Performance Limits of Spatial Multiplexing MIMO Systems

Luis García Ordóñez

Departament de Teoria
del Senyal i Comunicacions
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Ph.D Thesis:

Performance Limits of Spatial Multiplexing MIMO Systems

Author:
Luis García Ordóñez

Advisors:
Javier Rodríguez Fonollosa
Alba Pagès Zamora

Signal Processing in Communications Group
Department of Signal Theory and Communications (TSC)
Technical University of Catalonia (UPC)

Jordi Girona 1-3, Campus Nord, Edifici D5
08034 Barcelona, SPAIN

Barcelona, February 2009

A mis padres,

Abstract

Multiple-input multiple-output (MIMO) channels are an abstract and general way to model many different communication systems of diverse physical nature. In particular, wireless MIMO channels have been attracting a great interest in the last decade, since they provide significant improvements in terms of spectral efficiency and reliability with respect to single-input single-output (SISO) channels.

In this thesis we concentrate on spatial multiplexing MIMO systems with perfect channel state information (CSI) at both sides of the link. Spatial multiplexing is a simple MIMO transmit technique that does not require CSI at the transmitter and allows a high spectral efficiency by dividing the incoming data into multiple independent substreams and transmitting each substream on a different antenna. When perfect CSI is available at the transmitter, channel-dependent linear precoding of the data substreams can further improve performance by adapting the transmitted signal to the instantaneous channel eigen-structure. An example of practical relevance of this concept is given by linear MIMO transceivers, composed of a linear precoder at the transmitter and a linear equalizer at the receiver.

The design of linear MIMO transceivers has been extensively studied in the literature for the past three decades under a variety of optimization criteria. However, the performance of these schemes has not been analytically investigated and key performance measures such as the average bit error rate (BER) or the outage probability have been obtained through time-comsuming

Monte Carlo simulations. In contrast to numerical simulations, which do not provide any insight on the system behavior, analytical performance expressions help the system designer to identify the degrees of freedom and better understand their influence on the system performance. This thesis attempts to fill this gap by providing analytical average and outage performance characterizations in some common MIMO channel models. More exactly, we derive exact expressions or bounds (depending on the case) for the average BER and the outage probability of linear MIMO transceivers designed under a variety of design criteria. Special attention is given to the high signal-to-noise ratio (SNR) regime, where the system performance is investigated under two different perspectives. First, from a more practical point-of-view, we characterize the average BER and outage probability versus SNR curves in terms of two key parameters: the diversity gain and the array gain. Then, we focus on the diversity and multiplexing tradeoff framework in order to take into consideration the capability of the system to deal with the fading nature of the channel, but also its ability to accommodate higher data rates as the SNR increases.

The performance of linear MIMO transceivers is simultaneously analyzed for the most common wireless MIMO channel models such as the uncorrelated and semicorrelated Rayleigh, and the uncorrelated Rician MIMO fading channels. For this purpose, we have obtained a general formulation that unifies the probabilistic characterisation of the eigenvalues of Hermitian random matrices with a specific structure, which includes the previous channel distributions as particular cases, i.e., the uncorrelated and semicorrelated central Wishart, the uncorrelated noncentral Wishart, and the semicorrelated central Pseudo-Wishart distributions. Indeed, the proposed formulation and derived results provide a solid framework for the analytical performance evaluation of MIMO systems, but it could also find numerous applications in other fields of statistical signal processing and communications.

Finally, and as a consequence of our performance analysis, limitations inherent to all practical linear MIMO transceiver designs have been enlightened. Accordingly, new schemes have been proposed which achieve considerable performance enhancements with respect to classical linear MIMO transceivers.

Agradecimientos

In the elder days of art
Builder wrought with greatest care
Each minute and unseen part,
For the Gods are everywhere.

Wittgenstein once said that the previous bit of verse could serve him as a motto.

Contents

Notation

Vector and Matrices

$\mathbf{a}$ $(n \times 1)$	$\mathbf{a}$ is a column vector with n elements.
$\mathbf{A}$ $(n \times m)$	$\mathbf{A}$ is a matrix with n rows and m columns.
$[\mathbf{A}]_{i,j}$ or a_{ij}	(i,j)th element of $\mathbf{A}$.
$[\mathbf{A}]_i$	ith column of $\mathbf{A}$.
$\mathbf{A}'$	transpose of $\mathbf{A}$.
$\mathbf{A}^\dagger$	conjugate transpose of $\mathbf{A}$.
$\mathbf{A} > 0$	$\mathbf{A}$ is positive definite.
$\mathbf{A} \geq 0$	$\mathbf{A}$ is positive semidefinite.
$\lvert\mathbf{A}\rvert$	determinant of $\mathbf{A}$.
$\lVert\mathbf{a}\rVert$	Euclidean norm of vector $\mathbf{a}$: $\lVert\mathbf{a}\rVert = \sqrt{\mathbf{a}^\dagger\mathbf{a}}$.
$\lVert\mathbf{A}\rVert_{\mathrm{F}}$	Frobenius norm of matrix $\mathbf{A}$: $\lVert\mathbf{A}\rVert_{\mathrm{F}} = \sqrt{\mathrm{tr}\,(\mathbf{A}^\dagger\mathbf{A})}$.
$\mathrm{rank}(\mathbf{A})$	rank of $\mathbf{A}$.
$\mathrm{tr}(\mathbf{A})$	trace of $\mathbf{A}$.
$\mathbf{A}^{-1}$	inverse of $\mathbf{A}$.
$\mathrm{vec}(\mathbf{A})$	$(nm \times 1)$ vector obtained by stacking the columns of $\mathbf{A}$ $(n \times m)$.
$\mathbf{A} \otimes \mathbf{B}$	$(nm \times nm)$ Kronecker product between $\mathbf{A}$ $(n \times n)$ and $\mathbf{B}$ $(m \times m)$.

$\text{diag}(a_1, \ldots, a_n)$	$(n \times n)$ diagonal matrix with diagonal entries equal to $a_1, \ldots, a_n$.
$\mathbf{I}_n$	$(n \times n)$ identity matrix.
$\mathbf{0}_{n,m}$	$(n \times m)$ zero matrix.

Sets and Permutations

$\mathbb{N}$	set of natural numbers, i.e., positive integers.
$\mathbb{N}^n$	set of natural n-dimensional vectors.
$\mathbb{R}$	set of real numbers.
$\mathbb{R}^n$	set of real n-dimensional vectors.
$\mathbb{R}^{n \times m}$	set of real $n \times m$ matrices.
$\mathbb{C}$	set of complex numbers.
$\mathbb{C}^n$	set of complex n-dimensional vectors.
$\mathbb{C}^{n \times m}$	set of complex $n \times m$ matrices.
$a \in \mathcal{A}$	a belongs to set $\mathcal{A}$.
$\mathcal{A} \subseteq \mathcal{B}$	$\mathcal{A}$ is a subset of $\mathcal{B}$.
$\mathcal{A} \cup \mathcal{B}$	union of sets $\mathcal{A}$ and $\mathcal{B}$.
$\mathcal{A} \cap \mathcal{B}$	intersection of sets $\mathcal{A}$ and $\mathcal{B}$.
$\lvert\mathcal{A}\rvert$	cardinality of set $\mathcal{A}$, i.e., number of elements in $\mathcal{A}$.
$\boldsymbol{\mu}$	set of ordered elements $(\mu_1, \ldots, \mu_n)$.
$\pi(\boldsymbol{\mu})$	permutations of the elements of $\boldsymbol{\mu}$.
$\text{sgn}(\boldsymbol{\mu})$	sign of the permutation $\boldsymbol{\mu}$.

Probability and Statistics

$\mathsf{Pr}\{a \leq b\}$	probability of the event $(a \leq b)$.
$\mathbb{E}\{a\}$	expectation of the random variable a.
$\mathbb{E}_a\{f(a)\}$	expectation of the random variable $f(a)$ with respect to a.
$F_a(x)$	cumulative distribution function of the random variable a.
$f_a(x)$	probability density function of the random variable a.

Miscellanea

$\sim$	distributed as.
$\approx$	approximately equal.
$\triangleq$	defined as.
$\doteq$	exponentially equivalent, i.e., $f(x) \doteq g(x)$ denotes $\lim_{x\to\infty} \frac{\log f(x)}{\log x} = \lim_{x\to\infty} \frac{\log g(x)}{\log x}$.
$\dot{\leq}$, $\dot{\geq}$	exponentially smaller or equal, exponentially greater or equal.
lim	limit.
max, min	maximum and minimum.
arg	argument.
$\lvert a \rvert$	Modulus of the complex scalar a.
$(a)^+$	positive part of the real scalar a, i.e., $a = \max(0, a)$.
$\lceil a \rceil$	integer part of a, i.e., smallest integer greater than or equal to a.
$o(\cdot)$	Landau little-o, i.e., $f(x) = o(g(x))$, $g(x) > 0$, states that $f(x)/g(x) \to 0$ as $x \to 0$.

Distributions

$x \sim \mathcal{CN}(\mu, \sigma^2)$	random variable x follows a complex normal distribution with mean μ and variance σ^2.
$\mathbf{x} \sim \mathcal{CN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	random vector $\mathbf{x}$ $(n \times 1)$ follows a multivariate complex normal distribution with mean vector $\boldsymbol{\mu}$ $(n \times 1)$ and covariance matrix $\boldsymbol{\Sigma}$ $(n \times n)$.
$\mathbf{X} \sim \mathcal{CN}_{n,m}(\boldsymbol{\Theta}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$	random matrix $\mathbf{X}$ $(n \times m)$ follows a matrix variate complex normal distribution with mean matrix $\boldsymbol{\Theta}$ $(n \times m)$ and covariance matrix $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$ $(nm \times nm)$.
$\mathbf{X} \sim \mathcal{W}_n(m, \mathbf{0}_n, \boldsymbol{\Sigma})$	Hermitian random matrix $\mathbf{X}$ $(n \times n)$ follows a complex central Wishart distribution with m degrees of freedom and covariance matrix $\boldsymbol{\Sigma}$ $(n \times n)$.
$\mathbf{X} \sim \mathcal{W}_n(m, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$	Hermitian random matrix $\mathbf{X}$ $(n \times n)$ follows a complex noncentral Wishart distribution with m degrees of freedom, noncentrality parameter matrix $\boldsymbol{\Omega}$ $(n \times n)$, and covariance matrix $\boldsymbol{\Sigma}$ $(n \times n)$.
$\mathbf{X} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \boldsymbol{\Psi})$	Hermitian random matrix $\mathbf{X}$ $(m \times m)$ follows a complex central Pseudo-Wishart distribution with parameters n, m, and covariance matrix $\boldsymbol{\Psi}$ $(m \times m)$.
$\mathbf{X} \sim \mathcal{Q}_{n,m}(\mathbf{0}_n, \mathbf{A}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$	Hermitian random matrix $\mathbf{X}$ follows a central Quadratic form distribution with parameters n, m, and $\mathbf{A}$ $(m \times m)$ and covariance matrices $\boldsymbol{\Sigma}$ $(n \times n)$ ans $\boldsymbol{\Psi}$ $(m \times m)$.

Acronyms and Abbreviations

AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
BPSK	Binary Phase Shift Keying.
cdf	cumulative density function.
cf.	(from Latin *confer*) compare.
CSI	Channel State Information.
CSI-R	Channel State Information at the Receiver.
CSI-T	Channel State Information at the Transmitter.
DSL	Digital Subscriber Line.
e.g.	(from Latin *exempli gratia*) for example.
i.e.	(from Latin *id est*) that is.
i.i.d.	independent and identicallt distributed.
ISI	Inter-Symbol Interference.
mgf	moment generating function.
MIMO	Multiple-Input Multiple-Output.
MISO	Multiple-Input Single-Output.
MMSE	Minimum Mean Square Error.
MRC	Maximum Ratio Combining.

MRT	Maximum Ratio Transmission.
MSE	Mean Square Error.
OFDM	Orthogonal Frequency Division Multiplexing.
pdf	probability density function.
PSK	Phase Shift Keying.
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service.
QPSK	Quadrature Phase Shift Keying.
SIMO	Single-Input Multiple-Output.
SISO	Single-Input Single-Output.
SNR	Signal-to-Noise Ratio.
SVD	Singular Value Decomposition.
V-BLAST	Vertical Bell-Labs Layered Space-Time.
vs.	versus.

1

Introduction

Multiple-input Multiple-output (MIMO) technology constitutes a breakthrough in the design of wireless communication systems, and is already at the core of several wireless standards. The introduction of the spatial dimension (provided by the multiple antennas at the transmitter and the receiver) delivers significant performance enhancements in terms of data transmission rate and transmission reliability with respect to conventional single-antenna wireless systems. Hence, the design of MIMO systems has been traditionally posed under two different perspectives: either the increase of the data transmission rate through spatial multiplexing or the improvement of the system reliability through the increased antenna diversity. Actually, both types of gains can be simultaneously obtained subject to a fundamental tradeoff between the two. This dissertation focuses on spatial multiplexing MIMO systems when perfect channel state information is available at both sides of the link. Analytical studies of these schemes from a communication- and an information-theoretical point-of-view enlighten the implications of the tradeoff beween spatial multiplexing and diversity gain, or, in other words, the tradeoff between transmission rate and reliability. In this chapter we provide a brief overview on wireless MIMO systems to locate, justify, and motivate the analyses developed in the rest of the dissertation.

1.1 MIMO Wireless Communication

The use of multiple antennas at the transmitter and receiver, commonly known as multiple-input multiple-output (MIMO) technology, has rapidly gained in popularity over the past decade due to its powerful performance-enhancing capabilities. MIMO technology offers a number of benefits over conventional single-input single-output (SISO) systems that help to meet the challenges posed by both the impairments in the wireless channel as well as the strict resource (power and bandwidth) constraints. In addition to the time and frequency dimensions (the natural dimensions of digital communication data), the leverage of MIMO is realized by exploiting the spatial dimension inherent in the use of multiple spatially distributed antennas. As such MIMO systems can be viewed as an extension of the so-called smart antennas, a popular technology using antenna arrays either at the transmitter or at the receiver dating back several decades. Although pioneering work on wireless MIMO channels can be found as early as 1987 in [Win87], current interest in MIMO has been mainly inspired by the significant capacity benefits uncovered independently in [Tel99] and [Fos98]. In particular, it was shown that a MIMO channel formed by n_{T} antennas at the transmitter and n_{R} antennas at the receiver provides up to $\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$ times the capacity of a SISO channel without any increase of the required bandwidth or transmitted power. This has prompted progress in areas as diverse as channel modeling, information theory and coding, signal processing, antenna design and multiantenna-aware network design, fixed or mobile.

1.1.1 Benefits of MIMO Technology

MIMO channels provide a number of advantages over conventional SISO channels such as the array gain, the diversity gain, and the multiplexing gain. While the array and diversity gains are not exclusive of MIMO channels and also exist in single-input multiple-output (SIMO) and multiple-input single-output (MISO) channels, the multiplexing gain is a unique characteristic of MIMO channels. These gains are described in brief below [Böl02b, Big07].

Array Gain

Array gain denotes the improvement in receive signal-to-noise ratio (SNR) that results from a coherent combining effect of the information signals. The coherent combining may be realized

through spatial processing at the receive antenna array and/or spatial pre-processing at the transmit antenna array. Formally, the array gain characterizes the horizontal shift of the error probability versus transmitted or received power curve (in a log-log scale), due to the gain in SNR.

Spatial Diversity Gain

Diversity gain is the improvement in link reliability obtained by receiving replicas of the information signal through (ideally independent) fading links. With an increasing number of independent copies, the probability that at least one of the signals is not experiencing a deep fade increases, thereby improving the quality and reliability of reception. A MIMO channel with n_{T} transmit and n_{R} receive antennas offers potentially $n_{\mathsf{T}}n_{\mathsf{R}}$ independently fading links and, hence, a spatial diversity order of $n_{\mathsf{T}}n_{\mathsf{R}}$. Formally, the diversity gain characterizes the slope of the error probability versus transmitted or received power curve (in a log-log scale) in the high-SNR regime.

Spatial Multiplexing Gain

MIMO systems offer a linear increase in data rate through spatial multiplexing [Böl02a, Fos98, Tel99], i.e., transmitting multiple, independent data streams within the bandwidth of operation. Under suitable channel conditions, such as rich scattering in the environment, the receiver can separate the data streams. Furthermore, each data stream experiences at least the same channel quality that would be experienced by a SISO system, effectively enhancing the capacity by a multiplicative factor equal to the number of substreams. In general, the number of data streams that can be reliably supported by a MIMO channel coincides with the minimum of the number of transmit antennas n_{T} and the number of receive antennas n_{R}, i.e., $\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$.

1.1.2 A Fundamental Tradeoff

Essentially, different design criteria of MIMO communication schemes are based on exploiting the previous gains, especially the spatial diversity and multiplexing gains. Actually, both perspectives come from different ways of understanding the ever-present fading in wireless communications. Traditionally, fading is considered as a source of randomness that makes wireless

links unreliable. In response, a natural attempt is to use multiple antennas for compensating the random signal fluctuations and achieving a steady channel gain. The spatial dimension is exploited in this case to maximize diversity. Observe that each pair of transmit and receive antennas provides a different (possibly independent) signal path from transmitter to receiver. By sending signals that carry the same information over a number of different paths, multiple independent faded replicas of the data can be obtained at the receiver end, increasing, thus, the reliability of the reception process. Some examples of MIMO schemes which fall within this category are space-time codes [Tar98, Has02] and orthogonal designs [Ala98, Tar99].

A different line of thought suggests that in a MIMO channel, fading can in fact be beneficial through increasing the degrees of freedom available for communication [Fos98, Tel99]. Essentially, if the path gains between individual transmit and receive antenna pairs fade independently, the channel matrix is well-conditioned with high probability, in which case multiple spatial channels are created. Hence, the data rate can be increased by transmitting independent information in parallel through the available spatial channels. This spatial multiplexing phenomenon were first exploited in [Pau94] and by the BLAST and V-BLAST architectures [Fos96, Fos99, Gol99].

This dichotomic view of the fading process and by extension of the analysis and design MIMO systems is not appropriate. In fact, given a MIMO channel, both the spatial diversity and the multiplexing gains can be simultaneously obtained, but there is a tradeoff between how much of each type of gain any MIMO scheme can extract: higher spatial multiplexing comes at the price of sacrificing diversity. Several attempts were made to understand the diversity and multiplexing tradeoff in [Hea00, Hea01, Oym02, Oym03]. However, the complete picture of this tradeoff was given by Zheng and Tse in the excellent groundbreaking paper [Zhe03]. To be more specific, [Zhe03] focuses on the high-SNR regime and provides the fundamental tradeoff curve achievable by any scheme, where the spatial multiplexing gain is understood as the fraction of capacity attained at high SNR and the diversity gain indicates the high-SNR reliability of the system. The two previously commented design strategies correspond to the two extreme points of the curve: maximum diversity and no multiplexing gain and maximum multiplexing gain and no diversity gain. The fundamental tradeoff curve bridges the gap between these two extremes and offers insights to understand the overall resources provided by MIMO channels.

1.2 MIMO Communication Systems

Since the emergence of the key ideas in the mid-1990s, a great number of transceiver algorithms for MIMO systems have been proposed in the literature. The different MIMO communication techniques basically depend on the amount and quality of the channel state information (CSI) available at the receiver (CSI-R) and/or transmitter (CSI-T). Clearly, the more channel state information, the better the performance of the system [Pal07].

The most commonly studied situation is that of perfect CSI available at the receiver. CSI-R is traditionally acquired via the transmission of a training sequence, i.e., pilot symbols, that allows the channel estimation. It is also possible to envisage a situation in which CSI is known to both the receiver and the transmitter. CSI-T can be obtained either from a dedicated feedback channel, when the channel is sufficiently slow varying, or by exploiting the channel reciprocity that allows to infer the channel from previous receive measurements. Observe that the first option implies a loss in spectral efficiency due to the utilization of part of the bandwidth to transmit the channel state, whereas the latter requires a full-duplex transmission for the reciprocity principle to hold.

1.2.1 MIMO designs with No CSI-T

MIMO transmit techniques that do not require channel knowledge at the transmitter may be broadly classified into two categories: those designed to increase the transmission rate and those designed to increase reliability. The former are often collectively referred to as spatial multiplexing and the latter as transmit diversity schemes. As previously stated, spatial multiplexing and transmit diversity systems achieve either one of the two extremes in the diversity and multiplexing tradeoff curve but are clearly suboptimal at the other extreme. Techniques that achieve a flexible diversity and multiplexing tradeoff form an important topic of current research [Big07].

1.2.2 MIMO designs with Perfect CSI-T

When perfect CSI-T is available, the transmission can be adapted to each channel realization using signal processing techniques. A low-complexity approach with high potential is the use of linear MIMO transceivers, which are composed of a linear precoder at the transmitter and a lin-

ear equalizer at the receiver. The design of linear MIMO transceivers is generally quite involved since several substreams are typically established over MIMO channels. Precisely, the existence of several substreams, each with its own performance, makes the definition of a global measure of the system performance not clear; as a consequence, a wide span of different design criteria has been explored in the literature (see a historical perspective in [Pal07, Sec. 1.4]). In general, the linear transmitter and receiver are jointly designed to optimize a global cost function that takes into account the individual SNRs, mean square errors (MSEs), or bit error rates (BERs) of the established substreams, the number of which has been chosen beforehand. Palomar developed in [Pal03] a general unifying framework for the joint linear MIMO transceiver design embracing a wide range of different design criteria. In particular, the optimal solution was obtained for the family of Schur-concave and Schur-convex cost functions, which happens to diagonalize the channel (possibly after a rotation of the data symbols) and exploit the spatial multiplexing property of MIMO channels to establish several independent data substreams through the strongest channel eigenmodes. The available transmit power is then distributed among the established substreams according to the specific design criterion.

1.3 Motivation and Outline of the Dissertation

The design MIMO communication systems with perfect CSI has been extensively addressed in the literature based on the optimization of some metric of the system performance. Nevertheless, the performance of the linear MIMO transceivers under realistic channel models has only been evaluated numerically, due to the difficulty of finding closed-form expressions for the average error probability or the outage probability.

In this dissertation we exploit the common spatial multiplexing structure of linear MIMO transceivers to analytically characterize their performance. We consider both a general spatial multiplexing system with CSI and the practical designs of [Pal03]. This allows us to understand how design parameters such the cost function itself or the number of established substreams do impact on the degree of array gain, diversity gain, and multiplexing gain extracted from the MIMO channel. This enlightens the limitations of the traditional design formulation of MIMO linear transceivers and poses an additional optimization stage on top of the conventional one to fully exploit the MIMO channel benefits.

Characterizing the performance of a communication scheme by computing the average error probability as a function of the SNR, for a given data rate, may not be appropriate when comparing several systems transmitting at different rates. Hence, taking into consideration the capability of the system to combat the fading nature of the channel, but also its ability to accommodate higher data rates as the SNR increases, spatial multiplexing MIMO schemes with CSI are also analyzed in the diversity and multiplexing tradeoff framework of [Zhe03]. This allows us to uncover the limitations of the spatial multiplexing structure with respect to the fundamental tradeoff of the MIMO channel and to resolve in which situations CSI-T offers advantages with respect to similar schemes that do no rely on channel information.

The rest of the dissertation is organized as follows.

Chapter 2

The probabilistic characterization of the eigenvalues of the random channel matrix is critical in the performance evaluation of many MIMO communication schemes. In particular, the performance of MIMO systems without CSI-T demands the characterization of the unordered eigenvalues, whereas the techniques that employ CSI-T require the evaluation of probabilities associated with one or several of the eigenvalues in some specific order (often the highest or smallest but sometimes any one in particular within the ordered set). Many different contributions, as early as the sixties in the mathematical literature and much more recently in the signal processing community, provide partial characterizations for specific problems. In Chapter 2 we introduce a unified perspective that can, not only fill the gap of the currently unknown results, but even more importantly, provide a solid framework for the understanding and direct derivation of all the previously known results regarding the joint and marginal distributions of the eigenvalues. Indeed, we consider a general class of Hermitian random matrices that includes some particular cases of the Wishart and other closely related distributions, such as the Pseudo-Wishart or the quadratic form distribution. Since these are precisely the distributions induced by most common wireless MIMO channel models, the results derived in this chapter provide the essential mathematical tools needed for the analytical performance studies developed in the subsequent chapters.

Chapter 2 also provides a self-contained overview on random matrix theory. In addition to our

contributions to this field, it also includes a summary of relevant mathematical preliminaries, such as determinant definitions and properties, definitions of important functions in integral form, or definitions of hypergeometric functions of matrix arguments, as well as an introduction to random matrix distributions.

Chapter 3

The average error probability and outage of probability spatial multiplexing MIMO systems with CSI is analytically investigated in Chapter 3 under common MIMO channel models: uncorrelated Rayleigh, semicorrelated Rayleigh, and uncorrelated Ricean MIMO fading channels. Exact expressions are only given for a channel non-dependent power allocation, which is a case of special interest since it measures the performance of the channel eigenmodes. However, in order to provide more insight into the general system behaviour, the focus is then turned to the high-SNR regime. In particular, we study the average and outage performance of each channel eigenmode at high SNR by characterizing the average BER versus SNR and the outage probability versus SNR curves in terms of the diversity gain, which determines the slope of the curve at high SNR in a log-log scale, and the array gain, which determines the horizontal shift of the curve. We also extend this characterization to global performance measures that take into account all the established substreams. In addition, our general results applied to analyze the performance of a wide family of practical linear MIMO transceivers.

Chapter 4

Chapter 4 focuses mainly on improving the performance achieved by the classical minimum BER linear transceiver. For this purpose, the average BER of the minimum BER linear transceiver with fixed and equal constellations are characterized in an uncorrelated and a semicorrelated Rayleigh and in an uncorrelated Rician fading channel. It turns out that this classical minimum BER design has a diversity order limited by that of the worst eigenmode used, which can be far from the full diversity provided by the channel. This shows that fixing a priori the number of independent data streams to be transmitted, a very common assumption in the linear transceiver design literature, inherently limits the average BER performance of the system. Based on this observation, we propose the minimum BER linear transceiver with fixed rate and equal

constellations and show that it achieves the full diversity of the channel thanks to the joint optimization of the number of substreams and the linear precoder.

Chapter 5

The spatial multiplexing strategy considered in the previous chapters, i.e., dividing the incoming data stream into multiple independent substreams and transmitting them through the channel eigenmodes, is also optimal in the sense of achieving the ergodic channel capacity when perfect CSI is available at both sides of the link. Chapter 5 analyzes the optimality of this technique with respect to the diversity and multiplexing tradeoff. The approach we adopt is to analyze first the individual diversity and multiplexing tradeoff curves of the channel eigenmodes. Then, we obtain the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI by deriving the optimum rate allocation policy among these channel eigenmodes. Practical linear MIMO transceivers are also considered.

Chapter 6

Chapter 6 concludes and summarizes the results of this PhD thesis and presents future research lines.

2

Eigenvalues of a General Class of Hermitian Random Matrices

The performance of multiple-input multiple-output (MIMO) systems is usually related to the eigenstructure of the channel matrix, or, more exactly, to its nonzero eigenvalues. When communicating over MIMO fading channels, we have a random channel matrix that depends on the particular system architecture and propagation conditions. Hence, the probabilistic characterization of these eigenvalues for the underlying channel distribution is necessary in order to derive analytical expressions for the average and outage performance measures. This chapter presents a formulation that unifies the probabilistic characterization of Hermitian random matrices with a specific structure. Based on a general expression for the joint pdf of the ordered eigenvalues, we obtain, among other results, (i) the joint cdf, (ii) the marginal cdf's, and (iii) the marginal pdf's of the ordered eigenvalues, where (ii) and (iii) follow as simple particularizations of (i). Our formulation is shown to include some particular cases of the Wishart and other closely related distributions, such as the Pseudo-Wishart or the quadratic form distribution. Since these are precisely the distributions induced by most common wireless MIMO channel models, the proposed formulation and derived results provide a solid framework for the analytical performance analysis of MIMO systems.

2.1 Introduction

2.1.1 Historical Perspective on the Wishart Distribution

Multivariate analysis is the part of mathematical statistics that derives methods for obtaining and analyzing observations consisting of several vector measurements (see, e.g., [Roy58, Ksh72, Sri79, Mui82, And84, Gup00] for standard textbooks). Typically, samples from the multivariate observation of the physical phenomena under analysis are collected in the columns of the sample observation matrix. In such matrix, when sampling from a multivariate normal population, the columns are independently[1] and identically distributed as multivariate normal with common mean vector and covariance matrix. The assumption of normality, besides mathematical tractability, has a solid empirical and theoretical basis, since multivariate observations from many natural phenomena are normal distributed. Even when this is not the case, the mean of repeated samples become normal distributed due to the central limit theorem (the larger the sample size, the more accurate this approximation is). Hence, the multivariate normal distribution plays a central role in multivariate analysis. Indeed, many authors define classical multivariate analysis as the techniques, distribution and inferences based on the multivariate normal distribution [Mui82].

Most common multivariate techniques and inference procedures are based on the sample covariance matrix, defined as the covariance matrix of the observation vectors, i.e., of the columns of the sample observation matrix. A fundamental contribution to multivariate analysis was made by Wishart in 1928 [Wis28] by deriving the distribution of the sample covariance matrix when sampling from a multivariate normal population. The Wishart distribution and the proof itself is a generalization of the distribution of the two sample standard deviation and the correlation coefficient obtained by Fisher in 1915 [Fis15]. However, the geometrical methods used by Fisher and Wishart were not totally accepted in those days by the statistical community and the Wishart distribution was subsequently rederived in many different ways (see, e.g., [Ing33, Mad38, Hsu39a, Sve47, Ras48, Oga53, Olk54, Jam54, Ksh59], see [Wis48] for a review of some of these, or see [Gho02] for a simple proof).

---

[1] The assumption of independence of multivariate observations is not met in multivariate times series, stochastic processes, and repeated measurements on multivariate variables. In these cases, the matrix of observations leads to the introduction of the matrix variate normal distribution [Car83, Gup00].

Wishart considered in his original paper [Wis28] the distribution of the sample covariance matrix of observations from a zero-mean multivariate normal population, which is known as the central Wishart distribution. In the case of normal populations with nonzero means, the distribution is known as the noncentral Wishart distribution and was derived by James in 1955 [Jam55b, Jam55a, Jam61a].

All aforementioned results and references restrict to the case of real multivariate distributions. In fact, the multivariate complex normal distribution was introduced by Wooding in 1956 [Woo56], long time after its real conterpart[2] had been derived, after observing that it was advantageous for the treatment of certain estimation problems of the envelop of random noise signals. From this point on, owing to the importance of complex multivariate distributions in various emerging areas of research such as in the analysis of time series or stochastic processes (see, e.g., [Han70]) or in nuclear physics (see, e.g., [Wig67, Car83]), many results of multivariate analysis were extended to or directly derived for the complex case. In particular, the complex central and noncentral Wishart distribution were obtained by Goodman in 1963 [Goo63] and by James in 1964 [Jam64], respectively. In the rest of this chapter we deal exclusively with complex distributions, although not explicitly stated.

Closely related to the Wishart distribution is the distribution of its eigenvalues,[3] which is also of great interest for certain test statistics in multivariate analysis. The joint distribution of the eigenvalues of Wishart matrices were derived in [Fis39, Hsu39b, Roy39, Jam60, Jam61b, Dav80] for the real case and extended in [Jam64, Kha65, Rat05d] for the complex case. The main object of this chapter is to provide a solid framework for the obtention of currently unknown joint and marginal eigenvalue distributions as well as for the understanding and direct derivation of all the previously known results. For this purpose we focus on a general class of Hermitian random matrices which includes the Wishart and some other closely related distributions as particular cases.

[2] The distribution of several real normal variables was used for the first time by Bravais in 1846 [Bra46].

[3] The eigenvalues are referred to as latent roots in the multivariate analysis literature.

2.1.2 Connection between the Wishart Distribution and MIMO Channels

Multiple-input multiple-output (MIMO) channels are an abstract and general way to model many different communication systems of diverse physical nature; ranging from wireless multi-antenna channels [Fos96, Ral98, Fos98, Tel99], to wireline digital subscriber line (DSL) systems [Hon90], and to single-antenna frequency-selective channels [Sca99]. Assuming that the communication link has n_{T} transmit and n_{R} receive dimensions, the MIMO channel is mathematically described by an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ channel matrix $\mathbf{H}$, whose (i,j)th entry characterizes the path between the jth transmit and the ith receive element. In particular, when communicating over MIMO fading channels, $\mathbf{H}$ is a random matrix that depends on the particular system architecture and the particular propagation conditions. Hence, $\mathbf{H}$ is assumed to be drawn from a certain probability distribution, which characterizes the system and scenario of interest and is known as channel model. The system behavior is then evaluated on the average or outage sense, taking into account all possible channel states. Since the performance of MIMO systems is related to the eigenstructure of $\mathbf{H}$ (channel eigenmodes) or, more exactly, to the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ (or $\mathbf{H}^\dagger\mathbf{H}$), the probabilistic characterization of these eigenvalues for the adopted channel model becomes necessary.

In MIMO wireless communications $\mathbf{H}$ is commonly modeled with Gaussian distributed entries, leading to the MIMO generalization of the well-known single-input single-output (SISO) Rayleigh or Rician fading channels, depending on whether the entries are zero mean or not. Some important particular cases of the MIMO Rayleigh and Rician channel models result in $\mathbf{H}\mathbf{H}^\dagger$ (or $\mathbf{H}^\dagger\mathbf{H}$) being a Wishart random matrix. The Wishart distribution and some closely related distributions have been widely studied during the sixties and seventies in the mathematical literature (see[4] [Goo63, Jam64, Kha65, Sri65, Kha66, Tan68]). More recently, the statistical properties of the eigenvalues of Wishart matrices have been investigated and effectively applied to analyze the information theoretical limits of MIMO channels [Tel99, Gra02, Kan03a, Chi03,Smi03,Rat03,Alf04a,Alf04b,Jay05,McK05,Rat05d,Alf06,Kan06a,Kan06b,Sim06,Maa07b] as well as the performance of practical MIMO systems [Bur02, Dig03, Kan03b, Let04, Zan05, Chi05, Ord05b, Gar05, Maa06, Jin06, Jin08, Ord07b, Maa07a]. Some other interesting channel

[4] This bibliographical review is not exhaustive. We only include here some relevant references that focus on the complex Wishart distribution.

models entail the study of complex Pseudo-Wishart distributed matrices. However, this distribution and its eigenvalues have been only marginally considered in the MIMO literature [Smi03, Kan03a, Rat05b, Kan06b, Sim06, Maa06, Maa07a]. This is also the case of the more general complex quadratic form distributions [Shi03, Rat05c, Rat05a, Rat06, Shi06, Sim06, McK07], which include the Wishart and Pseudo-Wishart distributions as particular cases.

Most of these works deal with the joint pdf of the ordered eigenvalues [Kan03a, Chi03, Smi03, Rat03, Chi05, Rat05a, Rat05b, Rat05c, Rat05d, McK05, Kan06a, Kan06b, Rat06, Shi06], the marginal distribution of an unordered eigenvalue [Tel99, Gra02, Alf04a, Alf04b, Jay05, Alf06], or the distribution of the smallest eigenvalue [Bur02] to evaluate the system performance for the uninformed transmitter case. In contrast, when perfect channel state information is available at the transmitter, the weakest channel eigenmodes can be discarded, and the marginal statistics of the ordered eigenvalues become necessary to evaluate the system performance. In this context, useful closed-form expressions for the distribution of the largest eigenvalue have been derived in [Kan03a, Dig03, Kan03b, Let04, Zan05, Gra05, Maa05, Maa06, Maa07a] to analyze the performance of the beamforming scheme (also referred as maximum ratio transmission [Lo99]).

Nevertheless, an exhaustive analysis of the marginals of all ordered eigenvalues is still missing. Some initial contributions in this direction are [Ord05b, Jin06, Ord07b]. In particular, the first order Taylor expansion of the marginal pdf's of all the ordered eigenvalues was given in [Ord05b, Ord07b] for the uncorrelated central Wishart distribution to characterize the high-SNR performance of the individual MIMO channel eigenmodes and of linear MIMO transceivers (see e.g. [Pal03]). With the same purpose, [Jin06] derived the exact marginal cdf's and the first order Taylor expansion of the marginal pdf's of the ordered eigenvalues for the uncorrelated noncentral Wishart distribution.

2.1.3 Motivation and Contributions

Distribution results of random matrices are typically derived in the literature in terms of hypergeometric functions of matrix arguments. In particular, these functions arise in the distribution of Wishart matrices, as well as in the joint density of its eigenvalues. Hypergeometric functions of matrix arguments were defined by Herz in 1955 [Her55] using Laplace and inverse Laplace transforms. Constantine gave in 1963 [Con63] the power series representation in terms of an

infinite series of zonal polynomials, which are symmetric polynomials in the eigenvalues of the matrix arguments. The zonal polynomials series expasion became the standard representation of hypergeometric functions in multivariate analysis and, hence, appears when characterizing the Wishart distribution (see e.g. [Jam64]).

Unfortunately, the computational complexity of zonal polynomials is extremely high and the convergence of the series is often very slow [Gut00].[5] Hence, it is convenient to derive simpler and computationally efficient expressions avoiding zonal polynomials to represent hypergeometric functions of matrix arguments and, thus, the Wishart distribution and related results. The solution has been available in the literature since 1970, when Khatri gave in [Kha70] an alternative expression in terms of a quotient of determinants including generalized hypergeometric functions of scalar arguments. However, this result has been widely overlooked until it was independently derived by Gross and Richards in 1989 [Gro89].

In this chapter we present a general formulation that unifies the probabilistic characterization of the eigenvalues of Hermitian random matrices with a specific structure. Based on a unified expression for the joint pdf of the ordered eigenvalues, we obtain:

(i) the joint cdf of the ordered eigenvalues,
(ii) the marginal cdf's of the ordered eigenvalues,
(iii) the marginal pdf's of the ordered eigenvalues,
(iv) the cdf of the maximum weighted ordered eigenvalue, and
(v) the first order Taylor expansions of (ii), (iii), and (iv),

where (ii), (iii), and (iv) follow as simple particularizations of (i). In addition we also consider the unordered eigenvalues and derive:

(vi) the joint cdf of a set of unordered eigenvalues,
(vii) the joint pdf of a set of unordered eigenvalues, and
(viii) the marginal cdf of a single unordered eigenvalue.

Using Khatri's result, we particularize the derived distributions for uncorrelated and correlated central Wishart, correlated central Pseudo-Wishart, and uncorrelated noncentral Wishart matrices avoiding the non-convenient series expansions in terms of zonal polynomials. Complex

[5] An efficient algorithm for the computation of hypergeometric functions of matrix arguments has been recently presented in [Koe06].

central quadratic forms are also addressed, although, in this case, we are not able to handle with the infinite series. To the best of the author's knowledge, the joint cdf was unknown for all these distributions and the marginal cdf's and pdf's of all ordered eigenvalues were only available for the uncorrelated central and noncentral Wishart distributions. Recently, other unified treatments have been proposed in [Zan05, Alf06], including, however, only uncorrelated and correlated central Wishart and uncorrelated noncentral Wishart matrices. Furthermore, only the distribution of the largest and the smallest eigenvalue was derived in [Zan05] and the distribution of an unordered eigenvalue in both [Zan05, Alf06]. Simultaneously to the publication of this work in [Ord08b], the marginal cdf's of all the ordered eigenvalues were obtained in [Zan08] following the unified approach by the same authors in [Zan05] that includes uncorrelated and correlated central Wishart and uncorrelated noncentral Wishart matrices.

The joint analysis of a general class of distributions presented in this chapter settles the basis for a unified framework in the performance analysis of MIMO systems. Specifically, in this dissertation our results are applied to investigate the spatial multiplexing system that results from transmitting independent substreams through the strongest eigenmodes when perfect channel state information is available at both sides of the link (also termed as MIMO SVD systems). The motivation behind the analysis of this particular communication scheme is that it was proven to be optimal in the design of linear MIMO transceivers under a wide range of different optimization criteria [Pal03] (see Chapter 3 for details).

2.1.4 Outline

The rest of the chapter is organized as follows. Section 2.2 is devoted to introducing some mathematical definitions, results, and functions that are intensively used in the subsequent sections. In Section 2.3 we present the matrix variate complex normal distribution and some other closely related Hermitian matrix distributions, such as the Wishart, Pseudo-Wishart, and quadratic form distributions. Sections 2.4 and 2.5 contain the main contribution of this chapter, i.e., the derivations of the joint cdf and both the marginal cdf's and pdf's of the ordered and unordered eigenvalues of a general class of Hermitian random matrices. Then, in Section 2.6 we establish the matching between this class and some particular cases of the Wishart, Pseudo-Wishart, quadratic form distributions. Finally, in Section 2.7 we summarize the main results and provide the list of publications where they have been presented.

2.2 Mathematical Preliminaries and Definitions

This section presents preliminary definitions and lemmas that are useful when dealing with matrix-variate distributions and are required to follow the results and developments in the subsequent sections.

2.2.1 Matrices and Determinants

First we review some basic results on matrices and determinants.

Definition 2.1 (Sign of a permutation [Hor90, Sec. 0.3.2])**.** *The sign of a permutation* $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ *of the integers* $(1, \ldots, n)$*, denoted by* $\operatorname{sgn}(\boldsymbol{\mu})$*, is* $+1$ *or* -1 *according to whether the minimum number of transpositions, or pairwise interchanges, necessary to achieve* $(\mu_1, \ldots, \mu_n)$ *starting from* $(1, \ldots, n)$ *is even or odd.*

Definition 2.2 (Determinant [Hor90, Sec. 0.3.2])**.** *The determinant of matrix* $\mathbf{A}$ $(n \times n)$*, denoted by* $|\mathbf{A}|$*, is defined as*

$$|\mathbf{A}| = \sum_{\boldsymbol{\mu}} \operatorname{sgn}(\boldsymbol{\mu}) \prod_{k=1}^{n} [\mathbf{A}]_{\mu_k, k} = \operatorname{sgn}(\boldsymbol{\nu}) \sum_{\boldsymbol{\mu}} \operatorname{sgn}(\boldsymbol{\mu}) \prod_{k=1}^{n} [\mathbf{A}]_{\mu_k, \nu_k} \tag{2.1}$$

where the summation over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ *is for all permutations of the integers* $(1, \ldots, n)$*,* $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ *is any arbitrary fixed permutation of the integers* $(1, \ldots, n)$*, and* $\operatorname{sgn}(\cdot)$ *denotes the sign of the permutation.*

Alternatively, we also use the common compact notation of the determinant of matrix $\mathbf{A}$ in terms of its (i, j)th element, $[\mathbf{A}]_{i,j} = a_{ij}$,

$$|\mathbf{A}| = |a_{ij}|. \tag{2.2}$$

Of special interest in this chapter are Vandermonde matrices, since they admite a closed-form expression for their determinant that is used to manipulate the considered distributions.

Definition 2.3 (Vandermonde matrix [Hor91, eq. (6.1.32)])**.** *The* n*th order Vandermonde matrix in* $\mathbf{x} = (x_1, \ldots, x_n)$*, denoted by* $\mathbf{V}(\mathbf{x})$ $(n \times n)$*, is defined as*

$$[\mathbf{V}(\mathbf{x})]_{i,j} = x_j^{i-1} \qquad \text{for } i, j = 1, \ldots, n. \tag{2.3}$$

Lemma 2.1 (Vandermonde determinant [Hor91, eq. (6.1.33)])**.** *The determinant of the nth order Vandermonde matrix introduced in Definition* 2.3 *is given by*

$$|\mathbf{V}(\mathbf{x})| = \prod_{i<j}(x_j - x_i). \tag{2.4}$$

The following operator is useful to express a sum of determinants compactly.

Definition 2.4. *Operator* $\mathcal{T}\{\cdot\}$ *over a tensor* $\mathbf{T}$ $(n \times n \times n)$ *is defined as*[6]

$$\mathcal{T}\{\mathbf{T}\} = \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu})\mathrm{sgn}(\boldsymbol{\nu}) \prod_{k=1}^{n} [\mathbf{T}]_{\mu_k,\nu_k,k} \tag{2.5}$$

where the summation over $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ *and* $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ *is for all permutations of the integers* $(1, \ldots, n)$ *and* $\mathrm{sgn}(\cdot)$ *denotes the sign of the permutation.*

Remark 2.1. *Observe that the operator* $\mathcal{T}\{\cdot\}$ *introduced in Definition* 2.4 *can be alternatively expressed as*

$$\mathcal{T}\{\mathbf{T}\} = \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \prod_{k=1}^{n} [\mathbf{T}]_{\mu_k,k,\nu_k} = \sum_{\boldsymbol{\nu}} |\mathbf{A}(\boldsymbol{\nu})| \tag{2.6}$$

where matrix $\mathbf{A}(\boldsymbol{\nu})$ $(n \times n)$ *is defined in terms of the elements of tensor* $\mathbf{T}$ *as*

$$[\mathbf{A}(\boldsymbol{\nu})]_{i,j} = [\mathbf{T}]_{i,j,\nu_j} \qquad \text{for } i, j = 1, \ldots, n. \tag{2.7}$$

Lemma 2.2 (Derivative of a determinant [Hor91, eq. (6.5.9)])**.** *The derivative of the determinant of matrix* $\mathbf{A}(x)$ $(n \times n)$ *is given by*

$$\frac{d}{dx}|\mathbf{A}(x)| = \sum_{t=1}^{n} |\mathbf{A}^{(t)}(x)| \tag{2.8}$$

where $\mathbf{A}^{(t)}(x)$ *coincides with* $\mathbf{A}(x)$ *except that every entry in the t-th column is differentiated with respect to* x*.*

Lemma 2.3 (rth derivative of a determinant [Chr64, eq. (10)])**.** *The* r*th derivative of the determinant of matrix* $\mathbf{A}(x)$ $(n \times n)$ *is given by*

$$\frac{d^r}{dx^r}|\mathbf{A}(x)| = \sum_{\mathbf{r}} \frac{r!}{r_1! \cdots r_n!} |\mathbf{A}^{(\mathbf{r})}(x)| \tag{2.9}$$

where the summation over $\mathbf{r} = (r_1, \ldots, r_n)$ *is for all* $\mathbf{r}$ *such that* $r_i \in \mathbb{N} \cup \{0\}$ *and* $\sum_{i=1}^{n} r_i = r$*, and matrix* $\mathbf{A}^{(\mathbf{r})}(x)$ $(n \times n)$ *is defined as*

$$[\mathbf{A}^{(\mathbf{r})}(x)]_{i,j} = \frac{d^{r_j}}{dx^{r_j}} [\mathbf{A}(x)]_{i,j} \qquad \text{for } i, j = 1, \ldots, n. \tag{2.10}$$

[6] This operator is also introduced in [Chi03, Def. 1].

In the coming sections it is necessary to calculate limits of ratios of the form

$$\lim_{(x_{n+1},\ldots,x_m)\to(x_0,\ldots,x_0)} \frac{|f_i(x_j)|}{|\mathbf{V}(\mathbf{x})|} \tag{2.11}$$

where $\mathbf{x} = (x_1, \ldots, x_n, x_{n+1} \ldots, x_m)$ and $\mathbf{V}(\mathbf{x})$ is an mth order Vandermonde matrix. The required generalizations of the L'Hôpital rule are given in the next lemmas.

Lemma 2.4 ([Sim06, Lem. 6])**.** *The limit in* (2.11) *with* $x_0 < \infty$ *is given by*

$$\lim_{(x_{n+1},\ldots,x_m)\to(x_0,\ldots,x_0)} \frac{|f_i(x_j)|}{|\mathbf{V}(\mathbf{x})|} = \frac{|\mathbf{Z}(x_1,\ldots,x_n)|}{|\mathbf{V}(x_1,\ldots,x_n)| \prod_{i=1}^{n}(x_i - x_0)^{m-n} \prod_{j=1}^{m-n-1} j!} \tag{2.12}$$

where matrix $\mathbf{Z}(x_1, \ldots, x_n)$ $(m \times m)$ *is defined as*

$$[\mathbf{Z}(x_1,\ldots,x_n)]_{i,j} = \begin{cases} f_i(x_j) & 1 \le j \le n \\ f_i^{(j-n-1)}(x)\Big|_{x=x_0} & n < j \le m \end{cases} \qquad \text{for } i,j = 1,\ldots,m \tag{2.13}$$

where $f^{(r)}(\cdot)$ *denotes the* r*th derivative of function* $f(\cdot)$*.*

Lemma 2.5 ([Sim06, Lem. 7])**.** *The limit in* (2.11) *with* $x_0 = \infty$ *is given by*

$$\lim_{(x_{n+1},\ldots,x_m)\to(\infty,\ldots,\infty)} \frac{|f_i(x_j)|}{|\mathbf{V}(\mathbf{x})|} = \frac{|\widetilde{\mathbf{Z}}(x_1,\ldots,x_n)|}{|\mathbf{V}(x_1,\ldots,x_n)|} \tag{2.14}$$

where matrix $\widetilde{\mathbf{Z}}(x_1, \ldots, x_n)$ $(m \times m)$ *is defined as*

$$[\widetilde{\mathbf{Z}}(x_1,\ldots,x_n)]_{i,j} = \begin{cases} f_i(x_j) & 1 \le j \le n \\ \tilde{f}_i^{(j-n-1)} & n < j \le m \end{cases} \qquad \text{for } i,j = 1,\ldots,m \tag{2.15}$$

whenever the asymptotic behavior of $f_i(x)$ *as* $x \to \infty$ *is of the form*

$$f_i(x) = x^{m-1} \sum_{k=0}^{\infty} \tilde{f}_i^{(k)} x^{-k}. \tag{2.16}$$

2.2.2 Integral Functions

Now we introduce some functions defined in integral form that, due to its importance, have been tabulated and are available as build-in functions in most common mathematical software packages such as MATLAB® or Mathematica®. We also list some of their properties.

Definition 2.5 (Gamma function [Abr72, eq. (6.1.1)])**.** *The gamma function is defined as*

$$\Gamma(a) = \int_0^{\infty} e^{-x} x^{a-1} dx. \tag{2.17}$$

Definition 2.6 (Complex multivariate gamma function [Jam64, eq. (83)])**.** *The complex multivariate gamma function is defined as*

$$\tilde{\Gamma}_n(a) = \int_{\mathbf{A}=\mathbf{A}^\dagger>0} e^{-\mathrm{tr}(\mathbf{A})} |\mathbf{A}|^{a-n} d\mathbf{A} = \pi^{n(n-1)/2} \prod_{i=1}^{n} \Gamma(a-i+1). \tag{2.18}$$

Definition 2.7 (Lower incomplete gamma function [Abr72, eq. (6.5.2)])**.** *The lower incomplete gamma function is defined as*

$$\gamma(a,\lambda) = \int_0^\lambda e^{-x} x^{a-1} dx. \tag{2.19}$$

Definition 2.8 (Upper incomplete gamma function [Abr72, eq. (6.5.3)])**.** *The upper incomplete gamma function is defined as*

$$\Gamma(a,\lambda) = \int_\lambda^\infty e^{-x} x^{a-1} dx = \Gamma(a) - \gamma(a,\lambda). \tag{2.20}$$

Lemma 2.6. *Let $a \in \mathbb{N}$, then it holds that*

$$\Gamma(a) = (a-1)! \qquad \text{[Abr72, eq. (6.1.5)]} \tag{2.21}$$

$$\tilde{\Gamma}_n(a) = \pi^{n(n-1)/2} \prod_{i=1}^{n} (a-i)! \qquad \text{[from (2.18) and (2.21)]} \tag{2.22}$$

$$\gamma(a,\lambda) = \sum_{i=0}^{\infty} \frac{(-1)^i}{(a+i)i!} \lambda^{a+i} \qquad \text{[Abr72, eq. (6.5.29)]} \tag{2.23}$$

$$\Gamma(a,\lambda) = (a-1)! - \sum_{i=0}^{\infty} \frac{(-1)^i}{(a+i)i!} \lambda^{a+i} \qquad \text{[from (2.20), (2.21) and (2.23)].} \tag{2.24}$$

Definition 2.9 (Modified Bessel function of the first kind)**.** *The modified Bessel function of the first kind of integer order n is defined as* [Abr72, eq. (9.6.19) and eq. (9.6.10)]

$$I_n(\lambda) = \frac{1}{\pi} \int_0^\pi e^{\lambda \cos x} \cos(nx) dx = (\lambda/2)^n \sum_{k=0}^{\infty} \frac{(\lambda^2/4)^k}{k!\Gamma(n+k+1)} \tag{2.25}$$

Definition 2.10 (Generalized hypergeometric function [Gra00, eq. (9.14.1)])**.** *The generalized hypergeometric function is defined in terms of the hypergeometric series as*

$${}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;\lambda) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k \cdots (a_p)_k}{(b_1)_k (b_2)_k \cdots (b_q)_k} \frac{\lambda^k}{k!} \tag{2.26}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$ denotes the Pochhammer's symbol [Abr72, eq. (6.1.22)].

Lemma 2.7 (Special cases of hypergeometric functions)**.** *For the following generalized hypergeometric functions it holds that*

$${}_0F_0(\lambda) = e^\lambda \qquad \text{[from (2.26)]} \tag{2.27}$$

$${}_0F_1(n+1;\lambda) = \Gamma(n+1)\lambda^{-n/2} I_n(2\sqrt{\lambda}) \qquad \text{[Abr72, eq. (9.6.47)].} \tag{2.28}$$

Consequently, ${}_0F_0(\cdot)$ and ${}_0F_1(\cdot;\cdot)$ are known as exponential type and Bessel type hypergeometric functions, respectively.

Definition 2.11 (Marcum Q-function [Mar50] [Can87, eq. (1)])**.** *The (generalized) Marcum Q-function is defined as*

$$Q_n(a,b) = \frac{1}{a^{n-1}} \int_b^\infty x^n e^{-\left(\frac{x^2+a^2}{2}\right)} I_n(ax)dx. \tag{2.29}$$

Definition 2.12 (Nuttall Q-function [Nut72, eq. (86)])**.** *The Nuttall Q-function is defined as*

$$Q_{m,n}(a,b) = \int_b^\infty x^m e^{-\left(\frac{x^2+a^2}{2}\right)} I_n(ax)dx. \tag{2.30}$$

The Nuttall Q-function is not considered to be a tabulated function. However, if $m+n$ is odd, the Nuttall Q-function can be expressed as a weighted sum of $k+1$ generalized Marcum Q-functions and modified Bessel functions of the first kind as stated in following lemma.

Lemma 2.8 ([Sim02a, eq. (8)])**.** *For $m > n \geq 0$ and $m+n$ odd, i.e., $m = n+2k+1$ for $k = 0, 1, \ldots$, the Nuttall Q-function can be calculated as*

$$Q_{n+2k+1,n}(a,b) = \sum_{i=1}^{k+1} w_i(k) Q_{n+i}(a,b) + e^{\frac{a^2+b^2}{2}} \sum_{i=1}^{k} w_{k,i}(b^2) a^{i-1} b^{n+i+1} I_{n+i+1}(ab) \tag{2.31}$$

where

$$w_i(k) = 2^{k-i-j} \frac{k!}{(i-1)!} \binom{k+n}{k-i+1} \tag{2.32}$$

$$w_{k,i}(b^2) = \sum_{j=0}^{k-i} 2^{k-i+1} \frac{(k-1-j)!}{(i-1)!} \binom{k+n}{k-i-j} b^{2j} \tag{2.33}$$

where $\binom{n}{k} = n!/((n-k)!k!)$ denotes the binomial coefficient [Abr72, eq. (3.1.2)].

The definition of the Nuttall Q-function has been extended to encompass negative values of both parameters m and n and the corresponding series expansion has been given in [Maa08, App. II].

2.2.3 Functions of Matrix Arguments

Many multivariate distributions are described using hypergeometric functions of matrix arguments, which were introduced by Herz [Her55] in integral form and by Constantine [Con63] in

terms of infinite series of zonal polynomials (see [Gup00, Sec. 1.5 and 1.6] for a detailed explanation). In particular, we focus only on hypergeometric functions of Hermitian matrix arguments as defined by James in [Jam64, Sec. 8] (see also [Rat04], [Gra05, App. II], [McK06, Sec. 2.2]), since this is the class that arises when describing complex multivariate distributions.

2.2.3.1 Zonal Polynomials

Before coming to hypergeometric functions of matrix arguments it is convenient to give a brief introduction on zonal polynomials, since they are involved in the power series representation of hypergeometric functions which is the most common definition. Zonal polynomials are homogenous symmetric polynomials[7] in n variables derived from the group representation theory [Jam64, Sec. 4].

Definition 2.13 (Partition [Mui82, Sec. 7.2.1]). *A partition κ of k into n parts, where $k, n \in \mathbb{N}$, is defined as $\kappa = (k_1, \ldots, k_n)$ with $k_i \in \mathbb{N} \cup \{0\}$, $k_1 \geq \cdots \geq k_n$, and $\sum_{i=1}^{n} k_i = k$.*

Definition 2.14 (Zonal polynomial [Jam64, eq. (85)]). *The zonal polynomial of an Hermitian matrix $\mathbf{A}$ $(n \times n)$, denoted by $\tilde{C}_\kappa(\mathbf{A})$, is defined as*

$$\tilde{C}_\kappa(\mathbf{A}) = \chi_{[\kappa]}(1)\chi_{\{\kappa\}}(\mathbf{A}) \tag{2.34}$$

where κ is a partition of k into n parts, $\chi_{\{\kappa\}}(\mathbf{A})$ is the character representation $\{\kappa\}$ of the linear group and is given as a symmetric function of the eigenvalues of $\mathbf{A}$, $\lambda_1, \ldots, \lambda_n$ by

$$\chi_{\{\kappa\}}(\mathbf{A}) = \frac{\left|\lambda_i^{k_j+n-j}\right|}{\left|\lambda_i^{n-j}\right|} \tag{2.35}$$

and $\chi_{[\kappa]}(1)$ is the dimension of the representation $[\kappa]$ of the symmetric group:

$$\chi_{[\kappa]}(1) = k! \frac{\prod_{i<j}^{n}(k_i - k_j - i + j)}{\prod_{i=1}^{n}\Gamma(n + k_i - i + 1)}. \tag{2.36}$$

2.2.3.2 Hypergeometric Functions of Matrix Arguments

Definition 2.15 (Complex multivariate hypergeometric coefficient [Jam64, eq. (84)]). *Let κ be a partition of k into n parts. The complex multivariate hypergeometric coefficient, denoted by*

[7] A homogeneous symmetric polynomial of degree k in n variables $x_1, x_2, \ldots, x_n$ is a polynomial that is unchanged by permutations of the subscripts and every term in the polynomial has degree k [Mui82, Sec. 7.2.1, Rem. 1].

$[a]_\kappa$, *is defined as* [Jam64, eq. (84)]

$$[a]_\kappa = \prod_{i=1}^{n} (a-(i-1)/2)_{k_i} \tag{2.37}$$

where $(a)_k = a(a+1)\cdots(a+k-1)$ *denotes the Pochhammer's symbol* [Abr72, eq. (6.1.22)].

Definition 2.16 (Hypergeometric function of matrix argument [Jam64, eq. (87)])**.** *Let* $\mathbf{A}$ $(n\times n)$ *be an Hermitian matrix. The hypergeometric function of Hermitian matrix argument is defined in terms of zonal polynomials as*

$${}_p\tilde{F}_q\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{A}\right) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{[a_1]_\kappa\cdots[a_p]_\kappa}{[b_1]_\kappa\cdots[b_q]_\kappa}\frac{\tilde{C}_\kappa(\mathbf{A})}{k!} \tag{2.38}$$

where the summation over κ *is for all partitions of* k *into* n *parts and* $\tilde{C}_\kappa(\mathbf{A})$ *is the zonal polynomial of an Hermitian matrix* (*see Definition* 2.14).

Definition 2.17 (Hypergeometric function of two matrix arguments [Jam64, eq. (88)])**.** *Let* $\mathbf{A}$ $(n\times n)$ *and* $\mathbf{B}$ $(n\times n)$ *be two Hermitian matrices. The hypergeometric function of two Hermitian matrix arguments is defined in terms of zonal polynomials as*[8]

$${}_p\tilde{F}_q\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\mathbf{A},\mathbf{B}\right) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{[a_1]_\kappa\cdots[a_p]_\kappa}{[b_1]_\kappa\cdots[b_q]_\kappa}\frac{\tilde{C}_\kappa(\mathbf{A})\tilde{C}_\kappa(\mathbf{B})}{\tilde{C}_\kappa(\mathbf{I}_n)k!} \tag{2.39}$$

where the summation over κ *is for all partitions of* k *into* n *parts and* $\tilde{C}_\kappa(\mathbf{A})$ *is the zonal polynomial of an Hermitian matrix* (*see Definition* 2.14).

Remark 2.2. *In particular, the hypergeometric functions* ${}_0\tilde{F}_0(\mathbf{A})$ *and* ${}_0\tilde{F}_0(\mathbf{A},\mathbf{B})$, ${}_0\tilde{F}_1(\cdot;\mathbf{A})$ *and* ${}_0\tilde{F}_1(\cdot;\mathbf{A},\mathbf{B})$, *and* ${}_1\tilde{F}_1(\cdot;\cdot;\mathbf{A})$ *and* ${}_1\tilde{F}_1(\cdot;\cdot;\mathbf{A},\mathbf{B})$ *are referred to as exponential type, Bessel type, and confluent type, respectively.*

The convergence of the series in (2.38) and (2.39) (see convergence conditions in [Gup00, p. 34]) is often very slow. Hence, hypergeometric functions of matrix argument have acquired a reputation of being notoriously difficult to approximate even in the simplest cases [Gut00]. However, Koev and Edelman have recently proposed an algorithm in [Koe06] whose complexity is only linear in the size of the matrix.

[8] If matrices $\mathbf{A}$ and $\mathbf{B}$ have unequal dimensions, the hypergeometric function of matrix arguments $\mathbf{A}$ and $\mathbf{B}$ can be still analogously defined.

Lemma 2.9 ([Jam64, eq. (89)]). *Let* $\mathbf{A}$ $(n \times n)$ *be an Hermitian matrix. The exponential type hypergeometric function of Hermitian matrix argument* ${}_0\tilde{F}_0(\mathbf{A})$ *is given by*

$$ {}_0\tilde{F}_0(\mathbf{A}) = e^{\mathrm{tr}(\mathbf{A})}. \tag{2.40}$$

Lemma 2.10 ([Kha70, Lem. 3] [Gro89, Thm. 4.2]). *Let* $\mathbf{A}$ $(n\times n)$ *and* $\mathbf{B}$ $(n\times n)$ *be two Hermitian matrices with eigenvalues* $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ *and* $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$*, such that*[9] $(\lambda_1 > \cdots > \lambda_n)$ *and* $(\sigma_1 > \cdots > \sigma_n)$*. The hypergeometric function of two Hermitian matrix arguments given in Definition* 2.17 *in terms of zonal polynomials can be alternatively expressed as*

$$ {}_p\tilde{F}_q\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{A}, \mathbf{B}\right) = {}_pG_q \frac{|{}_p\mathbf{G}_q(\boldsymbol{\lambda}, \boldsymbol{\sigma})|}{|\mathbf{V}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\sigma})|} \tag{2.41}$$

with

$$ {}_pG_q = \frac{\tilde{\Gamma}_n(n)}{\pi^{n(n-1)/2}} \frac{\prod_{i=1}^{n}\prod_{j=1}^{q}(b_j - i + 1)^{i-1}}{\prod_{i=1}^{n}\prod_{j=1}^{p}(a_j - i + 1)^{i-1}} \tag{2.42}$$

where $\tilde{\Gamma}_n(\cdot)$ *denotes the complex multivariate gamma function (see Definition* 2.6*),* $\mathbf{V}(\boldsymbol{\lambda})$ $(n\times n)$ *and* $\mathbf{V}(\boldsymbol{\sigma})$ $(n \times n)$ *are Vandermonde matrices (see Definition* 2.3*), and* ${}_p\mathbf{G}_q(\boldsymbol{\lambda}, \boldsymbol{\sigma})$ *is defined as*

$$ [{}_p\mathbf{G}_q(\boldsymbol{\lambda}, \boldsymbol{\sigma})]_{i,j} = {}_pF_q\left(a_1 - n + 1, \ldots, a_p - n + 1; b_1 - n + 1, \ldots, b_q - n + 1; \lambda_i\sigma_j\right) \quad \text{for } i, j = 1, \ldots, n \tag{2.43}$$

where ${}_pF_q(\cdot;\cdot;\cdot)$ *is the generalized hypergeometric function of scalar arguments (see Definition* 2.10*). In particular, for the exponential and Bessel type it holds that*

$$ {}_0\tilde{F}_0(\mathbf{A}, \mathbf{B}) = \frac{\tilde{\Gamma}_n(n)}{\pi^{n(n-1)/2}} \frac{|{}_0\mathbf{G}_0(\boldsymbol{\lambda}, \boldsymbol{\sigma})|}{|\mathbf{V}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\sigma})|} \tag{2.44}$$

$$ {}_0\tilde{F}_1(m; \mathbf{A}, \mathbf{B}) = \frac{\tilde{\Gamma}_n(m)\tilde{\Gamma}_n(n)}{\pi^{n(n-1)}\Gamma(m - n + 1)^n} \frac{|{}_0\mathbf{G}_1(\boldsymbol{\lambda}, \boldsymbol{\sigma})|}{|\mathbf{V}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\sigma})|} \tag{2.45}$$

where ${}_0\mathbf{G}_0(\boldsymbol{\lambda}, \boldsymbol{\sigma})$ *and* ${}_0\mathbf{G}_1(\boldsymbol{\lambda}, \boldsymbol{\sigma})$ *are defined as*

$$ [{}_0\mathbf{G}_0(\boldsymbol{\lambda}, \boldsymbol{\sigma})]_{i,j} = {}_0F_0(\lambda_i\sigma_j) = e^{\lambda_i\sigma_j} \qquad \text{for } i, j = 1, \ldots, n \tag{2.46}$$

$$ [{}_0\mathbf{G}_1(\boldsymbol{\lambda}, \boldsymbol{\sigma})]_{i,j} = {}_0F_1(m - n + 1; \lambda_i\sigma_j) \qquad \text{for } i, j = 1, \ldots, n. \tag{2.47}$$

We have already given most of the results needed in this chapter. Several other results, in addition to these, are given in the text with the corresponding relevant references.

[9] When some of the λ_i's or σ_i's are equal, the result is obtained as the limiting case of the righthand side of (2.41) using Lemma 2.4.

2.3 Random Matrices Derived from the Complex Normal Distribution

This section introduces the matrix variate complex normal distribution and some other closely related Hermitian matrix distributions, such as Wishart, Pseudo-Wishart, and quadratic form distributions. Here we focus only on the density of these random matrices while the probabilistic characterization of its ordered and unordered eigenvalues is addressed in Section 2.6.

A matrix random phenomenon is an observable phenomenon which can be represented in matrix form. Repeated observations yield different outcomes of the observation matrix which are not deterministically predictable. Then, the degree of certainty with which a matrix event, i.e., a subset of the sample space $\mathcal{S}$, occurs can be measured by defining a function which assigns a probability to every matrix event in $\mathcal{S}$ according to the three postulates of Kolmogorov. Hence, the pdf of the random matrix is defined analogously to that of an scalar random variable.

Definition 2.18 (Probability density function of random matrix [Gup00, Def. 1.9.2])**.** *The pdf of a random matrix* $\mathbf{X}$ *is defined by a scalar function* $f_{\mathbf{X}}(\mathbf{X})$ *satisfying*

$$\text{(i)} \quad f_{\mathbf{X}}(\mathbf{X}) \geq 0 \tag{2.48}$$

$$\text{(ii)} \quad \int_{\mathbf{X}\in\mathcal{S}} f_{\mathbf{X}}(\mathbf{X})d\mathbf{X} = 1 \tag{2.49}$$

$$\text{(iii)} \quad \mathsf{Pr}(\mathbf{X} \in \mathcal{A}) = \int_{\mathbf{X}\in\mathcal{A}} f_{\mathbf{X}}(\mathbf{X})d\mathbf{X} \tag{2.50}$$

where $\mathcal{A}$ *is a subset of the sample space of* $\mathbf{X}$ *denoted by* $\mathcal{S}$*.*

2.3.1 Complex Normal Random Matrices

First we present basic definitions involving the generalization of the complex normal distribution to include random vectors and matrices.

Definition 2.19 (Complex normal random vector [Sri79, Sec. 2.1 and Def. 2.9.1])**.** *Let* $\mathbf{u} = (u_1, \ldots, u_n)^{'}$ *be a complex circular*[10] *random vector of* n *independent complex normal random variables with zero mean and unit variance, i.e.,* $u_i \sim \mathcal{CN}(0, 1)$*. Then, the random vector* $\mathbf{x} =$

[10] A complex random vector with real and imaginary parts $\mathbf{x}$ and $\mathbf{y}$, respectively, is circular [Nee93] (or proper [Pic94]) if $\mathbb{E}\{(\mathbf{x} - \mathbb{E}\{\mathbf{x}\})(\mathbf{y} - \mathbb{E}\{\mathbf{y}\})^{'}\} = \mathbf{0}$. Henceforth we always consider circular complex random vectors although not explicitly stated. In fact, the multivariate complex distribution was initially obtained in [Woo56] assuming circularity, and this assumption was maintained in the follow-up literature on the complex Wishart distribution.

$(x_1, \ldots, x_n)'$ *is said to have a multivariate complex normal distribution with mean vector* $\boldsymbol{\mu}$ $(n\times 1)$ *and covariance matrix* $\boldsymbol{\Sigma}$ $(n \times n)$*, denoted by* $\mathbf{x} \sim \mathcal{CN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, if* $\mathbf{x}$ *has the same distribution as* $\boldsymbol{\mu} + \mathbf{Au}$*, where* $\mathbf{A}$ *is any nonsingular factorization of* $\boldsymbol{\Sigma}$ *such that* $\boldsymbol{\Sigma} = \mathbf{AA}^\dagger$*.*

Definition 2.20 (Complex normal random matrix [Gup00, Def. 2.2.1][11])**.** *The random matrix* $\mathbf{X}$ $(n \times m)$ *is said to have a matrix variate complex normal distribution with mean matrix* $\boldsymbol{\Theta}$ $(n \times m)$ *and covariance matrix* $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$*, where* $\boldsymbol{\Sigma}$ $(n \times n) > 0$ *and* $\boldsymbol{\Psi}$ $(m \times m) > 0$*, denoted by* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\boldsymbol{\Theta}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$*, if* $\mathrm{vec}(\mathbf{X}') \sim \mathcal{CN}_{nm}(\mathrm{vec}(\boldsymbol{\Theta}'), \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$*.*

Lemma 2.11 (Matrix variate complex normal distribution [Woo56, eq. (17)] [Jam64, eq. (78)] [Kha66, eq. (56)] [Tan68, eq. (2.1)][12])**.** *Let* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\boldsymbol{\Theta}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$*. The pdf of* $\mathbf{X}$ *is given by*

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{\pi^{mn}|\boldsymbol{\Sigma}|^m|\boldsymbol{\Psi}|^n} e^{-\mathrm{tr}\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\Theta})\boldsymbol{\Psi}^{-1}(\mathbf{X}-\boldsymbol{\Theta})^\dagger}. \tag{2.51}$$

Observe that matrix $\boldsymbol{\Sigma}$ describes the covariance among the elements of any column of $\mathbf{X}$, i.e.,

$$\boldsymbol{\Sigma} = \mathbb{E}\big\{([\mathbf{X}]_t - [\boldsymbol{\Theta}]_t)([\mathbf{X}]_t - [\boldsymbol{\Theta}]_t)^\dagger\big\} \qquad \text{for } t = 1, \ldots, n \tag{2.52}$$

while $\boldsymbol{\Psi}$ describes the covariance among the elements of any row of $\mathbf{X}$, i.e.,

$$\boldsymbol{\Psi} = \mathbb{E}\big\{([\mathbf{X}']_t - [\boldsymbol{\Theta}']_t)([\mathbf{X}']_t - [\boldsymbol{\Theta}']_t)^\dagger\big\} \qquad \text{for } t = 1, \ldots, m. \tag{2.53}$$

Here we restrict to the case of $\boldsymbol{\Sigma} > 0$ and $\boldsymbol{\Psi} > 0$, i.e., $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi} > 0$. Otherwhise, the pdf of $\mathbf{X}$ in (2.51) does not exist and $\mathbf{X}$ is said to have a singular matrix variate complex normal distribution which is not considered here (see [Gup00, Sec. 2.4] for details).

2.3.2 Complex Wishart Random Matrices

Now we focus on the distribution of Hermitian random matrices of the form $\mathbf{W} = \mathbf{XX}^\dagger$ $(n \times n)$ with $\mathbf{X} \sim \mathcal{CN}_{n,m}(\boldsymbol{\Theta}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ for the particular case of $\mathbf{X}$ having independent columns, i.e., $\boldsymbol{\Psi} = \mathbf{I}_m$, and $n \leq m$. The random matrix $\mathbf{W}$ is said to follow the Wishart distribution, which was initially obtained in [Wis28] for the real case to describe the distribution of the sample covariance matrix of m independent observation vectors drawn from a normal population. This result was extended to the complex case in [Goo63, Jam64, Kha65].

[11] Reference [Gup00] gives the real counterpart of Definition 2.20.

[12] In [Woo56] $\boldsymbol{\Theta} = \mathbf{0}_n$ and $\boldsymbol{\Psi} = \mathbf{I}_m$, in [Jam64] $\boldsymbol{\Psi} = \mathbf{I}_m$, and in [Kha66] $\boldsymbol{\Theta} = \mathbf{0}_n$.

Definition 2.21 (Complex Wishart random matrix [Sri79, Sec. 3.1 and Sec. 3.7])**.** *Let* $\mathbf{X}$ $(n\times m)$ *with* $n \leq m$*, then the random matrix* $\mathbf{W}$ $(n \times n)$ *defined as* $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ *is said to have:*

(i) *a complex uncorrelated central Wishart distribution with* m *degrees of freedom, denoted by* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$*, if* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{0}_{n,m}, \mathbf{I}_n, \mathbf{I}_m)$.

(ii) *a complex correlated central Wishart distribution with* m *degrees of freedom and covariance matrix* $\mathbf{\Sigma}$ $(n \times n) > 0$*, denoted by* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$*, if* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{0}_{n,m}, \mathbf{\Sigma}, \mathbf{I}_m)$.

(iii) *a complex uncorrelated noncentral Wishart distribution with* m *degrees of freedom and noncentrality parameter matrix* $\mathbf{\Omega} = \mathbf{\Theta}\mathbf{\Theta}^\dagger$ $(n \times n) > 0$*, denoted by* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$*, if* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{\Theta}, \mathbf{I}_n, \mathbf{I}_m)$.

(iv) *a complex correlated noncentral Wishart distribution with* m *degrees of freedom, noncentrality parameter matrix* $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}\mathbf{\Theta}\mathbf{\Theta}^\dagger$ $(n \times n) > 0$*, and covariance matrix* $\mathbf{\Sigma}$ $(n \times n) > 0$*, denoted by* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$*, if* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{\Theta}, \mathbf{\Sigma}, \mathbf{I}_m)$.

It is worth pointing out that the mean of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ is given by

$$\mathbb{E}\{\mathbf{W}\} = m\mathbf{\Sigma} + \mathbf{\Theta}\mathbf{\Theta}^\dagger = \mathbf{\Sigma}(m\mathbf{I}_n + \mathbf{\Omega}) \tag{2.54}$$

and the covariance matrix is [Gup00, Cor. 7.7.2.1]

$$\begin{aligned}\mathbb{E}\big\{\big(\mathrm{vec}(\mathbf{W}) - \mathrm{vec}(\mathbb{E}\{\mathbf{W}\})\big)\big(\mathrm{vec}(\mathbf{W}) - \mathrm{vec}(\mathbb{E}\{\mathbf{W}\})\big)^\dagger\big\} = \\ = (m\mathbf{\Sigma}\otimes\mathbf{\Sigma} + \mathbf{\Sigma}\mathbf{\Omega}\otimes\mathbf{\Sigma} + \mathbf{\Sigma}\otimes\mathbf{\Sigma}\mathbf{\Omega})(\mathbf{I}_n + \mathbf{K}_{nn})\end{aligned} \tag{2.55}$$

where $\mathbf{K}_{nn}$ is the commutation matrix of order $n^2 \times n^2$ defined as [Mag79, Def. 3.1]

$$\mathbf{K}_{nn} = \sum_{i,j=1}^{n} \left(\mathbf{J}_{i,j} \otimes \mathbf{J}_{i,j}'\right) \tag{2.56}$$

where matrix $\mathbf{J}_{i,j}$ $(n \times n)$ has a unit element at the (i,j)th entry and zeros elsewhere.

From the pdf of the complex normal random matrix $\mathbf{X}$ and the definition $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$, it holds that $\mathbf{W} > 0$ with probability one [Gup00, Thm. 3.2.2]. Hence, the pdf of the complex Wishart matrix $\mathbf{W}$ is well defined in the set of all Hermitian positive definite matrices. This pdf is given in the following lemmas distinguishing between the central and noncentral cases (see [Sri65] for a simple derivation).

Lemma 2.12 (Complex central Wishart distribution [Goo63, eq. (1.6)] [Jam64, eq. (94)])**.** *The pdf of* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ *is given by*

$$f_{\mathbf{W}}(\mathbf{W}) = \frac{1}{\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m} e^{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{W}\right)} |\mathbf{W}|^{m-n} \tag{2.57}$$

where $\tilde{\Gamma}_n(\cdot)$ *denotes the complex multivariate gamma function* (*see Definition* 2.6).

Lemma 2.13 (Complex noncentral Wishart distribution [Jam64, eq. (99)])**.** *Let* $\mathbf{W}_c \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$. *The pdf of* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ *is given by*

$$f_{\mathbf{W}}(\mathbf{W}) = e^{-\mathrm{tr}(\mathbf{\Omega})}{}_0\tilde{F}_1(m; \mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{W}) f_{\mathbf{W}_c}(\mathbf{W}_c) \tag{2.58}$$

$$= \frac{e^{-\mathrm{tr}(\mathbf{\Omega})}}{\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m}{}_0\tilde{F}_1(m; \mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{W}) e^{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{W}\right)}|\mathbf{W}|^{m-n} \tag{2.59}$$

where $\tilde{\Gamma}_n(\cdot)$ *denotes the complex multivariate gamma function* (*see Definition* 2.6) *and* ${}_0\tilde{F}_1(\cdot;\cdot)$ *is the Bessel type hypergeometric function of Hermitian matrix argument* (*see Definition* 2.16).

2.3.3 Complex Pseudo-Wishart Random Matrices

In this section we consider the distribution of Hermitian random matrices of the form $\mathbf{W} = \mathbf{X}^\dagger\mathbf{X}$ $(m \times m)$ with $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{\Theta}, \mathbf{\Sigma}, \mathbf{\Psi})$ and $n < m$ for the particular case of $\mathbf{X}$ having independent rows, i.e., $\mathbf{\Sigma} = \mathbf{I}_n$. Recall from Section 2.3.2 that $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ $(n \times n)$ with $\mathbf{X}$ having independent columns satisfies $\mathbf{W} > 0$ with probability one and is Wishart distributed. To the contrary, in this case $\mathbf{W}$ is not full-rank and follows a Pseudo-Wishart distribution.[13]

In contrast to the Wishart distribution, the Pseudo-Wishart distribution has not been so extensively studied, since no practical applications were foreseen [Sri03]. Its distribution was obtained in [Uhl94, Gar97] for the real case and in [Mal03, Jan03, Rat05b] for the complex case.

Definition 2.22 (Complex Pseudo-Wishart random matrix [Sri79, Sec 3.1][14])**.** *Let* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{0}_n, \mathbf{I}_n, \mathbf{\Psi})$ *with* $n < m$. *The Hermitian random matrix* $\mathbf{W} = \mathbf{X}^\dagger\mathbf{X}$ $(m \times m)$ *is said to have a complex central Pseudo-Wishart distribution with parameters* n, m, *and covariance matrix* $\mathbf{\Psi}$ $(m \times m) > 0$, *denoted by* $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$.

Lemma 2.14 (Complex central Pseudo-Wishart distribution [Rat05b, Thm. 3])**.** *The pdf of* $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$ *is given by*

$$f_{\mathbf{W}}(\mathbf{W}) = \frac{\pi^{n(n-m)}}{\tilde{\Gamma}_m(m)|\mathbf{\Psi}|^m} e^{-\mathrm{tr}\left(\mathbf{\Psi}^{-1}\mathbf{W}\right)}|\mathbf{\Lambda}|^{m-n} \qquad \textit{for } \mathbf{W} \in \mathcal{CS}_{m,n} \tag{2.60}$$

---

[13] The Pseudo-Wishart distribution [Ksh72, Sri79, Gup00] is also referred to as singular Wishart [Uhl94] or anti-Wishart [Jan03] distribution. However, here we prefer the Pseudo-Wishart denomination to distinguish it from the case where the singularity of $\mathbf{W}$ is a direct consequence of $\mathbf{X}$ following a singular matrix variate normal distribution (see classification in [Sri79, Sec. 3.1]).

[14] Reference [Sri79] gives the real counterpart of Definition 2.22.

where $\tilde{\Gamma}_n(\cdot)$ *denotes the complex multivariate gamma function (see Definition* 2.6*),* $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *the contains the* n *nonzero eigenvalues of* $\mathbf{W}$*, and* $\mathcal{CS}_{m,n}$ *denotes the* $(2mn - n^2)$*-dimensional manifold of all* $m \times m$ *positive semidefinite Hermitian matrices of rank* n*.*

Observe that Definition 2.22 and Lemma 2.14 only contemplate the central case, i.e., the case of $\mathbf{X}$ being zero-mean. To the best of the author's knowledge, the distribution of complex noncentral Pseudo-Wishart matrices is not available in the literature.

2.3.4 Quadratic Forms in Complex Normal Matrices

In Sections 2.3.2 and 2.3.3, we consider random matrices of the form $\mathbf{W} = \mathbf{X}\mathbf{X}^\dagger$ or $\mathbf{X}^\dagger\mathbf{X}$ with $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{\Theta}, \mathbf{\Sigma}, \mathbf{\Psi})$ for the particular case of $\mathbf{X}$ having independent columns or rows, respectively. In this section we focus on the distribution of central quadratic forms $\mathbf{W} = \mathbf{X}\mathbf{A}\mathbf{X}^\dagger$, where $\mathbf{A}$ $(m \times m) > 0$ and $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{0}_{n,m}, \mathbf{\Sigma}, \mathbf{\Psi})$ with $\mathbf{\Sigma}$ $(n \times n)$ and $\mathbf{\Psi}$ $(m \times m)$ being any positive definite correlation matrices. Hence, it includes the former central random matrix distributions as particular cases. The distribution of $\mathbf{W}$ was derived by Khatri in 1966 [Kha66] simultaneously for the real and the complex case. The more general noncentral case in which $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{\Theta}, \mathbf{\Sigma}, \mathbf{\Psi})$ with $\mathbf{\Theta} \neq \mathbf{0}_{n,m}$ is mathematically more involved (see e.g. [Gup00, Sec. 7.6]) and is not addressed here.

Definition 2.23 (Complex central quadratic form distribution [Gup00, Sec 7.2][15])**.** *Let* $\mathbf{X}$ $(n \times m)$ *with* $n \leq m$*, then the random matrix* $\mathbf{W}$ $(n \times n)$ *defined as* $\mathbf{W} = \mathbf{X}\mathbf{A}\mathbf{X}^\dagger$ *is said to have central quadratic form distribution, denoted by* $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{0}_n, \mathbf{A}, \mathbf{\Sigma}, \mathbf{\Psi})$*, if* $\mathbf{X} \sim \mathcal{CN}_{n,m}(\mathbf{0}_{n,m}, \mathbf{\Sigma}, \mathbf{\Psi})$ *with* $\mathbf{\Sigma}$ $(n \times n) > 0$*,* $\mathbf{\Psi}$ $(m \times m) > 0$ *and* $\mathbf{A}$ $(m \times m) > 0$*.*

Lemma 2.15 (Complex central quadratic form distribution [Kha66, eq. (57)])**.** *The pdf of* $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{A}, \mathbf{\Sigma}, \mathbf{\Psi})$ *is given by*

$$f_{\mathbf{W}}(\mathbf{W}) = \frac{1}{\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m|\mathbf{A}\mathbf{\Psi}|^m} {}_0\tilde{F}_0(-\mathbf{W}\mathbf{\Sigma}^{-1}, \mathbf{\Psi}^{-1}\mathbf{A}^{-1})|\mathbf{W}|^{m-n} \tag{2.61}$$

where $\tilde{\Gamma}_n(\cdot)$ *denotes the complex multivariate gamma function (see Definition* 2.6*) and* ${}_0\tilde{F}_0(\cdot,\cdot)$ *is the exponential type hypergeometric function of two Hermitian matrix arguments (see Definition* 2.17*).*

[15] Reference [Gup00] gives the real counterpart of Definition 2.23.

2.4 Ordered Eigenvalues of a General Class of Hermitian Random Matrices

This section focuses on the ordered eigenvalues of a general class of Hermitian random matrix distributions which includes the Wishart distribution, the Pseudo-Wishart distribution, and the quadratic form distribution. Assuming a particular structure for the joint pdf of their ordered eigenvalues, we are able to perform the probabilistic characterization of the ordered eigenvalues simultaneously for all these distributions. More exactly, we derive the joint cdf and both the marginal cdf's and pdf's of the ordered eigenvalues of this general class of Hermitian random matrices. The corresponding first order Taylor expansions are also provided.

2.4.1 Joint pdf of the Ordered Eigenvalues

The joint pdf of the ordered eigenvalues of a complex Hermitian positive definite[16] random matrix is obtained from its pdf as shown in the next lemma.

Lemma 2.16 ([Jam64, eq. (93)])**.** *The joint pdf of the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of the complex Hermitian random matrix $\mathbf{W}$ $(n \times n) > 0$ with pdf $f_{\mathbf{W}}(\mathbf{W})$ is given by*

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)} \prod_{i<j} (\lambda_i - \lambda_j)^2 \int_{\mathcal{U}(n)} f_{\mathbf{W}}\left(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\dagger\right) d\mathbf{U} \tag{2.62}$$

where $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, $\mathbf{W} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\dagger$ is the eigendecomposition of $\mathbf{W}$, and $d\mathbf{U}$ is the invariant measure on the unitary group $\mathcal{U}(n)$ normalized to make the total measure unity.

Observing that the distribution of the Hermitian random matrices presented in Section 2.3 have all a very similar form and using Lemma 2.16, we can deduce how this common structure in the pdf of a random matrix is translated into a common expression for the joint pdf of its ordered eigenvalues. Following the investigations in Appendix 2.A.1, the general class of Hermitian random matrices addressed in this section (not necessarily positive definite) is formalized next in Assumption 2.1 by imposing a particular structure on the joint pdf of its nonzero ordered eigenvalues.

---

[16] We say that a random matrix is positive definite if it is positive definite with probability one, i.e., its eigenvalues are greater than zero with probability one. A similar procedure can be followed to deal with the nonzero eigenvalues of positive semidefinite Hermitian random matrix (see the case of the Pseudo-Wishart distribution in Appendix 2.A.1).

Assumption 2.1. *We consider the class of Hermitian random matrices, for which the joint pdf of its n nonzero ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, can be expressed as*

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = f_{\boldsymbol{\lambda}}(\lambda_1, \ldots, \lambda_n) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} |\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})| |\mathbf{V}(\boldsymbol{\lambda})| \prod_{t=1}^{n} \varphi(\lambda_t) \tag{2.63}$$

where $\boldsymbol{\iota}$ is a vector of indices and the summation is for all vectors $\boldsymbol{\iota}$ in the set $\mathcal{I}$, $\mathbf{V}(\boldsymbol{\lambda})$ $(n \times n)$ is a Vandermonde matrix (see Definition 2.3) and matrix $\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})$ $(n \times n)$ satisfies

$$[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \zeta_u^{(\boldsymbol{\iota})}(\lambda_v) \qquad \text{for } u, v = 1, \ldots, n. \tag{2.64}$$

The dimension of $\boldsymbol{\iota}$, the set $\mathcal{I}$, the constant $K_{m,n}^{(\boldsymbol{\iota})}$, and the functions $\zeta_u^{(\boldsymbol{\iota})}(\lambda)$ and $\varphi(\lambda)$ depend on the particular distribution of the random matrix.

2.4.2 Joint cdf of the Ordered Eigenvalues

Next we present the main theorem of this chapter, since most other distributions and results follow as (straightforward) particularizations of this one.

Theorem 2.1. *The joint cdf of the n nonzero ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of an Hermitian random matrix satisfying Assumption 2.1 is given by*

$$F_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \mathsf{Pr}\left(\lambda_1 \leq \eta_1, \ldots, \lambda_n \leq \eta_n\right) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{i} \in \mathcal{S}} \frac{1}{\tau(\boldsymbol{i})} \mathcal{T}\{\mathbf{T}^{(\boldsymbol{\iota})}(\boldsymbol{i}; \boldsymbol{\eta})\} \tag{2.65}$$

where $(\eta_1 \geq \cdots \geq \eta_n > 0)$,[17] the summation over $\boldsymbol{i} = (i_1, \ldots, i_n)$ is for all $\boldsymbol{i}$ in the set $\mathcal{S}$ and the normalization factor $\tau(\boldsymbol{i})$ are defined as[18]

$$\mathcal{S} = \{\boldsymbol{i} \in \mathbb{N}^n | \max(i_{s-1}, s) \leq i_s \leq n,\ i_s \neq r \text{ if } \eta_r = \eta_{r+1}\} \tag{2.66}$$

$$\tau(\boldsymbol{i}) = \prod_{u=1}^{n} \Big((1 - \delta_{i_u, i_{u+1}}) \sum_{v=1}^{u} \delta_{i_u, i_v} \Big)! \tag{2.67}$$

where $\delta_{u,v}$ denotes de Kronecker delta. Operator $\mathcal{T}\{\cdot\}$ is introduced in Definition 2.4, tensor $\mathbf{T}^{(\boldsymbol{\iota})}(\boldsymbol{i}; \boldsymbol{\eta})$ $(n \times n \times n)$ is defined as

$$[\mathbf{T}^{(\boldsymbol{\iota})}(\boldsymbol{i}; \boldsymbol{\eta})]_{u,v,t} = \int_{\eta_{i_t+1}}^{\eta_{i_t}} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda \qquad \text{for } u, v, t = 1, \ldots, n \tag{2.68}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda) \varphi(\lambda) \lambda^{v-1}$ (see Assumption 2.1).

Proof. See Appendix 2.A.2.

[17] If $\eta_{k-1} < \eta_k$ then $F_{\boldsymbol{\lambda}}(\eta_1, \ldots, \eta_{k-1}, \eta_k, \ldots, \eta_n) = F_{\boldsymbol{\lambda}}(\eta_1, \ldots, \eta_{k-1}, \eta_{k-1}, \ldots, \eta_n)$ and if some $\eta_k = 0$, then $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) = 0$.

[18] Note that $i_n = n$ and by definition $i_0 = 0$, $i_{n+1} = n + 1$ and $\eta_{n+1} = 0$.

2.4.3 Marginal cdf and pdf of the kth Largest Ordered Eigenvalue

In the following result we particularize the joint cdf of the ordered eigenvalues given in Theorem 2.1 to derive the marginal cdf and the marginal pdf of the kth largest eigenvalue.

Theorem 2.2. *The marginal cdf of the kth largest nonzero eigenvalue, λ_k $(1 \leq k \leq n)$, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$F_{\lambda_k}(\eta) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{i=1}^{k} \sum_{\boldsymbol{\mu} \in \mathcal{P}(i)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)| \tag{2.69}$$

where $\mathcal{P}(i)$ is the set of all permutations $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ of the integers $(1, \ldots, n)$ such that $(\mu_1 < \ldots < \mu_{i-1})$ and $(\mu_i < \ldots < \mu_n)$, matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)$ $(n \times n)$ is defined as

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)]_{u,v} = \begin{cases} \int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & 1 \leq \mu_v < i \\ \int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & i \leq \mu_v \leq n \end{cases} \qquad \text{for } u, v = 1, \ldots, n \tag{2.70}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda) \varphi(\lambda) \lambda^{v-1}$ (see Assumption 2.1*).*

Proof. See Appendix 2.B.1.

Theorem 2.2 can be further simplified when dealing with the marginal cdf of the largest and smallest eigenvalues.

Corollary 2.2.1. *The marginal cdf of the largest eigenvalue, λ_1, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$F_{\lambda_1}(\eta) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} |\mathbf{F}^{(\boldsymbol{\iota})}(\eta)| \tag{2.71}$$

where matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\eta)$ $(n \times n)$ is defined as

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\eta)]_{u,v} = \int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda \qquad \text{for } u, v = 1, \ldots, n \tag{2.72}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda) \varphi(\lambda) \lambda^{v-1}$ (see Assumption 2.1*).*

Proof. See Appendix 2.B.2.

Corollary 2.2.2. *The marginal cdf of the smallest nonzero eigenvalue, λ_n, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$F_{\lambda_n}(\eta) = 1 - \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} |\mathbf{F}^{(\boldsymbol{\iota})}(\eta)| \tag{2.73}$$

where matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\eta)$ $(n \times n)$ is defined as

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\eta)]_{u,v} = \int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda \qquad \text{for } u, v = 1, \ldots, n \tag{2.74}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda)\varphi(\lambda)\lambda^{v-1}$ (see Assumption 2.1*).*

Proof. See Appendix 2.B.3.

Similarly, the marginal pdf of the kth largest eigenvalue can be easily derived from Theorem 2.2 as we illustrate in the following corollary.

Corollary 2.2.3. *The marginal pdf of the kth largest nonzero eigenvalue, λ_k $(1 \leq k \leq n)$, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$f_{\lambda_k}(\eta) = \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{i=1}^{k} \sum_{\boldsymbol{\mu}\in\mathcal{P}(i)} \sum_{t=1}^{n} |\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i, t; \eta)| \tag{2.75}$$

where $\mathcal{P}(i)$ is the set of all permutations $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ of the integers $(1, \ldots, n)$ such that $(\mu_1 < \ldots < \mu_{i-1})$ and $(\mu_i < \ldots < \mu_n)$, matrix $\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i, t; \eta)$ $(n \times n)$ is defined as

$$[\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i, t; \eta)]_{u,v} = \begin{cases} \int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & 1 \leq \mu_v < i, v \neq t \\ -\xi_{u,v}^{(\boldsymbol{\iota})}(\eta) & 1 \leq \mu_v < i, v = t \\ \int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & i \leq \mu_v \leq n, v \neq t \\ \xi_{u,v}^{(\boldsymbol{\iota})}(\eta) & i \leq \mu_v \leq n, v = t \end{cases} \qquad \text{for } u, v = 1, \ldots, n \tag{2.76}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda)\varphi(\lambda)\lambda^{v-1}$ (see Assumption 2.1*).*

Proof. See Appendix 2.B.4.

Observe that the simplifications used in the proofs of Corollaries 2.2.1 and 2.2.2 can be straightforwardly applied to obtain simpler expressions for the marginal pdf of the largest and smallest eigenvalues than the ones given in Corollary 2.2.3.

2.4.4 Marginal cdf of the Maximum Weighted Ordered Eigenvalue

Based on the joint cdf of the ordered eigenvalues given in Theorem 2.1, we can easily obtain the cdf of the maximum weighted eigenvalue (out of a subset $\mathcal{K}$ of the ordered eigenvalues).

Theorem 2.3. *Let us define the random variable $\lambda_{\mathcal{K}}$ as*

$$\lambda_{\mathcal{K}} = \max_{k\in\mathcal{K}} \lambda_k/\theta_k \tag{2.77}$$

where λ_k is the kth largest eigenvalue of an Hermitian random matrix satisfying Assumption 2.1, the set $\mathcal{K}$ of cardinality $|\mathcal{K}|$ is such that $\{1\} \subseteq \mathcal{K} \subseteq \{1,\ldots,n\}$, and $\{\theta_k\}_{k\in\mathcal{K}}$ are some given positive constants $(\theta_1 \geq \cdots \geq \theta_{|\mathcal{K}|})$.[19] *Then, the cdf of $\lambda_{\mathcal{K}}$ is given by*

$$F_{\lambda_{\mathcal{K}}}(\eta) = F_{\boldsymbol{\lambda}}(\boldsymbol{\vartheta}(\eta)) \tag{2.78}$$

where $\boldsymbol{\vartheta}(\eta) = (\vartheta_1(\eta),\ldots,\vartheta_n(\eta))$ with

$$\vartheta_k(\eta) = \vartheta_k \cdot \eta = \begin{cases} \theta_k\eta & k \in \mathcal{K} \\ \vartheta_{k-1}(\eta) & k \notin \mathcal{K} \end{cases} \qquad \text{for } k = 1,\ldots,n \tag{2.79}$$

and $F_{\boldsymbol{\lambda}}(\cdot)$ is the joint cdf of the ordered eigenvalues given in Theorem 2.1.

Proof. See Appendix 2.C

2.4.5 Taylor Expansions

In the previous sections we have focused on the general class of Hermitian random matrices defined in Assumption 2.1. Here, we concentrate on a more restrictive class (as formalized in Assumption 2.2) but general enough to include the Wishart distribution.

Assumption 2.2. *We consider the class of Hermitian random matrices, for which the joint pdf of its n nonzero ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, can be expressed as*

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = f_{\boldsymbol{\lambda}}(\lambda_1,\ldots,\lambda_n) = K_{m,n}|\mathbf{E}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\lambda})|\prod_{t=1}^{n}\varphi(\lambda_t) \tag{2.80}$$

---

[19] Note that if $\{\theta_k\}_{k\in\mathcal{K}}$ are strictly decreasing, then $\lambda_{\mathcal{K}} = \lambda_1$.

where $\mathbf{V}(\boldsymbol{\lambda})$ $(n \times n)$ *is a Vandermonde matrix (see Definition* 2.3*) and matrix* $\mathbf{E}(\boldsymbol{\lambda})$ $(n \times n)$ *satisfies*

$$[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \zeta_u(\lambda_v) \qquad \text{for } u, v = 1, \ldots, n. \tag{2.81}$$

The constant $K_{m,n}$ *and the functions* $\zeta_u(\lambda)$ *and* $\varphi(\lambda)$ *depend on the particular distribution of the random matrix.*

Assuming that the Taylor expansion[20] of the function $\zeta_u(\lambda)\varphi(\lambda)$ is known and satisfies some mild conditions (see Assumption 2.3), we derive in the following the first order Taylor expansions of both the cdf and pdf of the kth largest eigenvalue and of the maximum weighted eigenvalue.

Assumption 2.3. *Let the Taylor expansion of the function* $\zeta_u(\lambda)\varphi(\lambda)$ *(see Assumption* 2.2*) be*

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c(u)}^{\infty} \frac{a_u(t)}{t!}\lambda^t \tag{2.83}$$

where $a_u(t)$ *is such that* $a_u(t) = 0$ *for* $t < c(u)$ *and let matrix* $\mathbf{M}$ $(n \times n)$*, defined as*

$$[\mathbf{M}]_{u,v} = \begin{cases} a_u(c+v) & 1 \le v \le n-k+1 \\ b_{u,v} & n-k+1 < v \le n \end{cases} \qquad \text{for } u = 1, \ldots, n \tag{2.84}$$

have nonzero determinant, where $c = \min_u c(u)$ *and*

$$b_{u,v} = \int_0^\infty \xi_{u,v}(\lambda)d\lambda = \int_0^\infty \zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1}d\lambda. \tag{2.85}$$

Theorem 2.4. *Under Assumption* 2.3*, the first order Taylor expansions of the marginal cdf and the marginal pdf of the* k*th largest nonzero eigenvalue,* λ_k $(1 \le k \le n)$*, of an Hermitian random matrix satisfying Assumption* 2.2 *are given by*

$$F_{\lambda_k}(\eta) = \Big(\frac{a_k}{d_k+1}\Big)\eta^{d_k+1} + o\big(\eta^{d_k+1}\big) \tag{2.86}$$

$$f_{\lambda_k}(\eta) = a_k\eta^{d_k} + o\big(\eta^{d_k}\big) \tag{2.87}$$

---

[20] Recall that the Taylor expansion of a function $f(x)$ around a point x_0 is [Abr72, eq. (25.2.24)]

$$f(x) = \sum_{t=0}^{\infty} \frac{f^{(t)}(x)|_{x=x_0}}{t!}(x-x_0)^t = \sum_{t=0}^{r} \frac{f^{(t)}(x)|_{x=x_0}}{t!}(x-x_0)^t + o\left((x-x_0)^r\right) \tag{2.82}$$

where $f^{(t)}(x)$ denotes the t-th derivative of $f(x)$ and we say that $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to 0$ [Bru81, eq. (1.3.1)].

with a_k and d_k defined as

$$a_k = K_{m,n}(d_k+1)\sum_{\boldsymbol{\nu}} \frac{1}{\tau(\boldsymbol{\nu})}|\mathbf{F}(\boldsymbol{\nu})| \tag{2.88}$$

$$d_k = (c-n-k-1)(n-k+1)-1 \tag{2.89}$$

where the summation over $\boldsymbol{\nu}=(\nu_1,\ldots,\nu_{n-k+1})$ is for all permutations of the integers $(1,\ldots,n-k+1)$, $\tau(\boldsymbol{\nu})$ is

$$\tau(\boldsymbol{\nu}) = \prod_{v=1}^{n-k+1}(c+v+\nu_v-1)! \tag{2.90}$$

and matrix $\mathbf{F}(\boldsymbol{\nu})$ $(n\times n)$ is defined as

$$[\mathbf{F}(\boldsymbol{\nu})]_{u,v} = \begin{cases} \frac{(c+v+\nu_v-2)!}{(c+\nu_v-1)!}a_u(c+\nu_v-1) & 1\le v\le n-k+1 \\ b_{u,v} & n-k+1<v\le n \end{cases} \qquad \text{for } u,v=1,\ldots,n. \tag{2.91}$$

Proof. See Appendix 2.D.1

Theorem 2.5. *Under Assumption 2.3, the first order Taylor expansions of the cdf and the pdf of the random variable $\lambda_{\mathcal{K}}$ introduced in Theorem 2.3 are given by*

$$F_{\lambda_{\mathcal{K}}}(\eta) = \Big(\frac{a_{\mathcal{K}}}{d_{\mathcal{K}}+1}\Big)\eta^{d_{\mathcal{K}}+1} + o\big(\eta^{d_{\mathcal{K}}+1}\big) \tag{2.92}$$

$$f_{\lambda_{\mathcal{K}}}(\eta) = a_{\mathcal{K}}\eta^{d_{\mathcal{K}}} + o\big(\eta^{d_{\mathcal{K}}}\big) \tag{2.93}$$

with $a_{\mathcal{K}}$ and $d_{\mathcal{K}}$ defined as

$$a_{\mathcal{K}} = K_{m,n}\sum_{\boldsymbol{i}\in\mathcal{S}}\frac{d_{\mathcal{K}}+1}{\tau(\boldsymbol{i})}\sum_{\boldsymbol{\nu}}\mathcal{T}\{\mathbf{T}(\boldsymbol{\nu},\boldsymbol{i};\boldsymbol{\vartheta})\} \tag{2.94}$$

$$d_{\mathcal{K}} = (c-n)n \tag{2.95}$$

where the summation over $\boldsymbol{i}\in\mathcal{S}$ and $\tau(\boldsymbol{i})$ are defined as in Theorem 2.3, the summation over $\boldsymbol{\nu}=(\nu_1,\ldots,\nu_n)$ is for all permutations of integers $(1,\ldots,n)$, tensor $\mathbf{T}(\boldsymbol{\nu},\boldsymbol{i};\boldsymbol{\vartheta})$ $(n\times n\times n)$ is defined as

$$[\mathbf{T}(\boldsymbol{\nu},\boldsymbol{i};\boldsymbol{\vartheta})]_{u,v,t} = \frac{1}{c(\nu_v+v)c(v)!}a_u(c(v))\big(\vartheta_{i_t}^{c(\nu_v+v)} - \vartheta_{i_t+1}^{c(\nu_v+v)}\big) \qquad \text{for } u,v,t=1,\ldots,n \tag{2.96}$$

and $c(v)=c+v-1$.

Proof. See Appendix 2.D.2

2.5 Unordered Eigenvalues of a General Class of Hermitian Random Matrices

In Section 2.4 we focus on the probabilistic characterization of the ordered eigenvalues of a general class of Hermitian random matrices formalized in Assumption 2.1. Altenatively, in this section we disregard the order and consider the probabilistic characterization of a set of or a single unordered eigenvalue.

2.5.1 Joint pdf and cdf of the Unordered Eigenvalues

Imposing the structure in Assumption 2.1 on the joint pdf of the ordered eigenvalues, the joint pdf of the corresponding unordered eigenvalues in the next theorem follows.

Theorem 2.6. *The joint pdf of the n nonzero unordered eigenvalues, $x_1, \ldots, x_n$, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$f_{\mathbf{x}}(\mathbf{x}) = f_{\mathbf{x}}(x_1, \ldots, x_n) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{n!} |\mathbf{E}^{(\boldsymbol{\iota})}(\mathbf{x})| |\mathbf{V}(\mathbf{x})| \prod_{t=1}^{n} \varphi(x_t) \tag{2.97}$$

where $\boldsymbol{\iota}$ is a vector of indices and the summation is for all vectors $\boldsymbol{\iota}$ in the set $\mathcal{I}$, $\mathbf{V}(\mathbf{x})$ $(n \times n)$ is a Vandermonde matrix (see Definition 2.3*) and matrix $\mathbf{E}^{(\boldsymbol{\iota})}(\mathbf{x})$ $(n \times n)$ satisfies*

$$[\mathbf{E}(\mathbf{x})]_{u,v} = \zeta_u^{(\boldsymbol{\iota})}(x_v) \qquad \text{for } u, v = 1, \ldots, n. \tag{2.98}$$

The dimension of $\boldsymbol{\iota}$, the set $\mathcal{I}$, the constant $K_{m,n}^{(\boldsymbol{\iota})}$, and the functions $\zeta_u^{(\boldsymbol{\iota})}(x)$ and $\varphi(x)$ depend on the particular distribution of the random matrix.

Proof. See Appendix 2.E.1

Theorem 2.7. *The joint cdf of a subset of p $(1 \leq p \leq n)$ arbitrary unordered eigenvalues, $\mathbf{x}_p = (x_1, \ldots, x_p)$, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$F_{\mathbf{x}_p}(\boldsymbol{\eta}) = F_{\mathbf{x}_p}(\eta_1, \ldots, \eta_p) = \binom{n}{p}^{-1} \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{\mu} \in \mathcal{P}(p+1)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| \tag{2.99}$$

where $\binom{n}{p}$ denotes the binomial coefficient [Abr72, eq. (3.1.2)]*, $\mathcal{P}(p)$ is the set of all permutations $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ of the integers $(1, \ldots, n)$ such that $(\mu_1 < \ldots < \mu_{p-1})$ and $(\mu_p < \ldots < \mu_n)$,*

matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})$ $(n \times n)$ *is defined as*

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})]_{u,v} = \begin{cases} \int_0^{\eta_v} \xi_{u,\mu_v}(x)dx & 1 \leq v \leq p \\ \int_0^{\infty} \xi_{u,\mu_v}(x)dx & p < v \leq n \end{cases} \quad \text{for } u, v = 1, \ldots, n \tag{2.100}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(x) = \zeta_u^{(\boldsymbol{\iota})}(x)\varphi(x)x^{v-1}$ (*see Assumption* 2.1).

Proof. See Appendix 2.E.2

Similarly to Theorem 2.7, we can also obtain the joint pdf of p arbitrary unordered eigenvalues.

Corollary 2.7.1. *The joint pdf of a subset of* p $(1 \leq p \leq n)$ *unordered eigenvalues,* $\mathbf{x}_p = (x_1, \ldots, x_p)$*, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$f_{\mathbf{x}_p}(\boldsymbol{\eta}) = f_{\mathbf{x}_p}(\eta_1, \ldots, \eta_p) = \binom{n}{p}^{-1} \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{\mu} \in \mathcal{P}(p+1)} |\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| \tag{2.101}$$

where $\binom{n}{p}$ *denotes the binomial coefficient* [Abr72, eq. (3.1.2)], $\mathcal{P}(p)$ *is the set of all permutations* $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ *of the integers* $(1, \ldots, n)$ *such that* $(\mu_1 < \ldots < \mu_{p-1})$ *and* $(\mu_p < \ldots < \mu_n)$*, matrix* $\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})$ $(n \times n)$ *is defined as*

$$[\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})]_{u,v} = \begin{cases} \xi_{u,\mu_v}(\eta_v) & 1 \leq v \leq p \\ \int_0^{\infty} \xi_{u,\mu_v}(x)dx & p < v \leq n \end{cases} \quad \text{for } u, v = 1, \ldots, n \tag{2.102}$$

and $\xi_{u,v}^{(\boldsymbol{\iota})}(x) = \zeta_u^{(\boldsymbol{\iota})}(x)\varphi(x)x^{v-1}$ (*see Assumption* 2.1).

Proof. See Appendix 2.E.3

2.5.2 Marginal cdf and pdf of the Unordered Eigenvalues

Observe that the marginal cdf or pdf of an arbitrary unordered eigenvalue can be easily derived by particularizing Theorem 2.7 or Corollary 2.7.1 for $p = 1$. However if the focus is on an arbitrary unordered eigenvalue x_q chosen from the set of the $1 \leq q \leq n$ largest eigenvalues, $\{\lambda_1, \ldots, \lambda_q\}$, its marginal cdf can be obtained using the marginal cdf of the kth largest eigenvalues for $k = 1, \ldots, q$ given in Theorem 2.2 as we show next. The marginal pdf of x_q can be also analogously calculated using the marginal pdf of the kth largest eigenvalues for $k = 1, \ldots, q$ given in Corollary 2.2.3.

Theorem 2.8. *The marginal cdf and pdf of an unordered eigenvalue x_q chosen from the set of the $1 \leq q \leq n$ largest eigenvalues, $\{\lambda_1, \ldots, \lambda_q\}$, of an Hermitian random matrix satisfying Assumption* 2.1 *is given by*

$$F_{x_q}(\eta) = \frac{1}{q} \sum_{k=1}^{q} F_{\lambda_k}(\eta) \tag{2.103}$$

$$f_{x_q}(\eta) = \frac{1}{q} \sum_{k=1}^{q} f_{\lambda_k}(\eta) \tag{2.104}$$

where $F_{\lambda_k}(\eta)$ and $f_{\lambda_k}(\eta)$ denote the marginal cdf and pdf of the kth largest eigenvalue given in Theorem 2.2 *and Corollary* 2.2.3*, respectively.*

2.6 Eigenvalues of General Class of Hermitian Random Matrices: Particular Cases

In this section we particularize the results in Sections 2.4 and 2.5 regarding the probabilitic characterization of the ordered and unordered eigenvalues for the Wishart, Pseudo-Wishart, and quadratic form distributions introduced in Section 2.3. This requires rewriting first the joint pdf of the orderered eigenvalues in the form of Assumption 2.1 and then calculating the corresponding expressions (see Table 2.1 for details).

2.6.1 Complex Uncorrelated Central Wishart Matrices

The joint pdf of the ordered eigenvalues of a real uncorrelated central Wishart matrix was simultaneously derived in 1939 in [Fis39], [Hsu39b] and [Roy39]. Then, based on the concepts[21] introduced in 1956 by Wooding [Woo56] to extend the real normal multivariate distribution to the complex case, and in 1963 by Goodman [Goo63], who, in addition, obtained the distribution of complex central Wishart matrices, James derived in 1964 [Jam64] the joint pdf of the ordered eigenvalues of complex Wishart matrices for cases (i), (ii), and (iii) of Definition 2.21.

Joint pdf: The joint pdf of the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of $\mathbf{W} \sim$

[21] A review of the required results of [Woo56, Goo63] is given in [Jam64].

Result	Distribution	Required Parameters
Ass. 2.1	Joint pdf of the ordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\zeta_u^{(\boldsymbol{\iota})}(\lambda)$, $\varphi(\lambda)$.
Thm. 2.1	Joint cdf of the ordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\int_{\eta_{i_t+1}}^{\eta_{i_t}} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Thm. 2.2	Marginal cdf of the ordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$, $\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Cor. 2.2.3	Marginal pdf of the ordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Thm. 2.4	First order Taylor expansions	$a_u(t)$, $b_{u,v}$, $c(u)$.
Thm. 2.7	Joint cdf of a set of p unordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$, $\int_{0}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Cor. 2.7.1	Joint pdf of a set of p unordered eigenvalues	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$, $\int_{0}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Thm. 2.8	Marginal cdf of an unordered eigenvalue (chosen from the q largest eigenvalues)	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$, $\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.
Thm. 2.8	Marginal pdf of an unordered eigenvalue (chosen from the q largest eigenvalues)	$\mathcal{I}$, $K_{m,n}^{(\boldsymbol{\iota})}$, $\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$, $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$, $\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$.

Table 2.1 List of general distributional results and the parameters required to particularize each result for a given random matrix distribution.

$\mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ (case (i) of Definition 2.21) is given by[22] [Jam64, eq. (95)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)} {}_0\tilde{F}_0(\mathbf{W})|\mathbf{W}|^{m-n} \prod_{i<j}(\lambda_i - \lambda_j)^2 \tag{2.105}$$

where $\tilde{\Gamma}_n(\cdot)$ is the complex multivariate gamma function (see Definition 2.6) and ${}_0\tilde{F}_0(\cdot)$ is the exponential type hypergeometric function of Hermitian matrix argument (see Lemma 2.9). Using Lemmas 2.1 and 2.9, we can rewrite (2.105) as [Kha65, eq. (7.1.7)] [Chi03, eq. (10)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)} |\mathbf{V}(\boldsymbol{\lambda})|^2 \prod_{i=1}^{n} e^{-\lambda_i}\lambda_i^{m-n} \tag{2.106}$$

where $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3). Identifying terms, the expression in (2.106) coincides with the general pdf given in Assumption 2.1 if we let $\mathcal{I}$ be a singleton (the

[22] Note the abuse of notation in writting ${}_0\tilde{F}_0(\mathbf{W})$ instead of ${}_0\tilde{F}_0(\mathrm{diag}\,(\boldsymbol{\lambda}))$ and $|\mathbf{W}|$ instead of $\prod_{i=1}^{n}\lambda_i$. A similar observation also holds for the remaining matrix distributions addressed in this section.

superindex ($\boldsymbol{\iota}$) can then be dropped), define the normalization constant $K_{m,n}$ as

$$K_{m,n} = \frac{\pi^{-n(n-1)}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)} = \prod_{i=1}^{n} \frac{1}{(m-i)!(n-i)!} \tag{2.107}$$

the function $\varphi(\lambda)$ as

$$\varphi(\lambda) = e^{-\lambda}\lambda^{m-n} \tag{2.108}$$

and matrix $\mathbf{E}(\boldsymbol{\lambda})$ $(n \times n)$, equal to $\mathbf{V}(\boldsymbol{\lambda})$, with entries given by

$$[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \zeta_u(\lambda_v) = \lambda_v^{u-1} \qquad \text{for } u, v = 1, \ldots, n. \tag{2.109}$$

Hence, it follows that

$$\xi_{u,v}(\lambda) = \zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1} = e^{-\lambda}\lambda^{d(u+v-1)} \tag{2.110}$$

where we have defined the function $d(v) = m - n + v - 1$.

Results Regarding the Ordered Eigenvalues: In order to derive the marginal cdf and pdf of the kth largest eigenvalue using the results presented in Section 2.4, we only have to particularize

$$\int_{\eta}^{\infty} \xi_{u,v}(\lambda)d\lambda = \Gamma(d(u+v), \eta) \tag{2.111}$$

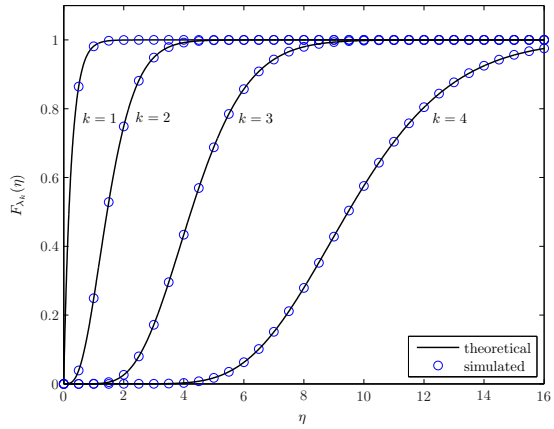
$$\int_{0}^{\eta} \xi_{u,v}(\lambda)d\lambda = \gamma(d(u+v), \eta) \tag{2.112}$$

where $\gamma(\cdot,\cdot)$ and $\Gamma(\cdot,\cdot)$ are the lower and upper incomplete gamma functions given in Definitions 2.7 and 2.8, respectively. The integrals in (2.111) and (2.112) can be conveniently combined to obtain the integrals needed in the computation of the joint cdf of the ordered eigenvalues.

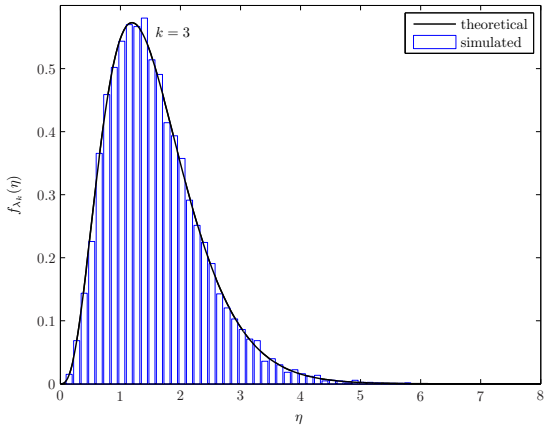
To the best of the author's knowledge the joint cdf of the ordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ was not available in the literature. The marginal cdf of the largest eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ was initially derived in [Kha64, Thm. 2] and extended to the marginal cdf of the kth largest eigenvalue in [AA68, eq. (16)]. Recently, the marginal cdf's of the largest and smallest eigenvalue were obtained in [Dig03, eq. (18)] [Kan03b, Cor. 2] [Let04, eq. (6)] [Gra05, Thm. 5] [Zan05, eq. (5)] and [Bur02, eq. (38)], respectively. Furthermore, the marginal pdf's of the largest and smallest eigenvalue were provided in [Dig03, eq. (22)] [Kan03b, Cor. 3] [Let04, eq. (7)] [Zan05, eq. (23)] and [Zan05, eq. (24)], respectively.

Results Regarding the Unordered Eigenvalues: The results concerning the unordered eigenvalues in Section 2.5 only require to calculate

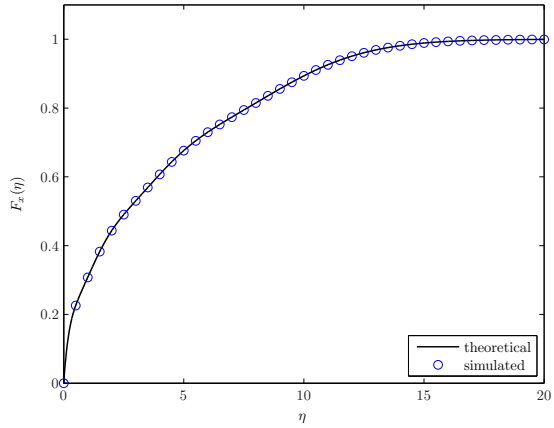
$$\int_{0}^{\infty} \xi_{u,v}(\lambda)d\lambda = \Gamma(d(u+v)) = d(u+v-1)! \tag{2.113}$$



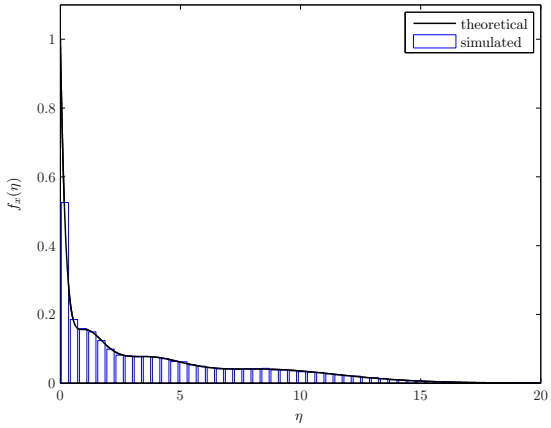
(i) marginal cdf of the kth ordered eigenvalue



(ii) marginal pdf of the kth ordered eigenvalue



(iii) marginal cdf of an unordered eigenvalue



(iv) marginal pdf of an unordered eigenvalue

Figure 2.1 Theoretical and simulated results for $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ with $n = m = 4$.

where $\Gamma(\cdot)$ is the gamma function in Definition 2.5, in addition to the integral $\int_0^\eta \xi_{u,v}(\lambda)d\lambda$, which is already given in (2.112).

The joint pdf of a set of unordered eigenvalues and the marginal pdf of an arbitrary unordered eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ were initially given in [Wig65, eq. (29)] and [Wig65, eq. (30)]. Recently, the joint cdf and pdf of a set of unordered eigenvalues were derived in [Maa07b, Cor. 3] and the marginal pdf of an arbitrary unordered eigenvalue in [Tel99, Sec 4.2] [Sca02a, Cor. 1] [Alf06, Thm. 1].

In Figure 2.1 we compare some examples of the marginal cdf's and pdf's of the ordered and unordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ obtained using the results in Sections 2.4 and 2.5 with the corresponding numerical simulations.

Taylor Expansions: Finally, in order to derive the first order Taylor expansion of the marginal cdf and pdf of the kth largest eigenvalue, we have to calculate the Taylor expansion of $\zeta_u(\lambda)\varphi(\lambda)$. Using the Taylor expansion of $e^{-\lambda}$ (see [Abr72, eq. (4.2.1)]), it follows that

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c(u)}^{\infty} \frac{a_u(t)}{t!}\lambda^t \tag{2.114}$$

where the function $c(u) = m - n + u - 1$, $c = \min_u c(u) = m - n$ and

$$a_u(t) = \frac{t!}{(t - c(u))!}(-1)^{t-c(u)}. \tag{2.115}$$

The expression for $b_{u,v}$ is given in (2.113).

The first order Taylor expansion of the marginal pdf of the kth largest eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ was recently obtained in [Ord05b, Thm. 1] [Ord07b, Thm. 1].

2.6.2 Complex Correlated Central Wishart Matrices

The joint pdf of the ordered eigenvalues of a real correlated central Wishart matrix was derived in 1960 by James [Jam60] for the real case and extended in 1964 [Jam64] to the complex case.

Joint pdf: The joint pdf of the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ (case (ii) of Definition 2.21) is given by [Jam64, eq. (95)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m} {}_0\tilde{F}_0(-\mathbf{\Sigma}^{-1}, \mathbf{W})|\mathbf{W}|^{m-n} \prod_{i<j}(\lambda_i - \lambda_j)^2 \tag{2.116}$$

where $\tilde{\Gamma}_n(\cdot)$ is the complex multivariate gamma function (see Definition 2.6) and ${}_0\tilde{F}_0(\cdot,\cdot)$ is the exponential type hypergeometric function of two Hermitian matrix arguments (see Lemma 2.10). Denoting by $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ the eigenvalues of $\mathbf{\Sigma}$ and using (2.44) in Lemma 2.10, it follows that for[23] $\sigma_1 > \cdots > \sigma_n > 0$ we can rewrite (2.116) as [Chi03, eq. (17)]

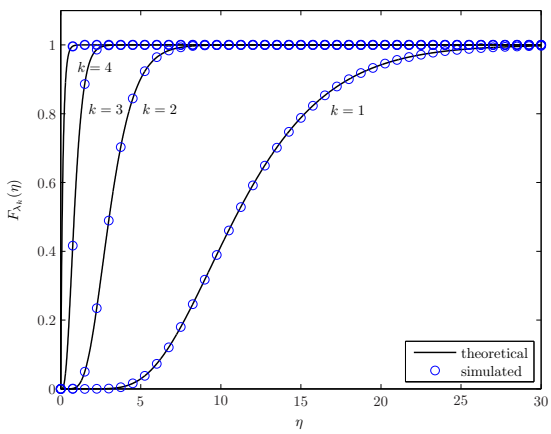
$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)/2}}{\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m|\mathbf{V}(-\boldsymbol{\sigma}^{-1})|}|\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})||\mathbf{V}(\boldsymbol{\lambda})| \prod_{i=1}^{n} \lambda_i^{m-n} \tag{2.117}$$

where $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3) and $\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})$ is defined as

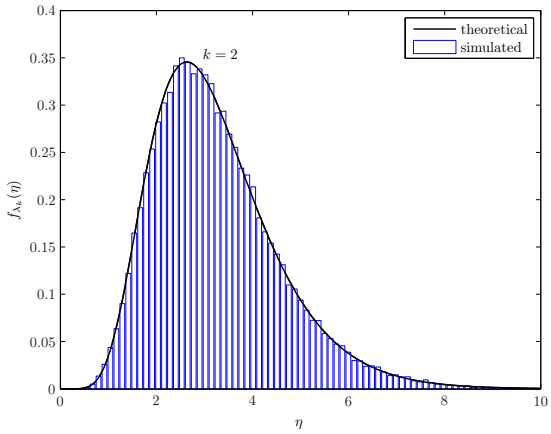
$$[\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})]_{u,v} = e^{-\lambda_v/\sigma_u} \qquad \text{for } u, v = 1, \ldots, n. \tag{2.118}$$

Then, identifying terms, the expression in (2.117) coincides with the general pdf given in Assumption 2.1 if we let $\mathcal{I}$ be a singleton (the superindex $(\boldsymbol{\iota})$ can then be dropped), define the

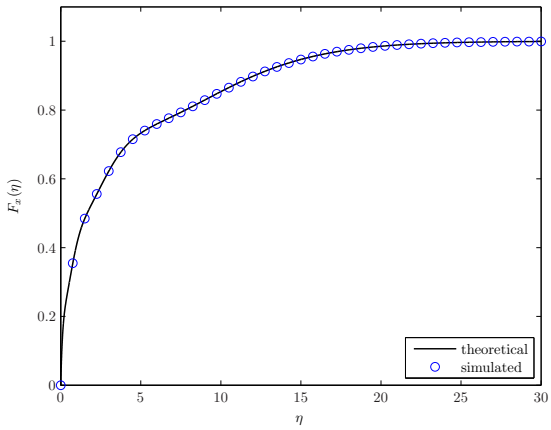
[23] If some σ_i's are equal the result is obtained by taking the limiting case of (2.117) using Lemma 2.4.



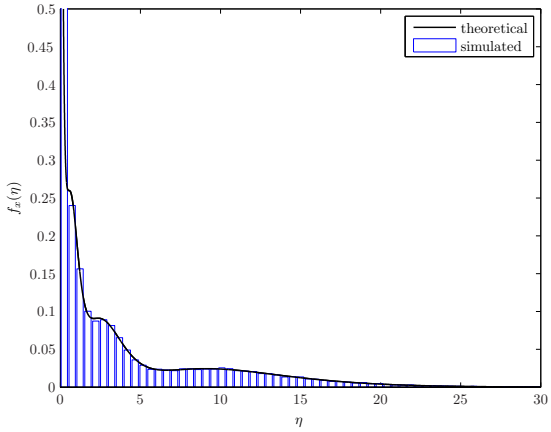
(i) marginal cdf of the kth ordered eigenvalue



(ii) marginal pdf of the kth ordered eigenvalue



(iii) marginal cdf of an unordered eigenvalue



(iv) marginal pdf of an unordered eigenvalue

Figure 2.2 Theoretical and simulated results for $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ with $n = m = 4$, and $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$ $(n \times n)$, $r = 0.6$.

normalization constant $K_{m,n}$ as

$$K_{m,n} = \frac{\pi^{n(n-1)/2}}{\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m|\mathbf{V}(-\boldsymbol{\sigma}^{-1})|} = \prod_{i=1}^{n} \frac{1}{\sigma_i^m (m-i)!} \prod_{i<j}^{n} \frac{\sigma_i \sigma_j}{(\sigma_j - \sigma_i)} \tag{2.119}$$

the function $\varphi(\lambda)$ as

$$\varphi(\lambda) = \lambda^{m-n} \tag{2.120}$$

and matrix $\mathbf{E}(\boldsymbol{\lambda})$ with entries $[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \zeta_u(\lambda)$ as given in (2.118). Hence, it follows that

$$\xi_{u,v}(\lambda) = \zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1} = e^{-\lambda/\sigma_u}\lambda^{d(v)} \tag{2.121}$$

where we have defined the function $d(v) = m - n + v - 1$.

Results Regarding the Ordered Eigenvalues: In order to derive the marginal cdf and pdf of the kth largest eigenvalue using the results presented in Section 2.4, we only have to

	$\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$	$\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$
$\mathcal{I}$	$\{1\}$	$\{1\}$
$K_{m,n}^{(\boldsymbol{\iota})}$	$\prod_{i=1}^{n} \frac{1}{(m-i)!(n-i)!}$	$\prod_{i=1}^{n} \frac{1}{\sigma_i^m (m-i)!} \prod_{i<j}^{n} \frac{\sigma_i \sigma_j}{\sigma_j - \sigma_i}$
$\varphi(\lambda)$	$e^{-\lambda}\lambda^{m-n}$	λ^{m-n}
$\zeta_u^{(\boldsymbol{\iota})}(\lambda)$	λ^{u-1}	$e^{-\lambda/\sigma_u}$
$\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$	$e^{-\lambda}\lambda^{d(u+v-1)}$	$e^{-\lambda/\sigma_u}\lambda^{d(v)}$
$\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$\Gamma(d(u+v), \eta)$	$\sigma_u^{d(v+1)}\Gamma(d(v+1), \eta/\sigma_u)$
$\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$\gamma(d(u+v), \eta)$	$\sigma_u^{d(v+1)}\gamma(d(v+1), \eta/\sigma_u)$
$\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$\gamma(d(u+v), \eta)$	$\sigma_u^{d(v+1)}\gamma(d(v+1), \eta/\sigma_u)$
$\int_{0}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$d(u+v-1)!$	$\sigma_u^{(c+v)}d(v)!$
$a_u(v)$	$\frac{v!}{(v-d(u))!}(-1)^{(v-d(u))}, \quad v \geq d(u)$	$\frac{v!}{(v-c)!}(-\sigma_u)^{c-t}, \quad v \geq c$
c	$m-n$	$m-n$
$d(v)$	$m-n+v-1$	$m-n+v-1$

Table 2.2 Parameters characterizing the eigenvalue distributions of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$ and $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ (case (i) and (ii) of Definition 2.21).

particularize

$$\int_{\eta}^{\infty} \xi_{u,v}(\lambda)d\lambda = \sigma_u^{d(v+1)}\Gamma(d(v+1), \eta/\sigma_u) \tag{2.122}$$

$$\int_{0}^{\eta} \xi_{u,v}(\lambda)d\lambda = \sigma_u^{d(v+1)}\gamma(d(v+1), \eta/\sigma_u) \tag{2.123}$$

where $\gamma(\cdot,\cdot)$ and $\Gamma(\cdot,\cdot)$ are the lower and upper incomplete gamma functions given in Definitions 2.7 and 2.8, respectively. The integrals in (2.122) and (2.123) can be conveniently combined to obtain the integrals needed in the computation of the joint cdf of the ordered eigenvalues.

The marginal cdf's of the largest and smallest eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ were recently derived in [Kan03a, Thm. 4.(1)] [Zan05, eq. (7)] [Maa06, eq. (9)] and in [Zan05, eq. (9)] [Maa06, eq. (13)], respectively. The corresponding marginal pdf's were obtained in [Zan05, eq. (26)] [Maa06, eq. (17)] and in [Maa06, eq. (18)]. To the best of the author's knowledge, the joint cdf of the ordered eigenvalues, the marginal cdf and pdf of the kth largest eigenvalue were not available in the literature.

Results Regarding the Unordered Eigenvalues: The results concerning the unordered

eigenvalues in Section 2.5 only require to calculate

$$\int_0^\infty \xi_{u,v}(\lambda)d\lambda = \sigma_u^{d(v+1)}\Gamma(d(v+1)) = \sigma_u^{d(v+1)}d(v)! \tag{2.124}$$

where $\Gamma(\cdot)$ is the gamma function in Definition 2.5, in addition to the integral $\int_0^\eta \xi_{u,v}(\lambda)d\lambda$, which is already given in (2.123).

The joint pdf of a set unordered eigenvalues and the marginal pdf of an arbitrary unordered eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ were initially given in [Wai72]. Recently, the joint cdf and pdf of a set of unordered eigenvalues were derived in [Maa06, eq. (3)] and [Maa06, eq. (4)] and the marginal pdf of an arbitrary unordered eigenvalue in [Alf04b, Thm. 1] [Alf06, Thm. 1] [Maa06, eq. (13)].

In Figure 2.2 we compare some examples of the marginal cdf's and pdf's of the ordered and unordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$ obtained using the results in Sections 2.4 and 2.5 with the corresponding numerical simulations.

Taylor Expansions: Finally, in order to derive the first order Taylor expansion of marginal cdf and pdf of the kth largest eigenvalue, we have to calculate the Taylor expansion of $\zeta_u(\lambda)\varphi(\lambda)$. Using the Taylor expansion of the $e^{-\lambda}$ (see [Abr72, eq. (4.2.1)]), it follows that

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c}^{\infty} \frac{a_u(t)}{t!}\lambda^t \tag{2.125}$$

where $c(u) = c = m - n$ and

$$a_{u,v}(t) = \frac{t!}{(t-c)!}(-\sigma_u)^{c-t}. \tag{2.126}$$

The expression for $b_{u,v}$ is given in (2.124).

2.6.3 Complex Uncorrelated Noncentral Wishart Matrices

The joint pdf of the ordered eigenvalues of a uncorrelated noncentral Wishart matrix was derived by James in [Jam61b] for the real case and extended in [Jam64] to the complex case.

Joint pdf: The joint pdf of the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ (case (iii) of Definition 2.21) is given by [Jam64, eq. (102)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{\pi^{n(n-1)}e^{-\mathrm{tr}(\mathbf{\Omega})}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)} {}_0\tilde{F}_1(m; \mathbf{\Omega}, \mathbf{W})e^{-\mathrm{tr}(\mathbf{W})}|\mathbf{W}|^{m-n}\prod_{i<j}(\lambda_i - \lambda_j)^2 \tag{2.127}$$

where $\tilde{\Gamma}_n(\cdot)$ is the complex multivariate gamma function (see Definition 2.6) and ${}_0\tilde{F}_1(\cdot;\cdot,\cdot)$ is the Bessel type hypergeometric function of two Hermitian matrix arguments (see Lemma 2.10). Denoting by $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)$ the eigenvalues of $\boldsymbol{\Omega}$ and using (2.45) in Lemma 2.10, we have that for[24] $(\omega_1 > \cdots > \omega_n > 0)$ we can rewrite (2.127) as [Kan03b, eq. (45)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{e^{-\mathrm{tr}(\boldsymbol{\Omega})}}{\Gamma(m-n+1)^n |\mathbf{V}(\boldsymbol{\omega})|} |\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\omega})| |\mathbf{V}(\boldsymbol{\lambda})| \prod_{i=1}^{n} e^{-\lambda_i} \lambda_i^{m-n} \tag{2.128}$$

where $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3) and $\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\omega})$ is defined as

$$[\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\omega})]_{u,v} = {}_0F_1(m-n+1; \omega_u \lambda_v) \qquad \text{for } u, v = 1, \ldots, n \tag{2.129}$$

where ${}_0F_1(\cdot;\cdot)$ is a generalized hypergeometric function (see Definition 2.10). Then, identifying terms, the expression in (2.128) coincides with the general pdf given in Assumption 2.1 if we let $\mathcal{I}$ be a singleton (the superindex $(\boldsymbol{\iota})$ can then be dropped), define the normalization constant $K_{m,n}$ as

$$K_{m,n} = \frac{e^{-\mathrm{tr}(\boldsymbol{\Omega})}}{\Gamma(m-n+1)^n |\mathbf{V}(\boldsymbol{\omega})|} = \frac{e^{-\sum_{i=1}^{n} \omega_i}}{((m-n)!)^n} \prod_{i<j}^{n} \frac{1}{(\omega_j - \omega_i)} \tag{2.130}$$

the function $\varphi(\lambda)$ as

$$\varphi(\lambda) = e^{-\lambda} \lambda^{m-n} \tag{2.131}$$

and matrix $\mathbf{E}(\boldsymbol{\lambda})$ with entries $[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \zeta_u(\lambda)$ as given in (2.129). Hence, it follows that

$$\xi_{u,v}(\lambda) = \zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1} = {}_0F_1(m-n+1; \omega_u\lambda) e^{-\lambda} \lambda^{d(v)} \tag{2.132}$$

where the function $d(v) = m - n + v - 1$.

Results Regarding the Ordered Eigenvalues: In order to derive the marginal cdf and pdf of the kth largest eigenvalue using the results presented in Section 2.4, we only have to particularize

$$\int_0^{\eta} \xi_{u,v}(\lambda) d\lambda = \int_0^{\eta} {}_0F_1(m-n+1; \omega_u\lambda) e^{-\lambda} \lambda^{d(v)} d\lambda \tag{2.133}$$

$$\int_{\eta}^{\infty} \xi_{u,v}(\lambda) d\lambda = \int_{\eta}^{\infty} {}_0F_1(m-n+1; \omega_u\lambda) e^{-\lambda} \lambda^{d(v)} d\lambda. \tag{2.134}$$

Using Lemma 2.7 (eq. (2.28)), it holds that

$$\int_{\eta}^{\infty} \xi_{u,v}(\lambda) d\lambda = \frac{\Gamma(m-n+1)}{\left(\sqrt{\omega_u}\right)^{m-n}} \int_{\eta}^{\infty} e^{-\lambda} \lambda^{(m-n)/2+v-1} I_{m-n}(2\sqrt{\omega_u \lambda}) d\lambda \tag{2.135}$$

$$= \frac{e^{\omega_u} 2^{1-v} (m-n)!}{\left(\sqrt{2\omega_u}\right)^{m-n}} Q_{d(2v),m-n}\left(\sqrt{2\omega_u}, \sqrt{2\eta}\right) \tag{2.136}$$

[24] If some ω_i's are equal the result is obtained by taking the limiting case of (2.128) using Lemma 2.4 (see [Kan03b, App. B]).

where $Q_{m,n}(\cdot,\cdot)$ is the Nuttall Q-function given in Definition 2.12. Similarly, using [Gra00, eq. (6.643.2)] and [Gra00, eq. (9.220.2)], it follows that

$$\int_0^\infty \xi_{u,v}(\lambda)d\lambda = \frac{\Gamma(m-n+1)}{\left(\sqrt{\omega_u}\right)^{m-n}} \int_0^\infty e^{-\lambda}\lambda^{(m-n)/2+v-1} I_{m-n}(2\sqrt{\omega_u\lambda})d\lambda \tag{2.137}$$

$$= \Gamma(d(v+1))_1F_1(d(v+1); m-n+1; \omega_u) \tag{2.138}$$

and we have that

$$\int_0^\eta \xi_{u,v}(\lambda)d\lambda = \int_0^\infty \xi_{u,v}(\lambda)d\lambda - \int_\eta^\infty \xi_{u,v}(\lambda)d\lambda \tag{2.139}$$

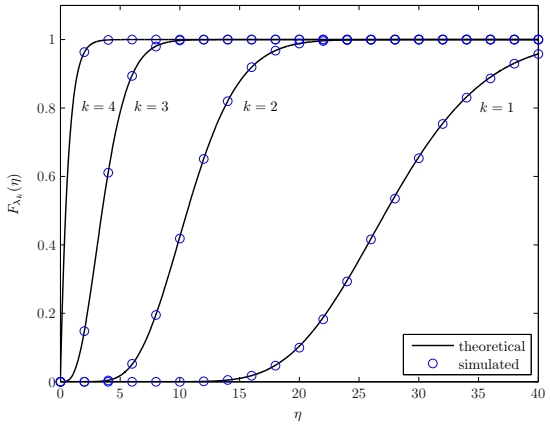
$$= d(v)!_1F_1(d(v+1); m-n+1; \omega_u) - \frac{e^{\omega_u}2^{1-v}(m-n)!}{\left(\sqrt{2\omega_u}\right)^{m-n}} Q_{d(2v),m-n}\left(\sqrt{2\omega_u}, \sqrt{2\eta}\right). \tag{2.140}$$

Observe that the sum of the two indices of the Nuttall Q-functions in (2.136) and (2.140) is always odd and, hence, Lemma 2.8 holds. The integrals in (2.136) and (2.140) can be conveniently combined to obtain the integrals needed in the computation of the joint cdf of the ordered eigenvalues.

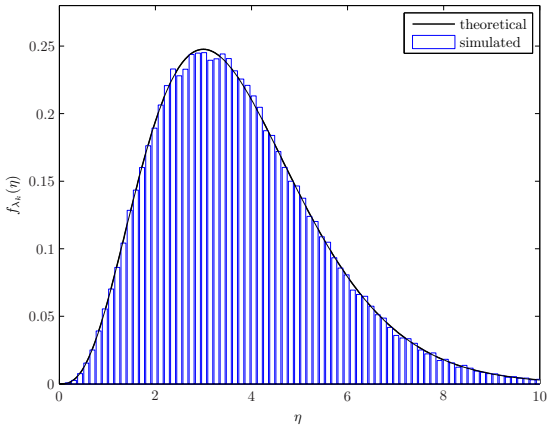
In Figure 2.3 we compare some examples of the marginal cdf's and pdf's of the ordered and unordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ obtained using the results in Sections 2.4 and 2.5 with the corresponding numerical simulations.

To the best of the author's knowledge the joint cdf of the ordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ was not available in the literature. The marginal cdf of the kth largest eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ was derived in [Kha69, eq. (9)] in terms of an infinite series of the zonal polynomials given in [Jam64]. However, the author conjectures that the given expressions can be also expressed in terms of a finite sum of determinants based on Lemma 2.10, which was proved in [Kha70, Lem. 3] by the same author (see also [Gro89, Thm. 4.2]). Recently, the marginal cdf the kth largest eigenvalue was obtained in terms of a finite sum of determinants in [Jin06, Thm. 3] and the particular cases of the largest and smallest eigenvalue in [Kan03b, Thm. 1] [Jin06, Thm. 2] and in [Jin06, Thm. 1], respectively. In addition, the marginal pdf of the maximum eigenvalue was given in [Kan03b, Cor. 3] and the case of $\mathbf{\Omega}$ being rank 1 was considered in [Kan03b, Cor. 3] and of $\mathbf{\Omega}$ with arbitrary rank in [Jin08].

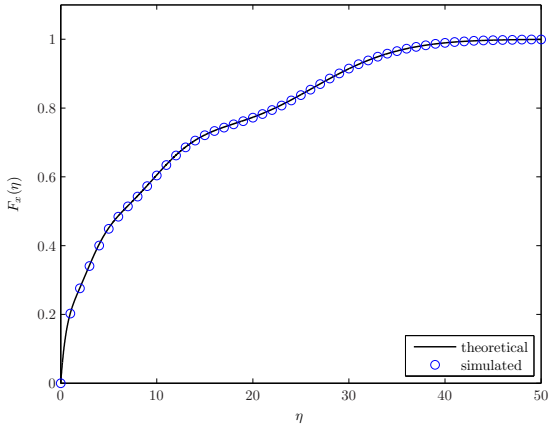
Results Regarding the Unordered Eigenvalues: The results concerning the unordered eigenvalues in Section 2.5 only require to calculate $\int_0^\infty \xi_{u,v}(\lambda)d\lambda$ and $\int_0^\eta \xi_{u,v}(\lambda)d\lambda$, which are



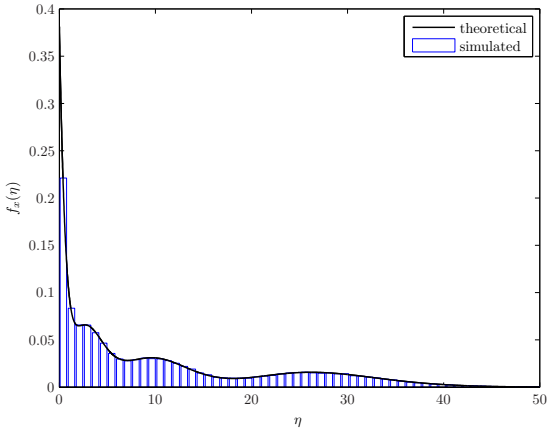
(i) marginal cdf of the kth ordered eigenvalue



(ii) marginal pdf of the kth ordered eigenvalue



(iii) marginal cdf of an unordered eigenvalue



(iv) marginal pdf of an unordered eigenvalue

Figure 2.3 Theoretical and simulated results for $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ with $n = m = 4$ and $\mathbf{\Omega}$ with eigenvalues $\boldsymbol{\omega}$=(19.1276, 5.8817, 1.4901, 0.6466).

given in (2.137) and (2.139), respectively.

The joint pdf of a few unordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_{n,m}, \mathbf{\Sigma})$ was initially given in [Maa07b, Thm. 1] for $\mathbf{\Omega}$ being full rank and in [Maa07b, Cor. 2] for $\mathbf{\Omega}$ being rank 1. Previously, the marginal pdf of an unordered eigenvalues had been derived in [Alf04a, Thm. 2] [Alf06, Thm. 1].

Taylor Expansions: Finally, in order to derive the first order Taylor expansion of marginal cdf and pdf of the kth largest eigenvalue, we have to obtain the Taylor expansion of

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c}^{\infty} \frac{a_u(t)}{t!}\lambda^t. \tag{2.141}$$

	$\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$
$\mathcal{I}$	$\{1\}$
$K_{m,n}^{(\boldsymbol{\iota})}$	$\frac{e^{-\sum_{i=1}^n \omega_i}}{((m-n)!)^n} \prod_{i<j}^n \frac{1}{(\omega_j - \omega_i)}$
$\varphi(\lambda)$	$e^{-\lambda}\lambda^{m-n}$
$\zeta_u^{(\boldsymbol{\iota})}(\lambda)$	${}_0F_1(m-n+1; \omega_u\lambda)$
$\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$	${}_0F_1(m-n+1; \omega_u\lambda)e^{-\lambda}\lambda^{d(v)}$
$\int_0^\eta \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$d(v)!{}_1F_1(d(v+1); m-n+1; \omega_u) - \frac{e^{\omega_u} 2^{1-v}(m-n)!}{\left(\sqrt{2\omega_u}\right)^{m-n}} Q_{d(2v),m-n}\left(\sqrt{2\omega_u}, \sqrt{2\eta}\right)$
$\int_\eta^\infty \xi_{i_u,v}^{(\boldsymbol{\iota})}(\lambda)d\lambda$	$\frac{e^{\omega_u} 2^{1-v}(m-n)!}{\left(\sqrt{2\phi_u}\right)^{m-n}} Q_{d(2v),m-n}\left(\sqrt{2\omega_u}, \sqrt{2\eta}\right)$
$d(v)$	$m-n+v-1$

Table 2.3 Parameters characterizing the eigenvalue distributions of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ (case (iii) of Definition 2.21).

Noting that [Abr72, eq. (9.6.10)]

$$I_n(\lambda) = \left(\frac{\lambda}{2}\right)^n \sum_{i=0}^\infty \frac{\left(\frac{\lambda^2}{4}\right)^i}{\Gamma(n+i+1)i!} \tag{2.142}$$

it follows that

$$\frac{d^t}{d\lambda^t}\left(\zeta_u(\lambda)\varphi(\lambda)\right) = \frac{\Gamma(m-n+1)}{\left(\sqrt{\omega_u}\right)^{m-n}} \frac{d^t}{d\lambda^t}\left(e^{-\lambda}\lambda^{(m-n)/2} I_{m-n}(2\sqrt{\omega_u\lambda})\right) \tag{2.143}$$

$$= \Gamma(c+1) \sum_{i=0}^\infty \frac{\omega_u^i}{\Gamma(c+i+1)i!} \frac{d^t}{d\lambda^t}\left(e^{-\lambda}\lambda^{c+i}\right) \tag{2.144}$$

where $c = m-n$. Then, using Leibniz's Rule (see [Abr72, eq. (3.3.8)]), we have that

$$\frac{d^t}{d\lambda^t}\left(e^{-\lambda}\lambda^{c+i}\right) = \sum_{r=0}^t \binom{t}{r} \frac{(c+i)!}{(c+i-r)!}(-1)^{t-r} e^{-\lambda}\lambda^{c+i-r} \tag{2.145}$$

where $\binom{n}{t}$ denotes the binomial coefficient [Abr72, eq. (3.1.2)] and, finally, using (2.144) and (2.145),

$$a_u(t) = \frac{d^t}{d\lambda^t}\left(\zeta_u(\lambda)\varphi(\lambda)\right)\Big|_{\lambda=0} = \begin{cases} 0 & t < c \\ c! \sum_{i=0}^{t-c} \binom{t}{c+i} \frac{(-1)^{t-(c+i)}}{i!} \omega_u^i & t \geq c \end{cases}. \tag{2.146}$$

The expression for $b_{u,v}$ is given in (2.137).

The first order Taylor expansion of the marginal pdf of the kth largest eigenvalue of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$ was obtained in [Jin06].

2.6.4 Complex Correlated Noncentral Wishart Matrices

The joint pdf of the ordered eigenvalues of a complex correlated noncentral Wishart matrix $\mathbf{W}$, i.e., $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ (case (iv) of Definition 2.21), has been largely unknown in the literature, due to the impossibility of expressing the joint pdf in terms of hypergeometric functions of Hermitian matrix arguments or their corresponding series expansion in terms of complex zonal polynomials. Only recently, Ratnarajah et al derived the joint pdf of the eigenvalues in [Rat05d] using an infinite series expansion of the invariant polynomials proposed by Davis in [Dav79, Dav80] to derive the real counterpart. However, analogously to the infinite series of zonal polynomials associated with hypergeometric functions, this infinite series is extremely difficult to compute. In this section, we present an approximation of the distribution of complex correlated central Wishart matrices, which allows us to apply the results presented in Sections 2.4 and 2.5.

Let $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$, using Lemma 2.16 and the pdf of $\mathbf{W}$ in Lemma 2.13, the joint pdf of the ordered strictly positive eigenvalues of $\mathbf{W}$, $\lambda_1 \geq \cdots \geq \lambda_n$, is given by

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{e^{-\mathrm{tr}(\mathbf{\Omega})}\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)\tilde{\Gamma}_n(m)|\mathbf{\Sigma}|^m} \prod_{i=1}^{n} \lambda_i^{m-n} \prod_{i<j}(\lambda_i - \lambda_j)^2 \int_{\mathcal{U}(n)} e^{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger\right)} {}_0\tilde{F}_1(m; \mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger) d\mathbf{U}. \quad (2.147)$$

In the case of uncorrelated noncentral Wishart matrices ($\mathbf{\Sigma} = \mathbf{I}_n$) we have that

$$e^{-\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger\right)} = e^{-\mathrm{tr}(\mathbf{\Lambda})}. \quad (2.148)$$

Thus, the exponential trace term can be dropped out of the integral and the joint pdf given in (2.127) follows from using the splitting property (see [Jam64, eq. (92)]) of hypergeometric functions of matrix arguments. This is not the case for the correlated noncentral case and the integral in (2.147) has to be obtained in terms of an infinite series expansion of invariant polynomials as done in [Rat05d].

The approximation of the noncentral Wishart distribution has, among others, been considered by [Ste72, Tan79, Tan82, Kol95].[25] In particular, [Ste72, Tan79, Tan82] perturb the covariance matrix of Wishart distribution so that the moments of the central and noncentral Wishart

[25] These references consider real Wishart distributions but the extension to the complex case is straightforward.

distributions are close to each other, while [Kol95] uses the idea of centering the noncentral Wishart distribution, which is not valid to approximate the eigenvalues.

The approach of [Ste72] can be briefly illustrated as follows [Gup00, Sec. 3.7]. Let $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ with $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}\mathbf{\Theta}\mathbf{\Theta}^\dagger$, then the first two moments of $\mathbf{W}$ are given by

$$\mathbb{E}\{[\mathbf{W}]_{i,j}\} = m\sigma_{ij} + \theta_{ij} \tag{2.149}$$

and

$$\mathbb{E}\{[\mathbf{W}]_{i,j}[\mathbf{W}]^*_{u,v}\} = (m\sigma_{ij} + \theta_{ij})(m\sigma_{uv} + \theta_{uv}) + m(\sigma_{iu}\sigma_{jv} + \sigma_{iv}\sigma_{ju}) + \sigma_{jv}\theta_{iu} + \sigma_{iv}\theta_{ju} + \sigma_{ju}\theta_{iv} + \sigma_{iu}\theta_{jv} \tag{2.150}$$

where $\theta_{ij} = [\mathbf{\Theta}\mathbf{\Theta}^\dagger]_{i,j}$ and $\sigma_{ij} = [\mathbf{\Sigma}]_{i,j}$. When $\mathbf{\Theta} = \mathbf{0}$, i.e, $\theta_{ij} = 0$, the above moments reduce to the moments the correlated central Wishart distribution:

$$\mathbb{E}\{[\mathbf{W}]_{i,j}\} = m\sigma_{ij} \tag{2.151}$$

and

$$\mathbb{E}\{[\mathbf{W}]_{i,j}[\mathbf{W}]^*_{u,v}\} = m^2\sigma_{ij}\sigma_{uv} + n(\sigma_{iu}\sigma_{jv} + \sigma_{iv}\sigma_{ju}). \tag{2.152}$$

Now consider a correlated central Wishart matrix $\widetilde{\mathbf{W}} \sim \mathcal{W}_n(m, \mathbf{0}_n, \widetilde{\mathbf{\Sigma}})$ with $\widetilde{\mathbf{\Sigma}} = \mathbf{\Sigma} + \frac{1}{m}\mathbf{\Theta}\mathbf{\Theta}^\dagger$. Then, from (2.151) and (2.152), we have that

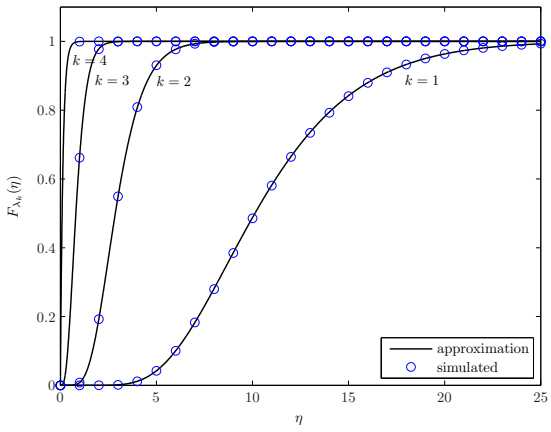
$$\mathbb{E}\{[\widetilde{\mathbf{W}}]_{i,j}\} = m\sigma_{ij} + \theta_{i,j} \tag{2.153}$$

and

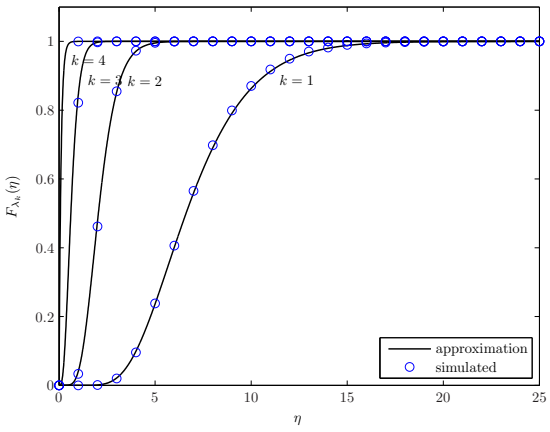
$$\mathbb{E}\{[\widetilde{\mathbf{W}}]_{i,j}[\widetilde{\mathbf{W}}]^*_{u,v}\} = (m\sigma_{ij} + \theta_{ij})(m\sigma_{uv} + \theta_{uv}) + m(\sigma_{iu}\sigma_{jv} + \sigma_{iv}\sigma_{ju}) + \sigma_{jv}\theta_{iu} + \sigma_{iv}\theta_{ju} + \sigma_{ju}\theta_{iv} + \sigma_{iu}\theta_{jv} + \frac{1}{m}(\theta_{iu}\theta_{jv} + \theta_{iv}\theta_{ju}). \tag{2.154}$$

Comparing (2.151) with (2.153) and (2.152) with (2.154), it can be seen that the first order moments of $\mathbf{W}$ and $\widetilde{\mathbf{W}}$ are equal, whereas the second order moments differ in terms of order $O(m^{-1})$. This suggests that the distribution of $\mathbf{W}$ can be approximated by the distribution of $\widetilde{\mathbf{W}}$ as summarized in the following lemma.

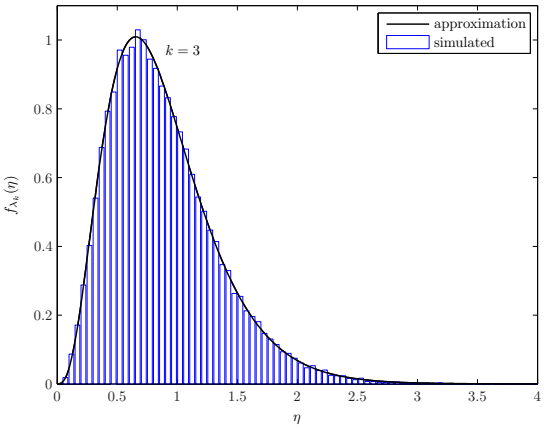
Lemma 2.17 (Correlated Noncentral Wishart Approximation [Ste72, eq. (3.4)])**.** *The distribution of* $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ *with* $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}\mathbf{\Theta}\mathbf{\Theta}^\dagger$ *can be approximated by the distribution of* $\widetilde{\mathbf{W}} \sim \mathcal{W}_n(m, \mathbf{0}_n, \widetilde{\mathbf{\Sigma}})$ *with* $\widetilde{\mathbf{\Sigma}} = \mathbf{\Sigma} + \frac{1}{m}\mathbf{\Theta}\mathbf{\Theta}^\dagger$.



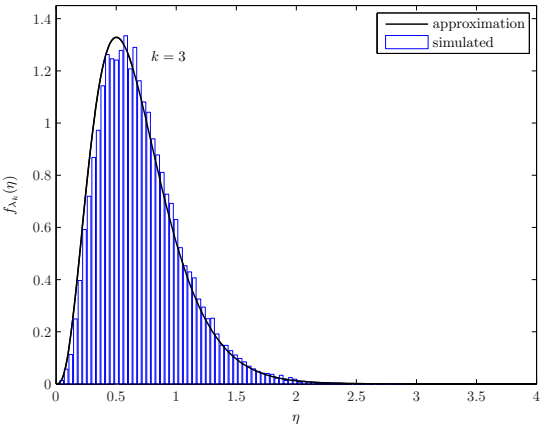
(i) marginal cdf of the kth ordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 0.1$



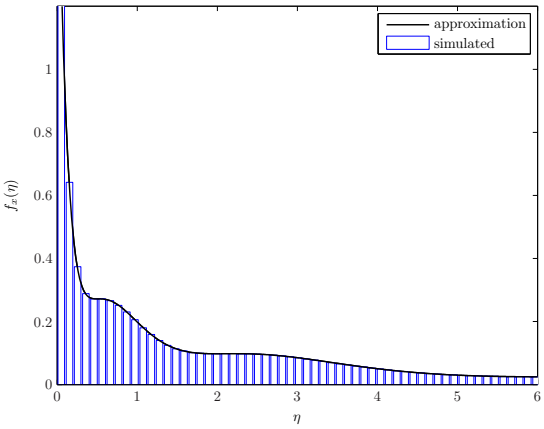
(ii) marginal cdf of the kth ordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 1$



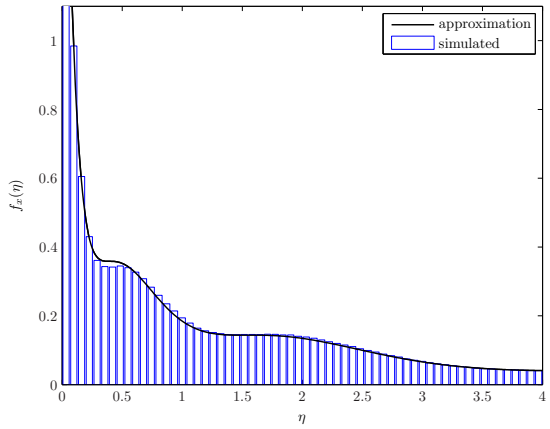
(iii) marginal pdf of the kth ordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 0.1$



(iv) marginal pdf of the kth ordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 1$



(v) marginal pdf of an unordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 0.1$



(vi) marginal pdf of an unordered eigenvalue, $\mathrm{tr}\,(\mathbf{\Omega}) = 1$

Figure 2.4 Examples of marginal cdf's and pdf's of the ordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ using the approximation in Lemma 2.17 with $n = m = 4$.

Hence, using Lemma 2.17, the eigenvalues of correlated noncentral Wishart matrices can be approximated by those of correlated central Wishart matrices addressed in Section 2.6.2. With illustration purposes, we provide in Figure 2.4 some examples of the simulated marginal cdf's and pdf's of the ordered and unordered eigenvalues of $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{\Sigma})$ and the approximation given Lemma 2.17. As it can be seen, this approximation is only accurate for a small degree of noncentrality, which is measured by $\mathrm{tr}\left(\mathbf{\Omega}\right) = \mathrm{tr}\left(\mathbf{\Theta}\mathbf{\Theta}^{\dagger}\right)/\mathrm{tr}\left(\mathbf{\Sigma}\right)$.

2.6.5 Complex Correlated Central Pseudo-Wishart Matrices

Recall from Section 2.3.3, that Pseudo-Wishart matrices are singular. More exactly, let $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$, then $\mathbf{W}$ has n strictly positive eigenvalues and $m-n$ zero eigenvalues with probability one. Furthermore, for the uncorrelated case, i.e., $\mathbf{\Psi} = \mathbf{I}_m$, it holds that the nonzero eigenvalues of $\mathbf{W}$ have the same distribution as the eigenvalues of a central uncorrelated Wishart matrix [Mal03, Sec. V], which have been analyzed in Section 2.6.1. Hence, here we focus only on complex correlated Pseudo-Wishart matrices.

The joint pdf of the ordered eigenvalues of a correlated central Pseudo-Wishart matrix $\mathbf{W}$, i.e., $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$ (see Definition 2.22) was derived in [Uhl94, Thm. 6] for the real case and in [Smi03, eq. (25)] [Rat05b, Thm. 4] for the complex case.

Joint pdf: The joint pdf of the ordered nonzero eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ of $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$ is given by [Smi03, eq. (25)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \prod_{i=1}^{n} \frac{1}{(n-i)!} \prod_{i<j}^{m} \frac{1}{(\psi_j - \psi_i)} |\mathbf{E}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\lambda})| \tag{2.155}$$

where $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3) and $\mathbf{E}(\boldsymbol{\lambda})$ is defined as

$$[\mathbf{E}(\boldsymbol{\lambda})]_{u,v} = \begin{cases} \psi_u^{v-1} & 1 \leq v \leq m-n \\ \psi_u^{m-n-1} e^{-\lambda_{v-m+n}/\psi_u} & m-n < v \leq m \end{cases} \quad \text{for } u, v = 1, \ldots, m \tag{2.156}$$

where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_m)$ are the eigenvalues of $\mathbf{\Psi}$ ordered such that $(\psi_1 > \cdots > \psi_m > 0)$. Performing the Laplace expansion (see e.g. [Ait83, Sec. 33]) over the first $m-n$ columns of $\mathbf{E}(\boldsymbol{\lambda})$, it follows that

$$|\mathbf{E}(\boldsymbol{\lambda})| = \sum_{\boldsymbol{\iota} \in \mathcal{I}} (-1)^{\sum_{i=1}^{m-n}(\iota_i + i)} |\mathbf{V}^{(\boldsymbol{\iota})}(\boldsymbol{\psi})||\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})| \tag{2.157}$$

where the summation over $\boldsymbol{\iota} = (\iota_1, \ldots, \iota_m)$ is for all permutation of integers $(1, \ldots, m)$ such that $(\iota_1 < \cdots < \iota_{m-n})$ and $(\iota_{m-n+1} < \cdots < \iota_m)$ and the matrices $\mathbf{V}^{(\boldsymbol{\iota})}(\boldsymbol{\psi})$ $((m-n) \times (m-n))$ and $\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})$ $(n \times n)$ are defined as

$$[\mathbf{V}^{(\boldsymbol{\iota})}(\boldsymbol{\sigma})]_{u,v} = \psi_{\iota_u}^{v-1} \qquad \text{for } u, v = 1, \ldots, m-n \tag{2.158}$$

$$[\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})]_{u,v} = \zeta_u^{(\boldsymbol{\iota})}(\lambda_v) = \psi_{\iota_{m-n+u}}^{m-n-1} e^{-\lambda_v/\psi_{\iota_{m-n+u}}} \qquad \text{for } u, v = 1, \ldots, n. \tag{2.159}$$

Observing that $\mathbf{V}^{(\boldsymbol{\iota})}(\boldsymbol{\sigma})$ is a Vandermonde matrix (see Lemma 2.1), we can finally rewrite the joint pdf as

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \prod_{i=1}^{n} \frac{1}{(n-i)!} \prod_{i<j}^{m} \frac{1}{(\psi_j - \psi_i)} \sum_{\boldsymbol{\iota} \in \mathcal{I}} (-1)^{\sum_{i=1}^{m-n}(\iota_i + i)} \prod_{i<j}^{m-n} (\psi_{\iota_j} - \psi_{\iota_i}) |\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})| |\mathbf{V}(\boldsymbol{\lambda})| \tag{2.160}$$

Identifying terms, the joint pdf of the ordered nonzero eigenvalues in (2.160) coincides with the general pdf given in Assumption 2.1 by defining the set $\mathcal{I}$ as

$$\mathcal{I} = \{(\iota_1, \ldots, \iota_m) = \pi(1, \ldots, m) | (\iota_1 < \cdots < \iota_{m-n}) \text{ and } (\iota_{m-n+1} < \cdots < \iota_n)\} \tag{2.161}$$

where $\pi(\cdot)$ denotes permutation, the constant $K_{m,n}^{(\boldsymbol{\iota})}$ as

$$K_{m,n}^{(\boldsymbol{\iota})} = \frac{(-1)^{\sum_{i=1}^{m-n}(\iota_i + i)}}{\prod_{i=1}^{n}(n-i)!} \frac{\prod_{i<j}^{m-n}(\psi_{\iota_j} - \psi_{\iota_i})}{\prod_{i<j}^{m}(\psi_j - \psi_i)} \tag{2.162}$$

the function $\varphi(\lambda) = 1$, and matrix $\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})$ with entries as given in (2.129). Hence, it follows that

$$\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda)\varphi(\lambda)\lambda^{v-1} = \psi_{\iota_{d(u+1)}}^{m-n-1} e^{-\lambda/\psi_{\iota_{d(u+1)}}} \lambda^{v-1} \tag{2.163}$$

where we have introduced the function $d(u) = m - n + u - 1$.

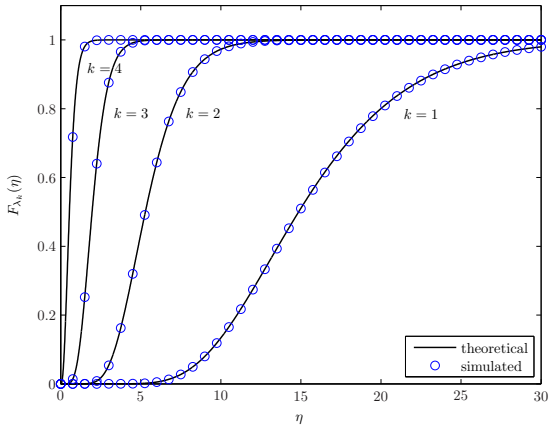
Results Regarding the Ordered Eigenvalues: In order to derive the marginal cdf and pdf of the kth largest eigenvalue using the results presented in Section 2.4, we only have to particularize

$$\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda = \psi_{\iota_{d(u+1)}}^{d(v)} \Gamma(v, \eta/\psi_{\iota_{d(u+1)}}) \tag{2.164}$$

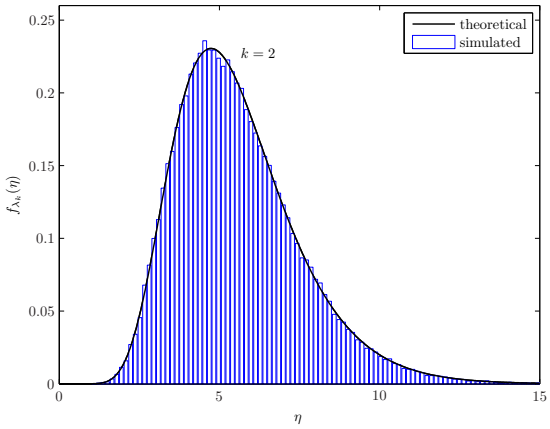
$$\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda = \psi_{\iota_{d(u+1)}}^{d(v)} \gamma(v, \eta/\psi_{\iota_{d(u+1)}}) \tag{2.165}$$

where $\Gamma(\cdot,\cdot)$ and $\gamma(\cdot,\cdot)$ are the upper and lower incomplete gamma functions given in Definitions 2.7 and 2.8, respectively. The integrals in (2.122) and (2.123) can be conveniently combined to obtain the integrals needed in the computation of the joint cdf of the ordered eigenvalues.

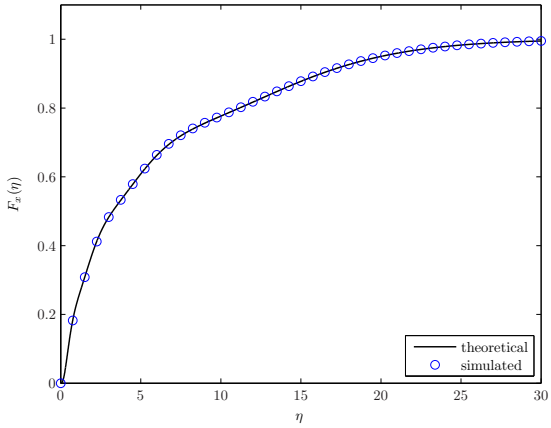
The marginal cdf of the largest eigenvalue and smallest eigenvalue of $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \Psi)$ was recently derived in [Kan03a, Thm. 4 (2)] [Maa06, eq. (21)] [Maa07a, eq. (40)] and [Maa07a,



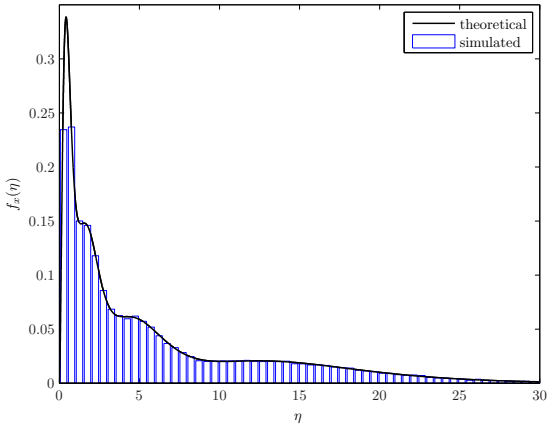
(i) marginal cdf of the kth ordered eigenvalue



(ii) marginal pdf of the kth ordered eigenvalue



(iii) marginal cdf of an unordered eigenvalue



(iv) marginal pdf of an unordered eigenvalue

Figure 2.5 Theoretical and simulated results for $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$ with $n = 4$ and $m = 6$, and $[\mathbf{\Psi}]_{i,j} = r^{|i-j|}$ $(m \times m)$, $r = 0.6$.

eq. (41)], respectively, and the marginal pdf's of the largest and smallest eigenvalue were calculated in [Maa06, eq. (22)] [Maa07a, eq. (42)] and in [Maa06, eq. (25)] [Maa07a, eq. (43)]. To the best of the author's knowledge, the marginal cdf and pdf of the kth largest eigenvalue were not available in the literature.

Results Regarding the Unordered Eigenvalues: The results concerning the unordered eigenvalues in Section 2.5 only require to calculate

$$\int_0^\infty \xi_{u,v}^{(\iota)}(\lambda)d\lambda = \psi_{\iota_{d(u+1)}}^{d(v)}\Gamma(v) = \psi_{\iota_{d(u+1)}}^{d(v)}(v-1)! \tag{2.166}$$

where $\Gamma(\cdot)$ is the gamma function in Definition 2.5, in addition to the integral $\int_0^\eta \xi_{u,v}^{(\iota)}(\lambda)d\lambda$, which is already given in (2.165). The joint cdf and pdf of a set of unordered eigenvalues were

	$\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Sigma})$
$\mathcal{I}$	$\{(\iota_1, \ldots, \iota_m) = \pi(1, \ldots, m) \vert (\iota_1 < \cdots < \iota_{m-n}) \text{ and } (\iota_{m-n+1} < \cdots < \iota_n)\}$
$K_{m,n}^{(\boldsymbol{\iota})}$	$\frac{(-1)^{\sum_{i=1}^{m-n}(\iota_i+i)}}{\prod_{i=1}^{n}(n-i)!} \frac{\prod_{i<j}^{m-n}(\psi_{\iota_j}-\psi_{\iota_i})}{\prod_{i<j}^{m}(\psi_j-\psi_i)}$
$\varphi(\lambda)$	1
$\zeta_u^{(\boldsymbol{\iota})}(\lambda)$	$\psi_{\iota_{d(u+1)}}^{m-n-1} e^{-\lambda_v/\psi_{\iota_{d(u+1)}}}$
$\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda)$	$\psi_{\iota_{d(u+1)}}^{m-n-1} e^{-\lambda_v/\psi_{\iota_{d(u+1)}}} \lambda^{v-1}$
$\int_\eta^\infty \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$	$\psi_{\iota_{d(u+1)}}^{d(v)} \Gamma(v, \eta/\psi_{\iota_{d(u+1)}})$
$\int_0^\eta \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$	$\psi_{\iota_{d(u+1)}}^{d(v)} \gamma(v, \eta/\psi_{\iota_{d(u+1)}})$
$\int_0^\infty \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$	$\psi_{\iota_{d(u+1)}}^{d(v)} (v-1)!$
$d(v)$	$m-n+v-1$

Table 2.4 Parameters characterizing the eigenvalue distributions of $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \Psi)$ (Definition 2.22).

recently derived in [Maa06, eq. (19)] and [Maa06, eq. (20)] and the marginal pdf of an unordered eigenvalue in [Maa07a, eq. (36)].

In Figure 2.5 we compare some examples of the marginal cdf's and pdf's of the ordered and unordered eigenvalues of $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Psi})$ obtained using the results in Sections 2.4 and 2.5 with the corresponding numerical simulations.

Taylor Expansions: The joint pdf of the ordered eigenvalues of $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \Psi)$ in (2.160) does not satisfy Assumption 2.2. Hence, the first order Taylor expansions of the marginal cdf and pdf of the kth largest eigenvalue cannot be obtained using the results presented in Section 2.4.5.

2.6.6 Complex Central Quadratic Form Matrices

The joint pdf of the ordered eigenvalues of a matrix $\mathbf{W}$ following the complex central quadratic form distribution, i.e., $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{I}_n, \mathbf{\Sigma}, \mathbf{\Psi})$ (see Definition 2.23) was recently derived in [Sim04, eq. (16)] [Sim06, eq. (56)] [McK07, eq. (57)].

Joint pdf ($n = m$): The joint pdf of the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of

$\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{I}_n, \mathbf{\Sigma}, \mathbf{\Psi})$ is given for $n = m$ by [Sim06, eq. (56)] [McK07, eq. (57)]

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{(-1)^{n(n-1)/2}}{\prod_{i=1}^{n}(\psi_i\sigma_i)^n} \sum_{k=0}^{\infty} \sum_{\boldsymbol{\kappa}\in\mathcal{K}(k)} \frac{\prod_{i=1}^{n}(-1)^{\kappa_i}}{\prod_{i=1}^{n}\kappa_i!|\mathbf{V}(\boldsymbol{\kappa})|} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\boldsymbol{\psi}^{-1})||\mathbf{K}^{(\boldsymbol{\kappa})}(\boldsymbol{\sigma}^{-1})|}{|\mathbf{V}(\boldsymbol{\psi}^{-1})||\mathbf{V}(\boldsymbol{\sigma}^{-1})|} |\mathbf{K}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\lambda})| \tag{2.167}$$

where the summation over $\boldsymbol{\kappa} \in \mathcal{K}(k)$ is for all strictly ordered partitions $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_n)$ with $\kappa_1 > \cdots > \kappa_n$ and $\kappa_1 + \cdots + \kappa_n = k$, $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3), matrix $\mathbf{K}^{(\boldsymbol{\kappa})}(\cdot)$ $(n \times n)$ is defined as

$$[\mathbf{K}^{(\boldsymbol{\kappa})}(\mathbf{x})]_{u,v} = x_u^{\kappa_v} \qquad \text{for } u, v = 1, \ldots, n \tag{2.168}$$

and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ and $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_n)$ denote the eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{\Psi}$ ordered such that $(\sigma_1 > \cdots > \sigma_n > 0)$ and $(\psi_1 > \cdots > \psi_n > 0)$. Identifying terms, the expression in (2.167) coincides with the general pdf given in Assumption 2.1 by defining the set $\mathcal{I}$ as

$$\mathcal{I} = \{(\iota_1, \ldots, \iota_n) \in \mathbb{N}^n | (\iota_1 > \cdots > \iota_n) \text{ and } (\iota_1 + \cdots + \iota_n = k) \text{ for } k = 0, 1, \ldots\} \tag{2.169}$$

with cardinality $|\mathcal{I}| = \sum_{k=0}^{\infty} |\mathcal{K}(k)|$, the constant $K_{m,n}^{(\boldsymbol{\iota})}$ as

$$K_{m,n}^{(\boldsymbol{\iota})} = \frac{(-1)^{n(n-1)/2}\prod_{i=1}^{n}(-1)^{\iota_i}}{\prod_{i=1}^{n}\iota_i!\prod_{i<j}^{n}(\iota_j - \iota_i)} \frac{|\mathbf{K}^{(\boldsymbol{\iota})}(\boldsymbol{\psi}^{-1})||\mathbf{K}^{(\boldsymbol{\iota})}(\boldsymbol{\sigma}^{-1})|}{|\mathbf{V}(\boldsymbol{\psi}^{-1})||\mathbf{V}(\boldsymbol{\sigma}^{-1})|} \prod_{i=1}^{n}(\psi_i\sigma_i)^{-n} \tag{2.170}$$

the function $\varphi(\lambda) = 1$, and matrix $\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda}) = \mathbf{K}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})$. Hence, it follows that

$$\xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) = \zeta_u^{(\boldsymbol{\iota})}(\lambda)\varphi(\lambda)\lambda^{v-1} = \lambda^{\iota_u+v-1}. \tag{2.171}$$

Joint pdf $(n < m)$: The case of $n < m$ was not explicitly addressed in [Sim06] [McK07]. However, we can use the approach followed in [Sim06, Sec. B] for correlated (Pseudo-)Wishart matrices and consider an auxiliary expanded system with m ordered eigenvalues denoted by $\tilde{\boldsymbol{\lambda}} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m) = (\lambda_1, \ldots, \lambda_n, \tilde{\lambda}_{n+1}, \ldots, \tilde{\lambda}_m)$ and an expanded correlation matrix $\widetilde{\mathbf{\Sigma}}$ $(m \times m)$ with eigenvalues denoted by $\tilde{\boldsymbol{\sigma}} = (\tilde{\sigma}_1, \ldots, \tilde{\sigma}_m) = (\sigma_1, \ldots, \sigma_n, \tilde{\sigma}_{n+1}, \ldots, \tilde{\sigma}_m)$. The joint pdf the ordered eigenvalues, $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, of $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{I}_n, \mathbf{\Sigma}, \mathbf{\Psi})$ for $n < m$ is then obtained by taking the limit as $(\tilde{\sigma}_{n+1}, \ldots, \tilde{\sigma}_m) \to (0, \ldots, 0)$ and $(\tilde{\lambda}_{n+1}, \ldots, \tilde{\lambda}_m) \to (0, \ldots, 0)$, i.e,

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = (-1)^{m(m-1)/2}|\mathbf{V}(\boldsymbol{\lambda})|^2 \prod_{i=1}^{n}\lambda_i^{m-n} \sum_{k=0}^{\infty} \sum_{\boldsymbol{\kappa}\in\mathcal{K}(k)} \frac{\prod_{i=1}^{m}(-1)^{\kappa_i}}{\prod_{i=1}^{m}\kappa_i!|\mathbf{V}(\boldsymbol{\kappa})|} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\boldsymbol{\psi}^{-1})|}{\prod_{i=1}^{m}\psi_i^m|\mathbf{V}(\boldsymbol{\psi}^{-1})|}$$
$$\lim_{(\tilde{\sigma}_{n+1},\ldots,\tilde{\sigma}_m,\tilde{\lambda}_{n+1},\ldots,\tilde{\lambda}_m)\to(0,\ldots,0)} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})|}{\prod_{i=1}^{m}\tilde{\sigma}_i^m|\mathbf{V}(\tilde{\boldsymbol{\sigma}}^{-1})|} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\lambda}})|}{|\mathbf{V}(\tilde{\boldsymbol{\lambda}})|} \tag{2.172}$$

where the summation over $\boldsymbol{\kappa} \in \mathcal{K}(k)$ is for all strictly ordered partitions $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)$ with $\kappa_1 > \cdots > \kappa_m$ and $\kappa_1 + \cdots + \kappa_m = k$, $\mathbf{V}(\cdot)$ is a Vandermonde matrix (see Definition 2.3), and

matrix $\mathbf{K}^{(\boldsymbol{\kappa})}(\cdot)$ $(m \times m)$ is defined as

$$[\mathbf{K}^{(\boldsymbol{\kappa})}(\mathbf{x})]_{u,v} = x_u^{\kappa_v} \qquad \text{for } u, v = 1, \ldots, m. \tag{2.173}$$

The limits in (2.180) can be easily calculated by applying the generalized L'Hopital rules in Lemmas 2.4 and 2.5 to the quotients $|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\lambda}})|/|\mathbf{V}(\tilde{\boldsymbol{\lambda}})|$ and $\prod_{i=1}^m \tilde{\sigma}_i^{-m}|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})|/|\mathbf{V}(\tilde{\boldsymbol{\sigma}}^{-1})|$, respectively. In the first case, it holds that (see Lemma 2.4)

$$\lim_{(\tilde{\lambda}_{n+1},\ldots,\tilde{\lambda}_m)\to(0,\ldots,0)} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\lambda}})|}{|\mathbf{V}(\tilde{\boldsymbol{\lambda}})|} = \frac{1}{\prod_{i=1}^{m-n-1} i!} \frac{|\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})|}{\prod_{i=1}^n \lambda_i^{m-n}|\mathbf{V}(\boldsymbol{\lambda})|} \tag{2.174}$$

where matrix $\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})$ $(m \times m)$ is defined as

$$[\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})]_{u,v} = \begin{cases} \lambda_u^{\kappa_v} & 1 \le u \le n \\ \prod_{i=1}^{u-n-1}(\kappa_v - i - 1) & n < u \le m, \kappa_v = u - n - 1 \\ 0 & n < u \le m, \kappa_v \ne u - n - 1 \end{cases} \tag{2.175}$$

for $u, v = 1, \ldots, m$. Before calculating the second limit it is convenient to observe that

$$\prod_{i=1}^m \tilde{\sigma}_i^{-m} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})|}{|\mathbf{V}(\tilde{\boldsymbol{\sigma}}^{-1})|} = \frac{|\overline{\mathbf{K}}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})|}{|\mathbf{V}(\tilde{\boldsymbol{\sigma}}^{-1})|} \tag{2.176}$$

where $[\overline{\mathbf{K}}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})]_{u,v} = \tilde{\sigma}_u^{-k_v-m}$. Then, by defining $\bar{\boldsymbol{\sigma}} = (\sigma_1^{-1}, \ldots, \sigma_n^{-1}, \tilde{\sigma}_{n+1}, \ldots, \tilde{\sigma}_m)$ we can apply Lemma 2.5 as

$$\lim_{(\tilde{\sigma}_{n+1},\ldots,\tilde{\sigma}_m)\to(0,\ldots,0)} \frac{|\widetilde{\mathbf{K}}^{(\boldsymbol{\kappa})}(\tilde{\boldsymbol{\sigma}}^{-1})|}{|\mathbf{V}(\tilde{\boldsymbol{\sigma}}^{-1})|} = \lim_{(\bar{\sigma}_{n+1},\ldots,\bar{\sigma}_m)\to(\infty,\ldots,\infty)} \frac{|\overline{\mathbf{K}}^{(\boldsymbol{\kappa})}(\bar{\boldsymbol{\sigma}})|}{|\mathbf{V}(\bar{\boldsymbol{\sigma}})|} \tag{2.177}$$

$$= \prod_{i=1}^n \sigma_i^{-m} \frac{|\widetilde{\mathbf{K}}^{(\boldsymbol{\kappa})}(\boldsymbol{\sigma}^{-1})|}{|\mathbf{V}(\boldsymbol{\sigma}^{-1})|} \tag{2.178}$$

where matrix $\widetilde{\mathbf{K}}^{(\boldsymbol{\kappa})}(\boldsymbol{\sigma}^{-1})$ $(m \times m)$ is defined as

$$[\widetilde{\mathbf{K}}^{(\boldsymbol{\kappa})}(\boldsymbol{\sigma}^{-1})]_{u,v} = \begin{cases} \sigma_u^{-\kappa_v} & 1 \le u \le n \\ 1 & n < u \le m, \kappa_v = u - n \\ 0 & n < u \le m, \kappa_v \ne u - n \end{cases} \qquad \text{for } u, v = 1, \ldots, m. \tag{2.179}$$

Finally, the joint pdf of the ordered eigenvalues is given by

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \frac{(-1)^{m(m-1)/2}}{\prod_{i=1}^{m-n-1} i! \prod_{i=1}^m \psi_i^m \prod_{i=1}^n \sigma_i^m} \sum_{k=0}^{\infty} \sum_{\boldsymbol{\kappa}\in\mathcal{K}(k)} \frac{\prod_{i=1}^n (-1)^{\kappa_i}}{\prod_{i=1}^n \kappa_i! |\mathbf{V}(\boldsymbol{\kappa})|} \frac{|\mathbf{K}^{(\boldsymbol{\kappa})}(\boldsymbol{\psi}^{-1})||\widetilde{\mathbf{K}}^{(\boldsymbol{\kappa})}(\boldsymbol{\sigma}^{-1})|}{|\mathbf{V}(\boldsymbol{\psi}^{-1})||\mathbf{V}(\boldsymbol{\sigma}^{-1})|} |\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\lambda})|. \tag{2.180}$$

Observe that the joint pdf in (2.180) has the form of the general pdf given in Assumption 2.1 except for the $m \times m$ matrix $\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})$. Nevertheless, we can use the same procedure as in Section 2.6.5 and perform the Laplace expansion of the determinant $|\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})|$ over the last $m-n$ rows of $\mathbf{Z}^{(\boldsymbol{\kappa})}(\boldsymbol{\lambda})$. Then, by direct identification we can obtain the expressions of the parameters describing the joint pdf of the ordered eigenvalues of $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{I}_n, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ with $n < m$ in terms of the pdf in Assumption 2.1 as done for the case $n = m$.

Results Regarding the Eigenvalues: In order to apply the distributional results in Sections 2.4 and 2.5, we have to calculate integrals of the form $\int_{\eta}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$, $\int_{0}^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$, and $\int_{0}^{\infty} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda$ (see details in Table 2.1). Unfortunately, integrating over the λ_k's before summing over $\iota \in \mathcal{I}$ seems problematic since some of the integrals are unbounded above (see expression in (2.171)). However, as noted in [Sim06, Lem. 5], the joint pdf of the ordered eigenvalues in (2.167) is bounded by an exponential function of any λ_k as λ_k becomes arbitrarily large and hence integrable. The problem lies in the interchange of sums and integrals performed in the derivations of Sections 2.4 and 2.5. To circumvent this discrepancy we can introduce a cutoff function $g(\lambda_k)$ which is unity as $\lambda_k \to 0$ and tends to zero faster than a power law as $\lambda_k \to \infty$, e.g. $g(\lambda_k) = e^{-\delta \lambda_k}$. Cutting off the integrals makes all terms finite and thus we can freely interchange the order of the summation and integration. Then, we should deal with infinite summation over $\iota \in \mathcal{I}$ by invoking, for instance, the Cauchy-Binet Formula in [Sim06, Lem. 3] and, finally, set $\delta = 0$ at the end of the calculation. Although we have not been able to go through this last step, we still conjeture that it is possible as done in [McK07] to calculate the marginal cdf of the largest ordered eigenvalue of $\mathbf{W} \sim \mathcal{Q}_{n,m}(\mathbf{I}_n, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$.

2.7 Conclusions and Publications

The probabilistic characterization of the eigenvalues of Wishart, Pseudo-Wishart and quadratic form distributions is critical in the performance evaluation of many communication and signal processing applications. Many different contributions, as early as the sixties in the mathematical literature and much more recently in the signal processing community, provided partial characterizations for specific problems. However, the unified perspective provided by this chapter was missing and can, not only fill the gap of the currently unknown results, but even more importantly, provide a solid framework for the understanding and direct derivation of all the already

existing results.

The main results contained in this chapter regarding the probabilistic characterization of the eigenvalues of a general class of Hermitian random matrices have been published in two journal papers and one conference paper:

[Ord09a] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 672–689, Feb. 2009.

[Ord08b] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 3846–3852, May 2008.

[Ord09b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Minimum BER linear MIMO transceivers with optimum number of substreams", *accepted in IEEE Trans. Signal Processing*, Jan. 2009.

2.A Appendix: Joint pdf and cdf of the Ordered Eigenvalues

2.A.1 Motivation of Assumption 2.1

The adoption of Assumption 2.1 can be motivated by investigating the form of the most common Hermitian matrix distributions and the form of the associated joint pdf of its ordered eigenvalues.

Typical univariate distributions such as the Chi-squared, Cauchy and Beta distributions involve Bessel and hypergeometric functions which can all be written as special cases, for particular integers p and q, of the generalized hypergeometric function of scalar arguments in Definition 2.10. The corresponding complex matrix variate distributions involve a generalization of this function to the case in which the variable x is replaced by an Hermitian matrix $\mathbf{X}$, known as generalized hypergeometric function of Hermitian matrix argument (see Definition 2.16). Observe that this statement holds in particular for the Wishart, Pseudo-Wishart, and quadratic form distributions introduced in Section 2.3, as is clarified in the following.

Let us consider first that the distribution of the Hermitian random matrix $\mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$ can be written as

$$f_{\mathbf{W}}(\mathbf{W}) = K_{\mathbf{W}}\, {}_p\tilde{F}_q\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{\Sigma}\mathbf{W}\right) \prod_{t=1}^{n} \varphi(\lambda_t) \tag{2.181}$$

where $K_{\mathbf{W}}$ is a normalization constant, ${}_p\tilde{F}_q(\cdot;\cdot;\cdot)$ denotes the hypergeometric function of Hermitian matrix argument (see Definition 2.16), $\mathbf{\Sigma}$ $(n \times n)$ is a deterministic Hermitian matrix with eigenvalues denoted by $\boldsymbol{\sigma}$, and $\varphi(\lambda)$ is an arbitrary function. By direct identification, is straightforward to see that the pdf in (2.181) admits as particular cases the pdf's of complex central Wishart and complex uncorrelated noncentral Wishart matrices (see Lemmas 2.12 and 2.13). In addition, this pdf expression also holds for some cases of the complex inverted Wishart distribution, complex matrix variate Cauchy, and Bessel distributions (see [Jam64, Sec. 8] and [Gup00]). Indeed, using Lemma 2.16, the joint pdf of the ordered eigenvalues of $\mathbf{W}$ is given by

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = K_{\mathbf{W}} \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)} |\mathbf{V}(\boldsymbol{\lambda})|^2 \prod_{t=1}^{n} \varphi(\lambda_t) \int_{\mathcal{U}(n)} {}_p\tilde{F}_q\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{\Sigma}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger\right) d\mathbf{U} \tag{2.182}$$

$$= K_{\mathbf{W}} \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)} |\mathbf{V}(\boldsymbol{\lambda})|^2 {}_p\tilde{F}_q\left(a_1, \ldots, a_p; b_1, \ldots, b_q; \mathbf{\Sigma}, \mathbf{\Lambda}\right) \prod_{t=1}^{n} \varphi(\lambda_t) \tag{2.183}$$

where ${}_p\tilde{F}_q\left(\cdot;\cdot;\cdot,\cdot\right)$ denotes the hypergeometric function of two Hermitian matrix arguments (see Definition 2.16) and (2.183) follows from the splitting property in [Jam64, eq. (92)]. Hypergeometric function of two Hermitian matrix arguments are defined as an infinite series of zonal

polynomials but can be alternatively expressed in terms of a quotient of determinants including generalized hypergeometric functions of scalar arguments. Using Lemma 2.10 it follows that

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = K_{\mathbf{W}} \frac{\pi^{n(n-1)}}{\tilde{\Gamma}_n(n)} \frac{\tilde{\Gamma}_n(n)}{\pi^{n(n-1)/2}} \frac{\prod_{i=1}^{n}\prod_{j=1}^{q}(b_j - i + 1)^{i-1}}{\prod_{i=1}^{n}\prod_{j=1}^{p}(a_j - i + 1)^{i-1}} \frac{|\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})|}{|\mathbf{V}(\boldsymbol{\lambda})||\mathbf{V}(\boldsymbol{\sigma})|} |\mathbf{V}(\boldsymbol{\lambda})|^2 \prod_{t=1}^{n} \varphi(\lambda_t)$$
$$= K_{\boldsymbol{\lambda}} |\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})||\mathbf{V}(\boldsymbol{\lambda})| \prod_{t=1}^{n} \varphi(\lambda_t) \tag{2.184}$$

where $|\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})|$ $(n \times n)$ is defined as

$$[\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})]_{i,j} = {}_pF_q\left(a_1 - n + 1, \ldots, a_p - n + 1; b_1 - n + 1, \ldots, b_q - n + 1; \lambda_i \sigma_j\right). \tag{2.185}$$

Observe that the joint pdf of the ordered eigenvalues in (2.184) coincide with Assumption 2.1 if we let the set $\mathcal{I}$ be a singleton.

Let us now consider an $m \times m$ Hermitian random matrices of rank n $(n < m)$ such as Pseudo-Wishart matrices. In this case, the previous procedure can be also followed but interchanging n by m and taking the limit $\{\lambda_i\}_{i=n+1,\ldots,m} \to 0$ [Sim04] [Sim06] [McK07]. Hence, matrix $\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})$ is $m \times m$ but only the first n rows depend on $\{\lambda_i\}_{i=1,\ldots,n}$. Performing the Laplace expansion of the determinant $|\mathbf{E}(\boldsymbol{\lambda}, \boldsymbol{\sigma})|$ over the last $(m - n)$ rows, the joint pdf of the ordered eigenvalues can be expressed as in Assumption 2.1 using the sum over the set $\mathcal{I}$ to include this sum of determinants.

Finally, in a more general case, such as the quadratic form distributions, the hypergeometric function of one Hermitian matrix argument in (2.181) is substituted by an hypergeometric function of more Hermitian matrix arguments [McK06, eq. (2.30)]. Since the splitting property of hypergeometric functions also holds, the structure of the joint pdf of the ordered eigenvalues is maintained (see e.g. [McK07, eq. (45)]). Although a determinantal expression for hypergeometric function of more than two Hermitian matrix arguments is not known, the corresponding series expansion in terms of zonal polynomials can still be included in the general expression of Assumption 2.1 using the summation over the set $\mathcal{I}$.

2.A.2 Proof of Theorem 2.1

Proof. The joint cdf of the ordered eigenvalues, $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta})$, can be obtained from the joint pdf of the ordered eigenvalues, $f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda})$, as

$$F_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \int_{\mathcal{D}_{\text{ord}}(\boldsymbol{\eta})} f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \tag{2.186}$$

where

$$\mathcal{D}_{\text{ord}}(\boldsymbol{\eta}) = \{0 \leq \lambda_1 \leq \eta_1, \ldots, 0 \leq \lambda_n \leq \eta_n\} \cap \{\lambda_1 \geq \cdots \geq \lambda_n\} \tag{2.187}$$

and recall that by assumption ($\eta_1 \geq \cdots \geq \eta_n > 0$).

Let $\boldsymbol{i} = (i_1, \ldots, i_n) \in \mathbb{N}^n$ and $\mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta}) = \mathcal{D}(\boldsymbol{i}; \boldsymbol{\eta}) \cap \{\lambda_1 \geq \cdots \geq \lambda_n\}$ with

$$\mathcal{D}(\boldsymbol{i}; \boldsymbol{\eta}) = \{\eta_{i_1+1} < \lambda_1 \leq \eta_{i_1}, \eta_{i_2+1} < \lambda_2 \leq \eta_{i_2}, \ldots, 0 < \lambda_n \leq \eta_{i_n}\}. \tag{2.188}$$

Then, $\mathcal{D}_{\text{ord}}(\boldsymbol{\eta})$ can be expressed as a union of non-overlapping domains $\mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta})$, i.e.,

$$\mathcal{D}_{\text{ord}}(\boldsymbol{\eta}) = \bigcup_{\boldsymbol{i} \in \mathcal{S}} \mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta}) \tag{2.189}$$

where the set $\mathcal{S}$ is defined as $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2$ with[26]

$$\mathcal{S}_1 = \{\boldsymbol{i} \in \mathbb{N}^n | \max(i_{s-1}, s) \leq i_s \leq n\} \tag{2.190}$$

$$\mathcal{S}_2 = \{\boldsymbol{i} \in \mathbb{N}^n | i_s \neq r \text{ if } \eta_r = \eta_{r+1}\}. \tag{2.191}$$

Observe that the set $\mathcal{S}_1$ is such that $\bigcup_{\boldsymbol{i} \in \mathcal{S}_1} \mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta})$ only includes domains $\mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta})$ in which each λ_k belongs to one of the $(n - k + 1)$ possible non-overlapping partitions of the interval $[0, \eta_k]$, i.e., $[0, \eta_n], (\eta_n, \eta_{n-1}], \ldots, (\eta_{k-1}, \eta_k]$ (see a representation for $n = 3$ in Figure 2.6). Then, by intersecting $\mathcal{S}_1$ with $\mathcal{S}_2$ we eliminate all domains $\mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta})$ with empty intervals, i.e., intervals such that $\eta_i = \eta_{i+1}$ (compare representations for $n = 3$ in Figure 2.6).

From (2.186) and (2.189), the joint cdf $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta})$ can be rewritten as

$$F_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \sum_{\boldsymbol{i} \in \mathcal{S}} \mathcal{I}(\boldsymbol{i}; \boldsymbol{\eta}) \tag{2.192}$$

where

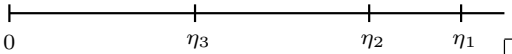
$$\mathcal{I}(\boldsymbol{i}; \boldsymbol{\eta}) = \int_{\mathcal{D}_{\text{ord}}(\boldsymbol{i}; \boldsymbol{\eta})} f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) d\boldsymbol{\lambda}. \tag{2.193}$$

Now, expanding the determinants of $\mathbf{E}(\boldsymbol{\lambda})$ and $\mathbf{V}(\boldsymbol{\lambda})$ (see Definition 2.2) in the joint pdf expression in (2.63), we can rewrite $f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda})$ as

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \text{sgn}(\boldsymbol{\mu}) \text{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} [\mathbf{E}^{(\boldsymbol{\iota})}(\boldsymbol{\lambda})]_{\mu_t, t} [\mathbf{V}(\boldsymbol{\lambda})]_{\nu_t, t} \varphi(\lambda_t) \tag{2.194}$$

$$= \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \text{sgn}(\boldsymbol{\mu}) \text{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \zeta_{\mu_t}^{(\boldsymbol{\iota})}(\lambda_t) \varphi(\lambda_t) \lambda_t^{\nu_t - 1} \tag{2.195}$$

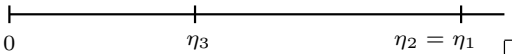
[26] Recall that by definition $i_0 = 0$, $i_{n+1} = n + 1$ and $\eta_{n+1} = 0$.



λ_3	λ_2	λ_1
$\lambda_3 \leq \lambda_2$		λ_1
λ_3	$\lambda_2 \leq \lambda_1$	
$\lambda_3 \leq \lambda_2$	λ_1	
$\lambda_3 \leq \lambda_2 \leq \lambda_1$		

$\boldsymbol{i}$	$\tau(\boldsymbol{i})$
$(1,2,3)$	1!1!1!
$(1,3,3)$	1!0!2!
$(2,2,3)$	0!2!1!
$(2,3,3)$	0!1!2!
$(3,3,3)$	0!0!3!

(i) $n = 3$ and $\eta_1 > \eta_2 > \eta_3$.



λ_3	$\lambda_2 \leq \lambda_1$
$\lambda_3 \leq \lambda_2$	λ_1
$\lambda_3 \leq \lambda_2 \leq \lambda_1$	

$\boldsymbol{i}$	$\tau(\boldsymbol{i})$
$(2,2,3)$	0!2!1!
$(2,3,3)$	0!1!2!
$(3,3,3)$	0!0!3!

(ii) $n = 3$ and $\eta_1 = \eta_2 > \eta_3$.

Figure 2.6 Integration region $\mathcal{D}_{\text{ord}}(\boldsymbol{i};\boldsymbol{\eta})$ and normalizing constant $\tau(\boldsymbol{i})$.

where the summation over $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ is for all permutations of the integers $(1, \ldots, n)$ and $\text{sgn}(\cdot)$ denotes the sign of the permutation. Substituting (2.195) back in (2.193) and defining $\xi_{u,v}^{(\iota)} = \zeta_u^{(\boldsymbol{\iota})}(\lambda_t)\varphi(\lambda_t)\lambda_t^{v-1}$, it follows that

$$\mathcal{I}(\boldsymbol{i};\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \int_{\mathcal{D}_{\text{ord}}(\boldsymbol{i};\boldsymbol{\eta})} \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \text{sgn}(\boldsymbol{\mu})\text{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda_t) d\boldsymbol{\lambda}. \tag{2.196}$$

Using the symmetry of the integrand in (2.196) with respect to $\boldsymbol{\lambda}$, i.e.,

$$\sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \text{sgn}(\boldsymbol{\mu})\text{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda_t) = \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \text{sgn}(\boldsymbol{\mu})\text{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda_{\sigma_t}) \tag{2.197}$$

where $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is any arbitrary fixed permutation of the integers $(1, \ldots, n)$, the domain $\mathcal{D}_{\text{ord}}(\boldsymbol{i};\boldsymbol{\eta})$ in (2.196) can be replaced with the unordered domain $\mathcal{D}(\boldsymbol{i};\boldsymbol{\eta})$ (see a similar result in [Kha69, Lem. 2]) by properly normalizing the result of the integral with

$$\tau(\boldsymbol{i}) = \prod_{s=1}^{o(\boldsymbol{i})} \tau_s(\boldsymbol{i})! = \prod_{u=1}^{n} (1 - \delta_{i_u,i_{u+1}}) \Big(\sum_{v=1}^{u} \delta_{i_u,i_v} \Big)! \tag{2.198}$$

where $o(\boldsymbol{i})$ denotes the number of different integration intervals in $\mathcal{D}(\boldsymbol{i};\boldsymbol{\eta})$, $\tau_s(\boldsymbol{i})$ is the number of ordered variables integrated in the sth one of these intervals (see Figure 2.6) and $\delta_{u,v}$ denotes de Kronecker delta. Then, it holds that

$$\mathcal{I}(\boldsymbol{i};\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota}\in\mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{\tau(\boldsymbol{i})} \int_{\mathcal{D}(\boldsymbol{i};\boldsymbol{\eta})} \prod_{t=1}^{n} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda_t) d\boldsymbol{\lambda} = \sum_{\boldsymbol{\iota}\in\mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{\tau(\boldsymbol{i})} \prod_{t=1}^{n} \int_{\eta_{i_t+1}}^{\eta_{i_t}} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda) d\lambda. \tag{2.199}$$

Finally, $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta})$ is given by

$$F_{\boldsymbol{\lambda}}(\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{i}\in\mathcal{S}} \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu})\mathrm{sgn}(\boldsymbol{\nu}) \frac{1}{\tau(\boldsymbol{i})} \prod_{t=1}^{n} \int_{\eta_{i_t+1}}^{\eta_{i_t}} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda) d\lambda \tag{2.200}$$

and the proof is completed by using operator $\mathcal{T}\{\cdot\}$ in Definition 2.4. □

2.B Appendix: Marginal cdf and pdf of the Ordered Eigenvalues

2.B.1 Proof of Theorem 2.2

Proof. The marginal cdf of the kth largest eigenvalue, $F_{\lambda_k}(\eta)$, is given by

$$F_{\lambda_k}(\eta) = \mathsf{Pr}(\lambda_k \leq \eta) = \mathsf{Pr}(\lambda_1 \leq \infty, \ldots, \lambda_{k-1} \leq \infty, \lambda_k \leq \eta, \ldots, \lambda_n \leq \eta). \tag{2.201}$$

Hence, it can be obtained from the joint cdf of the ordered eigenvalues $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta})$ for $\{\eta_i\}_{i=1,\ldots,k-1} = \infty$ and $\{\eta_i\}_{i=k,\ldots,n} = \eta > 0$. Particularizing the expression of $F_{\boldsymbol{\lambda}}(\boldsymbol{\eta})$ in Theorem 2.1, we have that the set $\mathcal{S}$ reduces to

$$\mathcal{S}_k = \{\boldsymbol{i} \in \{k-1, n\}^n | i_s = n \text{ if } s \geq k\} \tag{2.202}$$

which only contains k elements. Let us denote by an unique index i $(i = 1, \ldots, k)$ each element of $\mathcal{S}_k$ such that $\boldsymbol{i}(i)$ has the first $i-1$ components equal to $k-1$ and the rest equal to n. Noting that

$$\tau(\boldsymbol{i}(i)) = (i-1)!(n-i+1)! \tag{2.203}$$

and using the alternative expression of the operator $\mathcal{T}(\cdot)$ in Remark 2.1, it follows that

$$F_{\lambda_k}(\eta) = \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{i=1}^{k} \frac{1}{(i-1)!(n-i+1)!} \sum_{\boldsymbol{\mu}} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)| \tag{2.204}$$

where the summation over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ is for all permutations of the integers $(1, \ldots, n)$ and the $n \times n$ matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)$ is defined as (see (2.68))

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)]_{u,\mu_v} = [\mathbf{T}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, \boldsymbol{i}(i); \eta)]_{u,\mu_v,v} = \begin{cases} \int_{\eta}^{\infty} \xi_{u,\mu_v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & 1 \leq v < i \\ \int_{0}^{\eta} \xi_{u,\mu_v}^{(\boldsymbol{\iota})}(\lambda) d\lambda & i \leq v \leq n \end{cases} \quad \text{for } u, v = 1, \ldots, n. \tag{2.205}$$

Observing that $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta) = \mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\nu}, i; \eta)$ if $\boldsymbol{\nu} = (\pi(\mu_1, \ldots, \mu_{i-1}), \pi(\mu_i, \ldots, \mu_n))$ where $\pi(\cdot)$ denotes permutation, it suffices to calculate one determinant for all these $(i-1)!(n-i+1)!$

permutations, for instance, $|\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)|$ where $\boldsymbol{\mu}$ is such that $(\mu_1 < \ldots < \mu_{i-1})$ and $(\mu_i < \ldots < \mu_n)$. Finally, we have that

$$\frac{1}{(i-1)!(n-i+1)!} \sum_{\boldsymbol{\mu}} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)| = \sum_{\boldsymbol{\mu} \in \mathcal{P}(i)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, i; \eta)| \tag{2.206}$$

and this completes the proof. □

2.B.2 Proof of Corollary 2.2.1

Proof. Particularizing Theorem 2.2 for $k = 1$, it follows that

$$F_{\lambda_1}(\eta) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \sum_{\boldsymbol{\mu} \in \mathcal{P}(1)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, 1; \eta)|. \tag{2.207}$$

Observe now that $\mathcal{P}(1)$ only contains the element $(1, \ldots, n)$. Thus, using (2.70), we define the $n \times n$ matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\eta)$ as

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\eta)]_{u,v} = [\mathbf{F}^{(\boldsymbol{\iota})}((1, \ldots, n), 1; \eta)]_{u,v} = \int_0^{\eta} \xi_{u,v}^{(\boldsymbol{\iota})}(\lambda) d\lambda \qquad \text{for } u, v = 1, \ldots, n \tag{2.208}$$

and this completes the proof. □

2.B.3 Proof of Corollary 2.2.2

Proof. This proof could be done by particularizing Theorem 2.2 for $k = n$ and simplifying the resulting expression. However, it is easier to obtain $F_{\lambda_n}(\eta)$ directly as

$$F_{\lambda_n}(\eta) = 1 - \mathsf{Pr}\left(\lambda_n > \eta\right) = 1 - \mathsf{Pr}\left(\lambda_1 > \eta, \ldots, \lambda_n > \eta\right) = 1 - \int_{\mathcal{D}_{\mathrm{ord}}(\eta)} f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) d\boldsymbol{\lambda} \tag{2.209}$$

where

$$\mathcal{D}_{\mathrm{ord}}(\eta) = \{\lambda_1 > \eta, \ldots, \lambda_n > \eta\} \cap \{\lambda_1 > \cdots > \lambda_n\}. \tag{2.210}$$

Then, using the expression for the joint pdf $f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda})$ given in Assumption 2.2 and substituting operator $\mathcal{T}\{\cdot\}$ by its definition (Definition 2.4), it follows that

$$F_{\lambda_n}(\eta) = 1 - \sum_{\boldsymbol{\iota} \in \mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \int_{\mathcal{D}_{\mathrm{ord}}(\eta)} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \mathrm{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \xi_{\mu_t, \nu_t}^{(\boldsymbol{\iota})}(\lambda_t) d\boldsymbol{\lambda} \tag{2.211}$$

where the summation over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ is for all permutations of the integers $(1, \ldots, n)$ and $\mathrm{sgn}(\cdot)$ denotes the sign of the permutation. Noting the symmetry of the

integrand in (2.211)

$$F_{\lambda_n}(\eta) = 1 - \sum_{\boldsymbol{\iota}\in\mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{n!} \sum_{\boldsymbol{\mu},\boldsymbol{\nu}} \operatorname{sgn}(\boldsymbol{\mu})\operatorname{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \int_{\eta}^{\infty} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda_t) d\lambda \tag{2.212}$$

$$= 1 - \sum_{\boldsymbol{\iota}\in\mathcal{I}} K_{m,n}^{(\boldsymbol{\iota})} \operatorname{sgn}(\boldsymbol{\mu}) \sum_{\boldsymbol{\nu}} \operatorname{sgn}(\boldsymbol{\nu}) \prod_{t=1}^{n} \int_{\eta}^{\infty} \xi_{\mu_t,\nu_t}^{(\boldsymbol{\iota})}(\lambda) d\lambda \tag{2.213}$$

and using the definition of determinant (see Definition 2.2) the proof is completed. □

2.B.4 Proof of Corollary 2.2.3

Proof. The marginal pdf of the kth largest eigenvalue, $f_{\lambda_k}(\eta)$, can be obtained from its marginal cdf as

$$f_{\lambda_k}(\eta) = \frac{d}{d\eta} F_{\lambda_k}(\eta). \tag{2.214}$$

Then, the proof follows from using the expression for the marginal cdf of the kth largest largest eigenvalue in Theorem 2.2 and the derivative of a determinant in Lemma 2.2. □

2.C Appendix: Marginal cdf of the Maximum Weighted Ordered Eigenvalue. Proof of Theorem 2.3

Proof. The cdf of the random variable $\lambda_{\mathcal{K}}$ defined in (2.77), $F_{\lambda_{\mathcal{K}}}(\eta)$, can be obtained as

$$F_{\lambda_{\mathcal{K}}}(\eta) = \mathsf{Pr}\Big(\max_{k\in\mathcal{K}} \big(\lambda_k/\theta_k\big) \leq \eta\Big) = \mathsf{Pr}\big(\lambda_{\mathcal{K}_1} \leq \theta_1\eta, \ldots, \lambda_{\mathcal{K}_{|\mathcal{K}|}} \leq \theta_{|\mathcal{K}|}\eta\big) \tag{2.215}$$

where $\mathcal{K}_i$ denotes the ith element and $|\mathcal{K}|$ the cardinality of the set $\mathcal{K}$. Defining

$$\tilde{\vartheta}_k(\eta) = \begin{cases} \theta_k\eta & k \in \mathcal{K} \\ \infty & \text{otherwise} \end{cases} \qquad \text{for } k = 1, \ldots, n \tag{2.216}$$

we can rewritte the cdf in (2.215) as

$$F_{\lambda_{\mathcal{K}}}(\eta) = \mathsf{Pr}\big(\lambda_1 \leq \tilde{\vartheta}_1(\eta), \ldots, \lambda_n \leq \tilde{\vartheta}_n(\eta)\big) = F_{\boldsymbol{\lambda}}\big(\tilde{\vartheta}_1(\eta), \ldots, \tilde{\vartheta}_n(\eta)\big) \tag{2.217}$$

where $F_{\boldsymbol{\lambda}}(\cdot)$ denotes the joint cdf of the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. Precisely, due to the order of the eigenvalues, for $\eta_{k-1} < \eta_k$, it holds that

$$F_{\boldsymbol{\lambda}}(\eta_1, \ldots, \eta_{k-1}, \eta_k, \ldots, \eta_n) = F_{\boldsymbol{\lambda}}(\eta_1, \ldots, \eta_{k-1}, \eta_{k-1}, \ldots, \eta_n) \tag{2.218}$$

and, hence, we have that

$$F_{\lambda_{\mathcal{K}}}(\eta) = F_{\boldsymbol{\lambda}}\big(\vartheta_1(\eta), \ldots, \vartheta_n(\eta)\big) \tag{2.219}$$

where $\vartheta_k(\eta)$ is defined in (2.79) and this completes the proof. □

2.D Appendix: Taylor Expansions

2.D.1 Proof of Theorem 2.4

Proof. The first order Taylor expansion of $F_{\lambda_k}(\eta)$ is given by

$$F_{\lambda_k}(\eta) = \left(\frac{F^{(r)}_{\lambda_k}(\eta)|_{\eta=0}}{r!}\right)\eta^r + o(\eta^r) \tag{2.220}$$

where $F^{(r)}_{\lambda_k}(\eta)$ denotes the rth derivative of $F_{\lambda_k}(\eta)$ given in Theorem 2.2 and r is the smallest integer such that $F^{(r)}_{\lambda_k}(\eta)|_{\eta=0} \neq 0$. Using Lemma 2.3 for the rth derivative of a determinant, $F^{(r)}_{\lambda_k}(\eta)$ can be expressed as

$$F^{(r)}_{\lambda_k}(\eta) = K_{m,n}\sum_{i=1}^{k}\sum_{\boldsymbol{\mu}\in\mathcal{P}(i)}\sum_{\mathbf{r}}\frac{r!}{r_1!\cdots r_n!}|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu},i;\eta)| \tag{2.221}$$

where the summation over $\mathbf{r}=(r_1,\ldots,r_n)$ is for all $\mathbf{r}$ such that $r_s \in \mathbb{N}\cup\{0\}$ and $\sum_{s=1}^{n} r_s = r$, and the $n\times n$ matrix $\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu},i;\eta)$ is defined as (see (2.70))

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu},i;\eta)]_{u,\mu_v} = \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}}[\mathbf{F}(\boldsymbol{\mu},i;\eta)]_{u,\mu_v} = \begin{cases} \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}}\int_{\eta}^{\infty}\xi_{u,\mu_v}(\lambda)d\lambda & 1\le v<i \\ \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}}\int_{0}^{\eta}\xi_{u,\mu_v}(\lambda)d\lambda & i\le v\le n \end{cases} \tag{2.222}$$

for $u,v=1,\ldots,n$. The proof reduces to find the minimum integer r such that $F^{(r)}_{\lambda_k}(\eta)$ in (2.221) does not equal 0 when evaluated at $\eta=0$. First we determine, for a fixed i and a fixed permutation $\boldsymbol{\mu}$, the set $\{r_{\mu_v}\}_{v=1,\ldots,n}$ with minimum $r(\boldsymbol{\mu},i)=\sum_{v=1}^{n} r_{\mu_v}$ such that $|\mathbf{F}^{(\boldsymbol{\iota},\mathbf{r})}(\boldsymbol{\mu},i;\eta=0)|\neq 0$ and, then, we obtain r as

$$r = \min_{1\le i\le k}\min_{\boldsymbol{\mu}\in\mathcal{P}(i)} r(\boldsymbol{\mu},i). \tag{2.223}$$

Recall from the statement of the theorem that the Taylor expansion of $\zeta_u(\lambda)\varphi(\lambda)$ is given by

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c(u)}^{\infty}\frac{a_u(t)}{t!}\lambda^t \tag{2.224}$$

and, since $\xi_{u,v}(\lambda)=\zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1}$, we have that

$$\int_{\eta}^{\infty}\xi_{u,v}(\lambda)d\lambda = \int_{0}^{\infty}\xi_{u,v}(\lambda)d\lambda - \int_{0}^{\eta}\xi_{u,v}(\lambda)d\lambda = b_{u,v} - \sum_{t=c(u)+v}^{\infty}\frac{a_{u,v}(t)}{t!}\eta^t = b_{u,v}+o(\eta^0) \tag{2.225}$$

$$\int_{0}^{\eta}\xi_{u,v}(\lambda)d\lambda = \sum_{t=c(u)+v}^{\infty}\frac{a_{u,v}(t)}{t!}\eta^t = \frac{a_{u,v}(c(u)+v)}{(c(u)+v)!}\eta^{c(u)+v} + o(\eta^{c(u)+v}) \tag{2.226}$$

where we have defined

$$a_{u,v}(t) = \frac{(t-1)!}{(t-v)!} a_u(t-v) \tag{2.227}$$

$$b_{u,v} = \int_0^\infty \xi_{u,v}(\lambda) d\lambda. \tag{2.228}$$

From (2.225) and (2.226) we conclude that r_{μ_v}, such that the columns $[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)]_{\mu_v}$ do not have all entries equal to 0, satisfies

$$\begin{cases} r_{\mu_v} = 0 \text{ or } r_{\mu_v} \geq c + \mu_v & 1 \leq v < i \\ r_{\mu_v} \geq c + \mu_v & i \leq v \leq n \end{cases} \tag{2.229}$$

where $c = \min_u c(u)$.

Note that the condition in (2.229) is only a necessary condition, as we still have to guaratee that all columns of $\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta)$ are linearly independent in order to assure that $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)| \neq 0$. In fact, the condition in (2.229) is not sufficient, as in the following we show that the set $\{r_{\mu_v}\}_{v=1,\ldots,n}$ with minimum $r(\boldsymbol{\mu}, i)$ and $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)| \neq 0$ is given by

$$r_{\mu_v} = \begin{cases} 0 & 1 \leq v < i \\ c + \mu_v + \nu_{v-i+1} - 1 & i \leq v \leq n \end{cases} \tag{2.230}$$

where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_{n-i+1})$ is a permutation of integers $(1, \ldots, \nu_{n-i+1})$.

Let us focus first on the case $1 \leq v < i$ with $r_{\mu_v} = 0$, for which it holds that

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)]_{u,\mu_v} = \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}} \int_\eta^\infty \xi_{u,\mu_v}(\lambda) d\lambda \Big|_{\eta=0} = b_{u,\mu_v} \qquad \text{for } u = 1, \ldots, n. \tag{2.231}$$

For the case $i \leq v < n$ we have that

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)]_{\mu_v} = \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}} \int_\eta^\infty \xi_{u,\mu_v}(\lambda) d\lambda \Big|_{\eta=0} \tag{2.232}$$

$$= a_{u,\mu_v}(r_{\mu_v}) = \frac{(r_{\mu_v} - 1)!}{(r_{\mu_v} - \mu_v)!} a_u(r_{\mu_v} - \mu_v) \qquad \text{for } u = 1, \ldots, n. \tag{2.233}$$

and this, noting (2.229), shows that all r_{μ_v} in the set $\{r_{\mu_v}\}_{v=i,\ldots,n}$ have to be different. Since the set $\{r_{\mu_v}\}_{v=i,\ldots,n}$ with different elements and minimum $r(\boldsymbol{\mu}, i)$ is

$$r_{\mu_v} = c + \mu_v + \nu_{v-n-i+1} - 1 \tag{2.234}$$

this confirms (2.230) as long as $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)| \neq 0$ with

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)]_{u,\mu_v} = \begin{cases} b_{u,\mu_v} & 1 \leq v < i \\ a_{u,\mu_v}(c + \mu_v + \nu_{v-n-i+1} - 1) & i \leq v \leq n \end{cases}. \tag{2.235}$$

Now observe that $r = \min_{1\leq i\leq k} \min_{\boldsymbol{\mu}\in\mathcal{P}(i)} r(\boldsymbol{\mu}, i)$, with r_{μ_v} as given in (2.230), is obtained for $i = k$ and $\boldsymbol{\mu}$ such that

$$\mu_v = \begin{cases} n+v-k+1 & 1 \leq v < k \\ v-k+1 & k \leq v \leq n \end{cases} \tag{2.236}$$

and that, at least in this case, $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \eta = 0)| \neq 0$ by the assumption in the statement of the theorem. Thus, r is

$$r = \sum_{v=1}^{n} r_v = (c-1)(n-k+1) + 2\sum_{v=1}^{n-k+1} v = (c+n-k+1)(n-k+1) \tag{2.237}$$

where we have used [Gra00, eq. (0.121.1)]. Finally, we can rewrite $F_{\lambda_k}^{(r)}(\eta)|_{\eta=0}$ as

$$F_{\lambda_k}^{(r)}(\eta)\Big|_{\eta=0} = K_{m,n} \sum_{\boldsymbol{\nu}} \frac{r!}{\tau(\boldsymbol{\nu})} |\mathbf{F}(\boldsymbol{\nu})| \tag{2.238}$$

where

$$\tau(\boldsymbol{\nu}) = \prod_{v=1}^{n-k+1} (c+v+\nu_v-1)! \tag{2.239}$$

and the $n \times n$ matrix $\mathbf{F}(\boldsymbol{\nu})$ is defined, using (2.235) and (2.230), as

$$[\mathbf{F}(\boldsymbol{\nu})]_{u,\mu_v} = \begin{cases} b_{u,\mu_v} & 1 \leq v < k \\ a_{u,\mu_v}(c+\mu_v+\nu_{v-k+1}-1) & k \leq v \leq n \end{cases} \quad \text{for } u, v = 1, \ldots, n \tag{2.240}$$

or, equivalently, using (2.236) and (2.227), as

$$[\mathbf{F}(\boldsymbol{\nu})]_{u,v} = \begin{cases} \frac{(c+v+\nu_v-2)!}{(c+\nu_v-1)!} a_u(c+\nu_v-1) & 1 \leq v \leq n-k+1 \\ b_{u,v} & n-k+1 < v \leq n \end{cases} \quad \text{for } u, v = 1, \ldots, n. \tag{2.241}$$

Then, we complete the proof by substituting (2.238) back in (2.220). □

2.D.2 Proof of Theorem 2.5

Proof. The first order Taylor expansion of $F_{\lambda_\mathcal{K}}(\eta)$ is given by

$$F_{\lambda_\mathcal{K}}(\eta) = \left(\frac{F_{\lambda_\mathcal{K}}^{(r)}(\eta)|_{\eta=0}}{r!} \right) \eta^r + o(\eta^r) \tag{2.242}$$

where $F_{\lambda_\mathcal{K}}^{(r)}(\eta)$ denotes the rth derivative of $F_{\lambda_\mathcal{K}}(\eta)$ given in Theorem 2.3 and r is the smallest integer such that $F_{\theta_\mathcal{K}}^{(r)}(\eta)|_{\eta=0} \neq 0$. Using the alternative expression of operator $\mathcal{T}\{\cdot\}$ in Remark

2.1, we can rewrite $F_{\lambda_{\mathcal{K}}}(\eta)$ in (2.78) in terms of sum of determinants. Then, using Lemma 2.3 for the rth derivative of a determinant, it follows

$$F^{(r)}_{\lambda_{\mathcal{K}}}(\eta) = K_{m,n} \sum_{\boldsymbol{i}\in\mathcal{S}} \frac{1}{\tau(\boldsymbol{i})} \sum_{\boldsymbol{\mu}} \sum_{\mathbf{r}} \frac{r!}{r_1! \cdots r_n!} |\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta))| \tag{2.243}$$

where the summation over $\mathbf{r} = (r_1, \ldots, r_n)$ is for all $\mathbf{r}$ such that $r_s \in \mathbb{N} \cup \{0\}$ and $\sum_{s=1}^{n} r_s = r$, and the $n \times n$ matrix $\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta))$ is defined as (see (2.68))

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta))]_{u,\mu_v} = \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}} [\mathbf{T}(\boldsymbol{i}; \boldsymbol{\vartheta}(\eta))]_{u,\mu_v,v} = \frac{d^{r_{\mu_v}}}{d\eta^{r_{\mu_v}}} \int_{\vartheta_{i_v+1}\eta}^{\vartheta_{i_v}\eta} \xi_{u,\mu_v}(\lambda) d\lambda \tag{2.244}$$

for $u, v = 1, \ldots, n$. Then, the proof reduces to find the minimum integer r such that $F^{(r)}_{\lambda_{\mathcal{K}}}(\eta)$ in (2.243) does not equal 0 when evaluated at $\eta = 0$.

Recall from Assumption 2.3 that the Taylor expansion of $\zeta_u(\lambda)\varphi(\lambda)$ is given by

$$\zeta_u(\lambda)\varphi(\lambda) = \sum_{t=c(u)}^{\infty} \frac{a_u(t)}{t!} \lambda^t \tag{2.245}$$

and, since $\xi_{u,v}(\lambda) = \zeta_u(\lambda)\varphi(\lambda)\lambda^{v-1}$, we have that

$$\int_{\vartheta_{i_v+1}\eta}^{\vartheta_{i_v}\eta} \xi_{u,\mu_v}(\lambda) d\lambda = \int_{0}^{\vartheta_{i_v}\eta} \xi_{u,\mu_v}(\lambda) d\lambda - \int_{0}^{\vartheta_{i_v+1}\eta} \xi_{u,\mu_v}(\lambda) d\lambda \tag{2.246}$$

$$= \sum_{t=c(u)+\mu_v}^{\infty} \frac{a_u(t-\mu_v)}{t(t-\mu_v)!} (\vartheta^t_{i_v} - \vartheta^t_{i_v+1}) \eta^t \tag{2.247}$$

From (2.246) we conclude that r_{μ_v}, such that the columns $[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta = 0))]_{\mu_v}$ do not have all entries equal to 0, satisfies

$$r_{\mu_v} \geq c + \mu_v \tag{2.248}$$

where $c = \min_u c(u)$. Note that the condition in (2.248) is only a necessary condition, as we still have to guaratee that all columns of $\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta))$ are linearly independent in order to assure that $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta = 0))| \neq 0$. In fact, the condition in (2.248) is not sufficient, since we have that

$$[\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta}(\eta = 0))]_{\mu_v} = \frac{(r_{\mu_v} - 1)!}{(r_{\mu_v} - \mu_v)!} a_u(r_{\mu_v} - \mu_v) \left(\vartheta^{r_{\mu_v}}_{i_v} - \vartheta^{r_{\mu_v}}_{i_v+1}\right) \qquad \text{for } u = 1, \ldots, n \tag{2.249}$$

and this, noting (2.248), shows that all r_{μ_v} in the set $\{r_{\mu_v}\}_{v=i,\ldots,n}$ have to be different. The set $\{r_{\mu_v}\}_{v=i,\ldots,n}$ with different elements is

$$r_{\mu_v} = c + \mu_v + \nu_v - 1 \tag{2.250}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is a permutation of integers $(1, \ldots, n)$. In addition, due to Assumption 2.3, this set ensures that $|\mathbf{F}^{(\mathbf{r})}(\boldsymbol{\mu}, i; \boldsymbol{\vartheta}(\eta = 0))| \neq 0$, at least when $\boldsymbol{\mu} = \boldsymbol{\nu}$. Thus, independently of $\boldsymbol{i}$ and $\boldsymbol{\mu}$, r is

$$r = \sum_{v=1}^{n} r_{\mu_v} = (c-1)n + (n+1)n = (c+n)n \tag{2.251}$$

where we have used [Gra00, eq. (0.121.1)]. Finally, we can rewrite $F^{(r)}_{\lambda_K}(\eta)|_{\eta=0}$ as

$$F^{(r)}_{\lambda_K}(\eta)\Big|_{\eta=0} = K_{m,n} \sum_{\boldsymbol{i}\in\mathcal{S}} \frac{1}{\tau(\boldsymbol{i})} \sum_{\boldsymbol{\mu}} \sum_{\boldsymbol{\nu}} \frac{r!}{\tau(\boldsymbol{\mu},\boldsymbol{\nu})} |\mathbf{F}(\boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{i}; \boldsymbol{\vartheta})| \tag{2.252}$$

where

$$\tau(\boldsymbol{\mu}, \boldsymbol{\nu}) = \prod_{v=1}^{n} (c + \mu_v + \nu_v - 1)! \tag{2.253}$$

and the $n \times n$ matrix $\mathbf{F}(\boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{i}; \boldsymbol{\vartheta})$ is defined as

$$[\mathbf{F}(\boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{i}; \boldsymbol{\vartheta})]_{u,\mu_v} = \frac{(c + \mu_v + \nu_v - 2)!}{(c + \nu_v - 1)!} a_u (c + \nu_v - 1) \left(\vartheta_{i_v}^{r_{\mu_v}} - \vartheta_{i_v+1}^{r_{\mu_v}}\right) \qquad \text{for } u, v = 1, \ldots, n. \tag{2.254}$$

Finally, using again Remark 2.1, it follows

$$F^{(r)}_{\lambda_K}(\eta)\Big|_{\eta=0} = K_{m,n} \sum_{\boldsymbol{i}\in\mathcal{S}} \frac{r!}{\tau(\boldsymbol{i})} \sum_{\boldsymbol{\mu}} \mathcal{T}\{\mathbf{T}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta})\} \tag{2.255}$$

where the $n \times n \times n$ tensor $\mathbf{T}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta})$ is defined as

$$[\mathbf{T}(\boldsymbol{\mu}, \boldsymbol{i}; \boldsymbol{\vartheta})]_{u,v,t} = \frac{1}{(c + \mu_v + v - 1)(c + v - 1)!} a_u (c+v-1) \left(\vartheta_{i_t}^{r_v} - \vartheta_{i_t+1}^{r_v}\right) \qquad \text{for } u, v, t = 1, \ldots, n. \tag{2.256}$$

Then, we complete the proof by substituting (2.255) back in (2.242). □

2.E Appendix: Joint and Marginal cdf and pdf of the Unordered Eigenvalues

2.E.1 Proof of Theorem 2.6

Proof. The joint pdf of the unordered eigenvalues, $f_{\mathbf{x}}(\mathbf{x})$, is obtained by dividing the joint pdf of the ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, $f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda})$, by $n!$, i.e.,

$$f_{\mathbf{x}}(\mathbf{x}) = \frac{1}{n!} f_{\boldsymbol{\lambda}}(\mathbf{x}) \tag{2.257}$$

so that $f_{\mathbf{x}}(\mathbf{x})$ is a density function in the unordered domain $\{0 \leq x_1 \leq \infty, \ldots, 0 \leq x_n \leq \infty\}$. The proof is then completed by using the expression for the joint pdf of the ordered eigenvalues given in Assumption 2.1. □

2.E.2 Proof of Theorem 2.7

Proof. The joint cdf of $1 \leq p \leq n$ unordered eigenvalues, $F_{\mathbf{x}_p}(\boldsymbol{\eta})$, is obtained from the joint pdf of the unordered eigenvalues, $f_{\mathbf{x}}(\mathbf{x})$, as

$$F_{\mathbf{x}_p}(\boldsymbol{\eta}) = \int_{\mathcal{D}} f_{\mathbf{x}}(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n) d\mathbf{x} \tag{2.258}$$

where $\mathcal{D} = \{0 \leq x_1 \leq \eta_1, \ldots 0 \leq x_p \leq \eta_p, \ldots, 0 \leq x_n \leq \infty\}$. Using the expression for $f_{\mathbf{x}}(\mathbf{x})$ given in Theorem 2.6 and expanding the determinants (see Definition 2.2), it follows that

$$F_{\mathbf{x}_p}(\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K^{(\boldsymbol{\iota})}_{m,n}}{n!} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \mathrm{sgn}(\boldsymbol{\nu}) \int_{\mathcal{D}} \left[\prod_{t=1}^{p} \xi^{(\boldsymbol{\iota})}_{\mu_t, \nu_t}(x_t) \right] \left[\prod_{t=p+1}^{n} \xi^{(\boldsymbol{\iota})}_{\mu_t, \nu_t}(x_t) \right] d\mathbf{x} \tag{2.259}$$

$$= \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K^{(\boldsymbol{\iota})}_{m,n}}{n!} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \mathrm{sgn}(\boldsymbol{\nu}) \left[\prod_{t=1}^{p} \int_0^{\eta_t} \xi^{(\boldsymbol{\iota})}_{\mu_t, \nu_t}(x) dx \right] \left[\prod_{t=p+1}^{n} \int_0^{\infty} \xi^{(\boldsymbol{\iota})}_{\mu_t, \nu_t}(x) dx \right] \tag{2.260}$$

where the summation over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ is for all permutations of the integers $(1, \ldots, n)$ and $\mathrm{sgn}(\cdot)$ denotes the sign of the permutation. Now, using the alternative expression of the operator $\mathcal{T}\{\cdot\}$ in Remark 2.1, we can rewrite (2.266) as

$$F_{\mathbf{x}_p}(\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K^{(\boldsymbol{\iota})}_{m,n}}{n!} \sum_{\boldsymbol{\mu}} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| \tag{2.261}$$

where the summation over $\boldsymbol{\mu}$ is over all permutations of integers $(1, \ldots, n)$ and the $n \times n$ matrix $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})$ is defined as

$$[\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})]_{u, \mu_v} = \begin{cases} \int_0^{\eta_v} \xi^{(\boldsymbol{\iota})}_{u, \mu_v}(x) dx & 1 \leq v \leq p \\ \int_0^{\infty} \xi^{(\boldsymbol{\iota})}_{u, \mu_v}(x) dx & p < v \leq n \end{cases} \qquad \text{for } u, v = 1, \ldots, n. \tag{2.262}$$

Observing that $\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta}) = \mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\nu}, p; \boldsymbol{\eta})$ if $\boldsymbol{\nu} = (\pi(\mu_1, \ldots, \mu_p), \pi(\mu_{p+1}, \ldots, \mu_n))$ where $\pi(\cdot)$ denotes permutation, it suffices to calculate one determinant for all these $p!(n-p)!$ permutations, for instance, $|\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})|$ where $\boldsymbol{\mu}$ is such that $(\mu_1 < \ldots < \mu_p)$ and $(\mu_{p+1} < \ldots < \mu_n)$. Finally, we have that

$$\sum_{\boldsymbol{\mu}} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| = p!(n-p)! \sum_{\boldsymbol{\mu} \in \mathcal{P}(p+1)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| = n! \binom{n}{p}^{-1} \sum_{\boldsymbol{\mu} \in \mathcal{P}(p+1)} |\mathbf{F}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| \tag{2.263}$$

where $\binom{n}{p}$ denotes the binomial coefficient [Abr72, eq. (3.1.2)] and this completes the proof. □

2.E.3 Proof of Corollary 2.7.1

Proof. The joint pdf of $1 \leq p \leq n$ unordered eigenvalues, $f_{\mathbf{x}_p}(\boldsymbol{\eta})$, is obtained from the joint pdf of the unordered eigenvalues, $f_{\mathbf{x}}(\mathbf{x})$, by integrating out $n-p$ eigenvalues, i.e.,

$$f_{\mathbf{x}_p}(\boldsymbol{\eta}) = \int_{\mathcal{D}} f_{\mathbf{x}}(\eta_1, \ldots, \eta_p, x_{p+1}, \ldots, x_n) dx_{p+1} \cdots dx_n \tag{2.264}$$

where $\mathcal{D} = \{0 \leq x_{p+1} \leq \infty, \ldots, 0 \leq x_n \leq \infty\}$. Using the expression for $f_{\mathbf{x}}(\mathbf{x})$ given in Theorem 2.6 and expanding the determinants, it follows that

$$f_{\mathbf{x}_p}(\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{n!} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \mathrm{sgn}(\boldsymbol{\nu}) \left[\prod_{t=1}^{p} \xi_{\mu_t, \nu_t}^{(\boldsymbol{\iota})}(\eta_t) \right] \left[\int_{\mathcal{D}} \prod_{t=p+1}^{n} \xi_{\mu_t, \nu_t}^{(\boldsymbol{\iota})}(x_t) dx_{p+1} \cdots dx_n \right] \tag{2.265}$$

$$= \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{n!} \sum_{\boldsymbol{\mu}, \boldsymbol{\nu}} \mathrm{sgn}(\boldsymbol{\mu}) \mathrm{sgn}(\boldsymbol{\nu}) \left[\prod_{t=1}^{p} \xi_{\mu_t, \nu_t}^{(\boldsymbol{\iota})}(\eta_t) \right] \left[\prod_{t=p+1}^{n} \int_0^{\infty} \xi_{\mu_t, \nu_t}^{(\boldsymbol{\iota})}(x) dx \right] \tag{2.266}$$

where the summation over $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)$ is for all permutations of the integers $(1, \ldots, n)$ and $\mathrm{sgn}(\cdot)$ denotes the sign of the permutation. Now, using the alternative expression of the operator $\mathcal{T}\{\cdot\}$ in Remark 2.1, we can rewrite (2.266) as

$$f_{\mathbf{x}_p}(\boldsymbol{\eta}) = \sum_{\boldsymbol{\iota} \in \mathcal{I}} \frac{K_{m,n}^{(\boldsymbol{\iota})}}{n!} \sum_{\boldsymbol{\mu}} |\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})| \tag{2.267}$$

where the summation over $\boldsymbol{\mu}$ is over all permutations of integers $(1, \ldots, n)$ and the $n \times n$ matrix $\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})$ is defined as

$$[\mathbf{D}^{(\boldsymbol{\iota})}(\boldsymbol{\mu}, p; \boldsymbol{\eta})]_{u,v} = \begin{cases} \xi_{u,\mu_v}^{(\boldsymbol{\iota})}(\eta_v) & 1 \leq v \leq p \\ \int_0^{\infty} \xi_{u,\mu_v}^{(\boldsymbol{\iota})}(x) dx & p < v \leq n \end{cases} \qquad \text{for } u, v = 1, \ldots, n. \tag{2.268}$$

Finally, following the same procedure as in the proof of Theorem 2.7 (see 2.E.2) the proof is completed. □

3

Spatial Multiplexing MIMO Systems with CSI: Performance Analysis

Spatial multiplexing is a simple multiple-in multiple-out (MIMO) transmit technique that allows a high spectral efficiency by dividing the incoming data into multiple independent substreams which are simultaneously transmitted. When perfect channel state information (CSI) is available at both sides of the link, channel-dependent linear processing of the data substreams can improve reliability by adapting the transmitted signal to the instantaneous channel eigen-structure, i.e., by establishing different parallel channels through the strongest channel eigenmodes. In this chapter we investigate analytically the average and outage performance of a general spatial multiplexing MIMO system with CSI. With the final aim of establishing the performance limits of this particular transmission structure, the average bit error rate and outage probability versus signal-to-noise ratio curves of spatial multiplexing MIMO systems are characterized in terms of two key parameters: the array gain and the diversity gain. Finally, these results are applied to analyze the performance of practical linear MIMO transceiver designs.

3.1 Introduction

The gains obtained by the deployment of multiple antennas at both sides of the link are the array gain, the diversity gain, and the multiplexing gain (see Section 1.1.1 for details). The array gain is the improvement in signal-to-noise ratio (SNR) obtained by coherently combining the signals on multiple transmit or multiple receive dimensions while the diversity gain is the improvement in link reliability obtained by receiving replicas of the information signal through independently fading dimensions. These gains are not exclusive of multiple-input multiple-output (MIMO) channels and also exist in single-input multiple-output (SIMO) and multiple-input single-output (MISO) channels. In contrast, the multiplexing gain, which refers to the increase of rate at no additional power consumption, is a unique characteristic of MIMO channels. The basic idea is to exploit the multiple dimensions to open up several parallel subchannels within the MIMO channel, also termed channel eigenmodes. This allows the transmission of several symbols simultaneously or, in other words, the establishment of several substreams for communication.

In this chapter we focus on spatial multiplexing MIMO systems with perfect channel state information (CSI) at both sides of the link. Spatial multiplexing is a simple MIMO transmit technique that allows a high spectral efficiency by dividing the incoming data into multiple independent substreams which are simultaneously transmitted. When the transmitter has no CSI, each substream is transmitted on a different antenna resulting in the well-known V-BLAST scheme [Pau94,Fos99]. To the contrary, when perfect CSI is available at the transmitter, channel-dependent linear precoding of the data substreams can further improve performance by adapting the transmitted signal to the instantaneous channel eigen-structure. Different degrees of adaptation have been considered in the literature, namely:

(i) Adapt only the linear precoder and/or power allocation among the different substreams, keeping the number of substreams and the constellations fixed. This is by far the most widely considered scenario, e.g., [Lee76,Sal85,Yan94b,Yan94a,Sca99,Sam01, Sca02b,Ong03,Pal03,Din03a,Pal07].

(ii) Adapt the precoder/power allocation among the different substreams, the number of substreams, and choice of constellations, keeping the data rate fixed. This has been only partially considered in a few papers. For instance, the adaptation of the precoder and number of substreams is addressed in [Lov05b], and the precoder, constellations,

and number of substreams are designed in [Pal05a] to minimize the transmitted power under a bit error rate (BER) constraint.

This chapter concentrates on case (i) which embraces the schemes that have received more attention in the literature, while case (ii) is partially addressed in Chapter 4. More exactly, in this chapter we analyze analytically the uncoded performance of a general spatial multiplexing MIMO system with CSI, when the number of substreams and constellations have been chosen beforehand.

In order to simplify the study of MIMO systems, it is customary to divide them into an uncoded part, which transmits symbols drawn from some constellations, and a coded part that builds upon the uncoded system. Although the ultimate system performance depends on the combination of both parts, it is convenient to consider the uncoded and coded parts independently to simplify the analysis and design. In this chapter we analyze analytically the uncoded performance of a general spatial multiplexing MIMO system. The final objective is to provide the fundamental performance limits encountered when the establishing several parallel data streams through the channel eigenmodes, independently of how the available power is distributed among them.

First we study the exact individual performance of the substreams transmitted through the channel eigenmodes under common MIMO channel models: uncorrelated Rayleigh, semicorrelated Rayleigh, and uncorrelated Ricean MIMO fading channels. The resulting expressions turn out to be quite complicated and, consequently, we focus afterwards on the high-SNR performance in order to provide more insight into the system behaviour. In particular, we study the average and outage performance of each channel eigenmode at high SNR following the methodology introduced by Wang and Giannakis in [Wan03]. This allows us to characterize the curves corresponding to average BER versus SNR and outage probability versus SNR in terms of the diversity gain, which determines the slope of the curve at high SNR in a log-log scale, and the array gain, which determines the horizontal shift of the curve. We also extend this characterization to evaluate the global performance that takes into account all the established substreams for a fixed number of substreams according to scenario (i).

These general results are then applied to analyze the performance of linear MIMO transceivers which are low-complexity practical schemes composed of a linear precoder at the

transmitter and a linear equalizer at the receiver. The design of linear transceivers when the number of substream is fixed beforehand has been studied since the 1970s under different measures of performance based on the SNR, the mean square error (MSE), or the BER [Lee76, Sal85, Yan94b, Sca99, Sam01, Sca02b, Ong03, Pal03, Din03b, Din03a]. A general unifying framework that embraces most of these design criteria was proposed in [Pal03] (see an up-to-date overview in [Pal07]). Actually, based on the formulation in [Pal03], we are able to investigate the high-SNR global average performance of a wide family of practical linear MIMO transceivers.

The rest of the chapter is organized as follows. Section 3.2 is devoted to introducing the performance metrics of a general digital communication system in fading channels and presenting how these performance measures can be approximated in the high-SNR regime. The adopted fading channel models are described in Section 3.3 and the signal model corresponding to a general spatial multiplexing scheme with CSI in Section 3.4. Next, the the average BER and outage performance of spatial multiplexing MIMO systems is analytically investigated in Sections 3.5–3.7. Then, in Section 3.8, we apply these results to analyze the performance of linear MIMO transceivers. Finally, Section 3.9 summarizes the main results and provides the list of publications where they have been presented.

3.2 Performance Metrics of Digital Communication Systems

The ultimate performance of a digital communication system is given in terms of the symbol error probability or symbol error rate (SER), defined as the fraction of symbols in error, or in terms of the bit error probability or BER, defined as the fraction of bits in error. When communicating over fading channels, instantaneous performance measures[1] are random quantities and do not provide any insight on the behavior of the system. All possible realizations of the fading channel have to be taken into account, leading to the concepts of average and outage performance metrics. The average SNR and the average SER/BER measures the SNR and SER/BER averaged over the different channel states, whereas the outage probability is the probability that the system SNR performance is below a given acceptable threshold.

[1] The notion of instantaneousness is with respect to the channel, i.e., instantaneous performance denotes the performance for a given channel realization but averaging over the transmitted signal and noise.

Modulation	α	$\beta(i)$	β	γ
M-PAM	$2\big(1-\frac{1}{M}\big)$	$\frac{6}{M^2-1}$	$\frac{6}{M^2-1}$	1
M-QAM	$4\big(1-\frac{1}{\sqrt{M}}\big)$	$(2i-1)^2\big(\frac{3}{M-1}\big)$	$\frac{3}{M-1}$	$\sqrt{M}/2$
M-PSK	2	$2\sin^2\big(\frac{(2i-1)\pi}{M}\big)$	$2\sin^2\big(\frac{\pi}{M}\big)$	$\max(M/4,1)$

Table 3.1 Constellation parameters for M-PAM, M-QAM, and M-PSK modulations.

In this section, we provide the expressions of these performance measures assuming a flat-fading channel and additive white Gaussian noise (AWGN) as needed in subsequent subsections. For an exhaustive and comprehensive treatment of the subject see [Sim02b] or textbooks in digital communications [Sim95, Stü96, Ben99, Pro01, Tse05].

3.2.1 Instantaneous Performance

In the presence of additive white Gaussian noise, the instantaneous SER of a digital communication system (under coherent detection and for many different modulation formats) can be analytically approximated as[2] [Lu99, Sec. III] [Sim02b, Sec. 8.1.1]

$$\mathsf{SER}(\rho) \simeq \alpha \sum_{i=1}^{\gamma} \mathcal{Q}\left(\sqrt{\beta(i)\rho}\right) \tag{3.1}$$

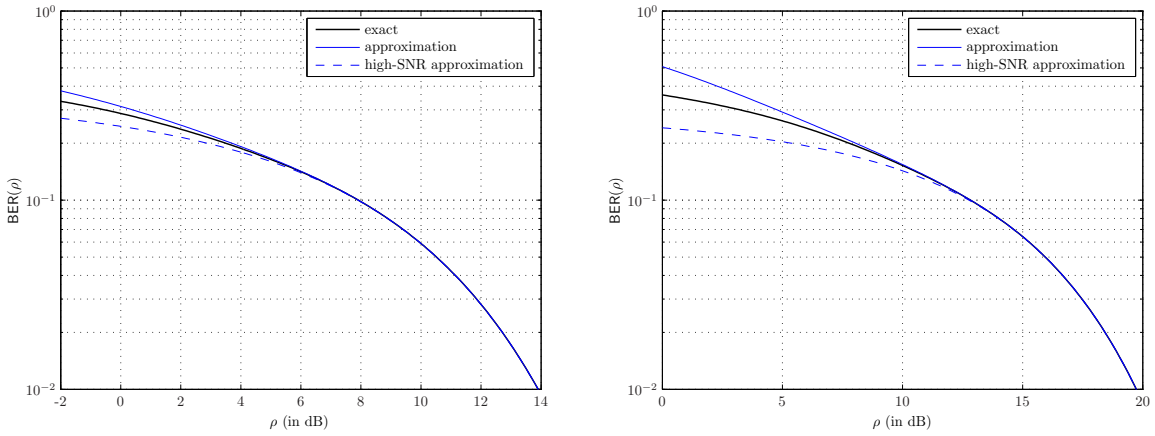
where ρ denotes the instantaneous SNR, $\mathcal{Q}(\cdot)$ is the Gaussian $\mathcal{Q}$-function defined as [Sim02b, eq. (4.1)]

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt \tag{3.2}$$

and the parameters $\alpha \triangleq \alpha(M)$, $\beta(i) \triangleq \beta(i,M)$, and $\gamma \triangleq \gamma(M)$ depend on the M-ary modulation used to map the source bits to symbols. When this modulation process includes a Gray code mapping, which has the property that in transitioning from one symbol to an adjacent symbol of the constellation only one out of the source $\log_2 M$ bits changes [Gra53] [Ben99, Sec. 5.1.3], the instantaneous BER can be approximately obtained from the instantaneous SER as [Sim02b, eq. (8.7)]

$$\mathsf{BER}(\rho) \simeq \frac{\mathsf{SER}(\rho)}{\log_2 M} \tag{3.3}$$

[2] See the exact expressions for QAM constellations in [Cho02] and for PSK constellations in [Las03].



(i) 16-QAM (ii) 64-QAM

Figure 3.1 Exact instantaneous BER, approximated instantaneous BER in (3.4), and high-SNR approximated instantaneous BER in (3.5).

since for relatively high SNR the only significant symbol errors occur when deciding in favor of adjacent symbols. Combining (3.1) and (3.3), we can express the instantaneous BER as

$$\mathsf{BER}(\rho) \simeq \frac{\alpha}{\log_2 M} \sum_{i=1}^{\gamma} \mathcal{Q}\left(\sqrt{\beta(i)\rho}\right). \tag{3.4}$$

The BER expression in (3.4) is shown in [Lu99] to be valid for all M and quite accurate at both low and high SNR. When focusing only on large SNR, one can further approximate (3.4) by the first term in the summation since $\beta(i)$ is strictly increasing in i and, thus, dominates the instantaneous BER:[3]

$$\mathsf{BER}(\rho) \simeq \frac{\alpha}{\log_2 M} \mathcal{Q}\left(\sqrt{\beta\rho}\right) \tag{3.5}$$

where we have defined $\beta \triangleq \beta(1)$.

The general BER expressions in (3.4) and (3.5) are particularized in Table 3.1 for the most common digital modulation formats (see [Sim02b, Sec. 8.1] for other modulations or non-coherent detection strategies). In Figure 3.1 we plot the exact instantaneous BER performance as a function of the instantaneous SNR together with the corresponding approximations in (3.4) and (3.5) when 16-QAM and 64-QAM constellations are used. As expected, the approximated BER curves predict very accurately the exact performance in the BER region of interest ($\mathsf{BER}(\rho) >> 10^{-1}$).

[3] Owing to its accurateness in the BER work-region of practical communication systems, the expression in (3.5) is henceforth regarded as the exact instantaneous BER.

3.2.2 Average Performance

Under the flat-fading assumption, the effect of the channel on the transmitted signal is in the form of a random multiplicative distortion [Big98, Sec. II.B]. According to this, the instantaneous SNR ρ is given by the product of a channel-dependent parameter μ and a deterministic positive quantity $\bar{\rho}$, i.e.,

$$\rho = \mu\bar{\rho} \tag{3.6}$$

where $\bar{\rho}$ is the average SNR at the receiver whenever $\mathbb{E}\{\mu\} = 1$. As we are interested in the average BER incurred by the system, we need to take the expectation over all possible channel states:

$$\overline{\mathsf{BER}}(\bar{\rho}) \triangleq \mathbb{E}_{\mu}\{\mathsf{BER}(\rho)\} = \int_0^{\infty} \mathsf{BER}\left(\bar{\rho}\mu\right) f_{\mu}(\mu) d\mu \tag{3.7}$$

where $f_{\mu}(\mu)$ is the pdf of the channel-dependent parameter μ.

The average BER in (3.7) has been analytically evaluated only in certain cases, such as when the channel is Rayleigh or Ricean, i.e., the channel parameter μ follows a χ^2-distribution (see [Sim02b, Sec. 8.2] for some available average BER closed-form expressions). A unifying procedure for finding the error of digital communication systems over general fading channels is presented in [Sim98] (see textbook [Sim02b] for further details). This method is based on an alternative representation of the Gaussian $\mathcal{Q}$-function and requires evaluating the moment generating function (mgf) of the instantaneous SNR.

In some cases, when even an exact statistical characterization of the channel parameter or, equivalently, of the SNR is not possible, it is convenient to allow a certain degree of approximation. The most common approach is to shift the focus from exact performance to high-SNR performance as done in Section 3.2.4. This high-SNR approximation is also meaningful when the exact performance evaluation involves complicate expressions and, thus, the potentially simpler high-SNR performance characterization can easier provide meaningful insight into the system behavior.

3.2.3 Outage Performance

The average performance metrics presented in Section 3.2.2 are useful performance measures when the transmission interval is long enough to reveal the long-term ergodic properties of the

fading channel. The ergodicity assumption, however, is not necessarily satisfied in practical communication systems operating over fading channels, because no significant channel variability may occur during the whole transmission. In these circumstances, the most convenient measure to capture the performance of the system is the outage BER, i.e., the minimum BER value guaranteed with a small given probability. However, the outage BER is difficult to calculate and the performance of communication systems over non-ergodic fading channels is instead commonly measured with the outage probability, defined as the probability that the instantaneous SNR ρ falls below a certain threshold ρ_{th} [Sim02b, eq. (1.4)]:

$$\mathsf{P}_{\mathsf{out}}(\bar{\rho}) \triangleq \mathsf{Pr}(\rho \leq \rho_{\text{th}}) = \int_0^{\rho_{\text{th}}/\bar{\rho}} f_\mu(\mu) d\mu. \tag{3.8}$$

Observe that, using the approximate instantaneous BER expression in (3.5), the SNR threshold ρ_{th} can be chosen as

$$\rho_{\text{th}} = \frac{1}{\beta} \mathcal{Q}^{-2}\Big(\frac{\log_2 M}{\alpha} \mathsf{BER}_{\text{th}}\Big) \tag{3.9}$$

where $\mathcal{Q}^{-1}(\cdot)$ is the inverse of the Gaussian $\mathcal{Q}$-function, and the outage probability can be then understood as the probability that the instantaneous BER fall below a target BER denoted by BER_{th}.

The outage probability in (3.8) is closely related to the information-theoretical outage probability defined as the probability that the instantaneous mutual information of the channel is below the transmitted code rate [Oza94, Cai99]. For instance, in a Gaussian channel the information-theoretical outage probability $\mathsf{P}_{\mathsf{out}}(\mathsf{R}) \triangleq \mathsf{Pr}(\log_2(1+\rho) \leq \mathsf{R})$ can be obtained from the outage probability in (3.8) by choosing the SNR threshold as $\rho_{\text{th}} = 2^{\mathsf{R}} - 1$.

3.2.4 High-SNR Performance

In the high-SNR regime, the average BER function when communicating over flat-fading channels can be approximated in most cases (clearly, a necessary condition is that $\overline{\mathsf{BER}}(\bar{\rho}) \to 0$ as $\rho \to \infty$) as[4]

$$\overline{\mathsf{BER}}(\bar{\rho}) = (G_{\mathsf{c}} \cdot \bar{\rho})^{-G_{\mathsf{d}}} + o\big(\bar{\rho}^{-G_{\mathsf{d}}}\big) \tag{3.10}$$

[4] We say that $f(x) = o(g(x))$, $g(x) > 0$, if $f(x)/g(x) \to 0$ as $x \to 0$ [Bru81, eq. (1.3.1)].

where G_{d} and G_{c} are referred to as the diversity and coding gains,[5] respectively, and $\bar{\rho}$ is the average SNR. The diversity gain determines the slope of the BER versus $\bar{\rho}$ curve at high SNR in a log-log scale and the coding gain determines the shift of the curve with respect to the benchmark BER curve $\bar{\rho}^{-G_{\mathsf{d}}}$. This leads to a simple parameterized average BER characterization in the high-SNR regime that can provide meaningful insights related to the system behavior.

As shown in [Wan03], the high-SNR average BER and the outage probability of a communication system with the instantaneous SNR given in (3.6) depend only on the behavior of the pdf of the channel dependent parameter μ near the origin ($\mu \to 0^+$). That is, given the first order Taylor expansion of the pdf, $f_\mu(\mu) = a\mu^d + o(\mu^d)$, parameterized expressions in terms of the diversity and coding gains for the average BER and outage probability can be straightforwardly obtained using [Wan03, Prop. 1] and [Wan03, Prop. 5], respectively. Due to their importance in this chapter, we reproduce these results in the following lemmas.

Lemma 3.1 ([Wan03, Prop. 1])**.** *The average BER of a communication system under AWGN and with instantaneous SNR given by $\rho = \bar{\rho}\mu$, where the pdf of the channel-dependent parameter μ can be written as $f_\mu(\mu) = a\mu^d + o(\mu^d)$, satisfies*

$$\overline{\mathsf{BER}}(\bar{\rho}) = (G_{\mathsf{c}} \cdot \bar{\rho})^{-G_{\mathsf{d}}} + o\left(\bar{\rho}^{-G_{\mathsf{d}}}\right) \tag{3.11}$$

where the diversity gain G_{d} and the coding gain G_{c} are given by

$$G_{\mathsf{d}} = d + 1 \tag{3.12}$$

$$G_{\mathsf{c}} = \beta \left(\frac{\alpha}{\log_2 M} \frac{a 2^d \Gamma(d + 3/2)}{\sqrt{\pi}(d+1)} \right)^{-1/(d+1)} \tag{3.13}$$

and $\Gamma(\cdot)$ denotes the gamma function (see Definition 2.5).

Lemma 3.2 ([Wan03, Prop. 5])**.** *The outage probability of a communication system under AWGN and with instantaneous SNR given by $\rho = \bar{\rho}\mu$, where the pdf of the channel-dependent parameter μ can be written as $f_\mu(\mu) = a\mu^d + o(\mu^d)$, satisfies*

$$\mathsf{P}_{\mathsf{out}}(\bar{\rho}) = (O_{\mathsf{c}} \cdot \bar{\rho})^{-G_{\mathsf{d}}} + o\left(\bar{\rho}^{-G_{\mathsf{d}}}\right) \tag{3.14}$$

[5] The coding gain is known as array gain in the context of multiantenna systems [And00a] and, hence, array gain is the nomenclature adopted in the next sections.

where the outage diversity gain G_{d} *and the outage coding gain* O_{c} *are given by*

$$G_{\mathsf{d}} = d + 1 \tag{3.15}$$

$$O_{\mathsf{c}} = \frac{1}{\rho_{\text{th}}} \left(\frac{a}{d+1} \right)^{-1/(d+1)} . \tag{3.16}$$

Lemma 3.1 and 3.2 offer a simple and unifying approach to evaluate the average and outage performance of communication systems over random fading channels and allow the interpretation of the effect of the system parameters in the performance. For convenience, we extend these results for the case in which the instantaneous SNR is given by

$$\rho = \bar{\rho}\mu + \phi \tag{3.17}$$

where ϕ is a fixed deterministic parameter. Although the instantaneous SNR in (3.17) does hardly correspond to a realistic setup, it is useful in the next sections to bound the performance of some practical linear transceiver MIMO systems. The following corollary shows the generalization of Lemma 3.1, whereas Lemma 3.2 can be generalized in the same way.

Corollary 3.1. *The average BER of a communication system under AWGN and with instantaneous SNR given by* $\rho = \bar{\rho}\mu + \phi$*, where the pdf of the channel-dependent parameter* μ *can be written as* $f_\mu(\mu) = a\mu^d + o(\mu^d)$*, is*

$$\overline{\text{BER}}(\bar{\rho}) = (G_{\mathsf{c}} \cdot \bar{\rho})^{-G_{\mathsf{d}}} + o\left(\bar{\rho}^{-G_{\mathsf{d}}}\right) \tag{3.18}$$

where the diversity gain G_{d} *and the coding gain* G_{c} *are given by*

$$G_{\mathsf{d}} = d + 1 \tag{3.19}$$

$$G_{\mathsf{c}} = \beta \left(\frac{\alpha}{\log_2 M} \frac{aI(d, \beta\phi)}{\sqrt{2\pi}(d+1)} \right)^{-1/(d+1)} \tag{3.20}$$

and $I(d, \beta\phi)$ *is defined as*[6]

$$I(d, \beta\phi) = \int_{\sqrt{\beta\phi}}^{\infty} e^{-\frac{x^2}{2}} \left(x^2 - \beta\phi\right)^{(d+1)} dx. \tag{3.21}$$

Proof. See Appendix 3.A.

[6] A closed-form expression for this integral does not exist for a general value of the parameter d; however, it can be easily evaluated for the most common values of d (integers).

3.3 MIMO Channel Model

Consider a communication link with n_{T} transmit and n_{R} receive dimensions, the resulting MIMO channel can be mathematically described by an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ channel matrix $\mathbf{H}$, whose (i,j)th entry characterizes the path between the jth transmit and the ith receive antenna. When communicating over MIMO fading channels, $\mathbf{H}$ is a random matrix that depends on the particular system architecture and the particular propagation conditions. Hence, $\mathbf{H}$ is considered to be drawn from a certain probability distribution, which characterizes the system and scenario of interest and is known as channel model. The system behavior is then evaluated on the average or outage sense as described in Section 3.2.

In this section we introduce the Rayleigh and Ricean flat-fading[7] MIMO channel models assumed in the analytical derivations and performance analysis of spatial multiplexing MIMO systems. In addition, we provide the required probabilistic characterization of its ordered eigenvalues.

3.3.1 Rayleigh and Ricean Fading MIMO Channels

In MIMO wireless communications, the large number of scatters in the channel that contribute to the signal at the receiver results in zero-mean Gaussian distributed channel matrix coefficients. Analogously to the single antenna case, this model is referred to as Rayleigh MIMO fading channel [Fos98].

In realistic environments the SISO channels connecting each pair of transmit and receive antenna elements are not independent due, for instance, to insufficient spacing between antenna elements or insufficient scattering. In such cases, a convenient approach is to construct a correlation model that can provide a reasonable description of the propagation environment and physical setup for the wireless application of interest (see [Yu02] for a review on MIMO channel models). The most common correlation model assumes that antenna correlation at the transmitter side and at the receiver side are caused by independent phenomena and is known as Kronecker model (see, e.g., [Chi00, Shi00, Chu02, Böl00, Böl03]). Thus, correlation can be separated and the

[7] Observe that in wideband MIMO systems a multicarrier approach is usually applied and the flat-fading assumption holds then for the channel seen by each subcarrier (see e.g. [Kon96]). Henceforth we use the term 'fading' instead of 'flat-fading', although a flat-fading channel is implicitly considered.

correlated MIMO Rayleigh channel can be modeled as

$$\mathbf{H}^{(\text{Rayleigh})} = \mathbf{\Sigma}_{\mathsf{R}}^{1/2}\mathbf{H}_{\mathsf{w}}\mathbf{\Sigma}_{\mathsf{T}}^{1/2} \tag{3.22}$$

where $\mathbf{\Sigma}_{\mathsf{T}} = \big(\mathbf{\Sigma}_{\mathsf{T}}^{1/2}\big)\big(\mathbf{\Sigma}_{\mathsf{T}}^{1/2}\big)^{\dagger}$ is the transmit correlation matrix, $\mathbf{\Sigma}_{\mathsf{R}} = \big(\mathbf{\Sigma}_{\mathsf{R}}^{1/2}\big)\big(\mathbf{\Sigma}_{\mathsf{R}}^{1/2}\big)^{\dagger}$ is the receive correlation matrix, and $\mathbf{H}_{\mathsf{w}}$ is the random channel matrix with i.i.d. zero-mean unit-variance circularly symmetric Gaussian entries, i.e., $[\mathbf{H}_{\mathsf{w}}]_{i,j} \sim \mathcal{CN}(0,1)$. Although this simple correlation model is not completely general (see, e.g., [Abd02, Ozc03] for environments where it does not apply), it has been validated experimentally in [Yu01, Ker02, McN02] as well as using the ray-tracing simulations in [Chu02]. Hence, it is widely accepted as an accurate representation of the fade correlation seen in actual cellular systems.

In addition, for scenarios where a line-of-sight or specular component is present, the channel matrix is modeled as having nonzero mean [Böl00, Böl03]:

$$\mathbf{H}^{(\text{Ricean})} = \sqrt{\frac{K_{\mathsf{c}}}{K_{\mathsf{c}}+1}}\overline{\mathbf{H}} + \sqrt{\frac{1}{K_{\mathsf{c}}+1}}\mathbf{\Sigma}_{\mathsf{R}}^{1/2}\mathbf{H}_{\mathsf{w}}\mathbf{\Sigma}_{\mathsf{T}}^{1/2} \tag{3.23}$$

where $K_{\mathsf{c}} \in [0,\infty)$ is power normalization factor known as the Ricean K_{c}-factor and $\overline{\mathbf{H}}$ is a deterministic $n_{\mathsf{R}} \times n_{\mathsf{T}}$ matrix containing the line-of-sight components of the channel. Analogously to the single antenna case, this model is referred to as MIMO Ricean fading channel [Dri99].

Observe that the MIMO Ricean fading model in (3.23) includes channels ranging from a fully random Rayleigh channel when $K_{\mathsf{c}} = 0$ to a fully deterministic channel when $K_{\mathsf{c}} \to \infty$. For a fair comparison of the different cases, the total average received power is assumed to be constant, i.e.,

$$\mathbb{E}\left\{\|\mathbf{H}\|^2\right\} = \frac{K_{\mathsf{c}}}{K_{\mathsf{c}}+1}\text{tr}\big(\overline{\mathbf{H}}\overline{\mathbf{H}}^{\dagger}\big) + \frac{1}{K_{\mathsf{c}}+1}\mathbb{E}\Big\{\text{tr}\big(\mathbf{\Sigma}_{\mathsf{R}}\mathbf{H}_{\mathsf{w}}\mathbf{\Sigma}_{\mathsf{T}}\mathbf{H}_{\mathsf{w}}^{\dagger}\big)\Big\} \tag{3.24}$$

$$= \frac{K_{\mathsf{c}}}{K_{\mathsf{c}}+1}\text{tr}\big(\overline{\mathbf{H}}\overline{\mathbf{H}}^{\dagger}\big) + \frac{1}{K_{\mathsf{c}}+1}\text{tr}\left(\mathbf{\Sigma}_{\mathsf{R}}\right)\text{tr}\left(\mathbf{\Sigma}_{\mathsf{T}}\right) = n_{\mathsf{R}}n_{\mathsf{T}} \tag{3.25}$$

and, hence, we can impose without loss of generality that

$$\text{tr}\left(\mathbf{\Sigma}_{\mathsf{T}}\right) = n_{\mathsf{T}}, \quad \text{tr}\left(\mathbf{\Sigma}_{\mathsf{R}}\right) = n_{\mathsf{R}}, \quad \text{and} \quad \text{tr}\big(\overline{\mathbf{H}}\overline{\mathbf{H}}^{\dagger}\big) = n_{\mathsf{R}}n_{\mathsf{T}}. \tag{3.26}$$

3.3.2 Particular Cases of Rayleigh and Ricean Fading MIMO Channels

The general channel model introduced in the previous section can be particularized to obtain some important special cases of the Rayleigh and Ricean MIMO channels (see physical justifications in [Ivr03]). In the following we present the adopted channel models as well as the

distributions of $\mathbf{H}\mathbf{H}^\dagger$ or $\mathbf{H}^\dagger\mathbf{H}$ induced by each particular model, which are required in the subsequent sections when analyzing the performance of spatial multiplexing MIMO systems with CSI.

Definition 3.1 (Uncorrelated Rayleigh MIMO channel)**.** *The uncorrelated Rayleigh MIMO fading channel model is defined as*

$$\mathbf{H} = \mathbf{H}_{\mathsf{w}} \tag{3.27}$$

where $\mathbf{H}_{\mathsf{w}}$ *is an* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *random channel matrix with i.i.d. zero-mean unit-variance complex Gaussian entries.*

Consider an uncorrelated Rayleigh fading MIMO channel $\mathbf{H}$ as given in Definition 3.1, then the random Hermitian matrix $\mathbf{W}$ $(n \times n)$ defined as

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger & n_{\mathsf{R}} \leq n_{\mathsf{T}} \\ \mathbf{H}^\dagger\mathbf{H} & n_{\mathsf{R}} > n_{\mathsf{T}} \end{cases} \tag{3.28}$$

follows a complex uncorrelated central Wishart distribution (see case (i) of Definition 2.21), denoted as $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{I}_n)$, where $n = \min(n_{\mathsf{T}}, n_{\mathsf{R}})$ and $m = \max(n_{\mathsf{T}}, n_{\mathsf{R}})$. Since the nonzero eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ and $\mathbf{H}^\dagger\mathbf{H}$ coincide, the statistical properties of the ordered channel eigenvalues can be analyzed without loss of generality by focusing on the ordered eigenvalues of $\mathbf{W}$.

Definition 3.2 (Min-semicorrelated Rayleigh MIMO channel)**.** *The semicorrelated Rayleigh fading MIMO channel model with correlation at the side with minimum number of antennas is defined as*

$$\mathbf{H} = \begin{cases} \mathbf{\Sigma}^{1/2}\mathbf{H}_{\mathsf{w}} & n_{\mathsf{R}} \leq n_{\mathsf{T}} \\ \mathbf{H}_{\mathsf{w}}\mathbf{\Sigma}^{1/2} & n_{\mathsf{R}} > n_{\mathsf{T}} \end{cases} \tag{3.29}$$

where $\mathbf{\Sigma} = \left(\mathbf{\Sigma}^{1/2}\right)\left(\mathbf{\Sigma}^{1/2}\right)^\dagger$ *is the* $n \times n$ *positive definite correlation matrix with* $n = \min(n_{\mathsf{T}}, n_{\mathsf{R}})$ *and* $\mathbf{H}_{\mathsf{w}}$ *is an* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *random channel matrix with i.i.d. zero-mean unit-variance complex Gaussian entries.*

Consider a min-semicorrelated Rayleigh fading MIMO channel $\mathbf{H}$ as given in Definition 3.2, then the random Hermitian matrix $\mathbf{W}$ $(n \times n)$ in (3.28) follows a complex correlated central

Wishart distribution (see case (ii) of Definition 2.21), denoted as $\mathbf{W} \sim \mathcal{W}_n(m, \mathbf{0}_n, \mathbf{\Sigma})$, where $n = \min(n_\mathsf{T}, n_\mathsf{R})$, $m = \max(n_\mathsf{T}, n_\mathsf{R})$, and $\mathbf{\Sigma}$ is the $n \times n$ positive definite correlation matrix.

Definition 3.3 (Max-semicorrelated Rayleigh MIMO channel)**.** *The semicorrelated Rayleigh fading MIMO channel model with correlation at the side with maximum number of antennas is defined as*

$$\mathbf{H} = \begin{cases} \mathbf{H}_\mathsf{w}\mathbf{\Sigma}^{1/2} & n_\mathsf{R} \leq n_\mathsf{T} \\ \mathbf{\Sigma}^{1/2}\mathbf{H}_\mathsf{w} & n_\mathsf{R} > n_\mathsf{T} \end{cases} \tag{3.30}$$

where $\mathbf{\Sigma} = \left(\mathbf{\Sigma}^{1/2}\right)\left(\mathbf{\Sigma}^{1/2}\right)^\dagger$ *is the* $m \times m$ *positive definite correlation matrix with* $m = \max(n_\mathsf{T}, n_\mathsf{R})$ *and* $\mathbf{H}_\mathsf{w}$ *is an* $n_\mathsf{R} \times n_\mathsf{T}$ *random channel matrix with i.i.d. zero-mean unit-variance complex Gaussian entries.*

Consider a max-semicorrelated Rayleigh fading MIMO channel $\mathbf{H}$ as given in Definition 3.3, then the random Hermitian matrix $\mathbf{W}$ $(m \times m)$ defined as

$$\mathbf{W} = \begin{cases} \mathbf{H}^\dagger\mathbf{H} & n_\mathsf{R} \leq n_\mathsf{T} \\ \mathbf{H}\mathbf{H}^\dagger & n_\mathsf{R} > n_\mathsf{T} \end{cases} \tag{3.31}$$

follows a complex correlated central Pseudo-Wishart distribution (see Definition 2.22) denoted as $\mathbf{W} \sim \mathcal{PW}_m(n, \mathbf{0}_m, \mathbf{\Sigma})$, where $n = \min(n_\mathsf{T}, n_\mathsf{R})$, $m = \max(n_\mathsf{T}, n_\mathsf{R})$, and $\mathbf{\Sigma}$ is the $m \times m$ positive definite correlation matrix.

Definition 3.4 (Uncorrelated Ricean MIMO channel)**.** *The uncorrelated Ricean fading MIMO channel model is defined as*

$$\mathbf{H} = \sqrt{\frac{K_\mathsf{c}}{K_\mathsf{c}+1}}\overline{\mathbf{H}} + \sqrt{\frac{1}{K_\mathsf{c}+1}}\mathbf{H}_\mathsf{w} \tag{3.32}$$

where $K_\mathsf{c} \in (0, \infty)$*,* $\overline{\mathbf{H}}$ *is an* $n_\mathsf{R} \times n_\mathsf{T}$ *deterministic matrix, and* $\mathbf{H}_\mathsf{w}$ *is an* $n_\mathsf{R} \times n_\mathsf{T}$ *random channel matrix with i.i.d. zero-mean unit-variance complex Gaussian entries.*

Consider an uncorrelated Ricean fading MIMO channel $\mathbf{H}$ as given in Definition 3.4, then the random Hermitian matrix $\widetilde{\mathbf{W}} = (K_\mathsf{c}+1)\mathbf{W}$ $(n \times n)$, where $\mathbf{W}$ is given in (3.28), follows a complex uncorrelated noncentral Wishart distribution (see case (iii) of Definition 2.21), denoted as $\widetilde{\mathbf{W}} \sim \mathcal{W}_n(m, \mathbf{\Omega}, \mathbf{I}_n)$, where $n = \min(n_\mathsf{T}, n_\mathsf{R})$ and $m = \max(n_\mathsf{T}, n_\mathsf{R})$, and the noncentrality

parameter $\mathbf{\Omega}$ is defined as

$$\mathbf{\Omega} = \begin{cases} K_{\mathsf{c}}\overline{\mathbf{H}\mathbf{H}}^{\dagger} & n_{\mathsf{R}} \leq n_{\mathsf{T}} \\ K_{\mathsf{c}}\overline{\mathbf{H}}^{\dagger}\overline{\mathbf{H}} & n_{\mathsf{R}} > n_{\mathsf{T}} \end{cases}. \tag{3.33}$$

Note that in this case the nonzero channel eigenvalues, i.e., the eigenvalues of $\mathbf{W}$, are a scaled version of the eigenvalues of the complex uncorrelated central Wishart distributed matrix $\widetilde{\mathbf{W}}$.

3.3.3 Ordered Channel Eigenvalues

The performance of MIMO systems is strongly related to the eigenstructure of the channel matrix $\mathbf{H}$, or more exactly, to the nonzero eigenvalues of $\mathbf{H}^{\dagger}\mathbf{H}$ (or $\mathbf{H}\mathbf{H}^{\dagger}$). Focusing on the spatial multiplexing MIMO systems with CSI investigated in this chapter, the instantaneous SNR of the substream transmitted through the kth strongest channel eigenmode depends (at least) on the kth largest channel eigenvalue (see Section 3.4). Hence, in order to obtain closed-form expressions for the performance measures introduced in Section 3.2 under the channel models in Definitions 3.1–3.4, the probabilistic characterization of the ordered eigenvalues of the corresponding random matrices is required. In Chapter 2 we introduced a unified probabilistic characterization of the ordered eigenvalues of a general class of Hermitian random matrices which includes Wishart and Pseudo-Wishart random matrices as particular cases. The results derived therein are, thus, intensively applied in the performance analysis of spatial multiplexing MIMO systems presented in this chapter.

3.4 Spatial Multiplexing MIMO Systems with CSI

Recall from the introduction that one of the salient and unique characteristics of MIMO channels is the multiplexing gain, which refers to the increase of rate at no additional power consumption. The multiple dimensions of the MIMO channel are exploited to open up several parallel subchannels which allow the transmission of several symbols simultaneously. When perfect CSI is available at the transmitter, the data signal is adapted to the instantaneous channel eigenstructure by transmitting the established substreams through the strongest channel eigenmodes. In this section we present a general spatial multiplexing MIMO system with perfect CSI at both sides of the link. This scheme is analyzed in Sections 3.5–3.7 and afterwards, in Section 3.8, it is shown to include most interesting MIMO linear transceiver designs as particular cases.

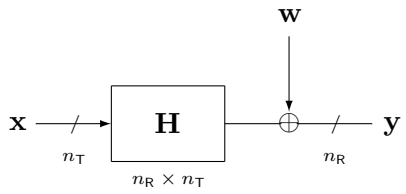


Figure 3.2 General MIMO system model.

3.4.1 System Model

The signal model corresponding to a transmission through a general MIMO channel with n_T transmit and n_R receive dimensions is (see Figure 3.2)

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{3.34}$$

where $\mathbf{x} \in \mathbb{C}^{n_\mathsf{T}}$ is the transmitted vector, $\mathbf{H} \in \mathbb{C}^{n_\mathsf{R} \times n_\mathsf{T}}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^{n_\mathsf{R}}$ is the received vector, and $\mathbf{w} \in \mathbb{C}^{n_\mathsf{R}}$ is a spatially white zero-mean circularly symmetric complex Gaussian noise vector normalized so that $\mathbb{E}\{\mathbf{w}\mathbf{w}^\dagger\} = \mathbf{I}_{n_\mathsf{R}}$. The channel matrix $\mathbf{H}$ contains the complex path gains $[\mathbf{H}]_{i,j}$ between every transmit and receive antenna pair.

Following the singular value decomposition (SVD), the channel matrix $\mathbf{H}$ can be written as

$$\mathbf{H} = \mathbf{U}\sqrt{\mathbf{\Lambda}}\mathbf{V}^\dagger \tag{3.35}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\sqrt{\mathbf{\Lambda}}$ is a diagonal matrix[8] containing the singular values of $\mathbf{H}$ sorted in descending order. This way, the channel matrix is effectively decomposed into rank $(\mathbf{H}) = \min\{n_\mathsf{T}, n_\mathsf{R}\}$ independent orthogonal modes of excitation, which are referred to as channel eigenmodes [Ral98, And00b, Böl02b].

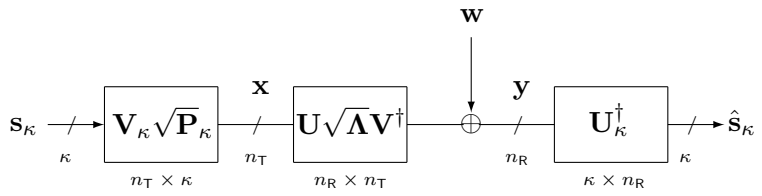
Assuming that the channel is perfectly known at the transmitter and that $\kappa \leq \min\{n_\mathsf{T}, n_\mathsf{R}\}$ data symbols per channel use have to be communicated, the transmitted vector can be expressed as

$$\mathbf{x} = \mathbf{V}_\kappa\sqrt{\mathbf{P}}_\kappa\mathbf{s}_\kappa \tag{3.36}$$

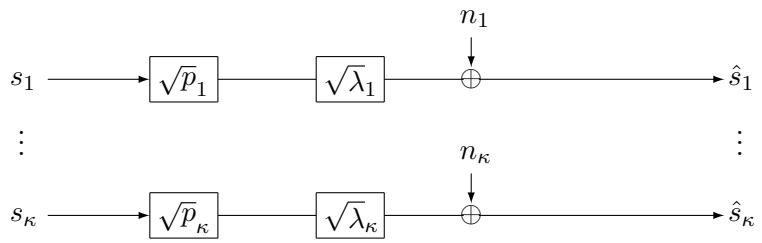
where $\mathbf{s}_\kappa \in \mathbb{C}^\kappa$ gathers the κ data symbols (zero-mean,[9] unit-energy and uncorrelated, i.e., $\mathbb{E}\{\mathbf{s}_\kappa{\mathbf{s}_\kappa}^\dagger\} = \mathbf{I}_\kappa$), $\mathbf{V}_\kappa$ is formed with the κ columns of $\mathbf{V}$ associated with the κ strongest channel

[8] We call this matrix diagonal even though it may be not square.

[9] The mean of the symbols does not carry any information and can always be set to zero saving power at the transmitter.



(i) System model.



(ii) Equivalent diagonal model.

Figure 3.3 System model of spatial multiplexing MIMO systems with CSI.

eigenmodes, and $\mathbf{P} = \mathrm{diag}\left(\{p_k\}_{k=1}^{\kappa}\right)$ is a diagonal matrix containing the power allocated to each established substream.

Assuming perfect channel knowledge also at the receiver, the symbols transmitted through the channel eigenmodes are recovered from the received signal $\mathbf{y}$ with matrix $\mathbf{U}_\kappa$, similarly defined to $\mathbf{V}_\kappa$, as (see Figure 3.3-(i))

$$\hat{\mathbf{s}}_\kappa = \mathbf{U}_\kappa^\dagger\big(\mathbf{H}\mathbf{V}_\kappa\sqrt{\mathbf{P}}_\kappa\mathbf{s}_\kappa + \mathbf{w}\big) = \sqrt{\mathbf{\Lambda}}_\kappa\sqrt{\mathbf{P}}_\kappa\mathbf{s}_\kappa + \mathbf{n}_\kappa \tag{3.37}$$

where $\mathbf{\Lambda}_\kappa = \mathrm{diag}\left(\{\lambda_k\}_{k=1}^{\kappa}\right)$ is a diagonal matrix that contains the κ largest channel eigenvalues (squared singular values) in descending order, $\lambda_1 \geq \cdots \geq \lambda_\kappa$, and the noise vector $\mathbf{n}_\kappa = \mathbf{U}_\kappa^\dagger\mathbf{w}$ has the same statistical properties as $\mathbf{w}$, possibly with a reduced dimension.

As shown in Figure 3.3-(ii), the spatial multiplexing MIMO system with CSI establishes κ independent data streams which are transmitted through the κ strongest channel eigenmodes experiencing an instantaneous SNR of

$$\rho_k = \lambda_k p_k \qquad \text{for } k = 1, \ldots, \kappa \tag{3.38}$$

where λ_k denotes the kth ordered channel eigenvalue and p_k defines the power allocation policy. This scheme is also known in the literature as MIMO SVD [Gar05, Jin06] or multichannel beamforming system [Jin08].

3.4.2 Power Constraint

The transmitted power is commonly constrained as

$$\mathbb{E}\{\|\mathbf{x}\|^2\} = \mathrm{tr}\big(\mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\}\big) = \sum_{k=1}^{\kappa} p_k \leq \mathsf{snr} \tag{3.39}$$

where the expectation is over the κ data symbols and snr is the average SNR per receive antenna. Note that (3.39) limits the total transmitted power in each channel state but does not impose explicitely any individual restriction on the elemental powers $\mathbb{E}\{|x_i|^2\}$ $(i = 1, \ldots, n_\mathsf{T})$ transmitted through each of the n_T antennas. Since it holds that $\mathbb{E}\{|x_i|^2\} \leq \lambda_{\mathsf{max}}\left\{\mathbb{E}\left\{\mathbf{x}\mathbf{x}^\dagger\right\}\right\}$, the elemental powers are usually controlled by imposing a peak-power constraint of the form [Sto02, Sec. II.B]

$$\lambda_{\mathsf{max}}\big\{\mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\}\big\} = \max_k p_k \leq \phi\mathsf{snr} \tag{3.40}$$

where ϕ is a given positive constant.

Whenever the constraints in (3.39) and (3.40) are simultaneously imposed, we have the following possible situations [Sto02, Sec. II.C]: (i) if $\phi \geq 1$, the sum-power constraint in (3.39) becomes more restrictive than the peak-power constraint in (3.40) and, hence, only (3.39) is active, (ii) if $\phi \leq \kappa^{-1}$, we have the opposite situation and only (3.40) is active, and (iii) if $\kappa^{-1} < \phi < 1$, both power constraints in (3.39) and (3.40) are active.

Both the sum-power constraint in (3.39) and the peak-power constraint in (3.40) are short-term power constraints, i.e., the power is limited for each channel state, as opposed to the less restrictive long-term power contraint in which the transmit power is limited in the average of all possible channel states (see e.g. [Big98, Sec. III.B] [Big01, eq. (9)]). Long-term power constraints are not considered in this thesis.

3.5 Individual Performance of Spatial Multiplexing MIMO Systems with CSI and Fixed Power Allocation

In this section we investigate the exact average and outage performance of the spatial multiplexing MIMO systems with CSI presented in Section 3.4 when a fixed (channel non-dependent) power allocation policy is adopted, i.e.,

$$p_k = \phi_k\mathsf{snr} \qquad \text{for } k = 1, \ldots, \kappa \tag{3.41}$$

where ϕ_k is a positive constant independent of the channel with $\sum_{i=1}^{\kappa} \phi_k = 1$. Under the fading channel models described in Section 3.3, we analyze the exact average BER and the outage probability of each individual substream. Furthermore we focus on the high-SNR regime and obtain parameterized analytic expressions for the previous individual performance measures.

3.5.1 Individual Average BER with Fixed Power Allocation

Under AWGN the instantaneous BER of the substream transmitted through the kth channel eigenmode can be analytically approximated as (see Section 3.2.2)

$$\mathsf{BER}_k(\rho_k) = \frac{\alpha_k}{\log_2 M_k} \mathcal{Q}\left(\sqrt{\beta_k \rho_k}\right) \qquad \text{for } k = 1, \ldots, \kappa \tag{3.42}$$

where α_k, β_k, and M_k are the constellation parameters (see Table 3.1), and ρ_k is the instantaneous SNR in (3.38), given in this case by

$$\rho_k = \lambda_k \phi_k \mathsf{snr} \qquad \text{for } k = 1, \ldots, \kappa. \tag{3.43}$$

The average BER is then obtained as

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \mathbb{E}_{\lambda_k}\{\mathsf{BER}_k(\rho_k)\} = \int_0^\infty \mathsf{BER}_k\left(\phi_k \lambda_k \mathsf{snr}\right) f_{\lambda_k}(\lambda_k) d\lambda_k \tag{3.44}$$

$$= \frac{\alpha_k}{\log_2 M_k} \int_0^\infty \mathcal{Q}\left(\sqrt{\beta_k \rho_k}\right) f_{\lambda_k}(\lambda_k) d\lambda_k \tag{3.45}$$

where $f_{\lambda_k}(\cdot)$ is the marginal pdf of the kth channel eigenvalue. Under the channel models in Definitions 3.1–3.4, the individual average BER given in the next theorem follows.

Theorem 3.1. *The individual average BER of the substream transmitted through the kth eigenmode of the $n_\mathsf{R} \times n_\mathsf{T}$ MIMO channels in Definitions* 3.1–3.4*, when the power allocation is fixed as in* (3.41)*, is given by*

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \frac{\alpha_k}{2 \log_2 M_k} \sqrt{\frac{\beta_k}{2\pi}} \int_0^\infty \frac{e^{-\frac{\beta_k \rho}{2}}}{\sqrt{\rho}} F_{\lambda_k}\Big(\frac{\rho}{\phi_k \mathsf{snr}}\Big) d\rho \tag{3.46}$$

where $F_{\lambda_k}(\cdot)$ is the marginal cdf of the kth largest eigenvalue in Theorem 2.2 *and the corresponding expressions for each channel model are given in Tables* 2.2–2.4.

Proof. Using integration by parts, we can rewrite the average BER in (3.44) as a function of the

cdf of λ_k:

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = -\int_0^\infty \frac{d}{d\rho}\mathsf{BER}_k\left(\rho\right) F_{\lambda_k}\Big(\frac{\rho}{\phi_k \mathsf{snr}}\Big) d\lambda_k \tag{3.47}$$

$$= \frac{\alpha_k}{2\log_2 M_k}\sqrt{\frac{\beta_k}{2\pi}}\int_0^\infty \frac{e^{-\frac{\beta_k \rho}{2}}}{\sqrt{\rho}} F_{\lambda_k}\Big(\frac{\rho}{\phi_k \mathsf{snr}}\Big) d\rho \tag{3.48}$$

and this completes the proof. □

In Section 3.2.2 we provided some efficient methods of calculating the average BER. In particular, the mgf-based approach of [Sim98] have been widely applied to calculate the average BER performance of different diversity combining techniques SIMO systems (see [Sim02b, Ch. 9] for a complete review). However, the mgf of the instantaneous SNR for the MIMO system presented in Section 3.4, is very difficult to obtain and, thus, the mgf method is not convenient without making further assumptions to simplify the analysis. For instance, it has been used in [Kan04,Maa06] to deal with the particular case of transmitting through the strongest eigenmode (beamforming or maximum ratio transmission (MRT) [Lo99]), i.e., $k = 1$, when $\min\{n_\mathsf{T}, n_\mathsf{R}\} = 2$.

The approach in Theorem 3.1 to express the average BER as a function of the cdf of the channel-dependent parameter was used in [Che04, eq. (32)] to obtain the average BER of selection combining under different SIMO fading channel models, and has been also recently applied in [Zan05, Sec. IV] [McK07, Sec. III.B] to analyze the average BER of MRT systems in uncorrelated, min-semicorrelated, and double-correlated Rayleigh, and uncorrelated MIMO Ricean channels and in [Jin08, Sec. IV.A] to analyze the individual average BER of the channel eigenmodes in an uncorrelated Ricean fading MIMO channel.

Although Theorem 3.1 provides an efficient numerical procedure to obtain the exact average BER without resorting to the time-consuming Monte Carlo simulations, it is still difficult to extract any conclusion on the inherent characteristics offered by the different eigenmodes. Hence we now focus on the high-SNR regime and characterize the average BER performance of the substream transmitted through the kth channel eigenmode in terms of array gain and diversity gain. Considering the approach of [Wan03] in Lemma 3.1, the array and diversity gains only depend on the fading distribution through its near-zero behaviour. Using the first order Taylor expansion of the marginal pdf of the kth largest channel eigenvalue

$$f_{\lambda_k}(\lambda_k) = a_k \lambda_k^{d_k} + o(\lambda_k^{d_k}) \tag{3.49}$$

derived in Theorem 2.4 for the distributions of the channels in Definitions 3.1, 3.2, and 3.4, the result in the next theorem follows.

Theorem 3.2. *The individual average BER of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1, 3.2, *and* 3.4, *when the power allocation is fixed as in* (3.41), *satisfies*

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \left(G_{\mathsf{a}}(k,\phi_k)\cdot \mathsf{snr}\right)^{-G_{\mathsf{d}}(k)} + o\left(\mathsf{snr}^{-G_{\mathsf{d}}(k)}\right) \tag{3.50}$$

where the diversity gain and the array gain are given by

$$G_{\mathsf{d}}(k) = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1) \tag{3.51}$$

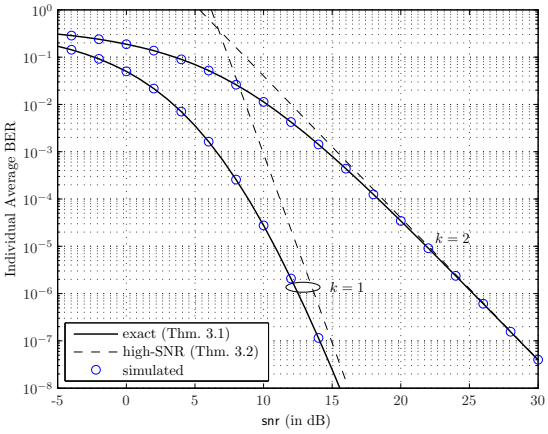
$$G_{\mathsf{a}}(k,\phi_k) = \beta_k \phi_k \left(\frac{\alpha_k}{\log_2 M_k} \frac{a_k 2^{d_k} \Gamma(d_k + 3/2)}{\sqrt{\pi}(d_k+1)} \right)^{-1/(d_k+1)} \tag{3.52}$$

where $\Gamma(\cdot)$ denotes the gamma function (see Definition 2.5*) and the parameters a_k and d_k model the fading distribution as done in Theorem* 2.4*. The corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

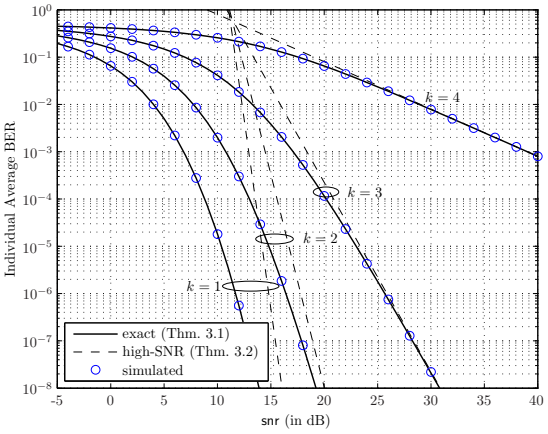
Proof. The proof follows from Lemma 3.1 and Theorem 2.4.

The average BER characterization given in Theorem 3.2 was initially presented in [Ord05b, Thm. 2] [Ord07b, Thm. 2] for uncorrelated Rayleigh fading MIMO channels and later in [Jin08, Thm. 4] for uncorrelated Ricean fading MIMO channels. It is interesting to note that when particularizing Theorem 3.7 to $k = 1$, i.e., a beamforming strategy, we obtain the full diversity of the channel $n_{\mathsf{T}} n_{\mathsf{R}}$, as has been widely observed in the literature, e.g., [Lo99, Dig03]. The diversity order of the case $k = \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$ has been also previously documented in [Bur02].

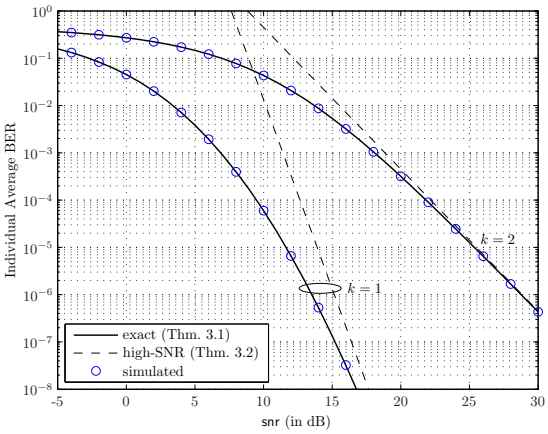
In Figure 3.4 we provide the average BER curves attained by the substreams transmitted through the eigenmodes of a MIMO channel with $n_{\mathsf{T}} = 2$ and $n_{\mathsf{R}} = 4$, and $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ in an uncorrelated Rayleigh, a semicorrelated Rayleigh and uncorrelated Ricean MIMO channel. We assume all substreams active ($\kappa = \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$), a uniform power allocation ($\phi_k = 1/\kappa$), and all data symbols drawn from a QPSK constellation. We can see that the result in Theorem 3.1 coincides with the average BER obtained through numerical Monte Carlo simulation and the result in Theorem 3.2 correctly predicts the diversity and array gain and, thus, approximates the average BER performance at medium to high SNR. Observe that, when the diversity gain



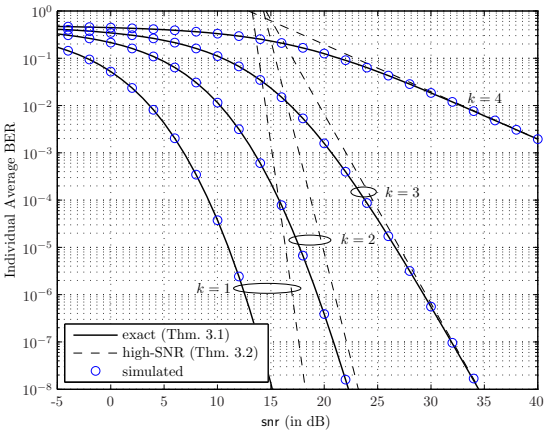
(i) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



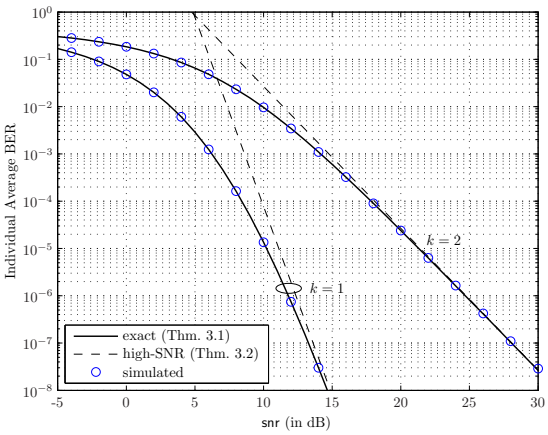
(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



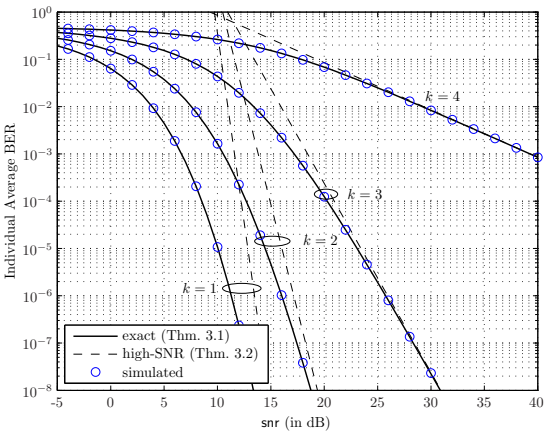
(iii) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, semicorrelated Rayleigh



(iv) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, semicorrelated Rayleigh

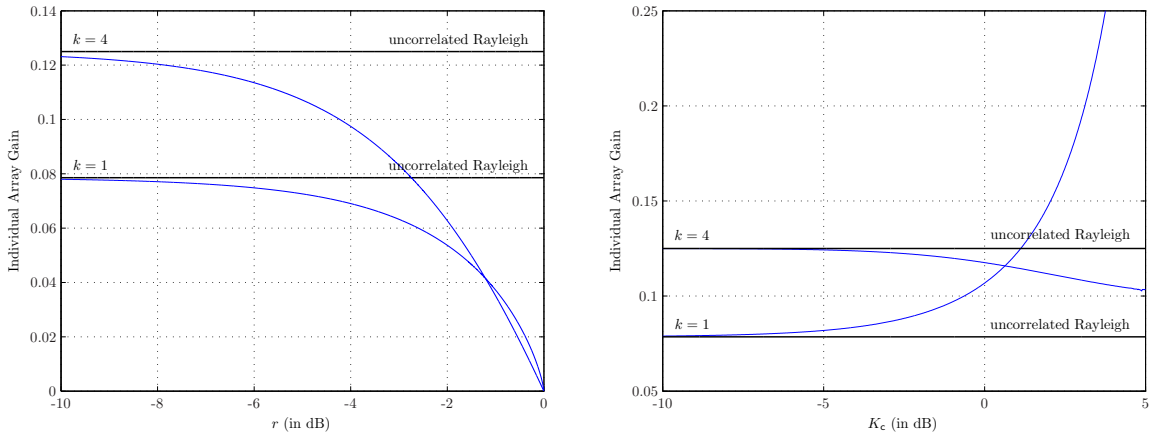


(v) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Ricean



(vi) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Ricean

Figure 3.4 Individual exact, simulated and parameterized average BER of the substreams transmitted by a spatial multiplexing MIMO systems with CSI in an uncorrelated Rayleigh, a semicorrelated Rayleigh (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$), and an uncorrelated Ricean (with Ricean factor $K_{\mathsf{c}} = 0$dB) MIMO channel.



(i) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, semicorrelated Rayleigh

(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Ricean

Figure 3.5 Individual array gain in a semicorrelated Rayleigh MIMO channel (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$) as a function of r and in an uncorrelated Ricean MIMO channel as a function of the Ricean factor K_{c}.

is high, as for the first substreams ($k = 1, 2$) in the cases $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, the BER decreases with the SNR so rapidly that the given approximation is only accurate for very small BER values.

For illustrative purposes, we plot in Figure 3.5 the individual array gain of the substreams transmitted through the strongest and weakest eigenmodes in (i) a semicorrelated Rayleigh and (ii) an uncorrelated Ricean MIMO channel as a function of (i) the correlation and (ii) the Ricean factor, respectively. Recall that the array gain depends on the fading distribution through the parameters d_k, which coincides for all investigated channels, and a_k, which is a function of (i) the eigenvalues of the correlation matrix or (ii) the eigenvalues of the line-of-sight matrix and the Ricean factor. As expected, when the correlation or the Ricean factor is low, both channel models become very close in distribution to an uncorrelated Rayleigh MIMO channel and so does the array gain. Despite this fact, these results are by no means representative of the behavior of a general system, since it can be substantially different for large correlation or large Ricean factor when the number of antennas, the correlation model, or line-of-sight matrix is modified.

3.5.2 Individual Outage Probability with Fixed Power Allocation

The outage probability as defined in (3.8) of the substream transmitted through the kth channel eigenmode is given by

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) \triangleq \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}, \rho_{\mathrm{th}}) = \mathsf{Pr}(\rho_k \leq \rho_{\mathrm{th}}) \tag{3.53}$$

Under the channel models in Definitions 3.1–3.4, the exact outage probability and the corresponding high-SNR characterization given in the next theorems follow.

Theorem 3.3. *The individual outage probability of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1–3.4*, when the power allocation is fixed as in* (3.41)*, is given by*

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = F_{\lambda_k}\Big(\frac{\rho_{\mathrm{th}}}{\phi_k \mathsf{snr}}\Big) \tag{3.54}$$

where $F_{\lambda_k}(\cdot)$ is the marginal cdf of the kth largest eigenvalue in Theorem 2.2 *and the corresponding expressions for each channel model are given in Tables* 2.2–2.4.

Proof. The proof follows from substituting instantaneous SNR in (3.41) in the outage probability definition of (3.53). □

Theorem 3.4. *The individual outage probability of the substream using the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1*,* 3.2*, and* 3.4*, when the power allocation is fixed as in* (3.41)*, satisfies*

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = \big(O_{\mathsf{a}}(k,\phi_k)\cdot \mathsf{snr}\big)^{-G_{\mathsf{d}}(k)} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}(k)}\big) \tag{3.55}$$

where the diversity and the outage array gain are given by

$$G_{\mathsf{d}}(k) = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1) \tag{3.56}$$

$$O_{\mathsf{a}}(k,\phi_k) = \frac{\phi_k}{\rho_{\mathrm{th}}}\left(\frac{a_k}{d_k+1}\right)^{-1/(d_k+1)} \tag{3.57}$$

and the parameters a_k and d_k model the fading distribution as done in Theorem 2.4 *and the corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

Proof. The proof follows from using Lemma 3.2 with the instantaneous SNR in (3.41). □

The individual outage probability when transmitting through the strongest eigenmode, i.e., for $\kappa = 1$ in (3.53) and $p_1 = \mathsf{snr}$, has been widely analyzed in the literature, since it corresponds to the outage probability of the MRT or beamforming scheme. In particular, the outage probability under uncorrelated Rayleigh fading was obtained in [Dig03, Sec. IV] [Kan03b, Sec. III] [Gra05, Sec. II] [Maa05, Sec. IV], under semicorrelated Rayleigh fading in [Kan03a, Sec. IV], and under uncorrelated Ricean fading in [Kan03b, Sec. III]. Additionally, the case of double-correlated

Rayleigh fading MIMO channels (not included in this chapter) has been recently addressed in [McK07, Sec. IV]. The individual outage probability of the κ substreams in an uncorrelated Ricean fading MIMO channel was previously investigated in [Jin08, Sec. IV.B].

In Figure 3.6 we provide the average individual probability curves attained by the substreams transmitted through the eigenmodes of a MIMO channel with $n_{\mathsf{T}} = 2$ and $n_{\mathsf{R}} = 4$, and $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ in an uncorrelated Rayleigh, a semicorrelated Rayleigh and uncorrelated Ricean MIMO channel. We assume all substreams active ($\kappa = \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$), a uniform power allocation ($\phi_k = 1/\kappa$), and all data symbols drawn from a QPSK constellation. Likewise the individual average BER plots, the theoretical expressions in Theorems 3.3 and 3.4 correctly predict the individual outage probability obtained through numerical Monte Carlo simulation.

3.6 Individual Performance of Spatial Multiplexing MIMO Systems with CSI and Non-Fixed Power Allocation

We have so far analyzed spatial multiplexing MIMO systems with fixed power allocation strategies as expressed in (3.41). We now consider power allocation strategies that depend on the eigenvalues associated with the κ active channel eigenmodes, i.e.,

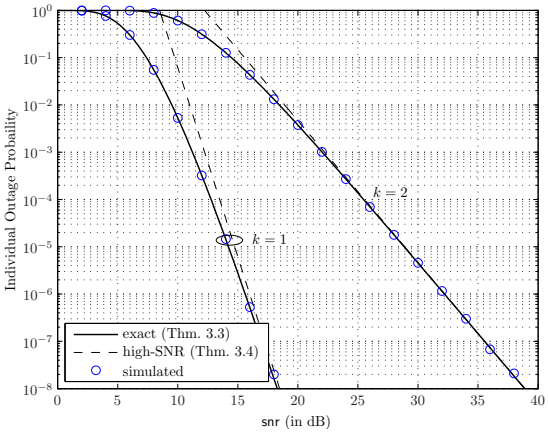
$$p_k \triangleq p_k(\lambda_1, \ldots, \lambda_\kappa, \mathsf{snr}) \qquad \text{for } k = 1, \ldots, \kappa \tag{3.58}$$

and satisfy the short-term power constraint in (3.39).

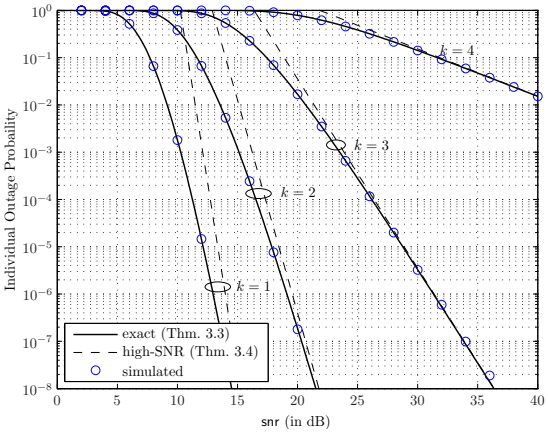
Owing to the difficulty of drawing any conclusion from an exact performance analysis while keeping the power allocation policy general, we concentrate in this section only on the high-SNR regime. This allows us to investigate the performance limits common to any fixed or non-fixed power allocation policy, i.e., the fundamental performance limits of the individual channel eigenmodes.

3.6.1 Individual Average BER with Non-Fixed Power Allocation

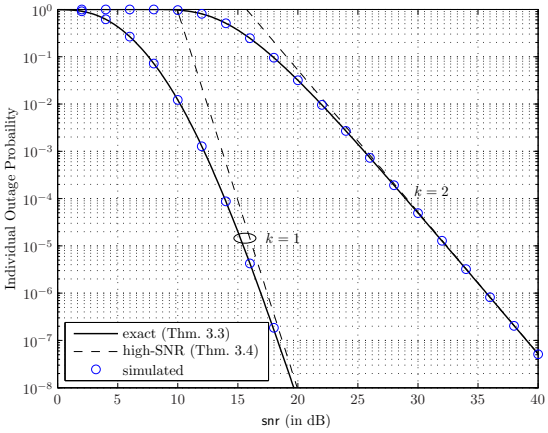
Most common channel-dependent power allocation policies discard transmission through the worst eigenchannels by setting the transmit power to zero for the corresponding substreams. The process by which the power is distributed among the established substreams reminds the way in which water distributes itself in a vessel and hence is referred to as waterfilling [Cov91,



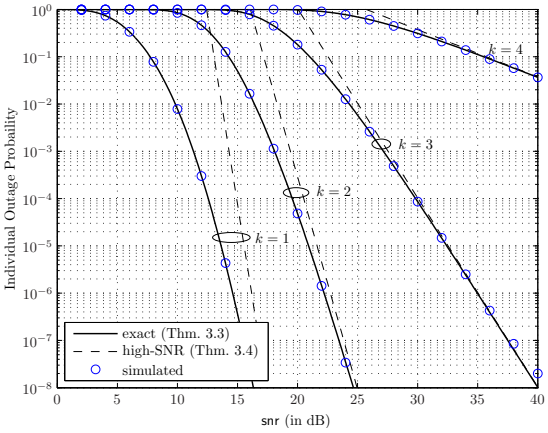
(i) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



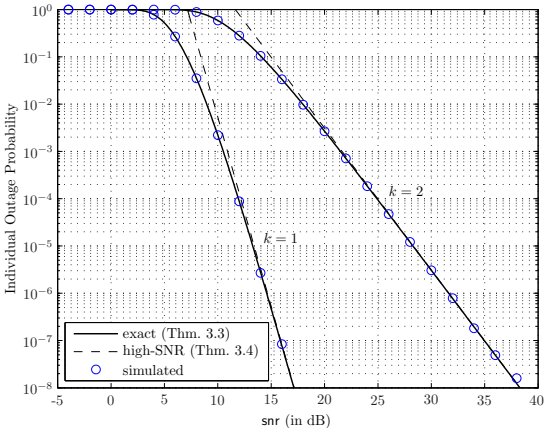
(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



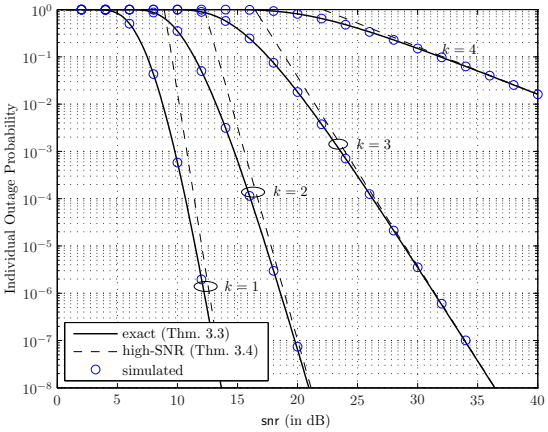
(iii) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, semicorrelated Rayleigh



(iv) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, semicorrelated Rayleigh



(v) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Ricean



(vi) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Ricean

Figure 3.6 Individual exact, simulated and parameterized outage probability of the substreams transmitted by a spatial multiplexing MIMO systems with CSI in an uncorrelated Rayleigh, a semicorrelated Rayleigh (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$), and an uncorrelated Ricean (with Ricean factor $K_{\mathsf{c}} = 0$dB) MIMO channel.

Sec. 10.4]. In order to take into account in the analysis this waterfilling mechanism, we introduce the following assumption.

Assumption 3.1. *Let us assume that the non-fixed power allocation in* (3.58) satisfies

$$\mathsf{Pr}\left(p_k < \xi_k \mathsf{snr}\right) = a(\xi_k)\mathsf{snr}^{-d(\xi_k)} + o\big(\mathsf{snr}^{-d(\xi_k)}\big) \tag{3.59}$$

where ξ_k, $a(\xi_k)$, and $d(\xi_k)$ are positive deterministic parameters.

Observe that fixed power allocations in (3.41) also satisfy Assumption 3.1 with $\xi_k = \phi_k$ ($\mathsf{Pr}\left(p_k < \xi_k \mathsf{snr}\right) = 0$ and $d(\xi_k) \to \infty$) and, hence, they are also included in the analysis.

Theorem 3.5. *Consider a non-fixed power allocation as in* (3.58) *under the short-term power constraint in* (3.39) *and Assumption* 3.1*. Then, the individual average BER of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1*,* 3.2*, and* 3.4 *satisfies*

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \left(G_{\mathsf{a}}(k) \cdot \mathsf{snr}\right)^{-G_{\mathsf{d}}(k)} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}(k)}\big) \tag{3.60}$$

where the diversity gain is given by

$$G_{\mathsf{d}}(k) = d_k + 1 = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1) \tag{3.61}$$

whenever[10] $d(\xi_k) \geq d_k + 1$ *and the outage array gain $O_{\mathsf{a}}(k)$ can be bounded by distinguishing between two cases:*

(i) *If there exists $0 < \xi_k < 1$ such that $d(\xi_k) > d_k + 1$, the array gain is bounded as*[11]

$$G_{\mathsf{a}}(k, \xi_k) \leq G_{\mathsf{a}}(k) < G_{\mathsf{a}}(k, 1). \tag{3.62}$$

(ii) *If $d(0) = d_k + 1$ and there exists $0 < \xi_k < 1$ such that $d(\xi_k) = d_k + 1$, the array gain is bounded as*

$$\left(G_{\mathsf{a}}(k, \xi_k)^{-(d_k+1)} + \frac{\alpha_k}{2\log_2 M_k} a(\xi_k)\right)^{-1/(d_k+1)} < G_{\mathsf{a}}(k) < G_{\mathsf{a}}(k, 1) \tag{3.63}$$

where $G_{\mathsf{a}}(k, \xi_k)$ is the array gain obtained when using a fixed power allocation for $\phi_k = \xi_k$ and is defined in (3.52)*.*

---

[10] The condition $d(\xi_k) \geq d_k + 1$ is satisfied by any reasonable power allocation. If $d(\xi_k) < d_k + 1$, the average BER performance is inherently limited by the power allocation and not by the statistical properties of the channel.

[11] Observe that an array gain lower bound provides an upper bound on the high-SNR average BER and vice versa.

The parameters a_k and d_k model the fading distribution as done in Theorem 2.4 *and the corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

Proof. See Appendix 3.B.1.

Theorem 3.5 shows that the diversity gain associated to a substream transmitted through the strongest kth channel eigenmode does not depend on the power allocation and is fundamentally limited to $(n_{\mathsf{T}}-k+1)(n_{\mathsf{R}}-k+1)$. The power allocation can only possibly increase the array gain as it will be investigated in Section 3.8, where practical power allocation policies are analyzed in in the context of linear MIMO transceivers.

3.6.2 Individual Outage Probability with Non-Fixed Power Allocation

Similarly to the average BER analysis in Theorem 3.5 we can bound the individual outage probability of a spatial multiplexing MIMO system with CSI and a non-fixed power allocation policy as shown in the next theorem.

Theorem 3.6. *Consider a non-fixed power allocation as in* (3.58) *under the short-term power constraint in* (3.39) *and Assumption* 3.1. *Then, the individual outage probability of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1, 3.2, *and* 3.4 *satisfies*

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = \left(O_{\mathsf{a}}(k) \cdot \mathsf{snr}\right)^{-G_{\mathsf{d}}(k)} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}(k)}\big) \tag{3.64}$$

where the diversity gain is given by

$$G_{\mathsf{d}}(k) = d_k + 1 = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1) \tag{3.65}$$

whenever $d(\phi_k) \geq d_k + 1$ and the outage array gain $O_{\mathsf{a}}(k)$ can be bounded by distinguishing between two cases:

(i) *If there exists $0 < \xi_k < 1$ such that $d(\xi_k) > d_k + 1$, the array gain is bounded as*

$$O_{\mathsf{a}}(k, \xi_k) \leq O_{\mathsf{a}}(k) < O_{\mathsf{a}}(k, 1). \tag{3.66}$$

(ii) *If $d(0) = d_k + 1$ and there exists $0 < \xi_k < 1$ such that $d(\xi_k) = d_k + 1$, the array gain is bounded as*

$$\left(O_{\mathsf{a}}(k, \xi_k)^{-(d_k+1)} + a(\xi_k)\right)^{-1/(d_k+1)} < O_{\mathsf{a}}(k) < O_{\mathsf{a}}(k, 1) \tag{3.67}$$

where $O_{\mathsf{a}}(k,\xi_k)$ is the array gain obtained when using a fixed power allocation for $\phi_k = \xi_k$ and is defined in (3.57).

The parameters a_k and d_k model the fading distribution as done in Theorem 2.4 *and the corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

Proof. See Appendix 3.B.2.

3.7 Global Performance of Spatial Multiplexing MIMO Systems with CSI

Finally, we investigate the global average BER and global outage probability of the general spatial multiplexing MIMO system with CSI described in Section 3.4, i.e., when transmitting over the κ strongest channel eigenmodes. The final aim of this section is to characterize the quality of the system with a single performance measure.

3.7.1 Global Average BER

We define first the global instantaneous BER as the arithmetic mean of the instantaneous BER of the κ established substreams:

$$\mathsf{BER}\left(\{\rho_k\}_{k=1}^{\kappa}\right) \triangleq \frac{1}{\kappa}\sum_{k=1}^{\kappa}\mathsf{BER}_k(\rho_k). \tag{3.68}$$

Note that the global instantaneous BER in (3.68) takes into account the instantaneous BER performance experienced by each one of the κ data symbols to be transmitted, i.e., the number of substreams κ is fixed regardless of the power assigned to each substream, which can even be zero for some power allocation strategies under poor propagation conditions. We obtain the global average BER performance of the spatial multiplexing MIMO system by averaging the global instantaneous BER in (3.68) over all possible channel states:

$$\overline{\mathsf{BER}}(\mathsf{snr}) \triangleq \mathbb{E}\left\{\mathsf{BER}\left(\{\rho_k\}_{k=1}^{\kappa}\right)\right\} = \frac{1}{\kappa}\sum_{k=1}^{\kappa}\overline{\mathsf{BER}}_k(\mathsf{snr}). \tag{3.69}$$

For spatial multiplexing MIMO systems with fixed power allocation policies as in (3.41), the exact global average BER in (3.69) can be obtained using Theorem 3.1 to calculate each one of the individual average BERs. In general, we can still characterize the global average BER in the high-SNR regime. Since $G_{\mathsf{d}}(1) > G_{\mathsf{d}}(2) > \cdots > G_{\mathsf{d}}(\kappa)$ (see Theorem 3.2 for fixed and Theorem

3.5 for non-fixed power allocations), the global average BER is dominated by the average BER associated with the κth substream:

$$\overline{\mathsf{BER}}(\mathsf{snr}) = \frac{1}{\kappa}\left(G_{\mathsf{a}}(\kappa)\cdot\mathsf{snr}\right)^{-G_{\mathsf{d}}(\kappa)} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}(\kappa)}\big) \tag{3.70}$$

where $G_{\mathsf{a}}(\kappa)$ and $G_{\mathsf{d}}(\kappa)$ denote the array and the diversity gain associated with the κth substream. This result is summarized in the following theorem.

Theorem 3.7. *The global average BER of a spatial multiplexing MIMO system transmitting through the κ strongest eigenmodes of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1*,* 3.2*, and* 3.4*, when the power allocation is either fixed as in* (3.41) *or non-fixed as in* (3.58), satisfies

$$\overline{\mathsf{BER}}(\mathsf{snr}) = \left(G_{\mathsf{a}}\cdot\mathsf{snr}\right)^{-G_{\mathsf{d}}} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}}\big) \tag{3.71}$$

where the diversity gain G_{d} and the array gain G_{a} are given by

$$G_{\mathsf{d}} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1) \tag{3.72}$$

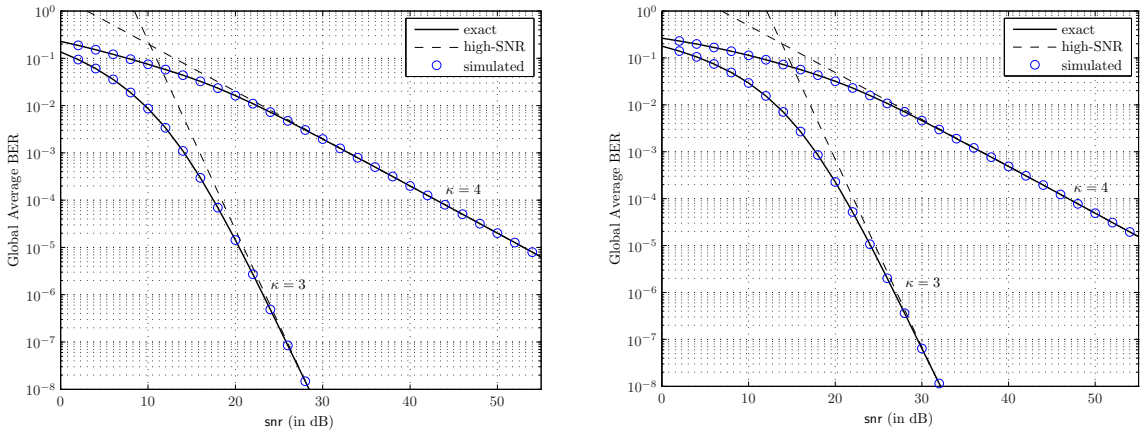
$$G_{\mathsf{a}} = \kappa^{1/G_{\mathsf{d}}} G_{\mathsf{a}}(\kappa) \tag{3.73}$$

and $G_{\mathsf{a}}(\kappa)$ is the array gain of the κth substream either given by (3.52) for fixed power allocations or bounded by (3.62) and (3.63) for non-fixed power allocations.

In Figure 3.7 we show the global average BER of a spatial multiplexing MIMO system with $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ and $\kappa = \{3, 4\}$ in an uncorrelated and a semicorrelated Rayleigh fading MIMO channel. The transmit power is uniform distributed ($\phi_k = 1/\kappa$) and all data streams use QPSK constellations. In all three cases the exact global average BER with the numerical performance is correctly approximated by the parameterized characterization proposed in Theorem 3.7.

3.7.2 Global Outage Probability

In this section we analyze the global outage probability of the spatial multiplexing MIMO system described in Section 3.4. Consider, for instance, that κ services or substreams with possibly different performance constraints are multiplexed by accommodating each service in a different channel eigenmode. The global outage probability can be defined in many different ways depending on how the application of interest takes into account the individual outages of the established substreams. In the following we provide two illustrative examples.



(i) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Rayleigh

(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, semicorrelated Rayleigh

Figure 3.7 Global exact, simulated and parameterized BER of a spatial multiplexing MIMO system with κ active substreams and a uniform power allocation among them in an uncorrelated Rayleigh and a semicorrelated Rayleigh (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$) MIMO channel.

All-Outage Probability: Assume that the communication process is regarded as successful if at least one of the substreams achieves the desired performance. Then, a global outage event is declared only when all used channel eigenmodes fail to offer their corresponding target performance and, hence, the global outage probability is defined as

$$\mathsf{P}_{\mathsf{out}}^{(\mathrm{all})}(\mathsf{snr}) \triangleq \mathsf{Pr}(\rho_1 \leq \rho_{\mathrm{th},1}, \ldots, \rho_\kappa \leq \rho_{\mathrm{th},\kappa}) = F_{\boldsymbol{\lambda}}\Big(\frac{\rho_{\mathrm{th},1}}{\phi_1 \mathsf{snr}}, \ldots, \frac{\rho_{\mathrm{th},\kappa}}{\phi_\kappa \mathsf{snr}}, \infty, \ldots, \infty\Big) \tag{3.74}$$

$$= F_{\boldsymbol{\lambda}}\Big(\frac{\rho_{\mathrm{th},1}}{\phi_1 \mathsf{snr}}, \ldots, \frac{\rho_{\mathrm{th},\kappa}}{\phi_\kappa \mathsf{snr}}, \frac{\rho_{\mathrm{th},\kappa}}{\phi_\kappa \mathsf{snr}}, \ldots, \frac{\rho_{\mathrm{th},\kappa}}{\phi_\kappa \mathsf{snr}}\Big) \tag{3.75}$$

where $\rho_{\mathrm{th},1}, \ldots, \rho_{\mathrm{th},\kappa}$ denote the individual target performances and $F_{\boldsymbol{\lambda}}(\cdot)$ is the joint cdf of the ordered channel eigenvalues. Under the MIMO channel models presented in Definitions 3.1–3.4, $F_{\boldsymbol{\lambda}}(\cdot)$ can be obtained particularizing Theorem 2.1 with the expression of the corresponding parameters in Tables 2.2–2.4.

When equal target performances are imposed on all substreams, i.e., $\rho_{\mathrm{th},k} = \rho_{\mathrm{th}}$ for $k = 1, \ldots, \kappa$, and a uniform power allocation is used, i.e., $\phi_k = 1/\kappa$, the outage probability in (3.75) is simply given by

$$\mathsf{P}_{\mathsf{out}}^{(\mathrm{all})}(\mathsf{snr}) = \mathsf{Pr}(\rho_1 \leq \rho_{\mathrm{th}}) = F_{\lambda_1}\Big(\frac{\kappa \rho_{\mathrm{th}}}{\mathsf{snr}}\Big) \tag{3.76}$$

where $F_{\lambda_1}(\cdot)$ denotes the marginal cdf of the largest channel eigenvalue given in Corollary 2.2.1.

Any-Outage Probability: Assume that the quality of all substreams has to be simultaneously guaranteed, then a global outage event is declared whenever at least one of the used channel

eigenmodes cannot offer the desired performance. In this case, the global outage probability is defined as

$$\mathsf{P}_{\mathsf{out}}^{(\mathrm{any})}(\mathsf{snr}) \triangleq 1 - \mathsf{Pr}(\rho_1 > \rho_{\mathrm{th},1}, \ldots, \rho_\kappa > \rho_{\mathrm{th},\kappa}) = 1 - CF_{\boldsymbol{\lambda}}\Big(\frac{\rho_{\mathrm{th},1}}{\phi_1 \mathsf{snr}}, \ldots, \frac{\rho_{\mathrm{th},\kappa}}{\phi_\kappa \mathsf{snr}}, 0, \ldots, 0\Big) \quad (3.77)$$

where $CF_{\boldsymbol{\lambda}}(\cdot)$ denotes the joint complementary cdf of the ordered channel eigenvalues and can be obtained with techniques similar to those used to derive the joint cdf in Theorem 2.1.

When equal target performances ρ_{th} are imposed on all substreams and a uniform power allocation is used, the outage probability in (3.77) is simply given by

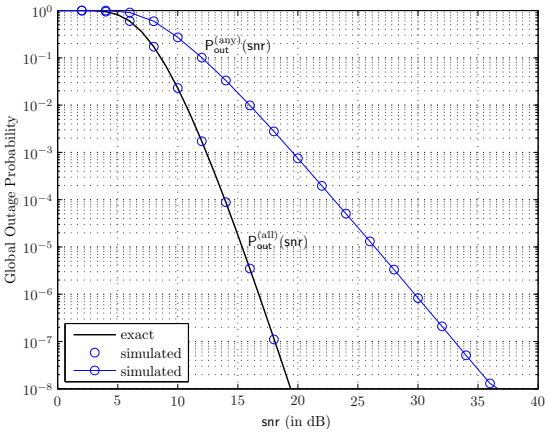
$$\mathsf{P}_{\mathsf{out}}^{(\mathrm{any})}(\mathsf{snr}) = 1 - \mathsf{Pr}(\rho_\kappa > \rho_{\mathrm{th}}) = F_{\lambda_\kappa}\Big(\frac{\kappa \rho_{\mathrm{th}}}{\mathsf{snr}}\Big) \quad (3.78)$$

where $F_{\lambda_\kappa}(\cdot)$ denotes the marginal cdf of the κth largest channel eigenvalue in Theorem 2.2.

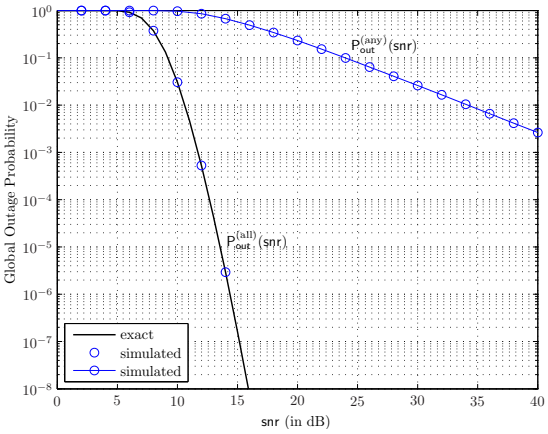
In Figure 3.8 we compare the global all-outage and any-outage probabilities of a spatial multiplexing MIMO system with $n_{\mathsf{T}} = 2$ and $n_{\mathsf{R}} = 4$, and $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ in an uncorrelated Rayleigh, a semicorrelated Rayleigh and uncorrelated Ricean MIMO channel. We assume all substreams active ($\kappa = \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$), a uniform power allocation ($\phi_k = 1/\kappa$), and all data symbols drawn from a QPSK constellation. The target performances $\rho_{\mathrm{th},1}, \ldots, \rho_{\mathrm{th},\kappa}$ have been chosen using (3.9) to guarantee the following target BERs: $\mathsf{BER}_{\mathrm{th},1} = 10^{-4}$ and $\mathsf{BER}_{\mathrm{th},2} = 10^{-2}$ for the case $\kappa = 2$; and $\mathsf{BER}_{\mathrm{th},1} = 10^{-4}$, $\mathsf{BER}_{\mathrm{th},2} = 10^{-3}$, $\mathsf{BER}_{\mathrm{th},3} = 10^{-2}$, and $\mathsf{BER}_{\mathrm{th},4} = 10^{-1}$ for the case $\kappa = 4$.

3.8 High-SNR Global Performance of Linear MIMO Transceivers

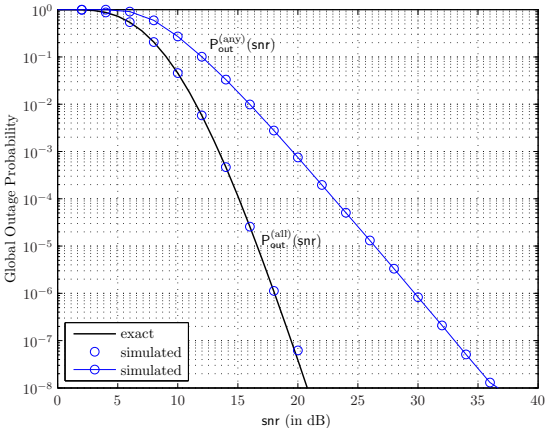
The performance of linear MIMO transceivers has been always analyzed numerically, due to the difficulty of finding a closed-form expression for the average bit error probability. For instance, in [Sam01] we can find the simulated average BER curves for linear precoding schemes designed under the minimum weighted MSE criterion in an uncorrelated Rayleigh fading MIMO channel. Other numerical results can also be found in [Yan94b, Sca99, Pal03]. The advantage of obtaining numerical results via computer simulation is that they provide the performance in realistic environments. However, they do not give insight into the behavior of the system as analytical expressions do. In this section we fill the gap by applying the results obtained in Section 3.7 to analytically characterize the high-SNR global average BER performance of the linear MIMO transceivers given in the unifying framework of [Pal03].



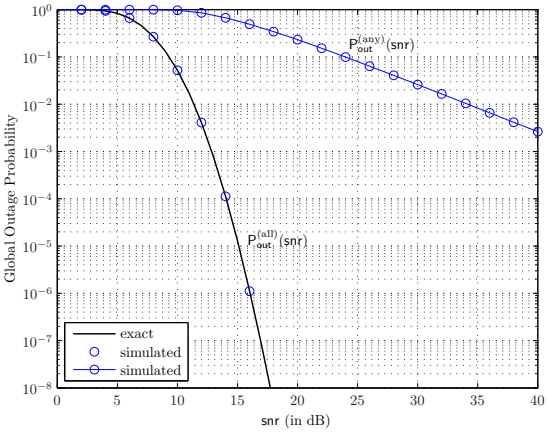
(i) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



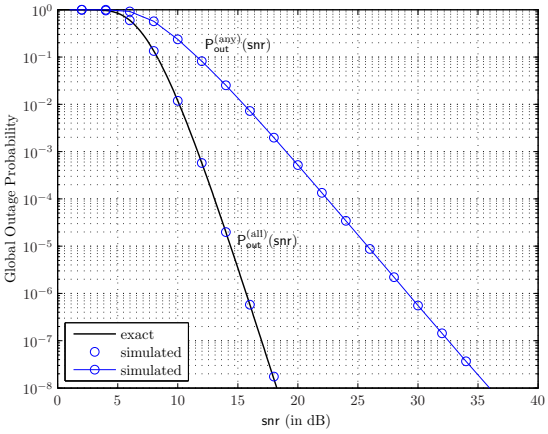
(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Rayleigh



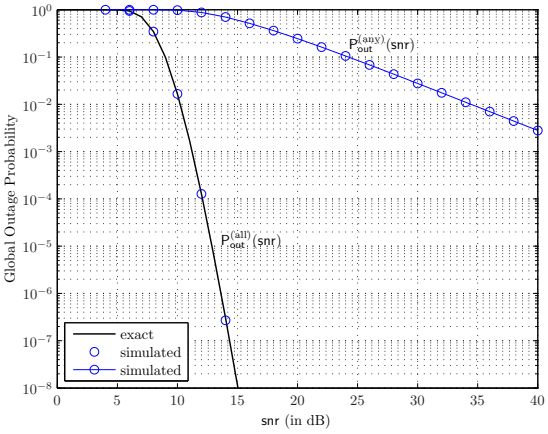
(iii) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, semicorrelated Rayleigh



(iv) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, semicorrelated Rayleigh



(v) $n_{\mathsf{T}} = 2$, $n_{\mathsf{R}} = 4$, uncorrelated Ricean



(vi) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, uncorrelated Ricean

Figure 3.8 Exact instantaneous BER, approximated instantaneous BER in (3.4), and high-SNR approximated instantaneous BER in (3.5).

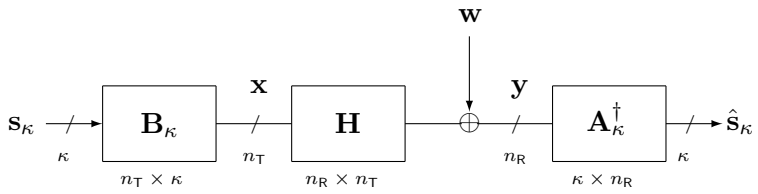


Figure 3.9 Linear MIMO transceivers system model.

3.8.1 Linear MIMO Transceivers Design

Suppose that the general MIMO communication system of (3.34) is equipped with a linear transceiver (linear precoder and linear equalizer) as shown in Figure 3.9. The transmitted vector is given by

$$\mathbf{x} = \mathbf{B}_\kappa \mathbf{s}_\kappa \tag{3.79}$$

where $\mathbf{B}_\kappa \in \mathbb{C}^{n_\mathsf{T} \times \kappa}$ is the transmit matrix (precoder), and $\mathbf{s}_\kappa \in \mathbb{C}^\kappa$ gathers the $\kappa \leq \min\{n_\mathsf{T}, n_\mathsf{R}\}$ data symbols to be transmitted (zero-mean, unit-energy and uncorrelated, i.e., $\mathbb{E}\{\mathbf{s}_\kappa \mathbf{s}_\kappa^\dagger\} = \mathbf{I}_\kappa$) drawn from a set of constellations. The average transmit power is constrained to satisfy

$$\mathbb{E}\big\{\|\mathbf{s}_\kappa\|^2\big\} = \mathrm{tr}\big(\mathbf{B}_\kappa \mathbf{B}_\kappa^\dagger\big) \leq \mathsf{snr} \tag{3.80}$$

where snr is the average SNR at each receive antenna. Similarly, the estimated data vector at the receiver is

$$\hat{\mathbf{s}}_\kappa = \mathbf{A}_\kappa^\dagger \mathbf{y} \tag{3.81}$$

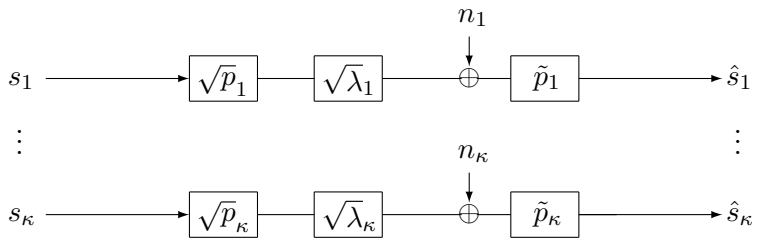
where $\mathbf{A}_\kappa^\dagger \in \mathbb{C}^{\kappa \times n_\mathsf{R}}$ is the receive matrix (equalizer).

The general problem of designing the optimal linear MIMO transceiver under perfect CSI knowledge is formulated in [Pal03] as the minimization of some cost function of the MSEs, since other common system quality measures such as the SNR, or the BER can be easily related to the MSE. Assuming that κ data symbols have to be communicated at each channel use, [Pal03] shows that (i) the optimum receive matrix $\mathbf{A}_\kappa$, for a given transmit matrix $\mathbf{B}_\kappa$, is the Wiener filter solution [Pal03, eq. (7)]:

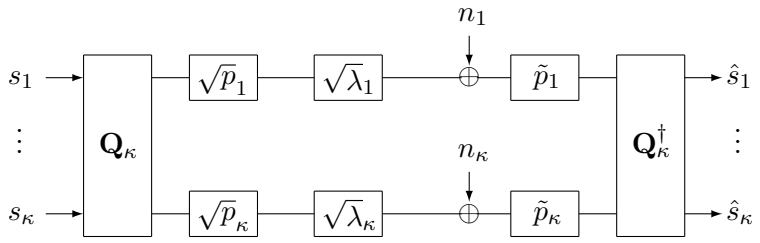
$$\mathbf{A}_\kappa = \big(\mathbf{H}\mathbf{B}_\kappa \mathbf{B}_\kappa^\dagger \mathbf{H}^\dagger + \mathbf{I}_{n_\mathsf{R}}\big)^{-1} \mathbf{H}\mathbf{B}_\kappa \tag{3.82}$$

and (ii) the optimum transmit matrix $\mathbf{B}_\kappa$, for a wide family of design criteria (with Schur-concave and Schur-convex cost functions), has the following form [Pal03, eq. (14) and (15)]:

$$\mathbf{B}_\kappa = \mathbf{U}_\kappa \sqrt{\mathbf{P}}_\kappa \mathbf{Q}_\kappa \tag{3.83}$$



(i) Schur-concave designs (diagonal transmission).



(ii) Schur-convex designs (non-diagonal transmission).

Figure 3.10 Spatial multiplexing MIMO systems with CSI model ($\tilde{p}_k = \sqrt{p_k \lambda_k}/(1 + p_k \lambda_k)$).

where $\mathbf{U}_\kappa \in \mathbb{C}^{n_\mathsf{T} \times \kappa}$ has as columns the eigenvectors of $\mathbf{H}^\dagger \mathbf{H}$ corresponding to the κ largest nonzero eigenvalues, $\mathbf{Q}_\kappa \in \mathbb{C}^{\kappa \times \kappa}$ is a unitary matrix, and $\mathbf{P} \in \mathbb{C}^{\kappa \times \kappa}$ is a diagonal matrix with diagonal entries equal to $\{p_k\}_{k=1}^{\kappa}$, that represent the power allocated to each established substream and depend on the particular design cost function. Owing to the power constraint in (3.80), the power allocation $\{p_k\}_{k=1}^{\kappa}$ satisfies

$$\sum_{k=1}^{\kappa} p_k = \mathsf{snr}. \tag{3.84}$$

For Schur-concave objective functions (see examples in Tables 3.2 and 3.3), $\mathbf{Q}_\kappa = \mathbf{I}_\kappa$ and the global communication process including pre- and post-processing is fully diagonalized as shown in Figure 3.10-(i). For Schur-convex objective functions (see examples in Table 3.4), however, $\mathbf{Q}_\kappa$ is a unitary matrix such that $(\mathbf{I}_\kappa + \mathbf{B}_\kappa^\dagger \mathbf{H}^\dagger \mathbf{H} \mathbf{B}_\kappa)^{-1}$ has identical diagonal elements (see [Pal03, Sec. IV.B] for details). In this case, the communication process is diagonalized up to a very specific rotation of the data symbols as shown in Figure 3.10-(ii).

Given the transmit matrix in (3.83) and the receive matrix in (3.82), the components of the estimated data signal $\hat{\mathbf{s}}$ are equal to (possibly with an additional pre- and post-processing of the data symbols s_k in the case of Schur-convex cost functions)

$$\hat{s}_k = \frac{p_k \lambda_k}{1 + p_k \lambda_k} s_k + \frac{\sqrt{p_k \lambda_k}}{1 + p_k \lambda_k} n_k \qquad \text{for } k = 1, \ldots, \kappa \tag{3.85}$$

Design criterion	Optimal power allocation
Maxim. weighted sum of SNRs [Pal03]	$p_k = \mathsf{snr}$ if $\omega_k \lambda_k$ is max. and 0 otherwise
Maxim. product of SNRs [Pal03]	$p_k = \mathsf{snr}/\kappa$
Maxim. weighted product of SNRs [Pal03]	$p_k = \mathsf{snr}\omega_k / \sum_{i=1}^{\kappa} \omega_i$

Table 3.2 Examples of diagonal schemes with fixed power allocation [Pal07, Tab. 3.1].

with instantaneous SNR given by

$$\rho_k = \lambda_k p_k \qquad \text{for } k = 1, \ldots, \kappa \tag{3.86}$$

where $\lambda_1, \ldots, \lambda_\kappa$ are the κ largest nonzero eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ in decreasing order and the complex κ-dimensional vector $\mathbf{n}_\kappa = (n_1, \ldots, n_\kappa)^{'}$ is a normalized equivalent noise vector with i.i.d. zero-mean, unit-variance, Gaussian entries.

In summary, linear MIMO transceivers transforms the MIMO channel into κ SISO channels, in which each signal component (possibly after a rotation) corresponds to a different substream transmitted in parallel through a different channel eigenmode. Hence, they can be treated as particular cases of the general spatial multiplexing MIMO system introduced in Section 3.4.

3.8.2 Performance of Diagonal Schemes with Fixed Power Allocation

Several design criteria with a Schur-concave cost function found in the literature fall within the class of diagonal schemes with fixed power allocations as summarized in Table 3.2. Examples are the maximization of the (weighted) sum of SNRs and the maximization of the (weighted) product of the SNRs. Furthermore, if the short-term power constraint is substituted by a peak-power constraint (see Section 3.4.2), the optimum spatial multiplexing system transmits always at full allowed power through each active channel eigenmode independently of the channel state [Sca02b, Lov05a]. For all these schemes, the exact individual performance is given in Theorems 3.1 and 3.3, the high-SNR individual performance in Theorems 3.2 and 3.4, and the global performance is addressed in Section 3.7.

Design criterion	Optimal power allocation
Minim. sum of the MSEs [Lee76, Sal85, Yan94b, Sca99, Pal03]	$p_k = (\mu\lambda_k^{-1/2} - \lambda_k^{-1})^+$
Minim. weighted sum of the MSEs [Lee76, Sam01, Pal03]	$p_k = (\mu\omega_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1})^+$
Minim. product of MSEs [Yan94a, Pal03]	$p_k = (\mu - \lambda_k^{-1})^+$
Minim. weighted product of MSEs [Pal03]	$p_k = (\mu\omega_k - \lambda_k^{-1})^+$
Maxim. mutual information [Cov91]	$p_k = (\mu - \lambda_k^{-1})^+$

Table 3.3 Examples of diagonal schemes with non-fixed power allocation [Pal07, Tab. 3.1].

3.8.3 Performance of Diagonal Schemes with Non-Fixed Power Allocation

Some other design criteria with Schur-concave cost functions found in the literature, which still have a diagonal structure, use non-fixed power allocations. From the summary in Table 3.3, we distinguish two different power allocation strategies which we analyze in the following.

For instance, we consider first design criteria that lead to waterfilling power allocations of the type

$$p_{\mathsf{wf},k} = \left(\mu\omega_k - \lambda_k^{-1}\right)^+ \qquad \text{for } k = 1, \ldots, \kappa \tag{3.87}$$

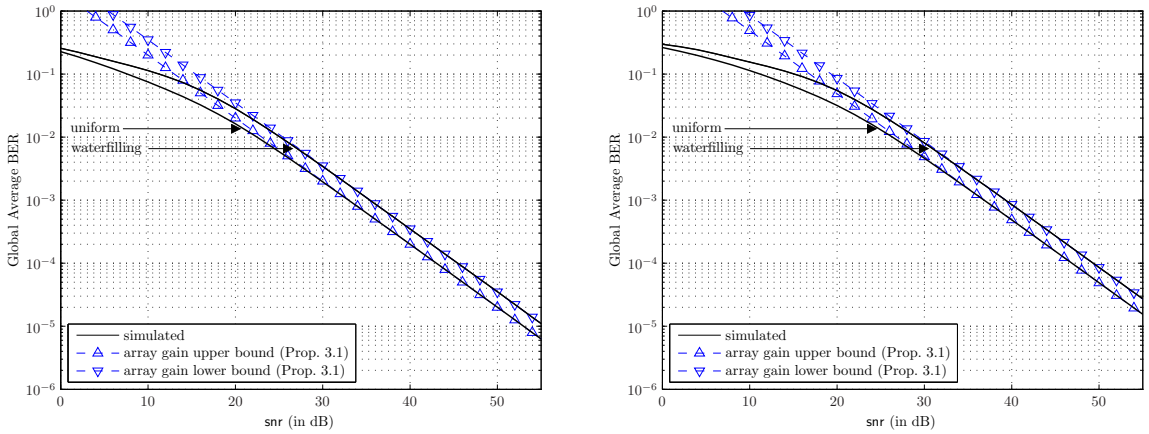
where μ is chosen to satisfy the power constraint in (3.84), $\{\omega_k\}_{k=1}^{\kappa}$ denote the weights, and $a^+ = \max(0, a)$. These criteria include the minimization of the determinant of the MSE matrix, the minimization of the (weighted) product of the MSEs, and the maximization of the mutual information (see Table 3.3). The global average BER performance achieved with the waterfilling in (3.87) can be analyzed combining Theorem 3.7 with the results for non-fixed power allocations given in Theorem 3.2-(ii). In addition, tighter bounds can be obtained as presented in the following result.

Proposition 3.1. *The global average BER of a diagonal MIMO linear transceiver when κ data symbols have to be communicated and the power is allocated in a waterfilling fashion as in* (3.87) *under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models in Definitions* 3.1, 3.2, *and* 3.4 *satisfies*

$$\overline{\mathsf{BER}}(\mathsf{snr}) = (G_{\mathsf{a,wf}} \cdot \mathsf{snr})^{-G_{\mathsf{d}}} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}}\big) \tag{3.88}$$

where the diversity gain is given by

$$G_{\mathsf{d}} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1) \tag{3.89}$$



(i) $n_{\mathsf{T}} = n_{\mathsf{R}} = \kappa = 4$, uncorrelated Rayleigh

(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = \kappa = 4$, semicorrelated Rayleigh

Figure 3.11 Global exact, simulated and parameterized average BER of linear MIMO transceivers in an uncorrelated Rayleigh and a semicorrelated Rayleigh MIMO channel (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$).

and the array gain can be bounded as

$$\left(G_{\mathsf{a}}\left(\{\omega_k\}_{k=1}^{\kappa}\right)^{-G_{\mathsf{d}}} + \left(\frac{\alpha_\kappa}{2\kappa \log_2 M_\kappa}\right)\left(\frac{a_\kappa}{d_\kappa + 1}\right)\sum_{k=1}^{\kappa-1}(\omega_k/\omega_\kappa)^{G_{\mathsf{d}}}\right)^{-1/G_{\mathsf{d}}} < G_{\mathsf{a,wf}} < G_{\mathsf{a}}\left(\{\omega_k\}_{k=1}^{\kappa}\right) \tag{3.90}$$

where $G_{\mathsf{a}}\left(\{\omega_k\}_{k=1}^{\kappa}\right)$ *is the global array gain obtained with a fixed power allocation with* $\phi_k = 1/\sum_{i=1}^{\kappa}(\omega_i/\omega_k)$ *(see* (3.73) *in Theorem* 3.7*). The parameters* a_κ *and* d_κ *model the fading distribution as given in Theorem* 2.4 *and the corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

Proof. See Appendix 3.C.1.

In Figure 3.11 we show the average BER attained by a linear MIMO transceiver with a uniform power allocation over the κ active substreams and with the waterfilling power allocation in (3.87). In particular, we provide the results for a MIMO system with $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, $\kappa = 4$ active substreams with equal QPSK constellations. The simulation results in Figure 3.11 demonstrate how the average BER curve for both power allocation policies is correctly approximated by the results presented in Theorem 3.7 and Proposition 3.1, respectively.

Now let us consider waterfilling power allocations of the type

$$p_{\mathsf{mse},k} = \left(\mu\omega_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1}\right)^+ \qquad \text{for } k = 1, \ldots, \kappa \tag{3.91}$$

where μ is chosen to satisfy the power constraint in (3.84) and $\{\omega_k\}_{k=1}^{\kappa}$ denote the weights.

The power allocation in (3.91) results from minimizing the (weighted) sum of the MSEs (see Table 3.3) and the corresponding global average BER performance is analyzed in the following proposition.

Proposition 3.2. *The global average BER of a diagonal MIMO linear transceiver when κ data symbols have to be communicated and the power is allocated in a waterfilling fashion as in* (3.91) *under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models in Definitions* 3.1, 3.2, *and* 3.4 *satisfies*

$$\overline{\mathsf{BER}}(\mathsf{snr}) = (G_{\mathsf{a,mse}} \cdot \mathsf{snr})^{-G_{\mathsf{d}}} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}}\big) \tag{3.92}$$

where the diversity gain is given by

$$G_{\mathsf{d}} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1) \tag{3.93}$$

and the array gain can be bounded as

$$G_{\mathsf{a}}\big(\{\omega_k^{1/2}\}_{k=1}^{\kappa}\big) \leq G_{\mathsf{a,mse}} < \kappa G_{\mathsf{a}} \tag{3.94}$$

where $G_{\mathsf{a}}\big(\{\omega_k^{1/2}\}_{k=1}^{\kappa}\big)$ and G_{a} are the global array gains obtained with a fixed power allocation with $\phi_k = 1/\sum_{i=1}^{\kappa}(\omega_i/\omega_k)^{1/2}$ and $\phi_k = 1/\kappa$ (uniform), respectively, (see (3.73) *in Theorem* 3.7*).*

Proof. It follows from Theorem 3.7 and Theorem 3.2-(i) with $\xi_\kappa = 1/\sum_{i=1}^{\kappa}(\omega_i/\omega_\kappa)^{1/2}$, since the exponent of $\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} \leq \xi_\kappa \mathsf{snr}\right)$ is greater than G_{d} (see Appendix 3.C.2). □

3.8.4 Performance of Non-Diagonal Schemes with Non-Fixed Power Allocation

In this section we complete the performance analysis of linear MIMO transceivers by focusing on the non-diagonal scheme with the non-fixed power allocation that results for Schur-convex cost functions. Let us consider, for instance, the minimum BER design with equal constellations independently derived in [Din03a] and [Pal03]. Other examples of design criteria with a Schur-convex cost function are the minimization of the maximum MSE, the maximization of the minimum SNR or the minimization of the maximum BER as summarized in Table 3.4.

When the cost function is Schur-convex, the global communication system including pre- and post-processing is diagonalized only up to a rotation of the data symbols (see Figure 3.10-(ii)), which ensures that all substreams have the same MSE, and the optimal power allocation is

Design criterion	Optimal power allocation
Minim. maximum of the MSEs [Pal03]	
Maxim. minimum of the SNRs [Pal03]	
Maxim. harmonic mean of SNRs [Pal03]	$p_k = (\mu\lambda_k^{-1/2} - \lambda_k^{-1})^+$
Minim. average BER [Din03a, Pal03] (equal constellations)	
Minim. maximum of BERs [Pal03]	

Table 3.4 Examples of non-diagonal schemes with non-fixed power allocation [Pal07, Tab. 3.2].

independent of the particular cost function and coincides with the power allocation in (3.91) (see [Pal03, Sec. IV.B]). Due to the rotation of the data symbols, Theorem 3.2 can not be directly applied and the global average BER performance is analyzed in the following proposition.

Proposition 3.3. *The global average BER of the non-diagonal MIMO linear transceiver derived from Schur-convex cost functions when κ data symbols have to be communicated under the $n_{\mathsf{R}}\times n_{\mathsf{T}}$ MIMO channel models in Definitions* 3.1*,* 3.2*, and* 3.4 *satisfies*

$$\overline{\mathsf{BER}}(\mathsf{snr}) = (G_{\mathsf{a,ber}} \cdot \mathsf{snr})^{-G_{\mathsf{d}}} + o\big(\mathsf{snr}^{-G_{\mathsf{d}}}\big) \tag{3.95}$$

where the diversity gain is given by

$$G_{\mathsf{d}} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1) \tag{3.96}$$

and the array gain can be bounded as

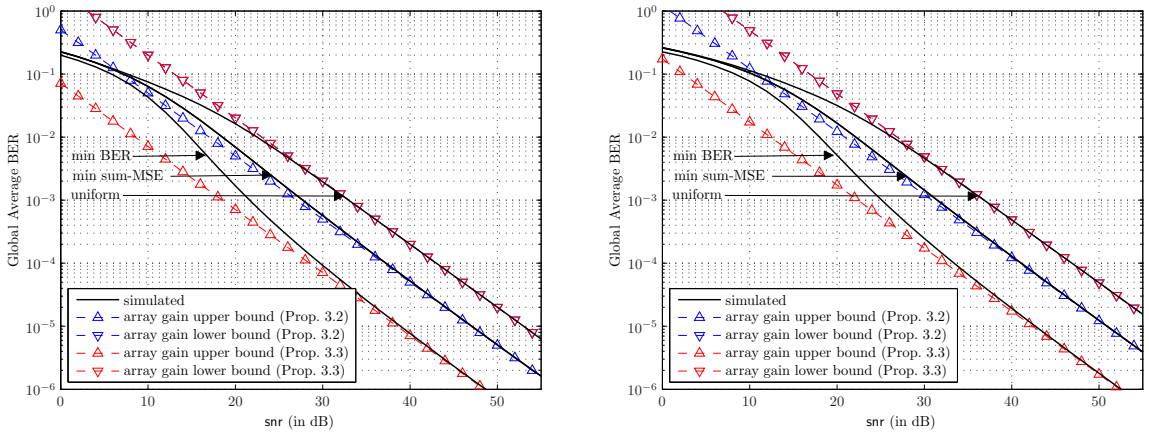
$$G_{\mathsf{a}} < G_{\mathsf{a,ber}} < G_{\mathsf{a}}^{(\mathrm{ub})} \tag{3.97}$$

where G_{a} is the global array gain obtained with a uniform power allocation (see (3.73) *in Theorem* 3.7*) and $G_{\mathsf{a}}^{(\mathrm{ub})}$ is defined as*

$$G_{\mathsf{a}}^{(\mathrm{ub})} = \beta_\kappa \left(\frac{\alpha_\kappa}{\log_2 M_\kappa} \frac{a_\kappa I(d_\kappa, \beta_\kappa(\kappa-1))}{\sqrt{2\pi}(d_\kappa+1)} \right)^{-1/(d_\kappa+1)} \tag{3.98}$$

where $I(\cdot,\cdot)$ is given in (3.21) *and the parameters a_κ and d_κ model the fading distribution as done in Theorem* 2.4 *and the corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

Proof. See Appendix 3.C.3.



(i) $n_{\mathsf{T}} = n_{\mathsf{R}} = \kappa = 4$, uncorrelated Rayleigh

(ii) $n_{\mathsf{T}} = n_{\mathsf{R}} = \kappa = 4$, semicorrelated Rayleigh

Figure 3.12 Global exact, simulated and parameterized average BER of linear MIMO transceivers in an uncorrelated Rayleigh, a semicorrelated Rayleigh MIMO channel (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$).

In summary, Propositions 3.1, 3.2 and 3.3 show that linear MIMO transceivers with non-fixed power allocation policies (with or without additional pre- and post-processing of the data symbols) do not provide any diversity advantage with respect to diagonal schemes with fixed power allocation policies but a possibly higher array gain, which results in non-negligible average performance differences. This statement is confirmed by Figure 3.12, where we show the global performance of linear MIMO transceivers with $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, all substreams active, and all symbols drawn from a QPSK modulation for the following cases: (i) the diagonal scheme with uniform power allocation, (ii) the diagonal scheme with the power allocation that minimizes the sum of the MSEs, and (iii) the non-diagonal scheme obtained for Schur-convex cost functions. Similarly to Figure 3.11, the average BER performance is always measured as the BER averaged over the κ transmitted data symbols even when the corresponding power allocation assigns zero power (or a very small amount of power) to the worst substreams. We also provide the high-SNR average BER parameterized upper and lower bounds derived in Propositions 3.2 and 3.3. It turns out that the proposed array gain upper bounds are in fact very tight and approximate perfectly the high-SNR performance of the diagonal and non-diagonal designs with the corresponding non-fixed power allocation.

3.9 Conclusions and Publications

The main contribution of this chapter is the analytical performance analysis of spatial multiplexing MIMO systems in Rayleigh and Rician MIMO channels. To summarize, in this chapter we obtain analytical expressions to calculate:

(i) the exact individual average BER and its corresponding parameterized high-SNR characterization when using a fixed power allocation,

(ii) the exact individual outage probability and its corresponding parameterized high-SNR characterization when using a fixed power allocation,

(iii) high-SNR parameterized upper and lower bounds for the individual average BER when using a non-fixed power allocation,

(iv) high-SNR parameterized upper and lower bounds for the individual outage probability when using a non-fixed power allocation,

(v) the global average BER and its corresponding parameterized high-SNR characterization, and

(vi) different global outage probability measures.

These general results are then applied to analyze the performance of most linear MIMO transceivers existing in the literature with adaptive linear precoder but fixed number of data symbols and fixed constellations. In particular, using the unifying formulation in [Pal03], we are able to investigate the high-SNR global average performance of a wide family of practical designs:

(vii) diagonal schemes with a fixed power allocation,

(viii) diagonal schemes with a non-fixed power allocation, and

(ix) non-diagonal schemes with a non-fixed power allocation.

The proposed parameterized characterization shows that spatial multiplexing MIMO systems have a diversity order limited by that of the worst eigenmode used, $(n_{\mathsf{T}}-\kappa+1)(n_{\mathsf{R}}-\kappa+1)$, which can be far from the full diversity of $n_{\mathsf{T}}n_{\mathsf{R}}$ provided by the channel. This enlightens that fixing a priori the number of independent data streams to be transmitted, a very common assumption in the linear transceiver design literature, inherently limits the performance of the system. As a consequence, it seems reasonable to optimize the number of substreams jointly with the linear precoder for each channel realization. This design is investigated in Chapter 4.

The main results contained in this chapter regarding the analytical performance analysis of spatial multiplexing MIMO systems with CSI and linear MIMO transceivers have been published in three journal papers and two conference papers:

[Ord05b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Analytical BER performance in spatial multiplexing MIMO systems", *Proc. IEEE Workshop Signal Process. Advan. Wireless Commun. (SPAWC)*, pp. 460–464, June 2005.

[Ord07b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "High SNR analytical performance of spatial multiplexing MIMO systems with CSI", *IEEE Trans. Signal Processing*, vol. 55, no. 11, pp. 5447–5463, Nov. 2007.

[Ord08b] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *Proc. IEEE Int. Conf. Commun. (ICC)*, pag. May, 2008.

[Ord09a] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 672–689, Feb. 2009.

[Ord09b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Minimum BER linear MIMO transceivers with optimum number of substreams", *accepted in IEEE Trans. Signal Processing*, Jan. 2009.

3.A Appendix: Proof of Corollary 3.1

Proof. This proof is strongly based on the proof given in [Wan03] to Lemma 3.1, thus some repetitive parts are omitted. Let ϵ be a small positive number so that $p_\mu(\mu)$ can be approximated by its first order expansion for $0 \leq \mu \leq \epsilon$, then the average BER can be written as

$$\begin{aligned}
\overline{\mathrm{BER}}(\bar{\rho}) &= \frac{\alpha}{\log_2 M}\int_0^{\epsilon} \mathcal{Q}\left(\sqrt{\beta(\bar{\rho}\mu+\phi)}\right) p_\mu(\mu)d\mu + \frac{\alpha}{\log_2 M}\int_{\epsilon}^{\infty} \mathcal{Q}\left(\sqrt{\beta(\bar{\rho}\mu+\phi)}\right) p_\mu(\mu)d\mu \\
&= \frac{\alpha}{\log_2 M}\int_{\epsilon}^{\infty} \mathcal{Q}\left(\sqrt{\beta(\bar{\rho}\mu+\phi)}\right) p_\mu(\mu)d\mu \qquad (3.99)\\
&- \frac{\alpha}{\log_2 M}\frac{1}{\sqrt{2\pi}}\int_{\epsilon}^{\infty}\int_{\sqrt{\beta(\bar{\rho}\mu+\phi)}}^{\infty} e^{-\frac{x^2}{2}}\left(a\mu^d + o(\mu^d)\right)dxd\mu \qquad (3.100)\\
&+ \frac{\alpha}{\log_2 M}\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\int_{\sqrt{\beta(\bar{\rho}\mu+\phi)}}^{\infty} e^{-\frac{x^2}{2}}\left(a\mu^d + o(\mu^d)\right)dxd\mu. \qquad (3.101)
\end{aligned}$$

The terms (3.99) and (3.100) are shown in [Wan03] to be of order $o\big(\bar{\rho}^{-(d+1)}\big)$ and, computing the integral in (3.101) by interchanging the integration order, it follows that

$$\begin{aligned}
\overline{\mathrm{BER}}(\bar{\rho}) &= \frac{\alpha}{\log_2 M}\frac{a}{\sqrt{2\pi}}\int_{\sqrt{\beta\phi}}^{\infty}\int_{0}^{\frac{x^2-\beta\phi}{\beta\bar{\rho}}} e^{-\frac{x^2}{2}}\mu^d d\mu dx + o\big(\bar{\rho}^{-(d+1)}\big)\\
&= \frac{\alpha}{\log_2 M}\frac{a}{\sqrt{2\pi}(d+1)(\beta\bar{\rho})^{d+1}}\int_{\sqrt{\beta\phi}}^{\infty} e^{-\frac{x^2}{2}}\left(x^2-\beta\phi\right)^{(d+1)}dx + o\big(\bar{\rho}^{-(d+1)}\big)\\
&= \Big(\frac{\alpha}{\log_2 M}\frac{aI\left(d,\beta\phi\right)}{\sqrt{2\pi}(d+1)\beta^{d+1}}\Big)\bar{\rho}^{-(d+1)} + o\big(\bar{\rho}^{-(d+1)}\big) \qquad (3.102)
\end{aligned}$$

where $I(d,\beta\phi)$ is defined as

$$I(d,\beta\phi) = \int_{\sqrt{\beta\phi}}^{\infty} e^{-\frac{x^2}{2}}\left(x^2-\beta\phi\right)^{(d+1)}dx \qquad (3.103)$$

and this completes the proof. □

3.B Appendix: Performance of Spatial Multiplexing MIMO Systems with CSI

3.B.1 Proof of Theorem 3.5

Proof. The average BER of the kth substream with $p_k \triangleq p_k(\lambda_1,\ldots,\lambda_\kappa,\mathsf{snr})$ can be expressed as

$$\begin{aligned}
\overline{\mathsf{BER}}_k(\mathsf{snr}) &= \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k\mathsf{snr})\left(1-\mathsf{Pr}\left(p_k < \xi_k\mathsf{snr}\right)\right) \qquad (3.104)\\
&+ \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k < \xi_k\mathsf{snr})\mathsf{Pr}\left(p_k < \xi_k\mathsf{snr}\right) \qquad (3.105)
\end{aligned}$$

where $\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k\mathsf{snr})$ denotes the average BER conditioned on $(p_k \geq \xi_k\mathsf{snr})$ and $\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k < \xi_k\mathsf{snr})$ is analogously defined. Using the expression for $\mathsf{Pr}\left(p_k < \xi_k\mathsf{snr}\right)$ given in

(3.59), the average BER can be rewritten as

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-d(\xi_k)} + o\big(\mathsf{snr}^{-d(\xi_k)}\big)\big) \tag{3.106}$$

$$+ \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k < \xi_k \mathsf{snr})\big(a(\xi_k)\mathsf{snr}^{-d(\xi_k)} + o\big(\mathsf{snr}^{-d(\xi_k)}\big)\big). \tag{3.107}$$

In the following we distinguish between two cases: (i) when $d(\xi_k) > d_k + 1$ for a given $\xi_k \in (0,1)$, and (ii) when $d(\xi_k) = d_k + 1$ for a given $\xi_k \in (0,1)$ and $d(0) = d_k + 1$.

(i) If there exists $0 < \xi_k < 1$ such that $d(\xi_k) > d_k + 1$, the term in (3.107) is of order $o\big(\mathsf{snr}^{-(d_k+1)}\big)$, since it holds that

$$\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k < \xi_k \mathsf{snr}) a(\xi_k)\mathsf{snr}^{-d(\xi_k)} < \frac{\alpha_k}{2\log_2 M_k} a(\xi_k)\mathsf{snr}^{-d(\xi_k)} = o\big(\mathsf{snr}^{-(d_k+1)}\big). \tag{3.108}$$

Using the short-term power constraint in (3.39), the power allocated to the kth substream is upper-bounded as $p_k \leq \mathsf{snr}$ and, hence, $\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k > \xi_k \mathsf{snr})$ satisfies

$$\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \mathsf{snr}) < \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr}) \leq \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \xi_k \mathsf{snr}). \tag{3.109}$$

The upper and lower bounds in (3.109) can be analyzed applying Theorem 3.1. Both result in the same diversity gain $G_{\mathsf{d}}(k) = d_k + 1$ and, thus,

$$\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-d(\xi_k)}\big) = \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big). \tag{3.110}$$

Finally, the average BER is

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) = \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \tag{3.111}$$

and the array gain bounds given in the theorem are obtained by deriving the array gain associated with the bounds in (3.109) recalling Theorem 3.1.

(ii) If $d(0) = d_k + 1$ and there exists $0 < \xi_k < 1$ such that $d(\xi_k) = d_k + 1$, the term in (3.107) can be bounded as

$$\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k < \xi_k \mathsf{snr}) a(\phi_k)\mathsf{snr}^{-(d_k+1)} < \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = 0) a(\xi_k)\mathsf{snr}^{-(d_k+1)} \tag{3.112}$$

$$= \frac{\alpha_k}{2\log_2 M_k} a(\xi_k)\mathsf{snr}^{-(d_k+1)}. \tag{3.113}$$

Then, proceeding as in case (i), it follows that

$$\overline{\mathsf{BER}}_k(\mathsf{snr}) < \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \xi_k \mathsf{snr}) + \frac{\alpha_k}{2\log_2 M_k} a(\xi_k)\mathsf{snr}^{-(d_k+1)} + o\big(\mathsf{snr}^{-(d_k+1)}\big) \tag{3.114}$$

and the array gain lower bound given in the theorem is obtained by combining the array gain corresponding to $\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \xi_k\mathsf{snr})$ with the term in (3.113). Similarly, we can lower-bound the average BER as

$$\begin{aligned}\overline{\mathsf{BER}}_k(\mathsf{snr}) &> \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-(d_k+1)}\big) \\ &\quad + \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \xi_k\mathsf{snr})\big(a(\xi_k)\mathsf{snr}^{-(d_k+1)}\big) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \\ &= \overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \end{aligned} \tag{3.115}$$

and the array gain corresponding to $\overline{\mathsf{BER}}_k(\mathsf{snr}|p_k = \mathsf{snr})$ is the array gain upper bound given in the theorem. □

3.B.2 Proof of Theorem 3.6

Proof. This proof uses similar techniques as the ones used in the proof of Theorem 3.5 in Appendix 3.B.1. Nevertheless, for the sake of completeness, we provide in the following all important steps. The outage probability of the kth substream can be expressed as

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k\mathsf{snr})(1 - \mathsf{Pr}(p_k < \xi_k\mathsf{snr})) \tag{3.116}$$

$$+ \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k < \xi_k\mathsf{snr})\mathsf{Pr}(p_k < \xi_k\mathsf{snr}) \tag{3.117}$$

where $\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \phi_k\mathsf{snr})$ denotes the outage probability conditioned on $p_k \geq \phi_k\mathsf{snr}$ and $\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k < \xi_k\mathsf{snr})$ is analogously defined. Using the expression for $\mathsf{Pr}\left(p_k < \xi_k\mathsf{snr}\right)$ given in (3.59), the outage probability can be rewritten as

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k\mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-d(\xi_k)} + o\big(\mathsf{snr}^{-d(\xi_k)}\big)\big) \tag{3.118}$$

$$+ \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k < \xi_k\mathsf{snr})\big(a(\xi_k)\mathsf{snr}^{-d(\xi_k)} + o\big(\mathsf{snr}^{-d(\xi_k)}\big)\big). \tag{3.119}$$

In the following we distinguish between two cases: (i) when $d(\xi_k) > d_k + 1$ for a given $\xi_k \in (0, 1)$ and (ii) when $d(\xi_k) = d_k + 1$ for a given $\xi_k \in (0, 1)$ and $d(0) = d_k + 1$.

(i) If there exists $0 < \xi_k < 1$ such that $d(\phi_k) > d_k+1$, the term in (3.119) is of order $o\big(\mathsf{snr}^{-(d_k+1)}\big)$, since it holds that

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k < \xi_k\mathsf{snr})a(\xi_k)\mathsf{snr}^{-d(\xi_k)} < a(\xi_k)\mathsf{snr}^{-d(\xi_k)} = o\big(\mathsf{snr}^{-(d_k+1)}\big). \tag{3.120}$$

Using the short-term power constraint in (3.59), the power allocated to the kth substream is upper-bounded as $p_k \leq \mathsf{snr}$ and, hence, $\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}, \rho_{\mathrm{th}}|p_k > \phi_k\mathsf{snr})$ satisfies

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \mathsf{snr}) < \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k\mathsf{snr}) \leq \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \xi_k\mathsf{snr}). \tag{3.121}$$

The upper and lower bound in (3.121) can be analyzed applying Theorem 3.3. Both result in the same diversity gain, $G_{\mathsf{d}}(k) = d_k + 1$ and

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-d(\xi_k)}\big) = \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big). \quad (3.122)$$

Finally, the outage probability is

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) = \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k \geq \xi_k \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \quad (3.123)$$

and the outage array gain bounds given in the theorem are obtained by deriving the outage array gain associated with the bounds in (3.121) recalling Theorem 3.3.

(ii) If $d(0) = d_k + 1$ and there exists $0 < \xi_k < 1$ such that $d(\phi_k) = d_k + 1$, the term in (3.119) can be bounded as

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k < \xi_k \mathsf{snr}) a(\xi_k)\mathsf{snr}^{-(d_k+1)} < \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = 0) a(\xi_k)\mathsf{snr}^{-(d_k+1)} \quad (3.124)$$

$$= a(\xi_k)\mathsf{snr}^{-(d_k+1)}. \quad (3.125)$$

Then, proceeding as in case (i), it follows that

$$\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) < \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \xi_k \mathsf{snr}) + a(\xi_k)\mathsf{snr}^{-(d_k+1)} + o\big(\mathsf{snr}^{-(d_k+1)}\big) \quad (3.126)$$

and the outage array gain lower bound given in the theorem is obtained by combining the outage array gain corresponding to $\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \xi_k \mathsf{snr})$ with the term in (3.124). Similarly, we can lower-bound the outage probability as

$$\begin{aligned}\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}) &> \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \mathsf{snr})\big(1 - a(\xi_k)\mathsf{snr}^{-(d_k+1)}\big) \\ &\quad + \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \xi_k \mathsf{snr})\big(a(\xi_k)\mathsf{snr}^{-(d_k+1)}\big) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \\ &= \mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \mathsf{snr}) + o\big(\mathsf{snr}^{-(d_k+1)}\big) \end{aligned} \quad (3.127)$$

and the array gain corresponding to $\mathsf{P}_{\mathsf{out},k}(\mathsf{snr}|p_k = \mathsf{snr})$ is the outage array gain upper bound given in the theorem. □

3.C Appendix: Performance of Linear MIMO Transceivers

3.C.1 Proof of Proposition 3.1

This proof is in essence the same as the proof of Theorem 3.2-(ii) in Appendix 3.B.1, but here we use an asymptotic equivalence of the power allocation, instead of bounding it. Observe that the

waterfilling in (3.87) tends to a fixed power allocation over the κ active substreams as $\mathsf{snr} \to \infty$, i.e.,

$$\lim_{\mathsf{snr}\to\infty} \frac{p_{\mathsf{wf},k}}{\mathsf{snr}/\sum_{i=1}^{\kappa}(\omega_i/\omega_k)} = 1 \tag{3.128}$$

whenever $p_{\mathsf{wf},k} > 0$. Hence, the average BER at high SNR (see Section 3.7.2) can be expressed as

$$\overline{\mathsf{BER}}(\mathsf{snr}) = \frac{1}{\kappa}\overline{\mathsf{BER}}_{\kappa}(\mathsf{snr})\left(1 - \mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right)\right) + \frac{\alpha_{\kappa}}{2\kappa\log_2 M_{\kappa}}\mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right) \tag{3.129}$$

where $\overline{\mathsf{BER}}_{\kappa}(\mathsf{snr})$ is the individual average BER of the substream transmitted through the κth eigenmode with fixed power $p_{\kappa} = \mathsf{snr}/\sum_{k=1}^{\kappa}(\omega_k/\omega_{\kappa})$ and $\mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right)$ denotes the probability of not allocating power to the κth substream. This probability is upper-bounded as

$$\begin{aligned}
\mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right) &= \mathsf{Pr}\left(\left(\sum_{k=1}^{\kappa-1}(\omega_k/\omega_{\kappa})\lambda_{\kappa}^{-1} - \sum_{k=1}^{\kappa-1}\lambda_k^{-1}\right) \geq \mathsf{snr}\right) \\
&\leq \mathsf{Pr}\left(\lambda_{\kappa}^{-1} \geq \frac{\mathsf{snr}}{\sum_{k=1}^{\kappa-1}(\omega_k/\omega_{\kappa})}\right) \\
&= \mathsf{Pr}\left(\lambda_{\kappa} \leq \frac{\sum_{k=1}^{\kappa-1}(\omega_k/\omega_{\kappa})}{\mathsf{snr}}\right) \qquad (3.130) \\
&= \frac{a_{\kappa}}{d_{\kappa}+1}\left(\frac{\mathsf{snr}}{\sum_{k=1}^{\kappa-1}(\omega_k/\omega_{\kappa})}\right)^{-(d_{\kappa}+1)} + o\left(\mathsf{snr}^{-(d_{\kappa}+1)}\right) \qquad (3.131)
\end{aligned}$$

where the last equality comes from the first order Taylor expansion of the marginal cdf of the κth strongest eigenvalue in Theorem 2.4. Then, substituting $\overline{\mathsf{BER}}_{\kappa}(\mathsf{snr})$ by its parameterized characterization applying Theorem 3.1 with $\phi_k = 1/\sum_{i=1}^{\kappa}(\omega_i/\omega_k)$ and $\mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right)$ by its upper bound derived in (3.131) back in (3.129), it follows that

$$\begin{aligned}
\overline{\mathsf{BER}}(\mathsf{snr}) < &\left(G_{\mathsf{a}}{}^{-(d_{\kappa}+1)} + \left(\frac{\alpha_{\kappa}}{2\kappa\log_2 M_{\kappa}}\right)\left(\frac{a_{\kappa}}{d_{\kappa}+1}\right)\left(\sum_{k=1}^{\kappa-1}(\omega_k/\omega_{\kappa})\right)^{(d_{\kappa}+1)}\right)\mathsf{snr}^{-(d_{\kappa}+1)} \\
&+ o\left(\mathsf{snr}^{-(d_{\kappa}+1)}\right)
\end{aligned} \tag{3.132}$$

where G_{a} denotes the global array gain achieved with a fixed power allocation with $\phi_k = 1/\sum_{i=1}^{\kappa}(\omega_i/\omega_k)$. Finally, the array gain lower bound can be directly obtained from (3.132) and the upper bound simply comes from setting $\mathsf{Pr}\left(p_{\mathsf{wf},\kappa} = 0\right) = 0$ in (3.129). □

3.C.2 Proof of Proposition 3.2: Exponent of $\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} \leq \xi_{\kappa}\mathsf{snr}\right)$

We want to prove that the power allocation in (3.91) satisfies the condition in Theorem 3.1-(ii), i.e.,

$$\delta_{\kappa} = \lim_{\mathsf{snr}\to\infty} -\frac{\log \mathsf{Pr}\left(p_{\mathsf{mse},\kappa} < \xi_{\kappa}\mathsf{snr}\right)}{\log \mathsf{snr}} > G_{\mathsf{d},\mathsf{mse}} = d_{\kappa} + 1 \tag{3.133}$$

with ξ_κ smaller but arbitrarily close to $1/\sum_{k=1}^{\kappa}(\omega_k/\omega_\kappa)^{1/2}$. The probability $\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} < \xi_\kappa \mathsf{snr}\right)$ is given by

$$\begin{aligned}\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} \le \xi_\kappa \mathsf{snr}\right) &= \mathsf{Pr}\left(\frac{(\omega_\kappa/\lambda_\kappa)^{1/2}}{\sum_{k=1}^{\kappa}(\omega_k/\lambda_k)^{1/2}}\left(\mathsf{snr}+\sum_{k=1}^{\kappa}\lambda_k^{-1}\right)-\lambda_\kappa^{-1} \le \xi_\kappa \mathsf{snr}\right)\\ &= \mathsf{Pr}\Bigg(\Bigg((\omega_\kappa\lambda_\kappa)^{-1/2}\sum_{k=1}^{\kappa}(\omega_k/\lambda_k)^{-1/2}-\sum_{k=1}^{\kappa}\lambda_k^{-1}\Bigg)\\ &\qquad \ge \mathsf{snr}\left(1-\xi_k\frac{(\omega_\kappa/\lambda_\kappa)^{1/2}}{\sum_{k=1}^{\kappa}(\omega_k/\lambda_k)^{1/2}}\right)\Bigg) \qquad (3.134)\end{aligned}$$

Observing that

$$\xi_\kappa(\omega_\kappa/\lambda_\kappa)^{-1/2}\sum_{k=1}^{\kappa}(\omega_k/\lambda_k)^{1/2} \le \xi_\kappa\sum_{k=1}^{\kappa}(\omega_k/\omega_\kappa)^{1/2} \qquad (3.135)$$

since $\lambda_\kappa/\lambda_k \le 1$ for $k = 1, \cdots, \kappa$, $\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} \le \xi_\kappa \mathsf{snr}\right)$ can be upper-bounded as

$$\begin{aligned}\mathsf{Pr}\left(p_{\mathsf{mse},\kappa} \le \xi_k \mathsf{snr}\right) &\le \mathsf{Pr}\left((\omega_\kappa\lambda_\kappa)^{-1/2}\sum_{k=1}^{\kappa-1}(\omega_k/\lambda_k)^{1/2} \ge \mathsf{snr}\left(1-\xi_\kappa\sum_{k=1}^{\kappa}(\omega_k/\omega_\kappa)^{1/2}\right)\right)\\ &\le \mathsf{Pr}\left(\lambda_\kappa\lambda_{\kappa-1} \le \left(\frac{\sum_{k=1}^{\kappa-1}(\omega_k/\omega_\kappa)^{1/2}}{1-\xi_\kappa\sum_{k=1}^{\kappa}(\omega_k/\omega_\kappa)^{1/2}}\right)^2 \mathsf{snr}^{-2}\right) \qquad (3.136)\end{aligned}$$

whenever $\xi_\kappa < 1/\sum_{k=1}^{\kappa}(\omega_k/\omega_\kappa)^{1/2}$. Hence, it follows that

$$\delta_\kappa \ge \lim_{\mathsf{snr}\to\infty}\frac{\log \mathsf{Pr}\left(\lambda_\kappa\lambda_{\kappa-1} \le \mathsf{snr}^{-2}\right)}{\log \mathsf{snr}^{-1}}. \qquad (3.137)$$

Then, by defining $\eta_\kappa = \lambda_{\kappa-1}\lambda_\kappa$, we can rewrite (3.137) as

$$\delta_\kappa \ge \lim_{x\to 0}\frac{\log \mathsf{Pr}\left(\eta_\kappa \le x^2\right)}{\log x} = \lim_{x\to 0}\frac{\log\left(\int_0^{x^2} f_{\eta_\kappa}(\eta)d\eta\right)}{\log x} \qquad (3.138)$$

where $f_{\eta_\kappa}(\eta)$ is the pdf of a product of two random variables and can be calculated as [Roh76, Sec. 4.4, Th. 7]

$$f_{\eta_\kappa}(\eta) = \int_0^\infty \frac{1}{x} f_{\lambda_{\kappa-1},\lambda_\kappa}(x, \eta/x)dx. \qquad (3.139)$$

We are interested in $f_{\eta_\kappa}(\eta)$ as $\eta \to 0$ and, thus, we only need to derive the joint pdf $f_{\lambda_{\kappa-1},\lambda_\kappa}(\lambda_{\kappa-1}, \lambda_\kappa)$ as $\lambda_\kappa \to 0$. Using the same procedure as in the proof of Theorem 2.4, it can be shown that

$$f_{\lambda_{\kappa-1},\lambda_\kappa}(\lambda_{\kappa-1}, \lambda_\kappa) = g(\lambda_{\kappa-1})\lambda_\kappa^{d_\kappa} + o(\lambda_\kappa^{d_\kappa}) \qquad (3.140)$$

where $g(\lambda_{\kappa-1})$ is a function of $\lambda_{\kappa-1}$. Then, substituting back this result in the expression of $f_{\eta_\kappa}(\eta)$ in (3.139), it follows that

$$f_{\eta_\kappa}(\eta) = a\eta^{d_\kappa} + o(\eta^{d_\kappa}) \qquad (3.141)$$

where a is a fixed constant in terms of η. Finally, the exponent of $\mathsf{Pr}\left(p_{\mathsf{mse},k} < \xi_\kappa \mathsf{snr}\right)$ can be bounded as

$$\delta_\kappa \geq \lim_{x\to 0} \frac{\log\left(a\int_0^{x^2} \eta^{d_\kappa} d\eta\right)}{\log x} = \lim_{\eta\to 0} \frac{\log\left(\eta^{2d_\kappa+2}\right)}{\log \eta} = 2d_\kappa + 2 \tag{3.142}$$

and this proves (3.133). □

3.C.3 Proof of Proposition 3.3

The minimum BER scheme distributes the available power over the κ active substreams in a waterfilling fashion as in (3.91). Since the exponent of the probability of not allocating power to the κth substream is greater than G_{d} (see Appendix 3.C.2 for details), we can just focus in the case when $\{p_k > 0\}_{k=1}^{\kappa}$.

First we lower-bound the instantaneous SNR of the minimum BER design using the uniform power allocation, because it leads to a higher sum MSE than the power allocation in (3.91):

$$\rho_{\mathsf{ber},\kappa} > \left(\frac{1}{\kappa}\sum_{k=1}^{\kappa} \frac{1}{(\lambda_k/\kappa)\mathsf{snr}+1}\right)^{-1} - 1 \tag{3.143}$$

$$> (\lambda_\kappa/\kappa)\mathsf{snr} \tag{3.144}$$

where in (3.143) we have forced all substreams to experience the same MSE (see Section 3.8.1) and (3.144) follows from lower-bounding each λ_k by λ_κ. The lower bound in (3.144) corresponds to the instantaneous SNR achieved by the κth substream of a diagonal scheme with a uniform power allocation and, hence, we can lower-bound the array gain by G_{a} as in Propositions 3.1 and 3.2.

Let us now consider a non-diagonal linear MIMO transceiver that allocates infinite power to all the substreams except to the κth one, to which it assigns $p_\kappa = \mathsf{snr}$. Due to the power constraint in (3.84), the instantaneous SNR of the BER minimizing design can be upper-bounded by the instantaneous SNR of this scheme:

$$\rho_{\mathsf{ber},\kappa} < \left(\frac{1}{\kappa\lambda_\kappa \mathsf{snr}+\kappa}\right)^{-1} - 1 = \kappa\lambda_\kappa \mathsf{snr} + (\kappa - 1). \tag{3.145}$$

Finally, using Corollary 3.1 with the upper bound in (3.145), it follows the array gain upper bound given in Proposition 3.3. □

4

Spatial Multiplexing MIMO Systems with CSI: Optimum Number of Substreams

Multiple-input multiple-output (MIMO) systems with perfect channel state information at both sides of the link can adapt to the instantaneous channel conditions to optimize the spectral efficiency and/or the reliability of the communication. A low-complexity approach is the use of linear MIMO transceivers, which are composed of a linear precoder at the transmitter and a linear equalizer at the receiver. The design of linear transceivers has been extensively studied in the literature with a variety of cost functions. In this chapter we focus on the minimum bit error rate (BER) design, and show that the common practice of fixing a priori the number of data symbols to be transmitted per channel use inherently limits the diversity gain of the system. By introducing the number of symbols in the optimization process, we propose a minimum BER linear precoding scheme that achieves the full diversity of the MIMO channel. For the cases of uncorrelated Rayleigh, semicorrelated Rayleigh, and uncorrelated Rician fading, the average BER performance of both schemes is analytically investigated and characterized in terms of two key parameters: the array gain and the diversity gain.

4.1 Introduction

One of the salient and unique characteristics of multiple-input multiple-output (MIMO) channels is the multiplexing gain, which refers to the increase of rate at no additional power consumption. The multiple dimensions of the MIMO channel are exploited to open up several parallel subchannels which allow the transmission of several symbols simultaneously. When perfect channel state information (CSI) is available at the transmitter, the data signal is adapted to the instantaneous channel eigenstructure by transmitting the established substreams through the strongest channel eigenmodes. These schemes are commonly known as spatial multiplexing MIMO systems with CSI (see Chapter 3 for details) and are practically implemented by using linear MIMO transceivers, which are composed of a linear precoder at the transmitter and a linear equalizer at the receiver.

The design of linear transceivers when perfect CSI is available at both sides of the link has been extensively studied in the literature according to a variety of criteria based on performance measures such as the signal-to-noise ratio (SNR), the mean square error (MSE), or the bit error rate (BER). The most common approach in the linear transceiver design literature is to adapt only the linear precoder/power allocation among the different substreams, assuming that the number of substreams κ and the corresponding constellations are fixed beforehand, e.g. [Lee76, Sal85, Yan94b, Yan94a, Sca99, Sam01, Sca02b, Ong03, Pal03, Din03a, Pal07].

In this chapter we show analytically for the cases of uncorrelated Rayleigh, semicorrelated Rayleigh and uncorrelated Rician fading that the diversity gain of these schemes with κ substreams is at most given by $(n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1)$, where n_{T} is the number of transmit and n_{R} the number of receive antennas. This diversity order can be far from the inherent diversity provided by the MIMO channel $n_{\mathsf{T}} n_{\mathsf{R}}$ [And00a]. To overcome this limitation, we consider the introduction of the number of active substreams in the design criterion by fixing the global rate but allowing the use of an adaptive symbol constellation to compensate for the change in the number of active substreams. Observe that this alternative restriction does not render ineffective the schemes available in the literature but simply implies an additional optimization stage on top of the classical design. The given diversity analysis confirms the intuitive notion that the full diversity $n_{\mathsf{T}} n_{\mathsf{R}}$ demands this final optimization step that has been commonly neglected in the literature. An exception is [Lov05b], where a similar scheme that adapts the number of

substreams under a fixed rate constraint was proposed in the context of limited feedback linear precoding.

More specifically, this chapter focuses on the minimum BER linear MIMO transceiver design. We first present the conventional design in which the linear transmitter and receiver are designed to minimize the BER under a transmit power constraint and assuming that the number of transmitted symbols and constellations are fixed. This scheme is denoted hereafter as minBER-fixed scheme and was derived independently in [Pal03] and [Din03a]. Suboptimal results, in which the linear precoder and receiver are designed under the same constraints but forced to diagonalize the channel can be found in [Ong03, Din03b]. The substream optimization stage is considered next, i.e., the linear MIMO transceiver design with adaptive number of substreams, denoted as minBER-adap scheme. Although the symbol constellation is jointly adapted with the number of substreams to keep the total transmission rate fixed, the design problem addressed here is substantially different from the classical problem formulation in the adaptive modulation literature [Chu01, Cat02, Zha03, Zho05], where, typically, the transmission rate is maximized under power and quality-of-service (QoS) constraints or the transmission power is minimized under QoS and rate constraints. For instance, in the context of MIMO linear transceivers, the design of both the constellations and the linear transceiver to minimize the transmit power under QoS constraints (given in terms of BER) is addressed in [Pal05a].

In addition, in this chapter we investigate the global performance of the minBER-fixed and the minBER-adap design in terms of the BER averaged over all transmitted data substreams and all possible channel states, when assuming an uncorrelated, a semicorrelated Rayleigh, and an uncorrelated Rician fading channel. We derive upper and lower bounds for the average BER performance that can be efficiently computed without resorting to time-consuming Monte Carlo simulations. In order to provide more insights into the system behavior, we also focus on the high-SNR regime and characterize the average BER vs. SNR curves in terms of two key parameters: the diversity gain and the array gain.

The rest of the chapter is organized as follows. Section 4.2 is devoted to introducing the system model and the design problem formulation. In addition, the average BER performance measure and the adopted channel models are presented. The design and performance analysis of the minimum BER linear transceiver with fixed and with adaptive constellations is addressed

in Sections 4.3 and 4.4, respectively. Finally, in Section 4.5 we summarize the main results and provide the list of publications where they have been presented.

4.2 Preliminaries

In this section we introduce the general signal model corresponding to linear MIMO transceivers and formulate the minimum BER design problem. Additionally, we briefly describe the procedure followed to analyze the performance of these schemes in fading channels.

4.2.1 System Model

The signal model corresponding to a transmission through a general MIMO channel with n_{T} transmit and n_{R} receive antennas is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{4.1}$$

where $\mathbf{x} \in \mathbb{C}^{n_{\mathsf{T}}}$ is the transmitted vector, $\mathbf{H} \in \mathbb{C}^{n_{\mathsf{R}} \times n_{\mathsf{T}}}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^{n_{\mathsf{R}}}$ is the received vector, and $\mathbf{w} \in \mathbb{C}^{n_{\mathsf{R}}}$ is a spatially white zero-mean circularly symmetric complex Gaussian noise vector normalized so that $\mathrm{E}\{\mathbf{w}\mathbf{w}^{\dagger}\} = \mathbf{I}_{n_{\mathsf{R}}}$.

Suppose that the MIMO communication system is equipped with a linear transceiver (see Figure 4.1), then the transmitted vector is given by

$$\mathbf{x} = \mathbf{B}_{\kappa}\mathbf{s}_{\kappa} \tag{4.2}$$

where $\mathbf{B}_{\kappa} \in \mathbb{C}^{n_{\mathsf{T}} \times \kappa}$ is the transmit matrix (linear precoder) and the data vector $\mathbf{s}_{\kappa} \in \mathbb{C}^{\kappa}$ gathers the $\kappa \leq \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$ data symbols to be transmitted (zero-mean, unit-energy, and uncorrelated, i.e., $\mathrm{E}\{\mathbf{s}_{\kappa}\mathbf{s}_{\kappa}^{\dagger}\} = \mathbf{I}_{\kappa}$). We consider a fixed-rate data transmission and, hence, each data symbol[1] $s_{k,\kappa}$ is drawn from an $M_{k,\kappa}$-dimensional constellation such that the total transmission rate

$$\mathsf{R} = \sum_{k=1}^{\kappa} \log_2 M_{k,\kappa} \tag{4.3}$$

is fixed for all channel realizations. The transmitted power is constrained such that

$$\mathrm{E}\left\{\|\mathbf{x}\|^2\right\} = \mathrm{tr}\{\mathbf{B}_{\kappa}\mathbf{B}_{\kappa}^{\dagger}\} \leq \mathsf{snr} \tag{4.4}$$

[1] Observe the slight change of notation with respect to Chapter 3. Now we use the subscript κ to emphasize that the number of data symbols to be transmitted is fixed to κ.

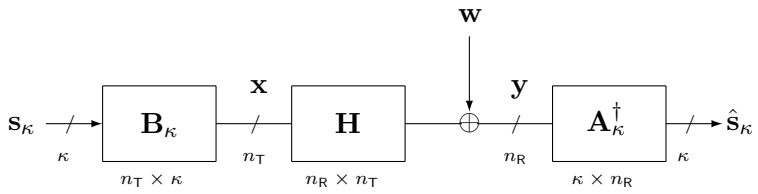


Figure 4.1 Linear MIMO transceivers system model.

where snr denotes the average SNR per receive antenna. The estimated data vector at the receiver is

$$\hat{\mathbf{s}}_\kappa = \mathbf{A}_\kappa^\dagger \mathbf{y} = \mathbf{A}_\kappa^\dagger \left(\mathbf{H}\mathbf{B}_\kappa \mathbf{s}_\kappa + \mathbf{w}\right) \tag{4.5}$$

where $\mathbf{A}_\kappa^\dagger \in \mathbb{C}^{\kappa \times n_\mathsf{R}}$ is the receive matrix (linear equalizer). Observe from (4.5) that κ data streams are established for communication over the MIMO channel, where the kth column of $\mathbf{B}_\kappa$ and $\mathbf{A}_\kappa$, denoted by $\mathbf{b}_{k,\kappa}$ and $\mathbf{a}_{k,\kappa}$, respectively, can be interpreted as the transmit and receive beamvectors associated with the kth data stream or symbol $s_{k,\kappa}$:

$$\hat{s}_{k,\kappa} = \mathbf{a}_{k,\kappa}^\dagger (\mathbf{H}\mathbf{b}_{k,\kappa} s_{k,\kappa} + \mathbf{n}_{k,\kappa}) \qquad \text{for } k = 1, \ldots, \kappa \tag{4.6}$$

where $\mathbf{n}_{k,\kappa} = \sum_{i=1, i\neq k}^{\kappa} \mathbf{H}\mathbf{b}_{i,\kappa} s_{i,\kappa} + \mathbf{w}$ is the interference-plus-noise seen at the kth substream.

4.2.2 Problem Statement

The linear MIMO transceiver design, i.e., the joint design of the receive and transmit matrices $(\mathbf{A}_\kappa, \mathbf{B}_\kappa)$ when perfect CSI is available at both sides of the link, is generally quite involved since several substreams are established over the MIMO channel (see (4.6)). Precisely, the existence of several substreams, each with its own performance, makes the definition of a global measure of the system not clear. As a consequence, a span of different design criteria has been explored in the literature since the 1970s, based on the MSEs, the SNRs, and the BERs. A general unifying framework that embraces most of these design criteria was proposed in [Pal03] (see an up-to-date overview in [Pal07]). The general design problem is formulated as the minimization of certain cost function $f_0(\cdot)$ of the MSEs, $\{\mathsf{mse}_{k,\kappa}\}_{k=1}^{\kappa}$:

$$(\mathbf{A}_\kappa, \mathbf{B}_\kappa) = \arg \min_{\mathbf{A}_\kappa, \mathbf{B}_\kappa} f_0 \left(\{\mathsf{mse}_{k,\kappa}\}_{k=1}^{\kappa}\right) \tag{4.7}$$

subject to the power constraint in (4.4), since any reasonable performance measure can be easily related to the MSEs.

Some examples of the design criteria included in the previous general formulation are given in Tables 3.2, 3.3, and 3.4. In this chapter, we focus only on the design that minimizes the BER averaged over the κ data symbols:

$$\mathsf{BER}_{\kappa}\left(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}\right) = \frac{1}{\kappa}\sum_{k=1}^{\kappa}\mathsf{BER}_{k,\kappa}(\rho_{k,\kappa}) \tag{4.8}$$

where $\rho_{k,\kappa}$ is the instantaneous SNR of the kth substream in (4.6) given by

$$\rho_{k,\kappa} = \frac{|\mathbf{a}_{k,\kappa}^{\dagger}\mathbf{H}\mathbf{b}_{k,\kappa}|^2}{\mathbf{a}_{k,\kappa}^{\dagger}\mathbf{a}_{k,\kappa}} \qquad \text{for } k = 1,\ldots,\kappa \tag{4.9}$$

and $\mathsf{BER}_{k,\kappa}(\rho_{k,\kappa})$ is the corresponding instantaneous BER. In the presence of additive white Gaussian noise and assuming a Gray coding mapping, $\mathsf{BER}_{k,\kappa}(\rho_{k,\kappa})$ can be approximated as (see details in Section 3.2)

$$\mathsf{BER}_{k,\kappa}(\rho_{k,\kappa}) = \frac{\alpha_{k,\kappa}}{\log_2 M_{k,\kappa}}\mathcal{Q}\left(\sqrt{\beta_{k,\kappa}\rho_{k,\kappa}}\right) \qquad \text{for } k = 1,\ldots,\kappa \tag{4.10}$$

where $\mathcal{Q}(\cdot)$ is the Gaussian $\mathcal{Q}$-function defined in (3.2) and the parameters $\alpha_{k,\kappa}$ and $\beta_{k,\kappa}$ depend on the $M_{k,\kappa}$-dimensional modulation used to map the source bits to symbols (see the expressions for the most common digital modulation formats in Table 3.1).

To summarize, the problem studied in this chapter is

$$\begin{aligned} \text{minimize} \quad & \mathsf{BER}_{\kappa}\left(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}\right) \\ \text{subject to} \quad & \mathrm{tr}\left(\mathbf{B}_{\kappa}\mathbf{B}_{\kappa}\right) \leq \mathsf{snr} \end{aligned} \tag{4.11}$$

in the following two cases:

(i) minBER-fixed design: fixed constellations (for some given rate R) and fixed number of substreams κ with optimization variables $(\mathbf{A}_{\kappa}, \mathbf{B}_{\kappa})$.

(ii) minBER-adap design: fixed rate R with optimization variables $(\kappa, \mathbf{A}_{\kappa}, \mathbf{B}_{\kappa})$.

4.2.3 Performance Evaluation

For fading channels, the instantaneous BER defined in (4.8) does not offer representative information about the overall system performance and all different realizations of the random channel have to be taken into account, leading to the concept of average BER:

$$\overline{\mathsf{BER}}_{\kappa}(\mathsf{snr}) \triangleq \mathbb{E}\{\mathsf{BER}_{\kappa}\left(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}\right)\}. \tag{4.12}$$

Given the limited availability of closed-form expressions for the average BER in (4.12), a convenient method to find simple performance measures is to allow a certain degree of approximation. In this respect, the most common approach is to shift the focus from exact performance to large SNR performance as done in [Wan03], where the average BER versus SNR curve is characterized in terms of two key parameters: the diversity gain and the coding gain (also known as the array gain in the context of multiantenna systems [And00a]). The diversity gain represents the slope of the BER curve at high SNR and the coding gain (or array gain) determines the horizontal shift of the BER curve. Interestingly, both parameters only depend on the channel statistics through the first order expansion of the pdf of the channel parameter (see details in Section 3.2.4).

In this chapter we analyze the average BER of both the minBER-fixed and the minBER-adap linear transceivers. More exactly, we find upper and lower bounds on the performance in uncorrelated/semicorrelated Rayleigh and uncorrelated Rician fading MIMO channels, with special emphasis on the high-SNR regime. The outage performance measures introduced in Chapter 3 (see Sections 3.2.3 and 3.2.4 for details) could be also analogously analyzed.

4.2.4 MIMO Channel Model

When analyzing the performance of a communication system over a MIMO fading channel, it is necessary to assume a certain channel fading distribution in order to obtain the average BER measure introduced in Section 4.2.3. In wireless communications, the large number of scatters in the channel that contributes to the signal at the receiver results in Gaussian distributed channel matrix coefficients. Analogously to the single antenna channel, this model is referred to as MIMO Rayleigh or Rician fading channel, depending whether the channel entries are zero-mean or not. In this chapter we adopt the uncorrelated Rayleigh, the semicorrelated Rayleigh, and the uncorrelated Rician MIMO channel models introduced in Definitions 3.1–3.4 (see Section 3.3 for details). Similarly to Chapter 3 we rely on the unified probabilistic characterization of the ordered eigenvalues of a general class of Hermitian random matrices presented in Chapter 2, since this class includes the distribution of the considered channel models.

4.3 MinBER Linear MIMO Transceiver with Fixed Number of Substreams

The linear MIMO transceiver design that minimizes the BER has been addressed in [Din03a] and [Pal03] when equal constellations are used on all substreams. In this section we present this minimum BER linear transceiver and analyze its average BER performance under the Rayleigh and Rician MIMO channel models introduced in Definitions 3.1–3.4.

4.3.1 Linear Transceiver Design

Following the approach in [Pal03], the optimum receive matrix $\mathbf{A}_\kappa$, for a given transmit matrix $\mathbf{B}_\kappa$, is the Wiener filter solution [Pal03, eq. (7)]:

$$\mathbf{A}_\kappa = \left(\mathbf{H}\mathbf{B}_\kappa\mathbf{B}_\kappa^\dagger\mathbf{H}^\dagger + \mathbf{I}_{n_\mathsf{R}}\right)^{-1}\mathbf{H}\mathbf{B}_\kappa \tag{4.13}$$

independently of the design cost function. Specifically, under the minimum BER design criterion when equal constellations are used in all κ substreams, the transmit matrix $\mathbf{B}_\kappa$ is given by [Pal03, Sec. V.C and eq. (15)]

$$\mathbf{B}_\kappa = \mathbf{U}_\kappa\sqrt{\mathbf{P}_\kappa}\mathbf{Q}_\kappa \tag{4.14}$$

where $\mathbf{U}_\kappa \in \mathbb{C}^{n_\mathsf{T}\times\kappa}$ has as columns the eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$ corresponding to the κ largest nonzero eigenvalues $\lambda_1 \geq \cdots \geq \lambda_\kappa$, $\mathbf{Q}_\kappa \in \mathbb{C}^{\kappa\times\kappa}$ is a unitary matrix such that $\left(\mathbf{I}_\kappa + \mathbf{B}_\kappa^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{B}_\kappa\right)^{-1}$ has identical diagonal elements (see [Pal03, Sec. IV.B] for details), and $\mathbf{P}_\kappa \in \mathbb{C}^{\kappa\times\kappa}$ is a diagonal matrix with diagonal entries equal to

$$p_{k,\kappa} = \left(\mu\lambda_k^{-1/2} - \lambda_k^{-1}\right)^+ \qquad \text{for } k = 1,\ldots,\kappa \tag{4.15}$$

where μ is chosen to satisfy the power constraint in (4.4) with equality, i.e.,

$$\sum_{k=1}^{\kappa} p_{k,\kappa} = \mathsf{snr}. \tag{4.16}$$

4.3.2 Analytical Performance

Given the optimum receive matrix in (4.13) and the optimum transmit matrix in (4.14), the communication process is diagonalized up to a specific rotation (see Figure 3.10-(ii)) that forces all κ data symbols to have the same MSE:

$$\mathsf{mse}_\kappa \triangleq \mathsf{mse}_{k,\kappa} = \frac{1}{\kappa}\sum_{i=1}^{\kappa}(1 + p_{i,\kappa}\lambda_i)^{-1} \tag{4.17}$$

and, hence, the same instantaneous SNR:

$$\rho_\kappa \triangleq \rho_{k,\kappa} = \mathsf{mse}_\kappa^{-1} - 1 = \left(\frac{1}{\kappa} \sum_{i=1}^{\kappa} (1 + p_{i,\kappa}\lambda_i)^{-1} \right)^{-1} - 1. \tag{4.18}$$

Thus, the minBER-fixed design transmits a rotated version of the κ data symbols through the κ strongest channel eigenmodes, so that all data symbols experience the same BER performance. The instantaneous BER averaged over the κ data symbols defined in (4.8) is then given by

$$\mathsf{BER}_\kappa(\mathsf{snr}) = \frac{\alpha_\kappa}{\log_2 M_\kappa} \mathcal{Q}\left(\sqrt{\beta_\kappa \rho_\kappa} \right) \tag{4.19}$$

where $M_\kappa \triangleq M_{k,\kappa}$, $\alpha_\kappa \triangleq \alpha_{k,\kappa}$, $\beta_\kappa \triangleq \beta_{k,\kappa}$ for $k = 1, \ldots, \kappa$, since all constellations are equal. Now, taking into account all possible channel states, the average BER is obtained as

$$\overline{\mathsf{BER}}_\kappa(\mathsf{snr}) = \mathbb{E}\{\mathsf{BER}_\kappa(\mathsf{snr})\} = \frac{\alpha_\kappa}{\log_2 M_\kappa} \int_0^\infty \mathcal{Q}\left(\sqrt{\beta_\kappa \rho} \right) f_{\rho_\kappa}(\rho) d\rho \tag{4.20}$$

where $f_{\rho_\kappa}(\rho)$ is the pdf of the instantaneous SNR, ρ_κ, given in (4.18). Observe that ρ_κ is a non-trivial function of the κ strongest eigenvalues of the channel matrix $\mathbf{H}^\dagger\mathbf{H}$. Thus, a closed-form expression for the marginal pdf $f_{\rho_\kappa}(\rho)$ and by extension for the average BER in (4.20) is extremely difficult to obtain. However, we can derive easily computable average BER bounds based only on the marginal cdf of the κth largest channel eigenvalue given in Theorem 2.2 as done in the following theorem.

Theorem 4.1. *The average BER attained by the minimum BER linear transceiver with fixed and equal constellations (assuming κ data symbols per channel use) under the $n_\mathsf{R} \times n_\mathsf{T}$ MIMO channel models in Definitions* 3.1–3.4 *can be bounded as*

$$\overline{\mathsf{BER}}_\kappa^{(\mathrm{lb})}(\mathsf{snr}) \leq \overline{\mathsf{BER}}_\kappa(\mathsf{snr}) \leq \overline{\mathsf{BER}}_\kappa^{(\mathrm{ub})}(\mathsf{snr}) \tag{4.21}$$

with

$$\overline{\mathsf{BER}}_\kappa^{(\mathrm{ub})}(\mathsf{snr}) = \frac{\alpha_\kappa}{2\log_2 M_\kappa} \sqrt{\frac{\beta_\kappa \mathsf{snr}}{2\pi\kappa}} \int_0^\infty \frac{e^{-\frac{\beta_\kappa \lambda \mathsf{snr}}{2\kappa}}}{\sqrt{\lambda}} F_{\lambda_\kappa}(\lambda)\, d\lambda \tag{4.22}$$

$$\overline{\mathsf{BER}}_\kappa^{(\mathrm{lb})}(\mathsf{snr}) = \frac{\alpha_\kappa}{2\log_2 M_\kappa} \frac{\beta_\kappa \kappa \mathsf{snr}}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-\frac{\beta_\kappa(\lambda\kappa\mathsf{snr}+\kappa-1)}{2}}}{\sqrt{\beta_\kappa(\lambda\kappa\mathsf{snr}+\kappa-1)}} F_{\lambda_\kappa}(\lambda)\, d\lambda \tag{4.23}$$

where $F_{\lambda_k}(\cdot)$ is the marginal cdf of the kth largest eigenvalue in Theorem 2.2 *and the corresponding expressions for each channel model are given in Tables* 2.2–2.4.

Proof. See Appendix 4.A.1.

Remark 4.1. *Observing that the average BER upper-bound in Theorem* 4.1 *follows from using a diagonal scheme with a uniform power allocation among the κ established substreams and recalling Theorem* 3.7*, the average BER of the minBER-fixed design can be tighter upper-bounded in the high-SNR regime by dividing $\overline{\mathsf{BER}}_{\kappa}^{(\mathrm{ub})}(\mathsf{snr})$ by κ.*

Although Theorem 4.1 provides a numerical procedure to bound the average BER performance of the minBER-fixed design without resorting to the time-comsuming Monte Carlo simulations, it is still difficult to extract any conclusion on how to improve the system performance. Thus, we focus now on the high-SNR regime and provide a simpler performance characterization in terms of the array gain and the diversity gain. In Chapter 3, the average BER versus SNR curves of the channel eigenmodes have been parameterized for an uncorrelated, a semicorrelated Rayleigh, and an uncorrelated Rician fading MIMO channel. In addition, the performance characterization of the channel eigenmodes has been applied to analyze the global average BER performance of practical linear MIMO transceivers. For the sake of completeness we provide the high-SNR average BER characterization of the minBER-fixed scheme in the following theorem.

Theorem 4.2. *The average BER attained by the minimum BER linear transceiver with fixed and equal constellations (assuming κ data symbols per channel use) under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models in Definitions* 3.1*,* 3.2*, and* 3.4 *satisfies*[2]

$$\overline{\mathsf{BER}}_{\kappa}(\mathsf{snr}) = \left(G_{\mathsf{a},\kappa} \cdot \mathsf{snr}\right)^{-G_{\mathsf{d},\kappa}} + o\left(\mathsf{snr}^{-G_{\mathsf{d},\kappa}}\right) \tag{4.24}$$

where the diversity gain is given by

$$G_{\mathsf{d},\kappa} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1) \tag{4.25}$$

and the array gain can be bounded as[3]

$$G_{\mathsf{a},\kappa}^{(\mathrm{lb})} < G_{\mathsf{a},\kappa} < G_{\mathsf{a},\kappa}^{(\mathrm{ub})} \tag{4.26}$$

with

$$G_{\mathsf{a},\kappa}^{(\mathrm{lb})} = \frac{\beta_\kappa}{\kappa}\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\frac{a_\kappa 2^{d_\kappa}\Gamma(d_\kappa + 3/2)}{\sqrt{\pi}\kappa(d_\kappa+1)}\right)^{-1/(d_\kappa+1)} \tag{4.27}$$

$$G_{\mathsf{a},\kappa}^{(\mathrm{ub})} = \kappa\beta_\kappa\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\frac{a_\kappa I(d_\kappa, \beta_\kappa(\kappa-1))}{\sqrt{2\pi}(d_\kappa+1)}\right)^{-1/(d_\kappa+1)} \tag{4.28}$$

[2] We say that $f(x) = o(g(x))$, $g(x) > 0$, if $f(x)/g(x) \to 0$ as $x \to 0$ [Bru81, eq. (1.3.1)].

[3] Observe that an array gain lower bound provides an upper bound on the high-SNR average BER and vice versa.

where $\Gamma(\cdot)$ *denotes the gamma function in Definition* 2.5*,* $I(d,\beta)$ *is given in* (3.21)*, and the parameters* a_κ *and* d_κ *model the pdf of the* κ*th largest eigenvalue as in Theorem* 2.4*. The corresponding expressions for each channel model are given in Tables* 2.2 and 2.3.

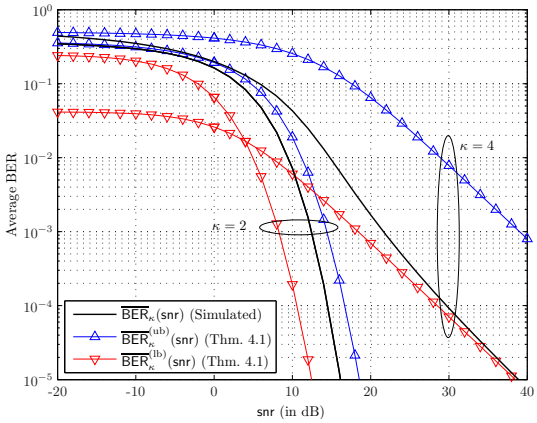
Proof. See Proof of Proposition 3.3 in Appendix 3.C.3.

In Figures 4.2-(i) and 4.2-(iii) we show the average BER performance of the minBER-fixed design and the average BER bounds derived in Theorem 4.1 in an uncorrelated and a semicorrelated Rayleigh MIMO channel, respectively. In both cases we consider the minBER-fixed scheme with $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, a target transmission rate of $\mathsf{R} = 8$ bits per channel use, and $\kappa = \{2, 4\}$. In Figures 4.2-(ii) and 4.2-(iv) we show the high-SNR performance and the parameterized upper and lower average BER bounds (dashed lines) corresponding respectively to the lower and upper array gain bounds derived in Theorem 4.2. We only include the beamforming strategy ($\kappa = 1$) in the high-SNR plots, as for this case the upper and lower bounds coincide with the exact average BER. It turns out that for $\kappa > 1$ the proposed average BER upper bound is more convenient to approximate the low SNR performance while the average BER lower bound is very tight in the high-SNR regime.

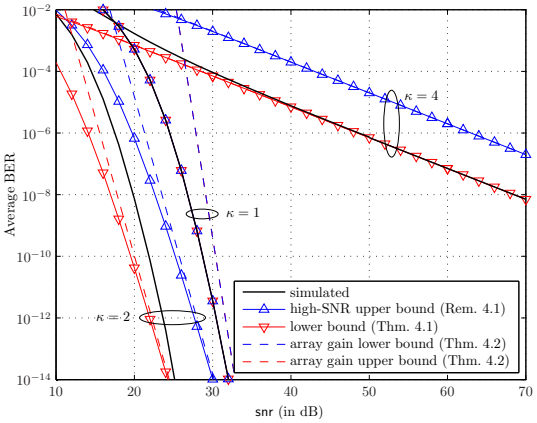
Finally, it is important to note that the diversity gain given in Theorem 4.2 coincides with the diversity gain achieved with the classical SVD transmission scheme without the additional rotation of the data symbols (cf. Theorems 3.2 and 3.5). Hence, Theorem 4.2 shows that the minBER-fixed design does not provide any diversity advantage with respect to diagonal schemes with simpler channel non-dependent power allocation policies but only a higher array gain. Actually, this statement is not exclusive of the investigated scheme but a common limitation of all linear MIMO transceivers whenever the number of symbols to be transmitted is fixed beforehand (even when using different constellations), as we show in the next theorem.

Theorem 4.3. *The diversity gain attained by any linear MIMO transceiver with fixed constellations* (*assuming* κ *data symbols per channel use*) *under the* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *MIMO channel models in Definitions* 3.1*,* 3.2*, and* 3.4 *satisfies*

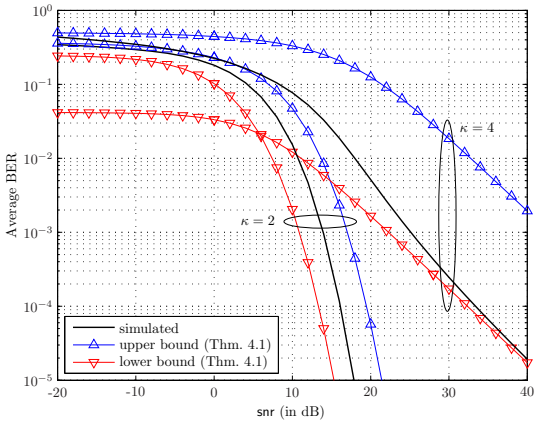
$$G_{\mathsf{d},\kappa} \leq (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1). \tag{4.29}$$



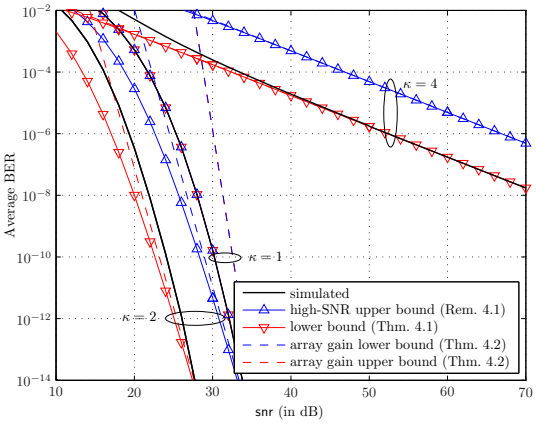
(i) Average BER, uncorrelated Rayleigh



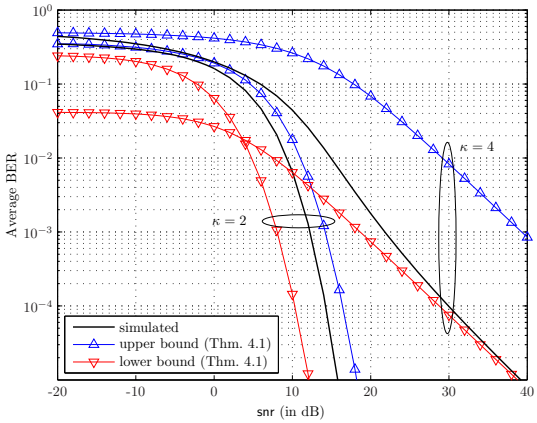
(ii) High-SNR Average BER, uncorrelated Rayleigh



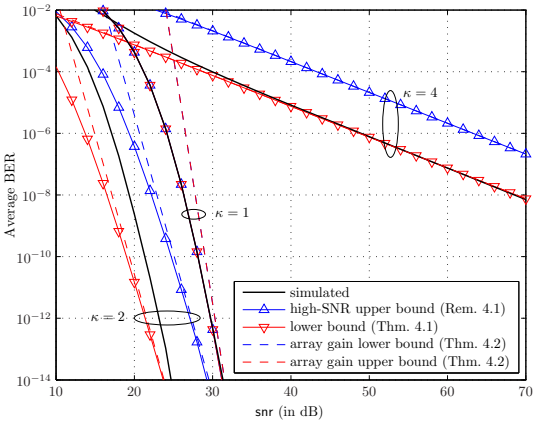
(iii) Average BER, semicorrelated Rayleigh



(iv) High-SNR Average BER, semicorrelated Rayleigh



(v) Average BER, uncorrelated Rician



(vi) High-SNR Average BER, uncorrelated Rician

Figure 4.2 Simulated average BER of the minBER-fixed design and bounds ($n_\mathsf{T} = 4, n_\mathsf{R} = 4, \kappa = \{1, 2, 4\}, \mathsf{R} = 8$) in an uncorrelated Rayleigh, a semicorrelated Rayleigh (with correlation matrix $[\mathbf{\Sigma}]_{i,j} = r^{|i-j|}$, $r = 0.7$), and an uncorrelated Ricean (with Ricean factor $K_\mathsf{c} = 0$dB) MIMO channel.

Proof. The BER averaged over the κ data symbols to be transmitted of any linear MIMO transceiver given in (4.8) can be lower-bounded as

$$\mathsf{BER}_{\kappa}(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}) \geq \frac{1}{\kappa} \max_{1\leq k\leq \kappa} \mathsf{BER}_{k,\kappa}(\rho_{k,\kappa}) \geq \frac{1}{\kappa} \min_{\mathbf{A}_{\kappa},\mathbf{B}_{\kappa}} \max_{1\leq k\leq \kappa} \mathsf{BER}_{k,\kappa}(\rho_{k,\kappa}). \tag{4.30}$$

The linear MIMO transceiver $(\mathbf{A}_{\kappa}, \mathbf{B}_{\kappa})$ that minimizes the maximum of the BERs as in (4.30) coincides with the optimum receive and transmit matrices given in (4.13) and (4.14), respectively (see Table 3.4). Hence, as the factor $1/\kappa$ does not have any influence on the SNR exponent, the diversity gain of any linear MIMO transceiver is upper-bounded by the one provided in Theorem 4.2. □

Intuitively, the performance of any linear MIMO transceiver is inherently limited by the performance of κth strongest channel eigenmode, since the design cost function (see (4.11) for the minimum BER design) is evaluated for the κ data symbols to be transmitted, regardless of whether transmission power is allocated to all κ channel eigenmodes during the effective transmission or not. This reveals that the average BER can be improved by introducing the parameter κ into the design criterion, as analyzed in the following section.

4.4 MinBER Linear MIMO Transceiver with Adaptive Number of Substreams

In this section we derive the minimum BER design with fixed rate and adaptive constellations and examine analytically its performance.

4.4.1 Linear Transceiver Design

The precoding process is in this case slightly different from classical linear precoding, where the number of data symbols to be transmitted per channel use κ is fixed beforehand. In the following, the parameter κ and the M_{κ}-dimensional constellations (assumed equal for simplicity) are adapted to the instantaneous channel conditions to minimize the BER by allowing κ to vary between 1 and $n = \min(n_{\mathsf{T}}, n_{\mathsf{R}})$ while keeping the total transmission rate $\mathsf{R} = \kappa \log_2 M_{\kappa}$ fixed. Usually, only a subset $\mathcal{K}$ of all n possible values of κ is supported, since the number of bits per symbol R/κ has to be an integer (see examples in Table 4.1). This simple optimization of κ suffices to exploit the full diversity of the channel whenever $\kappa = 1$ is included in $\mathcal{K}$, as shown in the next theorem.

R	$\kappa = 1$	$\kappa = 2$	$\kappa = 3$	$\kappa = 4$
1	BPSK	$\star$	$\star$	$\star$
2	QPSK	BPSK	$\star$	$\star$
3	8-QAM	$\star$	$\star$	$\star$
4	16-QAM	QPSK	$\star$	BPSK
6	64-QAM	8-QAM	QPSK	$\star$
8	256-QAM	16-QAM	$\star$	QPSK

Table 4.1 Examples of supported number of active substreams κ and the corresponding modulations for different rates R (in bits per channel use).

Theorem 4.4. *The diversity gain attained by any linear MIMO transceiver with adaptive number of substreams (assuming that the number of data symbols per channel use is chosen optimally from the set* $\{1\} \subseteq \mathcal{K} \subseteq \{1, \ldots, n\}$*) under the* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *MIMO channel models in Definitions* 3.1*,* 3.2*, and* 3.4 *satisfies*

$$G_{\mathsf{d},\mathcal{K}} = n_{\mathsf{T}} n_{\mathsf{R}} \tag{4.31}$$

provided that the linear transceiver design reduces to the optimum beamforming scheme for $\kappa = 1$*.*

Proof. The average BER of any linear MIMO transceiver when κ is optimized to minimize the BER can be upper-bounded using Jensen's inequality [Gra00, Sec. 12.411] as

$$\overline{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) = \mathbb{E}\{\min_{\kappa \in \mathcal{K}} \mathsf{BER}_{\kappa}(\{\rho_{k,\kappa}\}_{k=1}^{\kappa})\} \leq \min_{\kappa \in \mathcal{K}} \mathbb{E}\{\mathsf{BER}_{\kappa}(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}\} \leq \overline{\mathsf{BER}}_1(\mathsf{snr}) \tag{4.32}$$

where $\overline{\mathsf{BER}}_1(\mathsf{snr})$ denotes the average BER obtained for $\kappa = 1$. If, in this case, the optimum beamforming scheme is selected, it follows that

$$n_{\mathsf{T}} n_{\mathsf{R}} \geq G_{\mathsf{d},\mathcal{K}} \geq G_{\mathsf{d},1} = n_{\mathsf{T}} n_{\mathsf{R}} \tag{4.33}$$

where we have used Theorem 4.2 for $\kappa = 1$. Hence, the full diversity of the channel is achieved. □

Observe that a more general setup would also adapt the individual modulations without the constraint of equal constellations. However, the proposed minBER-adap scheme achieves already the full diversity of the channel with low complexity. On top of that, not even the minimum

BER linear transceiver with fixed and unequal constellations can be optimally obtained in closed form [Pal05b], and this dramatically increases the complexity of the system.

The linear transceiver $(\mathbf{A}_\kappa, \mathbf{B}_\kappa)$ and κ are designed to minimize the BER averaged over the data symbols to be transmitted for all supported values of κ:

$$\{\kappa, \mathbf{A}_\kappa, \mathbf{B}_\kappa\} = \arg \min_{\kappa \in \mathcal{K}, \mathbf{A}_\kappa, \mathbf{B}_\kappa} \mathsf{BER}_\kappa(\{\rho_{k,\kappa}\}_{k=1}^{\kappa}) \tag{4.34}$$

where $\mathbf{B}_\kappa$ has to satisfy the power-constraint in (4.4) and $\mathsf{BER}_\kappa(\{\rho_{k,\kappa}\}_{k=1}^{\kappa})$ is defined in (4.8). The optimum linear transceiver $(\mathbf{A}_\kappa, \mathbf{B}_\kappa)$ for a fixed κ and equal constellations has been presented and analyzed in Section 4.3. Using the resulting BER expression in (4.19), the optimum κ should be selected as

$$\kappa = \arg \min_{\kappa \in \mathcal{K}} \frac{\alpha_\kappa}{\log_2 M_\kappa} \mathcal{Q}\left(\sqrt{\beta_\kappa \rho_\kappa}\right) \tag{4.35}$$

or, neglecting the contribution of $\alpha_\kappa / \log_2 M_\kappa$ (since it is not in the argument of the Gaussian $\mathcal{Q}$-function), as[4]

$$\kappa = \arg \max_{\kappa \in \mathcal{K}} \beta_\kappa \rho_\kappa \tag{4.36}$$

where ρ_κ is given in (4.18).

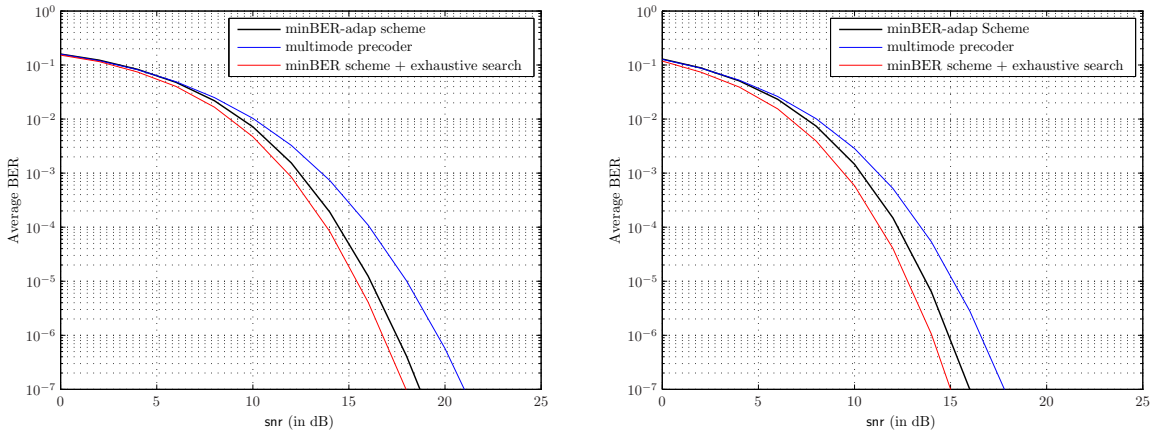
A similar scheme, named multimode precoder, that adapts the number of substreams under a fixed rate constraint was proposed in [Lov05b] in the context of limited feedback linear precoding. Even assuming perfect CSI, the multimode precoder designed in [Lov05b] is still suboptimum, since it does not perform the rotation to ensure equal BER on all substreams, the power is uniformly allocated among the established substreams, and parameter κ is suboptimally chosen as

$$\kappa = \arg \max_{\kappa \in \mathcal{K}} \beta_\kappa \lambda_\kappa / \kappa. \tag{4.37}$$

A different approach to overcome the diversity limitation of classical linear MIMO transceiver has been also recently given in [Vri08], where the precoder is designed to maximize the minimum Euclidean distance between symbols. However, the diversity order is only increased to $(n_\mathsf{T} - \kappa/2 + 1)(n_\mathsf{R} - \kappa/2 + 1)$.

In Figure 4.3 we compare the performance of the proposed minBER-adap scheme, the multimode precoder of [Lov05b], and the optimum minimum BER linear MIMO transceiver of [Pal05b]

[4] Numerical simulations do not show appreciable average BER differences between the selection functions in (4.35) and in (4.36) in the BER region of practical interest ($\mathsf{BER} << 10^{-1}$).



(i) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ (ii) $n_{\mathsf{T}} = 4$, $n_{\mathsf{R}} = 6$

Figure 4.3 Simulated average BER of the minBER-adap design ($\kappa = \{1, 2, 4\}$, $\mathsf{R} = 8$), the multimode precoder ($\kappa = \{1,2,4\}$, $\mathsf{R} = 8$), and the optimum minimum BER linear transceiver with an exhaustive search over all possible number of substreams and QAM constellations such that $\mathsf{R} = 8$ in an uncorrelated Rayleigh fading channel.

combined with an exhaustive search over all possible combinations of number of substreams and (possibly unequal) modulation orders which satisfy the rate constraint. We have obtained the average BER performance by numerical simulation in an uncorrelated Rayleigh fading channel, a target transmission rate of $\mathsf{R} = 8$ bits per channel use, and (i) $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$ and (ii) $n_{\mathsf{T}} = 4$, $n_{\mathsf{R}} = 6$. For the minBER-adap design and the multimode precoder, the number of substreams has been adapted with $\mathcal{K} = \{1, 2, 4\}$ and the corresponding constellations {256-QAM, 16-QAM, QPSK} (see Table 4.1), while for the optimum minimum BER system all 11 feasible combinations of number of substreams $\{1, 2, 3, 4\}$ and modulations {256-QAM, 128-QAM, 64-QAM, 32-QAM, 16-QAM, 8-QAM, QPSK, BPSK} have been taken into account. As expected, the minBER-adap design offers a better BER performance than the multimode precoder but it is still outperformed by the optimum minimum BER linear transceiver. However, this performance improvement over the proposed scheme does not justify in any case the prohibitive increase of complexity in the optimum design, which implies numerical linear transceiver design and exhaustive search over all possible combinations of substreams and modulations.

4.4.2 Analytical Performance

The minimum BER linear transceiver with fixed rate and adaptive constellations has the same structure as the minBER-fixed design presented in Section 4.3.1 but the number of symbols to

be transmitted is optimally adapted to minimize the BER. Analogously to Section 4.3.2 for the minBER-fixed scheme, we analyze in the following theorems the average BER performance of the minBER-adap design.

Theorem 4.5. *The average BER attained by the minimum BER linear transceiver with fixed rate and equal constellations (assuming that the number of data symbols per channel use is chosen from the set $\{1\} \subseteq \mathcal{K} \subseteq \{1,\ldots,n\}$ as in (4.36)) under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models in Definitions 3.1–3.4 can be bounded as*

$$\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{lb})}(\mathsf{snr}) \leq \overline{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) \leq \overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{ub})}(\mathsf{snr}) \tag{4.38}$$

with

$$\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{ub})}(\mathsf{snr}) = \min_{\kappa\in\mathcal{K}} \left(\frac{\alpha_\kappa}{2\log_2 M_\kappa}\right) \sqrt{\frac{\mathsf{snr}}{2\pi}} \int_0^\infty \frac{e^{-\frac{\lambda \mathsf{snr}}{2}}}{\sqrt{\lambda}} F_{\lambda_{\mathcal{K}}^{(\mathrm{ub})}}(\lambda) d\lambda \tag{4.39}$$

$$\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{lb})}(\mathsf{snr}) = \max_{\kappa\in\mathcal{K}} \left(\frac{\alpha_\kappa}{2\log_2 M_\kappa}\right) \frac{\mathsf{snr}}{\sqrt{2\pi}} \int_0^\infty \frac{e^{-\frac{\lambda \mathsf{snr}+\beta_{\mathcal{K}}}{2}}}{\sqrt{\lambda \mathsf{snr}+\beta_{\mathcal{K}}}} F_{\lambda_{\mathcal{K}}^{(\mathrm{lb})}}(\lambda) d\lambda \tag{4.40}$$

where $\beta_{\mathcal{K}} = \max_{\kappa\in\mathcal{K}}(\kappa-1)\beta_\kappa$ and we have defined

$$\lambda_{\mathcal{K}}^{(\mathrm{ub})} = \max_{\kappa\in\mathcal{K}}(\beta_\kappa \lambda_\kappa/\kappa) \quad \textit{and} \quad \lambda_{\mathcal{K}}^{(\mathrm{lb})} = \max_{\kappa\in\mathcal{K}}(\kappa\beta_\kappa\lambda_\kappa). \tag{4.41}$$

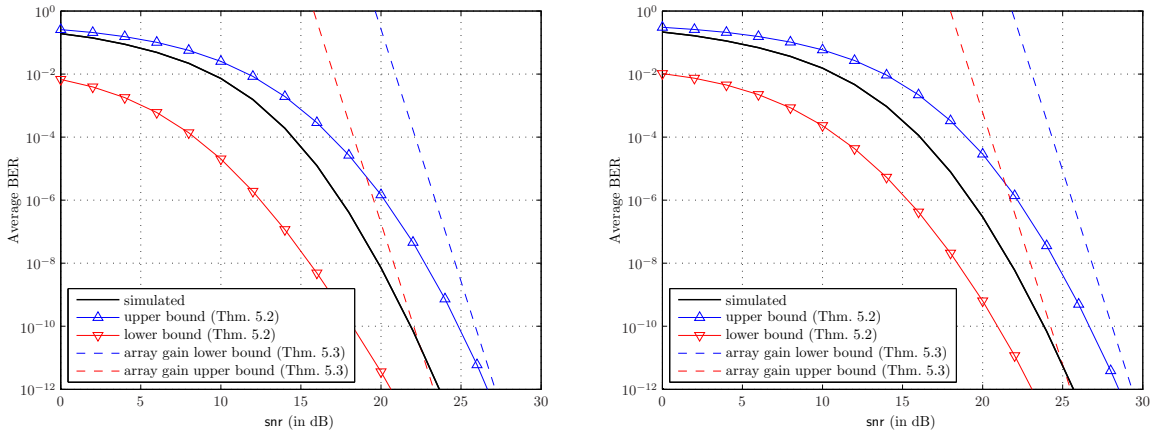
The corresponding cdfs, $F_{\lambda_{\mathcal{K}}^{(\mathrm{ub})}}(\cdot)$ and $F_{\lambda_{\mathcal{K}}^{(\mathrm{lb})}}(\cdot)$, can be obtained using Theorem 2.3 with the expressions for each channel model given in Tables 2.2–2.4.

Proof. See Appendix 4.A.2.

Remark 4.2. *Observe that $\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{ub})}(\mathsf{snr})$ in (4.39) coincides with the exact average BER performance of the multimode precoder of* [Lov05b] *except for the factor $\min_{\kappa\in\mathcal{K}}\left(\frac{\alpha_\kappa}{2\log_2 M_\kappa}\right)$ which is an upper bound. The corresponding lower bound can be obtained using (4.50) in Appendix 4.A.2.*

Theorem 4.6. *The average BER attained by the minimum BER linear transceiver with fixed rate and equal constellations (assuming that the number of data symbols per channel use is chosen from the set $\{1\} \subseteq \mathcal{K} \subseteq \{1,\ldots,n\}$ as in (4.36)) under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models in Definitions 3.1, 3.2, and 3.4 satisfies*

$$\overline{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) = (G_{\mathsf{a},\mathcal{K}} \cdot \mathsf{snr})^{-G_{\mathsf{d},\mathcal{K}}} + o\left(\mathsf{snr}^{-G_{\mathsf{d},\mathcal{K}}}\right) \tag{4.42}$$



(i) Uncorrelated Rayleigh Fading Channel (ii) Semicorrelated Rayleigh Fading Channel

Figure 4.4 Simulated average BER of the minBER-adap design and bounds ($n_{\mathsf{T}} = 4, n_{\mathsf{R}} = 4, \mathcal{K} = \{1, 2, 4\}, \mathsf{R} = 8$).

where the diversity gain is given by

$$G_{\mathsf{d},\mathcal{K}} = n_{\mathsf{T}} n_{\mathsf{R}} \tag{4.43}$$

the array gain can be bounded as

$$G_{\mathsf{a},\mathcal{K}}^{(\mathrm{lb})} < G_{\mathsf{a},\mathcal{K}} < G_{\mathsf{a},\mathcal{K}}^{(\mathrm{ub})} \tag{4.44}$$

with

$$G_{\mathsf{a},\mathcal{K}}^{(\mathrm{ub})} = \left(\max_{\kappa \in \mathcal{K}} \left(\frac{\alpha_\kappa}{\kappa \log_2 M_\kappa} \right) \frac{a_{\mathcal{K}}^{(\mathrm{ub})} 2^{d_{\mathcal{K}}}}{\sqrt{\pi}(d_{\mathcal{K}} + 1)} \right)^{-1/(d_{\mathcal{K},\mathrm{u}}+1)} \tag{4.45}$$

$$G_{\mathsf{a},\mathcal{K}}^{(\mathrm{lb})} = \left(\min_{\kappa \in \mathcal{K}} \left(\frac{\alpha_\kappa}{\log_2 M_\kappa} \right) \frac{a_{\mathcal{K}}^{(\mathrm{lb})} I(d_{\mathcal{K}}, \beta_{\mathcal{K}})}{\sqrt{2\pi}(d_{\mathcal{K}} + 1)} \right)^{-1/(d_{\mathcal{K}}+1)} \tag{4.46}$$

where $I(d, \beta)$ is defined (3.21) *and $\beta_{\mathcal{K}} = \max_{\kappa \in \mathcal{K}} (\kappa - 1)\beta_\kappa$. The parameters $\{a_{\mathcal{K}}^{(\mathrm{ub})}, d_{\mathcal{K}}\}$ and $\{a_{\mathcal{K}}^{(\mathrm{lb})}, d_{\mathcal{K}}\}$ model the pdfs of $\lambda_{\mathcal{K}}^{(\mathrm{ub})}$ and $\lambda_{\mathcal{K}}^{(\mathrm{lb})}$ defined in* (4.41). *They can be obtained using Theorem* 2.5 *with the expressions for each channel model given in Tables* 2.2 *and* 2.3.

Proof. The proof follows from using the high-SNR average BER characterizations in Lemma 3.1 and in Corollary 3.1 with the instantaneous SNR bounds derived in the proof of Theorem 4.5. □

Theorem 4.6 shows that the minimum BER linear transceiver with fixed rate and equal constellations effectively exploits the maximum diversity offered by the MIMO channel whenever $\kappa = 1$ is contained in $\mathcal{K}$. In Figure 4.4 we show the average BER performance of the minBER-adap design and of the average BER bounds derived in Theorem 4.5 in an uncorrelated and

a semicorrelated Rayleigh fading channel. We have considered the minBER-fixed scheme with $n_{\mathsf{T}} = n_{\mathsf{R}} = 4$, a target transmission rate of $\mathsf{R} = 8$ bits per channel use, and $\mathcal{K} = \{1, 2, 4\}$. As expected, the proposed design outperforms the classical minBER-fixed linear transceiver (cf. Figure 4.2).

4.5 Conclusions and Publications

The linear MIMO transceiver design has been addressed in the literature with the typical underlying assumption that the number of data symbols to be transmitted per channel use is chosen beforehand. In this chapter we have proved that, under this assumption, the diversity order of any linear MIMO transceiver is at most driven by that of the weakest channel eigenmode employed, which can be far from the diversity intrinsically provided by the channel. Based on this observation, we have fixed the rate (instead of the number of data symbols) and we have optimized the number of substreams and constellations jointly with the linear precoder. This procedure implies only an additional optimization stage upon the classical design which suffices to extract the full diversity of the channel. Since the ultimate performance of a communication system is given by the BER, we have focused on the minimum BER design. Nevertheless, a similar procedure holds for any of the practical linear MIMO transceiver designs presented in Section 3.8. The implications of the proposed optimization have been then illustrated by means of analytical performance analysis of the minimum BER linear MIMO transceiver with fixed and with adaptive number of substreams.

The main results contained in this chapter regarding the minimum BER linear MIMO transceiver design and performance analysis have been published in one journal paper and two conference papers:

[Ord07c] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "On equal constellation minimum BER linear MIMO transceivers", *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. III, pp. 221–224, Apr. 2007.

[Ord07a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "On the design of minimum BER linear MIMO transceivers with perfect or partial side channel state information", *Proc. IST Summit on Mobile and Wireless Comm.*, pp. 1–5, July 2007.

[Ord09b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Minimum BER linear MIMO transceivers with optimum number of substreams", *accepted in IEEE Trans. Signal Processing*, Jan. 2009.

4.A Appendix: Performance of Minimum BER Linear MIMO Transceivers

4.A.1 Proof of Theorem 4.1

Proof. The instantaneous SNR of the minBER-fixed design in (4.18) can be bounded as (see proof of Proposition 3.3 in Appendix 3.C.3)

$$\lambda_\kappa \mathsf{snr}/\kappa < \rho_\kappa < \kappa\lambda_\kappa \mathsf{snr} + (\kappa - 1). \tag{4.47}$$

The proof follows then from calculating the average BER attained with the bounds in (4.47):

$$\overline{\mathsf{BER}}_\kappa^{(\mathrm{ub})}(\mathsf{snr}) = \frac{\alpha_\kappa}{\log_2 M_\kappa} \int_0^\infty \mathcal{Q}\left(\sqrt{\beta_\kappa \lambda \mathsf{snr}/\kappa}\right) f_{\lambda_\kappa}(\lambda)\, d\lambda \tag{4.48}$$

$$\overline{\mathsf{BER}}_\kappa^{(\mathrm{lb})}(\mathsf{snr}) = \frac{\alpha_\kappa}{\log_2 M_\kappa} \int_0^\infty \mathcal{Q}\left(\sqrt{\beta_\kappa(\kappa\lambda \mathsf{snr} + (\kappa - 1))}\right) f_{\lambda_\kappa}(\lambda)\, d\lambda. \tag{4.49}$$

Finally, using integration by parts (see details in the proof of Theorem 3.1), we can rewrite (4.48) and (4.49) in terms of the cdf of of the κth largest eigenvalue, $F_{\lambda_\kappa}(\cdot)$, and this completes the proof. □

4.A.2 Proof of Theorem 4.5

Proof. The average BER of the minBER-adap design with κ chosen from $\mathcal{K}$ as in (4.36) can be bounded as

$$\min_{\kappa\in\mathcal{K}}\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\right)\widetilde{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) \le \overline{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) \le \max_{\kappa\in\mathcal{K}}\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\right)\widetilde{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) \tag{4.50}$$

where $\widetilde{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr})$ is

$$\widetilde{\mathsf{BER}}_{\mathcal{K}}(\mathsf{snr}) = \int_0^\infty \mathcal{Q}\left(\sqrt{\rho}\right) f_{\rho_{\mathcal{K}}}(\rho) d\rho \tag{4.51}$$

we have defined $\rho_{\mathcal{K}} = \max_{\kappa\in\mathcal{K}}(\beta_\kappa\rho_\kappa)$, and $f_{\rho_{\mathcal{K}}}(\cdot)$ denotes its pdf. Using the bounds of the instantaneous SNR ρ_κ in (4.47), it holds that

$$\max_{\kappa\in\mathcal{K}}(\kappa\beta_\kappa\lambda_\kappa)\mathsf{snr} + \max_{\kappa\in\mathcal{K}}\beta_\kappa(\kappa-1) \ge \max_{\kappa\in\mathcal{K}}(\kappa\beta_\kappa\lambda_\kappa\mathsf{snr} + \beta_\kappa(\kappa-1)) > \rho_{\mathcal{K}} > \max_{\kappa\in\mathcal{K}}(\beta_\kappa\lambda_\kappa/\kappa)\mathsf{snr}. \tag{4.52}$$

Let us define $\lambda_{\mathcal{K}}^{(\mathrm{ub})} = \max_{\kappa\in\mathcal{K}}(\beta_\kappa\lambda_\kappa/\kappa)$ and $\lambda_{\mathcal{K}}^{(\mathrm{lb})} = \max_{\kappa\in\mathcal{K}}(\kappa\beta_\kappa\lambda_\kappa)$ and denote their pdf by $f_{\lambda_{\mathcal{K}}^{(\mathrm{ub})}}(\cdot)$ and $f_{\lambda_{\mathcal{K}}^{(\mathrm{lb})}}(\cdot)$, respectively. The proof follows then from calculating the average BER attained with the bounds in (4.52) and combining them with (4.50):

$$\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{ub})}(\mathsf{snr}) = \max_{\kappa\in\mathcal{K}}\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\right)\int_0^\infty \mathcal{Q}\left(\sqrt{\lambda\mathsf{snr}}\right) f_{\lambda_{\mathcal{K}}^{(\mathrm{ub})}}(\lambda)\, d\lambda \tag{4.53}$$

$$\overline{\mathsf{BER}}_{\mathcal{K}}^{(\mathrm{lb})}(\mathsf{snr}) = \min_{\kappa\in\mathcal{K}}\left(\frac{\alpha_\kappa}{\log_2 M_\kappa}\right)\int_0^\infty \mathcal{Q}\left(\sqrt{\lambda\mathsf{snr} + \beta_{\mathcal{K}}}\right) f_{\lambda_{\mathcal{K}}^{(\mathrm{lb})}}(\lambda)\, d\lambda. \tag{4.54}$$

Finally, we can rewrite (4.53) and (4.54) in terms of the cdfs $F_{\lambda_{\mathcal{K}}^{(\mathrm{ub})}}(\cdot)$ and $F_{\lambda_{\mathcal{K}}^{(\mathrm{lb})}}(\cdot)$, respectively, using again integration by parts (see details in the proof of Theorem 3.1). □

5

Spatial Multiplexing MIMO Systems with CSI: Diversity and Multiplexing Tradeoff

Following the seminal work of Zheng and Tse, this chapter investigates the fundamental diversity and multiplexing tradeoff of spatial multiplexing multiple-input multiple-output (MIMO) systems in which knowledge of the channel state at both sides of the link is employed to transmit independent data streams through the channel eigenmodes. First, the fundamental diversity and multiplexing tradeoff of each of the individual substreams is obtained and this result is then used to derive a tradeoff optimal scheme for rate allocation among channel eigenmodes. The tradeoff of spatial multiplexing is finally compared to the fundamental tradeoff of the MIMO channel and to the tradeoff of both space only codes and V-BLAST which do not require channel state information at the transmit side.

5.1 Introduction

The usual performance characterization of a communication scheme based on computing the average error probability as a function of the signal-to-noise ratio (SNR) for a fixed data rate, may not be appropriate when comparing several systems with different data rates. In order to compare these schemes fairly, Forney and Ungerboeck proposed in [For98] to plot the average error probability against the normalized SNR, defined as the nominal SNR divided by the SNR needed to achieved the actual data rate as predicted by the channel capacity formula. Under this philosophy, the performance of a system could be also evaluated alternatively by obtaining the average error probability as function of the data rate, for a fixed SNR level. Analogously, the data rate should be normalized by the capacity of the channel to take into account the effect of the SNR. Based on these considerations, Zheng and Tse proposed a new framework for comparison among MIMO systems [Zhe03], in which the performance is characterized by the tradeoff between the diversity and multiplexing gains when both, the SNR and the transmission rate, increase without bound.

Indeed, Zheng and Tse showed in [Zhe03] that both diversity and multiplexing gains can be simultaneously obtained, but there is a fundamental tradeoff between how much of each type of gain any coding scheme can extract. The main result of [Zhe03] is a simple characterization of the optimal tradeoff curve between diversity and spatial multiplexing gains for any coding scheme when perfect channel state information (CSI) is available only at the receiver. This tradeoff framework does not only enlighten the fundamental limits of MIMO channels but also provides a very interesting procedure to compare the performance of existing diversity-based and multiplexing-based practical MIMO schemes by jointly analyzing their reliability and rate accommodation properties.

In this chapter we deal with the diversity and multiplexing tradeoff of MIMO systems with perfect CSI at the transmitter (CSI-T) and at the receiver (CSI-R)[1] in contrast to [Zhe03], where only perfect CSI-R is assumed. More specifically, we concentrate on spatial multiplexing MIMO systems, which divide the incoming data stream into multiple independent substreams without any temporal coding of the data symbols. When the transmitter is not aware of the channel realization, each substream is transmitted on a different antenna, such as the well-known V-

[1] The combination of CSI-T and CSI-R is henceforth referred to as just CSI.

BLAST scheme [Fos99]. However, when perfect CSI-T is available, performance can be further improved by transmitting the established substreams through the strongest channel eigenmodes. The resulting spatial multiplexing MIMO system with CSI is optimal in the sense of achieving the ergodic channel capacity [Tel99], and also arises in the joint linear transmitter-receiver design of practical MIMO systems, e.g., [Lee76, Sal85, Yan94b, Sca99, Sam01, Sca02b, Ong03, Pal03].

The approach we adopt in this chapter is to analyze the individual diversity and multiplexing tradeoff curves of the channel eigenmodes. Then, the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI is obtained by deriving the optimum rate allocation policy among these channel eigenmodes.

The rest of the chapter is organized as follows. Section 5.2 is devoted to introducing the system and MIMO channel model and Section 5.3 to briefly reviewing the diversity and multiplexing tradeoff framework. Section 5.4 describes the signal model corresponding to spatial multiplexing MIMO systems with CSI and motivates their analysis. In Section 5.5 we obtain the individual tradeoff curves of the different channel eigenmodes. Section 5.6 provides the fundamental diversity and multiplexing tradeoff of spatial multiplexing systems with CSI, which is compared to the fundamental limits offered by the channel in Section 5.7. Finally, the last section summarizes the main results of the chapter and provide the list of publications where they have been presented.

5.2 Preliminaries

5.2.1 System Model

We consider a wireless communication system with n_{T} transmit and n_{R} receive antennas, in which the channel matrix $\mathbf{H}$ remains constant within a block of T symbols, i.e., the block length is significantly smaller than the channel coherence time. In this situation, the received signal within one block can be gathered in an $n_{\mathsf{R}} \times \mathsf{T}$ matrix $\mathbf{Y}$ related to the $n_{\mathsf{T}} \times \mathsf{T}$ transmitted matrix $\mathbf{X}$ as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W} \tag{5.1}$$

where $\mathbf{W}$ is the additive white Gaussian noise and has i.i.d. entries with zero mean and unit variance, $[\mathbf{W}]_{ij} \sim \mathcal{CN}(0,1)$. The transmitted signal $\mathbf{X}$ is normalized forcing the transmit power

per channel use to satisfy

$$\frac{1}{\mathsf{T}}\mathbb{E}\left\{\|\mathbf{X}\|_{\mathrm{F}}^2\right\} \leq \mathsf{snr} \tag{5.2}$$

where snr is the average SNR at each receive antenna.

In the work by Zheng and Tse [Zhe03], the channel is assumed to be perfectly known at the receiver only. In contrast, we focus on the situation where the instantaneous channel gains are perfectly known at both transmitter and receiver, so that the transmitter can adapt its transmission strategy relative to the instantaneous channel state under the short-term power constraint in (5.2).

5.2.2 MIMO Channel Model

Recall that the MIMO channel with n_{T} transmit and n_{R} receive dimensions is described by an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ channel matrix $\mathbf{H}$, whose (i,j)th entry characterizes the propagation path between the jth transmit and the ith receive antenna. In wireless communications, the large number of scatters in the channel that contributes to the signal at the receiver results in Gaussian distributed channel matrix coefficients. Analogously to the single antenna channel, this model is referred to as MIMO Rayleigh or Rician fading channel, depending whether the channel entries are zero-mean or not. In this chapter we adopt the uncorrelated Rayleigh, the min-semicorrelated Rayleigh, and the uncorrelated Rician MIMO channel models introduced in Definitions 3.1, 3.2, and 3.4, respectively (see Section 3.3 for details).

5.2.3 Spatial Diversity and Spatial Multiplexing

The traditional role of multiple antennas was to provide spatial diversity to overcome channel fading by supplying the receiver with several independently faded replicas of the transmitted signal and, thus, increasing the link reliability. A MIMO channel with n_{T} transmit and n_{R} receive antennas has a maximal diversity order of $n_{\mathsf{T}}n_{\mathsf{R}}$, since there are a maximum of $n_{\mathsf{T}}n_{\mathsf{R}}$ random fading coefficients to be averaged over in the reception process of one symbol. Mathematically, we define the diversity gain as the slope of the average error probability versus SNR curve at high SNR, i.e., the SNR exponent of the average error probability.

Additionally, the simultaneous use of multiple antennas at the transmitter and receiver en-

ables also the exploitation of multiple parallel channels which can operate independently. This property is a consequence of the MIMO channel ergodic capacity which can be approximated in the high-SNR regime as [Fos98]:

$$C(\mathsf{snr}) \approx \min\{n_\mathsf{T}, n_\mathsf{R}\} \log(\mathsf{snr}) \,. \tag{5.3}$$

The channel capacity increases with the SNR as $\log(\mathsf{snr})$ with the prelog factor $\min\{n_\mathsf{T}, n_\mathsf{R}\}$, in contrast to the prelog factor 1 corresponding to SISO channels. In order to achieve a certain non-trivial fraction of the capacity in the high-SNR regime, we consider schemes that support a data rate which also increases with the SNR. Hence, we define a scheme as a family of codes $\{\mathcal{C}(\mathsf{snr})\}$ of block length T, which employs a different code $\mathcal{C}(\mathsf{snr})$ for each SNR level. Let $\mathsf{R}(\mathsf{snr})$ be the rate of a code $\mathcal{C}(\mathsf{snr})$, then a coding scheme is said to achieve a spatial multiplexing gain r if the supported data rate can be approximated in the high-SNR regime as

$$\mathsf{R}(\mathsf{snr}) \approx r \log \mathsf{snr}. \tag{5.4}$$

These intuitive definitions regarding spatial diversity and spatial multiplexing gain can be formalized as follows.

Definition 5.1. *A MIMO coding scheme $\{\mathcal{C}(\mathsf{snr})\}$ is said to achieve a spatial multiplexing gain r and a diversity gain d if the data rate satisfies*

$$\lim_{\mathsf{snr}\to\infty} \frac{\mathsf{R}(\mathsf{snr})}{\log \mathsf{snr}} = r \tag{5.5}$$

and the error probability follows[2]

$$\lim_{\mathsf{snr}\to\infty} \frac{\log \mathsf{P_e}(\mathsf{R}(\mathsf{snr}))}{\log \mathsf{snr}} = -d. \tag{5.6}$$

For each r, we define $d^\star(r)$ to be the supremum of the diversity gain achieved over all schemes.

Notation: Let us define the symbol $\doteq$ to denote exponential equality, i.e., $\mathsf{P}_i(\mathsf{snr}) \doteq \mathsf{P}_j(\mathsf{snr})$ denotes

$$\lim_{\mathsf{snr}\to\infty} \frac{\log \mathsf{P}_i(\mathsf{snr})}{\log \mathsf{snr}} = \lim_{\mathsf{snr}\to\infty} \frac{\log \mathsf{P}_j(\mathsf{snr})}{\log \mathsf{snr}}. \tag{5.7}$$

---

[2] This definition (introduced in [Zhe03, Def. 1]) differs from the standard definition of diversity gain widely used in the MIMO (see Section 3.2.4) and space-time coding literature (see e.g. [Tar98]) in the fact that we consider the error probability of a family of coding schemes $\{\mathcal{C}(\mathsf{snr})\}$, which employs a different code for each SNR level, instead of considering the error probability of a fixed coding scheme.

The symbols $\dot{\geq}$ and $\dot{\leq}$ are similarly defined. Hence, equation (5.6), for instance, can be expressed as

$$\mathsf{P_e}(\mathsf{R}(\mathsf{snr})) \doteq \mathsf{snr}^{-d}. \tag{5.8}$$

5.3 Fundamental Diversity and Multiplexing Tradeoff

5.3.1 Fundamental Diversity and Multiplexing Tradeoff with Perfect CSI-R

The design of MIMO systems has been traditionally tackled from two different perspectives: either the maximization of the diversity gain (to increase the transmission reliability) or the maximization of the spatial multiplexing gain (to approach the capacity limits). The diversity and multiplexing tradeoff was conceived as a unified framework to deal simultaneously with both design criteria. In fact, given a MIMO channel, both gains can be simultaneously obtained, and the fundamental tradeoff curve shows how much of each one any coding scheme can potentially extract or, equivalently, it provides the fundamental relation between the error probability and the normalized data rate in a system.

The tradeoff curve in [Zhe03] is based on analyzing the behavior of the error probability $\mathsf{P_e}(\mathsf{R})$ as function of the data rate R in the high-SNR regime by deriving tight (exponentially equal) upper and lower bounds. In particular, the lower bound is given by the outage probability, denoted by $\mathsf{P_{out}}(\mathsf{R})$ and defined as the probability that the mutual information between the input and the output of the channel is smaller than the data rate R [Oza94, Cai99, Big01]. On the other hand, the upper bound is obtained by conditioning the error probability on the outage event as

$$\mathsf{P_e}(\mathsf{R}) = \mathsf{P_{out}}(\mathsf{R})\mathsf{Pr}(\text{error}|\text{outage}) + \mathsf{Pr}(\text{error}, \text{no outage}) \tag{5.9}$$

$$\leq \mathsf{P_{out}}(\mathsf{R}) + \mathsf{Pr}(\text{error}, \text{no outage}). \tag{5.10}$$

The second term in (5.10) can be upper-bounded by the pairwise error probability averaged over the non-outage channel states and the resulting SNR exponent coincides with that of the lower bound whenever the coding length T satisfies

$$\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1. \tag{5.11}$$

This illustrates that, under the condition in (5.11), the typical error occurrence is caused by the channel outage event. The resulting tradeoff curve is given in the following lemma and plotted in Figure 5.1.

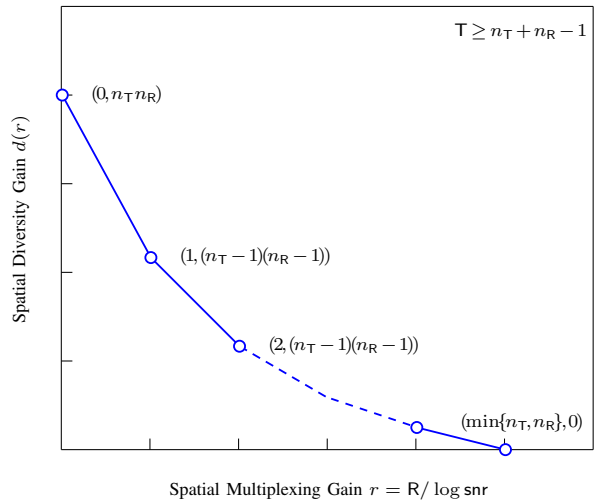


Figure 5.1 Fundamental diversity and multiplexing tradeoff of MIMO channels with perfect CSI-R.

Lemma 5.1 ([Zhe03, Thm. 2])**.** *Consider the MIMO system in* (5.1) *with perfect CSI-R only. The fundamental diversity and multiplexing tradeoff* $d^\star(r)$ *in an uncorrelated Rayleigh* $n_\mathsf{R} \times n_\mathsf{T}$ *MIMO fading channel* (*see Definition* 3.1) *is given by the piecewise-linear function connecting the points*

$$(k, d^\star(k)) \qquad k = 0, 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} \tag{5.12}$$

where

$$d^\star(k) = (n_\mathsf{T} - k)\,(n_\mathsf{R} - k) \tag{5.13}$$

whenever the coding block length condition in (5.11) *is satisfied.*

If the block length does not fulfill (5.11), the given upper bound can not be guaranteed to be tight anymore, since the outage occurrence is not the dominant event in the probability of error. Hence, we have to take into account the three circumstances leading to a detection error: (i) the channel matrix is atypically ill-conditioned, (ii) the additive noise is atypically large, and (iii) some codewords are atypically close together. When the block length is finite and satisfies (5.11), the two last events are averaged out and the probability of error is dominated by the bad channel occurrence. Otherwise, all three events come into play and the optimal tradeoff curve can be only partially obtained. In any case, the fundamental tradeoff given in Lemma 5.1 provides an upper-bound for the tradeoff achievable by any coding scheme with $\mathsf{T} < n_\mathsf{T} + n_\mathsf{R} - 1$.

In particular, when $\mathsf{T} = 1$ (space-only coding schemes), tight bounds on the average error

probability can be derived. The tradeoff curve is indeed completely characterized in [Zhe03] when no CSI-T is available. This case is of particular interest in this chapter, since we analyze spatial multiplexing schemes with CSI which do not perform any temporal encoding strategy, and, thus, it is introduced in Section 5.7 for comparison purposes.

The fundamental tradeoff of the channel originally derived in [Zhe03] and presented in Lemma 5.1 only holds for uncorrelated Rayleigh MIMO channels. This results have been however generalized in the literature for many different channel models [Cha06,Poo06,Cor07,Zha07,Shi08]. Interestingly, [Zha07] extend Lemma 5.1 to more general fading conditions, which include Rayleigh, Rician, Nakagami-m, Weibull, and Nakagami-q i.i.d. distributed fading coefficients. The effects of correlation and nonidentical distribution among the channel elements are also included in the analysis of [Zha07].

Remark 5.1 ([Zha07, Thm. 2, Thm. 3, and Cor. 3])**.** *The fundamental diversity and multiplexing tradeoff in Lemma* 5.1 *also holds for the semicorrelated Rayleigh and uncorrelated Rician MIMO fading channel models* (*see Definitions* 3.2 *and* 3.4)*.*

5.3.2 Fundamental Diversity and Multiplexing Tradeoff with Perfect CSI

The diversity and multiplexing tradeoff is still a meaningful framework when perfect CSI-T is also available. For instance, it is useful to analyze the performance of delay limited systems in which the data rate cannot depend on the channel variations except in outage states, where the channel cannot support the desired rate and the data to be transmitted is lost. In fact, during outages the wisest strategy is to stop transmission and do not waste power (see [Big01, Rem. 3]) but, of course, generating unrecoverable errors at the receiver.

When the channel is perfectly known at the transmitter, the coding strategy can be adapted to the instantaneous channel state by properly tuning the input distribution. As in [Zhe03], the input distribution can be taken to be Gaussian with a covariance matrix $\mathbf{R} \triangleq \mathbf{R}(\mathbf{H})$, where the optimum $\mathbf{R}$ maximizes the mutual information, i.e.,

$$\mathbf{R}^\star = \arg \max_{\mathbf{R}\geq 0, \mathrm{tr}(\mathbf{R})\leq \mathsf{snr}} \log\det(\mathbf{I}_{n_\mathsf{R}} + \mathbf{H}\mathbf{R}\mathbf{H}^\dagger). \tag{5.14}$$

Then, the outage probability [Big01, Def. 2]

$$\mathsf{P}_{\mathsf{out}}(\mathsf{R}) = \inf_{\mathbf{R}\geq 0, \mathrm{tr}(\mathbf{R})\leq \mathsf{snr}} \mathsf{Pr}\left(\log\det(\mathbf{I}_{n_\mathsf{R}} + \mathbf{H}\mathbf{R}\mathbf{H}^\dagger) \leq \mathsf{R}\right) \tag{5.15}$$

$$= \mathsf{Pr}\left(\log\det(\mathbf{I}_{n_\mathsf{R}} + \mathbf{H}\mathbf{R}^\star\mathbf{H}^\dagger) \leq \mathsf{R}\right) \tag{5.16}$$

can be upper-bounded by choosing

$$\mathbf{R}^{(\mathrm{ub})} = \frac{\mathsf{snr}}{n_\mathsf{T}}\mathbf{I}_{n_\mathsf{T}} \tag{5.17}$$

without exploiting the available CSI-T and lower-bounded by choosing

$$\mathbf{R}^{(\mathrm{lb})} = \mathsf{snr}\mathbf{I}_{n_\mathsf{T}} \tag{5.18}$$

since $(\mathsf{snr}\mathbf{I}_{n_\mathsf{T}} - \mathbf{R}^\star) \geq 0$, due to the short-term power constraint in (5.2), and this implies that $\log\det(\mathbf{I}_{n_\mathsf{R}} + \mathsf{snr}\mathbf{H}^\dagger\mathbf{H}) \geq \log\det(\mathbf{I}_{n_\mathsf{T}} + \mathsf{snr}\mathbf{H}\mathbf{R}^\star\mathbf{H}^\dagger)$. Both bounds were shown in [Zhe03] to be tight and the corresponding exponent was derived. Finally, using Fano's inequality, the outage probability was shown to provide a lower bound on the error probability in the high-SNR regime, independently of $\mathbf{R}$.

In consequence, the fundamental tradeoff presented in [Zhe03] for $\mathsf{T} \geq n_\mathsf{T} + n_\mathsf{R} - 1$ holds for any $\mathbf{R}$ bounded by (5.3.2) and (5.18) with or without CSI-T. This result is not surprising, since the diversity and multiplexing tradeoff framework focuses on the high-SNR regime, where the system is degree-of-freedom limited [Zhe03] and, thus, the additional power gain obtained by adapting $\mathbf{R}$ to the instantaneous channel does not modify the fundamental tradeoff of the channel whenever $\mathsf{T} \geq n_\mathsf{T} + n_\mathsf{R} - 1$.

5.4 Spatial Multiplexing MIMO Systems with CSI

5.4.1 Signal Model

Let us consider the block-fading MIMO system presented in (5.1), i.e.,

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}. \tag{5.19}$$

Following the singular value decomposition (SVD), the channel matrix $\mathbf{H}$ can be expressed as

$$\mathbf{H} = \mathbf{U}\sqrt{\mathbf{\Lambda}}\mathbf{V}^\dagger \tag{5.20}$$

where $\mathbf{U} \in \mathbb{C}^{n_{\mathsf{R}} \times n_{\mathsf{R}}}$ and $\mathbf{V} \in \mathbb{C}^{n_{\mathsf{T}} \times n_{\mathsf{T}}}$ are unitary matrices, and $\sqrt{\mathbf{\Lambda}} \in \mathbb{R}^{n_{\mathsf{R}} \times n_{\mathsf{T}}}$ is a diagonal[3] matrix containing the singular values of $\mathbf{H}$ sorted in descending order. Assuming that perfect CSI is available, we can rewrite the signal model in (5.19) without loss of generality as

$$\hat{\mathbf{S}} = \sqrt{\mathbf{\Lambda}}\mathbf{S} + \mathbf{N} \tag{5.21}$$

where $\hat{\mathbf{S}} = \mathbf{U}^{\dagger}\mathbf{Y}$, $\mathbf{S} = \mathbf{V}^{\dagger}\mathbf{X}$, and $\mathbf{N} = \mathbf{U}^{\dagger}\mathbf{W}$ with i.i.d. Gaussian entries with zero mean and unit variance, i.e., $[\mathbf{N}]_{i,j} \sim \mathcal{CN}(0,1)$, as the distribution of $\mathbf{W}$ is invariant under unitary transformations [Tul04, Ex. 2.4]. Since $\mathbb{E}\left\{\|\mathbf{X}\|_{\mathrm{F}}^2\right\} = \mathbb{E}\left\{\|\mathbf{S}\|_{\mathrm{F}}^2\right\}$, the power constraint in (5.2) can be equivalently expressed in terms of $\mathbf{S}$ as

$$\frac{1}{\mathsf{T}}\mathbb{E}\left\{\|\mathbf{S}\|_{\mathrm{F}}^2\right\} \leq \mathsf{snr} \tag{5.22}$$

where snr is the average SNR per receive antenna. Henceforth, we restrict our attention to spatial multiplexing MIMO systems with CSI, as formalized in the following assumptions.

Assumption 5.1. *The t-th transmitted vector, denoted by $\mathbf{s}_t \in \mathbb{C}^{n_{\mathsf{T}}}$, within a given block $\mathbf{S} = [\mathbf{s}_1\ \mathbf{s}_2\ \cdots\ \mathbf{s}_{\mathsf{T}}]$ is generated independently from all other transmitted vectors in $\mathbf{S}$, $\{\mathbf{s}_i\}_{i=1,\, i\neq t}^{\mathsf{T}}$, or, equivalently, coding is performed only across space.*

In this sense, we can restrict our attention to the case $\mathsf{T} = 1$, although our results hold for any given block length $\mathsf{T} < \infty$.

Assumption 5.2. *Let $\kappa \leq \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$ i.i.d., zero-mean, and unit energy data symbols, denoted by $\{z_k\}_{k=1}^{\kappa}$, be transmitted per channel use, such that (omitting the index t within a block):*

$$\mathbf{s} = \sqrt{\mathbf{P}_{\kappa}}\mathbf{z}_{\kappa} \tag{5.23}$$

where $\mathbf{z}_{\kappa} \in \mathbb{C}^{\kappa}$ gathers the κ data symbols, and $\mathbf{P}_{\kappa} \triangleq \mathbf{P}_{\kappa}(\mathbf{H}) \in \mathbb{R}^{n_{\mathsf{T}} \times \kappa}$ is a non-negative matrix whose off-diagonal entries are zero and the κ nonzero diagonal entries, $\{p_k \triangleq [\mathbf{P}_{\kappa}]_{k,k}\}_{k=1}^{\kappa}$, contain the power allocated among the κ established substreams and satisfy

$$\sum_{k=1}^{\kappa} p_k \leq \mathsf{snr} \tag{5.24}$$

due to the power constraint in (5.22).

---

[3] We call this matrix diagonal even though it may be not square.

The signal model corresponding to spatial multiplexing MIMO systems with CSI (see Assumptions 5.1 and 5.2) is then given by

$$\hat{\mathbf{s}} = \sqrt{\mathbf{\Lambda}\mathbf{P}_\kappa}\mathbf{z}_\kappa + \mathbf{n} \tag{5.25}$$

or, component-wise,[4]

$$\hat{s}_k = \sqrt{\lambda_k p_k} z_k + n_k \qquad \text{for } k = 1, \ldots, \kappa \tag{5.26}$$

where λ_k is the kth ordered ($\lambda_1 \geq \cdots \geq \lambda_\kappa$) eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$ (squared modulus of the kth channel singular value) and n_k is the kth component of the noise vector $\mathbf{n}$.

5.4.2 Motivation of Spatial Multiplexing MIMO system with CSI

The analysis of spatial multiplexing MIMO systems with CSI is mainly motivated by their optimality in terms of the channel capacity. Telatar showed in [Tel99] that the ergodic MIMO channel capacity with perfect CSI can be achieved by splitting the incoming data stream into $\min\{n_\mathsf{T}, n_\mathsf{R}\}$ substreams, coding these substreams separately using i.i.d. Gaussian codes, and waterfilling the available transmit power as

$$p_k = \left(\mu - \lambda_k^{-1}\right)^+ \qquad \text{for } k = 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} \tag{5.27}$$

where the water level μ is selected to satisfy the power constraint in (5.24) with equality. Note that this scheme can be described using the general signal model of spatial multiplexing MIMO systems in (5.26) with $\kappa = \min\{n_\mathsf{T}, n_\mathsf{R}\}$ and the power allocation given in (5.27).

It is worth pointing out that, sacrificing the low-complexity of the previous coding strategy (and the low-complexity of the corresponding optimum decoding), one can approach capacity with lower error probabilities using multidimensional codes, as can be inferred from the theory of error exponents of parallel channels [Gal68, Sec. 7.5]. As it will be seen in Section 5.7, this fact has important consequences in terms of the diversity and multiplexing tradeoff[5] achievable by spatial multiplexing MIMO schemes with perfect CSI.

However, the motivation behind the analysis of spatial multiplexing MIMO systems with CSI is not only supported by their optimality from the channel capacity point-of-view. Let us

[4] Observe that $\{\hat{s}_k\}_{k=\kappa+1}^{n_\mathsf{R}}$ only contain noise.

[5] The similitudes and differences between the theory of error exponents and the diversity and multiplexing tradeoff framework are investigated in [Zhe03, Sec. VI].

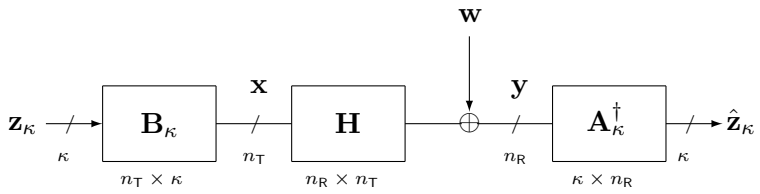


Figure 5.2 Linear MIMO transceivers system model.

assume that the ideal Gaussian codes are substituted with practical constellations (e.g. QAM) and that the MIMO system is equipped with a linear transmitter $\mathbf{B}_\kappa \in \mathbb{C}^{n_\mathsf{T} \times \kappa}$ and a linear receiver $\mathbf{A}_\kappa \in \mathbb{C}^{n_\mathsf{R} \times \kappa}$ (linear transceiver) as shown in Figure 5.2. The resulting signal model is given by

$$\hat{\mathbf{z}}_\kappa = \mathbf{A}_\kappa^\dagger(\mathbf{H}\mathbf{B}_\kappa \mathbf{z}_\kappa + \mathbf{w}) \tag{5.28}$$

where $\hat{\mathbf{z}}_\kappa \in \mathbb{C}^\kappa$ is the estimated data vector and $\mathbf{z}_\kappa \in \mathbb{C}^\kappa$ contains the κ data symbols as in (5.23). The uncoded linear transceiver optimization with perfect CSI, i.e., the joint optimization of $\mathbf{A}_\kappa$ and $\mathbf{B}_\kappa$ in (5.28) subject to the short term power constraint (equivalent to (5.24))

$$\mathrm{tr}\left(\mathbf{B}_\kappa \mathbf{B}_\kappa^\dagger\right) \leq \mathsf{snr} \tag{5.29}$$

has been largely studied in the literature under practical design criteria based on performance measures such as the SNR, the mean square error (MSE), or the bit error rate (BER), e.g., [Lee76, Sal85, Yan94b, Yan94a, Sca99, Sam01, Sca02b, Ong03, Pal03, Din03a, Pal07]. In most cases (see details in Section 5.5.3) the optimum strategy results in transmitting the κ independent data substreams through the κ strongest eigenmodes with a particular power allocation that depends on the specific design criterion. Hence, most of the linear MIMO transceiver designs proposed in the literature satisfy Assumptions 5.1 and 5.2 and, hence, can be also expressed as in (5.26).

5.5 Diversity and Multiplexing Tradeoff of the Individual Substreams

In this section we analyze the diversity and multiplexing tradeoff behavior of each individual substream transmitted in parallel through the channel eigenmodes, given the spatial multiplexing signal model presented in the previous section. We focus first on the capacity-achieving solution and derive the fundamental tradeoff of each individual MIMO eigenchannel. Then, we extend our analysis to include a more general class of power allocation policies and show that the given

individual diversity and multiplexing tradeoff curves also hold for some interesting practical linear transceiver designs.

5.5.1 Capacity-Achieving Spatial Multiplexing MIMO System with CSI

Let us consider the spatial multiplexing MIMO system in (5.26) with the power allocation $\{p_k\}_{k=1}^{\kappa}$ given by the capacity-achieving waterfilling in (5.27). Following the same procedure as in [Zhe03], the diversity and multiplexing tradeoff can be derived by bounding the individual error probability of the kth substream, $\mathsf{P_e}^{(k)}(\mathsf{R}_k)$, as

$$\mathsf{P_{out}}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{PEP}^{(k)}(\mathsf{R}_k) \tag{5.30}$$

where $\mathsf{P_{out}}^{(k)}(\mathsf{R}_k)$ denotes the outage probability and $\mathsf{PEP}^{(k)}(\mathsf{R}_k)$ denotes the pairwise error probability averaged over the non-outage channel states (see (5.10)) of the kth substream. The outage probability is given by [Tse05, Sec. 5.4]

$$\mathsf{P_{out}}^{(k)}(\mathsf{R}_k) = \mathsf{Pr}\left(\log(1+p_k\lambda_k) \leq r_k \log \mathsf{snr}\right) \tag{5.31}$$

where $0 \leq r_k \leq 1$ can not possibly exceed the value of 1, since it represents the rate of the kth eigenchannel in which capacity scales as $\log(1+p_k\lambda_k)$. The short-term power constraint in (5.24) implies that $p_k \leq \mathsf{snr}$ and, hence, $\mathsf{P_{out}}^{(k)}(\mathsf{R}_k)$ can be lower-bounded with the outage probability obtained when allocating all available power to the kth substream independently of the channel state:

$$\mathsf{P_{out}}^{(k)}(\mathsf{R}_k) \geq \mathsf{Pr}\left(\log(1+\mathsf{snr}\lambda_k) \leq r_k \log \mathsf{snr}\right) \tag{5.32}$$

$$\dot{\geq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right). \tag{5.33}$$

On the other hand, the pairwise error probability $\mathsf{PEP}^{(k)}(\mathsf{R}_k)$ can be upper-bounded by substituting the ideal Gaussian codes with some suboptimal coding strategy. Let us assume, for instance, that the data symbols are drawn from a QAM constellation of size snr^{r_k} in order to sustain a data rate of $\mathsf{R}_k = r_k \log \mathsf{snr}$. Hence, the minimum distance between symbols is $\mathsf{snr}^{-r_k/2}$. The pairwise error probability between two arbitrary symbols transmitted in the kth substream satisfies[6] [Zhe03, Sec. VII.A]

$$\mathsf{PEP}^{(k)}(\mathsf{R}_k) \doteq \mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1\right). \tag{5.34}$$

[6] The only difference with [Zhe03] is that in our case the channel power gain is λ_k instead of $\|\mathbf{H}\|_{\mathrm{F}}^2$.

In order to take into account the waterfilling mechanism, we rewrite (5.34) as

$$\mathsf{PEP}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1 \middle| p_k > 0\right)(1 - \mathsf{Pr}(p_k = 0)) + \mathsf{Pr}(p_k = 0) \tag{5.35}$$

$$\leq \mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1 \middle| p_k > 0\right) + \mathsf{Pr}(p_k = 0) \tag{5.36}$$

and the upper bound in (5.36) is shown in Appendix 5.A.1 to satisfy

$$\mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1 \middle| p_k > 0\right) + \mathsf{Pr}(p_k = 0) \dot{\leq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right). \tag{5.37}$$

Finally, combining (5.33) and (5.37) we conclude that

$$\mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1}) \dot{\leq} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1}) \tag{5.38}$$

and, hence,

$$\mathsf{P_e}^{(k)}(\mathsf{R}_k) \doteq \mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1}). \tag{5.39}$$

The resulting individual diversity and multiplexing tradeoff curves are characterized in the following theorem and plotted in Figure 5.3.

Theorem 5.1. *Consider a spatial multiplexing MIMO system with CSI (see Assumptions* 5.1 *and* 5.2*) and consider that the power allocated to the substream transmitted through the kth ordered channel eigenmode is given by the capacity-achieving waterfilling, i.e., $p_k = \left(\mu - \lambda_k^{-1}\right)^+$. The individual diversity and multiplexing tradeoff $d_{\mathsf{S}}^{(k)}(r)$ of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1*,* 3.2*, and* 3.4 *is given by*

$$d_{\mathsf{S}}^{(k)}(r_k) = d_k(1 - r_k) \qquad 0 \leq r_k \leq 1 \tag{5.40}$$

where d_k is defined as

$$d_k = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1). \tag{5.41}$$

Proof. See Appendix 5.A.2.

Note that the individual diversity and multiplexing tradeoff curves presented in Theorem 5.1 are, in fact, the fundamental tradeoff curves of the eigenmodes of the adopted $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channel models.

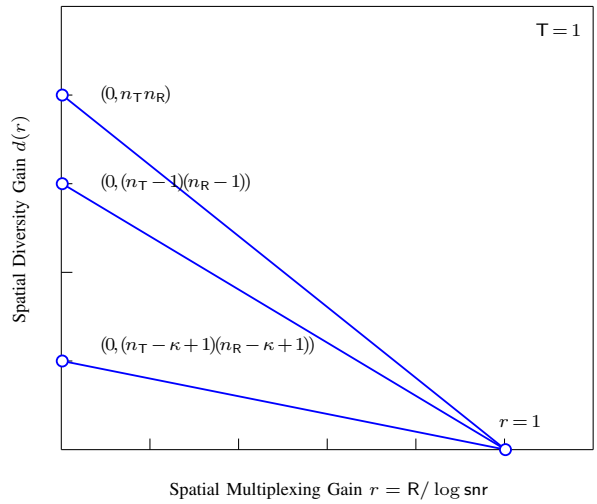


Figure 5.3 Fundamental diversity and multiplexing tradeoff of the MIMO channel eigenmodes.

5.5.2 General Spatial Multiplexing MIMO Systems with CSI

Let us now consider a spatial multiplexing MIMO system with CSI such that the power allocated to the kth substream is a function of the SNR and possibly a function of the κ strongest channel eigenvalues, i.e, $p_k = g_k(\lambda_1, \ldots, \lambda_\kappa, \mathsf{snr})$, and satisfies the short-term power constraint in (5.24). We also assume that there exits a deterministic strictly positive quantity, ϕ_k, such that

$$\mathsf{Pr}\left(p_k \leq \phi_k \mathsf{snr}\right) \stackrel{\cdot}{\leq} \mathsf{P_{out}}^{(k)}(\mathsf{R}_k). \tag{5.42}$$

This condition ensures that the power allocation does not inherently limit the achievable individual tradeoff either by allocating with large probability zero or a small amount of power in terms of the SNR. Observe that, due to the short-term power constraint, it is not possible to design a power allocation which privileges the individual tradeoffs of some substreams by penalizing the tradeoffs of the remaining substreams. The individual diversity and multiplexing tradeoff curves given in Theorem 5.1 for the capacity-achieving waterfilling also hold for this general power allocation policy as presented in the following corollary.

Corollary 5.1.1. *Consider a spatial multiplexing MIMO system with CSI (see Assumptions* 5.1 *and* 5.2*) and consider that the power allocated to the substream transmitted through the kth ordered channel eigenmode is given by $p_k = g_k(\lambda_1, \ldots, \lambda_\kappa, \mathsf{snr})$ under the condition in* (5.42) *and the short-term power constraint in* (5.24)*. The individual diversity and multiplexing tradeoff*

$d_{\mathsf{S}}^{(k)}(r)$ of the substream transmitted through the kth eigenmode of the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions 3.1*,* 3.2*, and* 3.4 *is given by*

$$d_{\mathsf{S}}^{(k)}(r_k) = d_k(1 - r_k) \qquad 0 \leq r_k \leq 1 \tag{5.43}$$

where d_k is defined in Theorem 5.1*.*

Proof. See Appendix 5.A.3.

Corollary 5.1.1 shows that the power allocated to an individual substream (under a short-term power constraint) cannot improve the SNR exponent of the error probability, which is on the contrary fixed by the order of the channel eigenmode used for communication.

5.5.3 Practical Spatial Multiplexing MIMO Systems with CSI

In Section 5.4.2 we introduced linear MIMO transceivers as a practical implementation of spatial multiplexing systems with CSI. Palomar developed in [Pal03] a unifying framework, in which the design problem of linear MIMO transceivers under perfect CSI is formulated as the minimization of some cost function of the MSEs of the individual substreams, since the other common practical system quality measures such as the SNR, or the BER can be easily related to the MSEs. More exactly, the optimum transmit matrix $\mathbf{B}_\kappa$ and receive matrix $\mathbf{A}_\kappa$ (see (5.28)) are obtained as

$$\{\mathbf{A}_\kappa, \mathbf{B}_\kappa\} = \quad \arg \min_{\mathbf{A}_\kappa, \mathbf{B}_\kappa} \qquad f_0\left(\{\mathsf{mse}_k\}_{k=1}^{\kappa}\right) \tag{5.44}$$

$$\text{subject to} \qquad \operatorname{tr}\left(\mathbf{B}_\kappa \mathbf{B}_\kappa^\dagger\right) \leq \mathsf{snr} \tag{5.45}$$

where $f_0(\cdot)$ denotes the design cost function and mse_k is the kth diagonal element of the MSE matrix $\mathbf{E} = \mathbb{E}\left\{(\hat{\mathbf{z}}_\kappa - \mathbf{z}_\kappa)(\hat{\mathbf{z}}_\kappa - \mathbf{z}_\kappa)^\dagger\right\}$. In particular, [Pal03] shows that the optimum linear receiver is always given by the MMSE solution [Pal03, eq. (7)]:

$$\mathbf{A}_\kappa = \left(\mathbf{H}\mathbf{B}_\kappa \mathbf{B}_\kappa^\dagger \mathbf{H}^\dagger + \mathbf{I}_{n_{\mathsf{R}}}\right)^{-1} \mathbf{H}\mathbf{B}_\kappa \tag{5.46}$$

and provides the optimum linear transmitter for the class of Schur-concave and Schur-convex cost functions. Interestingly, this framework embraces most of the linear transceiver schemes previosly proposed in the literature, e.g., [Lee76, Sal85, Yan94b, Yan94a, Sca99, Sam01, Sca02b, Din03a].

In the case of Schur-concave cost functions (see Tables 3.2 and 3.3 for a list of design criteria), the optimum transmit strategy establishes κ independent data streams through the κ strongest

channel eigenmodes with a power allocation that depends on the particular cost function [Pal03, eq. (14)]:

$$\mathbf{B}_\kappa = \mathbf{U}_\kappa \sqrt{\mathbf{P}_\kappa} \tag{5.47}$$

where $\mathbf{U}_\kappa \in \mathbb{C}^{n_\mathsf{T} \times \kappa}$ has as columns the eigenvectors of $\mathbf{H}^\dagger\mathbf{H}$ corresponding to the κ largest eigenvalues and $\mathbf{P}_\kappa \in \mathbb{C}^{\kappa \times \kappa}$ is a diagonal matrix containing the power allocation policy $\{p_k\}_{k=1}^{\kappa}$. For instance, when the design criterion is the minimization of the product of MSEs, the optimum power allocation is given by the capacity-achieving waterfilling in (5.27) [Pal03, eq. (24)]. When the focus is on the minimization of the weighted sum of MSEs, the optimum power allocation is given by [Pal03, eq. (22)]

$$p_{\mathsf{mse},k} = \left(\mu\omega_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1}\right)^+ \qquad \text{for } k = 1, \ldots, \kappa \tag{5.48}$$

where ω_k is a strictly positive constant (weight) and μ is chosen to fulfill the power constraint in (5.45), and when it is on the maximization of the weighted product of SNRs, the optimum power allocation is given by [Pal03, eq. (40)]

$$p_{\mathsf{snr},k} = \frac{\omega_k}{\sum_{i=1}^{\kappa}\omega_i}\mathsf{snr} \qquad \text{for } k = 1, \ldots, \kappa \tag{5.49}$$

where $\{\omega_k\}_{k=1}^{\kappa}$ is a strictly positive constant (weight). In the following propositions we show that the power allocation policies in (5.48) and (5.49) satisfy the conditions of Corollary 5.1.1. Observe that the optimum power allocation for the design criteria in Tables 3.2 and 3.3 is given by (5.27), (5.48), or (5.49). Hence, all these practical linear MIMO transceivers are included in our analysis using either Theorem 5.1 or Corollary 5.1.1.

Proposition 5.1. *The power allocation that minimizes the weighted sum of MSEs given in* (5.48) *satisfies the conditions of Corollary* 5.1.1.

Proof. See Appendix 5.B.

Proposition 5.2. *The power allocation that maximizes the weighted product of SNRs given by* (5.49) *satisfies the conditions of Corollary* 5.1.1.

Proof. The power allocation in (5.49) is channel non-dependent and, thus, the condition in (5.42) is directly satisfied. □

Finally, since the linear transceivers of (5.28) with $\mathbf{A}_\kappa$ and $\mathbf{B}_\kappa$ as given in (5.46) and (5.47), respectively, can be described by the general signal model presented in (5.26) for spatial multiplexing MIMO system with CSI (see Section 3.8.1 for details), it follows that the tradeoff curves given in Corollary 5.1.1 also characterize the individual substreams of practical linear transceiver designed under Schur-concave cost functions.

In the case of Schur-convex cost functions (see Table 3.4 for a list of design criteria), however, the optimum transmit strategy transmits a unitary transformation of the κ independent data streams through the κ strongest channel eigenmodes [Pal03, eq. (14)]:

$$\mathbf{B}_\kappa = \mathbf{U}_\kappa \sqrt{\mathbf{P}_\kappa} \mathbf{Q}_\kappa \tag{5.50}$$

where $\mathbf{U}_\kappa$ and $\mathbf{P}_\kappa$ are defined as in (5.47), and $\mathbf{Q}_\kappa \in \mathbb{C}^{\kappa \times \kappa}$ is a unitary matrix such that all κ substreams experience the same equivalent channel (see [Pal03] for details). Furthermore, the optimum power allocation is always given by (5.48). Schur-convex cost functions appear, for instance, in the minimization of the maximum of the MSEs, or in the minimization of the sum of BERs (under equal constellations). These schemes cannot be described by the signal model in (5.26), since the individual data symbols are not transmitted in parallel through the channel eigenmodes and, hence, Corollary 5.1.1 cannot be applied.

5.6 Diversity and Multiplexing Tradeoff of Spatial Multiplexing Systems with CSI

In the previous section we derived the SNR exponent of the error probability associated to the individual substreams transmitted in parallel through the channel eigenmodes when using the capacity-achieving waterfilling or a general power allocation that satisfies the condition in (5.42). Based on this result, we obtained the individual diversity and multiplexing tradeoff curves of the established substreams in Theorem 5.1 and Corollary 5.1.1, respectively. In contrast, in this section we are interested in the global diversity and multiplexing tradeoff of the spatial multiplexing MIMO system with CSI that results from the combination of these individual substreams.

Recall that the κ-dimensional data symbol vector $\mathbf{z}_\kappa \in \mathbb{C}^\kappa$ is formed by the κ i.i.d. individual data symbols $\{z_k\}_{k=1}^{\kappa}$ transmitted through each channel eigenmode (see Assumption 5.2). We define the global error probability of spatial multiplexing MIMO systems with CSI, denoted by $\mathsf{P_e}(\mathsf{R})$, $\mathsf{R} = \sum_{k=1}^{\kappa} r_k \log \mathsf{snr}$, as the probability of having an error in the detection of at least one

of the κ individual symbols in the data symbol vector $\mathbf{z}_\kappa$. Thus, $\mathsf{P_e}(\mathsf{R})$ can be bounded as

$$\max_{1\leq k\leq \kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \leq \mathsf{P_e}(\mathsf{R}) \leq \sum_{k=1}^{\kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k). \tag{5.51}$$

The SNR exponent of $\sum_{k=1}^{\kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k)$ is dominated by the term with the lowest exponent and this term corresponds precisely to the substream with the worst error probability, i.e, $\max_{1\leq k\leq \kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k)$. Assuming that the power allocation among the κ substreams is either given by the capacity-achieving waterfilling (or satisfies the condition in (5.42)), we can use Theorem 5.1 (or Corollary 5.1.1) and (5.51) to obtain that

$$\mathsf{P_e}(\mathsf{R}) \doteq \mathsf{snr}^{-\min_\kappa d_{\mathsf{S}}^{(\kappa)}(r_\kappa)} \tag{5.52}$$

where $d_{\mathsf{S}}^{(\kappa)}(r_\kappa)$ is defined in (5.40) and[7] $\kappa \geq \lceil r \rceil$, since the individual multiplexing gains $\{r_k\}_{k=1}^{\kappa}$ cannot exceed the value 1. Hence, as it becomes apparent later on in this section, the diversity and multiplexing tradeoff curve depends on both the number of active substreams κ and the rate allocation policy adopted by the transmitter, $\{r_k\}_{k=1}^{\kappa}$. In the following, we present first the tradeoff curve obtained when using a uniform rate allocation among the active substreams. Then, we derive the optimum rate allocation, which results in the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI formalized in Assumptions 5.1 and 5.2.

5.6.1 Diversity and Multiplexing Tradeoff with Uniform Rate Allocation

Let us denote by $\kappa(r)$ the number of active substreams as a function of the spatial multiplexing gain $r = \sum_{k=1}^{\kappa(r)} r_k$ with $0 < r_k \leq 1$ and let us assume a uniform rate allocation policy among these $\kappa(r)$ substreams, i.e.,

$$r_k = \frac{r}{\kappa(r)} \qquad \text{for } k = 1, \ldots, \kappa(r). \tag{5.53}$$

Then, as shown in (5.52), the diversity and multiplexing tradeoff is obtained by deriving the optimum number of active substreams:

$$\kappa^\star(r) = \arg \max_{\kappa \geq \lceil r \rceil} \min_{1\leq k\leq \kappa} d_{\mathsf{S}}^{(k)}(r/\kappa). \tag{5.54}$$

This problem is implicitly solved in the following theorem and the resulting diversity and multiplexing tradeoff curve $d_{\mathsf{SU}}^\star(r)$ is plotted in Figure 5.4.

---

[7] $\lceil a \rceil$ denotes the smallest integer bigger than or equal to a.

Theorem 5.2. *Consider a spatial multiplexing MIMO system with CSI (see Assumptions* 5.1 *and* 5.2*) and consider that the power allocation among the active substreams is either given by the capacity-achieving waterfilling in* (5.27) *or satisfies the condition in* (5.42)*. The diversity and multiplexing tradeoff $d^{\star}_{\mathsf{SU}}(r)$ achievable with a uniform rate allocation policy under the $n_{\mathsf{R}} \times n_{\mathsf{T}}$ MIMO channels in Definitions* 3.1*,* 3.2*, and* 3.4 *is given by the piecewise-linear function connecting the points $(r(\kappa), d^{\star}_{\mathsf{SU}}(\kappa))$ and $(\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}, 0)$, where*

$$\begin{aligned} r(\kappa) &= \kappa - \frac{\kappa d_{\kappa+1}}{(\kappa+1)d_{\kappa} - \kappa d_{\kappa+1}} \\ d^{\star}_{\mathsf{SU}}(\kappa) &= \frac{d_{\kappa} d_{\kappa+1}}{(\kappa+1)d_{\kappa} - \kappa d_{\kappa+1}} \end{aligned} \qquad \text{for } \kappa = 0, \ldots, \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\} - 1 \tag{5.55}$$

with $d_{\kappa} = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1)$ and $r(\kappa)$ denoting the values of r at which the number of active substreams is increased from κ to $\kappa + 1$.

Proof. Assuming a uniform rate allocation among the κ active eigenchannels, the SNR exponent of the kth substream is equal to

$$d^{(k)}_{\mathsf{S}}(r/\kappa) = d_k(1 - r/\kappa) \qquad \text{for } k = 1, \ldots, \kappa \tag{5.56}$$

where $d_k = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1)$. Since it holds that $d_k > d_{k+1}$ for $k = 1, \ldots, \kappa - 1$, the global performance of the spatial multiplexing scheme with κ active channel eigenmodes is dictated by the individual performance of the κth substream. Hence, for a given multiplexing gain r, we can maximize the diversity gain by optimizing the number of active substreams κ

$$\kappa^{\star}(r) = \arg \max_{\kappa \geq \lceil r \rceil} d^{(\kappa)}_{\mathsf{S}}(r/\kappa) \tag{5.57}$$

and the global diversity and multiplexing tradeoff is given by the supremum of the individual tradeoff curves, $\{d^{(\kappa)}_{\mathsf{S}}(r/\kappa)\}_{\kappa=1,\ldots,\min\{n_{\mathsf{T}},n_{\mathsf{R}}\}}$,

$$d^{\star}_{\mathsf{SU}}(r) = \max_{\kappa \geq \lceil r \rceil} d^{(\kappa)}_{\mathsf{S}}(r/\kappa). \tag{5.58}$$

Clearly, the resulting tradeoff curve is a piecewise-linear function connecting the points $r(\kappa)$, defined as the values of r at which the optimum number of active substreams is increased from κ to $\kappa + 1$, i.e., the discontinuity points of $\kappa^{\star}(r)$. Forcing equal diversity gains in these points:

$$d^{(\kappa)}_{\mathsf{S}}(r/\kappa) = d^{(\kappa+1)}_{\mathsf{S}}(r/(\kappa+1)) \tag{5.59}$$

it follows that

$$r(\kappa) = \kappa - \frac{\kappa d_{\kappa+1}}{(\kappa+1)d_{\kappa} - \kappa d_{\kappa+1}} \qquad \text{for } \kappa = 1, \ldots, \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\} - 1 \tag{5.60}$$

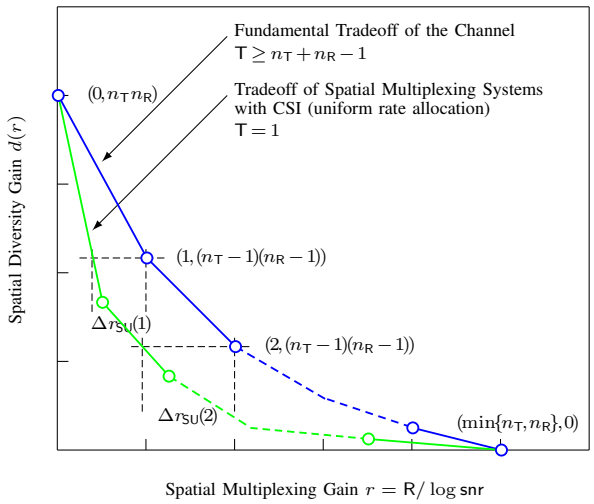


Figure 5.4 Diversity and multiplexing tradeoff of spatial multiplexing systems with CSI (uniform rate allocation).

and, finally, substituting (5.60) back in (5.56), we obtain the corresponding diversity gains

$$d^\star_{\mathsf{SU}}(\kappa) = d^{(\kappa)}_{\mathsf{S}}(r(\kappa)/\kappa) = \frac{d_\kappa d_{\kappa+1}}{(\kappa+1)d_\kappa - \kappa d_{\kappa+1}} \qquad \kappa = 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} - 1 \tag{5.61}$$

which completes the proof. □

5.6.2 Diversity and Multiplexing Tradeoff with Optimal Rate Allocation

The derivation of the fundamental tradeoff of spatial multiplexing systems with CSI requires maximizing the minimum SNR exponent not only over the number of active channel eigenmodes, $\kappa(r)$, but also over the rate allocation policy among the established substreams, $\{r_k\}_{k=1}^{\kappa(r)}$:

$$d^\star_{\mathsf{S}}(r) = \max_{\kappa, \{r_k\}_{k=1}^{\kappa}} \min_{1 \leq k \leq \kappa} d^{(k)}_{\mathsf{S}}(r_k) \tag{5.62}$$

$$\text{subject to} \quad \kappa \geq \lceil r \rceil \tag{5.63}$$

$$\sum_{k=1}^{\kappa} r_k = r \tag{5.64}$$

$$0 < r_k \leq 1. \tag{5.65}$$

This problem is implicitly solved in the following theorem and the resulting diversity and multiplexing tradeoff curve $d^\star_{\mathsf{S}}(r)$ is plotted in Figure 5.5.

Theorem 5.3. *Consider a spatial multiplexing MIMO system with CSI (see Assumptions* 5.1 *and* 5.2*) and consider that the power allocation among the active substreams is either given by the*

capacity-achieving waterfilling in (5.27) *or satisfies the condition in* (5.42)*. The fundamental diversity and multiplexing tradeoff* $d_{\mathsf{S}}^{\star}(r)$ *under the* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *MIMO channels in Definitions* 3.1, 3.2, *and* 3.4 *is given by the piecewise-linear function connecting the points* $(0, n_{\mathsf{T}} n_{\mathsf{R}})$, $(r(\kappa), d_{\mathsf{S}}^{\star}(\kappa))$, *and* $(\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}, 0)$, *where*

$$\begin{aligned} r(\kappa) &= \kappa - d_{\kappa+1}\left(\sum_{k=1}^{\kappa} 1/d_k\right) \\ d_{\mathsf{S}}^{\star}(\kappa) &= (n_{\mathsf{T}} - \kappa)(n_{\mathsf{R}} - \kappa) \end{aligned} \qquad \text{for } \kappa = 1, \ldots, \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\} - 1 \tag{5.66}$$

with $d_\kappa = (n_{\mathsf{T}} - \kappa + 1)(n_{\mathsf{R}} - \kappa + 1)$ *and* $r(\kappa)$ *denoting the values of* r *at which the number of active substreams is increased from* κ *to* $\kappa + 1$*. The fundamental diversity and multiplexing tradeoff is achieved if the rate is allocated among the optimum number of active substreams* $\kappa^{\star}(r)$ *as*

$$r_k = 1 - \frac{1/d_k}{\sum_{i=1}^{\kappa^{\star}(r)} 1/d_i}(\kappa^{\star}(r) - r) \qquad \text{for } k = 1, \ldots, \kappa^{\star}(r). \tag{5.67}$$

Proof. The constrained optimization problem in (5.62) is equivalent to first imposing a rate allocation that assures the same SNR exponent $d(\kappa, r)$ for the κ active substreams:

$$d(\kappa, r) \triangleq d_{\mathsf{S}}^{(k)}(r_k) = d_k(1 - r_k) \qquad \text{for } k = 1, \ldots, \kappa \tag{5.68}$$

and, then, maximizing the resulting SNR exponent $d(\kappa, r)$ over κ subject to the constraint in (5.63), i.e.,

$$d_{\mathsf{S}}^{\star}(r) = \max_{\kappa \geq \lceil r \rceil} d(\kappa, r) \tag{5.69}$$

$$\kappa^{\star}(r) = \arg\max_{\kappa \geq \lceil r \rceil} d(\kappa, r). \tag{5.70}$$

From (5.68) we obtain the individual rates that force all κ active substreams to have the same SNR exponent:

$$r_k = 1 - \frac{d(\kappa, r)}{d_k} \qquad \text{for } k = 1, \ldots, \kappa \tag{5.71}$$

and, combining (5.71) and (5.64), we have that

$$d(\kappa, r) = \frac{\kappa - r}{\sum_{k=1}^{\kappa} 1/d_k}. \tag{5.72}$$

The fundamental diversity and multiplexing tradeoff $d_{\mathsf{S}}^{\star}(r)$ is then given by the supremum of the individual curves, $\{d(\kappa, r)\}_{\kappa=1,\ldots,\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}}$. As in Theorem 5.2, the resulting tradeoff curve is a piecewise-linear function connecting the points $r(\kappa)$, defined as the values of r at which the

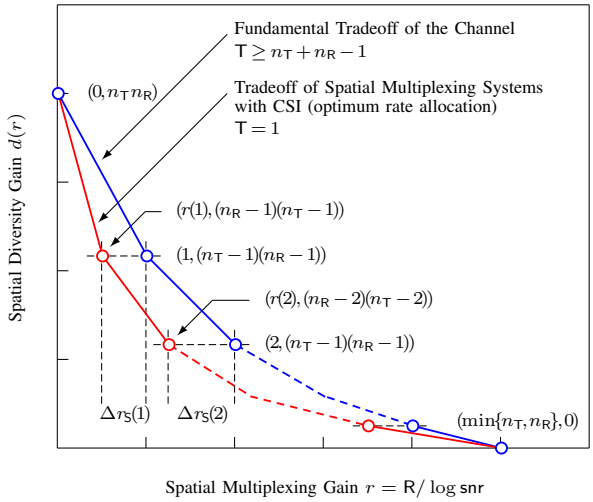


Figure 5.5 Diversity and multiplexing tradeoff of spatial multiplexing systems with CSI (optimum rate allocation).

optimum number of active substreams is increased from κ to $\kappa+1$, i.e., the discontinuity points of $\kappa^\star(r)$. Forcing equal diversity gains in these points:

$$d(\kappa, r(\kappa)) = d(\kappa+1, r(\kappa)) \tag{5.73}$$

it follows that

$$r(\kappa) = \kappa - d_{\kappa+1}\left(\sum_{k=1}^{\kappa} 1/d_k\right) \qquad \text{for } \kappa = 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} - 1 \tag{5.74}$$

and, substituting (5.74) back in (5.72), we obtain the corresponding diversity gains

$$d_\mathsf{S}^\star(\kappa) = d(\kappa, r(\kappa)) = (n_\mathsf{T} - \kappa)(n_\mathsf{R} - \kappa) \qquad \text{for } \kappa = 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} - 1. \tag{5.75}$$

Finally, the optimal rate allocation among substreams comes simply from combinig (5.71) and (5.72) and this completes the proof. □

5.6.3 Achievability of the Diversity and Multiplexing Tradeoff

In the previous section we obtained the diversity and multiplexing tradeoff of spatial multiplexing systems with CSI based on the individual tradeoff curves derived in Theorem 5.1. Thus, the diversity and multiplexing tradeoff curves provided in Theorems 5.2 and 5.3 for the uniform and the optimal rate allocation policy, respectively, are achieved whenever the individual tradeoffs in Theorem 5.1 are achieved. In addition to the capacity-achieving spatial multiplexing scheme

analyzed in Section 5.5.1, we considered spatial multiplexing systems with a general power allocation in Section 5.5.2 and practical linear MIMO transceivers in Section 5.5.3. Consequently, the fundamental diversity and multiplexing tradeoff of spatial multiplexing systems with CSI can be achieved with practical linear MIMO transceiver designs even with a channel non-dependent power allocation and using QAM constellations.

Recall that in Section 5.5.3 we proved the individual tradeoff achievability for the class linear transceivers designed under Schur-concave cost functions. However, the individual tradeoff performance of the linear MIMO transceivers obtained under Schur-convex cost functions could not be analyzed using Theorem 5.1 due to the unitary transformation applied to the data symbols before transmission. This preprocessing of the data symbols forces all κ established substreams to experience the same equivalent channel (see Section 4.3.2 for details) and, hence, it holds that

$$\max_{1\leq k\leq \kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \geq \max_{1\leq k\leq \kappa} \mathsf{P_{out}}^{(k)}(\mathsf{R}_k) = \max_{r_k} \mathsf{Pr}(\log(1+\rho_k) \leq r_k \log \mathsf{snr}) \tag{5.76}$$

$$> \max_{r_k} \mathsf{Pr}(\log(\kappa + \kappa \mathsf{snr}\lambda_\kappa) \leq r_k \log \mathsf{snr}) \tag{5.77}$$

$$\doteq \max_{r_k} \mathsf{Pr}(\lambda_\kappa \leq \mathsf{snr}^{r_k-1}) \doteq \mathsf{snr}^{-\min_{r_k} d_{\mathsf{S}}^{(\kappa)}(r_k)} \tag{5.78}$$

where in (5.77) we have used the upper bound on the instantaneous SNR derived in Appendix 3.C.3. Similarly, using the lower bound on the instantaneous SNR given also in Appendix 3.C.3, it is not difficult to show that

$$\sum_{k=1}^{\kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \leq \kappa \max_{1\leq k\leq \kappa} \mathsf{P_e}^{(k)}(\mathsf{R}_k) \doteq \mathsf{snr}^{-\min_{r_k} d_{\mathsf{S}}^{(\kappa)}(r_k)} \tag{5.79}$$

and, noting (5.51), we can finally conclude that

$$\mathsf{P_e}(\mathsf{R}) \doteq \mathsf{snr}^{-\min_{r_k} d_{\mathsf{S}}^{(\kappa)}(r_k)}. \tag{5.80}$$

The diversity and multiplexing tradeoff of the linear MIMO transceivers obtained under Schur-convex cost functions follows then from maximizing the exponent of the error probability in (5.80) with respect to the number of active substreams κ and the rate allocation among them, $\{r_k\}_{k=1}^{\kappa}$. In this case the optimum strategy is to allocate uniformly the rate among the established substreams and, thus, the resulting tradeoff coincides with that presented in Theorem 5.2.

5.7 Analysis of the Results

In this section we analyze the diversity and multiplexing tradeoffs of spatial multiplexing MIMO systems with CSI with respect to the fundamental tradeoff of the channel and the achievable tradeoffs under different assumptions when space-only coding is performed and no CSI-T is available. Observe that space-only coding schemes imply $\mathsf{T} = 1$ or the adoption of Assumption 5.1, whereas the fundamental tradeoff requires $\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1$ and, hence, provides an upper bound on the achievable tradeoff for any block length T.

5.7.1 Tradeoff of Spatial Multiplexing with CSI ($\mathsf{T} = 1$) vs. Fundamental Tradeoff of the Channel ($\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1$)

In this section we compare the diversity and multiplexing tradeoff of spatial multiplexing systems with CSI with the fundamental tradeoff offered by the channel for $\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1$ as illustrated in Figures 5.4 and 5.5 for the uniform and the optimum rate allocation policies, respectively. It can be observed that, for a given diversity gain, spatial multiplexing schemes suffer from a degradation with respect to the diversity and multiplexing limits of the channel, due to the lack of coding between substreams (see Assumption 5.2). In fact, the established substreams are coded independently and transmitted through the κ strongest channel eigenmodes and, hence, the outage event must be understood under the perspective of the individual eigenmodes, since spatial multiplexing systems experiences an outage whenever at least one of the κ established substreams is in outage.

Let us consider the fundamental tradeoff of the channel and the fundamental tradeoff of spatial multiplexing systems with CSI (see Figure 5.5) and let us focus on values of r close to 0, such that only the first linear part of both curves is taken into consideration. This requires $0 \leq r \leq 1$ in the fundamental tradeoff of the channel and $0 \leq r \leq 1 - \frac{(n_{\mathsf{T}}-1)(n_{\mathsf{R}}-1)}{n_{\mathsf{T}} n_{\mathsf{R}}}$ in the spatial multiplexing case (see Figure 5.6). In this region, the tradeoff exponent of the spatial multiplexing system, $d^{\star}_{\mathsf{S}}(r)$, is derived by evaluating the outage probability for the transmission through the eigenmode associated with the largest eigenvalue of $\mathbf{H}^{\dagger}\mathbf{H}$, denoted by λ_1. More exactly, applying Theorem 5.1 and noting that for $0 \leq r \leq 1 - \frac{(n_{\mathsf{T}}-1)(n_{\mathsf{R}}-1)}{n_{\mathsf{T}} n_{\mathsf{R}}}$ only one substream is used, it follows that $r_1 = r$ and

$$\mathsf{P}_{\mathsf{out}}(\mathsf{R}) \doteq \mathsf{Pr}(\lambda_1 \leq \mathsf{snr}^{-(1-r)}) \doteq \mathsf{snr}^{-d^{\star}_{\mathsf{S}}(r)} = \mathsf{snr}^{-n_{\mathsf{T}} n_{\mathsf{R}}(1-r)}. \tag{5.81}$$

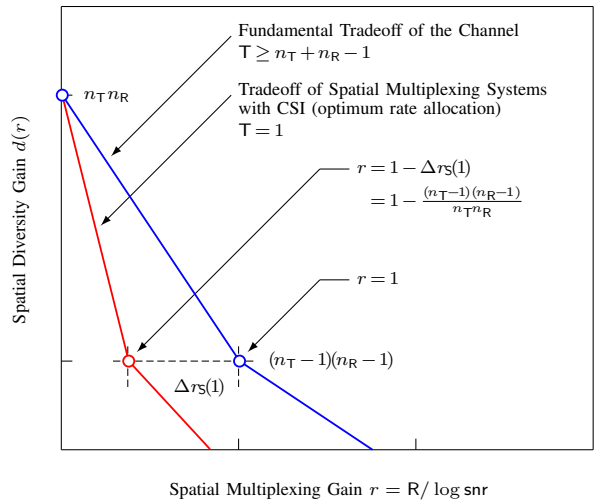


Figure 5.6 Diversity and multiplexing tradeoff of the channel and of spatial multiplexing systems with CSI (low spatial multiplexing regime).

On the other hand, the fundamental tradeoff curve is obtained evaluating the dominant term of the SNR exponent of the outage probability, which in the interval $0 \leq r \leq 1$ can be expressed as [Zhe03, Sec. III.B]

$$\mathsf{P}_{\mathsf{out}}(r) \doteq \mathsf{Pr}(\lambda_1 \leq \mathsf{snr}^{-(1-r)}, \lambda_2 \leq \mathsf{snr}^{-1}, \cdots, \lambda_{\min\{n_\mathsf{T}, n_\mathsf{R}\}} \leq \mathsf{snr}^{-1}) \doteq \mathsf{snr}^{-d^\star(r)} \tag{5.82}$$

where $d^\star(r) = (n_\mathsf{T} - 1)(n_\mathsf{R} - 1) + (n_\mathsf{T} + n_\mathsf{R} - 1)(1 - r)$. Comparing the expressions in (5.81) and (5.82), it is now apparent that the probability of outage of the strongest eigenmode is higher than the dominant outage event probability of the MIMO channel, which in addition requires all other $\min\{n_\mathsf{T}, n_\mathsf{R}\} - 1$ eigenmodes to be fully ineffective. A similar argument can be used in different regions of the tradeoff curve (see Figure 5.5) and this explains the diversity and multiplexing loss.

This performance degradation can be characterized more exactly by defining the spatial multiplexing loss $\Delta r(\kappa)$ as

$$\Delta r(\kappa) = \kappa - r(d^\star(\kappa)) \qquad \text{for } \kappa = 1, \ldots, \min\{n_\mathsf{T}, n_\mathsf{R}\} - 1 \tag{5.83}$$

where $r(d^\star(\kappa))$ is the spatial multiplexing gain for which the spatial multiplexing system achieves a diversity gain of $(n_\mathsf{T} - \kappa)(n_\mathsf{R} - \kappa)$. Using Theorems 5.2 and 5.3, we obtain

$$\Delta r_{\mathsf{SU}}(\kappa) = \min_{0 \leq k < \kappa} k + d_{\kappa+1} \frac{\kappa - k}{d_{\kappa - k}} \tag{5.84}$$

for the uniform rate allocation and

$$\Delta r_{\mathsf{S}}(\kappa) = d_{\kappa+1} \sum_{k=1}^{\kappa} \frac{1}{d_k}. \tag{5.85}$$

for the optimum rate allocation. Observe that in (5.84) the multiplexing loss is dominated by $\kappa - k$ times the inverse of the maximum diversity of the worst active eigenmode, $(\kappa - k)/d_{\kappa-k}$, whereas in (5.85) the multiplexing loss depends on the sum of the inverse of the maximum diversities associated the κ active channel eigenmodes, $\sum_{k=1}^{\kappa} 1/d_k$. By allocating the same rate among all substreams, the performance of the system is limited by the performance of the $(\kappa - k)$th channel eigenmode, since it is the worst active subchannel and is the first one to become in outage. Therefore, in this case, not even the active subchannels are fully exploited. On the contrary, by using the optimal rate allocation, we force all active channel eigenmodes to become in outage simultaneously or, equivalently, we transmit always (for any given diversity) at the maximum individual rate supported by each channel eigenmode. Moreover, with the optimal rate allocation the number of active substreams is optimal in the sense that it coincides with the number of channel eigenmodes which are typically not in outage (see geometrical interpretation in [Zhe03, Sec. III.B]) and this is not always the case when using the uniform rate allocation, where the term k in (5.84) can further increase the multiplexing loss. As we illustrate in the following section, however, spatial multiplexing with CSI often provides an advantage with respect to space-only codes and spatial multiplexing with CSI-R only.

5.7.2 Tradeoff of Spatial Multiplexing with CSI ($\mathsf{T} = 1$) vs. Tradeoff of Space-only Codes with CSI-R ($\mathsf{T} = 1$)

In this section we focus on the tradeoff achieved by any space-only coding scheme ($\mathsf{T} = 1$ or Assumption 5.1 is satisfied) and CSI-R only. The fundamental diversity and multiplexing tradeoff of space-only codes was derived in [Zhe03] as given in the following lemma.

Lemma 5.2 ([Zhe03, Sec. IV.D] and Rem. 5.1)**.** *Consider the MIMO system in* (5.1) *with* $\mathsf{T} = 1$ *and perfect CSI-R only. The fundamental diversity and multiplexing tradeoff* $d^{\star}_{S_0}(r)$ *under the* $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *MIMO channels in Definitions* 3.1, 3.2, *and* 3.4 *with* $n_{\mathsf{T}} \leq n_{\mathsf{R}}$ *is given by*

$$d^{\star}_{\mathsf{S}_0}(r) = n_{\mathsf{R}} \left(1 - r/\min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}\right) \qquad \text{for } 0 \leq r \leq \min\{n_{\mathsf{T}}, n_{\mathsf{R}}\}. \tag{5.86}$$

Observe that the fundamental tradeoff of spatial multiplexing MIMO systems with CSI obtained in Theorem 5.3 can be expressed as

$$d^\star_{\mathsf{S}}(r) = \max_{\kappa,\kappa\geq\lceil r\rceil} \frac{\kappa}{\sum_{k=1}^{\kappa} 1/d_k}\,(1-r/\kappa) \qquad \text{for } 0\leq r\leq \min\{n_{\mathsf{T}},n_{\mathsf{R}}\}. \tag{5.87}$$

Clearly, it holds that $d^\star_{\mathsf{S}}(r)\geq d^\star_{\mathsf{S}_0}(r)$ as long as $n_{\mathsf{T}}\neq n_{\mathsf{R}}$. Hence, CSI-T exploitation, or more exactly, the diversity gain increase obtained by selecting only the strongest channel eigenmodes and optimally allocating the rate overcomes the limitations inherent to the independent coding procedure of spatial multiplexing systems. Only when $n_{\mathsf{T}}=n_{\mathsf{R}}$ and for

$$r\geq \max_{\kappa,\kappa\geq r} \frac{\kappa - n_{\mathsf{T}}\sum_{k=1}^{\kappa} 1/d_k}{1-\sum_{k=1}^{\kappa} 1/d_k} \tag{5.88}$$

higher diversity gains can be achieved with space-only codes. Intuitively, in the high multiplexing gain regime, all channel eigenmodes are used to transmit independent data streams, rendering CSI-T less useful. On the contrary, an arbitrary space coding scheme can potentially benefit from coding among different streams. This turns out to be more beneficial than spatial multiplexing with CSI when $n_{\mathsf{T}}=n_{\mathsf{R}}$, i.e., when the weakest eigenmode has diversity 1. In fact, when $n_{\mathsf{T}}=n_{\mathsf{R}}$ and $r>\min\{n_{\mathsf{T}},n_{\mathsf{R}}\}-1$, $d^\star_{S_0}(r)$ coincides with the fundamental tradeoff of the channel $d^\star(r)$.

Observe that it is false to conclude from this result that there are circumstances under which is better to ignore the perfect CSI at the transmitter. Indeed, the comparison is not completely fair, since space-only codes can transmit any data vector with arbitrary correlation among components while spatial multiplexing MIMO systems are restricted to transmit data vectors with i.i.d. data symbols due to Assumption 5.2. This low-complexity spatial multiplexing strategy, although capacity-achieving, is suboptimal in terms of error probability (see e.g. [Tel99, Ex. 1]).

5.7.3 Tradeoff of Spatial Multiplexing with CSI ($\mathsf{T}=1$) vs. Tradeoff of V-BLAST ($\mathsf{T}=1$)

Finally, in this section we compare the fundamental tradeoff of spatial multiplexing systems with CSI and with CSI-R only, in order to evaluate the benefits of having perfect channel knowledge at the transmitter. Let us consider the spatial multiplexing system with CSI-R only, in which an independent substream is transmitted through each transmit antenna ($n_{\mathsf{T}}\leq n_{\mathsf{R}}$) with the same rate and the receiver uses the nulling and canceling algorithm. This scheme is commonly known as V-BLAST [Fos99] and the corresponding diversity and multiplexing tradeoff curve is presented in the following Theorem.

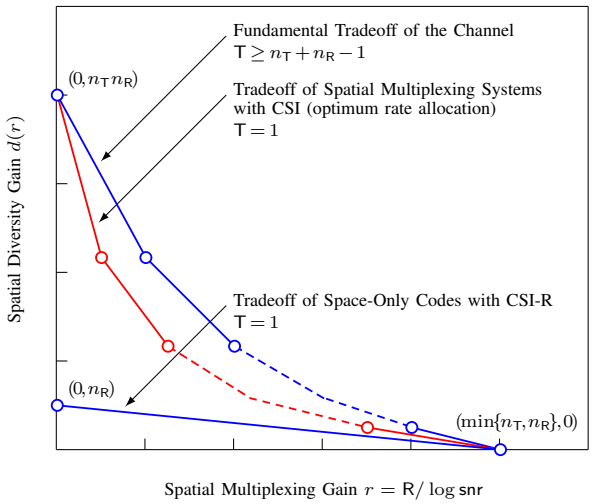


Figure 5.7 Comparison of the tradeoff curves of spatial multiplexing systems with CSI and space-only codes with CSI-R only.

Theorem 5.4. *Consider the spatial multiplexing MIMO system with CSI-R in* [Fos99]. *The fundamental diversity and multiplexing tradeoff curve* $d_\mathsf{V}(r)$ *under the* $n_\mathsf{R} \times n_\mathsf{T}$ *MIMO channels in Definitions* 3.1, 3.2, *and* 3.4 *with* $n_\mathsf{T} \leq n_\mathsf{R}$ *is given by*

$$d_\mathsf{V}(r) = (n_\mathsf{R} - n_\mathsf{T} + 1)\,(1 - r/\min\{n_\mathsf{T}, n_\mathsf{R}\}) \qquad 0 \leq r \leq \min\{n_\mathsf{T}, n_\mathsf{R}\}. \tag{5.89}$$

Proof. The diversity and multiplexing tradeoff of V-BLAST was obtained in [Zhe03, Sec. VII.B] for the uncorrelated Rayleigh MIMO channel model and $n_\mathsf{T} = n_\mathsf{R}$. Using the procedure presented in [Zhe03] and Remark 5.1 along with the high-SNR performance analysis of V-BLAST in [Jia05], the tradeoff curve given in Theorem 5.4 follows. □

Observing (5.87), it is not difficult to show that $d_\mathsf{S}^\star(r) > d_\mathsf{V}(r)$ for all values or r. This result is not surprising, since the spatial multiplexing scheme with CSI exploits the channel information at the transmitter to select only the best eigenmodes and perform the optimal rate allocation among them, whereas the V-BLAST architecture must rely blindly on all eigenchannels. In fact, this is precisely the reason why the tradeoff curves for spatial multiplexing with CSI and a uniform rate allocation and for V-BLAST coincide in the high multiplexing gain regime, i.e., when r is such that $\kappa^\star(r) = \min\{n_\mathsf{T}, n_\mathsf{R}\}$. In this particular situation, the spatial multiplexing scheme with CSI does not use the CSI at the transmitter for either discarding substreams or

allocating the target rate optimally.

5.8 Conclusions and Publications

This chapter addresses the study of the fundamental tradeoff of MIMO systems when not only the receiver but also the transmitter has access to the channel matrix. First we show that the fundamental tradeoff is not altered by channel knowledge at the transmit side, as long as the duration of the encoding blocks satisfies $\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1$. The chapter then concentrates on the analysis of spatial multiplexing MIMO systems with CSI, a family of transceiver schemes which has been subject to study in the literature owing to its simplicity and capacity achieving capabilities. These systems transmit independent symbols through the MIMO channel eigenmodes. The fundamental tradeoff of each of these substreams is determined and allows the formulation of a fundamental tradeoff optimal rate allocation strategy for linearly combining different parallel channels as done by most of the spatial multiplexing schemes. The fact that channel knowledge at the transmitter does not increase the diversity versus multiplexing tradeoff of the MIMO channel was highly predictable. However, what was not known so far, is how much loss would simple spatial multiplexing schemes with CSI incur with respect to the fundamental limits of the channel and how much would they be able to benefit from channel knowledge at the transmitter. Precise answer to these two questions has been given in this chapter comparing the fundamental tradeoff of spatial multiplexing schemes with the fundamental tradeoff of the channel, the tradeoff of space only codes and V-BLAST.

The main results contained in this chapter regarding the diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI and linear MIMO transceivers have been published in two conference papers and one journal paper:

[PZ04] A. Pages-Zamora, J. R. Fonollosa, and L. G. Ordóñez, "Diversity and multiplexing tradeoff of beamforming for MIMO channels", *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 536–540, July 2004.

[Ord05a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "Diversity and multiplexing tradeoff of multiple beamforming in MIMO channels", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 1808–1812, 2005.

[Ord08a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "Diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI", *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2959–2975, July. 2008.

5.A Appendix: Individual Diversity and Multiplexing Tradeoff

5.A.1 Exponent of the Individual Pairwise Error Probability

Proof. In this appendix we derive the exponent of bound on the pairwise error probability of the kth substream given in (5.36):

$$\mathsf{PEP}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1 \Big| p_k > 0\right) + \mathsf{Pr}(p_k = 0) \tag{5.90}$$

where $\mathsf{Pr}(p_k = 0)$ denotes the probability of not transmitting power through the kth channel eigenmode with the capacity-achieving waterfilling. Recall that p_k is given by (see (5.27))

$$p_k = \left(\mu - \lambda_k^{-1}\right)^+ \tag{5.91}$$

where the water level μ is chosen such that

$$\sum_{k=1}^{\kappa} p_k = \sum_{k=1}^{\kappa} \left(\mu - \lambda_k^{-1}\right)^+ = \mathsf{snr}. \tag{5.92}$$

Hence, $p_k = 0$, if $(\mu - \lambda_k^{-1}) \leq 0$ for the water level μ calculated as in (5.92) with $\kappa = k$, then $\mathsf{Pr}(p_k = 0)$ is given by

$$\mathsf{Pr}\left(p_k = 0\right) = \mathsf{Pr}\left((k-1)\lambda_k^{-1} - \sum_{i=1}^{k} \lambda_i^{-1} \geq \mathsf{snr}\right) \tag{5.93}$$

This probability can be upper-bounded as

$$\begin{aligned} \mathsf{Pr}\left(p_k = 0\right) &\leq \mathsf{Pr}\Big(\lambda_k^{-1} \geq \frac{\mathsf{snr}}{k-1}\Big) \\ &\doteq \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{-1}\right). \end{aligned} \tag{5.94}$$

Now, let us assume that $\tilde{\kappa}$ substreams are transmitted with nonzero power, such that $k < \tilde{\kappa} \leq \kappa$. Then, the power allocated to the kth substream is

$$p_k = \frac{\mathsf{snr}}{\tilde{\kappa}} + \frac{1}{\tilde{\kappa}} \sum_{i=1}^{\tilde{\kappa}} \lambda_i^{-1} - \lambda_k^{-1} \tag{5.95}$$

and the first term in (5.90) satifies

$$\begin{aligned} \mathsf{PEP}^{(k)}(\mathsf{R}_k | p_k > 0) &\dot{\leq} \mathsf{Pr}\Big(\frac{p_k\lambda_k}{2}\mathsf{snr}^{-r_k} < 1 \Big| p_k > 0\Big) \\ &= \mathsf{Pr}\left(\left(\frac{\mathsf{snr}}{\tilde{\kappa}}\lambda_k + \frac{1}{\tilde{\kappa}}\sum_{i=1}^{\tilde{\kappa}} \frac{\lambda_k}{\lambda_i} - 1\right)\mathsf{snr}^{-r_k} < 1\right) \\ &\leq \mathsf{Pr}\Big(\frac{\mathsf{snr}}{\tilde{\kappa}}\lambda_k < \mathsf{snr}^{r_k}\Big) \\ &\doteq \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k - 1}\right). \end{aligned} \tag{5.96}$$

Finally, combining (5.94) and (5.96), it follows that

$$\mathsf{PEP}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right) \tag{5.97}$$

since $0 \leq r_k \leq 1$ and this completes the proof. □ □

5.A.2 Proof of Theorem 5.1

Proof. Theorem 5.1 presents the diversity and multiplexing tradeoff of the the k^{th} channel eigenmode, assuming an uncorrelated flat fading Rayleigh channel. In Section 5.5.1 we show that

$$\mathrm{P}_e^{(k)}(\mathsf{R}_k) \doteq \mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1}) \tag{5.98}$$

where λ_k is the k^{th} ordered eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$. Thus, the diversity and multiplexing tradeoff can be obtained as

$$d_{\mathsf{S}}^{(k)}(r) = -\lim_{\mathsf{snr}\to\infty} \frac{\log \mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1})}{\log \mathsf{snr}} = \lim_{\lambda\to 0} \frac{\log \mathsf{Pr}(\lambda_k \leq \lambda^{1-r_k})}{\log \lambda} \tag{5.99}$$

since $0 \leq r_k \leq 1$, and the proof of Theorem 5.1 reduces to obtaining the exponent[8] of the marginal pdf of the k^{th} ordered eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$. Using the first order Taylor expansions of the marginal cdf of the kth ordered eigenvalue in Theorem 2.4 for the uncorrelated Rayleigh, the semicorrelated Rayleigh, and the uncorrelated Rician MIMO channel models (see Section 3.3 for details), the exponent of $\mathsf{Pr}(\lambda_k \leq \lambda)$ is given by

$$\mathsf{Pr}(\lambda_k \leq \lambda) \doteq \lambda^{(n_\mathsf{T}-k+1)(n_\mathsf{R}-k+1)}. \tag{5.101}$$

Now, returning to the equivalence presented in (5.99), we have that

$$\mathsf{P_e}^{(k)}(\mathsf{R}_k) \doteq \mathsf{snr}^{-(n_\mathsf{T}-k+1)(n_\mathsf{R}-k+1)(1-r_k)} \tag{5.102}$$

which completes the proof of the theorem. □

5.A.3 Proof of Corollary 5.1.1

Proof. In this proof we derive the exponent of the error probability of spatial multiplexing MIMO system with a general power allocation satisfying the condition in (5.42). Since this proof is

[8] In this appendix the exponent is defined as $\lambda_k \to 0$ and $g(\lambda_k) \doteq h(\lambda_k)$ denotes

$$\lim_{\lambda_k\to 0} \frac{g(\lambda_k)}{\log \lambda_k} = \lim_{\lambda_k\to 0} \frac{h(\lambda_k)}{\log \lambda_k}. \tag{5.100}$$

strongly based on the procedure used for the capacity-achieving waterfilling power allocation in Section 5.5.1, some repetitive parts are omitted.

The individual error probability of the kth substream, $\mathsf{P_e}^{(k)}(\mathsf{R}_k)$ with the general power allocation described in Corollary 5.1.1 can be lower-bounded by the outage probability obtained when allocating all available power to the kth substream:

$$\mathsf{P_e}^{(k)}(\mathsf{R}_k) \geq \mathsf{P_{out}}^{(k)}(\mathsf{R}_k) \dot{\geq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right) \tag{5.103}$$

since $p_k \leq \mathsf{snr}$, due to the short-term power constraint. In addition, $\mathsf{P_e}^{(k)}(\mathsf{R}_k)$ can be upper-bounded as (see Section 5.5.1)

$$\begin{aligned}
\mathsf{PEP}^{(k)}(\mathsf{R}_k) &\leq \mathsf{Pr}\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{snr}^{-r_k/2} < 1\right) \\
&= \mathsf{Pr}\Big(\frac{p_k\lambda_k}{2}\mathsf{snr}^{-r_k} < 1\Big|p_k > \phi_k\mathsf{snr}\Big)\left(1 - \mathsf{Pr}\left(p_k \leq \phi_k\mathsf{snr}\right)\right) \\
&+ \mathsf{Pr}\Big(\frac{p_k\lambda_k}{2}\mathsf{snr}^{-r_k} < 1\Big|p_k \leq \phi_k\mathsf{snr}\Big)\mathsf{Pr}\left(p_k \leq \phi_k\mathsf{snr}\right) \\
&\leq \mathsf{Pr}\Big(\frac{p_k\lambda_k}{2}\mathsf{snr}^{-r_k} < 1\Big|p_k = \phi_k\mathsf{snr}\Big) + \mathsf{Pr}\left(p_k \leq \phi_k\mathsf{snr}\right) \\
&\leq \mathsf{Pr}\left(\phi_k\lambda_k\mathsf{snr}^{-r_k+1} < 1\right) + \mathsf{Pr}\left(p_k \leq \phi_k\mathsf{snr}\right) \\
&\dot{\leq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right)
\end{aligned} \tag{5.104}$$

where we have used that the power allocation satisfies the condition in (5.42).

$$\mathsf{Pr}\left(p_k \leq \phi_k\mathsf{snr}\right) \dot{\leq} \mathsf{P_{out}}^{(k)}(\mathsf{R}_k) \dot{\leq} \mathsf{Pr}\left(\lambda_k \leq \mathsf{snr}^{r_k-1}\right). \tag{5.105}$$

Finally, combining (5.103) and (5.104), it follows that

$$\mathsf{P_e}^{(k)}(\mathsf{R}_k) \doteq \mathsf{Pr}(\lambda_k \leq \mathsf{snr}^{r_k-1}) \tag{5.106}$$

which coincides with the exponent derived for the capacity-achieving spatial multiplexing MIMO system in Theorem 5.1 and this completes the proof. ss □

5.B Proof of Proposition 5.1

Proof. We want to prove that power allocation in (5.48) satisfies the condition in (5.42). Since the exponent of the error probability is given by (see Corollary 5.1.1)

$$d_{\mathsf{S}}^{(k)}(r_k) = d_k(1 - r_k) \tag{5.107}$$

where $d_k = (n_{\mathsf{T}} - k + 1)(n_{\mathsf{R}} - k + 1)$ and $0 \leq r_k \leq 1$, we have to show that

$$\delta_k = \lim_{\mathsf{snr}\to\infty} -\frac{\log \mathsf{Pr}\left(p_{\mathsf{mse},k} \leq \phi_k \mathsf{snr}\right)}{\log \mathsf{snr}} > d_k \tag{5.108}$$

where ϕ_k is a strictly positive constant. The individual power p_k is maximum when the weakest $\kappa - k$ substreams are discarded, i.e., $\{p_k\}_{i=k+1}^{\kappa} = 0$, and, thus, we can assume without loss of generality that only $k \leq \kappa$ substreams are transmitted with nonzero power. Then, $\mathsf{Pr}\left(p_{\mathsf{mse},k} \leq \phi_k \mathsf{snr}\right)$ is given by

$$\mathsf{Pr}\left(p_{\mathsf{mse},k} \leq \phi_k \mathsf{snr}\right) = \mathsf{Pr}\left(\frac{(\omega_k/\lambda_k)^{1/2}}{\sum_{i=1}^{k}(\omega_i/\lambda_i)^{1/2}}\left(\mathsf{snr} + \sum_{i=1}^{k}\lambda_i^{-1}\right) - \lambda_k^{-1} \leq \phi_k \mathsf{snr}\right) \tag{5.109}$$

$$\begin{aligned} &= \mathsf{Pr}\Bigg((\omega_k\lambda_k)^{-1/2}\sum_{i=1}^{k}(\omega_i/\lambda_i)^{1/2} - \sum_{i=1}^{k}\lambda_i^{-1} \\ &\qquad \geq \mathsf{snr}\Big(1 - \phi_k(\omega_k/\lambda_k)^{-1/2}\sum_{i=1}^{k}(\omega_i/\lambda_i)^{1/2}\Big)\Bigg). \end{aligned} \tag{5.110}$$

Observing that

$$\phi_k(\omega_k/\lambda_k)^{-1/2}\sum_{i=1}^{k}(\omega_i/\lambda_i)^{1/2} \leq \phi_k \sum_{i=1}^{k}(\omega_i/\omega_k)^{1/2} \tag{5.111}$$

since $\lambda_k/\lambda_i \leq 1$ for $i = 1, \ldots, k$, $\mathsf{Pr}\left(p_{\mathsf{mse},k} \leq \phi_k \mathsf{snr}\right)$ can be upper-bounded as

$$\mathsf{Pr}\left(p_k \leq \phi_k \mathsf{snr}\right) \leq \mathsf{Pr}\left(\lambda_k^{-1/2}\sum_{i=1}^{k-1}(\omega_i/\lambda_i)^{1/2} \geq \mathsf{snr}\Big(1 - k\phi_k\sum_{i=1}^{k}\left(\omega_i/\omega_k\right)^{1/2}\Big)\right) \tag{5.112}$$

$$\leq \mathsf{Pr}\left(\lambda_k\lambda_{k-1} \leq \left(\frac{k-1}{1 - k\phi_k\sum_{i=1}^{k}\left(\omega_i/\omega_k\right)^{1/2}}\right)^2 \mathsf{snr}^{-2}\right) \tag{5.113}$$

whenever $\phi_k < 1/\left(k\sum_{i=1}^{k}\left(\omega_i/\omega_k\right)^{1/2}\right)$. Hence, it follows that

$$\delta_k \geq \lim_{\mathsf{snr}\to\infty} \frac{\log \mathsf{Pr}\left(\lambda_k\lambda_{k-1} \leq \mathsf{snr}^{-2}\right)}{\log \mathsf{snr}^{-1}}. \tag{5.114}$$

Then, by defining $\eta_k = \lambda_{k-1}\lambda_k$, we can rewrite (5.114) as

$$d_{\mathsf{mse},k} \geq \lim_{x\to 0} \frac{\log \mathsf{Pr}\left(\eta_k \leq x^2\right)}{\log x} = \lim_{x\to 0} \frac{\log\left(\int_0^{x^2} f_{\eta_k}(\eta)d\eta\right)}{\log x} \tag{5.115}$$

where $f_{\eta_k}(\eta)$ is the pdf of a product of two random variables and is given by [Roh76, Sec. 4.4, Th. 7]

$$f_{\eta_k}(\eta) = \int_0^{\infty} \frac{1}{x} f_{\lambda_{k-1},\lambda_k}(x, \eta/x)dx. \tag{5.116}$$

We are interested in $f_{\eta_k}(\eta)$ as $\eta \to 0$ and, thus, we only need to derive the joint pdf $f_{\lambda_{k-1},\lambda_k}(\lambda_{k-1}, \lambda_k)$ as $\lambda_k \to 0$. Using the same procedure as in the proof of Theorem 2.4, it

can be shown that[9]

$$f_{\lambda_{k-1},\lambda_k}(\lambda_{k-1},\lambda_k) = g(\lambda_{k-1})\lambda_k^{d_k-1} + o(\lambda_k^{d_k-1}) \tag{5.117}$$

where $g(\lambda_{k-1})$ is a function of λ_{k-1}. Then, substituting back this result in the expression of $f_{\eta_k}(\eta)$ in (5.116), it follows that

$$f_{\eta_k}(\eta) = a\eta^{d_k-1} + o(\eta^{d_k-1}) \tag{5.118}$$

where a is a fixed constant in terms of η. Finally, the exponent of $\mathsf{Pr}\left(p_{\mathsf{mse},k} \leq \phi_k \mathsf{snr}\right)$ can be bounded as

$$\delta_k \geq \lim_{x\to 0} \frac{\log\left(a\int_0^{x^2} \eta^{d_k-1} d\eta\right)}{\log x} = \lim_{\eta\to 0} \frac{\log\left(\eta^{2d_k}\right)}{\log \eta} = 2d_k > d_k \tag{5.119}$$

and this completes the proof. □

[9] We say that $f(x) = o(g(x))$, $g(x) > 0$, if $f(x)/g(x) \to 0$ as $x \to 0$ [Bru81, eq. (1.3.1)].

6

Conclusions and Future Work

This dissertation has focused on the evaluation of the performance limits of the spatial multiplexing MIMO systems that result from transmitting independent substreams through the strongest channel eigenmodes when perfect channel state information is available at both sides of the link. Analytical studies of these schemes from a communication- and an information-theoretical point-of-view have enlightened the implications of the tradeoff beween spatial multiplexing and diversity gain, or, in other words, the tradeoff between transmission rate and reliability. For this purpose, the probabilistic characterization of the eigenvalues of Wishart, Pseudo-Wishart and quadratic form distributions has been proven to be critical. In particular, the performance of spatial multiplexing MIMO systems with CSI demanded the evaluation of probabilities associated with one or several of the eigenvalues in some specific order. In this work we have developed a unified formulation that can, not only fill the gap of the currently unknown results, but even more importantly, provide a solid framework for the understanding and direct derivation of all the already existing results.

6.1 Summary of Results

Let us give a detailed summary of the main results contained in this PhD thesis. The open issues or lines for future research that have been singled out during the elaboration of the present dissertation are also highlighted.

Chapter 2

In Chapter 2 we have proposed a general formulation that unifies the probabilistic characterization of the eigenvalues of Hermitian random matrices with a specific structure. Based on a unified expression for the joint pdf of the ordered eigenvalues, we obtained:

(i) the joint cdf of the ordered eigenvalues,
(ii) the marginal cdf's of the ordered eigenvalues,
(iii) the marginal pdf's of the ordered eigenvalues,
(iv) the cdf of the maximum weighted ordered eigenvalue, and
(v) the first order Taylor expansions of (ii), (iii), and (iv),

where (ii), (iii), and (iv) follow as simple particularizations of (i). In addition we have also considered the unordered eigenvalues and we have derived:

(vi) the joint cdf of a set of unordered eigenvalues,
(vii) the joint pdf of a set of unordered eigenvalues, and
(viii) the marginal cdf of a single unordered eigenvalue.

The former distributions and results have been particularized for uncorrelated and correlated central Wishart, correlated central Pseudo-Wishart, and uncorrelated noncentral Wishart matrices avoiding the non-convenient series expansions in terms of zonal polynomials. Central quadratic forms have been also addressed, although, in this case, we have not been able to handle with the infinite series. In this respect, there are some issues which are still open:

(i) Derivation of the first order Taylor expansion of a more general class of Hermitian random matrices than the one formalized in Assumption 2.2. In particular, this will allow to deal also with Pseudo-Wishart and quadratic forms distributions.
(ii) Evaluation of the joint and marginal distributions of the ordered eigenvalues of complex quadratic forms distributions.

(iii) Evaluation of the joint and marginal distributions of the ordered eigenvalues of complex correlated Wishart distributions, since they are not included in the proposed unified formulation.

The joint analysis of a general class of distributions presented in this chapter settles the basis for a unified framework in the performance analysis of MIMO systems. Specifically, in this dissertation our results are applied to investigate the performance limits of spatial multiplexing MIMO systems with CSI. Our unified framework could, however, be also useful in other communication and signal processing applications.

Chapter 3

The main contribution of Chapter 3 is the analytical performance analysis of spatial multiplexing MIMO systems in Rayleigh and Rician MIMO channels. More exactly, in this chapter we have obtained analytical expressions to calculate:

(i) the exact individual average BER and its corresponding parameterized high-SNR characterization when using a fixed power allocation,

(ii) the exact individual outage probability and its corresponding parameterized high-SNR characterization when using a fixed power allocation,

(iii) high-SNR parameterized upper and lower bounds for the individual average BER when using a non-fixed power allocation,

(iv) high-SNR parameterized upper and lower bounds for the individual outage probability when using a non-fixed power allocation,

(v) the global average BER and its corresponding parameterized high-SNR characterization, and

(vi) different global outage probability measures.

These general results have been applied to analyze the performance of linear MIMO transceivers existing in the literature with adaptive linear precoder but fixed number of data symbols and fixed constellations. In particular, using the unifying formulation in [Pal03], we have been able to investigate the high-SNR global average performance of a wide family of practical designs:

(vii) diagonal schemes with a fixed power allocation,

(viii) diagonal schemes with a non-fixed power allocation, and

(ix) non-diagonal schemes with a non-fixed power allocation.

Chapter 4

The linear MIMO transceiver design has been addressed in the literature with the typical underlying assumption that the number of data symbols to be transmitted per channel use is chosen beforehand. In Chapter 4 we have proved that, under this assumption, the diversity order of any linear MIMO transceiver is at most driven by that of the weakest channel eigenmode employed, which can be far from the diversity intrinsically provided by the channel. Based on this observation, we have fixed the rate (instead of the number of data symbols) and we have optimized the number of substreams and constellations jointly with the linear precoder. This procedure implies only an additional optimization stage upon the classical design which suffices to extract the full diversity of the channel. Since the ultimate performance of a communication system is given by the BER, we have focused on the minimum BER design. The implications of the proposed optimization have been then illustrated by means of analytical performance expressions of the minimum BER linear MIMO transceiver with fixed and with adaptive number of substreams, including:

(i) upper and lower bounds for the global average BER, and

(ii) high-SNR parameterized upper and lower bounds for the global average BER.

Chapter 5

Chapter 5 addresses the study of the fundamental tradeoff of MIMO systems when not only the receiver but also the transmitter has access to the channel matrix. First, it has been shown that the fundamental tradeoff is not altered by channel knowledge at the transmit side, as long as the duration of the encoding blocks satisfies $\mathsf{T} \geq n_{\mathsf{T}} + n_{\mathsf{R}} - 1$. Then we have concentrated on the analysis of spatial multiplexing MIMO systems with CSI ($\mathsf{T} = 1$ is implicitly assumed) and we have derived:

(i) the individual diversity and multiplexing tradeoff of the channel eigenmodes,

(ii) the global diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI and a uniform rate allocation, and

(iii) the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI and optimum rate allocation.

The fact that channel knowledge at the transmitter does not increase the diversity versus multiplexing tradeoff of the MIMO channel was highly predictable. However, what was not known so far, is how much loss would simple spatial multiplexing MIMO schemes with CSI incur with respect to the fundamental limits of the channel and how much would they be able to benefit from channel knowledge at the transmitter. Precise answer to these two questions has been given in this chapter comparing the fundamental tradeoff of spatial multiplexing schemes with the fundamental tradeoff of the channel, the tradeoff of space only codes and V-BLAST.

6.2 Summary of Key Insights

This dissertation has focused on the evaluation of the performance limits of the spatial multiplexing MIMO systems with CSI with the final aim of understanding the implications of the tradeoff beween spatial multiplexing and diversity gain in those schemes. In the following we present the key insights extracted from the two perspectives adopted in this dissertation. As a conclusion a new line of research to merge both approaches is briefly commented.

Communication-Theoretic Analysis

First, we have considered a rather practical setup, in which the transmission rate is fixed (practical modulations are assumed) and the reliability of the communication is measured by the average BER vs. SNR curves. In the high-SNR regime the performance have been characterized in terms of the diversity and the array gain. The following key insights have been observed:

(i) The global diversity gain is limited by that of the worst eigenmode use, independently of unitary transformations and power allocations.

(ii) Channel-dependent power allocation policies can only possibly increase the array gain but not the diversity gain of the system.

(iii) Pre-equalization of the equivalent channels by means of a rotation of the data symbols before transmission through the channel eigenmodes increases the array gain but not the diversity gain of the system.

This enlightens that fixing a priori the number of independent data streams to be transmitted, a very common assumption in the linear transceiver design literature, inherently limits the average BER performance of the system. However, an additional optimization of the number of substreams suffices to obtain the full diversity of the channel.

Information-Theoretic Analysis

In order to take also into consideration the rate accommodation capabilities of spatial multiplexing MIMO systems, we have next analyzed their performance in terms of the diversity and multiplexing tradeoff. The following key insights have been observed:

(i) The fundamental tradeoff of the channel is not achieved due to the lack of coding between substreams. The outage event is dictated by the individual outages of the used channel eigenmodes.

(ii) The fundamental tradeoff of spatial multiplexing is achieved whenever the number of substreams is optimally chosen and the transmission rate is optimally allocated among them.

(iii) Channel-dependent power allocation policies among the established substreams do not modify the diversity and multiplexing tradeoff.

(iv) Pre-equalization of the equivalent channels by means of a rotation of the data streams before transmission suffers from a degradation with respect to the fundamental tradeoff due to the impossibility of optimally allocating the data rate.

Future Work: Unified Analysis

The perspectives under which spatial multiplexing MIMO systems have been analyzed in this thesis are distinct in their motivations and implications. On the one hand, the communication-theoretic analysis gives clear answers on how the reliability of the communication is modified by an SNR increase when the data rate is fixed. On the other hand, the information-theoretic analysis indicates how the reliability of the communication is decreased as the data rate approaches the ergodic capacity of the channel. By means of the diversity and multiplexing tradeoff, a more global point-of-view is given, since it shows how an SNR increase can be used to lower the error probability or to increase the transmission rate or a combination of the two. However, although

elegant, this formulation is also quite loose and observations extracted from the tradeoff curve are difficult to be translated into implications in the error probability curve of a particular scheme with a finite SNR. For instance, the diversity and multiplexing tradeoff curve suggest for the fixed rate case that the maximum diversity order can be achieved by beamforming over the strongest channel eigenmode. We know, to the contrary, that for finite SNR, schemes with lower error probabilities exist. On the other extreme of the curve, i.e., for a fixed error probability, the high-SNR scaling of the ergodic capacity is shown to be achieved by transmitting independent substreams though all the channel eigenmodes but nothing is said about the optimal power allocation among them. In this respect, a combination of both analysis would be preferable. This would require to measure more exactly the error probability, for instance in terms of array and diversity gain, and perhaps redefine the multiplexing gain. Several attempts to deal with the limitations of the diversity and multiplexing tradeoff framework, such as the non-asymptotic framework of [Nar06] or the throughput-reliability tradeoff of [Aza07] have been recently developed. However, this unified analysis is left as a future line of research.

References

[AA68] A. Al-Ani, "On the distribution of the ith largest latent root under null hyphotheses concerning complex multivariate normal populations", *Mimeo Series 145, Purdue University*, 1968.

[Abd02] A. Abdi, and M. Kaveh, "A space–time correlation model for multielement antenna systems in mobile fading channels", *IEEE J. Select. Areas Commun.*, vol. 20, no. 3, pp. 550–561, Apr. 2002.

[Abr72] M. Abramowitz, and I. A. Stegun, *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*, Dover, 1972.

[Ait83] A. C. Aitken, *Determinants and Matrices*, Greenwood Press, 1983.

[Ala98] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[Alf04a] G. Alfano, A. Lozano, A. Tulino, and S. Verdu, "Mutual information and eigenvalue distribution of MIMO Ricean channels", *Proc. IEEE Int. Symp. Inform. Theory and its Appl. (ISITA)*, Oct. 2004.

[Alf04b] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, "Capacity of MIMO channels with one-sided correlation", *Proc. IEEE Int. Symp. Spr. Spect. Techn. and Appl. (ISSSTA)*, pp. 515–519, Sep. 2004.

[Alf06] G. Alfano, A. Tulino, A. Lozano, and S. Verdú, "Eigenvalue statistics of finite-dimensional random matrices for MIMO wireless communications", *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 4125–4129, Jun. 2006.

[And84] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, $2^{\underline{\text{nd}}}$ ed., 1984.

[And00a] J. B. Andersen, "Antenna arrays in mobile communications: Gain, diversity, and channel capacity", *IEEE Antennas Propagat. Mag.*, vol. 42, no. 2, pp. 12–16, Apr. 2000.

[And00b] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends", *IEEE J. Select. Areas Commun.*, vol. 18, no. 11, pp. 2172–2178, Nov. 2000.

[Aza07] K. Azarian, and H. E. Gamal, "The throughput-reliability tradeoff in block-fading MIMO channels", *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 488–501, Feb. 2007.

[Ben99] S. Benedetto, and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*, New York: Kluwer Academic, 1999.

[Big98] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects", *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[Big01] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas", *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 1273–1289, May 2001.

[Big07] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor (eds.), *MIMO Wireless Communications*, Cambridge Univ. Press, 2007.

[Böl00] H. Bölcskei, and A. J. Paulraj, "Performance of space-time codes in the presence of spatial fading correlation", *Proc. Asilomar Conf. Signals, Systems, and Computers*, pp. 687–693, Oct. 2000.

[Böl02a] H. Bölcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems", *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.

[Böl02b] H. Bölcskei, and A. J. Paulraj, "Multiple-input multiple-output (MIMO) wireless systems", J. Gibson (ed.), *The Communications Handbook*, pag. 90.1–90.14, CRC Press, 2$^{\underline{\text{nd}}}$ ed., 2002.

[Böl03] H. Bölcskei, M. Bergmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM", *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 427–439, Apr. 2003.

[Bra46] A. Bravais, "Analyse mathématique sur les probabilités des erreurs de situation d'un point", *Mém. prés. Acad. Sci., Paris*, vol. 9, pp. 255–332, 1846.

[Bru81] N. G. Bruijn, *Asymptotic Methods in Analysis*, Dover Publications Inc., 3$^{\underline{\text{rd}}}$ ed., 1981.

[Bur02] G. Burel, "Statistical analysis of the smallest singular value in MIMO transmission systems", *Proc. WSEAS Int. Conf. on Signal, Speech and Image Processing (ICOSSIP)*, Sep. 2002.

[Cai99] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.

[Can87] P. Cantrell, and A. Ojha, "Comparison of generalized Q- function algorithms", *IEEE Trans. Inf. Theory*, vol. 33, no. 4, pp. 591–596, July 1987.

[Car83] M. Carmeli, *Statistical Theory and Random Matrices*, vol. 74 of *Monographs and textbooks in pure and applied mathematics*, Marcel Dekker, 1983.

[Cat02] S. Catreux, V. Erceg, D. Gesbert, and R. W. Heath, Jr., "Adaptive modulation and MIMO coding for broadband wireless data networks", *IEEE Commun. Magazine*, vol. 40, pp. 108–115, June 2002.

[Cha06] W. Chang, S.-Y. Chung, and Y. H. Lee, "Diversity-multiplexing tradeoff in rank-deficient and spatially correlated MIMO channels", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 1144–1148, July 2006.

[Che04] Y. Chen, and C. Tellambura, "Distribution functions of selection combiner output in equally correlated Rayleigh, Rician, and Nakagami-m fading channels", *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1948–1956, Nov. 2004.

[Chi00] D. Chizhik, F. R. Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels", *IEEE Comm. Lett.*, vol. 4, no. 11, pp. 337–339, Nov. 2000.

[Chi03] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels", *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[Chi05] M. Chiani, M. Z. Win., and A. Zanella, "On optimum combining of m-PSK signals with unequal-power interferers and noise", *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 44–47, Jan. 2005.

[Cho02] 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, July 2002.

[Chr64] J. G. Christiano, and J. E. Hall, "On the n-th derivative of the j-th order determinant", *Mathem. Mag.*, vol. 37, no. 4, pp. 215–217, Sep. 1964.

[Chu01] S. T. Chung, and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view", *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1561–1571, Sep. 2001.

[Chu02] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading", *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[Con63] A. G. Constantine, "Some noncentral distributions problems in multivariate analysis", *Ann. of Math. Statist.*, vol. 34, pp. 1270–1285, 1963.

[Cor07] P. Coronel, and H. Bölcskei, "Diversity-multiplexing tradeoff in selective-fading MIMO channels", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 2841–2845, June 2007.

[Cov91] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[Dav79] A. W. Davis, "Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate disrtribution theory", *Ann. Inst. Statist. Math.*, vol. 31, pp. 465–485, 1979.

[Dav80] A. W. Davis, "Invariant polynomials with two matrix arguments extending the zonal polynomials", *Multivariate Analysis V (ed. P. R. Krishnaiah)*, pp. 287–299, 1980.

[Dig03] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading", *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[Din03a] Y. Ding, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization", *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2410–2423, Sep. 2003.

[Din03b] Y. Ding, T. N. Davidson, and K. M. Wong, "On improving the BER performance of rate-adaptive block-by-block transceivers with application to DMT", *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2003.

[Dri99] P. F. Driessen, and G. F. Foschini, "On the capacity formula for multiple input–multiple output wireless channels: A geometric approach", *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, Feb. 1999.

[Fis15] R. A. Fisher, "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population", *Biometrika*, vol. 10, pp. 507–521, 1915.

[Fis39] R. A. Fisher, "The sampling distribution of some statistics obtained from non-linear equations", *Ann. Eugenics*, vol. 9, pp. 238–249, 1939.

[For98] G. D. Forney, Jr., and G. Ungerboeck, "Modulation and coding for linear Gaussian channels", *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[Fos96] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[Fos98] G. Foschini, and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[Fos99] G.J. Foschini, G.D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays", *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.

[Gal68] R. B. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, 1968.

[Gar97] J. A. Díaz García, and R. Gutierrez Jáimez, "Wishart and Pseudo-Wishart distributions and some applications to shape theory", *J. Multivar. Anal.*, vol. 63, pp. 73–87, Oct. 1997.

[Gar05] L. M. Garth, P. J. Smith, and M. Shafi, "Exact symbol error probabilities for SVD transmission of BPSK data over fading channels", *Proc. IEEE Int. Conf. Communications (ICC)*, vol. 4, pp. 2271– 2276, May 2005.

[Gho02] M. Ghosh, and B. K. Sinha, "A simple derivation of the Wishart distribution", *The Americ. Static.*, vol. 56, no. 2, pp. 100, May 2002.

[Gol99] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture", *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[Goo63] N. R. Goodman, "Statistical analysis based on certain multivariate complex Gaussian distribution", *Ann. of Math. Statist.*, vol. 34, no. 1, pp. 152–176, Mar. 1963.

[Gra53] F. Gray, "Pulse code communications", U.S. Patent 2 632 058, Mar. 17, 1953.

[Gra00] I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, $6^{\underline{\text{th}}}$ ed., 2000.

[Gra02] A. J. Grant, "Rayleigh fading multiple-antenna channels", *EURASIP J. Appl. Signal Processing*, vol. 2002, no. 3, pp. 316–329, Mar. 2002.

[Gra05] A. J. Grant, "Performance analysis of transmit beamforming", *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 738–744, Apr. 2005.

[Gro89] K. I. Gross, and D. S. P. Richards, "Total positivity, spherical series, and hypergeometric functions of matrix argument", *J. Approx. Theory*, vol. 59, no. 2, pp. 224–246, 1989.

[Gup00] A.K. Gupta, and D. K. Nagar, *Matrix Variate Distributions*, Monographs and Surveys in Pure and Applied Mathematics, Chapman & Hall/CRC, 2000.

[Gut00] R. Gutíerrez, J. Rodríguez, and J. A. Sáez, "Approximation of hypergeometric functions with matricial argument through their development in series of zonal polynomials", *Electron. Trans. Num. Anal.*, vol. 11, pp. 121–130, 2000.

[Han70] E. J. Hannan, *Multiple Time Series*, John Wiley & Sons, 1970.

[Has02] B. Hassibi, and B. M. Hochwald, "High-rate codes that are linear in space and time", *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804–1824, July 2002.

[Hea00] R. W. Heath, Jr., and A. Paulraj, "Switching between spatial multiplexing and transmit diversity based on constellation distance", *Proceedings of the Allerton Conf. Commun. Control, Comput.*, Oct. 2000.

[Hea01] R. W. Heath, Jr., and A. Paulraj, "Characterization of MIMO channels for spatial multiplexing systems", *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 2, pp. 591–595, June 2001.

[Her55] C. S. Herz, "Bessel functions of matrix argument", *Ann. of Math.*, vol. 61, no. 474–523, 1955.

[Hon90] M. L. Honig, K. Steiglitz, and B. Gopinath, "Multi channel signal processing for data communications in the presence of crosstalk", *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 551–558, Apr. 1990.

[Hor90] R. A. Horn, and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[Hor91] R. A. Horn, and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[Hsu39a] P. L. Hsu, "A new proof of the joint product moment distribution", *Proc. Camb. Phil. Soc.*, vol. 35, pp. 336–338, 1939.

[Hsu39b] P. L. Hsu, "On the distribution of the roots of certain determinantal equations", *Annals of Eugenics*, vol. 9, pp. 250–258, 1939.

[Ing33] A. E. Ingham, "An integral which occurs in statistics", *Proc. Camb. Phil. Soc.*, vol. 29, pp. 271–276, 1933.

[Ivr03] M. T. Ivrlac, W. Utschick, and J. A. Nossek, "Fading correlations in wireless MIMO communication systems", *IEEE J. Select. Areas Commun.*, vol. 21, pp. 819–.828, Jun. 2003.

[Jam54] A. T. James, "Normal multivariate analysis and the orthogonal group", *Ann. of Math. Statist.*, vol. 25, no. 1, pp. 40–75, Mar. 1954.

[Jam55a] A. T. James, "A generating function for averages over the orthogonal group", *Proc. Roy. Soc. Lond. Ser. A*, vol. 229, pp. 367–375, 1955.

[Jam55b] A. T. James, "The noncentral Wishart distribution", *Proc. Camb. Roy. Soc. Lond.*, vol. A229, pp. 364–366, 1955.

[Jam60] A. T. James, "The distribution of the latent roots of covariance matrix", *Ann. of Math. Statist.*, vol. 31, no. 1, pp. 151–158, Mar. 1960.

[Jam61a] A. T. James, "The distribution of noncentral means with known covariance the distribution of noncentral means with known covariance", *Ann. of Math. Statist.*, vol. 32, no. 3, pp. 874–882, Sep. 1961.

[Jam61b] A. T. James, "Zonal polynomials of the real positive definite symmetric matrices", *Ann. of Math.*, vol. 74, no. 3, pp. 456–469, Nov. 1961.

[Jam64] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples", *Ann. of Math. Statist.*, vol. 35, no. 2, pp. 475–501, Jun. 1964.

[Jan03] R. A. Janik, and M. A. Nowak, "Wishart and anti-Wishart random matrices", *J. Phys. A: Math. Gen.*, vol. 36, no. 12, pp. 3629–3637, Mar. 2003.

[Jay05] S. K. Jayaweera, and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading", *IEEE Trans. Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.

[Jia05] Y. Jiang, X. Zheng, and J. Li, "Asymptotic performance analysis of V-BLAST", *Proc. IEEE*, vol. 6, pp. 3882–3886, Dec. 2005.

[Jin06] S. Jin, X. Gao, and M. R. McKay, "Ordered eigenvalues of complex noncentral Wishart matrices and performance analysis of SVD MIMO systems", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 1564–1568, July 2006.

[Jin08] S. Jin, M. R. McKay, X. Gao, and I. B. Collings, "MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral Wishart matrices", *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 424–434, Mar. 2008.

[Kan03a] M. Kang, and M. S. Alouini, "Impact of correlation on the capacity of MIMO channels", *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 2623–2627, May 2003.

[Kan03b] 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. 418–425, Apr. 2003.

[Kan04] M. Kang, and M. S. Alouini, "Quadratic forms in complex gaussian matrices and performance analysis of MIMO systems with cochannel interference", *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 418–431, Mar. 2004.

[Kan06a] M. Kang, and M. S. Alouini, "Capacity of correlated MIMO Rayleigh channels", *IEEE Trans. Commun.*, vol. 5, no. 1, pp. 143–155, Jan. 2006.

[Kan06b] M. Kang, and M. S. Alouini, "Capacity of MIMO Rician channels", *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[Ker02] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation", *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[Kha64] C. G. Khatri, "Distribution of the largest or smallest characteristic root under null hypothesis concerning complex multivariate normal populations", *Ann. of Math. Statist.*, vol. 35, no. 4, pp. 1807–1810, Dec. 1964.

[Kha65] C. G. Khatri, "Classical statistical analysis based on certain multivariate complex Gaussian distribution", *Ann. of Math. Statist.*, vol. 36, no. 1, pp. 98–114, Feb. 1965.

[Kha66] C. G. Khatri, "On certain distribution problems based on positive quadratic forms in normal vectors", *Ann. of Math. Statist.*, vol. 37, no. 2, pp. 468–470, Apr. 1966.

[Kha69] C. G. Khatri, "Non-central distributions of the ith largest characteristic roots of three matrices concerning complex multivariate normal populations", *Ann. Inst. Statist. Math.*, vol. 21, no. 1, pp. 23–32, Dec. 1969.

[Kha70] C. G. Khatri, "On the moments of traces of two matrices in three situations for complex multivariate normal populations", *Sankhya Ser. A*, vol. 32, pp. 65–80, 1970.

[Koe06] P. Koev, and A. Edelman, "The efficient evaluation of the hypergeometric function of a matrix argument", *Mathematics of Computation*, vol. 75, no. 254, pp. 833–846, Jan. 2006.

[Kol95] T. Kollo, and D. von Rosen, "Approximating by the wishart distribution", *Ann. Inst. Statist. Math.*, vol. 47, no. 4, pp. 767–783, Dec. 1995.

[Kon96] S. Kondo, and L. B. Milstein, "Performance of multicarrier DS-CDMA systems", *IEEE Trans. Commun.*, vol. 44, pp. 238–246, Feb. 1996.

[Ksh59] A. M. Kshirsagar, "Barlett decomposition and Wishart distribution", *Ann. of Math. Statist.*, vol. 30, no. 1, pp. 239–241, Mar. 1959.

[Ksh72] A. M. Kshirsagar, *Multivariate Analysis*, Marcel Dekker, 1972.

[Las03] J. Lassing, E. G. Ström, E. Agrell, and T. Ottosson, "Computation of the exact bit-error rate of coherent m-ary PSK with Gray code bit mapping", *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1758–1760, Nov. 2003.

[Lee76] K. H. Lee, and D. P. Petersen, "Optimal linear coding for vector channels", *IEEE Trans. Commun.*, vol. COM-24, no. 12, pp. 1283–1290, Dec. 1976.

[Let04] J. Letessier, P. Rostang, and G. Burel, "Performance analysis of maximum-SNR design in Rayleigh fading MIMO channels", *Proc. IEEE Symp. Per. Ind. and Mob. Radio Commun. (PIMRC)*, Sep. 2004.

[Lo99] T. K. Y. Lo, "Maximum ratio transmission", *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.

[Lov05a] D. J. Love, and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems", *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[Lov05b] D. J. Love, and R. W. Heath, Jr., "Multimode precoding for MIMO wireless systems", *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3674–3687, Oct. 2005.

[Lu99] J. Lu, K. B. Letaief, J. C.-I. Chuang, and M. L. Liou, "m-PSK and m-QAM BER computation using signal-space concepts", *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 181–184, Feb. 1999.

[Maa05] A. Maaref, and S. Aïssa, "Closed-form expressions for the outage and ergodic Shannon capacity of MIMO MRC systems", *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1092–1095, July 2005.

[Maa06] A. Maaref, and S. Aïssa, "Eigenvalue distributions of Wishart-type random matrices and error probability analysis of dual maximum-ratio transmission in semicorrelated Rayleigh fading", *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 4130–4136, June 2006.

[Maa07a] A. Maaref, and S. Aïssa, "Eigenvalue distributions of Wishart-type random matrices with application to the performance analysis of MIMO MRC systems", *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2679–2686, July 2007.

[Maa07b] A. Maaref, and S. Aïssa, "Joint and marginal eigenvalue distributions of (non) central Wishart matrices and PDF-based approach for characterizing the capacity statistics of MIMO Rician and Rayleigh fading channels", *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3607–3619, Oct 2007.

[Maa08] A. Maaref, and S. Aïssa, "Capacity of MIMO Rician fading channels with transmitter and receiver channel state information", *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1687–1698, May 2008.

[Mad38] W. A. Madow, "Contribution to the theory of multivariate statistical analysis", *Trans. Amer. Math.*, vol. 44, no. 3, pp. 454–495, Nov. 1938.

[Mag79] J. R. Magnus, and H. Neudecker, "The commutation matrix: Some properties and applications", *Ann. of Statist.*, vol. 7, no. 2, pp. 381–394, Mar. 1979.

[Mal03] R. K. Mallik, "The Pseudo-Wishart distribution and its application to MIMO systems", *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2761–2769, Oct. 2003.

[Mar50] J. I. Marcum, "Table of Q functions", U.S. Air Force RAND Research Memorandum M-339, ASTIA Document AD 1165451, Rand Corp., Santa Monica, CA, Jan. 1950.

[McK05] M. R. McKay, and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels", *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3121–3145, Sep. 2005.

[McK06] M. R. McKay, *Random Matrix Theory Analysis of Multiple Antenna Communications Systems*, PhD Thesis, School of Electrical and Information Engineering, Telecommunications Laboratory, University of Sydney, Oct. 2006.

[McK07] M. R. McKay, A. Grant, and I. B. Collings, "Performance analysis of MIMO-MRC in double-correlated Rayleigh environments", *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 497–507, Mar. 2007.

[McN02] D.P. McNamara, M.A. Beach, and P.N. Fletcher, "Spatial correlation in indoor MIMO channels", *Proc. IEEE Symp. Per. Ind. and Mob. Radio Commun. (PIMRC)*, pp. 290–294, Sep. 2002.

[Mui82] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wisley series in probability and mathematical statistics, John Wiley & Sons, 1982.

[Nar06] R. Narasimhan, "Finite-SNR diversity-multiplexing tradeoff for correlated Rayleigh and Rician MIMO channels", *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3965–3979, Sep. 2006.

[Nee93] F. D. Neeser, and J. L. Massey, "Proper complex random processes with applications to information theory", *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[Nut72] A. H. Nuttall, "Some integrals involving the Q function", Naval Underwater Systems Center, New London, CT, 4297, Apr. 1972.

[Oga53] J. Ogawa, "On the sampling distribution of classical statistics in multivariate analysis", *Osaka Math. J.*, vol. 5, pp. 13–52, 1953.

[Olk54] I. Olkin, and S. N. Roy, "On multivariate distribution theory", *Ann. of Math. Statist.*, vol. 25, no. 2, pp. 329–333, June 1954.

[Ong03] E. N. Onggosanusi, A. M. Sayeed, and B. D. V. Veen, "Efficient signaling schemes for wideband space-time wireless channels using channel state information", *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 1–13, Jan. 2003.

[Ord05a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "Diversity and multiplexing tradeoff of multiple beamforming in MIMO channels", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 1808–1812, Sep. 2005.

[Ord05b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Analytical BER performance in spatial multiplexing MIMO systems", *Proc. IEEE Workshop Signal Process. Advan. Wireless Commun. (SPAWC)*, pp. 460–464, June 2005.

[Ord07a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "On the design of minimum BER linear MIMO transceivers with perfect or partial side channel state information", *Proc. IST Summit on Mobile and Wireless Comm.*, pp. 1–5, July 2007.

[Ord07b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "High SNR analytical performance of spatial multiplexing MIMO systems with CSI", *IEEE Trans. Signal Processing*, vol. 55, no. 11, pp. 5447–5463, Nov. 2007.

[Ord07c] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "On equal constellation minimum BER linear MIMO transceivers", *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. III, pp. 221–224, Apr. 2007.

[Ord08a] L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa, "Diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI", *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2959–2975, July 2008.

[Ord08b] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 3846–3852, May 2008.

[Ord09a] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of Hermitian random matrices with application to the performance analysis of MIMO systems", *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 672–689, Feb. 2009.

[Ord09b] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "Minimum BER linear MIMO transceivers with optimum number of substreams", *accepted in IEEE Trans. Signal Processing*, Jan. 2009.

[Oym02] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels: Insights into diversity-multiplexing tradeoff", *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2002.

[Oym03] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels", *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.

[Oza94] 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.

[Ozc03] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of the 'Kronecker' MIMO radio channel model", *IEE Elec. Lett.*, vol. 30, no. 16, pp. 1209–1210, Aug. 2003.

[Pal03] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization", *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[Pal05a] D. P. Palomar, and S. Barbarossa, "Designing MIMO communication systems: Constellation choice and linear transceiver design", *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3804–3818, Oct. 2005.

[Pal05b] D. P. Palomar, M. Bengtsson, and B. Ottersten, "Minimum BER linear transceivers for MIMO channels via primal decomposition", *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2866–2882, Aug. 2005.

[Pal07] D. P. Palomar, and Y. Jiang, *MIMO Transceiver Design via Majorization Theory*, Now Publishers, Hanover, MA, 2007.

[Pau94] A. Paulraj, and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission / directional reception (DTDR)", U. S. Patent no. 5,345,599, Sep. 1994.

[Pic94] B. Picinbono, "On circularity", *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[Poo06] A. S. Y. Poon, D. N. C. Tse, and R. W. Brodersen, "Impact of scattering on the capacity, diversity, and propagation range of multiple-antenna channels", *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 1087–1100, Mar. 2006.

[Pro01] J. G. Proakis, *Digital Communications*, McGraw-Hill, $4^{\underline{\text{th}}}$ ed., 2001.

[PZ04] A. Pages-Zamora, J. R. Fonollosa, and L. G. Ordóñez, "Diversity and multiplexing tradeoff of beamforming for MIMO channels", *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 536–540, July 2004.

[Ral98] G. G. Raleigh, and J. M. Cioffi, "Spatio-temporal coding for wireless communication", *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–365, Mar. 1998.

[Ras48] G. Rasch, "A functional equation for Wishart's distribution", *Ann. of Math. Statist.*, vol. 19, no. 2, pp. 262–266, June 1948.

[Rat03] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Complex random matrices and Rayleigh channel capacity", *Commun. in Inform. and Syst.*, vol. 3, no. 2, pp. 119–138, Oct. 2003.

[Rat04] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Jacobians and hypergeometric functions in complex multivariate analysis", *Canadian Applied Mathematics Quarterly*, vol. 12, no. 2, pp. 213–239, Summer 2004.

[Rat05a] T. Ratnarajah, "Non-central quadratic forms in complex random matrices and applications", *Proc. IEEE Works. Stat. Signal Process. (SSP)*, pp. 555–559, July 2005.

[Rat05b] T Ratnarajah, and R Vaillancourt, "Complex singular Wishart matrices and applications", *Comp. Math. Applic.*, vol. 50, no. 3-4, pp. 399–411, Aug. 2005.

[Rat05c] T. Ratnarajah, and R. Vaillancourt, "Quadratic forms on complex random matrices and multiple-antenna systems", *IEEE Trans. Inf. Theory*, vol. 51, no. 2976-2984, pp. 2976, Aug. 2005.

[Rat05d] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Complex random matrices and Ricean channel capacity", *Probl. of Inform. Trans.*, vol. 41, no. 1, pp. 1–22, Jan. 2005.

[Rat06] T. Ratnarajah, "Spatially correlated multiple-antenna channel capacity distributions", *IEE Proc. Commun.*, vol. 153, no. 2, pp. 264–271, Apr. 2006.

[Roh76] V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons, 1976.

[Roy39] S. N. Roy, "p-statistics, or some generalizations in analysis of variance appropriate to multivariate problems", *Sankhya*, vol. 61, pp. 15–34, 1939.

[Roy58] S. N. Roy, *Some Aspects of Multivariate Analysis*, John Wiley & Sons, 1958.

[Sal85] J. Salz, "Digital transmission over cross-coupled linear channels", *AT & T Technical Journal*, vol. 64, no. 6, pp. 1147–1159, July-Aug. 1985.

[Sam01] 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.

[Sca99] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs", *IEEE Trans. Signal Processing*, vol. 47, no. 5, pp. 1988–2006, July 1999.

[Sca02a] A. Scaglione, "Statistical analysis of capacity of MIMO frequency selective Rayleigh fading channels with arbitrary number of inputs and outputs", *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pag. 278, July 2002.

[Sca02b] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders", *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[Shi00] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems", *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[Shi03] H. Shin, and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole", *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[Shi06] H. Shin, M.Z. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels", *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2264, Aug. 2006.

[Shi08] W.-Y. Shin, S.-Y. Chung, and Y. H. Lee, "Diversity-multiplexing tradeoff and outage performance for rician MIMO channels", *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1186–1196, Mar. 2008.

[Sim95] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communications Techniques: Signal Design and Detection*, Prentice Hall, 1995.

[Sim98] M. K. Simon, and M. S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels", *Proc. of the IEEE*, vol. 86, no. 9, pp. 1860–1877, Sep. 1998.

[Sim02a] M. K. Simon, "The Nuttall Q function- Its relation to the Marcum Q function and its application in digital communication performance evaluation", *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1712–1715, Nov. 2002.

[Sim02b] M. K. Simon, and M. S. Alouini, *Digital communication over fading channels*, Wiley series in telecommunications and signal processing, John Wiley & Sons, $2^{\underline{\text{nd}}}$ ed., 2002.

[Sim04] S. H. Simon, and A. L. Moustakas, "Eigenvalue density of correlated complex random Wishart matrices", *Physical Review E*, vol. 69, no. 6, pp. 1–4, June 2004.

[Sim06] 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.

[Smi03] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading", *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[Sri65] M. S. Srivastava, "On the complex Wishart distribution", *Ann. of Math. Statist.*, vol. 36, no. 1, pp. 313–315, Feb. 1965.

[Sri79] M. S. Srivastava, and C. G. Khatri, *An Introduction to Multivariate Statistics*, Elsevier North Holland, 1979.

[Sri03] M. S. Srivastava, "Singular Wishart and multivariate Beta distributions", *Ann. of Statist.*, vol. 31, no. 5, pp. 1537–1560, Oct. 2003.

[Ste72] H. S. Steyn, and J. J. J. Roux, "Approximations for the non-central Wishart distributions", *South African Statist. J.*, vol. 6, pp. 165–173, 1972.

[Sto02] P. Stoica, and G. Ganesan, "Maximum-SNR spatial-temporal formatting designs for MIMO channels", *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3036–3042, Dec. 2002.

[Stü96] G. L. Stüber, *Principles of Mobile Communications*, Kluwer Academic, 1996.

[Sve47] E. Sverdrup, "Derivation of the Wishart distribution of the second order sample moments by straight forward integration of a multiple integral", *Skandinavisk Aktuarietidskrift*, vol. 30, pp. 151–166, 1947.

[Tan68] W. Y. Tan, "Some distribution theory associated with complex Gaussian distribution", *Tamkang J.*, vol. 7, pp. 263–301, 1968.

[Tan79] W. Y. Tan, "On the approximation of noncentral Wishart distribution by Wishart distribution", *Metron*, vol. 37, no. 3, pp. 49–58, 1979.

[Tan82] W. Y. Tan, and R. P. Gupta, "On approximating the non-central Wishart distribution by central Wishart distribution a Monte Carlo studytion by central wishart distribution a monte carlo study", *Commun. Statist. Simul. Comp.*, vol. 11, no. 1, pp. 47–64, Jan. 1982.

[Tar98] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction", *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[Tar99] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[Tel99] I. E. Telatar, "Capacity of multi-antenna Gaussian channel", *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[Tse05] D. Tse, and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge Univ. Press, 2005.

[Tul04] A. M Tulino, and S. Verdú, *Random Matrix Theory and Wireless Communications*, Now Publishers, Hanover, MA, 2004.

[Uhl94] H. Uhlig, "On singular Wishart and singular multivariate beta distributions", *Ann. of Statist.*, vol. 22, no. 1, Mar. 1994.

[Vri08] B. Vrigneau, J. Letessier, P. Rostaing, L. Collin, and G. Burel, "Extension of the MIMO precoder based on the minimum Euclidean distance: a cross-form matrix", *IEEE J-STSP Special Issue: MIMO-Optimized Transmission Systems for Delivering Data and Rich Content*, vol. 2, no. 2, pp. 135–146, Apr. 2008.

[Wai72] V. B. Waikar, T. C. Chang, and P. R. Krishnaiah, "Exact distribution of few arbitrary roots of some complex random matrices", *Austral. J. Statist.*, vol. 14, no. 1, pp. 84–88, 1972.

[Wan03] Z. Wang, and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels", *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[Wig65] E. P. Wigner, "Distribution laws for the roots of a random Hermitian matrix", *in Statistical Theories of Spectra: Fluctuations*, pp. 446–461, 1965.

[Wig67] E. P. Wigner, "Random matrices in physics", *SIAM Review*, vol. 9, no. 1, pp. 1–23, Jan. 1967.

[Win87] J. H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment", *IEEE J. Select. Areas Commun.*, vol. 5, no. 5, pp. 871–878, Jun. 1987.

[Wis28] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population", *Biometrika*, vol. 20A, pp. 32–52, 1928.

[Wis48] J. Wishart, "Proofs of the distribution law of the second order moment statistics", *Biometrika*, vol. 35, pp. 55–57, 1948.

[Woo56] R. A. Wooding, "The multivariate distribution of complex normal variates", *Biometrika*, vol. 43, pp. 212–215, 1956.

[Yan94a] J. Yang, and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback", *IEEE Trans. Inform. Theory*, vol. 40, no. 5, pp. 1334–1347, Sep. 1994.

[Yan94b] J. Yang, and S. Roy, "On joint transmitter and receiver optimization for multiple-input multiple-output (MIMO) transmission systems", *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[Yu01] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Second order statistics of NLOS indoor MIMO channels based on 5.2 GHz measurements", *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 20, pp. 156–60, Nov. 2001.

[Yu02] K. Yu, and B. Ottersten, "Models for MIMO propagation channels: a review", *Wirel. Commun. Mob. Comput.*, vol. 2, pp. 653–666, Nov. 2002.

[Zan05] A. Zanella, M. Chiani, and M. Z. Win, "Performance analysis of MIMO MRC in correlated Rayleigh fading environments", *Proc. IEEE Veh. Tech. Conf. Spring (VTC-S)*, Jun. 2005.

[Zan08] A. Zanella, M. Chiani, and M. Z. Win, "A general framework for the distribution of the eigenvalues of wishart matrices", *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2008.

[Zha03] X. Zhang, and B. Ottersten, "Power allocation and bit loading for spatial multiplexing in MIMO systems", *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. V, pp. 53–56, Apr. 2003.

[Zha07] L. Zhao, W. Mo, Y. Ma, and Z. Wang, "Diversity and multiplexing tradeoff in general fading channels", *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1549–1557, Apr. 2007.

[Zhe03] 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.

[Zho05] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li, "MIMO systems with adaptive modulation", *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1828–1842, Sep. 2005.