. The curves are plotted by using (5.9), and the symbols are computer simulation results. As observed in Figure 5.1, a perfect agreement exists between the analytical results and the simulation results.

◊

In the case of equal eigenvalues for the matrix $\mathbf{G}_0\mathbf{G}_0^*$, one can use Lemma A.1 in Appendix A to obtain the appropriate MGF. One such example occurs when the MIMO channel is an i.i.d. Ricean fading channel.

5.5 i.i.d. Ricean MIMO Channel

The i.i.d. Ricean MIMO channel is a special case of the uncorrelated Ricean MIMO channel when all elements of the channel matrix $\mathbf{H}$ are i.i.d. $\mathcal{CN}(g, 1)$ random variables; i.e., $\mathbf{H} \in \mathcal{CN}(g, 1)^{N_r \times N_t}$. Therefore, $\gamma = N_r N_t\, |g|^2$ is the *only* nonzero eigenvalue of the matrix $\mathbf{G}_0\mathbf{G}_0^*$. As a result, the matrix $\mathbf{HH}^*$ becomes the i.i.d. noncentral Wishart matrix whose joint eigenvalue distribution is given in (3.40). Substituting $G_\nu(z)$ for this case from (4.30) into (5.4) results in

$$\begin{aligned} g(z) &= \rho^{zM}\, \mathcal{G}_\gamma \det_M \left[\int_0^\infty d\lambda \left(\lambda - \left(\frac{-1}{\rho}\right)\right)^z \lambda^{M-i} f_j(\lambda) \right] \\ &= \mathcal{G}_\gamma \det_M \left[\int_0^\infty d\lambda\, (1+\rho\lambda)^z\, \lambda^{M-i} f_j(\lambda) \right] \end{aligned} \tag{5.10}$$

where the parameter $\mathcal{G}_\gamma$ and the function $f_j(\lambda)$ are defined in (4.28) and (4.29), respectively.

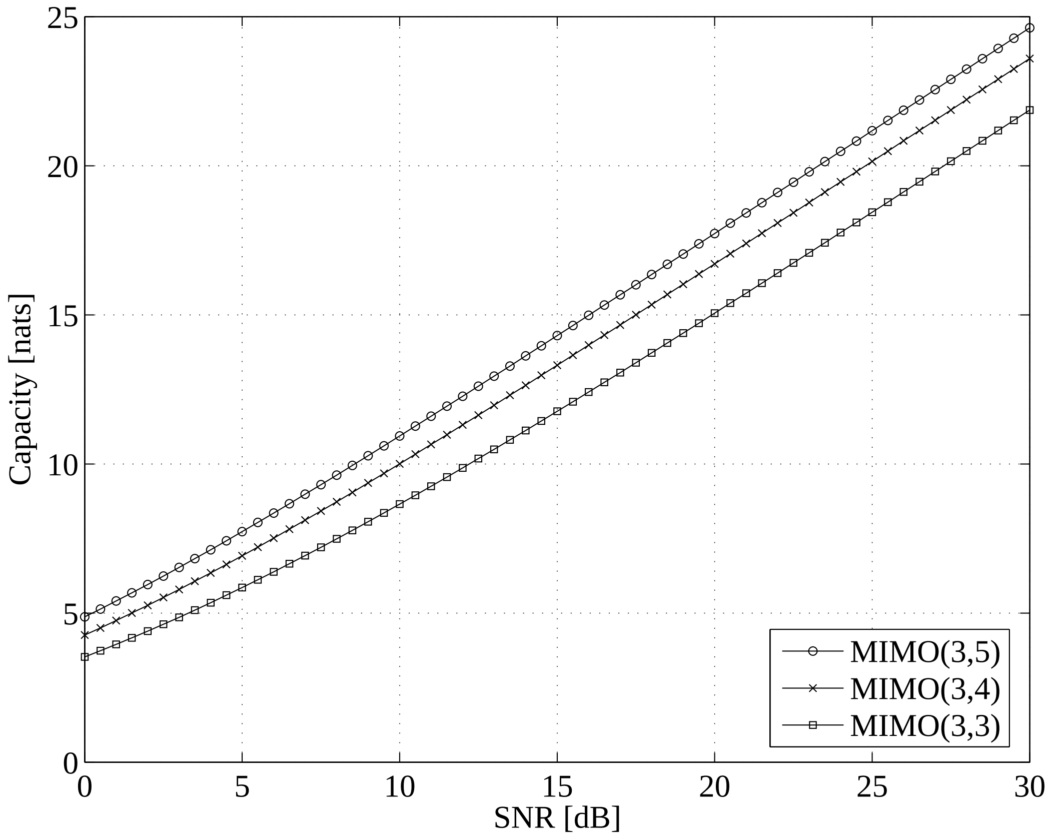

Figure 5.1: The capacity of uncorrelated Ricean MIMO channels with eigenvalues $\boldsymbol{\gamma} = (0.1, 0.3, 0.7)^T$ for the matrix $\mathbf{G}_0\mathbf{G}_0^*$.

Differentiating (5.10) with respect to z (Appendix C.2) gives

$$C = g'(0) = \mathcal{G}_\gamma \sum_{m=1}^{M} \det_M [\, \mathbf{L}_m \,] \tag{5.11}$$

where

$$L_{m,ij} = \begin{cases} \displaystyle\int_0^\infty d\lambda \log(1+\rho\lambda)\, \mathrm{e}^{-\lambda}\, \lambda^{\frac{N+M}{2}-i}\, I_{N-M}\Big(2\sqrt{\gamma\,\lambda}\Big) & , \text{if} \quad i = m\,,\, j = 1\,; \\ \displaystyle\int_0^\infty d\lambda \log(1+\rho\lambda)\, \mathrm{e}^{-\lambda}\, \lambda^{N+M-i-j} & , \text{if} \quad i = m\,,\, j > 1\,; \\ \dfrac{(N-i)!}{(N-M)!}\, \gamma^{\frac{N-M}{2}}\, \Phi\Big(N-i+1, N-M+1; \gamma\Big) & , \text{if} \quad i \neq m\,,\, j = 1\,; \\ (N+M-i-j)! & , \text{if} \quad i \neq m\,,\, j > 1\,. \end{cases}$$

■

Example 5.2 : Consider MIMO wireless systems with $N_t = 3$ transmit antennas and $N_r = 3$, 4 and 5 receive antennas. Assuming that the SNR is fixed at $\rho = 10$dB, Figure 5.2 shows the capacity of MIMO systems when the channel is i.i.d. Ricean fading where all elements of the channel matrix have $\mathcal{CN}(g, 1)$ distribution. In this figure, the curves are obtained by using (5.11), and the symbols are from computer simulations.

◊

5.6 Full–Correlated Rayleigh MIMO Channel

In this scenario, the channel matrix is assumed to be correlated at both sides of the communication link. Thus, it can be modeled as $\mathbf{H} = \mathbf{R}^{\frac{1}{2}}\mathbf{G}\mathbf{T}^{\frac{1}{2}}$ where $\mathbf{R} \in \mathrm{GL}(N_r, \mathcal{C})$ and $\mathbf{T} \in \mathrm{GL}(N_t, \mathcal{C})$ denote the receiver and the transmitter correlation matrices, respectively, and $\mathbf{G} \in \mathcal{CN}(0,1)^{N_r \times N_t}$ is a standard Rayleigh matrix. Therefore, the matrix $\mathbf{HH}^*$ becomes the full–correlated central Wishart matrix whose joint eigenvalue distribution is given in (3.43). Substituting $G_\nu(z)$ for this case from (4.40) into (5.4) results in

$$g(z) = \rho^{zM} \frac{\mathcal{X}\left(\frac{-1}{\rho}\right)^{-\frac{M(M-1)}{2}}}{\prod_{i=1}^{M-1}(z+i)^i} \det_N \begin{bmatrix} F\left(x_i y_j\,, z+M, \dfrac{-1}{\rho}\right) \Big|_M^{i=1} \\ y_j^{N+M-i} \qquad\qquad \Big|_N^{i=M+1} \end{bmatrix}$$

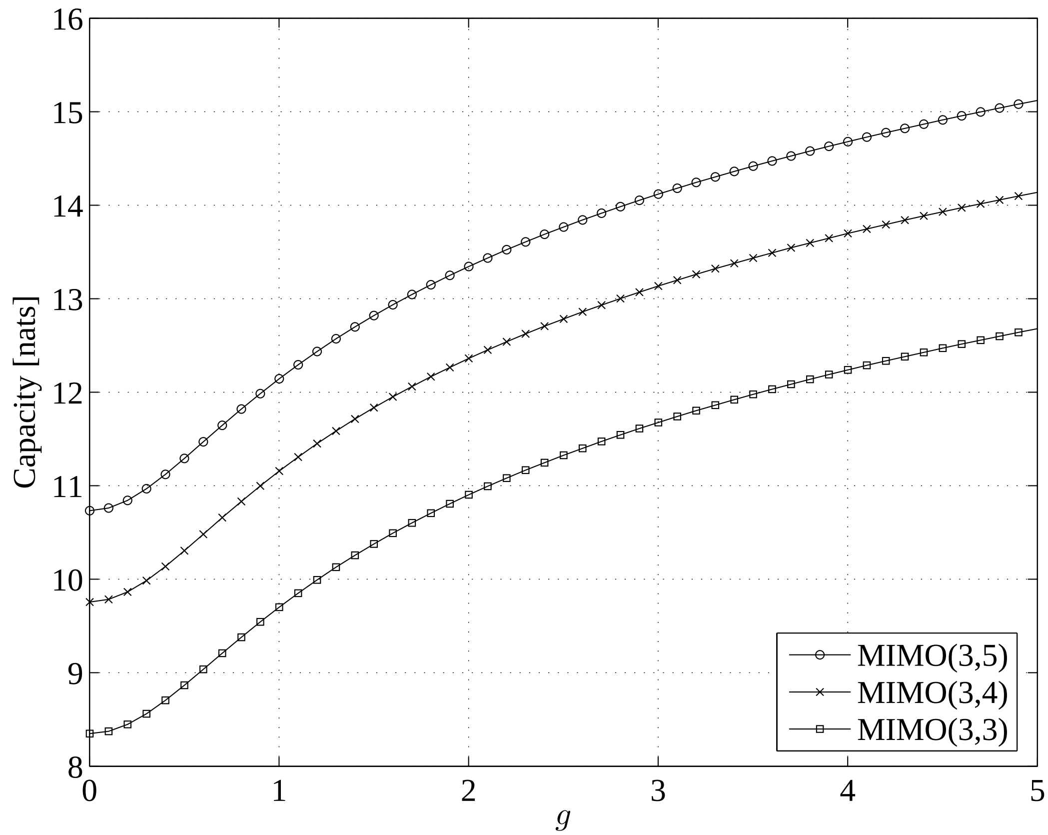

Figure 5.2: The capacity of i.i.d. Ricean MIMO channels at SNR=10dB with $\mathcal{CN}(g,1)$ distribution.

where $M = \min(N_t, N_r)$, $N = \max(N_t, N_r)$, and the parameter $\mathcal{X}$ and the function $F\big(x, \alpha, z\big)$ are defined in (4.42) and (4.41), respectively. The M–dimensional vector $\mathbf{x}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension M, and the N–dimensional vector $\mathbf{y}$ represents the eigenvalues of the matrix $\mathbf{R}^{-1}$ or $\mathbf{T}^{-1}$, whichever has the dimension N ($N \geqslant M$).

Considering the fact that

$$F\left(x, \alpha, \frac{-1}{\rho}\right) = \rho^{M-\alpha}\, F\left(\frac{x}{\rho}, \alpha, -1\right),$$

we have

$$g(z) = \frac{\mathcal{Y}}{\prod_{i=1}^{M-1}(z+i)^i} \det_N \begin{bmatrix} F\left(\dfrac{x_i y_j}{\rho}, z+M, -1\right) \Big|_{M}^{i=1} \\ y_j^{N+M-i} \qquad \Big|_{N}^{i=M+1} \end{bmatrix} \tag{5.12}$$

where we define

$$\mathcal{Y} \triangleq \frac{(-1)^{M(N-M)}\, \rho^{\frac{M(M-1)}{2}}}{\underset{M}{\Delta}(\mathbf{x})\, \underset{N}{\Delta}(\mathbf{y})}\,. \tag{5.13}$$

In the event of equal elements in $\mathbf{x}$ and/or $\mathbf{y}$, one can use Lemma A.1 in Appendix A to obtain the appropriate MGF. The semi–correlated and i.i.d. Rayleigh fading channels are two such examples where in the first case $\mathbf{x} = \mathbf{1}$, and in the second case $\mathbf{x} = \mathbf{1}$ and $\mathbf{y} = \mathbf{1}$.

Differentiating (5.12) with respect to z (Appendix C.2) results in

$$C = g'(0) = 1 - M + \frac{\mathcal{Y}}{\prod_{i=1}^{M-1} i^i} \sum_{m=1}^{M} \det_N[\,\mathbf{L}_m\,] \tag{5.14}$$

where

$$L_{m,ij} = \begin{cases} \rho\left(\dfrac{x_i\, y_j}{\rho}\right)^M \displaystyle\int_0^\infty d\lambda \log(1+\rho\lambda)\, (1+\rho\lambda)^{M-1}\, \mathrm{e}^{-\lambda x_i y_j} & \text{, if} \quad i = m \leqslant M\,; \\ F\left(\dfrac{x_i\, y_j}{\rho}, M, -1\right) & \text{, if} \quad i \neq m \leqslant M\,; \\ y_j^{N+M-i} & \text{, if} \quad i > M\,. \end{cases}$$

The index m in (5.14) goes up to M since the rest of the rows in (5.12) are independent of z and result in zero determinants after differentiation.

■

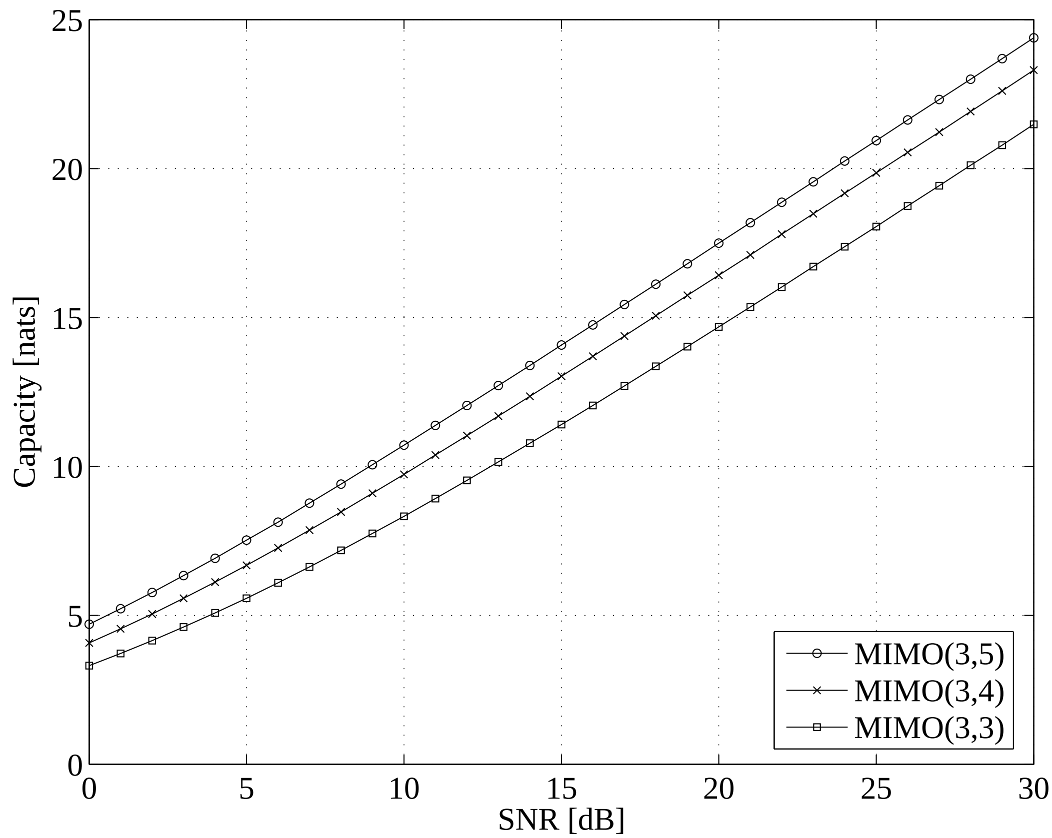

Figure 5.3: The capacity of full–correlated Rayleigh MIMO channels with $d_\lambda = 2$ and $\delta = 10°$.

Example 5.3 : Consider MIMO systems with $N_t = 3$ transmit antennas and $N_r = 3$, 4 and 5 receive antennas where similar to the Example 4.2, both the transmitter and the receiver antennas are correlated in a linear antenna array deployment. Therefore, the correlation matrices $\mathbf{R}$ and $\mathbf{T}$ can be obtained from (4.19).

Figure 5.3 shows the capacity of MIMO systems versus SNR (ρ) when the channel is full–correlated Rayleigh fading. The results in Figure 5.3 are obtained by assuming $d_\lambda = 2$ and $\delta = 10°$ in (4.19).

Figure 5.4 shows the capacity of MIMO systems at $\rho = 10$dB when the channel is full–correlated Rayleigh fading. In this figure, the results are obtained by assuming $\delta = 10°$ in (4.19).

In both figures, the solid curves are from (5.14), and the symbols are obtained by computer simulations. As observed in both figures, the analytical results and simulation results are in perfect accord.

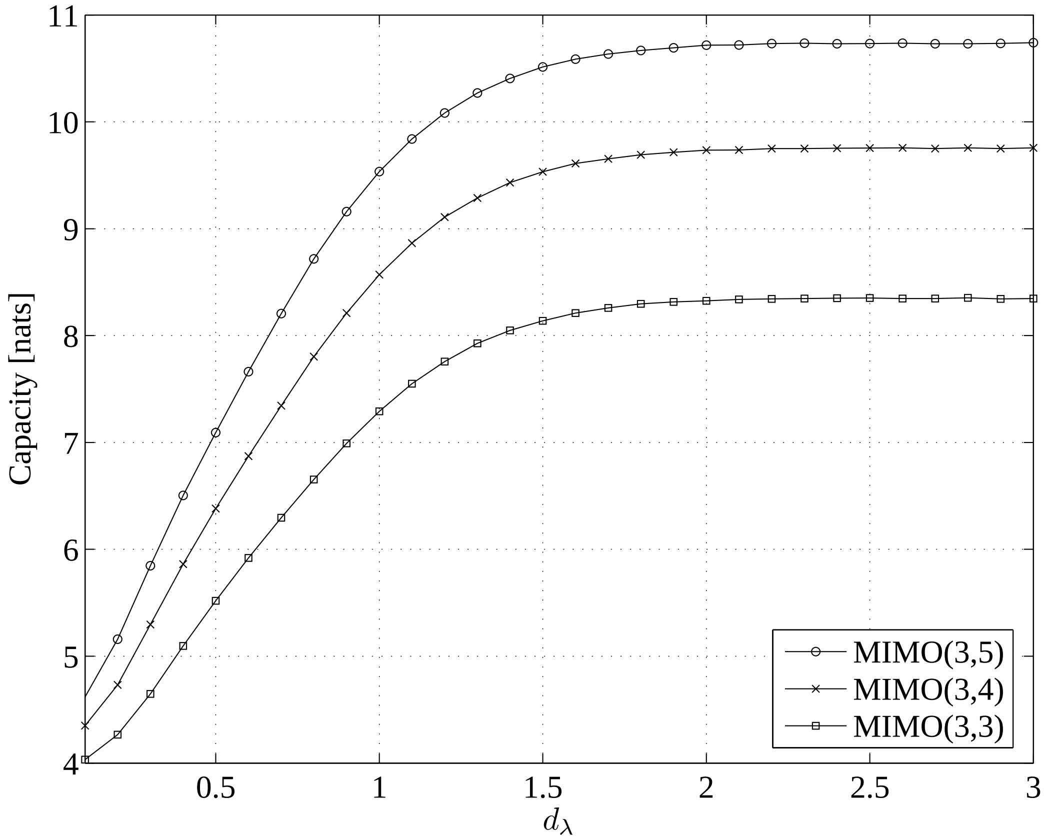

Figure 5.4: The capacity of full–correlated Rayleigh MIMO channels at SNR=10dB with $\delta = 10°$.

5.7 Summary

In this chapter, we used the results derived in Chapter 4 to obtain the capacity of MIMO channels. Compared to the previous approaches in the literature, the approach presented in this thesis is exceptionally simple, compact and straightforward. Interested readers can compare the capacity derivation in this thesis with the derivations in [20], [104], and [112]– [114] to observe the differences. For instance, the authors of [104], [112] and [113] manipulate the joint eigenvalue distributions of Wishart matrices, derived by James [59], to obtain similar determinant–form expressions as in this thesis and, then, use them for the capacity analysis of MIMO channels.

The character expansion method provides new insights into the way we can obtain the eigenvalue distributions of Wishart matrices and, consequently, the capacity of MIMO wireless channels.

Chapter 6

Design, Selection Algorithm and Performance Analysis of Limited Feedback Transmit Beamforming

Popular MIMO techniques commonly assume the availability of channel state information at the receiver, but not at the transmitter. However, in a slow fading environment, complete [60]– [62] or partial CSI [67]– [71] may be available at the transmitter. CSI at the transmitter may be exploited in two ways: antenna subset selection [63]– [66] and precoding. The optimum precoder matrix can be obtained based on the eigenvalue structure of the channel matrix [60].

In transmit beamforming, the optimal beamformer selects the subchannel corresponding to the largest singular value of the channel matrix by weighting the transmit signal with the corresponding eigenvector [115]. Transmit beamforming can achieve full diversity and array gain with a simple receiver structure. However, the bandwidth limits on the feedback channel [116] restrict the availability of full CSI at the transmitter. Hence, limited feedback beamforming techniques are of interest.

In limited feedback beamforming, the transmitter and receiver share a codebook of beamformers. Codebook design can be based on either vector quantization and the Lloyd algorithm [67], [117]– [119] or the maximization of the minimum distance between each pair of beamformers in the codebook [68], [70], [71]. The simulation results show that the codebooks obtained by both methods perform identically in Rayleigh fading channels [120], [121]. The design complexity of these methods is large when the size of the codebook is large.

The authors in [70] map the design problem into the Grassmannian line packing (GLP) problem [122], and use the unitary structure presented in [24] for codebook design. This structure, which was originally proposed for differential unitary space–time modulation, consists of a diagonal matrix and a rectangular sub–matrix of the Discrete Fourier Transform matrix. The diagonal terms are points on the unit circle in the complex plane where their angles are defined by integers that should be optimized offline. These angles are the only parameters that should be saved at the transmitter and receiver. Therefore, the implementation of the codebooks based on GLP needs small resources (memory), whereas in other design methods, the whole codebook should be saved at the transmitter and receiver.

This thesis uses the same unitary structure as in [24], [70] and [71]. In these papers, the optimum rotation matrix is obtained via an exhaustive search for small dimensions (number of antennas and/or codebook size), and a random search for large dimensions. We propose to use the genetic algorithm [123] to find the optimum parameters. For this purpose, the design parameters are relaxed from positive integers [71] to positive real values. The simulation results show that the genetic codebooks not only achieve a larger minimum distance than those of [71], but also reduce the optimization complexity.

In limited feedback beamforming, the receiver selects the best beamformer for each realization of the channel by exhaustive search over the codebook. However, by exploiting the specific structure of the GLP codebooks, we present an order and bound algorithm to reduce the receiver's search complexity. In this algorithm, the beamformers of the codebook are ordered based on their vicinity to the optimal beamformer. The original beamformer selection, which is a maximization problem, is converted to a minimization problem, allowing the use of bounding techniques. Metric Bounding is the basic scheme used by minimization algorithms such as sphere decoders [44] to avoid unnecessary computations.

The performance of optimal transmit beamforming has been analyzed in the literature. In [124], the cdf and pdf of the largest singular value of the channel matrix, presented in the form of generalized hypergeometric functions, are used to calculate the outage probability in a system with optimal transmit beamforming. The expressions in [125] are in the form of infinite series of averaged BER over the distribution of the largest eigenvalue of the i.i.d. central Wishart matrix, which itself is in the form of a hypergeometric function with matrix arguments. This fact motivates another representation for the pdf of the largest eigenvalue of the i.i.d. central Wishart matrix. We use the joint eigenvalue distribution of the i.i.d. central Wishart, derived in section 3.2.2, to obtain the pdf of the largest eigenvalue in a finite series expression, which is appropriate for the BER analysis of the optimal and limited feedback transmit beamforming.

The performance analysis of limited feedback beamforming is complicated for general MIMO systems. However, for multiple–input single–output (MISO) systems, the outage probability of transmit beamforming has been studied in [68], the symbol error rate with transmit correlation in [126], and a lower bound on the symbol error rate in [120]. In [127], a framework using high–resolution quantization theory is proposed for distortion analysis of MIMO systems with feedback, where the distortion function is the capacity loss in a MISO system with limited feedback. Similarly, the authors in [121], use the SNR loss and outage capacity for distortion analysis of MIMO systems with limited feedback beamforming. In this thesis, however, by assuming a large–size codebook as in [127] and [121] (high–resolution analysis), we analyze the BER of limited feedback MIMO beamforming. By employing a geometrical approach, we derive an approximate BER of limited feedback MIMO beamforming in closed–form. The simulation results show that the approximate BER is comparatively tight even for small–size codebooks.

6.1 System Model

Recalling the MIMO(N_t, N_r) wireless system from (1.7), and setting $T = 1$, we model the linear transformation between the transmit and receive antennas in a transmit beamforming scenario as

$$\mathbf{y} = \sqrt{\rho}\,\mathbf{H}\mathbf{f}s + \mathbf{n} \tag{6.1}$$

where the vector $\mathbf{y} \in \mathcal{C}^{Nr}$ is the complex received vector, $s \in \mathcal{C}$ is the transmitted signal, $\mathbf{f} \in \mathcal{C}^{N_t}$ is the beamformer vector, $\mathbf{n} \in \mathcal{CN}(0,1)^{N_r}$ is the additive noise vector, $\mathbf{H} \in \mathcal{CN}(0,1)^{N_r \times N_t}$ is the standard Rayleigh channel matrix, and ρ is the total transmit power at each signaling interval. Entries of $\mathbf{H}$ and $\mathbf{n}$ are independent and identically distributed. For each transmission, according to the input information, s is selected from a signal constellation (e.g., PAM or QAM) with unit average energy. The transmit signal is weighted and parallelized by the beamformer $\mathbf{f}$ to be sent over N_t transmit antennas. To ensure that the transmit power on each signaling interval

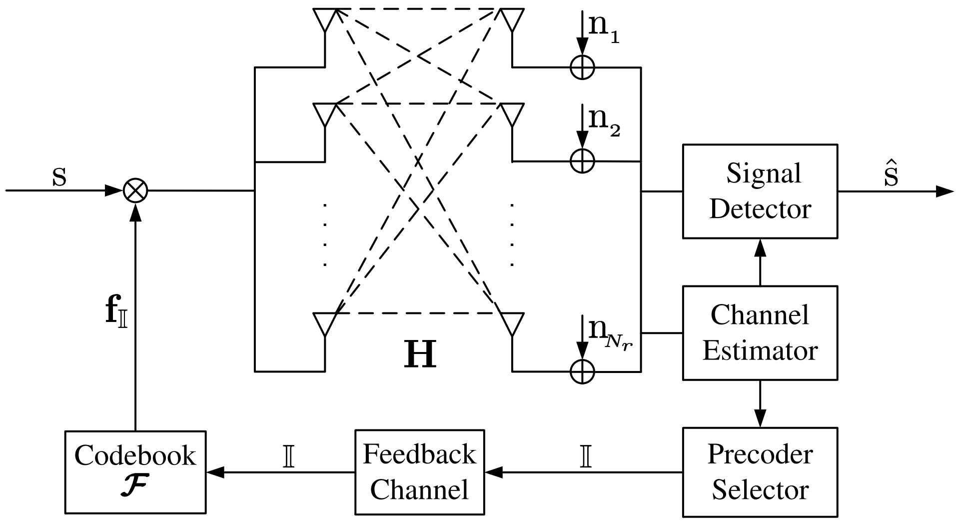

Figure 6.1: The block diagram of a MIMO system with limited feedback transmit beamforming.

is ρ, the beamformer vector should satisfy the power constraint: $|\mathbf{f}^*\mathbf{f}| = 1$ where $|x|$ denotes the absolute value of x.

To obtain the equivalent SISO system model of the transmit beamforming, we combine the received signals by multiplying (6.1) with $\mathbf{w}^*$. The optimum combiner vector $\mathbf{w}$, which maximizes the received SNR for each transmitted symbol, is $\mathbf{w} = \mathbf{Hf}$ [13]. Therefore, the equivalent SISO model of the transmit beamforming is

$$\hat{s} = \sqrt{\rho}\, \|\mathbf{Hf}\|^2\, s + \hat{n} \tag{6.2}$$

where $\hat{n}$ is the additive white noise with $\mathcal{CN}\left(0, \|\mathbf{Hf}\|^2\right)$ distribution. Based on (6.2), the received SNR (γ) for each transmitted symbol is $\gamma = \rho\, \|\mathbf{Hf}\|^2$.

Considering the ordered singular value decomposition (SVD) of the channel matrix as $\mathbf{H} = \mathbf{U\Sigma V}^*$ where $\mathbf{U} \in \mathcal{U}(N_r)$, $\mathbf{V} \in \mathcal{U}(N_t)$, and $\mathbf{\Sigma} \in \mathcal{R}_{\geqslant 0}^{N_r \times N_t}$ is a diagonal matrix with decreasing order, i.e., $\sigma_i \geqslant \sigma_{i+1}$, $i = 1, 2, ..., M$, we have:

$$\begin{aligned} \|\mathbf{Hf}\|^2 &= \|\mathbf{\Sigma V}^*\mathbf{f}\|^2 \\ &= \sigma_1^2\, |\mathbf{v}_1^*\mathbf{f}|^2 + \sum_{i=2}^{m} \sigma_i^2\, |\mathbf{v}_i^*\mathbf{f}|^2 \end{aligned} \tag{6.3}$$

where $M = \min\{N_t, N_r\}$, and $\mathbf{v}_i$, $i = 1, ..., N_t$, is the i-th right eigenvector of the channel matrix corresponding to the i-th largest singular value. The optimal beamformer (which maximizes the received SNR) is $\mathbf{f}_{opt} = \mathbf{v}_1$, and, consequently, $\gamma_{opt} = \rho\, \sigma_1^2$. Optimal beamforming is applicable in very slow fading environments, or when the system uses the time division duplex (TDD) technique for transmission, where the transmitter and receiver use the same bandwidth.

However, in limited feedback systems where $\mathbf{f}_{opt}$ is not available at the transmitter, we use a pre–designed set (codebook) $\boldsymbol{\mathcal{F}}$ of $L = 2^{N_b}$ beamformers (Fig. 6.1), where N_b is the number of feedback bits. For a given $\mathbf{H}$, the only feedback parameter is $\mathbb{I}$, which is the index of $\mathbf{f}_{\mathbb{I}} \in \boldsymbol{\mathcal{F}}$ that maximizes $\|\mathbf{H}\mathbf{f}_k\|^2$, for all $\mathbf{f}_k \in \boldsymbol{\mathcal{F}}$.

The following issues are investigated in this chapter:

- How should the codebook $\boldsymbol{\mathcal{F}}$ be designed?
- How should the appropriate beamformer $\mathbf{f}_{\mathbb{I}}$ be selected from $\boldsymbol{\mathcal{F}}$ for each realization of the channel?
- How well will a given codebook $\boldsymbol{\mathcal{F}}$ with L beamformers perform?

6.2 Codebook Design

In suboptimal beamforming, a vector $\mathbf{f}$ is used instead of the optimal vector $\mathbf{v}_1$ so that the received SNR per symbol is related to $\|\mathbf{H}\mathbf{f}\|^2$. Thus, it is convenient to define the following distortion minimization problem [71] as a figure of merit for codebook $\boldsymbol{\mathcal{F}}$:

$$\mathrm{E}_{\mathbf{H}}\left\{\min_{\mathbf{f}\in\boldsymbol{\mathcal{F}}}\left(\|\mathbf{H}\mathbf{v}_1\|^2 - \|\mathbf{H}\mathbf{f}\|^2\right)\right\}. \tag{6.4}$$

It is shown in [71] that the minimization in (6.4) leads to the maximization of the chordal distance between any pairs of precoders in $\boldsymbol{\mathcal{F}}$. The chordal distance between $\mathbf{f}_i$ and $\mathbf{f}_j$ is defined as

$$\begin{aligned} d_c(\mathbf{f}_i,\mathbf{f}_j) &\triangleq \frac{1}{\sqrt{2}}\|\mathbf{f}_i\mathbf{f}_i^* - \mathbf{f}_j\mathbf{f}_j^*\| \\ &= \sqrt{1-\lambda^2(\mathbf{f}_i^*\mathbf{f}_j)} \\ &= \sqrt{1-|\mathbf{f}_i^*\mathbf{f}_j|^2} \\ &\triangleq \sin(\theta_{ij}) \end{aligned} \tag{6.5}$$

where $0 \leqslant i \neq j < L$, and θ_{ij} denotes the angle between pair vectors $(\mathbf{f}_i,\mathbf{f}_j) \in \boldsymbol{\mathcal{F}}$. The optimum codebook is the one with the maximum θ_{min} defined as

$$\theta_{min} \triangleq \min_{\forall \mathbf{f}_i \neq \mathbf{f}_j \in \boldsymbol{\mathcal{F}}} \arcsin\left(\sqrt{1 - |\mathbf{f}_i^* \mathbf{f}_j|^2}\right) \tag{6.6}$$

$$\leqslant \arcsin\left(\sqrt{\frac{L}{L-1} \times \frac{N_t - 1}{N_t}}\right) \tag{6.7}$$

where the inequality (6.7) is called the Welch bound [128] or Rankin bound [122] for $L \geqslant N_t$.

Finding a codebook (a pack) $\boldsymbol{\mathcal{F}}$ of L vectors in N_t–dimensional complex space with maximum possible θ_{min} (6.6), is called the Grassmannian line packing problem in applied mathematics and information theory [70], [122]. For arbitrary L and N_t, line packings that achieve equality in (6.7) are often impossible to design. The most practical method for generating packings is to use the unitary matrix structure proposed in [24] for non–coherent space–time modulation. Similar to the codebook structure proposed in [24], which can be easily implemented and yields codebooks with large minimum distances, the codebook $\boldsymbol{\mathcal{F}}$ in [70] is constructed as follows:

$$\boldsymbol{\mathcal{F}} = \left\{ \mathbf{f}_k \middle| \, \mathbf{f}_k = \frac{1}{\sqrt{N_t}} \left[\mathrm{e}^{j\frac{2\pi}{L}ku_1}, \mathrm{e}^{j\frac{2\pi}{L}ku_2}, \ldots, \mathrm{e}^{j\frac{2\pi}{L}ku_{N_t}} \right]^T, k = 0, \ldots, L-1 \right\}$$

where $0 \leqslant u_i < L$ are the integer design parameters and should be optimized as follows

$$\mathbf{u} = \arg\max_{\{u_i\}} \min_{1 \leqslant k < L} d_c(\mathbf{f}_0, \mathbf{f}_k)$$

where $\mathbf{u} = \left[u_1, u_2, \ldots, u_{N_t}\right]^T$.

In previous works [24], [71], the design parameters, $\{u_i\}$, are restricted to integers. Exhaustive computer search or random search for their optimum values is employed since analytical determination of the optimum appears impossible. Moreover, because the computational complexity increases exponentially with dimensions (N_t and L), it is impossible to find the optimum parameters for large dimensions with exhaustive search.

To overcome these problems, we propose to employ the genetic algorithm [123]. Although it does not guarantee the global optimality, we find that genetic solutions have larger θ_{min} than the optimum values from exhaustive search for integers $\{u_i\}$. This seemingly contradictory result is obtained by relaxing the design parameters to be real rather than integer numbers; i.e., the codebook $\boldsymbol{\mathcal{F}}$ is constructed as follows:

$$\boldsymbol{\mathcal{F}} = \left\{ \mathbf{f}_k \middle| \, \mathbf{f}_k = \frac{1}{\sqrt{N_t}} \left[\mathrm{e}^{jk\alpha_1}, \mathrm{e}^{jk\alpha_2}, \ldots, \mathrm{e}^{jk\alpha_{N_t}} \right]^T, k = 0, \ldots, L-1 \right\} \tag{6.8}$$

where $0 \leqslant \alpha_i < 2\pi$ are the design parameters and should be optimized as follows:

$$\boldsymbol{\alpha} = \arg \max_{\{\alpha_i\}} \min_{1 \leqslant k < L} d_c(\mathbf{f}_0, \mathbf{f}_k) \tag{6.9}$$

where $\boldsymbol{\alpha} = \left[\alpha_1, \alpha_2, \ldots, \alpha_{N_t}\right]^T$. This relaxation increases the search space of the design parameters, thereby improving the chance to obtain codebooks with larger θ_{min}.

6.2.1 Genetic Algorithm

Genetic algorithms are adaptive heuristic search algorithms based on the evolutionary ideas of natural selection and genetics, where an intelligent exploitation of a random search is used to solve optimization problems. This algorithm exploits historical information to direct the search into the region of better performance within the search space. A genetic algorithm simulates the survival of the fittest among individuals over consecutive generations for solving a problem. In our problem, the fitness of a solution is determined by θ_{min} (6.6). Each generation consists of a population of bit (gene) strings that are analogous to the chromosome of the DNA. Each individual represents a point in the search space and a possible solution. By assuming N_g bits (genes) for each α_i in (6.9), we define a string of $N_g \times N_t$ bits for each solution (individual in the population). The individuals in the population are first generated randomly and then are made to go through a process of evolution. A fitness score (θ_{min}) is assigned to each solution, representing the ability of an individual to compete. Individuals with higher fitness scores are selected as parents for the next generation with higher probability.

In a basic genetic algorithm, the next generation is composed of three types of children as follows:

- *Elite Children*: Children in the current generation are selected for the next generation based on their fitness values. Since the selection rule here is probabilistic not deterministic, fitter solutions (larger θ_{min}) are typically more likely to be selected.
- *Crossover Children*: These children are created by combining pairs of parents in the current population. Generally, the crossover operation recombines selected solutions (parents) by swapping parts of them for producing divergent solutions to explore the search space. Many crossover techniques exist

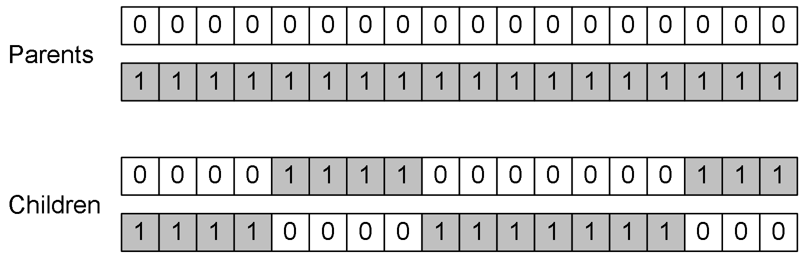

Figure 6.2: Multi–point crossover.

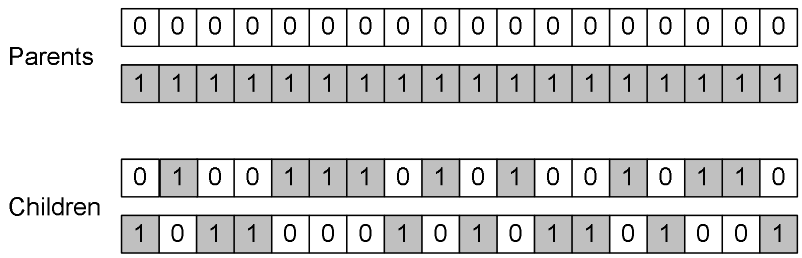

Figure 6.3: Scattered crossover.

to produce a child of a pair parents. However, all of them are surprisingly simple to implement, involving random number generation and some partial string exchange. Fig. 6.2 and Fig. 6.3 illustrate two different techniques used in crossover generation. Scattered crossover is a popular technique used for crossover generation. This method first creates a random binary vector with the same size of parents. Then if the i-th bit is 0, the corresponding gene is selected from the first parent; otherwise, it is selected from the second parent. Ultimately, all selected genes are combined to form the child.

- *Mutation Children*: The algorithm generates mutation children by randomly changing the bits of individual parent in the current population. This process can be done by adding a random vector from a Gaussian distribution to the parent. The aim of mutation in the algorithm is to avoid local optima by preventing the population from becoming too similar to each other, thus slowing or even stopping the evolution.

As a result, new mutated members along with new crossed–over members and the rest of those selected from the previous population form the new generation. The genetic algorithm terminates when there is no improvement in the objective function

Table 6.1: The θ_{min} (in degrees) of beamformers obtained by genetic algorithm and exhaustive search, and their corresponding Welch bounds.

N_t	N_b	L	θ_{min}(exhaustive)	θ_{min}(genetic)	θ_{min}(Welch)
4	3	8	60.00	64.67	81.78
4	4	16	54.43	57.32	78.46
4	5	32	45.00	47.68	76.95
4	6	64	37.12	38.93	76.22
4	7	128	29.92	31.33	75.87
6	3	8	72.06	75.76	87.27
6	4	16	64.18	67.27	83.62
6	5	32	61.87	64.11	81.96
6	6	64	55.92	57.65	81.17
6	7	128	50.75	51.97	80.78

for a specific number of successive iterations. Interested readers are referred to [123] and [129] for sophisticated genetic algorithms and more information about the algorithm such as robustness, marginal and large dimensional behavior, comparisons with and benefits over other optimization techniques.

Table 6.1 shows the θ_{min} obtained by using the exhaustive search and by using the genetic algorithm. For comparison, the Welch bound for θ_{min} is also included. Our simulations show that using $N_g = 8$ bits for each α_i is enough to obtain the results in Table 6.1. The genetic solutions not only have a larger θ_{min} than those from exhaustive search, but also are obtained much faster due to the computational complexity of exhaustive search in large dimensions (N_t and L). Genetic optimization can also be used for other applications such as "precoder design for multiplexing" and "code design for differential unitary space–time modulation" [28].

Although other methods are proposed in the literature for codebook design, particularly based on vector quantization (VQ) and Lloyd algorithm [67], [117]– [119], simulation results show that the codebooks obtained by other methods perform the same as GLP codebooks in the Rayleigh fading channel [120], [121], and that the design complexity of all previous methods is large when the size of the codebook is high. However, codebook design using the structure proposed in (6.8) and by exploiting the genetic algorithm has the following benefits:

- Genetic solutions have larger θ_{min}.

- The genetic algorithm reduces the design complexity effectively, especially in large dimensions.

- The only parameters that should be saved at the transmitter and receiver are $\{\alpha_1, \alpha_2, ..., \alpha_{N_t}\}$. Therefore, this method is easy to implement, while in other methods, the whole codebook should be saved at the transmitter and receiver.
- The structure proposed in (6.8) allows us to propose an algorithm that works faster than the exhaustive search to reduce the beamformer selection complexity at the receiver side. This algorithm is presented in the next section.

6.3 Beamformer Selection

For every realization of the channel matrix $\mathbf{H}$, the best beamformer $\mathbf{f}_{\mathbb{I}} \in \boldsymbol{\mathcal{F}}$ is selected, and only the index $\mathbb{I}$ is fed back to the transmitter for beamforming the transmit signals by $\mathbf{f}_{\mathbb{I}}$. Since the received SNR for each symbol is $\gamma = \rho \|\mathbf{H}\mathbf{f}_{\mathbb{I}}\|^2$, $\mathbb{I}$ should be selected from the following optimization problem:

$$\mathbb{I} = \arg \max_{0 \leqslant k < L} \|\mathbf{H}\mathbf{f}_k\|^2 \tag{6.10}$$

where $\mathbf{f}_k$ is defined in (6.8).

The maximization problem in (6.10) is simply an exhaustive search over all members of $\boldsymbol{\mathcal{F}}$, which can be computationally complex for large dimensions (L, N_t and also N_r). Moreover, the beamformers in $\boldsymbol{\mathcal{F}}$ should be stored at the receiver and transmitter, which needs a dedicated memory, specially for large dimensions. Therefore, it is expedient to find an intelligent algorithm to solve (6.10) efficiently, with reasonable memory, particularly when the structure of the beamformers is known (6.8).

According to the Rayleigh–Ritz inequality [48],

$$\|\mathbf{H}\mathbf{a}\|^2 \leqslant \sigma^2_{max}(\mathbf{H}) \quad \Longrightarrow \quad \sigma^2_{max}(\mathbf{H}) - \|\mathbf{H}\mathbf{a}\|^2 \geqslant 0$$

when $\|\mathbf{a}\|^2 = 1$. Thus, the maximization problem (6.10) can be changed to the following minimization problem:

$$\begin{aligned}
\mathbb{I} &= \arg \min_{0 \leqslant k < L} \sigma^2_{max}(\mathbf{H}) - \|\mathbf{H}\mathbf{f}_k\|^2 \\
&= \arg \min_{0 \leqslant k < L} \sigma^2_{max}(\mathbf{H}) \mathbf{f}_k^* \mathbf{I}_{N_t}^* \mathbf{I}_{N_t} \mathbf{f}_k - \mathbf{f}_k^* \mathbf{H}^* \mathbf{H} \mathbf{f}_k \\
&= \arg \min_{0 \leqslant k < L} \mathbf{f}_k^* \mathbf{G} \mathbf{f}_k
\end{aligned}$$

where $\mathbf{G} = \lambda_{max}(\mathbf{H}^*\mathbf{H})\mathbf{I}_{N_t} - \mathbf{H}^*\mathbf{H}$.

By Cholesky decomposition of $\mathbf{G}$ as $\mathbf{G} = \mathbf{R}^*\mathbf{R}$ where $\mathbf{R}$ is an upper triangular matrix, we have

$$\mathbb{I} = \arg \min_{0 \leqslant k < L} \|\mathbf{R}\mathbf{f}_k\|^2 . \tag{6.11}$$

Due to the upper triangularity of $\mathbf{R}$ and by expanding $\|\mathbf{R}\mathbf{f}_k\|^2$ to its scalar form

$$\|\mathbf{R}\mathbf{f}_k\|^2 = \sum_{q=N_t}^{1} \left| \sum_{t=q}^{N_t} r_{q,t}\, f_{k,t} \right|^2 , \tag{6.12}$$

it can be seen that (6.12) consists of an outer sum of non–negative real terms, where the computational load of each term is increased when the index of the outer sum (q) is decreased. Thus, if we know that $\|\mathbf{R}\mathbf{f}_{\mathbb{I}}\|^2 \leqslant \mathcal{B}$ where $\mathcal{B}$ is a bound, we can compare the outer sum value in (6.12), index by index, to the bound $\mathcal{B}$, and if it is greater than the bound, the rest of the computations (for the given $\mathbf{f}_k$) are discarded. By this bounding technique, the computational complexity of (6.11) is reduced efficiently.

Clearly, the bound $\mathcal{B}$ plays a critical role in the complexity reduction. Initially, $\mathcal{B}$ is set to infinity for the first $\mathbf{f}_k$; i.e., $\mathbf{f}_0$, but for the rest of $\mathbf{f}_k$'s, $k = 1, 2, \ldots, L-1$, $\mathcal{B}$ is set to the minimum $\|\mathbf{R}\mathbf{f}_k\|^2$ obtained thus far during the algorithm. Consequently, it is expedient to run the proposed algorithm in a rational ordering for k (not simply k from 0 to $L-1$) so that the probability of obtaining as small $\mathcal{B}$ as possible in the primary k's is as high as possible.

We resort to the geometry of N_t–dimensional vectors in $\boldsymbol{\mathcal{F}}$. We define $\mathbf{f}_{k_0}$ as the reference vector (for any arbitrary $0 \leqslant k_0 < L$), and the reference set $\Theta = \{\theta_0, \theta_1, \ldots, \theta_{L-1}\}$ where

$$\theta_i \triangleq \arcsin\left(\sqrt{1 - |\mathbf{f}_{k_0}^* \mathbf{f}_i|^2}\right)$$

is the angle between the pair of vectors $(\mathbf{f}_{k_0}, \mathbf{f}_i)$. For a realization of the channel matrix $\mathbf{H}$, we compute

$$\theta_{\mathbf{H}} \triangleq \arcsin\left(\sqrt{1 - |\mathbf{f}_{k_0}^* \mathbf{v}_1|^2}\right)$$

where $\mathbf{v}_1$ is the right eigenvector corresponding to the largest singular value of $\mathbf{H}$. Then, an ordered set $\mathcal{K}$ of the angle indexes in Θ is constructed based on their vicinity to $\theta_{\mathbf{H}}$, or, equivalently,

$$\mathcal{K} = \text{Index}\,\{\, \text{Sort}\,\{\, |\, \Theta - \theta_{\mathbf{H}} | \,\}\,\} \tag{6.13}$$

where Index$\{|\theta_k - \theta_{\mathbf{H}}|\} = k$, and the operator Sort$\{\cdot\}$ sorts its argument set from minimum to maximum. Thus, the search for the minimum in (6.11) is re–ordered as follows:

$$\mathbb{I} = \arg\min_{k \in \mathcal{K}} \|\mathbf{R}\mathbf{f}_k\|^2. \tag{6.14}$$

Therefore, for a given codebook, the angle set Θ is calculated and stored at the receiver. For a given channel matrix $\mathbf{H}$, $\theta_{\mathbf{H}}$ is computed, and the ordered set $\mathcal{K}$ (6.13) is constructed. Then, the minimization in (6.14) is executed with respect to the structure of $\|\mathbf{R}\mathbf{f}_k\|^2$ (6.12) and the bound $\mathcal{B}$. This algorithm is referred to as the order and bound algorithm. Algorithm 1 presents the semi–code of the algorithm where $\mathbf{f}_0$ is adopted as the reference vector.

Example 6.1 : In this example, we verify the computational performance of the order and bound algorithm proposed for beamformer selection at the receiver side. Figure 6.4 shows the average flops (floating point operations) of beamformer selection with exhaustive search and with the order and bound algorithm. By exploiting the proposed ordering and bounding method, the beamformer selection complexity is reduced by 68% for $N_b = 7$, and 80% for $N_b = 12$, in a MIMO system with $N_t = 4$ transmit antennas and $N_r = 5$ receive antennas. The flops of the order and bound algorithm include the flops consumed by the singular value decomposition and Cholesky decomposition. Since the complexity order of both decompositions is $\mathcal{O}(N_t^3)$, for small–size codebooks, specifically $N_b = 3$ in Figure 6.4, the order and bound algorithm is more complex than the exhaustive search.

The complexity of exhaustive search depends mainly on the number of receive antennas, N_r, and the codebook size, L, while the complexity of the order and bound algorithm depends mainly on the number of transmit antennas, N_t, and codebook size, L, and is almost independent of the number of receive antennas, N_r. Therefore, due to the dependency of the order and bound algorithm to SVD and Cholesky decomposition with $\mathcal{O}(N_t^3)$ order of complexity, when the number of transmit antennas is larger than the number of receive antennas, the complexity of the proposed algorithm exceeds the complexity of exhaustive search (Figure 6.5). Clearly, this issue depends on the codebook size since the sensitivity of exhaustive search to the codebook size is significant.

$\Diamond$

Algorithm 1 : The order and bound algorithm

```
Data   : H, Θ, f_0
Result: 𝕀

[U, Σ, V] ← svd(H);
λ_max ← max{ diag(Σ) }^2 ;
v_1 ← V(:, 1);
θ_H ← arcsin( sqrt(1 − | f_0^* v_1 |^2) ) ;
𝒦 ← Index{Sort{ |Θ − θ_H| }} ;
R ← chol(λ_max I_{N_t} − H^* H);
L ← size(Θ);
N ← size(R);
ℬ ← ∞;
d_0 = | R(N, N) |^2;
for i = 1 : L , do
    k ← 𝒦(i);
    f_k ← f_0^k ;
    d ← d_0 ;
    for q = N − 1 : −1 : 1 , do
        d ← d + | R(q, q : N) f_k(q : N) |^2 ;
        if d > ℬ then
            break;
        else
            if q == 1 then
                ℬ ← d;
                𝕀 ← k;
            end
        end
    end
end
```

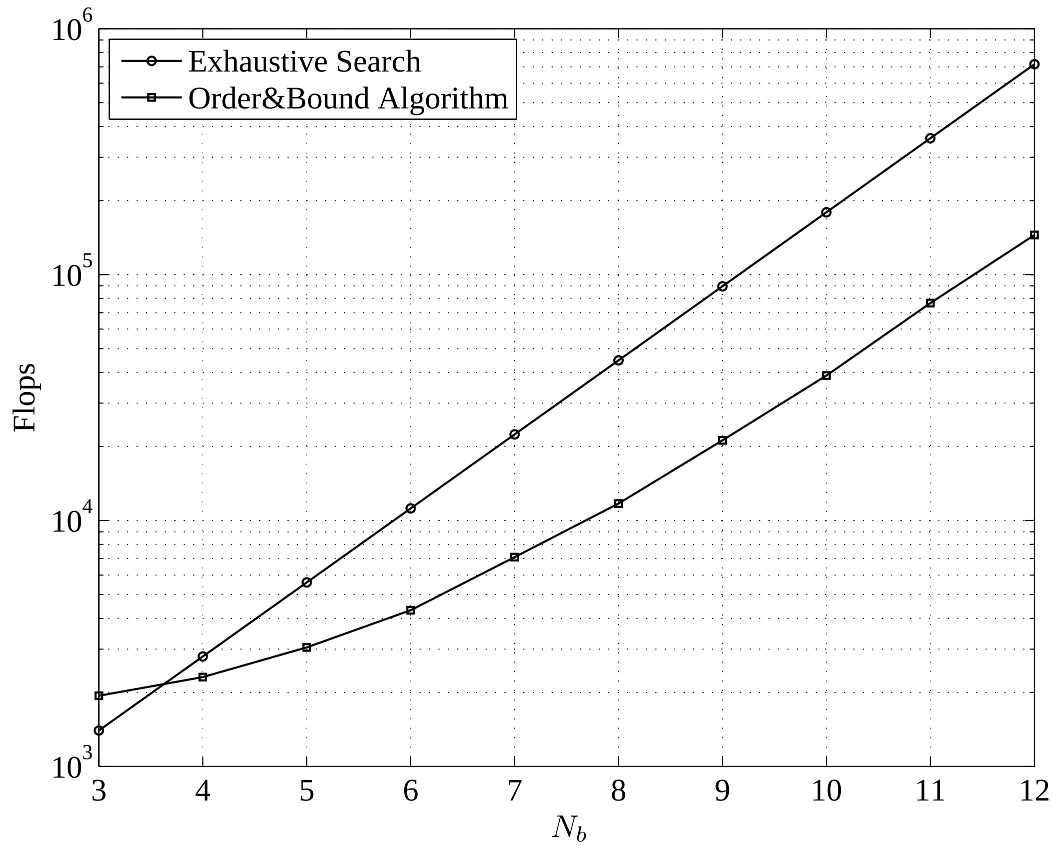

Figure 6.4: Average flops of beamformer selection with the exhaustive search and the order and bound algorithm for MIMO(4,5) versus codebook size.

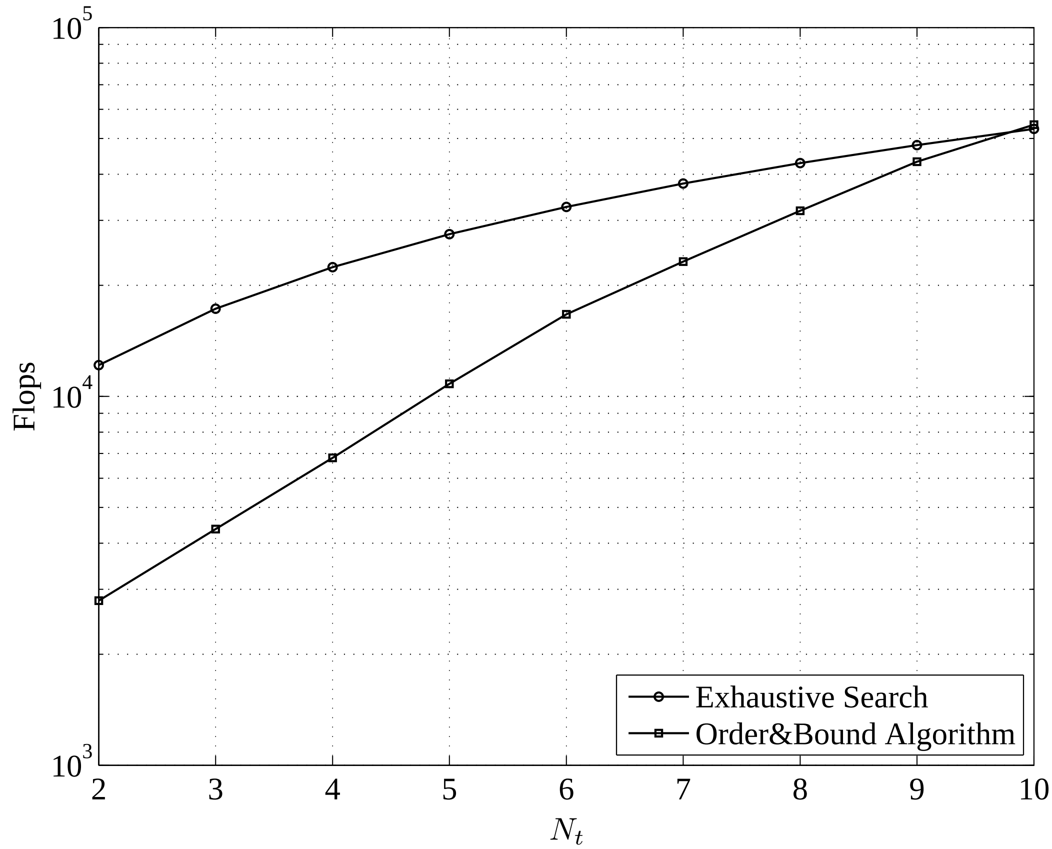

Figure 6.5: Average flops of beamformer selection with the exhaustive search and the order and bound algorithm for MIMO system with $N_r = 5$ receive antennas and $N_b = 7$ bits of feedback versus the number of transmit antennas.

6.4 Performance Analysis of Optimal Transmit Beamforming

In order to perform the BER analysis of limited feedback beamforming, we first present the exact BER of optimal transmit beamforming. Although the performance of optimal transmit beamforming has been analyzed in the literature, the expressions are infinite series and, therefore, are not suitable for our BER analysis of the limited feedback transmit beamforming presented in section 6.5.

Since in optimal beamforming, the right eigenvector corresponding to the largest singular value of the channel matrix ($\mathbf{v}_1$) is used for beamforming, by substituting (6.3) into (6.2), the system model for optimal beamforming becomes

$$\hat{s} = \sqrt{\rho}\,\sigma_1^2\,s + \hat{n} \tag{6.15}$$

where $\hat{n}$ is the additive white noise sample with $\mathcal{CN}\left(0, \sigma_1^2\right)$ distribution.

6.4.1 Exact BER Expression for PAM and QAM

Assume the transmitted signal s in (6.15) is selected from a $I \times J$ rectangular QAM constellation with unit average energy and a Gray code mapping [130]. I and J denote the number of in-phase and quadrature amplitudes, respectively. We define

$$P_{I|\sigma_1^2}(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i)\,Q\left(\sqrt{\eta(i)\sigma_1^2}\right) \tag{6.16}$$

where

$$\beta_I(k,i) = \frac{2\,(-1)^{\left\lfloor \frac{i\,2^{k-1}}{I} \right\rfloor}}{I}\left(2^{k-1} - \left\lfloor \frac{i\,2^{k-1}}{I} + \frac{1}{2} \right\rfloor\right), \tag{6.17}$$

$$\eta(i) = \frac{6\,(2i+1)^2\rho}{I^2 + J^2 - 2}, \tag{6.18}$$

$Q(x)$ is the Q–function defined in (1.5), and $\lfloor x \rfloor$ denotes the largest integer to x. Now the average BER of the $I \times J$ rectangular QAM conditioned on σ_1^2 is expressed as [131]

$$P_{b|\sigma_1^2} = \frac{1}{\log_2(I \times J)}\left[\sum_{k=1}^{\log_2 I} P_{I|\sigma_1^2}(k) + \sum_{l=1}^{\log_2 J} P_{J|\sigma_1^2}(l)\right]. \tag{6.19}$$

Note that (6.19) reduces to the BER of BPSK for $I = 2$ and $J = 1$, I–array PAM for $J = 1$, and M–array square QAM for $I = J = \sqrt{M}$. Thus, the exact BER is obtained by averaging (6.19) over the distribution of σ_1^2. Interested readers can refer to [131] for more details on the derivation of (6.16) and (6.19).

6.4.2 Largest Eigenvalue Distribution of i.i.d. Central Wishart

The singular values of $\mathbf{H}$ correspond to the nonzero eigenvalues of the matrix $\mathbf{HH}^*$ by $\lambda_i = \sigma_i^2$, $i = 1, \ldots, M$. Therefore, $\lambda_1 = \lambda_{max}(\mathbf{HH}^*)$ and σ_1^2 have the same distribution. When the elements of $\mathbf{H}$ are i.i.d. $\mathcal{CN}(0,1)$ random variables, the matrix $\mathbf{HH}^*$ becomes a central Wishart matrix, whose joint ordered eigenvalue distribution can be simply obtained from (3.31) as

$$P(\boldsymbol{\lambda}) = \mathcal{K}_{N,M} \underset{M}{\Delta}(\boldsymbol{\lambda})^2 \prod_{i=1}^{M} e^{-\lambda_i} \lambda_i^{N-M} \tag{6.20}$$

where $\mathcal{K}_{N,M}$ is defined in (3.28), and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_M$. By integrating (6.20) over $\lambda_2, \ldots, \lambda_M$, the pdf of λ_1 is expressed as

$$\begin{aligned} f_{\lambda_1}(\lambda_1) = \mathcal{K}_{N,M}\, \mathrm{e}^{-\lambda_1} \lambda_1^{N-M} \int_0^{\lambda_1} d\lambda_2\, \mathrm{e}^{-\lambda_2} \lambda_2^{N-M} (\lambda_1 - \lambda_2)^2 \int_0^{\lambda_2} d\lambda_3 \ldots \\ \times \int_0^{\lambda_{M-1}} d\lambda_M\, \mathrm{e}^{-\lambda_M} \lambda_M^{N-M} \prod_{i=1}^{M-1} (\lambda_i - \lambda_M)^2 . \end{aligned} \tag{6.21}$$

The calculation of (6.21) is a complicated process for general N and M, and although it is available in the form of a hypergeometric function with matrix arguments (e.g., [132]), this form is not computationally useful, especially for the BER analysis of the limited feedback transmit beamforming presented in section 6.5. However, in MIMO systems where N and M are relatively small numbers, (6.21) can be easily calculated.

Example 6.2 : Assuming $M = 1$, the pdf of λ_1 is obtained from (6.21) as

$$f_{\lambda_1}(\lambda_1) = \frac{1}{\Gamma(N)}\, \mathrm{e}^{-\lambda_1} \lambda_1^{N-1}$$

which is a central Chi–square distribution with N degrees of freedom.

◊

Example 6.3 : For $M = 2$, from (6.21), we have

$$f_{\lambda_1}(\lambda_1) = \frac{1}{\Gamma(N)\Gamma(N-1)}\, \mathrm{e}^{-\lambda_1} \lambda_1^{N-2} \int_0^{\lambda_1} d\lambda_2\, \mathrm{e}^{-\lambda_2} \lambda_2^{N-2} (\lambda_1^2 - 2\lambda_1\lambda_2 + \lambda_2^2) .$$

Since

$$\int_0^x dt\, \frac{1}{\Gamma(N)}\, t^{N-1}\, \mathrm{e}^{-t} = 1 - \mathrm{e}^{-x} \sum_{i=0}^{N-1} \frac{x^i}{i!} ,$$

and by some manipulations, we obtain

$$
\begin{aligned}
f_{\lambda_1}(\lambda_1) = \mathrm{e}^{-\lambda_1} &\left\{ \frac{\lambda_1^{N-2}}{\Gamma(N)} \Big[N(N-1) - 2(N-1)\lambda_1 + \lambda_1^2 \Big] \right\} \\
+ \mathrm{e}^{-2\lambda_1} &\Bigg\{ - \frac{\lambda_1^{N-2}}{\Gamma(N)} \Big[N(N-1) + (N-1)(N-2)\lambda_1 \\
&\qquad + \sum_{n=2}^{N} \left(\frac{N(N-1)}{n(n-1)} - 2\frac{(N-1)}{(n-1)} + 1 \right) \frac{\lambda_1^n}{n!} \Big] \Bigg\}.
\end{aligned}
$$

$\Diamond$

In general, it is easy to verify [82] that the pdf of λ_1 can be represented by

$$
f_{\lambda_1}(\lambda_1) = \sum_{m=1}^{M} \mathrm{e}^{-m\lambda_1} G_m(\lambda_1) \tag{6.22}
$$

where

$$
G_m(\lambda_1) = \sum_{j=0}^{D_m} a_{m,j} \lambda_1^j \tag{6.23}
$$

denotes the corresponding polynomial coefficient of $e^{-m\lambda_1}$, D_m is the degree of $G_m(\lambda_1)$, and the coefficients $\{a_{m,j}\}$ can be tabulated by simple integrations.

Example 6.4 : For $M = 3$ and $N = 4$, we have

$$
\begin{aligned}
G_1(\lambda_1) &= 6\lambda_1 - 8\lambda_1^2 + \frac{9}{2}\lambda_1^3 - \lambda_1^4 + \frac{1}{12}\lambda_1^5 , \\
G_2(\lambda_1) &= -12\lambda_1 + 4\lambda_1^2 + \lambda_1^3 - \lambda_1^4 - \frac{1}{12}\lambda_1^5 - \frac{1}{12}\lambda_1^6 , \\
G_3(\lambda_1) &= 6\lambda_1 + 4\lambda_1^2 + \frac{1}{2}\lambda_1^3 .
\end{aligned}
$$

$\Diamond$

6.4.3 BER of Optimal Transmit Beamforming

From (6.16) and (6.19), it is clear that to find $P_b = \mathrm{E}\{P_{b|\lambda_1}\}$, we need to calculate $P_I(k) = \mathrm{E}\{P_{I|\lambda_1}(k)\}$ or, equivalently,

$$
P_I(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i)\, \mathrm{E}\Big\{ Q\Big(\sqrt{\eta(i)\lambda_1}\Big) \Big\} .
$$

Therefore, by considering the distribution of λ_1 (6.22) and (6.23), we obtain

$$P_I(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i) \sum_{m=1}^{M} \sum_{j=0}^{D_m} a_{m,j} \int_0^\infty d\lambda\, \lambda_1^j\, \mathrm{e}^{-m\lambda_1}\, Q\Big(\sqrt{\eta(i)\lambda_1}\Big)\ . \tag{6.24}$$

It has been shown [130] that

$$\int_0^\infty dx\, m^L\, x^{L-1}\, \mathrm{e}^{-mx}\, Q\big(\sqrt{\alpha x}\big) = \phi\big(L, m, \alpha\big)$$

where

$$\phi\big(L, m, \alpha\big) = \Gamma(L) \left[\frac{1-\mu}{2}\right]^L \sum_{r=0}^{L-1} \binom{L-1+r}{r} \left[\frac{1+\mu}{2}\right]^r \tag{6.25}$$

and

$$\mu = \sqrt{\frac{\alpha}{2m+\alpha}}\ .$$

Hence, from (6.24) and (6.25), we achieve

$$P_I(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i) \sum_{m=1}^{M} \sum_{j=0}^{D_m} \frac{a_{m,j}}{m^{j+1}}\ \phi\Big(j+1, m, \eta(i)\Big). \tag{6.26}$$

Consequently, the average BER of the $I \times J$ rectangular QAM signal transmitted through the system model in (6.15) is expressed as

$$P_b = \frac{1}{\log_2(I \times J)} \left[\sum_{k=1}^{\log_2 I} P_I(k) + \sum_{l=1}^{\log_2 J} P_J(l)\right]. \tag{6.27}$$

Example 6.5 : Figure 6.6 shows the BER performance of the optimal beamforming for MIMO systems with $N_t = 4$ transmit antennas and $N_r = 1$, 2 and 3 receive antennas. The transmit signal is selected from a 16–QAM constellation. In this figure, the solid lines are the results from (6.27), and the symbols are from computer simulations[1], which verify our analytical results.

$\Diamond$

6.5 Performance Analysis of Limited Feedback Transmit Beamforming

The performance analysis of limited feedback beamforming has been studied in [126] for MISO systems with transmit antenna correlations. The analysis involves an $L-$tuple integration with infinite limits, which is computationally difficult even for

[1] Semi–analytic simulations.

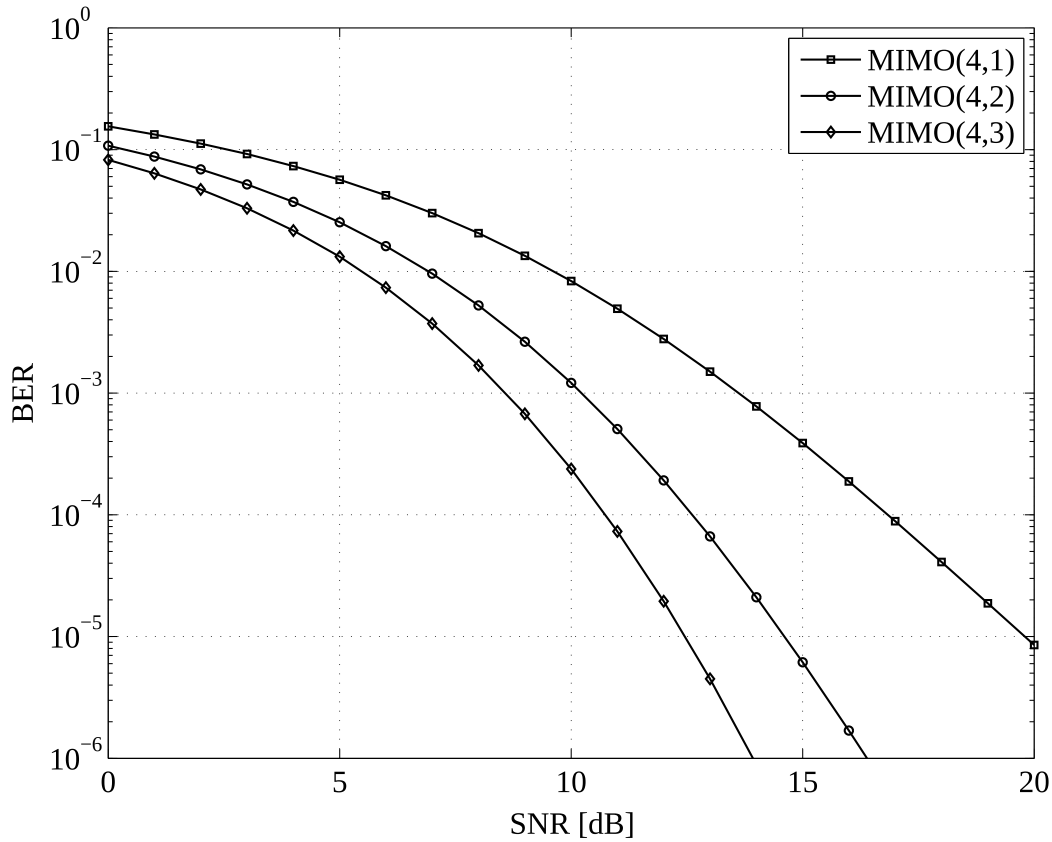

Figure 6.6: BER of optimal transmit beamforming.

a small L (codebook size). On the other hand, a geometric approach has been proposed in [68] for the outage probability analysis, and used in [120] for a lower bound on the symbol error rate of transmit beamforming, both for MISO systems. By using high–resolution quantization theory, the capacity loss in a MISO system with limited feedback has been derived in [127]. Similarly, the SNR loss and outage capacity of MIMO systems with limited feedback beamforming have been approximated in [121]. In this section, however, without distortion analysis, we analyze the BER of limited feedback transmit beamforming for general MIMO systems.

Assume that codebook $\boldsymbol{\mathcal{F}}$ with L beamformers is obtained by maximizing the minimum distance (6.6), where each vector in $\boldsymbol{\mathcal{F}}$ represents a point on the N_t–dimensional complex unit hypersphere. For a large–size codebook, the vector points are uniformly distributed over the surface of the unit hypersphere. On the other hand, the optimal beamformer $\mathbf{v}_1$ uniformly rotates on the surface of the unit hypersphere when the channel gains are i.i.d. distributed. Therefore, for a large–size codebook, we have

$$\lim_{L\to\infty} |\mathbf{v}_1^* \mathbf{f}_{\mathbb{I}}| = 1 \quad , \quad \lim_{L\to\infty} |\mathbf{v}_i^* \mathbf{f}_{\mathbb{I}}| = 0 \, , \; i = 2, \ldots, N_t \tag{6.28}$$

where $\mathbf{f}_{\mathbb{I}}$ is selected from (6.10), and $\mathbf{v}_i$ is the right eigenvector corresponding to the i-th largest singular value of the channel. Considering (6.28) and that $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_M$, we assume the following approximation to simplify the analysis:

$$\sum_{i=2}^{M} \sigma_i^2 \; |\mathbf{v}_i^* \mathbf{f}_{\mathbb{I}}|^2 \approx 0 \tag{6.29}$$

With this approximation, from (6.2) and (6.3) the equivalent channel model for limited feedback transmit beamforming will be

$$\hat{s} = \sqrt{\rho}\, \lambda_1 (1 - X)\, s + \hat{n} \tag{6.30}$$

where

$$X \triangleq 1 - |\mathbf{v}_1^* \mathbf{f}_{\mathbb{I}}|^2 = d_c^2(\mathbf{v}_1, \mathbf{f}_{\mathbb{I}}) \, , \tag{6.31}$$

$\hat{n} \in \mathcal{CN}(0, \lambda_1(1 - X))$, and $\lambda_1 = \lambda_{max}(\mathbf{HH}^*)$. Assuming the transmitted signal s in (6.30) is selected from a $I \times J$ rectangular QAM constellation with unit average energy and a Gray code mapping, the average BER of the system conditioned on λ_1 and X is expressed as

$$P_{b|\lambda_1, X} = \frac{1}{\log_2(I \times J)} \left[\sum_{k=1}^{\log_2 I} P_{I|\lambda_1, X}(k) + \sum_{l=1}^{\log_2 J} P_{J|\lambda_1, X}(l) \right] \tag{6.32}$$

where

$$P_{I|\lambda_1,\,X}(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i)\, Q\Big(\sqrt{\eta(i)\lambda_1(1-X)}\Big)\,, \tag{6.33}$$

and $\beta_I(k,i)$ and $\eta(i)$ are defined in (6.17) and (6.18), respectively. Since the eigenvalues and eigenvectors of the central Wishart matrix are independent [82], λ_1 and X are independent for the i.i.d. channel matrix. Thus, we first average $P_{I|\lambda_1,\,X}(k)$ over λ_1 by using the distribution of λ_1 presented in (6.22) to obtain

$$P_{I|X}(k) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i) \sum_{m=1}^{M}\sum_{j=0}^{D_m} \frac{a_{m,j}}{m^{j+1}}\ \phi\Big(j+1, m, (1-X)\eta(i)\Big) \tag{6.34}$$

where $\phi\Big(j+1, m, (1-X)\eta(i)\Big)$, defined in (6.25), should be averaged over the distribution of X.

To obtain the distribution of X, we use the geometrical method presented in [68]. For each vector $\mathbf{f}_k \in \boldsymbol{\mathcal{F}}$, a spherical cap is defined on the surface of the hypersphere $S_k(x) = \big\{\mathbf{v}_1 \mid d_c^2(\mathbf{v}_1, \mathbf{f}_k) \leqslant x\big\}$ where $0 \leqslant x \leqslant 1$. By defining $A\{S_k(x)\}$ as the area of the cap $S_k(x)$, it has been shown [68] that

$$A\{S_k(x)\} = \frac{2\pi^{N_t} x^{N_t-1}}{(N_t-1)!}\,. \tag{6.35}$$

Equation (6.35) shows that the surface of the unit hypersphere grows exponentially with N_t. Therefore, when N_t is increased, L should be increased accordingly so that the limits in (6.28) and the approximation in (6.29) hold tightly.

According to the definition of X in (6.31), we have

$$F_X(x) = \Pr\bigg\{\, \big[d_c^2(\mathbf{v}_1, \mathbf{f}_0) \leqslant x\big] \text{ or } \big[d_c^2(\mathbf{v}_1, \mathbf{f}_1) \leqslant x\big] \text{ or } \cdots \text{ or } \big[d_c^2(\mathbf{v}_1, \mathbf{f}_{L-1}) \leqslant x\big] \,\bigg\}$$

where $F_X(x)$ denotes the cdf of X. When the channel matrix entries are i.i.d., the optimal vector $\mathbf{v}_1$ is uniformly distributed on the surface of the unit hypersphere. Therefore, we have

$$F_X(x) = \frac{A\left\{\cup_{k=0}^{L-1} S_k(x)\right\}}{A\left\{S_k(1)\right\}}$$

where $\cup_{k=0}^{L-1} S_k(x)$ denotes the union of the regions $\{S_k(x)\}$, $k = 0, \ldots, L-1$. Using (6.35) and the fact that

$$A\left\{\cup_{k=0}^{L-1} S_k(x)\right\} \leqslant \sum_{k=0}^{L-1} A\left\{S_k(x)\right\},$$

we obtain

$$F_X(x) \leqslant \frac{\sum_{k=0}^{L-1} A\left\{S_k(x)\right\}}{A\left\{S_k(1)\right\}} = L\, x^{N_t-1}. \tag{6.36}$$

Finally, by taking into account that $F_X(x) \leqslant 1$, the following approximate cdf and pdf for X can be defined:

$$F_X(x) \approx \begin{cases} L\, x^{N_t-1}\,, & 0 \leqslant x \leqslant X_0 \\ 1\,, & x > X_0 \end{cases} \tag{6.37}$$

$$f_X(x) = L\,(N_t-1)\, x^{N_t-2}\,, \quad 0 \leqslant x \leqslant X_0 \tag{6.38}$$

where $X_0 = \left(\frac{1}{L}\right)^{\frac{1}{N_t-1}}$.

Now, by using the pdf of X (6.38), we can calculate the average of $P_{I|X}(k)$ over the distribution of X. By defining

$$\phi_L\Big(N, m, \alpha\Big) = \Gamma(N) \sum_{r=0}^{N-1} \binom{N-1+r}{r} \frac{L(N_t-1)}{2^{N+r}} \times \int_0^{X_0} dx\, x^{N_t-2} \Big[1+\mu(x)\Big]^r \Big[1-\mu(x)\Big]^N \tag{6.39}$$

where

$$\mu(x) = \sqrt{\frac{\alpha(1-x)}{2m+\alpha(1-x)}}\,,$$

we conclude that

$$P_I(k, L) = \sum_{i=0}^{(1-2^{-k})I-1} \beta_I(k,i) \sum_{m=1}^{M} \sum_{j=0}^{D_m} \frac{a_{m,j}}{m^{j+1}}\ \phi_L\Big(j+1, m, \eta(i)\Big), \tag{6.40}$$

and consequently, the approximate BER of a MIMO system with L beamformers is expressed as

$$P_b(L) \approx \frac{1}{\log_2(I \times J)} \left[\sum_{k=1}^{\log_2 I} P_I(k, L) + \sum_{l=1}^{\log_2 J} P_J(l, L) \right]. \tag{6.41}$$

The accuracy of this approximate BER is verified in the following example.

Example 6.6 : The approximate BER of MIMO(4,2) and MIMO(4,3) systems with $L = 8, 32$ and 128 beamformers is illustrated for 4–QAM and 16–QAM constellations in Figures 6.7 and 6.8, respectively. In these figures, SNR represents ρ in system model (6.1). For comparison, the exact simulated[2] BER and the BER of

[2] Semi–analytic simulations.

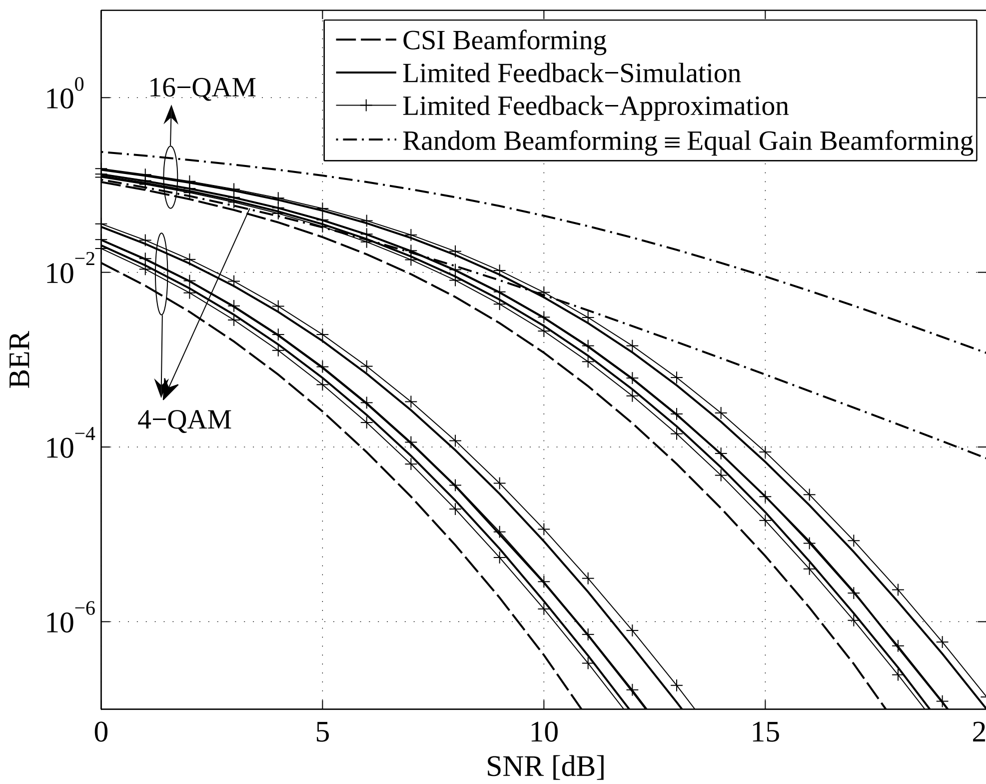

Figure 6.7: BER of MIMO(4,2) system obtained from simulation and closed-form approximation for 4–QAM and 16–QAM signalings. Pair curves from left to right are for $N_b = 7, 5$ and 3 feedback bits, respectively.

the optimal transmit beamforming are included. GLP codebooks have been used for the simulations. Although our approximate analysis was for large–size codebooks, the simulation results show that the approximate BER expression in (6.41) is satisfactorily tight, even for small–size codebooks.

For small–size codebooks, the approximation made in (6.29) is not tight, and, consequently, the effective SNR used in the equivalent system model (6.30) is less than the actual SNR used by the system. Therefore, we expect the approximate BER curves for small L to be an upper bound for the curves obtained by simulations. On the other hand, for large–size codebooks, the approximation in (6.29) is tight enough, but since we used an upper bound approximation for $F_X(x)$, going from (6.36) to (6.37), the effective SNR used in the equivalent system model (6.30) is larger than the actual SNR used by the system. Therefore, we expect the approximate BER curves for large L to be a lower bound for the curves obtained by simulations. These

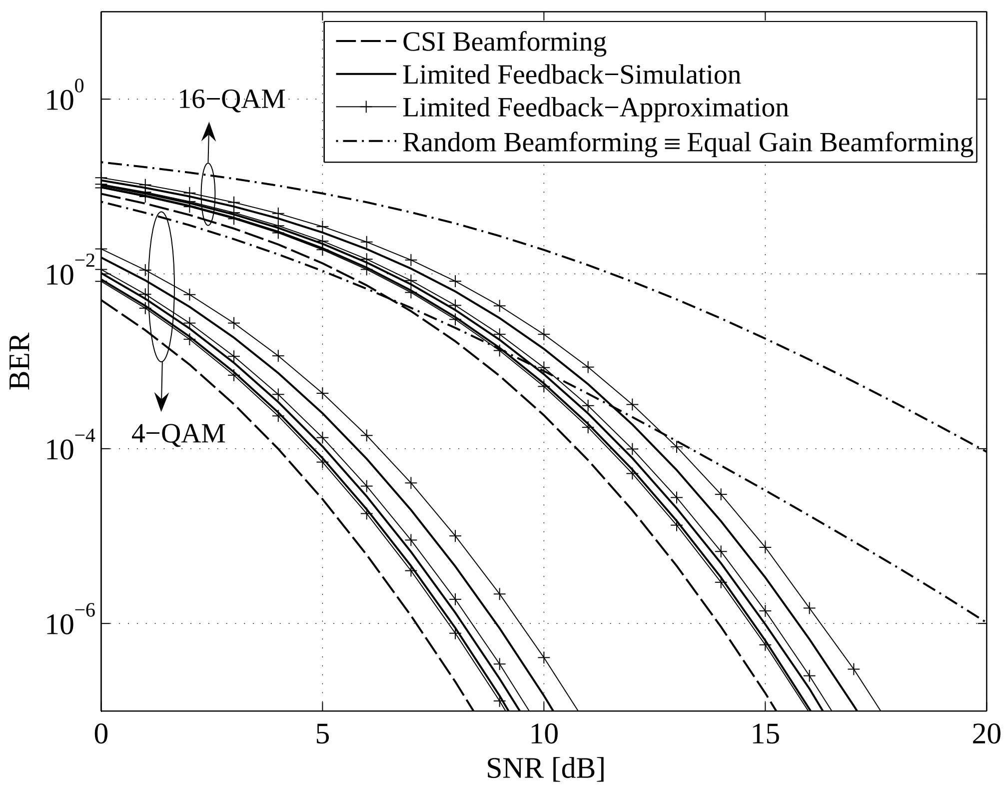

Figure 6.8: BER of MIMO(4,3) system obtained from simulation and closed-form approximation for 4–QAM and 16–QAM signalings. Pair curves from left to right are for $N_b = 7, 5$ and 3 feedback bits, respectively.

behaviors are clearly observable in Figures 6.7 and 6.8.

$\diamond$

6.6 Summary

In this chapter, the genetic GLP beamformer codebooks were developed. The design parameters were relaxed from positive integers to positive real values (angles). The examples showed that the genetic GLP codebooks achieve a larger minimum distance than those of [70], and reduce the optimization complexity. Although this approach was applied for single user scenario, it can be applied for multiuser case as well. In single user case, the aim is to maximize the received SNR, which leads to the maximization of minimum chordal distance of the codebook. However, in multiuser case, the transmitter should select the beamformer that maximizes the throughput of whole system (not just a single user). As a result, the cost function that should be defined for genetic algorithm is changed.

By exploiting the specific structure of the GLP codebooks, the order and bound algorithm was proposed to reduce the beamformer selection complexity at the receiver. This algorithm used bounding techniques to avoid unnecessary computations.

By employing the joint eigenvalue distribution of the i.i.d. central Wishart, the distribution of the largest eigenvalue was derived and used to obtain the exact closed–form expression of the BER performance for the optimal beamforming in finite summations for PAM and QAM constellations. The resulting expression was used to simplify the BER analysis of limited feedback transmit beamforming. By assuming a large–size codebook (high–resolution analysis) and employing a geometrical approach, an approximate BER performance for limited feedback beamforming was derived. The examples showed that the approximate BER is satisfactorily tight, even for small–size codebooks.

Chapter 7

Conclusions and Future Work

For the exact capacity and performance analysis of MIMO wireless systems, the eigenvalue distributions of Wishart random matrices are required for different channel models. Deriving the joint eigenvalue distributions leads to unitary integrals, which play an important role in other applications, particularly in physics. Balantekin [86] introduced the character expansions of the unitary group and used this technique to simplify the integrations over the unitary group when the coefficient matrices appearing in the integrand are nonzero–determinant square matrices. However, due to the nature of some applications, including the joint eigenvalue distribution of Wishart matrices, the resulting matrix in the integrand is not always full rank or, equivalently, a group member to apply the character expansions.

On the other hand, the simplicity of Balantekin's approach, as well as the convenient mathematical form of the unitary integral results for further analysis, motivated researchers to generalize the method for rectangular complex matrices in the integrand. However, the only paper that actually claims to do the generalization [95] fails to obtain correct results, even for the previously known unitary integrals, because the authors simply follow Balantekin's approach by assuming $N = M$, and in the end, they take the limit of the final result.

In this thesis, we showed that there is a difference between assuming $N = M$, and assuming the matrix integrand is full rank. As explained in Remark 2.1, the coefficient matrices in the integrand should be properly grouped, and Propositions 2.1 and 2.2 should be carefully applied to unitary integrals. Hence, we proposed a universal integration framework to use the character expansions for any unitary integral with general rectangular complex matrices in the integrand, which prevents the possible errors due to misusing Propositions 2.1 and 2.2.

We employed our proposed integration framework to solve three generalized unitary integrals. These integrals have been solved before only for special cases, by using considerably more complicated methods [85]. We used the results of the unitary integrals to obtain the joint eigenvalue distributions of Wishart matrices for common statistical assumptions. Due to the convenient mathematical form of the joint eigenvalue distributions derived in this thesis, we derived the exact eigenvalue densities of Wishart matrices in a unified approach. Our results are in the form of finite summations of determinants, which can be easily calculated and also used for further analysis, such as the capacity and performance analysis of MIMO systems with precoding. Accordingly, we showed that by using the joint eigenvalue distributions derived in this thesis, the MIMO capacity derivation is significantly simplified in comparison to the previous derivations in the literature.

Finally, as another application of the eigenvalue distributions, we used a more appropriate pdf expression for the distribution of the largest eigenvalue to obtain the exact closed–form expression of the BER performance for the optimal beamforming in finite summations for PAM and QAM constellations. The resulting expression was used to derive an approximate BER performance for limited feedback transmit beamforming in finite–series expressions.

We included some applications of the joint eigenvalue distributions derived in this thesis, but they have other applications as well. By using our results, the distributions of a subset of eigenvalues, the largest eigenvalue and the smallest eigenvalue of the Wishart matrices can be derived much more easily than the available ones in the literature.

According to our investigations, and also based on the comments we received from the reviewers of our publication [133], there are more applications in physics involving unitary integrals, which can be developed if those integrals can be solved. On the other hand, we considered only the most practical channel models in this thesis. However, other channel models can be considered such as the semi–correlated and full–correlated Ricean channels; the i.i.d., semi–correlated and full–correlated Hoyt channels; and the i.i.d., semi–correlated and full–correlated Ricean–Hoyt channels. (In Hoyt fading, the real and imaginary parts of the complex Gaussian channel gains have zero means but different variances.) For our analysis in this thesis, we considered the core standard Gaussian matrix $\mathbf{G}$ with identical distribution for its elements. However, in some applications, including the capacity analysis of cooperative MIMO networks, the elements of the channel matrix are not identically distributed.

According to our best knowledge, the above cases are not considered in the literature because the corresponding joint eigenvalue distributions lead to complicated multiple unitary integrals that cannot currently be solved, even by the character expansion method. To be more specific, such unitary integrals require a more general form of Lemma 2.4, which represents the orthogonality relation between the unitary matrix elements. The generalization of Lemma 2.4, however, requires a deep knowledge of the representation theory of unitary group and could be subject for future research.

Appendix A

Generalization of l'Hôpital Rule

A.1

Lemma A.1 : If we define

$$R(x_1,\ldots,x_N)=\frac{\det\limits_{N}\Big[\,f_i(x_j)\,\Big]}{\Delta(x_1,\ldots,x_N)}$$

where $i,j=1,\ldots,N$, then

$$\lim_{\{x_{M+1},\ldots,x_N\}\to x_0} R(x_1,\ldots,x_N)=\frac{\det\limits_{N}[\,\mathbf{F}\,]}{\Delta(x_1,\ldots,x_M)\prod_{i=1}^{M}(x_i-x_0)^{N-M}\prod_{j=1}^{N-M-1}j!}$$

where

$$\mathbf{F}=\left[f_i(x_j)\Big|_{j=1}^{M},\ f_i^{(N-j)}(x_0)\Big|_{j=M+1}^{N}\right],$$

$i=1,2,\ldots,N$ generates all rows of $\mathbf{F}$, and $f^{(k)}$ denotes the k-th derivative of the function f.

▼

Proof: If only x_N approaches to x_0, we can define

$$A^{(1)}=\frac{\det\limits_{N}\left[\,f_i(x_j)\Big|_{j=1}^{N-1},\ f_i(x_0)\,\right]}{\Delta(x_1,\ldots,x_{N-1})\prod_{i=1}^{N-1}(x_i-x_0)} \tag{A.1}$$

$$=\frac{\det\limits_{N}\left[\,f_i(x_j)\Big|_{j=1}^{N-2},\ f_i(x_{N-1}),f_i(x_0)\,\right]}{\Delta(x_1,\ldots,x_{N-2})\prod_{i=1}^{N-2}(x_i-x_0)\prod_{i=1}^{N-2}(x_i-x_{N-1})\,(x_{N-1}-x_0)}\,. \tag{A.2}$$

Now if x_{N-1} approaches x_0, $A^{(1)}$ will have a first–order ambiguity since the determinant in the numerator and the denominator will both go to zero. By using the

Laplacian determinant expansion over the $(N-1)$-th column and taking only the parts in (A.2) which cause the ambiguity, we can write

$$\begin{aligned}
\lim_{x_{N-1}\to x_0} & \frac{\det_N \left[f_i(x_j)\Big|_{j=1}^{N-2}, f_i(x_{N-1}), f_i(x_0) \right]}{(x_{N-1}-x_0)} \\
&= \lim_{x_{N-1}\to x_0} \frac{\sum_{i=1}^{N}(-1)^{i+N-1}\, m_{i,N-1}\, f_i(x_{N-1})}{(x_{N-1}-x_0)} \\
&= \sum_{i=1}^{N}(-1)^{i+N-1}\, m_{i,N-1}\, f_i'(x_0) \\
&= \det_N \left[f_i(x_j)\Big|_{j=1}^{N-2}, f_i'(x_0), f_i(x_0) \right] \qquad \text{(A.3)}
\end{aligned}$$

where $m_{i,j}$ is the determinant of the matrix obtained by eliminating the i-th row and j-th column of the argument matrix in the numerator. Here, the second equality is obtained by using l'Hôpital rule. Therefore, from (A.2) and (A.3), we obtain

$$\begin{aligned}
A^{(2)} &= \lim_{x_{N-1}\to x_0} A^{(1)} \\
&= \frac{\det_N \left[f_i(x_j)\Big|_{j=1}^{N-2}, f_i'(x_0), f_i(x_0) \right]}{\Delta(x_1,\ldots,x_{N-2}) \prod_{i=1}^{N-2}(x_i-x_0)^2} \\
&= \frac{\det_N \left[f_i(x_j)\Big|_{j=1}^{N-3}, f_i(x_{N-2}), f_i'(x_0), f_i(x_0) \right]}{\Delta(x_1,\ldots,x_{N-3}) \prod_{i=1}^{N-3}(x_i-x_0)^2 \prod_{i=1}^{N-3}(x_i-x_{N-2})\,(x_{N-2}-x_0)^2} .
\end{aligned}$$

Now if x_{N-2} approaches to x_0, $A^{(2)}$ will have a second–order ambiguity since the determinant in the numerator and the denominator will both go to zero if l'Hôpital rule is applied once. By applying l'Hôpital rule twice, we have

$$\begin{aligned}
A^{(3)} &= \lim_{x_{N-2}\to x_0} A^{(2)} \\
&= \frac{\det_N \left[f_i(x_j)\Big|_{j=1}^{N-3}, f_i^{(2)}(x_0), f_i^{(1)}(x_0), f_i(x_0) \right]}{\Delta(x_1,\ldots,x_{N-3}) \prod_{i=1}^{N-3}(x_i-x_0)^3 \; 2!} .
\end{aligned}$$

By repeating the above procedure, Lemma A.1 is proved.

■

A.2

Let us define

$$\Phi(x_1,\ldots,x_N)=\frac{\det\limits_{N}\left[\, x_j^{-L} I_L(\alpha_i\, x_j)\,\right]}{\Delta(x_1^2,\ldots,x_N^2)} \tag{A.4}$$

where $I_L(\cdot)$ denotes the modified Bessel function. If only x_N approaches zero in (A.4), we can define

$$\begin{aligned}\Phi^{(1)}(x_1,\ldots,x_{N-1}) &= \lim_{x_N\to 0}\Phi(x_1,\ldots,x_N)\\ &= \frac{\det\limits_{N}\left[\left[x_j^{-L}I_L(\alpha_i\,x_j)\right]_{j=1}^{N-1},\lim\limits_{x_N\to 0}\frac{I_L(\alpha_i\,x_N)}{x_N^L}\right]}{\Delta(x_1^2,\ldots,x_{N-1}^2)\prod_{j=1}^{N-1}x_j^2}\end{aligned} \tag{A.5}$$

where a L order of ambiguity exists in the limit. By applying the l'Hôpital rule L times, we can write

$$\begin{aligned}\lim_{x_N\to 0}\frac{I_L(\alpha_i\,x_N)}{x_N^L} &= \frac{\alpha_i^L}{L!}\, I_L^{(L)}(0)\\ &= \frac{\alpha_i^L}{L!\,2^L}\binom{L}{0}\end{aligned} \tag{A.6}$$

where $I_L^{(k)}(\cdot)$ denotes the k-th derivative of the Bessel function. Here, the second equality comes from the facts that [134]

$$I_L^{(k)}(z)=\frac{1}{2^k}\sum_{i=0}^{k}\binom{k}{i}I_{L-k+2i}(z), \tag{A.7}$$

$I_{L\neq 0}(0)=0$ and $I_0(0)=1$. By substituting (A.6) into (A.5), we have

$$\Phi(x_1,\ldots,x_{N-1})=\frac{\binom{L}{0}}{L!\,2^L}\times\frac{\det\limits_{N}\left[\left[x_j^{-L-2}I_L(\alpha_i\,x_j)\right]_{j=1}^{N-1},\ \alpha_i^L\right]}{\Delta(x_1^2,\ldots,x_{N-1}^2)}. \tag{A.8}$$

Now if x_{N-1} approaches zero in (A.8), we have

$$\begin{aligned}\Phi^{(2)}(x_1,\ldots,x_{N-2}) &= \lim_{x_{N-1}\to 0}\Phi^{(1)}(x_1,\ldots,x_{N-1})\\ &= \frac{\binom{L}{0}}{L!\,2^L}\times\frac{\det\limits_{N}\left[\left[x_j^{-L-2}I_L(\alpha_i\,x_j)\right]_{j=1}^{N-2},\ \lim\limits_{x_{N-1}\to 0}\frac{I_L(\alpha_i\,x_{N-1})}{x_{N-1}^{L+2}},\ \alpha_i^L\right]}{\Delta(x_1^2,\ldots,x_{N-2}^2)\prod_{j=1}^{N-2}x_j^2}\end{aligned} \tag{A.9}$$

where the limit has a $L+2$ order of ambiguity. By applying the l'Hôpital rule $L+2$ times, we can write

$$\begin{aligned}\lim_{x_{N-1}\to 0} \frac{I_L(\alpha_i\, x_{N-1})}{x_{N-1}^{L+2}} &= \frac{\alpha_i^{L+2}}{(L+2)!}\, I_L^{(L+2)}(0) \\ &= \frac{\alpha_i^{L+2}}{(L+2)!\, 2^{L+2}} \binom{L+2}{1}. \end{aligned} \tag{A.10}$$

Therefore,

$$\Phi^{(2)}(x_1,\ldots,x_{N-2}) = \frac{\binom{L}{0}\binom{L+2}{1}}{L!(L+2)!\, 2^L 2^{L+2}} \times \frac{\det_N\left[\left[x_j^{-L-4} I_L(\alpha_i\, x_j)\right]_{j=1}^{N-2},\ \alpha_i^{L+2},\ \alpha_i^L\right]}{\Delta(x_1^2,\ldots,x_{N-2}^2)}. \tag{A.11}$$

Finally, by induction, we can conclude the following lemma:

Lemma A.2 : If

$$\Phi(x_1,\ldots,x_N) = \frac{\det_N\left[\, x_j^{-L} I_L(\alpha_i\, x_j)\,\right]}{\Delta(x_1^2,\ldots,x_N^2)},$$

then

$$\begin{aligned}\Phi^{(P)}(x_1,\ldots,x_{N-P}) &= \lim_{\{x_{N-P+1},\ldots,x_N\}\to 0} \Phi(x_1,\ldots,x_N) \\ &= 2^{-P(L+P-1)} \prod_{i=0}^{P-1}\left[\frac{1}{i!(L+i)!}\right] \\ &\quad \times \frac{\det_N\left[I_L(\alpha_i\, x_j)\Big|_{j=1}^{N-P},\ \alpha_i^{L+2(N-j)}\Big|_{j=N-P+1}^{N}\right]}{\Delta(x_1^2,\ldots,x_{N-P}^2)\prod_{j=1}^{N-P} x_j^{L+2P}}. \end{aligned} \tag{A.12}$$

▼

Appendix B

Generalization of Cauchy–Binet Formula

For square matrices, the following lemma has been proved in [96] as the Cauchy–Binet formula:

Lemma B.1 : Given vectors $\mathbf{x}$ and $\mathbf{y}$ with dimension N, and a power series expansion $f(z) = \sum_{i=0}^{\infty} a_i z^i$ convergent for $|z| < \xi$, then if $\left|x_i y_j\right| < \xi$ for all $1 \leqslant i, j \leqslant N$, one can write

$$\sum_{\mathbf{k}_N \geqslant 0} \det_N \left[x_i^{k_j} \right] \det_N \left[y_i^{k_j} \right] \prod_{i=1}^{N} a_{k_i} = \det_N \left[f(x_i y_j) \right] \tag{B.1}$$

where $\mathbf{k}_N$ is an irreducible representation of $\mathrm{GL}(N, \mathcal{C})$.

▼

If only x_N approaches zero in (B.1), then $\det_N \left[x_i^{k_j} \right]$ goes to zero except when $k_N = 0$ and $k_{N-1} > 0$. Thus, for $x_N = 0$ and $k_N = 0$ we have

$$a_0 \sum_{\mathbf{k}_{N-1} > 0} \det_{N-1} \left[x_i^{k_j} \right] \det_N \left[y_i^{k_j} \Big|_{j=1}^{N-1}, 1 \right] \prod_{i=1}^{N-1} a_{k_i} = \det_N \begin{bmatrix} f(x_i y_j) \Big|_{N-1}^{i=1} \\ a_0 \end{bmatrix} \tag{B.2}$$

or, equivalently,

$$\sum_{\mathbf{k}_{N-1} > 0} \det_{N-1} \left[x_i^{k_j} \right] \det_N \left[y_i^{k_j} \Big|_{j=1}^{N-1}, 1 \right] \prod_{i=1}^{N-1} a_{k_i} = \det_N \begin{bmatrix} f(x_i y_j) \Big|_{N-1}^{i=1} \\ 1 \end{bmatrix}. \tag{B.3}$$

If we define $\mathbf{k}'_{N-1} = \mathbf{k}_{N-1} - 1$, then $\mathbf{k}'_{N-1}$ will be an irreducible representation since now $\mathbf{k}'_{N-1} \geqslant 0$. Therefore,

$$\sum_{\mathbf{k}'_{N-1}\geqslant 0} \det_{N-1}\left[x_i^{k'_j+1} \right] \det_N \left[y_i^{k'_j+1}\Big|_{j=1}^{N-1} , 1 \right] \prod_{i=1}^{N-1} a_{k'_i+1} =$$

$$\prod_{i=1}^{N-1} x_i \sum_{\mathbf{k}'_{N-1}\geqslant 0} \det_{N-1}\left[x_i^{k'_j} \right] \det_N \left[y_i^{k'_j+1}\Big|_{j=1}^{N-1} , 1 \right] \prod_{i=1}^{N-1} a_{k'_i+1} . \quad \text{(B.4)}$$

From (B.3) and (B.4), and by replacing $\mathbf{k}'_{N-1}$ with $\mathbf{k}_{N-1}$, we obtain

$$\sum_{\mathbf{k}_{N-1}\geqslant 0} \det_{N-1}\left[x_i^{k_j} \right] \det_N \left[y_i^{k_j+1}\Big|_{j=1}^{N-1} , 1 \right] \prod_{i=1}^{N-1} a_{k_i+1} = \frac{1}{\prod_{i=1}^{N-1} x_i} \det_N \begin{bmatrix} f(x_i y_j)\Big|_{N-1}^{i=1} \\ 1 \end{bmatrix} . \quad \text{(B.5)}$$

Now if we set $x_{N-1} = 0$, then $\det_{N-1}\left[x_i^{k_j} \right]$ goes to zero except when $k_{N-1} = 0$ and $k_{N-2} > 0$. By repeating the same procedure as described above, we have

$$a_1 \prod_{i=1}^{N-2} x_i \sum_{\mathbf{k}_{N-2}\geqslant 0} \det_{N-2}\left[x_i^{k_j} \right] \det_N \left[y_i^{k_j+2}\Big|_{j=1}^{N-2} , y_i \, , 1 \right] \prod_{i=1}^{N-2} a_{k_i+2}$$

$$= \lim_{x_{N-1}\to 0} \frac{\det_N \begin{bmatrix} f(x_i y_j) \Big|_{N-2}^{i=1} \\ f(x_{N-1} y_j) \\ 1 \end{bmatrix}}{x_{N-1} \prod_{i=1}^{N-2} x_i} \quad \text{(B.6)}$$

where a first–order ambiguity exists in the right–hand side of (B.6) since both the numerator and denominator go to zero when $x_{N-1} \to 0$. By using the Laplacian determinant expansion (A.3) and applying the l'Hôpital rule once, we attain

$$\lim_{x_{N-1}\to 0} \frac{\det_N \begin{bmatrix} f(x_i y_j) \Big|_{N-2}^{i=1} \\ f(x_{N-1} y_j) \\ 1 \end{bmatrix}}{x_{N-1} \prod_{i=1}^{N-2} x_i} = \lim_{x_{N-1}\to 0} \frac{\sum_{j=1}^{N} (-1)^{N-1+j} \, m_{N-1,j} \, f(x_{N-1} y_j)}{x_{N-1} \prod_{i=1}^{N-2} x_i}$$

$$= \lim_{x_{N-1}\to 0} \frac{\sum_{j=1}^{N} (-1)^{N-1+j} \, m_{N-1,j} \, y_j \, f^{(1)}(x_{N-1} y_j)}{\prod_{i=1}^{N-2} x_i}$$

$$= \frac{a_1 \sum_{j=1}^{N} (-1)^{N-1+j} \, m_{N-1,j} \, y_j}{\prod_{i=1}^{N-2} x_i}$$

$$
= \frac{a_1 \det_N \begin{bmatrix} f(x_i y_j)\Big|_{N-2}^{i=1} \\ y_j \\ 1 \end{bmatrix}}{\prod_{i=1}^{N-2} x_i} \tag{B.7}
$$

where $m_{i,j}$ is the determinant of the matrix obtained by crossing out the i-th row and j-th column of the argument matrix, and $f^{(k)}$ denotes the k-th derivative of the function f. Note that since $f(z) = \sum_{i=0}^{\infty} a_i z^i$, thus $f^{(k)}(0) = k!\, a_k$.

By substituting (B.7) into (B.6), we have

$$
\sum_{\mathbf{k}_{N-2} \geqslant 0} \det_{N-2}\left[x_i^{k_j} \right] \det_N \left[y_i^{k_j+2}\Big|_{j=1}^{N-2}, y_i, 1 \right] \prod_{i=1}^{N-2} a_{k_i+2} = \frac{\det_N \begin{bmatrix} f(x_i y_j)\Big|_{N-2}^{i=1} \\ y_j \\ 1 \end{bmatrix}}{\prod_{i=1}^{N-2} x_i^2}. \tag{B.8}
$$

By following the above procedure when x_{N-2} is set to zero, we have a second–order ambiguity on the right–hand side of (B.8) so that we need to perform the l'Hôpital rule twice. Since $f^{(2)}(0) = 2!\, a_2$, the 2! coefficient that appears in the numerator will be eliminated by the 2! coefficient generated in the denominator after derivations. Similarly, the factor a_2 will be omitted by the one that appears in the left–hand side of (B.8). Consequently,

$$
\sum_{\mathbf{k}_{N-3} \geqslant 0} \det_{N-3}\left[x_i^{k_j} \right] \det_N \left[y_i^{k_j+3}\Big|_{j=1}^{N-3}, y_i^2, y_i, 1 \right] \prod_{i=1}^{N-3} a_{k_i+3} = \frac{\det_N \begin{bmatrix} f(x_i y_j)\Big|_{N-3}^{i=1} \\ y_j^2 \\ y_j \\ 1 \end{bmatrix}}{\prod_{i=1}^{N-3} x_i^3}. \tag{B.9}
$$

By induction, we conclude the following lemma:

Lemma B.2 : *(Generalized Cauchy–Binet Formula)* Given vectors $\mathbf{x}$ and $\mathbf{y}$ with dimensions M and N ($M \leqslant N$), respectively, and a power series expansion $f(z) = \sum_{i=0}^{\infty} a_i z^i$ convergent for $|z| < \xi$, then if $\left| x_i y_j \right| < \xi$ for all $1 \leqslant i \leqslant M$ and $1 \leqslant j \leqslant N$, one can write

$$\sum_{\mathbf{k}_M \geqslant 0} \det_M \left[x_i^{k_j} \right] \det_N \left[y_i^{k_j+N-M} \Big|_{j=1}^{M} , \, y_i^{N-j} \Big|_{j=M+1}^{N} \right] \prod_{i=1}^{M} a_{k_i+N-M}$$

$$= \frac{1}{\prod_{i=1}^{M} x_i^{N-M}} \det_N \begin{bmatrix} f(x_i y_j) \Big|_{M}^{i=1} \\ y_j^{N-i} \Big|_{N}^{i=M+1} \end{bmatrix} \quad \text{(B.10)}$$

where $\mathbf{k}_M$ is an irreducible representation of $\mathrm{GL}(M, \mathcal{C})$.

▼

Appendix C

Auxiliary Expressions

C.1

Lemma C.1 : Assuming $\mathbf{k} = (k_1, \ldots, k_N)^T$ is a vector of N non-negative integers, we have

$$\det_N \left[\frac{k_i!}{(k_i - N + j)!} \right] = \Delta_N(\mathbf{k}) \tag{C.1}$$

where the matrix elements with $(k_i - N + j) < 0$ are zero.

▼

Proof: From (C.1), we have

$$\begin{aligned}
\det_N \left[\frac{k_i!}{(k_i - N + j)!} \right] &= \det_N [\, k_i(k_i-1)\cdots(k_i - N + 2), \ldots, k_i(k_i-1), k_i, 1\,] \\
&= \det_N [\, k_i(k_i-1)\cdots(k_i - N + j + 1)\,] \\
&= \det_N \left[\, k_i^{N-j} + \mathcal{P}(N-j-1, k_i) \right]
\end{aligned} \tag{C.2}$$

where $\mathcal{P}(n, x)$ denotes a polynomial of x with degree n. From (C.2), it is clear that if we start from the $(N-2)$-th column, we can omit the term $\mathcal{P}(N-j-1, k_i)$ in the j-th column, by a linear combination of the columns $(j+1), \ldots, N$, without any change in the determinant value. Consequently,

$$\det_N \left[\frac{k_i!}{(k_i - N + j)!} \right] = \det_N \left[k_i^{N-j} \right] = \Delta_N(\mathbf{k})\,. \tag{C.3}$$

■

By applying Lemma C.1 into (2.1) we have

$$
\begin{aligned}
d_{\mathbf{r}_N} &= \frac{1}{\prod_{i=1}^{N}(N-i)!} \det_N\left[\frac{(r_i+N-i)!}{(r_i+N-i-N+j)!}\right] \\
&= \frac{1}{\prod_{i=1}^{N}(N-i)!} \det_N\left[\frac{k_i!}{(k_i-N+j)!}\right] \\
&= \frac{1}{\prod_{i=1}^{N}(N-i)!} \Delta_N(\mathbf{k})
\end{aligned}
$$

where $k_i \triangleq r_i + N - i$.

C.2 Leibniz Formula for Determinants

The Leibniz formula for the determinant expansion [48] is as follows:

$$
\begin{aligned}
\det_M\left[\,X_{ij}\,\right] &= \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \prod_{i=1}^{M} X_{ia_i} && \text{(C.4)} \\
&= \frac{1}{M!} \sum_{\mathbf{a}} \sum_{\mathbf{b}} \mathrm{S}(\mathbf{a})\, \mathrm{S}(\mathbf{b}) \prod_{i=1}^{M} X_{a_i b_i} && \text{(C.5)}
\end{aligned}
$$

where the vector $\mathbf{a} = (a_1, a_2, \ldots, a_M)^T$ is a permutation of integers $(1, 2, \ldots, M)$, $\mathrm{S}(\mathbf{a}) = +1$ if the permutation is even, $\mathrm{S}(\mathbf{a}) = -1$ if the permutation is odd, and the summation is over all possible permutations.

From (C.4), we have

$$
\begin{aligned}
\frac{\partial}{\partial z} \det_M\left[\,X_{ij}\,\right] &= \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \frac{\partial}{\partial z} \prod_{i=1}^{M} X_{ia_i} \\
&= \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \sum_{m=1}^{M} \frac{\partial}{\partial z} X_{ma_m} \prod_{i=1, i\neq m}^{M} X_{ia_i} \\
&= \sum_{m=1}^{M} \sum_{\mathbf{a}} \mathrm{S}(\mathbf{a}) \frac{\partial}{\partial z} X_{ma_m} \prod_{i=1, i\neq m}^{M} X_{ia_i} \\
&= \sum_{m=1}^{M} \det_M\left[\,X_{m,ij}\,\right]
\end{aligned}
$$

where

$$
X_{m,ij} \triangleq \begin{cases} \dfrac{\partial}{\partial z} X_{ij} & , \text{if} \quad i = m \quad ; \\ X_{ij} & , \text{otherwise.} \end{cases} \qquad \text{(C.6)}
$$

C.3

From the definition of the Vandermonde determinant in (2.4), note that

$$\begin{aligned}\underset{N}{\Delta}(\boldsymbol{\lambda}) &= \prod_{i<j}(\lambda_i - \lambda_j)\\ &= \prod_{i<j}\Big((\lambda_i - z) - (\lambda_j - z)\Big)\\ &= \underset{N}{\Delta}(\boldsymbol{\lambda} - z) \end{aligned} \tag{C.7}$$

and

$$\begin{aligned}\underset{N}{\Delta}(\boldsymbol{\lambda}) &= \prod_{i<j}(\lambda_i - \lambda_j)\\ &= \prod_{i<j}\frac{1}{z}\,(z\lambda_i - z\lambda_j)\\ &= \frac{1}{z^{\frac{N(N-1)}{2}}}\,\prod_{i<j}(z\lambda_i - z\lambda_j)\\ &= \frac{1}{z^{\frac{N(N-1)}{2}}}\,\underset{N}{\Delta}(z\boldsymbol{\lambda}) \end{aligned} \tag{C.8}$$

where z is a constant. Moreover,

$$\begin{aligned}\underset{N}{\Delta}(\boldsymbol{\lambda}) &= \prod_{i<j}(\lambda_i - \lambda_j)\\ &= \prod_{i<j}\lambda_i\lambda_j\left(\frac{-1}{\lambda_i} - \frac{-1}{\lambda_j}\right)\\ &= \underset{N}{\Delta}\left(\frac{-1}{\boldsymbol{\lambda}}\right)\prod_{i=1}^{N}\lambda_i^{N-1}\,. \end{aligned} \tag{C.9}$$

Therefore,

$$\begin{aligned}\underset{N}{\Delta}(\boldsymbol{\lambda}) &= \underset{N}{\Delta}(\boldsymbol{\lambda} - z)\\ &= \underset{N}{\Delta}\left(\frac{-1}{\boldsymbol{\lambda} - z}\right)\prod_{i=1}^{N}(\lambda_i - z)^{N-1}\\ &= \underset{N}{\Delta}\left(\frac{-1}{\boldsymbol{\lambda} - z} - \frac{1}{z}\right)\prod_{i=1}^{N}(\lambda_i - z)^{N-1}\\ &= \underset{N}{\Delta}\left(\frac{-\boldsymbol{\lambda}}{z(\boldsymbol{\lambda} - z)}\right)\prod_{i=1}^{N}(\lambda_i - z)^{N-1}\\ &= \frac{1}{(-z)^{\frac{N(N-1)}{2}}}\,\underset{N}{\Delta}\left(\frac{\boldsymbol{\lambda}}{\boldsymbol{\lambda} - z}\right)\prod_{i=1}^{N}(\lambda_i - z)^{N-1}\,. \end{aligned} \tag{C.10}$$

Consequently, by setting $z = -1$, we have

$$\underset{N}{\Delta}(\boldsymbol{\lambda}) = \underset{N}{\Delta}\left(\frac{\boldsymbol{\lambda}}{\boldsymbol{\lambda}+1}\right)\prod_{i=1}^{N}(\lambda_i + 1)^{N-1}\,. \tag{C.11}$$

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, 1948.

[2] G. Forney, *Cocatenated Codes.* Cambridge, MA: M.I.T. Press, 1966.

[3] R. G. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[4] G. Clark and J. B. Cain, *Error-Correction Coding for Digital Communications.* New York: Plenum, 1981.

[5] E. Biglieri, *Introduction to Trellis-Coded Modulation, with Applications.* New York: Macmillan, 1991.

[6] D. J. C. S. Lin, *Error Control Coding*, 2nd ed. NJ: Prentice Hall, 2004.

[7] G. L. Stuber, *Principles of Mobile Communications.* Norwell, MA: Kluwer, 1996.

[8] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. New Jersey: Prentice Hall, 2001.

[9] M. Patzold, *Mobile Fading Channels.* England: John Wiley, 2002.

[10] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 1st ed. New York: Wiley, 2000.

[11] R. A. Mozingo and T. W. Miller, *Introduction to Adaptive Arrays.* New York: Wiley, 1980.

[12] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications.* New York: Cambridge University Press, 2003.

[13] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications.* New York: Cambridge University Press, 2003.

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

[15] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communication," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

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

[17] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[18] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.

[19] H. E. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1097–1119, May 2003.

[20] E. Telatar, "Capacity of multiantenna Gaussian channels," *AT&T Bell Laboratories, Tech. Memo.*, 1995.

[21] B. Hochwald and T. Marzetta, "Space-time modulation for unknown fading," *Proc. SPIE - Int. Soc. Opt. Eng. (USA)*, vol. 3708, pp. 10 – 19, 1999.

[22] ——, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 543 – 564, Mar. 2000.

[23] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, 2000.

[24] B. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[25] B. Hassibi, B. Hochwald, A. Shokrollahi, and W. Sweldens, "Multiple antennas and representation theory," in *IEEE Int. Symposium on Inform. Theory*, Sorrento, Italy, 2000, p. 337.

[26] B. L. Hughes, "Differential space-time modulation," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2567–2578, 2000.

[27] X. Liang and X. Xia, "Unitary signal constellations for differential spacetime modulation with two transmit antennas: Parametric codes, optimal designs, and bounds," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2291 – 2322, Aug. 2002.

[28] A. Ghaderipoor, M. Hajiaghayi, and C. Tellambura, "Unitary matrix design via genetic search for differential space-time modulation and limited feedback precoding," in *Proc. IEEE Int. Symposium on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Helsinki, Sep. 2006, pp. 1–5.

[29] K. L. Clarkson, W. Sweldens, and A. Zheng, "Fast multiple-antenna differential decoding," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 253–261, Feb. 2001.

[30] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[31] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002.

[32] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[33] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. Commun.*, vol. 49, pp. 1–4, Jan. 2001.

[34] A. Sezgin and E. Jorswieck, "On optimal constellations for quasi-orthogonal space-time codes," in *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 4, Apr. 2003, pp. IV – 345–8.

[35] N. Sharma and C. Papadias, "Improved quasi-orthogonal codes through constellation rotation," *IEEE Trans. Commun.*, vol. 51, pp. 332–335, Mar. 2003.

[36] W. Su and X.-G. Xia, "Signal constellations for quasi-orthogonal space-time block codes with full diversity," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2331 – 2347, Oct. 2004.

[37] M. O. Damen, A. Tewfik, and J. C. Belflore, "A construction of a space-time code based on number theory," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 753–760, Mar. 2002.

[38] M. Damen, H. Gamal, and N. Beaulieu, "Systematic construction of full diversity algebraic constellations," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3344 – 3349, Dec. 2003.

[39] E. Bayer-Fluckiger, F. Oggier, and E. Viterbo, "New algebraic constructions of rotated Z/sup n/-lattice constellations for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 50, no. 4, pp. 702 – 714, Apr. 2004.

[40] J. C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: a 2/spl times/2 full-rate space-time code with nonvanishing determinants," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[41] A. Ghaderipoor, M. Hajiaghayi, and C. Tellambura, "On the design of 2x2 full-rate full-diversity space-time block codes," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Washington, DC, USA, Nov. 26–30, 2007, pp. 3406–3410.

[42] G. D. Golden, G. J. Foschini, R. A.Valenzuela, and P. W.Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *IEE Elect. Lett.*, vol. 35, pp. 14–15, Jan. 1999.

[43] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wirelesscommunication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.

[44] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," vol. 48, pp. 2201–2214, Aug. 2002.

[45] M. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389–2402, 2003.

[46] R. Gowaikar and B. Hassibi, "Efficient near-ML decoding via statistical pruning," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*, 2003, pp. 274–274.

[47] A. Ghaderipoor and C. Tellambura, "A statistical pruning strategy for schnorr-euchner sphere decoding," *IEEE Commun. Lett.*, vol. 12, no. 2, pp. 121–123, Feb. 2008.

[48] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York: Cambridge University Press, 2001.

[49] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[50] B. Hassibi and H. Vikalo, "On the expected complexity of sphere decoding," *Asilomar COnf. on Signals, Systems and computers*, vol. 2, pp. 1051–1055, Nov. 2001.

[51] X. Zhu and R. D. Murch, "Performance analysis of maximum likelihood detection in a MIMOantenna system," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 187–191, Feb. 2002.

[52] G. Taricco and E. Biglieri, "Exact pairwise error probability of space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 510–513, Feb. 2002.

[53] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: an analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, Jul. 2004.

[54] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 271–286, Jan. 2006.

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

[56] E. Biglieri, G. Taricco, and A. Tulino, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1794–1803, Jul. 2002.

[57] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications.* Hanover, MA: Now Publishers Inc., 2004.

[58] N. Prasad and M. K. Varanasi, "Throughput analysis for MIMO systems in the high SNR regime," in *IEEE Int. Symposium on Inform. Theory*, Seattle, WA, Jul. 9–14, 2006, pp. 1954–1958.

[59] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 475–501, 1964.

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

[61] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[62] L. Collin, O. Berder, P. Rostaing, and G. Burel, "Optimal minimum distance-based precoder for MIMO spatial multiplexing systems," *IEEE Trans. Signal Processing*, vol. 52, no. 3, pp. 617–627, 2004.

[63] D. Gore and A. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2580–2588, Oct. 2002.

[64] R. J. Heath, S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[65] R. W. J. Heath and D. J. Love, "Multimode antenna selection for spatial multiplexing systems with linear receivers," *IEEE Trans. Signal Processing*, vol. 53, pp. 3042–3056, Aug. 2005.

[66] A. Ghaderipoor and C. Tellambura, "Optimal limited feedback technique for small size and low rate MIMO systems," in *Proc. Int. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, Vancouver, BC, Canada, July 3–6, 2006.

[67] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct 1998.

[68] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.

[69] S. Zhou and G. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[70] D. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735 – 2747, 2003.

[71] D. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 64–73, 2005.

[72] A. Ghaderipoor and C. Tellambura, "Minimum distance-based limited-feedback precoder for MIMO spatial multiplexing systems," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, Montreal, QC, Canada, Sep. 2006, pp. 1–4.

[73] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20A, no. 1/2, pp. 32–52, 1928.

[74] R. A. Fisher, "The sampling distribution of some statistics obtained from non-linear equations," *Ann. of Eugenics*, vol. 9, pp. 238–249, 1939.

[75] P. L. Hsu, "On the distribution of roots of certain determinantal equations," *Ann. of Eugenics*, vol. 9, pp. 250–258, 1939.

[76] S. N. Roy, "p–Statistics or some generalizations in the analysis of variance appropriate to multivariate problems," *Sankhya: Indian J. of Statistics*, vol. 4, pp. 381–396, 1939.

[77] M. A. Girshick, "On the sampling theory of roots of determinantal equations," *Ann. Math. Statist.*, vol. 10, pp. 203–204, 1939.

[78] E. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. Math.*, vol. 62, pp. 546–564, 1955.

[79] ——, "On the distribution of roots of certain symmetric matrices," *Ann. Math.*, vol. 67, pp. 325–327, 1958.

[80] ——, "Statistical properties of real symmetric matrices with many dimensions," in *4th Canadian Math. Congress*, 1959, pp. 174–176.

[81] ——, "Random matrices in physics," *SIAM Review*, vol. 9, pp. 1–123, 1967.

[82] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Department of Mathematics, MIT, May 1989.

[83] P. Koev and A. Edelman, "The efficient evaluation of the hypergeometric function of a matrix argument," *Math. Comp.*, vol. 75, no. 254, pp. 833–846, Jan. 2006.

[84] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, July 2003.

[85] B. Schlittgen and T. Wettig, "Generalizations of some integrals over the unitary group," *J. Phys. A: Math. and General*, vol. 36, no. 12, pp. 3195–3201, 2003.

[86] A. B. Balantekin, "Character expansions, Itzykson-Zuber integrals and the QCD partition action," *Phys. Rev. D*, vol. 62, no. 8, p. 5017, Oct. 2000.

[87] H. Leutwyler and A. Smilga, "Spectrum of Dirac operator and role of winding number in QCD," *Phys. Rev. D*, vol. 46, pp. 5607–5632, 1992.

[88] J. Miller and J. Wang, "Passive scalars, random flux, and chiral phase fluids," *Phys. Rev. Lett.*, vol. 76, pp. 1461–1464, 1996.

[89] R. Grobe, F. Haake, and H. J. Sommers, "Quantum distinction of regular and chaotic dissipative motion," *Phys. Rev. Lett.*, vol. 61, pp. 1899–1902, 1988.

[90] M. R. Zirnbauer, "Supersymmetry for systems with unitary disorder: Circular ensembles," *J. PHYS. A*, vol. 29, p. 7113, 1996.

[91] Harish-Chandra, "Differential operators on a semi-simple Lie algebra," *Amer. J. Math.*, vol. 79, pp. 87–120, 1957.

[92] M. L. Mehta, *Random Matrices*, 3rd ed. NewYork: Academic Press, 2004.

[93] A. B. Balantekin, "Character expansion for U(N) groups and U(N/M) supergroups," *J. Math. Phys.*, vol. 25, no. 6, pp. 2028–2030, June 1984.

[94] C. Itzykson and J. B. Zuber, "Planar approximation 2," *J. Math. Phys.*, vol. 21, pp. 411–421, 1980.

[95] S. H. Simon, A. L. Moustakas, and L. Marinelli, "Capacity and character expansions: moment-generating function and other exact results for MIMO correlated channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5336–5351, Dec. 2006.

[96] L. K. Hua, *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains.* Providence, RI: American Mathematical Society, 1963.

[97] S. Sternberg, *Group Theory and Physics.* Cambridge, UK: Cambridge University Press, 1995.

[98] G. James and M. Liebeck, *Representations and Chaaracters of Groups*, 2nd ed. Cambridge. UK: Cambridge University Press, 2001.

[99] A. B. Balantekin and I. Bars, "Dimension and character formulas for Lie supergroups," *J. Math. Phys.*, vol. 22, no. 6, pp. 1149–1162, 1981.

[100] H. Weyl, *The Classical Groups.* Princeton, NJ: Princeton University Press, 1948.

[101] T. Guhr and T. Wettig, "An ItzyksonZuber-like integral and diffusion for complex ordinary and supermatrices," *J. Math. Phys.*, vol. 37, pp. 6395–6413, Dec 1996.

[102] E. Khan and C. Heneghan, "Capacity of fully correlated MIMO system using character expansion of groups," *International Journal of Mathematics and Mathematical Sciences*, vol. 2005, pp. 2461–2471, 2005.

[103] R. G. Bartle, *The Elements of Integration and Lebesgue Measure.* Wiley Interscience, 1995.

[104] M. Kang and M. S. Alouini, "Capacity of MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[105] P. J. Forrester and T. D. Hughes, "Complex wishart matrices and conductance in mesoscopic systems: Exact results," *J. Math. Phys.*, vol. 35, no. 12, pp. 6736–6747, 1994.

[106] Y. Demasure and R. A. Janik, "Effective matter superpotentials from wishart random matrices," *Phys. Lett. B*, vol. 553, no. 1–2, pp. 105–108, 2003.

[107] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, "Random matrix approach to cross correlations in financial data," *Phys. Rev. E*, vol. 65, no. 6, p. 066126, Jun 2002.

[108] S. H. Simon and A. L. Moustakas, "Eigenvalue density of correlated complex random wishart matrices," *Phys. Rev. E*, vol. 69, no. 6, p. 065101, Jun 2004.

[109] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. Academic Press, Inc., 1994.

[110] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[111] Z. Wang and G. B. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 4, pp. 657–662, Apr. 2004.

[112] M. Chiani, 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.

[113] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[114] M. Kiessling and J. Speidel, "Mutual information of MIMO channels in correlated Rayleigh fading environments - a general solution," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, vol. 2, Jun. 20–24, 2004, pp. 814–818.

[115] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

[116] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[117] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1101–1112, Mar 2006.

[118] P. Xia, S. Zhou, and G. B. Giannakis, "Multiantenna adaptive modulation with beamforming based on bandwidth-constrained feedback," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 526–536, Mar. 2005.

[119] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Trans. Signal Processing*, vol. 54, no. 5, pp. 1853–1863, May 2006.

[120] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit beamforming relies on finite-rate feedback," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1948–1957, Jul. 2005.

[121] B. Mondal and R. W. Heath, Jr., "Performance analysis of quantized beamforming MIMO systems," *IEEE Trans. Signal Processing*, vol. 54, no. 12, pp. 4753–4766, Dec. 2006.

[122] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.

[123] M. Mitchell, *An Introduction to Genetic Algorithms.* Cambridge, MA: The MIT Press, 1996.

[124] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 418–426, Apr. 2003.

[125] A. Grant, "Performance analysis of transmit beamforming," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 738–744, Apr. 2005.

[126] C. K. Au-Yeung and D. J. Love, "A performance analysis framework for limited feedback beamforming in correlated fading," *IEEE Commun. Lett.*, vol. 10, no. 5, pp. 344–346, May 2006.

[127] J. Zheng, E. R. Duni, and B. D. Rao, "Analysis of multiple-antenna systems with finite-rate feedback using high-resolution quantization theory," *IEEE Trans. Signal Processing*, vol. 55, no. 4, pp. 1461–1476, Apr. 2007.

[128] L. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.

[129] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd ed. Berlin ; New York : Springer-Verlag, 1992.

[130] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[131] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, 2002.

[132] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Eigenvalues and condition numbers of complex random matrices," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 441–456, 2005.

[133] A. Ghaderipoor and C. Tellambura, "Generalization of some integrals over unitary matrices by character expansion of groups," *J. Math. Phys.*, vol. 49, no. 7, p. 073519, July 2008.

[134] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover, 1972.