



<b>Project Number:</b>	IST-026905
<b>Project Title:</b>	Multiple-Access Space-Time Coding Testbed
<b>Project Coordinator:</b>	C.F. Mecklenbräuker
<b>Deliverable Number:</b>	D3.1.3

<b>Title of Deliverable:</b>	Capacity regions of MU-MIMO systems
<b>Workpackage:</b>	WP-3
<b>Nature:</b>	R
<b>Dissemination level:</b>	PU
<b>Editors:</b>	Giorgio Taricco, Jialai Weng, Giulio Coluccia
<b>Authors:</b>	see list inside
<b>Contractual Date of Deliverable:</b>	Feb. 28, 2009
<b>Actual Date of Delivery:</b>	Feb. 18, 2009

**Abstract:**

The focus of this deliverable is on the study of capacity regions (i.e., the sets of jointly achievable rates) of multiuser MIMO systems by means of analytic and simulation tools. Chapter 1 introduces the main concepts about the analytic tools used. Chapters 3 to 4 address the case of separately-correlated Rician fading for narrowband MIMO channels by considering the following topics: *i*) transmitted signal covariance optimization in the presence of intended and interfering signals; *ii*) second-order statistics of the mutual information in the presence of interference; *iii*) ergodic capacity region for a multiuser channel. Chapter 5 introduces interference functions to analyse capacity regions (and also other performance regions) by expressing them as sub-level sets of a certain interference function. Chapters 6 and 7 address the case of wideband MIMO channels with separately-correlated Rician fading and considers the following topics: *i*) capacity and optimum signal covariance matrix for a point-to-point MIMO link; *ii*) ergodic capacity region for a multiuser channel.

# Contents

<b>1</b>	<b>Introduction (J. Weng)</b>	<b>9</b>
1.1	Multiuser MIMO fading channels . . . . .	9
1.2	Introduction to Asymptotic Analysis of MIMO Channel . . . . .	11
<b>2</b>	<b>Ergodic Capacity with Interference (G. Taricco and E. Riegler)</b>	<b>14</b>
2.1	Introduction . . . . .	14
2.2	System model and basic results . . . . .	15
2.2.1	Definition of $E_b/N_0$ and Shannon's limit . . . . .	18
2.3	Ergodic channel capacity . . . . .	18
2.3.1	Jensen approximations . . . . .	20
2.4	Numerical results . . . . .	20
2.4.1	Comparison with the results from [50] . . . . .	21
2.4.2	Impact of interference . . . . .	21
2.4.3	Consistency with Shannon's limit . . . . .	21
<b>3</b>	<b>Second-Order Statistics... (G. Taricco and E. Riegler)</b>	<b>28</b>
3.1	Introduction . . . . .	28
3.2	System Model . . . . .	29
3.3	The Cumulant Generating Function . . . . .	30
3.3.1	Rewriting determinants . . . . .	31
3.3.2	Calculating expectations . . . . .	31
3.3.3	Disentangling products and integration . . . . .	32
3.3.4	Saddle point approximation . . . . .	32
3.4	Asymptotic mean of the mutual information . . . . .	33
3.5	Asymptotic variance of mutual information . . . . .	33
3.6	Numerical results . . . . .	34
3.7	Conclusion . . . . .	35
3.8	Appendix . . . . .	35
3.8.1	Identities for complex valued matrices . . . . .	35
3.8.2	Identities for Grassmann valued matrices . . . . .	36
3.8.3	Supermatrices, superdeterminant and supertrace . . . . .	37

3.8.4	Definition of functions and matrices . . . . .	37
<b>4</b>	<b>Ergodic capacity region (G. Taricco and E. Riegler)</b>	<b>40</b>
4.1	Introduction . . . . .	40
4.2	System model . . . . .	41
4.3	Ergodic Capacity Region . . . . .	42
4.3.1	Jensen approximation . . . . .	45
4.4	Numerical results . . . . .	47
4.4.1	First scenario ( $K = 2$ ), symmetric . . . . .	47
4.4.2	Second scenario ( $K = 4$ , symmetric) . . . . .	49
4.4.3	Third scenario ( $K = 4$ , asymmetric) . . . . .	50
4.5	Conclusions . . . . .	50
<b>5</b>	<b>Interference Functions (H. Boche and M. Schubert)</b>	<b>54</b>
5.1	Introduction . . . . .	54
5.1.1	Comprehensive Performance Sets . . . . .	55
5.1.2	Level Sets and Interference Functions . . . . .	56
5.1.3	Discussion . . . . .	58
<b>6</b>	<b>Wideband MIMO Channels (G. Taricco and E. Riegler)</b>	<b>59</b>
6.1	Introduction . . . . .	59
6.2	System model . . . . .	60
6.2.1	OFDM signaling . . . . .	61
6.2.2	Frequency-domain equivalent channel . . . . .	62
6.3	Asymptotic Analysis . . . . .	63
6.3.1	Asymptotic CGF . . . . .	64
6.3.2	Ergodic mutual information . . . . .	65
6.3.3	Ergodic capacity . . . . .	66
6.4	Numerical results . . . . .	67
6.5	Conclusions . . . . .	68
<b>7</b>	<b>Capacity Region of OFDM MIMO Channels (G. Taricco, E. Riegler)</b>	<b>72</b>
7.1	Introduction . . . . .	72
7.2	System model . . . . .	73
7.3	Achievable rate region . . . . .	75
7.4	Optimization . . . . .	77
7.4.1	Simultaneous water-filling . . . . .	80
7.5	Numerical simulations . . . . .	81
7.6	Conclusion . . . . .	83

# Authors

*Holger Boche*

Fraunhofer Institute for Telecommunications HHI

E-mail: boche@hhi.fhg.de

*Erwin Riegler*

Forschungszentrum Telekommunikation Wien

E-mail: riegler@ftw.at

*Martin Schubert*

Fraunhofer Institute for Telecommunications HHI

E-mail: schubert@hhi.fhg.de

*Giorgio Taricco*

Politecnico di Torino

e-Mail: taricco@polito.it

*Jialai Weng*

Politecnico di Torino

e-Mail: jialai@tlc.polito.it

# Executive Summary

This deliverable summarizes the contributions in the MASCOT WP3 concerning the study of capacity regions for the multiuser MIMO channel. The main target of WP3 is to establish and investigate the fundamental performance limits of the multi-user (MU) MIMO systems.

We focus on MIMO channel capacity in the Shannon theoretic sense. According to Shannon's Theorem, the Shannon capacity of a single-user time-invariant channel is the maximum mutual information between the channel input and output. This maximum mutual information is shown to be the maximum *achievable* data rate, i.e., the maximum data rate that can be transmitted over the channel with arbitrarily small error probability. When the channel is time-varying, channel capacity has different definitions, depending on what is known about the channel state or its distribution at the transmitter and/or receiver and whether capacity is measured based on averaging the rate over all channel states/distributions or maintaining a constant fixed or minimum rate. In this case, the Shannon (ergodic) capacity is the maximum mutual information averaged over all channel states.

The research study described in this report has been influenced by the methods of statistical physics, developed in the last century to study the interactions of particles in gases, fluids, and solids. Statistical physics and multiuser communications show strong analogies from a conceptual point of view. As a result, we derive the asymptotic capacity of single-user and MU MIMO channel by using the *replica methods* and *superanalysis*, which origin from statistical physics.

We study the ergodic capacity of the asymptotic separately-correlated Rician fading MIMO channel with interference in in Chapter 2. In this chapter we consider the separately-correlated Rician fading MIMO channel with narrowband interference and calculate its channel capacity with the only limitation that the receive correlation matrix is common for both the intended user signal and interference. A simple method to derive the ergodic capacity and the corresponding capacity-achieving covariance matrix for a MIMO fading channel with multiuser interference is provided. The method applies when the fading distribution is based on the separately-correlated (Kronecker) Rician fading model (with common receive correlation), as the number of antennas grow asymptotically large. Numeri-

cal results are provided to assess the accuracy of the asymptotic analytic method. The results are compared to those obtained by more complex optimization algorithms proposed by Vu and Paulraj [50] in the interference-free case. Next, the analysis is extended to other MIMO channels affected by interference in order to assess the benefits of covariance optimization versus independent and uniform power allocation (corresponding to iid transmitted symbols).

The second-order statistics of the mutual information of the asymptotic separately-correlated Rician fading MIMO channel with interference are the main subject of Chapter 3. This chapter aims at finding an analytic expression for the *moment generating function* of the asymptotic mutual information. The results are based on the *replica method* and *superanalysis*, powerful tools developed in the context of theoretical physics. Our initial findings are based on the *replica method*, which turned out in the recent past to be a powerful tool to handle similar problems [58, 60, 61, 64]. Basically, we extend the approach used by Moustakas *et al.* in [3], though in a nontrivial manner, to the correlated Rician fading case and derive the mean and the variance of the mutual information. The mean and variance of the mutual information of a separately-correlated Rician fading MIMO channel are obtained in the presence of multi-access interference, when the number of transmit and receive antennas grows asymptotically large and perfect receive channel-state information is assumed. The presence of line-of-sight components induces additional coupling between the signal and interference parts (off-diagonal matrix blocks in equations (3.31)) in the saddle point approximation, which makes the calculations considerably complex. This problem is solved by applying the methods of *superanalysis* developed in the context of theoretical physics [52]. Analytic asymptotic results are compared by Monte-Carlo simulation in order to assess the accuracy of the method even when the number of antennas is small.

In Chapter 4 we study the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel. This chapter presents an asymptotic analytic method, still based on the replica method, to calculate the *ergodic capacity region* of a multiple-access MIMO channel with correlated Rician fading. The method applies when the number of antennas is very large but provides very accurate approximations even with a small number of antennas. We assume that full channel state information at the receiver (CSIR) is available and the transmitter knows the statistics of the channel, i.e., the channel distribution information at the transmitter (CDIT), but not the full channel state information. Based on these assumptions, we provide an algorithm to find the maximum ergodic sum-rate achieving covariance matrices of a multiple-access MIMO channel when the number of transmit and receive antennas grow asymptotically large with finite asymptotic ratios and the number of users and their SNR's are finite. In this context we assume that the multiple-access communication channel is affected by Rician

fading with separate spatial correlation (with a common receive part and different transmit parts). Our results rely on a previous work [25] where the ergodic capacity achieving covariance matrix was obtained for a separately-correlated Rician fading MIMO channel with multiple-access interference, which extended previous results due to Moustakas *et al.* [3] relevant to the case of Rayleigh fading. It is shown by numerical results that this asymptotic approach is very accurate even when the number of antennas is as low as a few units. The ergodic capacity achieving covariance matrices for all users are derived according to the algorithm provided and the corresponding capacity is compared with the mutual information achieved by iid power allocation. Monte-Carlo simulations are also reported in order to verify the accuracy of the asymptotic results. Similar results for the separately-correlated Rician fading MIMO channel (without multiple-access interference) were obtained independently by Dumont *et al.* [31], using an asymptotic method based on Stieltjes transforms, and by Vu and Paulraj [50], using an interior point with barrier optimization method [14].

The analysis of capacity regions is complicated by interference between the communication links(users). The achievable capacity of one link can depend on the transmission strategies of other links. This typically results in a coupled system with many degrees of freedom. Thus, well-established communication strategies for point-to-point links are not always applicable to multiuser systems. Characterization of wireless capacity regions can be quite awkward, especially when additional system constraints are considered. This motivates an abstract approach based on *interference functions*, which focuses on some core properties.

Chapter 5 investigates the system performance limits based on interference functions. One main contribution of this chapter is to show that interference functions can be used to analyse capacity regions (and also other performance regions). In particular, every comprehensive capacity region can be expressed as a sub-level set of an interference function. This facilitates a general framework for analyzing performance trade-offs in multiuser networks.

The results of this work, published in [75, 77, 80], show that there is a direct correspondence between comprehensive performance regions and interference functions fulfilling the core properties of the interference functions. Further properties, like convexity or log-convexity, can be added. This theoretical framework facilitates a general and unifying approach for the analysis of different kinds of capacity regions. By focusing on core properties, we are able to develop a rigorous framework which allows for an analytical treatment. The results provide intuition and a roadmap for the development of algorithms in WP1. An application example is the iterative algorithm for max-min balancing published in [74]. Other resource allocation strategies are currently being investigated in WP1. For example, interference functions were successfully applied to the analysis of resource allocation strategies based on cooperative game theory in [76, 79].

Chapter 6 extends the analysis carried out for the narrowband MIMO channel to the wideband MIMO channel based on OFDM signalling to characterize the ergodic channel capacity. Although this is again an asymptotic analysis, numerical results show that there is very good agreement with the asymptotic analytic results. The advantage of having analytic results consists mainly in the fact that it allows to optimize the input signal covariance matrix in order to determine the ergodic capacity. The chapter focuses on a separately correlated Rician fading MIMO channel under some mild technical assumptions, the key one consisting in the fact that only the shortest delay communication path has a line-of-sight components while the remaining ones have only scattered components. This chapter provides the basic tools on which the next Chapter 7 builds its developments aimed to obtaining the full ergodic capacity region of a multiuser multiple-access wideband MIMO-OFDM separately correlated Rician fading channel. It is worth noticing that the true ergodic capacity region is obtained in this case and not only the rate region optimizing the ergodic sum rate. In order to achieve this result, somewhat sophisticated convex optimization techniques have been applied (interior point with barrier optimization method from [14]) and the resulting region contains all possible rate combinations that are achievable under the given power constraint. Again, these results have been obtained by using the asymptotic approach, and numerical (Monte-Carlo) results have been also reported in order to establish the accuracy of the analytic approach. It turned out that there is very good agreement between the asymptotic and the numerical results with a very small number of antennas.



# Chapter 1

## Introduction

*Jialai Weng*

### 1.1 Multiuser MIMO fading channels

Wireless systems continue to strive for ever higher data rates. This goal is particularly challenging for systems that are power, bandwidth, and complexity limited. However, another domain can be exploited to significantly increase channel capacity: the spatial domain based on the use of multiple transmit and receive antennas. Pioneering work by Foschini, and Telatar ignited much interest in this area by predicting remarkable spectral efficiencies for wireless systems with multiple antennas when the channel exhibits rich scattering and its variations can be accurately tracked. This initial promise of exceptional spectral efficiency resulted in an explosion of research activity to characterize the theoretical and practical issues associated with multiple-input multiple-output (MIMO) wireless channels and to extend these concepts to multiuser systems. We introduce some recent work focused on the capacity of MIMO systems for both single-users and multiple users under different assumptions about spatial correlation and channel information available at the transmitter and receiver.

The large spectral efficiencies associated with MIMO channels are based on the premise that a rich scattering environment provides independent transmission paths from each transmit antenna to each receive antenna. Therefore, for single-user systems, a transmission and reception strategy that exploits this structure achieves capacity on approximately separate channels, where  $N_t$  is the number of transmit antennas and  $N_r$  is the number of receive antennas. Thus, capacity scales linearly with relative to a system with just one transmit and one receive antenna. This capacity increase requires a scattering environment such that the matrix of channel gains between transmit and receive antenna pairs has full rank and independent en-

tries and that perfect estimates of these gains are available at the receiver. Perfect estimates of these gains at both the transmitter and receiver provides an increase in the constant multiplier associated with the linear scaling. Much subsequent work has been aimed at characterizing MIMO channel capacity under more realistic assumptions about the underlying channel model and the channel estimates available at the transmitter and receiver. The main question from both a theoretical and practical standpoint is whether the enormous capacity gains initially predicted by Winters, Foschini, and Telatar can be obtained in more realistic operating scenarios and what specific gains result from adding more antennas and/or a feedback link to feed receiver channel information back to the transmitter.

MIMO channel capacity depends heavily on the statistical properties and antenna element correlations of the channel. Recent work has developed both analytical and measurement-based MIMO channel models along with the corresponding capacity calculations for typical indoor and outdoor environments. Antenna correlation varies drastically as a function of the scattering environment, the distance between transmitter and receiver, the antenna configurations, and the Doppler spread. As we shall see, the effect of channel correlation on capacity depends on what is known about the channel at the transmitter and receiver: correlation sometimes increases capacity and sometimes reduces it. Moreover, channels with very low correlation between antennas can still exhibit a keyhole effect where the rank of the channel gain matrix is very small, leading to limited capacity gains. Fortunately, this effect is not prevalent in most environments. The impact of channel statistics in the low-power(wide-band) regime has interesting properties as well.

We focus on MIMO channel capacity in the Shannon theoretic sense. The Shannon capacity of a single-user time-invariant channel is defined as the maximum mutual information between the channel input and output. This maximum mutual information is shown by Shannon capacity theorem to be the maximum data rate that can be transmitted over the channel with arbitrarily small error probability. When the channel is time-varying channel capacity has multiple definitions, depending on what is known about the channel state or its distribution at the transmitter and/or receiver and whether capacity is measured based on averaging the rate over all channel states/distributions or maintaining a constant fixed or minimum rate. In this case, the Shannon (ergodic) capacity is the maximum mutual information averaged over all channel states. This ergodic capacity is typically achieved using an adaptive transmission policy where the power and data rate vary relative to the channel state variations. Other capacity definitions for time-varying channels with perfect transmitter and receiver CSI include outage capacity and minimum-rate capacity. These capacities require a fixed or minimum data rate in all non-outage channel states, which is needed for applications with delay-constrained data where the data rate cannot depend on channel variations (except in outage states, where no data is transmitted). The average rate

associated with outage or minimum rate capacity is typically smaller than ergodic capacity due to the additional constraints associated with these definitions. We will focus on ergodic capacity.

When only the channel distribution is known at the transmitter (receiver) the transmission (reception) strategy is based on the channel distribution instead of the instantaneous channel state. The channel coefficients are typically assumed to be jointly Gaussian, so the channel distribution is specified by the channel mean and covariance matrices. We will refer to knowledge of the channel distribution as channel distribution information (CDI). We assume throughout the work that CDI is always perfect, so there is no mismatch between the CDI at the transmitter or receiver and the true channel distribution. When only the receiver has perfect CSI the transmitter must maintain a fixed-rate transmission strategy optimized with respect to its CDI. In this case, ergodic capacity defines the rate that can be achieved based on averaging over all channel states. Alternatively, the transmitter can send at a rate that cannot be supported by all channel states: in these poor channel states the receiver declares an outage and the transmitted data is lost. In this scenario, each transmission rate has an outage probability associated with it and capacity is measured relative to outage probability (capacity CDF). For single-user MIMO channels with perfect transmitter and receiver CSI the ergodic and outage capacity calculations are straightforward since the capacity is known for every channel state.

In multiuser channels, capacity becomes a  $K$  dimensional region defining the set of all rate vectors simultaneously achievable by all users. The multiple capacity definitions for time-varying channels under different transmitter and receiver CSI and CDI assumptions extend to the capacity region of the multiple-access channel (MAC) in the obvious way. However, these MIMO multiuser capacity regions, even for time-invariant channels, are difficult to find. Few capacity results exist for time varying multiuser MIMO channels, especially under the realistic assumption that the transmitter(s) and/or receiver(s) have CDI only.

In this work we studied the ergodic capacity of single user and multiuser in a correlated Rician fading channel.

## 1.2 Introduction to Asymptotic Analysis of MIMO Channel

Wireless communication systems are designed to work in environments with the minimum possible amount of infrastructure. Their goal is to provide the users with the freedom to communicate with whomever they want regardless wherever they are. Since electromagnetic waves, the most popular carriers of digital com-

munications, propagate to almost any place, each user, though communicating with only a single other user, interacts, in principle, with all other users in the network. Such a setting is hard to press in formulas, in particular if the environment is arbitrary. The failure of the so-called 3-rd generation of wireless technology is, to a large extent, due to a lack of understanding of the fundamental principle governing wireless communications in the presence of many users operating simultaneously. This knowledge gap has become a severe obstacle to the further penetration of wireless communication devices into modern society and lifestyle and, therefore, must be overcome.

Research on the behavior of systems where many bodies mutually interact with many others has been driven forward by physicists for more than a century studying the interactions of particles in gases, fluids, and solids. Statistical physics and multiuser communications show strong analogies from a conceptual point of view. In both cases many objects interact with each other through variables that are constrained in a certain way. These inter-disciplinary analogies can be exploited to advance the understanding and design of future wireless communication systems. Though the analogies between the two fields do not extend too far and, in real world communication systems, statistical physics results cannot be applied directly, the engineering community can strongly benefit from the analytical tool-boxes developed by physicists. So far random matrix theory, originally studied to describe spacings of nuclear energy levels, has received the most attention in wireless system analysis and design. In addition, the replica method developed in statistical physics has entered wireless communication to cope with the often binary nature of wireless communication signals. In wireless communications, random matrix theory and statistical mechanics tools have overwhelmingly been used for performance analysis. Many works have used these large system tools for actual design of communication systems.

Communication via MIMO channel allows a significant increase in spectral efficiency the information rate per communication link. While many recent research works aim to utilize this advantage, it is still not sufficiently understood how the physical properties of these channels translate into achievable signal-to-interference-and-noise ratios (SINRs) and therefore the supported information rates. On the physical side, channel models are based on propagation measurements. They provide statistics of the propagation between a pair of transmitter arrays and receiver arrays in terms of delays, received powers, and directions of arrival and departure. Statements about the information rates capable in the channel, however, are given in terms of the eigenvalues related to the matrix algebraic description of the communication link. Many works aim to build a bridge between propagation scenarios and the eigenvalues of the covariance matrices of the channel in order to allow for predictions of channel capacity based on the morphology of the physical medium. It is natural to describe a linear time invariant MIMO

system by its matrix valued impulse response. The matrix valued taps of the impulse response of the antenna array channel depend on various parameters such as the exact locations of all antenna elements and all scattering objects, which are usually modeled as random variables in mobile communications. The quality of the communication link, however, is mainly determined by the singular values of these matrix taps. It is well known that the singular values of a large class of random matrix ensembles show fewer random fluctuations the larger the matrices are, and become deterministic in the limit of infinite matrix size. In the large matrix limit, the influences of many properties of the matrix entries are lost, such as the shapes of their distributions, and in some cases even the statistical dependencies among them. Though the asymptotic distribution of singular values is only an approximation to the distribution in the case of finite-dimensional matrices, it offers two important advantages.

1) In contrast with finite-dimensional matrices, the singular value distribution of asymptotically large random matrices can be calculated analytically in many cases. 2) In the asymptotic limit, only those physical parameters survive that show significant influence on the singular value distribution.

With these two properties, the limiting singular value distribution can help to analytically extract which physical parameters of the radio propagation channel largely determine the quality of a MIMO communication link.

## Chapter 2

# Ergodic Capacity of the Asymptotic Separately-Correlated Rician Fading MIMO Channel with Interference

*Giorgio Taricco, Erwin Riegler*

### 2.1 Introduction

Multiple-input-multiple-output (MIMO) channels have attracted considerable attention during the last decade because of the promise of very high information rates at an affordable cost. Seminal works by Winters [51], Telatar [46], Foschini and Gans [32, 33] illustrated the principles of MIMO communications and how to derive the channel capacity under the *rich scattering* assumption, corresponding to spatially-uncorrelated Rayleigh fading. More recently, experimental and theoretical works showed that the rich scattering assumption is often inadequate to encompass all the channel characteristic and more sophisticated correlated Rician fading models have been proposed to describe more realistic MIMO channels [34, 35, 44]. Additionally, multiuser interference is known to have a considerable impact on the achievable information rate and its *Gaussian approximation* is known to produce unduly pessimistic results [27, 28].

In the current literature, separately-correlated Rayleigh fading with interference has been considered by Moustakas *et al.* [3] as far as concerns the computation of the mean and variance of the mutual information. These results have been extended in [8, 9] to the separately-correlated Rician fading case without interference and, more recently, with interference [43]. All these results address the

evaluation of the mutual information when the number of antennas grow asymptotically large. Many works address the derivation of capacity for specific MIMO channels such as MISO with separately-correlated Rayleigh and uncorrelated Rician fading [49], subsequently extended to the MIMO case by [37, 48]. These works derive the eigenvectors of the ergodic capacity achieving covariance matrix while the eigenvalue derivation is based on other numerical algorithms [39]. Other works in the area of MIMO capacity derivation are [2, 38, 47]. In this context, but using an asymptotic approach, Dumont *et al.* provide an algorithm for the evaluation of the *ergodic capacity* for the separately-correlated Rician fading MIMO channel without interference [31]. The authors compare their results to those obtained by Vu and Paulraj [50].

In this work we consider the separately-correlated Rician fading MIMO channel with narrowband interference and calculate its channel capacity with the only limitation that the receive correlation matrix is common for both the intended user signal and interference. We compare our results with those obtained numerically by Vu and Paulraj [50] in the interference-free case. Then, we extend our analysis to other MIMO channels affected by interference and assess the effect of covariance optimization against iid power allocation (iid transmitted symbols).

## 2.2 System model and basic results

We consider a narrowband block fading channel with  $r$  receive antennas,  $t$  transmit antennas from an intended user, and  $\vartheta$  transmit antennas from an interfering source. The channel is specified by the following equation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{H}_I\mathbf{x}_I + \mathbf{z}. \quad (2.1)$$

Here,  $\mathbf{x} \in \mathbb{C}^{t \times 1}$  is the transmitted signal vector,  $\mathbf{x}_I \in \mathbb{C}^{\vartheta \times 1}$  is the interfering signal vector,  $\mathbf{H} \in \mathbb{C}^{r \times t}$  is the signal channel matrix,  $\mathbf{H}_I \in \mathbb{C}^{r \times \vartheta}$  is the interference channel matrix,  $\mathbf{z} \in \mathbb{C}^{r \times 1}$  is the additive noise vector, and  $\mathbf{y} \in \mathbb{C}^{r \times 1}$  is the received signal vector. Both  $\mathbf{x}$  and  $\mathbf{x}_I$  are assumed to have zero mean.

We assume that the additive noise vector contains has zero mean and covariance matrix  $\mathbf{Q}_Z = \mathbb{E}[\mathbf{z}\mathbf{z}^H]$ .

The channel matrices  $\mathbf{H}$  and  $\mathbf{H}_I$  are assumed to be of separately (or Kronecker) correlated Rician fading type *with common receive correlation*. Thus, they can be written as

$$\mathbf{H} = \bar{\mathbf{H}} + \mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}$$

and

$$\mathbf{H}_I = \bar{\mathbf{H}}_I + \mathbf{R}^{1/2}\mathbf{W}_I\mathbf{T}_I^{1/2},$$



where  $\bar{\mathbf{H}}$  and  $\bar{\mathbf{H}}_I$  represent the average channel matrices related to the presence of a line-of-sight signal component in the multipath fading channel, the Hermitian positive definite matrices  $\mathbf{R}$ ,  $\mathbf{T}$ ,  $\mathbf{T}_I$  are the receive and transmit (signal and interference) correlation matrices, and  $\mathbf{W}$  and  $\mathbf{W}_I$  have iid  $\mathcal{N}_c(0, 1)$  entries.

We define the signal and interference covariance matrices as  $\mathbf{Q}$  and  $\mathbf{Q}_I$ , respectively. In the following we assume that the interference covariance matrix is kept fixed and optimization is carried out on the signal covariance matrix  $\mathbf{Q}$  under a power constraint. Following standard conventions [1, 8], we define the Rician factors as

$$K \triangleq \frac{\|\bar{\mathbf{H}}\|^2}{\text{Tr}(\mathbf{T}) \text{Tr}(\mathbf{R})} \quad \text{and} \quad K_I \triangleq \frac{\|\bar{\mathbf{H}}_I\|^2}{\text{Tr}(\mathbf{T}_I) \text{Tr}(\mathbf{R})}. \quad (2.2)$$

We also define the signal-to-noise power ratio (SNR), the interference-to-noise power ratio (INR), and the signal-to-interference power ratio (SIR) at the receiver as

$$\begin{aligned} \text{SNR} &\triangleq \frac{(K+1) \text{Tr}(\mathbf{T}) \text{Tr}(\mathbf{R}) \text{Tr}(\mathbf{Q}/t)}{\text{Tr}(\mathbf{Q}_Z)}; \\ \text{INR} &\triangleq \frac{(K_I+1) \text{Tr}(\mathbf{T}_I) \text{Tr}(\mathbf{R}) \text{Tr}(\mathbf{Q}_I/\vartheta)}{\text{Tr}(\mathbf{Q}_Z)}; \\ \text{SIR} &\triangleq \frac{(K+1) \text{Tr}(\mathbf{T}) \text{Tr}(\mathbf{Q}/t)}{(K_I+1) \text{Tr}(\mathbf{T}_I) \text{Tr}(\mathbf{Q}_I/\vartheta)}. \end{aligned} \quad (2.3)$$

These power ratios coincide with the corresponding *received* power ratios when  $\mathbf{Q} = (P/t)\mathbf{I}_t$  (iid transmitted symbols).

We know [29] that the random mutual information for a given channel realization is given by

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}) &= \ln \det(\mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H + \mathbf{Q}_Z) \\ &\quad - \ln \det(\mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H + \mathbf{Q}_Z) \text{ nat/s/Hz}. \end{aligned} \quad (2.4)$$

Then, the ergodic capacity under a power constraint  $P$  is obtained as

$$C = \max_{\text{Tr}(\mathbf{Q}) \leq P} \mathbb{E}[I(\mathbf{x}; \mathbf{y})]. \quad (2.5)$$

In order to calculate the capacity (2.5) we resort to a recent result allowing to calculate the average mutual information of a separately-correlated Rician MIMO channel when the number of transmit/receive antennas grows asymptotically large [8, 9]. Summarizing, the average capacity when the channel matrix is  $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}$ , and the noise and signal covariance matrices are  $\mathbf{I}_r$  and  $\mathbf{Q}$ , respectively, is given by

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] \sim \mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}) \quad (2.6)$$



nat/complex dimension, where we defined

$$\mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}) \triangleq \ln \det \begin{pmatrix} \mathbf{I}_r + w\mathbf{R} & \tilde{\mathbf{H}} \\ -\tilde{\mathbf{H}}^H & \mathbf{I}_t + z\tilde{\mathbf{T}} \end{pmatrix} - wz \quad (2.7)$$

where  $\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{1/2}\mathbf{T}\mathbf{Q}^{1/2}$ ,  $\tilde{\mathbf{T}}_I \triangleq \mathbf{Q}_I^{1/2}\mathbf{T}_I\mathbf{Q}_I^{1/2}$ , and  $w, z$  can be obtained by solving the equations

$$\begin{cases} w = \text{Tr} \left\{ [z\mathbf{I}_t + \tilde{\mathbf{T}}^{-1} + \tilde{\mathbf{T}}^{-1}\tilde{\mathbf{H}}^H(\mathbf{I}_r + w\mathbf{R})^{-1}\tilde{\mathbf{H}}]^{-1} \right\} \\ z = \text{Tr} \left\{ [w\mathbf{I}_r + \mathbf{R}^{-1} + \mathbf{R}^{-1}\tilde{\mathbf{H}}(\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}^H]^{-1} \right\} \end{cases} \quad (2.8)$$

with  $\tilde{\mathbf{H}} \triangleq \bar{\mathbf{H}}\mathbf{Q}^{1/2}$ ,  $\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{1/2}\mathbf{T}\mathbf{Q}^{1/2}$ .

This result can be applied to the calculation of the terms in (2.4). In fact, to calculate

$$\mathbf{I}_1 \triangleq \mathbb{E}[\ln \det(\mathbf{I}_r + \mathbf{Q}_Z^{-1}(\mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H))], \quad (2.9)$$

we can consider the channel defined by

$$\begin{aligned} \mathbf{Q}_Z^{-1/2}\mathbf{y} &= \left\{ \mathbf{Q}_Z^{-1/2}(\bar{\mathbf{H}}, \bar{\mathbf{H}}_I) \right. \\ &\quad \left. + \mathbf{Q}_Z^{-1/2}\mathbf{R}^{1/2}(\mathbf{W}, \mathbf{W}_I) \begin{pmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_I \end{pmatrix}^{1/2} \right\} \begin{pmatrix} \mathbf{x} \\ \mathbf{x}_I \end{pmatrix} + \mathbf{Q}_Z^{-1/2}\mathbf{z}. \end{aligned} \quad (2.10)$$

Then, we get

$$\mathbf{I}_1 = \mu_I(\bar{\mathbf{H}}_1, \mathbf{R}_1, \mathbf{T}_1, \mathbf{Q}_1)$$

where  $\bar{\mathbf{H}}_1 \triangleq \mathbf{Q}_Z^{-1/2}(\bar{\mathbf{H}}, \bar{\mathbf{H}}_I)$ ,  $\mathbf{R}_1 \triangleq \mathbf{Q}_Z^{-1/2}\mathbf{R}\mathbf{Q}_Z^{-1/2}$ ,  $\mathbf{T}_1 \triangleq \text{diag}(\mathbf{T}, \mathbf{T}_I)$ , and  $\mathbf{Q}_1 \triangleq \text{diag}(\mathbf{Q}, \mathbf{Q}_I)$ . Similarly, defining

$$\mathbf{I}_2 \triangleq \mathbb{E}[\ln \det(\mathbf{I}_r + \mathbf{Q}_Z^{-1}\mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H)], \quad (2.11)$$

we get

$$\mathbf{I}_2 = \mu_I(\bar{\mathbf{H}}_2, \mathbf{R}_1, \mathbf{T}_I, \mathbf{Q}_I)$$

where  $\bar{\mathbf{H}}_2 \triangleq \mathbf{Q}_Z^{-1/2}\bar{\mathbf{H}}_I$ . Thus, we have

$$\begin{aligned} C &= \max_{\text{Tr}(\mathbf{Q}) \leq P} (\mathbf{I}_1 - \mathbf{I}_2) \\ &= \max_{\text{Tr}(\mathbf{Q}) \leq P} \{ \mu_I(\bar{\mathbf{H}}_1, \mathbf{R}_1, \mathbf{T}_1, \mathbf{Q}_1) - \mu_I(\bar{\mathbf{H}}_2, \mathbf{R}_1, \mathbf{T}_I, \mathbf{Q}_I) \}. \end{aligned}$$

### 2.2.1 Definition of $E_b/N_0$ and Shannon's limit

Here we assume that the received noise vector is uncorrelated (spatially white) and hence  $\mathbf{Q}_Z = N_0 \mathbf{I}_r$ .

We define  $E_b$  as the average received signal energy per bit obtained by dividing the total average received signal energy by the average achievable bit rate. Thus, we have

$$\frac{E_b}{N_0} \triangleq \frac{\mathbb{E}[\text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)]}{N_0 \mathbb{E}[I(\mathbf{x}; \mathbf{y})]} \ln 2. \quad (2.12)$$

When the average signal and interference powers approach zero with constant ratio, we can use the approximation  $\ln \det(\mathbf{I} + \mathbf{X}) \approx \text{Tr}(\mathbf{X})$  (holding for any nonnegative matrix  $\mathbf{X}$  when  $\text{Tr}(\mathbf{X}) \rightarrow 0$ ) to approximate the average achievable bit rate (2.4) as follows:

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] \approx N_0^{-1} \mathbb{E}[\text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^H)].$$

Then, inserting the approximation above into (2.12), we obtain

$$\frac{E_b}{N_0} \rightarrow \ln 2 \quad \text{as} \quad \text{Tr}(\mathbf{Q}) \rightarrow 0,$$

which is the well known Shannon limit.

**Remark 2.2.1** *In most research works in MIMO communications, the SNR is preferred to the  $E_b/N_0$  ratio as a system cost indicator as it effectively describes the channel reliability. However, only the  $E_b/N_0$  ratio at the receiver provides a precise description of the channel behavior in its asymptotically low power regime.*

## 2.3 Ergodic channel capacity

The ergodic capacity and the corresponding optimum covariance matrix are determined by maximizing the mutual information under the power constraint considered. Following the approach in the previous section, we can obtain the mutual information  $\mathbf{I}_1$  for any given transmit covariance matrix  $\mathbf{Q}$  in the asymptotic antenna setting. Since  $\mathbf{I}_2$  is independent of  $\mathbf{Q}$ , the channel capacity can be obtained by maximizing  $\mathbf{I}_1$  only, i.e.,

$$C = \left\{ \max_{\text{Tr}(\mathbf{Q}) \leq P} \mathbf{I}_1 \right\} - \mathbf{I}_2, \quad (2.13)$$

where we assume that the interference covariance matrix  $\mathbf{Q}_I$  is fixed.

In order to solve the maximization problem we need to maximize  $\mu_I(\bar{\mathbf{H}}_1, \mathbf{R}_1, \mathbf{T}_1, \mathbf{Q}_1)$ , where  $\bar{\mathbf{H}}_1 \triangleq \mathbf{Q}_Z^{-1/2}(\bar{\mathbf{H}}, \bar{\mathbf{H}}_I)$ ,  $\mathbf{R}_1 \triangleq \mathbf{Q}_Z^{-1/2}\mathbf{R}\mathbf{Q}_Z^{-1/2}$ ,  $\mathbf{T}_1 \triangleq \text{diag}(\mathbf{T}, \mathbf{T}_I)$ , and  $\mathbf{Q}_1 \triangleq \text{diag}(\mathbf{Q}, \mathbf{Q}_I)$ , with respect to  $\mathbf{Q}$  and under the constraint  $\text{Tr}(\mathbf{Q}) \leq P$ . In order to proceed, we let  $\mathbf{S} = \mathbf{Q}^{1/2}$ ,  $\mathbf{S}_I = \mathbf{Q}_I^{1/2}$ , and  $\mathbf{S}_1 = \mathbf{Q}_1^{1/2} = \text{diag}(\mathbf{S}, \mathbf{S}_I)$  and we write the Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{S}) = & \ln \det \begin{pmatrix} \mathbf{I}_r + w\mathbf{R}_1 & \bar{\mathbf{H}}_1\mathbf{S}_1 \\ -\mathbf{S}_1\bar{\mathbf{H}}^H & \mathbf{I}_{t+\varnothing} + z\mathbf{S}_1\mathbf{T}_1\mathbf{S}_1 \end{pmatrix} \\ & - wz - \lambda[\text{Tr}(\mathbf{S}^2) - \tilde{P}] \end{aligned} \quad (2.14)$$

where  $0 \leq \tilde{P} \leq P$ . Next, we calculate the first-order total variation of (2.14):

$$\begin{aligned} \delta\mathcal{L} = & \text{Tr} \left[ \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ \mathbf{C}_1 & \mathbf{D}_1 \end{pmatrix} \right. \\ & \cdot \left( \begin{array}{cc} \mathbf{R}_1\delta w & \bar{\mathbf{H}}_1\delta\mathbf{S}_1 \\ -\delta\mathbf{S}_1\bar{\mathbf{H}}^H & \mathbf{S}_1\mathbf{T}_1\mathbf{S}_1\delta z + z(\delta\mathbf{S}_1\mathbf{T}_1\mathbf{S}_1 + \mathbf{S}_1\mathbf{T}_1\delta\mathbf{S}_1) \end{array} \right) \\ & \left. - w\delta z - z\delta w - 2\lambda \text{Tr}(\mathbf{S}\delta\mathbf{S}) \right] \end{aligned} \quad (2.15)$$

where  $\delta\mathbf{S}_1 = \text{diag}(\delta\mathbf{S}, \mathbf{0}_{\varnothing \times \varnothing})$  and [8]:

$$\begin{cases} \mathbf{A}_1 = [\mathbf{I}_r + w\mathbf{R}_1 + \tilde{\mathbf{H}}_1(\mathbf{I}_{t+\varnothing} + z\tilde{\mathbf{T}}_1)^{-1}\tilde{\mathbf{H}}_1^H]^{-1} \\ \mathbf{B}_1 = -(\mathbf{I}_r + w\mathbf{R}_1)^{-1}\tilde{\mathbf{H}}_1\mathbf{D}_1 \\ \mathbf{C}_1 = (\mathbf{I}_{t+\varnothing} + z\tilde{\mathbf{T}}_1)^{-1}\tilde{\mathbf{H}}_1^H\mathbf{A}_1 \\ \mathbf{D}_1 = [\mathbf{I}_{t+\varnothing} + z\tilde{\mathbf{T}}_1 + \tilde{\mathbf{H}}_1^H(\mathbf{I}_r + w\mathbf{R}_1)^{-1}\tilde{\mathbf{H}}_1]^{-1} \end{cases} \quad (2.16)$$

where  $\tilde{\mathbf{H}}_1 \triangleq \bar{\mathbf{H}}_1\mathbf{Q}_1^{1/2}$  and  $\tilde{\mathbf{T}}_1 \triangleq \mathbf{Q}_1^{1/2}\mathbf{T}_1\mathbf{Q}_1^{1/2}$ . Expanding (2.15) we get:

$$\begin{aligned} \delta\mathcal{L} = & [\text{Tr}(\mathbf{A}_1\mathbf{R}_1) - z]\delta w + [\text{Tr}(\mathbf{D}_1\tilde{\mathbf{T}}_1) - w]\delta z \\ & + \text{Tr}[(-\tilde{\mathbf{H}}_1^H\mathbf{B}_1 + \mathbf{C}_1\bar{\mathbf{H}}_1 + z(\mathbf{T}_1\mathbf{S}_1\mathbf{D}_1 + \mathbf{D}_1\mathbf{S}_1\mathbf{T}_1))\delta\mathbf{S}_1] \\ & - 2\lambda \text{Tr}(\mathbf{S}_1\delta\mathbf{S}_1) \end{aligned} \quad (2.17)$$

Since the first two terms in (2.17) are zero when  $w, z$  satisfy (2.8), we can see that the total variation  $\delta\mathcal{L}$  is null, provided that the following equation holds:

$$\begin{aligned} & \{\bar{\mathbf{H}}_1^H(\mathbf{I}_r + w\mathbf{R}_1)^{-1}\tilde{\mathbf{H}}_1\mathbf{D}_1 + (\mathbf{I}_{t+\varnothing} + z\tilde{\mathbf{T}}_1)^{-1}\tilde{\mathbf{H}}_1^H\mathbf{A}_1\bar{\mathbf{H}}_1 \\ & + z(\mathbf{T}_1\mathbf{S}_1\mathbf{D}_1 + \mathbf{D}_1\mathbf{S}_1\mathbf{T}_1) - 2\lambda\mathbf{S}_1\}_{1:t,1:t} = 0 \end{aligned} \quad (2.18)$$

where  $(\mathbf{A})_{a:b,c:d}$  is the submatrix of  $\mathbf{A}$  obtained by extracting the elements of rows  $a$  to  $b$  and columns  $c$  to  $d$ . Then, we can show that:

$$(\mathbf{I}_r + w\mathbf{R}_1)^{-1}\tilde{\mathbf{H}}_1\mathbf{D}_1 = [(\mathbf{I}_{t+\varnothing} + z\tilde{\mathbf{T}}_1)^{-1}\tilde{\mathbf{H}}_1^H\mathbf{A}_1]^H$$

Therefore, we can simplify (2.18) and obtain the following fixed-point equation:

$$\mathbf{S} = \lambda^{-1}\mathcal{H}[(\bar{\mathbf{H}}_1^H(\mathbf{I}_r + w\mathbf{R}_1)^{-1}\bar{\mathbf{H}}_1 + z\mathbf{T}_1)\mathbf{S}_1\mathbf{D}_1]_{1:t,1:t}, \quad (2.19)$$

where  $\mathcal{H}(\mathbf{A}) \triangleq (\mathbf{A} + \mathbf{A}^H)/2$ .

### 2.3.1 Jensen approximations

In this section we extend the Jensen bound approach proposed in [50] relevant to the interference-free case. It is known from [29, 30, 40] that the following inequalities hold for positive definite matrices  $\mathbf{A}$ ,  $\mathbf{X}$  ( $\mathbf{X}$  random):

$$\mathbb{E}[\ln \det(\mathbf{X})] \leq \ln \det(\mathbb{E}[\mathbf{X}]), \quad (2.20)$$

$$\mathbb{E}[\ln \det(\mathbf{I} + \mathbf{X}^{-1}\mathbf{A})] \geq \ln \det(\mathbf{I} + \mathbb{E}[\mathbf{X}]^{-1}\mathbf{A}) \quad (2.21)$$

Then, applying (2.21), we have

$$\begin{aligned} \mathbb{E}[I(\mathbf{x}; \mathbf{y}) \mid \mathbf{H}] &= \mathbb{E}[\ln \det(\mathbf{I}_r + (\mathbf{H}_I \mathbf{Q}_I \mathbf{H}_I^H + \mathbf{Q}_Z)^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^H) \mid \mathbf{H}] \\ &\geq \ln \det(\mathbf{I}_r + \mathbf{Q}_{IZ}^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^H) \end{aligned}$$

where we defined  $\mathbf{Q}_{IZ} \triangleq \mathbb{E}[\mathbf{H}_I \mathbf{Q}_I \mathbf{H}_I^H + \mathbf{Q}_Z]$ . Thus,

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] \geq \mathbb{E}[\ln \det(\mathbf{I}_r + \mathbf{Q}_{IZ}^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^H)], \quad (2.22)$$

i.e., the average mutual information of a MIMO channel with interference is lower bounded by the average mutual information of the same channel where interference is replaced by Gaussian noise with the same covariance matrix (see also [27]). Applying sequentially inequalities (2.21) and (2.20) (even though they have opposite directions), we get the following *Jensen approximation*:

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] \approx \ln \det(\mathbf{I}_t + \mathbb{E}[\mathbf{H}^H \mathbf{Q}_{IZ}^{-1} \mathbf{H}] \mathbf{Q}). \quad (2.23)$$

Hence, we can maximize this approximate mutual information by applying the standard water-filling approach [29]. If

$$\mathbb{E}[\mathbf{H}^H \mathbf{Q}_{IZ}^{-1} \mathbf{H}] = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H,$$

the optimum covariance matrix is given by

$$\mathbf{Q}_J = \mathbf{U}(\mu \mathbf{I}_t - \mathbf{\Lambda}^{-1})_+ \mathbf{U}^H \quad (2.24)$$

where  $\mu$  is obtained by solving  $\text{Tr}[(\mu \mathbf{I}_t - \mathbf{\Lambda}^{-1})_+] = P$  and  $(\mathbf{A})_+$  is the matrix with entries  $\max\{0, (\mathbf{A})_{ij}\}$ .

## 2.4 Numerical results

In this section we present numerical examples based on the method developed before.

### 2.4.1 Comparison with the results from [50]

In order to validate the method proposed in this paper, we compare the results obtained versus those presented in [50]. Our results are presented in Fig. 2.1 (channel capacity) and 2.2 (normalized covariance matrix eigenvalues). Specifically, Fig. 2.1 reports the capacity obtained numerically in [50] (dotted line with circles), the capacity obtained by our asymptotic approach (solid line), and the capacity obtained by Monte-Carlo simulation and using the transmit covariance matrix obtained by our asymptotic approach (diamonds). Fig. 2.2 reports the normalized covariance matrix eigenvalues obtained numerically in [50] (dotted lines with circles) and those obtained by our asymptotic approach (solid lines with diamonds). In both cases, our results are in close agreement with [50].

### 2.4.2 Impact of interference

We consider a MIMO system with the following parameters:  $t = r = \infty = 4$ ,  $K = K_I = 10$  dB,  $\mathbf{R}, \mathbf{T}, \mathbf{T}_I$  exponential matrices ( $\alpha^{|i-j|}$  with base  $\alpha = 0.7$  in all cases),  $\bar{\mathbf{H}}, \bar{\mathbf{H}}_I$  with rank one (all-equal entries). Figs. 2.3 to 2.5 plot the average mutual information for  $\mathbf{SIR} = -10, 0, 10$  dB, respectively versus the SNR and the normalized optimum  $\mathbf{Q}$  eigenvalues. The three mutual information curves correspond to different  $\mathbf{Q}$ : 1) optimum based on the proposed algorithm; 2) iid symbols ( $\mathbf{Q} = q\mathbf{I}_t$ ); 3) Jensen approximation; and 4) optimum with Gaussian interference.

It can be noticed that the gain of the optimum versus iid covariance matrix may exceed 10 dB in some cases. Moreover, the Jensen approximation provides extremely accurate results up to some SNR threshold, above which it starts to degrade (as can be noticed from Fig. 2.3). The goodness of the Jensen approximation decreases (slightly) as the Rician factor  $K$  gets lower. The channel capacity corresponding to the Gaussian approximation of interference is much lower than the true channel capacity at  $\mathbf{SIR} = -10, 0$  dB, as also evidenced in [28] for a Rayleigh fading MIMO channel. The difference is more limited in the  $\mathbf{SIR} = 10$  dB case. In all cases, we have an SNR threshold, depending on the  $\mathbf{SIR}$ , above which the two capacity curves diverge.

### 2.4.3 Consistency with Shannon's limit

Previous results reported the mutual information versus the SNR defined in (2.3). Fig. 2.6 reports instead the mutual information versus the received  $E_b/N_0$  ratio defined in (2.12). The diagrams show that the results obtained are consistent with the Shannon's limit, as discussed in Section 2.2.1.

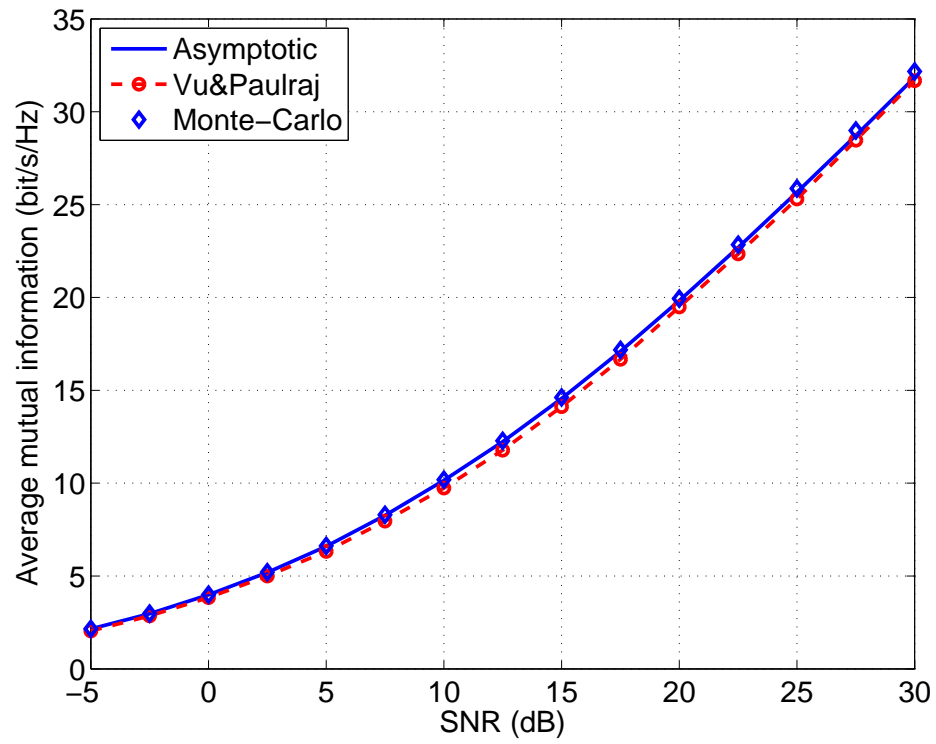


Figure 2.1: Mutual information of the  $4 \times 4$  MIMO channel proposed in [50].

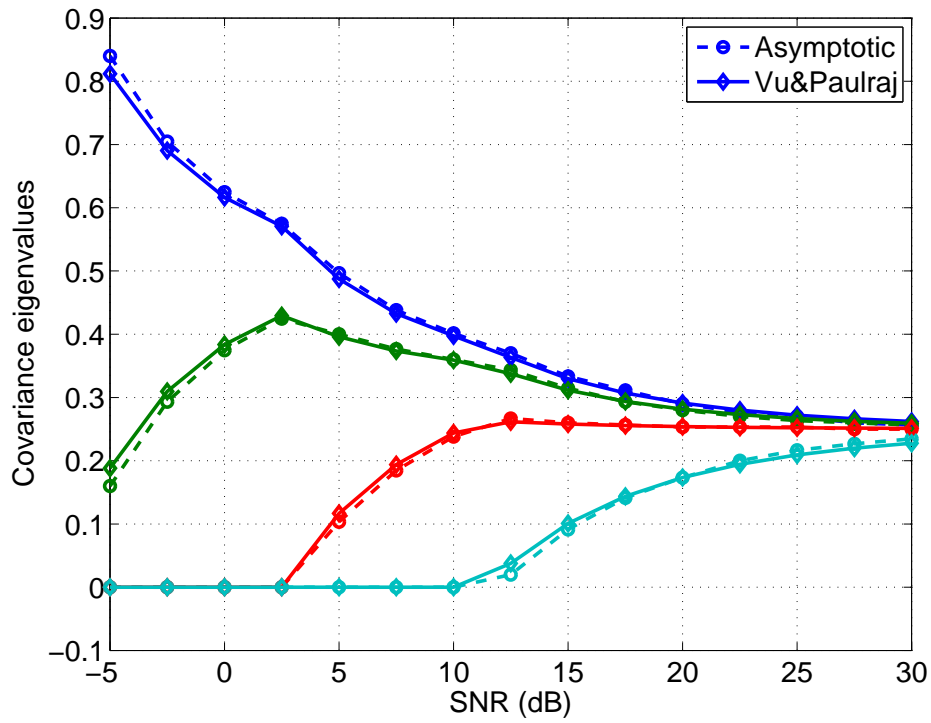


Figure 2.2: Normalized eigenvalues of the optimum transmit covariance matrix for the  $4 \times 4$  MIMO channel proposed in [50].

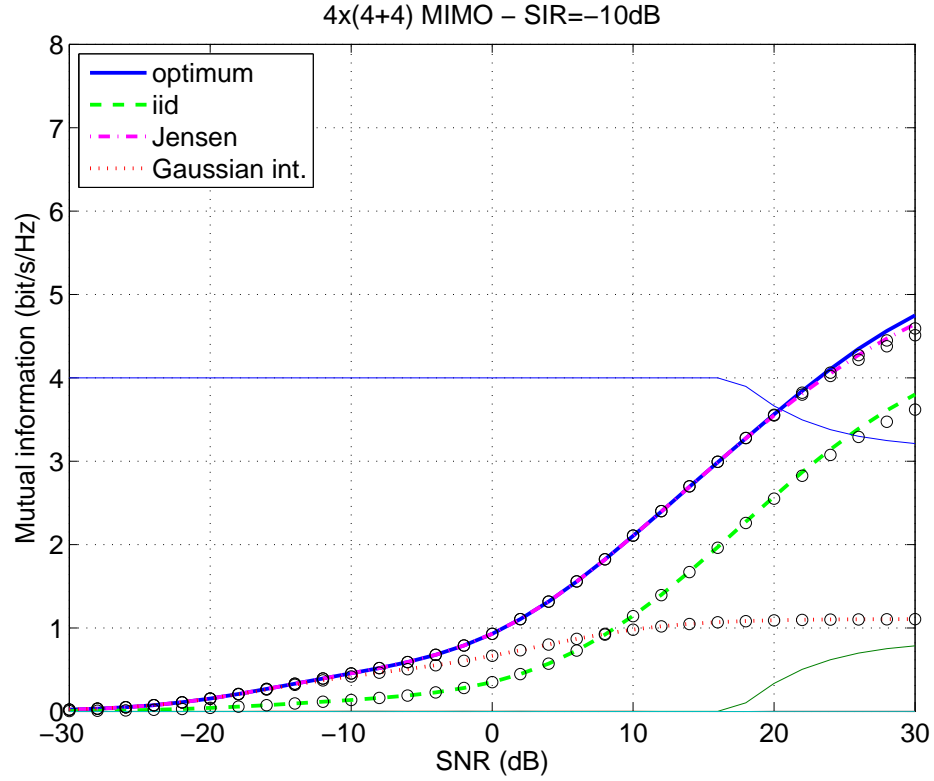


Figure 2.3: Average mutual information of a MIMO channel with interference and parameters  $t = r = \vartheta = 4$ ,  $\text{SIR} = -10$  dB,  $K = K_I = 10$  dB, exponential  $\mathbf{R}, \mathbf{T}, \mathbf{T}_I$  with base 0.7,  $\bar{\mathbf{H}}, \bar{\mathbf{H}}_I$  with all-equal entries. Optimum, Jensen, and iid covariance curves (solid lines: asymptotic, circles: Monte-Carlo sim.). Optimum normalized eigenvalues are also plotted.



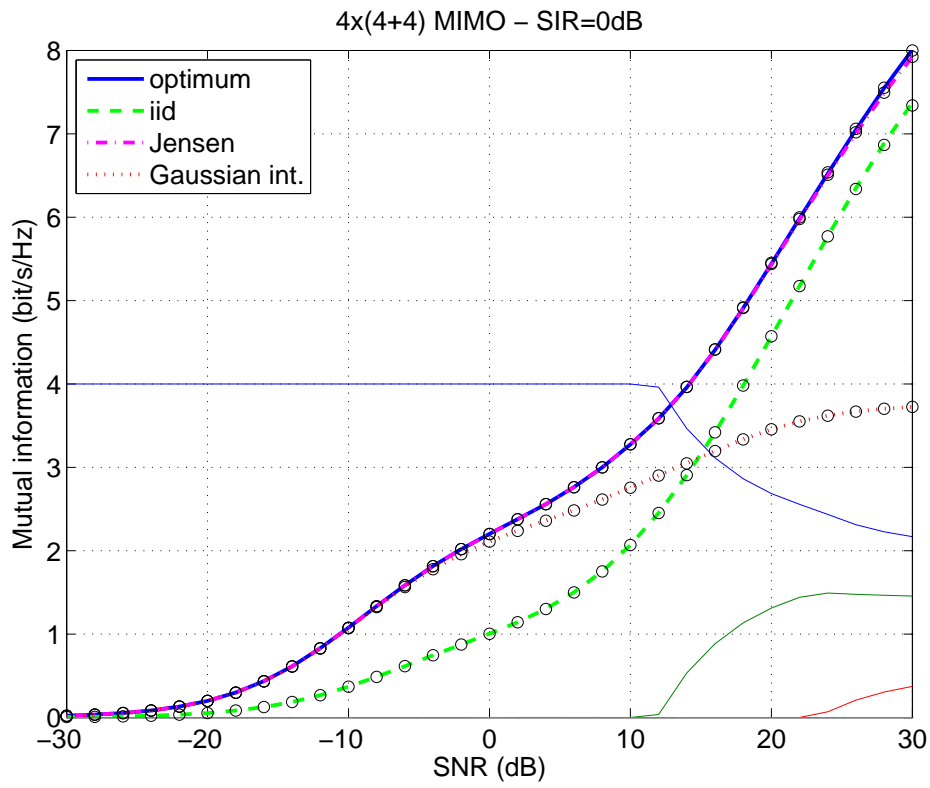


Figure 2.4: Same as Fig. 2.3 but  $\text{SIR} = 0$  dB.

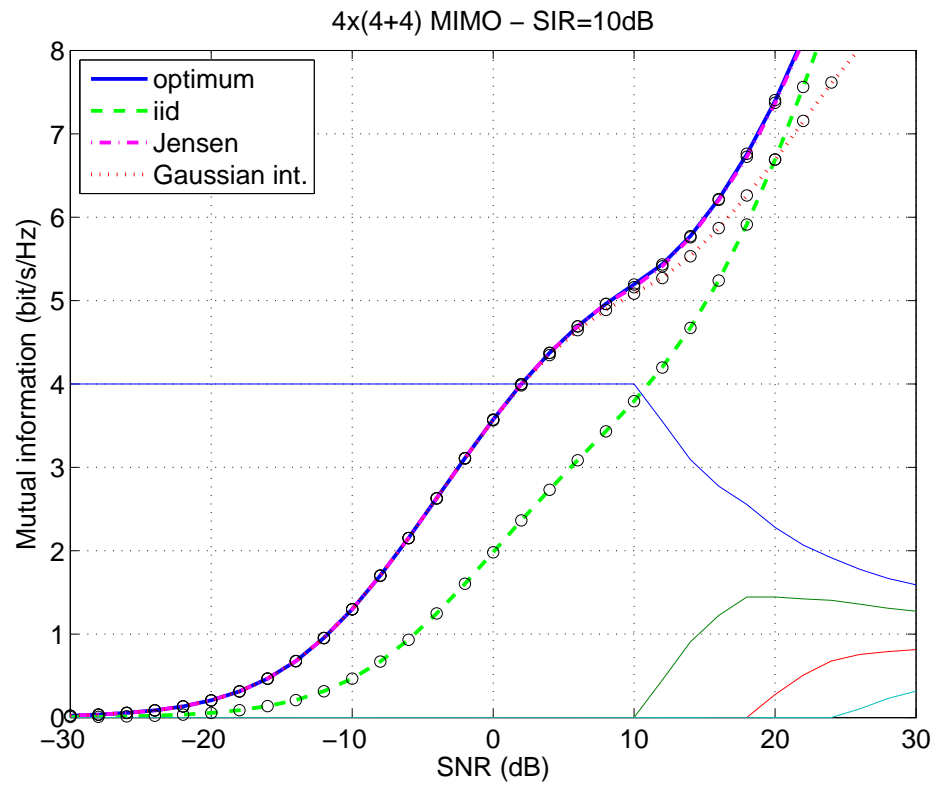


Figure 2.5: Same as Fig. 2.3 but  $\text{SIR} = 10$  dB.

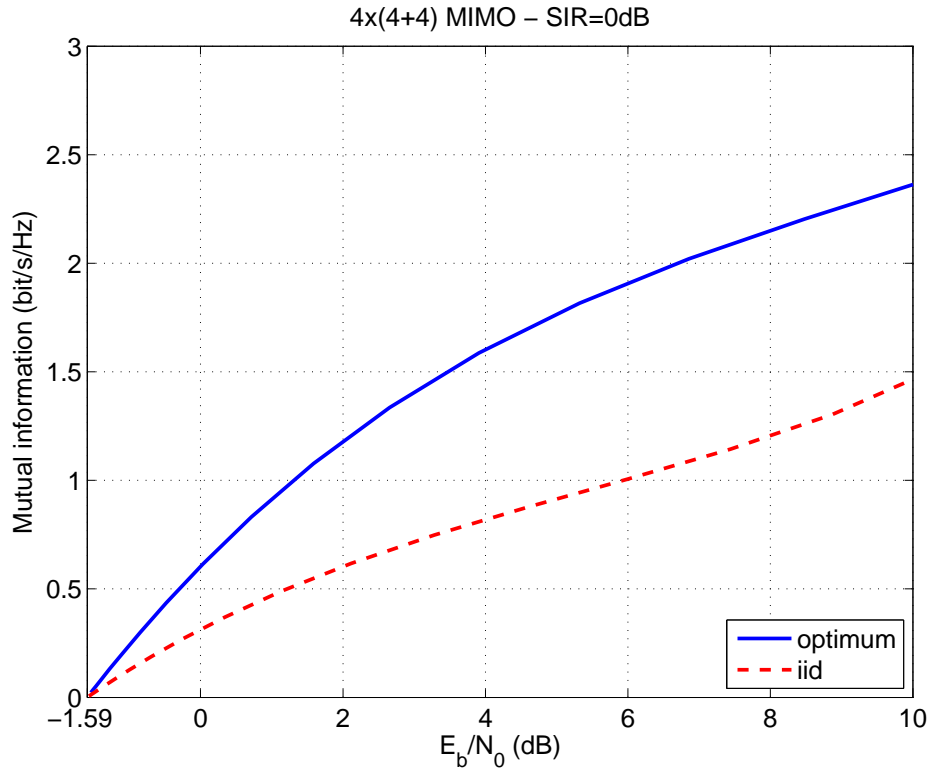


Figure 2.6: Same as Fig. 2.3 but  $\text{SIR} = 0$  dB. Plot of mutual information (optimum and iid covariance) versus received  $E_b/N_0$  ratio defined in (2.12).

## Chapter 3

# Second-Order Statistics of the Mutual Information of the Asymptotic Separately-Correlated Rician Fading MIMO Channel with Interference

*Giorgio Taricco, Erwin Riegler*

### 3.1 Introduction

During the last decade, much attention has been devoted to the analysis of the capacity of multiple-input multiple-output (MIMO) channels [57, 62, 63] but only a few papers considered the presence of multi-access interference, corresponding to a more realistic multiuser MIMO scenario [3, 53–55]. Recently, Chiani *et al.* found a closed-form expression of the exact mean capacity for an uncorrelated Rayleigh fading MIMO channel with interference [28]. Notice that assimilating interference into additive noise *leads to incorrect capacity results*, as properly evidenced in [28].

Though many studies focused on the case of Rayleigh fading (with different correlation structures), several experimental and theoretical works showed that Rician fading needs to be considered to describe more realistic MIMO channels [34, 35, 44]. The interference-free separately-correlated Rician fading MIMO channel has been recently studied in [9]<sup>1</sup>. The results therein turned out to be a

---

<sup>1</sup> Similar results were obtained independently by Dumont *et al.* [31, 56] by using an asymptotic

powerful tool in order to calculate the capacity achieving covariance matrices in a multiple access scenario [25, 43].

In this work we are interested in finding an analytic expression of the *moment generating function* of the asymptotic mutual information. Our findings are based on the *replica method*, which turned out in the recent past to be a powerful tool to handle similar problems [58, 60, 61, 64]. Basically, we extend the approach used by Moustakas *et al.* in [3] to the correlated Rician fading case and derive the mean and the variance of the mutual information. However, the presence of line-of-sight components induces an additional coupling between the signal and interference parts (off-diagonal matrix blocks in equations (3.31)) in the saddle point approximation, which makes the calculations highly nontrivial. This problem is solved by applying the methods of *superanalysis* developed in the context of theoretical physics [52].

In this paper we use Latin letters for complex variables  $x \in \mathbb{C}$  and Greek letters for Grassmann variables  $\psi \in \mathbb{G}$  [52]. Vectors and matrices are written with lowercase and uppercase boldface characters, respectively. For the (exponential of the) trace and the determinant of a matrix  $\mathbf{X}$  we use the symbol  $(\text{etr}(\mathbf{X}))$   $\text{tr}(\mathbf{X})$  and  $\det(\mathbf{X})$ , respectively. definite variance

## 3.2 System Model

We consider a narrowband block fading channel with  $t$  transmit,  $r$  receive and  $i$  interfering antennas characterized by the equation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{H}_I\mathbf{x}_I + \mathbf{z}. \quad (3.1)$$

Here,  $\mathbf{y} \in \mathbb{C}^r$  is the received signal vector,  $\mathbf{x} \in \mathbb{C}^t$  is the transmitted signal vector with zero mean and covariance  $\mathbf{Q} = \mathbb{E}[\mathbf{x}\mathbf{x}^H]$ ,  $\mathbf{x}_I \in \mathbb{C}^i$  is the narrowband interference signal vector with zero mean and covariance  $\mathbf{Q}_I = \mathbb{E}[\mathbf{x}_I\mathbf{x}_I^H]$ , and  $\mathbf{z} \in \mathbb{C}^r$  is the additive zero-mean noise vector with iid entries  $z_a \sim \mathcal{N}_c(0, 1)$ . The channel matrices  $\mathbf{H} \in \mathbb{C}^{r \times t}$  and  $\mathbf{H}_I \in \mathbb{C}^{r \times i}$  model separately-correlated Rician fading *with common receive correlation*. Thus, they can be written as

$$\begin{aligned} \mathbf{H} &= \bar{\mathbf{H}} + \mathbf{R}^{1/2} \mathbf{H}_w \mathbf{T}^{1/2}, \\ \mathbf{H}_I &= \bar{\mathbf{H}}_I + \mathbf{R}^{1/2} \mathbf{H}_{w,I} \mathbf{T}_I^{1/2}, \end{aligned} \quad (3.2)$$

where  $\bar{\mathbf{H}}$  and  $\bar{\mathbf{H}}_I$  represent the mean values and are related to the presence of line-of-sight components,  $\mathbf{R}$ ,  $\mathbf{T}$ , and  $\mathbf{T}_I$  are receive and transmit (signal and interference) correlation matrices, and  $\mathbf{H}_w$  and  $\mathbf{H}_{w,I}$  have iid  $\mathcal{N}_c(0, 1)$  entries. To

---

method based on Stieltjes transforms.

simplify notations, we define

$$\tilde{\mathbf{H}} \triangleq \bar{\mathbf{H}}\mathbf{Q}^{1/2}, \quad \tilde{\mathbf{H}}_I \triangleq \bar{\mathbf{H}}_I\mathbf{Q}_I^{1/2}, \quad (3.3)$$

$$\tilde{\mathbf{T}} \triangleq \mathbf{T}^{1/2}\mathbf{Q}\mathbf{T}^{1/2}, \quad \tilde{\mathbf{T}}_I \triangleq \mathbf{T}_I^{1/2}\mathbf{Q}_I\mathbf{T}_I^{1/2}. \quad (3.4)$$

Splitting the total received power components into direct and diffuse parts we obtain the Rician factors

$$K = \frac{\|\tilde{\mathbf{H}}\|^2}{\text{tr}(\mathbf{R}) \text{tr}(\tilde{\mathbf{T}})} \quad \text{and} \quad K_I = \frac{\|\tilde{\mathbf{H}}_I\|^2}{\text{tr}(\mathbf{R}) \text{tr}(\tilde{\mathbf{T}}_I)}. \quad (3.5)$$

The signal-to-noise (SNR) and interference-to noise (INR) ratio reads as

$$\begin{aligned} \text{SNR} &= \frac{(K+1) \text{tr}(\mathbf{R}) \text{tr}(\tilde{\mathbf{T}})}{r}, \\ \text{INR} &= \frac{(K_I+1) \text{tr}(\mathbf{R}) \text{tr}(\tilde{\mathbf{T}}_I)}{r} \end{aligned} \quad (3.6)$$

with signal-to-interference ratio  $\text{SIR} = \text{SNR}/\text{INR}$ . We assume that the receiver knows *perfectly* the channel matrices  $\mathbf{H}$  and  $\mathbf{H}_I$ . Thus, the mutual information *conditioned* to  $\mathbf{H}$  and  $\mathbf{H}_I$  is given by [29], [28]:

$$\begin{aligned} \mathcal{I} &= I(\mathbf{y}; \mathbf{x} \mid \mathbf{H}, \mathbf{H}_I) \\ &= \mathcal{H}(\mathbf{y} \mid \mathbf{H}, \mathbf{H}_I) - \mathcal{H}(\mathbf{y} \mid \mathbf{x}, \mathbf{H}, \mathbf{H}_I) \\ &= \ln \frac{\det(\mathbf{I}_r + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H + \mathbf{H}\mathbf{Q}\mathbf{H}^H)}{\det(\mathbf{I}_r + \mathbf{H}_I\mathbf{Q}_I\mathbf{H}_I^H)}, \end{aligned} \quad (3.7)$$

where  $\mathcal{H}$  denotes entropy. Clearly, when considering the *statistical variations* of the channel, the matrices  $\mathbf{H}$  and  $\mathbf{H}_I$  are random variables defined in equation (3.2). Therefore, the statistical behavior of  $\mathcal{I}$  is given by a random variable of the form (3.7) with first two cumulant moments, i.e. mean and variance

$$\mu_{\mathcal{I}} \triangleq \mathbb{E}_{\mathbf{H}_w, \mathbf{H}_{w,I}}[\mathcal{I}], \quad \sigma_{\mathcal{I}}^2 \triangleq \mathbb{E}_{\mathbf{H}_w, \mathbf{H}_{w,I}}[\mathcal{I}^2] - \mu_{\mathcal{I}}^2. \quad (3.8)$$

Note that  $\mu_{\mathcal{I}}$  is linear in  $\mathcal{I}$ , which implies that the influence of interference can be reduced to the case without interference [25, 28]. However for higher moments, and in particular for the variance  $\sigma_{\mathcal{I}}^2$ , this is *no longer the case*.

### 3.3 The Cumulant Generating Function

In order to calculate the cumulant moments of  $\mathcal{I}$  we introduce the generating function

$$G(a) \triangleq \mathbb{E}[\exp(-a\mathcal{I})] \quad \forall a \in \mathbb{C}_+ \quad (3.9)$$

and the cumulant generating function  $g(a) \triangleq \ln(G(a))$ .

Assuming that  $G(a)$  is analytic in a real right neighborhood of  $0^+$  allows to derive the mean and the variance of  $\mathcal{I}$  [3]:

$$\mu_{\mathcal{I}} = -g'(0^+), \quad \sigma_{\mathcal{I}}^2 = g''(0^+). \quad (3.10)$$

### 3.3.1 Rewriting determinants

In order to calculate (3.9) we rewrite the determinants in eq. (3.7). Repeated use of identity (3.23) yields:

$$\begin{aligned} & \det(\mathbf{I}_r + \mathbf{H}_I \mathbf{Q}_I \mathbf{H}_I^H + \mathbf{H} \mathbf{Q} \mathbf{H}^H)^{-a} \\ &= \int_{\mathbb{C}^{r \times a}} D_c \mathbf{U} \int_{\mathbb{C}^{t \times a}} D_c \mathbf{V} \int_{\mathbb{C}^{n_{oni} \times a}} D_c \mathbf{W} \\ & \quad \times \text{etr}(-\pi(\mathbf{U}^H \mathbf{U} + \mathbf{V}^H \mathbf{V} + \mathbf{W}^H \mathbf{W})) \\ & \quad \times \tilde{f}(\mathbf{U}, \mathbf{V}) f_w(\mathbf{U}, \mathbf{V}) \tilde{f}_I(\mathbf{U}, \mathbf{W}) f_{w,I}(\mathbf{U}, \mathbf{W}), \end{aligned} \quad (3.11)$$

with  $\tilde{f}$ ,  $f_w$ ,  $\tilde{f}_I$ , and  $f_{w,I}$  defined by (3.27). Similarly, repeated use of identity (3.25) yields:

$$\begin{aligned} & \det(\mathbf{I}_r + \mathbf{H}_I \mathbf{Q}_I \mathbf{H}_I^H)^a \\ &= \int_{\mathbb{C}^{r \times a}, \mathbb{G}^{a \times r}} D_g(\Psi, \bar{\Psi}) \int_{\mathbb{G}^{i \times a}, \mathbb{G}^{a \times i}} D_g(\Omega, \bar{\Omega}) \\ & \quad \text{etr}(\bar{\Psi} \Psi + \bar{\Omega} \Omega) \tilde{h}_I(\Psi, \bar{\Psi}, \Omega, \bar{\Omega}) h_{w,I}(\Psi, \bar{\Psi}, \Omega, \bar{\Omega}), \end{aligned} \quad (3.12)$$

with  $\tilde{h}_I$  and  $h_{w,I}$  defined in equations (3.28). The appearance of the functions  $\tilde{f}$ ,  $\tilde{f}_I$  and  $\tilde{h}_I$  is due to the line-of-sight components and responsible for the coupling of the determinants in the saddle point approximation 3.3.4.

### 3.3.2 Calculating expectations

Using identity (3.23), we obtain:

$$\begin{aligned} & \mathbb{E}_{\mathbf{H}_w, \mathbf{H}_{w,I}} [f_w(\mathbf{U}, \mathbf{V}) f_{w,I}(\mathbf{U}, \mathbf{W}) h_{w,I}(\Psi, \bar{\Psi}, \Omega, \bar{\Omega})] \\ &= \mathbb{E}_{\mathbf{H}_w} [f_w(\mathbf{U}, \mathbf{V})] \mathbb{E}_{\mathbf{H}_{w,I}} [f_{w,I}(\mathbf{U}, \mathbf{W}) h_{w,I}(\Psi, \bar{\Psi}, \Omega, \bar{\Omega})] \\ &= h_I(\mathbf{U}, \mathbf{W}; \Psi, \bar{\Psi}, \Omega, \bar{\Omega}) f(\mathbf{U}, \mathbf{V}) f_I(\mathbf{U}, \mathbf{W}), \end{aligned} \quad (3.13)$$

with  $f$ ,  $f_I$ , and  $h_I$  defined in equations (3.29) and (3.30).

### 3.3.3 Disentangling products and integration

In order to integrate out the Grassmann valued matrices  $\Psi, \bar{\Psi}, \Omega, \bar{\Omega}$  and the complex valued matrices  $U, V, W$  we have to disentangle the products of matrices in the exponents of  $f, f_I$ , and  $h_I$ , which can be done by using identity (3.25) and (3.24). After some algebra (omitted for space limitations) we get the following integral:

$$G(a) = \int d\mu(T_1, R_1) \int d\mu(T_2, R_2) \int d\mu(T_3, R_3) \int_{\mathbb{G}^{a \times a}, \mathbb{G}^{a \times a}} D_g(\Theta_1, \bar{\Theta}_1) \int_{\mathbb{G}^{a \times a}, \mathbb{G}^{a \times a}} D_g(\Theta_2, \bar{\Theta}_2) \exp[-F(T_1, R_1, T_2, R_2, T_3, R_3, \Theta_1, \bar{\Theta}_1, \Theta_2, \bar{\Theta}_2)] \quad (3.14)$$

where  $K, M, \Gamma$ , and  $\bar{\Gamma}$  are defined in (3.31) and

$$\begin{aligned} F &\triangleq -\text{tr}(R_1 T_1 + R_2 T_2 + R_3 T_3) + \text{tr}(\bar{\Theta}_1 \Theta_1 + \bar{\Theta}_2 \Theta_2) \\ &\quad - \ln \text{sdet}(\mathfrak{X}) \\ &= -\text{tr}(R_1 T_1 + R_2 T_2 + R_3 T_3) + \text{tr}(\bar{\Theta}_1 \Theta_1 + \bar{\Theta}_2 \Theta_2) \\ &\quad - \ln \det(M) + \ln \det(K + \bar{\Gamma} M^{-1} \Gamma) \end{aligned} \quad (3.15)$$

since  $\mathfrak{X}$  is given by

$$\mathfrak{X} = \begin{pmatrix} M & -\Gamma \\ \bar{\Gamma} & K \end{pmatrix}.$$

For this derivation we resorted to the concepts of supermatrix and superdeterminant developed in [52], and to the superdeterminant rule of eq. (3.26) in Appendix 3.8.3.

### 3.3.4 Saddle point approximation

Contrary to the case  $K = K_I = 0$  developed in [3], matrices  $K$  and  $M$  are not block-diagonal so that the determinants in (3.15) do not factor. This makes the task of evaluating the saddle point approximation much more complex.

Now, we assume that (3.15) has a replica-symmetric and real saddle point, hereafter denoted by  $\mathcal{S}$ , corresponding to

$$\begin{aligned} T_1 &= t_1 I_a & T_2 &= t_2 I_a & T_3 &= t_3 I_a \\ R_1 &= r_1 I_a & R_2 &= r_2 I_a & R_3 &= -r_3 I_a \\ \Theta_1 &= 0 & \Theta_2 &= 0 & \bar{\Theta}_1 &= 0 & \bar{\Theta}_2 &= 0 \end{aligned} \quad (3.16)$$



Using the fact that  $\ln \text{sdet}(\mathfrak{X}) = \text{str}(\ln \mathfrak{X})$  [52, pp. 112] and the multiplicative property of superdeterminants [52, p. 101], we have the following total variation expansion:

$$\delta \ln \text{sdet}(\mathfrak{X}) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \text{str}((\mathfrak{X}^{-1} \delta \mathfrak{X})^k), \quad (3.17)$$

where

$$\mathfrak{X}^{-1} |_{\mathcal{S}} \delta \mathfrak{X} = \begin{pmatrix} M^{-1} |_{\mathcal{S}} \delta M & -M^{-1} |_{\mathcal{S}} \delta \Gamma \\ K^{-1} |_{\mathcal{S}} \delta \bar{\Gamma} & K^{-1} |_{\mathcal{S}} \delta K \end{pmatrix}. \quad (3.18)$$

The explicit form of matrices  $M^{-1} |_{\mathcal{S}}$  and  $K^{-1} |_{\mathcal{S}}$  can be found in eqs. (3.32) and (3.34), respectively.

Nulling the first-order terms in the expansion (3.17) of (3.15) yields the following saddle point equations:

$$\begin{aligned} t_1 &= \text{tr}(\mathbf{A}_K \mathbf{R}) & t_2 &= \text{tr}(\mathbf{A}_K \mathbf{R}) & t_3 &= \text{tr}(\mathbf{A}_M \mathbf{R}) \\ r_1 &= \text{tr}(\mathbf{E}_K \tilde{\mathbf{T}}) & r_2 &= \text{tr}(\mathbf{I}_K \tilde{\mathbf{T}}_I) & r_3 &= \text{tr}(\mathbf{D}_M \tilde{\mathbf{T}}_I), \end{aligned} \quad (3.19)$$

with  $\mathbf{A}_M$  and  $\mathbf{D}_M$  from equations (3.33) and  $\mathbf{A}_K$ ,  $\mathbf{E}_K$ , and  $\mathbf{I}_K$  from equations (3.35).

### 3.4 Asymptotic mean of the mutual information

The leading term of the expansion of  $F$  at the saddle point  $\mathcal{S}$  is given by

$$\begin{aligned} F_0 &\triangleq -\ln \text{sdet}(\mathfrak{X}) |_{\mathcal{S}} - a(t_1 r_1 + t_2 r_2 - t_3 r_3) \\ &= \ln \det \mathbf{K} |_{\mathcal{S}} - \ln \det \mathbf{M} |_{\mathcal{S}} - a(t_1 r_1 + t_2 r_2 - t_3 r_3), \end{aligned}$$

after defining the matrices  $M |_{\mathcal{S}} = \mathbf{M}_0 \otimes \mathbf{I}_a$  and  $K |_{\mathcal{S}} = \mathbf{K}_0 \otimes \mathbf{I}_a$  with

$$\mathbf{M}_0 \triangleq \begin{pmatrix} \mathbf{I}_r - r_3 \mathbf{R} & -\tilde{\mathbf{H}}_I \\ \tilde{\mathbf{H}}_I^H & \mathbf{I}_i + t_3 \tilde{\mathbf{T}}_I \end{pmatrix}$$

and

$$\mathbf{K}_0 \triangleq \begin{pmatrix} \hat{\mathbf{R}} & \tilde{\mathbf{H}} & \tilde{\mathbf{H}}_I \\ -\tilde{\mathbf{H}}^H & \hat{\mathbf{T}} & \mathbf{0} \\ -\tilde{\mathbf{H}}_I & \mathbf{0} & \hat{\mathbf{T}}_I \end{pmatrix}$$

where  $\hat{\mathbf{R}} \triangleq \mathbf{I}_r + r_1 \mathbf{R} + r_2 \mathbf{R}$ ,  $\hat{\mathbf{T}} \triangleq \mathbf{I}_t + t_1 \tilde{\mathbf{T}}$ , and  $\hat{\mathbf{T}}_I \triangleq \mathbf{I}_i + t_2 \tilde{\mathbf{T}}_I$ . Finally, we obtain the asymptotic mean:

$$\mu_{\mathcal{I}} \sim (\ln \det(\mathbf{K}_0) - t_1 r_1 - t_2 r_2) - (\ln \det \mathbf{M}_0 - t_3 r_3). \quad (3.20)$$

$$\left\{ \begin{array}{l} \mathbf{M}_2 \triangleq \begin{pmatrix} \text{tr}((\mathbf{A}_M \mathbf{R})^2) & 1 + \text{tr}(\mathbf{C}_M \mathbf{R} \mathbf{B}_M \tilde{\mathbf{T}}_I) \\ 1 + \text{tr}(\mathbf{B}_M \tilde{\mathbf{T}}_I \mathbf{C}_M \mathbf{R}) & \text{tr}((\mathbf{D}_M \tilde{\mathbf{T}}_I)^2) \end{pmatrix} \\ \mathbf{K}_2 \triangleq - \begin{pmatrix} \text{tr}((\mathbf{A}_K \mathbf{R})^2) & \text{tr}(\mathbf{D}_K \mathbf{R} \mathbf{B}_K \tilde{\mathbf{T}}) + 1 & \\ \text{tr}(\mathbf{B}_K \tilde{\mathbf{T}} \mathbf{D}_K \mathbf{R}) + 1 & \text{tr}((\mathbf{E}_K \tilde{\mathbf{T}})^2) & \dots \\ \text{tr}(\mathbf{A}_K \mathbf{R} \mathbf{A}_K \mathbf{R}) & \text{tr}(\mathbf{D}_K \mathbf{R} \mathbf{B}_K \tilde{\mathbf{T}}) & \\ \text{tr}(\mathbf{C}_K \tilde{\mathbf{T}}_I \mathbf{G}_K \mathbf{R}) & \text{tr}(\mathbf{F}_K \tilde{\mathbf{T}}_I \mathbf{H}_K \tilde{\mathbf{T}}) & \\ \text{tr}(\mathbf{A}_K \mathbf{R} \mathbf{A}_K \mathbf{R}) & \text{tr}(\mathbf{G}_K \mathbf{R} \mathbf{C}_K \tilde{\mathbf{T}}_I) & \\ \text{tr}(\mathbf{B}_K \tilde{\mathbf{T}} \mathbf{D}_K \mathbf{R}) & \text{tr}(\mathbf{H}_K \tilde{\mathbf{T}} \mathbf{F}_K \tilde{\mathbf{T}}_I) & \\ \dots & \text{tr}((\mathbf{A}_K \mathbf{R})^2) & \text{tr}(\mathbf{G}_K \mathbf{R} \mathbf{C}_K \tilde{\mathbf{T}}_I) + 1 \\ \text{tr}(\mathbf{C}_K \tilde{\mathbf{T}}_I \mathbf{G}_K \mathbf{R}) + 1 & \text{tr}((\mathbf{I}_K \tilde{\mathbf{T}}_I)^2) & \\ \mathbf{G}_2 \triangleq - \begin{pmatrix} \text{tr}(\mathbf{C}_K \tilde{\mathbf{T}}_I \mathbf{C}_M \mathbf{R}) - 1 & \text{tr}(\mathbf{I}_K \tilde{\mathbf{T}}_I \mathbf{D}_M \tilde{\mathbf{T}}_I) \\ \text{tr}(\mathbf{A}_K \mathbf{R} \mathbf{A}_M \mathbf{R}) & \text{tr}(\mathbf{G}_K \mathbf{R} \mathbf{B}_M \tilde{\mathbf{T}}_I) - 1 \end{pmatrix} \end{pmatrix} \right. \quad (3.22)$$


---

### 3.5 Asymptotic variance of mutual information

The second-order term in the expansion of (3.15) at the saddle point  $\mathcal{S}$  is given by:

$$\begin{aligned} F_2 \triangleq & \frac{1}{2} \text{str}((\mathfrak{X}^{-1} \delta \mathfrak{X})^2) |_{\mathcal{S}} - \text{tr}(\delta \bar{\Theta}_1 \delta \Theta_1 + \delta \bar{\Theta}_2 \delta \Theta_2) \\ & - \text{tr}(\delta \mathbf{R}_1 \delta \mathbf{T}_1 \delta \mathbf{R}_2 \delta \mathbf{T}_2 - \delta \mathbf{R}_3 \delta \mathbf{T}_3). \end{aligned}$$

Defining the matrices  $\mathbf{M}_2$ ,  $\mathbf{K}_2$ , and  $\mathbf{G}_2$  as in eq. (3.22) on page 34, the contour integrals can be evaluated and yield the following result:

$$G(a) \sim \exp(-F_0) \left( \frac{\det(\mathbf{G}_2)^2}{-\det(\mathbf{M}_2) \det(\mathbf{K}_2)} \right)^a.$$

Therefore, the asymptotic variance is given by:

$$\sigma_{\mathcal{I}}^2 \sim -\ln \det(\mathbf{K}_2) - \ln(-\det(\mathbf{M}_2)) + 2 \ln \det(\mathbf{G}_2). \quad (3.21)$$

### 3.6 Numerical results

We consider a MIMO channel with  $t = r = i = 4$  antennas, SNR of 10 dB,  $K = 10$  dB, and  $K_I = 5$  dB. We assume that the average channel matrices  $\bar{\mathbf{H}}$  and

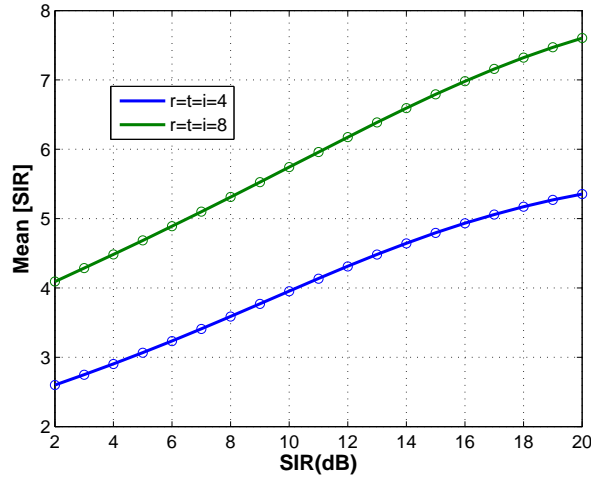


Figure 3.1: Average mutual information mean for MIMO as a function of SIR.  $T$ ,  $T_I$ , and  $R$  have base  $\alpha = 0.7$ . SNR = 10dB,  $K = 10$ dB, and  $K_I = 5$ dB. Solid lines:  $Q$  is proportional to  $I_t$ . Dashed lines:  $Q$  is optimized to achieve the asymptotic ergodic capacity [25]. Thin lines: interference is regarded as Gaussian noise. Circles: Monte-Carlo simulations.

$\bar{H}_I$  are multiple of the all-1 matrices and the spatial correlation matrices  $T$ ,  $T_I$ , and  $R$  have base  $\alpha = 0.7$ . Figs. 3.1 and 3.2 plot the average mutual information mean and standard deviation versus the SIR. Thin lines are obtained by regarding interference as Gaussian noise. They show a considerable reduction of the mean and of the standard deviation in the low SIR regime. In both cases solid lines correspond to  $Q$  proportional to  $I_t$  and dashed lines are obtained by optimizing  $Q$  to achieve the asymptotic ergodic capacity [25]. It can be noticed that both the mean and of the standard deviation are higher when  $Q$  is optimized in the low SIR regime. Circle markers, corresponding to Monte-Carlo simulations, show an excellent agreement with the asymptotic results even for a small number of antennas.

### 3.7 Conclusion

We calculated the asymptotic mean *and* variance of the mutual information of a separately-correlated Rician MIMO channel with interference in a very general setting. Our development involves the use of *superanalysis* [52] to deal with the coupling of the signal and interference part induced by the line-of-sight components in the saddle point approximation.

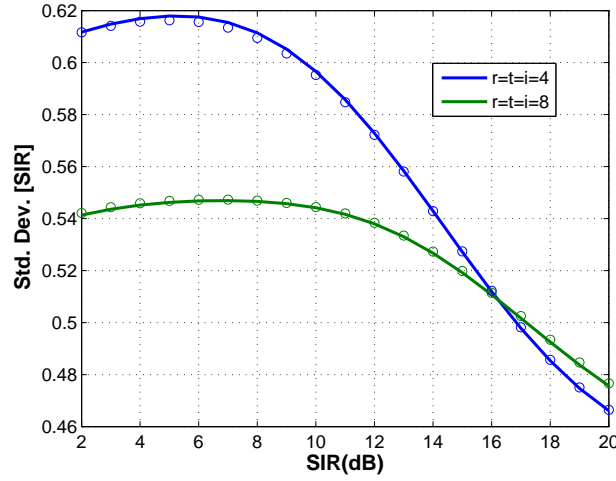


Figure 3.2: Average mutual information standard deviation as a function of SIR. Channel parameters and curves as in Fig. 3.1.

## 3.8 Appendix

### 3.8.1 Identities for complex valued matrices

**Identity 3.8.1** [3,9] (Completing the Square) Let  $\mathbf{A}$ ,  $\mathbf{B}$  be a complex, Hermitian, positive definite  $m \times m$  and  $n \times n$  matrix, respectively, and  $\mathbf{C}$ ,  $\mathbf{D}$  be  $n \times m$  complex matrices. Then

$$\begin{aligned} \int_{\mathbb{C}^{n \times m}} D_c \mathbf{U} \operatorname{etr}(-\pi(\mathbf{A} \mathbf{U}^H \mathbf{B} \mathbf{U} + \mathbf{C}^H \mathbf{U} + \mathbf{U}^H \mathbf{D})) \\ = \det(\mathbf{A})^{-n} \det(\mathbf{B})^{-m} \operatorname{etr}(\pi \mathbf{A}^{-1} \mathbf{C}^H \mathbf{B}^{-1} \mathbf{D}). \end{aligned} \quad (3.23)$$

**Identity 3.8.2** [3] (Hubbard-Stratonovich Transformation) Let  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{R}$ , and  $\mathbf{T}$  be complex  $m \times m$  matrices. Define the contours  $\mu(\mathbf{T}, \mathbf{R})$  for the elements of  $\mathbf{R}$  and  $\mathbf{T}$  over the real and imaginary axis, respectively. Then,

$$\operatorname{etr}(-\mathbf{A} \mathbf{B}) = \int d\mu(\mathbf{T}, \mathbf{R}) \operatorname{etr}(\mathbf{R} \mathbf{T} - \mathbf{A} \mathbf{T} - \mathbf{R} \mathbf{B}). \quad (3.24)$$

### 3.8.2 Identities for Grassmann valued matrices

**Identity 3.8.3** [3] (Completing the Square) Let  $\mathbf{A}$ ,  $\mathbf{B}$  be a complex, Hermitian, positive definite  $m \times m$  and  $n \times n$  matrix, respectively, and  $\Phi$ ,  $\bar{\Xi}^T$  be Grassmann

valued  $n \times m$  matrices. Then

$$\begin{aligned} & \int_{\mathbb{G}^{n \times m}, \mathbb{G}^{m \times n}} D_g(\Psi, \bar{\Psi}) \text{etr}(\mathbf{A} \bar{\Psi} \mathbf{B} \Psi + \bar{\Xi} \Psi + \bar{\Psi} \Phi) \\ &= \det(\mathbf{A})^n \det(\mathbf{B})^m \text{etr}(-\mathbf{A}^{-1} \bar{\Xi} \mathbf{B}^{-1} \Phi). \end{aligned} \quad (3.25)$$

It is important to note that matrices  $\Psi$  and  $\bar{\Psi}$  are *independent* and *not* related by an involution, like Hermitian conjugation in the complex case.

### 3.8.3 Supermatrices, superdeterminant and supertrace

Borrowing from [52, p. 82], we call *supermatrix* a matrix composed in part by complex entries and in part by Grassmann variables. We are interested in particular to the following block supermatrix:

$$\mathfrak{X} = \begin{pmatrix} \mathbf{M} & -\mathbf{\Gamma} \\ \bar{\mathbf{\Gamma}} & \mathbf{K} \end{pmatrix}.$$

In this case, the superdeterminant  $\text{sdet}(\mathfrak{X})$  and the supertrace  $\text{str}(\mathfrak{X})$  are given by [52, p. 99]:

$$\begin{cases} \text{sdet}(\mathfrak{X}) &= \det(\mathbf{M}) \det(\mathbf{K} + \bar{\mathbf{\Gamma}} \mathbf{M}^{-1} \mathbf{\Gamma})^{-1} \\ \text{str}(\mathfrak{X}) &= \text{tr}(\mathbf{M}) - \text{tr}(\mathbf{K}). \end{cases} \quad (3.26)$$

### 3.8.4 Definition of functions and matrices

$$\begin{aligned} \tilde{f}(U, V) &\triangleq \text{etr}(-\pi(U^H \tilde{\mathbf{H}} V - V^H \tilde{\mathbf{H}}^H U)) \\ \tilde{f}_I(U, W) &\triangleq \text{etr}(-\pi(U^H \tilde{\mathbf{H}}_I W - W^H \tilde{\mathbf{H}}_I^H U)) \\ f_{w,I}(U, W) &\triangleq \text{etr}(-\pi(\mathbf{T}_I^{1/2} \mathbf{Q}_I^{1/2} W U^H \mathbf{R}^{1/2} \\ &\quad \times \mathbf{H}_{w,I} - \mathbf{H}_{w,I}^H \mathbf{R}^{1/2} U W^H \mathbf{Q}_I^{1/2} \mathbf{T}_I^{1/2})). \end{aligned} \quad (3.27)$$

$$\begin{aligned} \tilde{h}_I(\Psi, \bar{\Psi}, \Omega, \bar{\Omega}) &\triangleq \text{etr}(\bar{\Omega} \tilde{\mathbf{H}}_I^H \Psi - \bar{\Psi} \tilde{\mathbf{H}}_I \Omega) \\ h_{w,I}(\Psi, \bar{\Psi}, \Omega, \bar{\Omega}) &\triangleq \text{etr}(\mathbf{T}_I^{1/2} \mathbf{Q}_I^{1/2} \Omega \bar{\Psi} \mathbf{R}^{1/2} \mathbf{H}_{w,I} \\ &\quad - \mathbf{H}_{w,I}^H \mathbf{R}^{1/2} \Psi \bar{\Omega} \mathbf{Q}_I^{1/2} \mathbf{T}_I^{1/2}). \end{aligned} \quad (3.28)$$

$$\begin{aligned} f(U, V) &\triangleq \text{etr}(-\pi^2(V^H \tilde{\mathbf{T}} V U^H \mathbf{R} U)) \\ f_I(U, W) &\triangleq \text{etr}(-\pi^2(W^H \tilde{\mathbf{T}}_I W U^H \mathbf{R} U)). \end{aligned} \quad (3.29)$$

$$\begin{aligned} h_I(U, W; \Psi, \bar{\Psi}, \Omega, \bar{\Omega}) &\triangleq \text{etr}(\bar{\Omega} \tilde{\mathbf{T}}_I \Omega \bar{\Psi} \mathbf{R} \Psi) \\ &\quad \times \text{etr}(-\pi(\bar{\Psi} \mathbf{R} U W^H \tilde{\mathbf{T}}_I \Omega + \bar{\Omega} \tilde{\mathbf{T}}_I W U^H \mathbf{R} \Psi)). \end{aligned} \quad (3.30)$$

$$\begin{aligned}
K &\triangleq \begin{pmatrix} I_{ra} + \tilde{R} & \tilde{H} \otimes I_a & \tilde{H}_I \otimes I_a \\ -\tilde{H}^H \otimes I_a & I_{ta} + \tilde{T} \otimes T_1^T & 0 \\ -\tilde{H}_I^H \otimes I_a & 0 & I_{ia} + \tilde{T}_I \otimes T_2^T \end{pmatrix} \\
M &\triangleq \begin{pmatrix} I_{ra} - R \otimes R_3^T & -\tilde{H}_I \otimes I_a \\ \tilde{H}_I^H \otimes I_a & I_{ia} + \tilde{T}_I \otimes T_3^T \end{pmatrix} \\
\Gamma &\triangleq \begin{pmatrix} R \otimes \Theta_1^T & 0 & 0 \\ 0 & 0 & \tilde{T}_I \otimes \Theta_2^T \end{pmatrix} \\
\bar{\Gamma} &\triangleq \begin{pmatrix} R \otimes \bar{\Theta}_2^T & 0 \\ 0 & 0 \\ 0 & \tilde{T}_I \otimes \bar{\Theta}_1^T \end{pmatrix}
\end{aligned} \tag{3.31}$$

with  $\tilde{R} \triangleq R \otimes R_1^T + R \otimes R_2^T$ .

$$M^{-1}|_S = \begin{pmatrix} A_M & B_M \\ C_M & D_M \end{pmatrix} \otimes I_a, \tag{3.32}$$

with

$$\begin{aligned}
A_M &\triangleq (I_r + r_3 R + \tilde{H}_I (I_i + t_3 \tilde{T}_I)^{-1} \tilde{H}_I^H)^{-1} \\
B_M &\triangleq (I_r + r_3 R)^{-1} \tilde{H}_I D_M \\
C_M &\triangleq - (I_i + t_3 \tilde{T}_I)^{-1} \tilde{H}_I^H A_M \\
D_M &\triangleq (I_i + t_3 \tilde{T}_I + \tilde{H}_I^H (I_r + r_3 R)^{-1} \tilde{H}_I)^{-1}.
\end{aligned} \tag{3.33}$$

$$K^{-1}|_S = \begin{pmatrix} A_K & B_K & C_K \\ D_K & E_K & F_K \\ G_K & H_K & I_K \end{pmatrix} \otimes I_a, \tag{3.34}$$

with

$$\begin{aligned}
A_K &\triangleq (\hat{R} + \tilde{H}\hat{T}^{-1}\tilde{H}^H + \tilde{H}_I\hat{T}_I^{-1}\tilde{H}_I^H)^{-1} \\
B_K &\triangleq -\hat{R}^{-1}(\tilde{H}E_K + \tilde{H}_IH_K) \\
C_K &\triangleq -\hat{R}^{-1}(\tilde{H}F_K + \tilde{H}_II_K) \\
D_K &\triangleq \hat{T}^{-1}\tilde{H}^HA_K \\
E_K &\triangleq (\hat{T} + \tilde{H}^H\hat{R}^{-1}\tilde{H} - \tilde{H}^H\hat{R}^{-1}\tilde{H}_I \\
&\quad (\hat{T}_I + \tilde{H}_I^H\hat{R}^{-1}\tilde{H}_I)^{-1}\tilde{H}_I^H\hat{R}^{-1}\tilde{H})^{-1} \\
F_K &\triangleq -(\hat{T} + \tilde{H}^H\hat{R}^{-1}\tilde{H})^{-1}\tilde{H}^H\hat{R}^{-1}\tilde{H}_II_K \\
G_K &\triangleq \hat{T}_I^{-1}\tilde{H}_I^HA_K \\
H_K &\triangleq -(\hat{T}_I + \tilde{H}_I^H\hat{R}^{-1}\tilde{H}_I)^{-1}\tilde{H}_I^H\hat{R}^{-1}\tilde{H}E_K \\
I_K &\triangleq (\hat{T}_I + \tilde{H}_I^H\hat{R}^{-1}\tilde{H}_I - \tilde{H}_I^H\hat{R}^{-1}\tilde{H} \\
&\quad (\hat{T} + \tilde{H}^H\hat{R}^{-1}\tilde{H})^{-1}\tilde{H}^H\hat{R}^{-1}\tilde{H}_I)^{-1}.
\end{aligned} \tag{3.35}$$

## Chapter 4

# On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel

*Giorgio Taricco, Erwin Riegler*

### 4.1 Introduction

An important problem of network information theory is the derivation of the achievable rate region of a multiple-access channel. In spite of intense research efforts carried out through the recent decades, there are several open problems in this area that have wide implications for the theory of communications and computation [15].

The multiple access channel achievable rate region is the set of rate vectors that are achievable by the different channel users and has been studied exhaustively in the literature [15]. The achievable rate region admits a simple expression for the Gaussian multiple access channel that has been extended to the Gaussian multiple access MIMO channel by Yu *et al.* [26] who provided an iterative water-filling algorithm aimed at finding the optimum user signal covariance matrices that maximize the sum-rate of the channel. Their result applies when the multiple access MIMO channel is perfectly known at the transmitter and at the receiver. However, when only the channel state information at the receiver (CSIR) is available, the problem of finding the maximum ergodic sum-rate achieving covariance matrices is still open in the general setting [16]. Nevertheless, a notable result in this area has been provided by Hösli *et al.* [2], who proved that the ergodic and outage achievable rate region increases (as sets) monotonically with the singular values of the line-of-sight component of the channel matrix.



In this work we assume that the receiver has full CSIR and the transmitter knows the statistics of the channel, i.e. the transmitter has channel distribution information (CDIT). Based on these assumptions we provide an algorithm to find the maximum ergodic sum-rate achieving covariance matrices of a multiple-access MIMO channel when the number of transmit and receive antennas grow asymptotically large with finite asymptotic ratios and the number of users and their SNR's are finite. In this context we assume that the multiple-access communication channel is affected by Rician fading with separate spatial correlation (with a common receive part and different transmit parts). Our results rely on a previous work [25] where the ergodic capacity achieving covariance matrix was obtained for a separately-correlated Rician fading MIMO channel with multiple-access interference, which extended previous results due to Moustakas *et al.* [3] relevant to the case of Rayleigh fading. Similar results for the separately-correlated Rician fading MIMO channel (without multiple-access interference) were obtained independently by Dumont *et al.* [31], using an asymptotic method based on Stieltjes transforms, and by Vu and Paulraj [50], using an interior point with barrier optimization method [14].

assumptions illustrates the and

## 4.2 System model

We consider a narrowband multiple access separately-correlated block Rician fading MIMO channel with  $K$  users with  $t_k$  transmit antennas for each user  $k \in \mathcal{K} \triangleq \{1, \dots, K\}$  and a single receiver with  $r$  receive antennas. The channel is characterized by the following equation:

$$\mathbf{y} = \sum_{k \in \mathcal{K}} \mathbf{H}_k \mathbf{x}_k + \mathbf{z}. \quad (4.1)$$

Here,  $\mathbf{x}_k \in \mathbb{C}^{t_k \times 1}$  is the transmitted signal vector of user  $k$ ,  $\mathbf{H}_k \in \mathbb{C}^{r \times t_k}$  is the channel matrix of user  $k$ ,  $\mathbf{z} \in \mathbb{C}^{r \times 1}$  is the additive noise vector, and  $\mathbf{y} \in \mathbb{C}^{r \times 1}$  is the received signal vector. Each  $\mathbf{x}_k$  is assumed to have zero mean with covariance matrix  $\mathbf{Q}_k$ . We assume that the additive noise vector has zero mean and covariance matrix  $\mathbf{Q}_Z = \mathbb{E}[\mathbf{z}\mathbf{z}^H]$ .

The channel matrices  $\mathbf{H}_k$  are assumed to be of separately (or Kronecker) correlated Rician fading type *with common receive correlation*. Thus, they can be written as

$$\mathbf{H}_k = \bar{\mathbf{H}}_k + \mathbf{R}^{1/2} \mathbf{W}_k \mathbf{T}_k^{1/2} \quad \forall k \in \mathcal{K},$$

where, for each user  $k$ ,  $\bar{\mathbf{H}}_k$  represents the average channel matrix related to the presence of a line-of-sight signal component in the multipath fading channel, the

Hermitian positive definite matrix  $\mathbf{R}$  characterizes the spatial receive correlation component, and the Hermitian positive definite matrices  $\mathbf{T}_k$  characterize the spatial transmit correlation component. The random matrices  $\mathbf{W}_k$  have iid  $\mathcal{N}_c(0, 1)$  entries.

Following standard assumptions [1, 9, 25], we define the Rice Factor of user  $k$  as

$$K_k^R \triangleq \frac{\|\bar{\mathbf{H}}_k\|^2}{\text{Tr}(\mathbf{T}_k) \text{Tr}(\mathbf{R})} \quad (4.2)$$

and the signal-to-noise power ratio  $\text{SNR}_k$  of user  $k$  as

$$\text{SNR}_k \triangleq \frac{(K_k^R + 1) \text{Tr}(\mathbf{T}_k) \text{Tr}(\mathbf{R}) \text{Tr}(\mathbf{Q}_k/t_k)}{\text{Tr}(\mathbf{Q}_Z)}. \quad (4.3)$$

Notice that the latter assumption allows to overcome the difficulties related with the consideration of the receive SNR

$$\text{SNR}_{k,\text{RX}} \triangleq \frac{\mathbb{E}[\text{Tr}(\mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H)]}{\text{Tr}(\mathbf{Q}_Z)}$$

which arise when we try to make any optimization based on a set of transmit power constraints of the type  $\text{Tr}(\mathbf{Q}_k) \leq P_k$  since the receive SNR  $\text{SNR}_{k,\text{RX}}$  is not proportional to  $\text{Tr}(\mathbf{Q}_k)$  (unless  $\mathbf{Q}_k$  is a multiple of the identity matrix). The assumption  $\mathbf{Q}_k = \frac{P_k}{t_k} \mathbf{I}_{t_k}$  will be referred to in the following as *iid power allocation*.

### 4.3 Ergodic Capacity Region

Let  $\mathbf{R}_K \triangleq (R_1, \dots, R_K) \in \mathbb{R}_+^K$  be the vector of rates for all users  $k \in \mathcal{K}$  and denote by  $\mathcal{S} \triangleq \{k_1, \dots, k_{|\mathcal{S}|}\} \subset \mathcal{K}$  subsets of  $\mathcal{K}$  with complement  $\mathcal{S}^c \triangleq \mathcal{K} \setminus \mathcal{S}$ . Then, the achievable rate region of the multiple access channel can be defined by extension of the results from [15] as the set:

$$\mathcal{R} = \left\{ \mathbf{R}_K \in \mathbb{R}_+^K \mid \sum_{k \in \mathcal{S}} R_k \leq \mathbb{E}[I(\mathbf{x}_{\mathcal{S}}; \mathbf{y} | \mathbf{x}_{\mathcal{S}^c})] \forall \mathcal{S} \subset \mathcal{K} \right\}, \quad (4.4)$$

where  $\mathbf{x}_{\mathcal{S}} \triangleq (\mathbf{x}_k \mid k \in \mathcal{S})$  and

$$\begin{aligned} \mathbb{E}[I(\mathbf{x}_{\mathcal{S}}; \mathbf{y} | \mathbf{x}_{\mathcal{S}^c})] &\triangleq \mathbb{E}[\ln \det(\mathbf{H}_{\mathcal{S}} \mathbf{Q}_{\mathcal{S}} \mathbf{H}_{\mathcal{S}}^H + \mathbf{Q}_Z)] \\ &\quad - \ln \det(\mathbf{Q}_Z), \end{aligned} \quad (4.5)$$

with

$$\mathbf{H}_{\mathcal{S}} \triangleq (\mathbf{H}_k \mid k \in \mathcal{S}) \quad \text{and} \quad \mathbf{Q}_{\mathcal{S}} \triangleq \text{diag}(\mathbf{Q}_k \mid k \in \mathcal{S}). \quad (4.6)$$

We restrict our attention to the dominant face

$$\mathcal{D} \triangleq \mathcal{R} \cap \left\{ \mathbf{R}_{\mathcal{K}} \in \mathbb{R}_+^K \mid \sum_{k \in \mathcal{K}} R_k = \mathbb{E}[I(\mathbf{x}; \mathbf{y})] \right\}, \quad (4.7)$$

where  $\mathbf{x} \triangleq \mathbf{x}_{\mathcal{K}}$  and

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] = \mathbb{E}[\ln \det(\mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{Q}_Z)] - \ln \det(\mathbf{Q}_Z) \quad (4.8)$$

with  $\mathbf{H} \triangleq \mathbf{H}_{\mathcal{K}}$  and  $\mathbf{Q} \triangleq \mathbf{Q}_{\mathcal{K}}$ . The reason for this is that every point  $\mathbf{R}_{\mathcal{K}} \in \mathcal{R}$  is dominated by (i.e., it has no rate component higher than) a point in  $\mathcal{D}$ . The ergodic capacity of the multiple access MIMO channel is defined as

$$\mathcal{C} \triangleq \max_{\mathbf{Q}_{\mathcal{K}}: \text{Tr}(\mathbf{Q}_{\mathcal{K}}) \leq P_{\mathcal{K}}} \mathbb{E}[I(\mathbf{x}; \mathbf{y})], \quad (4.9)$$

with power constraints  $\text{Tr}(\mathbf{Q}_k) \leq P_k$  ( $k \in \mathcal{K}$ ). The corresponding achievable rate region will be said *ergodic capacity region*.

In order to solve this optimization problem we resort to a recent result allowing to calculate the average mutual information of a separately-correlated Rician MIMO channel when the number of transmit and receive antennas grow asymptotically large [8, 9]. Summarizing, as  $t_k, r \rightarrow \infty$  with  $0 < t_k/r < \infty$ , the asymptotic average mutual information with channel matrix  $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}$ , noise covariance  $\mathbf{Q}_Z = \mathbf{I}_r$ , and signal covariance  $\mathbf{Q}$ , is given by

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] \sim \mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}) \quad (4.10)$$

nat/complex dimension, where we defined

$$\mu_I(\bar{\mathbf{H}}, \mathbf{R}, \mathbf{T}, \mathbf{Q}) \triangleq \ln \det \begin{pmatrix} \mathbf{I}_r + w\mathbf{R} & \tilde{\mathbf{H}} \\ -\tilde{\mathbf{H}}^H & \mathbf{I}_t + z\tilde{\mathbf{T}} \end{pmatrix} - wz \quad (4.11)$$

where  $\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{1/2}\mathbf{T}\mathbf{Q}^{1/2}$  and  $w, z$  can be obtained by solving the equations

$$\begin{cases} w = \text{Tr} \left\{ [z\mathbf{I}_t + \tilde{\mathbf{T}}^{-1} + \tilde{\mathbf{T}}^{-1}\tilde{\mathbf{H}}^H(\mathbf{I}_r + w\mathbf{R})^{-1}\tilde{\mathbf{H}}]^{-1} \right\} \\ z = \text{Tr} \left\{ [w\mathbf{I}_r + \mathbf{R}^{-1} + \mathbf{R}^{-1}\tilde{\mathbf{H}}(\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}^H]^{-1} \right\} \end{cases} \quad (4.12)$$

with  $\tilde{\mathbf{H}} \triangleq \bar{\mathbf{H}}\mathbf{Q}^{1/2}$  and  $\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{1/2}\mathbf{T}\mathbf{Q}^{1/2}$ . The solution of eqs. (4.12) can be shown to be a pair of positive real numbers. Eqs. (4.11) and (4.12) have been established independently in [65] for the uncorrelated Rician MIMO channel.

This result can be applied to calculate the average mutual information (4.8):

$$\mathbb{E}[I(\mathbf{x}; \mathbf{y})] = \mu_I(\bar{\mathbf{H}}_Z, \mathbf{R}_Z, \mathbf{T}, \mathbf{Q}),$$

where we defined  $\bar{\mathbf{H}}_Z \triangleq \mathbf{Q}_Z^{-1/2}(\bar{\mathbf{H}}_k \mid k \in \mathcal{K})$ ,  $\mathbf{R}_Z \triangleq \mathbf{Q}_Z^{-1/2}\mathbf{R}\mathbf{Q}_Z^{-1/2}$ , and  $\mathbf{T} \triangleq \text{diag}(\mathbf{T}_k \mid k \in \mathcal{K})$ . The optimum covariance matrix is obtained by maximizing the average mutual information (4.8) under the power constraints  $\text{Tr}(\mathbf{Q}_k) \leq P_k$  ( $k \in \mathcal{K}$ ).

Then we notice that the constraints of this optimization problem satisfy Slater's condition so that the Karush-Kuhn-Tucker (KKT) condition is necessary and sufficient for optimality [14]. Therefore, we define the Lagrangian dual function:

$$\begin{aligned} \mathcal{L}(\mathbf{S}_k \mid k \in \mathcal{K}) = & \ln \det \begin{pmatrix} \mathbf{I}_r + w\mathbf{R}_Z & \tilde{\mathbf{H}}_Z \\ -\tilde{\mathbf{H}}_Z^H & \mathbf{I}_t + z\tilde{\mathbf{T}} \end{pmatrix} \\ & - wz - \sum_{k \in \mathcal{K}} \lambda_k [\text{Tr}(\mathbf{S}_k^2) - \tilde{P}_k], \end{aligned} \quad (4.13)$$

where  $0 \leq \tilde{P}_k \leq P_k$ ,  $\mathbf{S}_k \triangleq \mathbf{Q}_k^{1/2}\mathbf{S}$ ,  $\mathbf{S} \triangleq \mathbf{Q}^{1/2}$ ,  $\tilde{\mathbf{H}}_Z \triangleq \bar{\mathbf{H}}_Z\mathbf{S}$ ,  $\tilde{\mathbf{T}} \triangleq \mathbf{S}\mathbf{T}\mathbf{S}$ , and  $t \triangleq \sum_{k \in \mathcal{K}} t_k$ . Next, the KKT condition is derived by calculating the first-order total variation of (4.13):

$$\begin{aligned} \delta \mathcal{L} = & \text{Tr} \left[ \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ \mathbf{C}_1 & \mathbf{D}_1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{R}_Z \delta w & \tilde{\mathbf{H}}_Z \delta \mathbf{S} \\ -\delta \mathbf{S} \tilde{\mathbf{H}}_Z^H & \mathbf{S} \mathbf{T} \mathbf{S} \delta z + z(\delta \mathbf{S} \mathbf{T} \mathbf{S} + \mathbf{S} \mathbf{T} \delta \mathbf{S}) \end{pmatrix} \right] \\ & - w \delta z - z \delta w - 2 \sum_{k \in \mathcal{K}} \lambda_k \text{Tr}(\mathbf{S}_k \delta \mathbf{S}_k), \end{aligned} \quad (4.14)$$

where [8]

$$\begin{cases} \mathbf{A}_1 &= [\mathbf{I}_r + w\mathbf{R}_Z + \tilde{\mathbf{H}}_Z(\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}_Z^H]^{-1} \\ \mathbf{B}_1 &= -(\mathbf{I}_r + w\mathbf{R}_Z)^{-1}\tilde{\mathbf{H}}_Z\mathbf{D}_1 \\ \mathbf{C}_1 &= (\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}_Z^H\mathbf{A}_1 \\ \mathbf{D}_1 &= [\mathbf{I}_t + z\tilde{\mathbf{T}} + \tilde{\mathbf{H}}_Z^H(\mathbf{I}_r + w\mathbf{R}_Z)^{-1}\tilde{\mathbf{H}}_Z]^{-1}. \end{cases} \quad (4.15)$$

Since we have from (4.12)  $w = \text{Tr}(\mathbf{D}_1\tilde{\mathbf{T}})$ ,  $z = \text{Tr}(\mathbf{A}_1\mathbf{R}_Z)$ , and [25]

$$(\mathbf{I}_r + w\mathbf{R}_Z)^{-1}\tilde{\mathbf{H}}_Z\mathbf{D}_1 = [(\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\tilde{\mathbf{H}}_Z^H\mathbf{A}_1]^H,$$

we end up with:

$$\begin{aligned} \delta \mathcal{L} = & 2 \text{Tr}(\mathcal{H}[(\tilde{\mathbf{H}}_Z^H(\mathbf{I}_r + w\mathbf{R}_Z)^{-1}\tilde{\mathbf{H}}_Z + z\mathbf{T})\mathbf{S}\mathbf{D}_1]\delta \mathbf{S}) \\ & - 2 \sum_{k \in \mathcal{K}} \lambda_k \text{Tr}(\mathbf{S}_k \delta \mathbf{S}_k), \end{aligned} \quad (4.16)$$

where  $\mathcal{H}[\mathbf{A}] \triangleq (\mathbf{A} + \mathbf{A}^H)/2$ . Thus, the first order total variation  $\delta L$  is null (and the KKT condition is satisfied) provided that

$$\mathbf{S}_k = \lambda_k^{-1} \mathcal{H}[(\bar{\mathbf{H}}_Z^H (\mathbf{I}_r + w \mathbf{R}_Z)^{-1} \bar{\mathbf{H}}_Z + z \mathbf{T}) \mathbf{S} \mathbf{D}_1] \big|_k, \quad (4.17)$$

where  $\mathbf{A}|_k$  is the submatrix of  $\mathbf{A}$  obtained by extracting the elements of the rows and columns with indexes from  $\sum_{i=1}^{k-1} t_i + 1$  to  $\sum_{i=1}^k t_i$ , and  $\lambda_k$  is obtained from the constraint  $\text{Tr}(\mathbf{S}_k^2) = P_k$ , namely,<sup>1</sup>

$$\lambda_k = \sqrt{\frac{\text{Tr}\{\{\mathcal{H}[(\bar{\mathbf{H}}_Z^H (\mathbf{I}_r + w \mathbf{R}_Z)^{-1} \bar{\mathbf{H}}_Z + z \mathbf{T}) \mathbf{S} \mathbf{D}_1] \big|_k\}^2\}}{P_k}} \quad (4.18)$$

If all  $\mathbf{S}_k$  ( $k \in \mathcal{K}$ ) satisfy equation (4.17), we define

$$\mathbf{Q}_O \triangleq \text{diag}(\mathbf{S}_k^2 \mid k \in \mathcal{K}) \quad (4.19)$$

the *optimum* covariance matrix, which can be found using the following algorithm:

**Algorithm 1** (*Iterative water-filling – optimum covariance*)

initialize  $\mathbf{S} \big|_k = \sqrt{P_k/t_k} \mathbf{I}_{t_k}$ ,  $k \in \mathcal{K}$

repeat

  for  $k = 1$  to  $K$

    obtain  $\lambda_k$  from (4.18)

    set  $\mathbf{S} \big|_k = \lambda_k^{-1} \mathcal{H}[(\bar{\mathbf{H}}_Z^H (\mathbf{I}_N + w \mathbf{R}_Z)^{-1} \bar{\mathbf{H}}_Z + z \mathbf{T}) \mathbf{S} \mathbf{D}_1] \big|_k$

  end

until  $\mathbf{S}$  converges. Set  $\mathbf{Q}_O = \mathbf{S}^2$ .

### 4.3.1 Jensen approximation

Here we consider an approximation of the mutual information based on Jensen's inequality and derive the covariance matrices that maximize it. We know from [15, 30, 40] that the following inequalities hold for positive definite Hermitian matrices  $\mathbf{A}$ ,  $\mathbf{X}$  (with  $\mathbf{X}$  random):

$$\mathbb{E}[\ln \det(\mathbf{X})] \leq \ln \det(\mathbb{E}[\mathbf{X}]), \quad (4.20)$$

$$\mathbb{E}[\ln \det(\mathbf{I} + \mathbf{X}^{-1} \mathbf{A})] \geq \ln \det(\mathbf{I} + \mathbb{E}[\mathbf{X}]^{-1} \mathbf{A}). \quad (4.21)$$

---

<sup>1</sup> The positive definiteness of  $\mathbf{S}_k$  derives from the positivity of  $w, z$ .

Applying inequality (4.21) to the mutual information (4.8) gives:

$$\begin{aligned}
\mathbb{E}[I(\mathbf{x}; \mathbf{y})] &= \mathbb{E}[\ln \det(\mathbf{H}\mathbf{Q}\mathbf{H}^H + \mathbf{Q}_Z)] - \ln \det(\mathbf{Q}_Z) \\
&= \mathbb{E}[\ln \det(\mathbf{H}_{\hat{k}}\mathbf{Q}_{\hat{k}}\mathbf{H}_{\hat{k}}^H + \mathbf{Q}_Z)] - \ln \det(\mathbf{Q}_Z) \\
&+ \mathbb{E}[\ln \det(\mathbf{I}_r + (\mathbf{H}_{\hat{k}}\mathbf{Q}_{\hat{k}}\mathbf{H}_{\hat{k}}^H + \mathbf{Q}_Z)^{-1}\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H)] \\
&\geq \mathbb{E}[\ln \det(\mathbf{H}_{\hat{k}}\mathbf{Q}_{\hat{k}}\mathbf{H}_{\hat{k}}^H + \mathbf{Q}_Z)] - \ln \det(\mathbf{Q}_Z) \\
&+ \mathbb{E}[\ln \det(\mathbf{I}_r + \mathbf{Q}_{IZ,k}^{-1}\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H)], \tag{4.22}
\end{aligned}$$

where we defined  $\mathbf{H}_{\hat{k}} \triangleq \mathbf{H}_{\mathcal{K} \setminus \{k\}}$ ,  $\mathbf{Q}_{\hat{k}} \triangleq \mathbf{Q}_{\mathcal{K} \setminus \{k\}}$ , and  $\mathbf{Q}_{IZ,k} \triangleq \mathbb{E}[\mathbf{H}_{\hat{k}}\mathbf{Q}_{\hat{k}}\mathbf{H}_{\hat{k}}^H + \mathbf{Q}_Z]$  for  $k \in \mathcal{K}$ . Then, applying (4.20) to the last term in (4.22), we obtain

$$\begin{aligned}
&\mathbb{E}[\ln \det(\mathbf{I}_r + \mathbf{Q}_{IZ,k}^{-1}\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^H)] \\
&\leq \ln \det(\mathbf{I}_{t_k} + \mathbb{E}[\mathbf{H}_k^H \mathbf{Q}_{IZ,k}^{-1} \mathbf{H}_k] \mathbf{Q}_k) \tag{4.23}
\end{aligned}$$

Finally, if we treat inequalities as approximations, the mutual information (4.8) is approximately given by

$$\begin{aligned}
\mathbb{E}[I(\mathbf{x}; \mathbf{y})] &\approx \mathbb{E}[\ln \det(\mathbf{H}_{\hat{k}}\mathbf{Q}_{\hat{k}}\mathbf{H}_{\hat{k}}^H + \mathbf{Q}_Z)] - \ln \det(\mathbf{Q}_Z) \\
&+ \ln \det(\mathbf{I}_{t_k} + \mathbb{E}[\mathbf{H}_k^H \mathbf{Q}_{IZ,k}^{-1} \mathbf{H}_k] \mathbf{Q}_k) \tag{4.24}
\end{aligned}$$

for any  $k \in \mathcal{K}$  and the last term can then be maximized by water-filling [15].

Thus, we call *Jensen approximation* the set of covariance matrices  $\{\mathbf{Q}_{J,k} \mid k \in \mathcal{K}\}$  that maximize

$$\ln \det(\mathbf{I}_{t_k} + \mathbb{E}[\mathbf{H}_k^H \mathbf{Q}_{IZ,k}^{-1} \mathbf{H}_k] \mathbf{Q}_k) \tag{4.25}$$

for all  $k \in \mathcal{K}$  and denote by

$$\mathbf{Q}_J \triangleq \text{diag}(\mathbf{Q}_{J,k} \mid k \in \mathcal{K}) \tag{4.26}$$

the *Jensen covariance matrix* where the matrices  $\mathbf{Q}_{J,k}$  can be found by using the following algorithm:

**Algorithm 2** (*Iterative water-filling – Jensen approx.*)

*initialize*  $\mathbf{Q} \mid_k = (P_k/t_k)\mathbf{I}_{t_k}$ ,  $k \in \mathcal{K}$

*repeat*

*for*  $k = 1$  *to*  $K$

*factor*  $\mathbb{E}[\mathbf{H}_k^H \mathbf{Q}_{IZ,k}^{-1} \mathbf{H}_k] = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^H$

*solve*  $\text{Tr}[(\mu_k \mathbf{I}_{t_k} - \mathbf{\Lambda}_k^{-1})_+] = P_k$  *for*  $\mu_k$ <sup>2</sup>

*set*  $\mathbf{Q} \mid_k = \mathbf{U}_k (\mu_k \mathbf{I}_{t_k} - \mathbf{\Lambda}_k^{-1})_+ \mathbf{U}_k^H$

*end*

*until*  $\mathbf{Q}$  *converges and then all the terms* (4.25) *are maximized.*

---

<sup>2</sup> $(\mathbf{A})_+$  is the matrix with entries  $\max\{0, (\mathbf{A})_{ij}\}$

## 4.4 Numerical results

In this section we present numerical results illustrating the accuracy of the asymptotic analytic optimization of the sum-rate capacity. We consider the following scenarios:

1. Multiple-access channel with  $K = 2$  users with  $t_k = 4$  transmit antennas and with  $r = 4$  receive antennas, all users have the same SNR.
2. Multiple-access channel with  $K = 4$  users with  $t_k = 4$  transmit antennas and with  $r = 4$  receive antennas, all users have the same SNR.
3. Multiple-access channel with  $K = 4$  users with  $t_k = 4$  transmit antennas and with  $r = 4$  receive antennas. There are three *strong users* with the same SNR and one *weak user* with SNR 10 dB below the strong user SNR. This scenario allows to assess the *near-far effect* of the multiuser MIMO system considered in [67].

In all scenarios it is assumed that the line-of-sight matrices  $\bar{\mathbf{H}}_k$  have all constant entries (i.e.,  $(\bar{\mathbf{H}}_k)_{ij} = 1$  for all  $k \in \mathcal{K}$ ), the spatial correlation matrices are of exponential type (i.e.,  $(\mathbf{R})_{ij} = \alpha^{|i-j|}$  and  $(\mathbf{T}_k)_{ij} = \alpha_k^{|i-j|}$  with bases  $\alpha = 0.7$  and  $\alpha_k = 0.7$  for all  $k \in \mathcal{K}$ ), and the Rice factors are  $K_k^R = 10$  dB for all  $k \in \mathcal{K}$ .

### 4.4.1 First scenario ( $K = 2$ ), symmetric

Figs. 4.1 and 4.2 refer to the first scenario described above and illustrate the achievable rate regions corresponding to the optimum covariance and iid power allocation (Fig. 4.1) and to the optimum and Jensen covariance matrices (Fig. 4.2) for several user SNR values. Solid and dashed lines are based on the analytic asymptotic approximations with the corresponding covariance matrices. Markers are based on Monte-Carlo simulations corresponding to the same covariance matrices. It can be noticed that there is always an excellent agreement between analytic asymptotic approximations and Monte-Carlo simulations. This also confirms the accuracy of the optimum covariance matrices derived with the analytic asymptotic approximation.

Fig. 4.1 shows a considerable difference between the achievable rate region corresponding to the optimum covariance matrix and that corresponding to iid power allocation. As expected, the differences get lower as the user SNR's get larger but are still notable at 20 dB. Fig. 4.2 instead shows that Jensen approximation is very close to the optimum for moderate SNR's though the difference becomes noticeable above 15 dB.

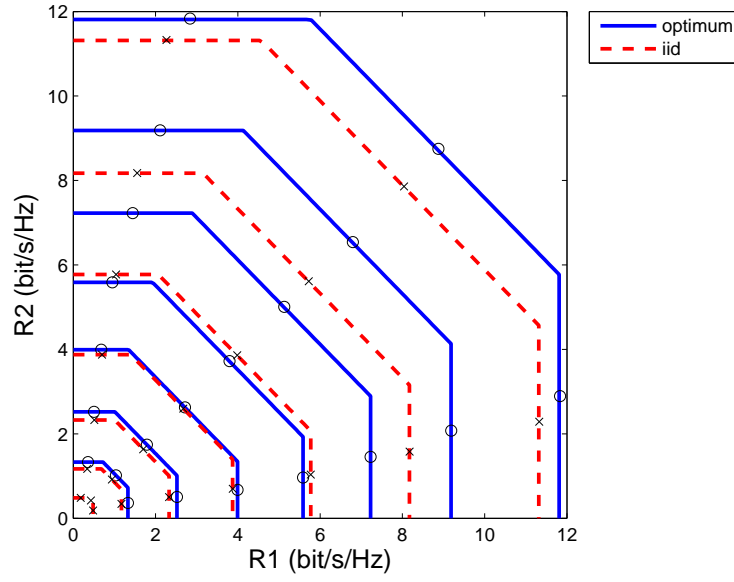


Figure 4.1: Achievable rate region of the MIMO multiple access channel for scenario 1 ( $K = 2$  users),  $\text{SNR}_k \in \{-10, -5, \dots, 15, 20\}$  dB. Solid lines correspond to the optimum covariance matrix  $\mathbf{Q}_O$  from eq. (4.19). Dashed lines correspond to iid power allocation. Markers correspond to Monte-Carlo simulations with the corresponding covariance matrices.



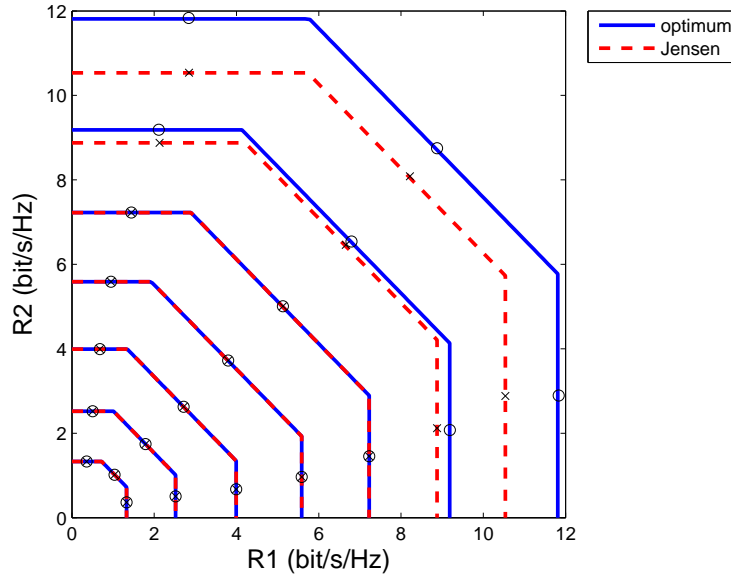


Figure 4.2: Same as Fig. 4.1 but dashed lines correspond to the Jensen approximation covariance matrix  $\mathbf{Q}_J$ .

#### 4.4.2 Second scenario ( $K = 4$ , symmetric)

The capacity of the second scenario ( $K = 4$ ) is reported in Fig. 4.3. In this case the achievable rate region cannot be illustrated directly, as it would require a four-dimensional picture. However, it is completely determined by the four achievable sum-rates  $\mathbb{E}[\mathbf{I}(\mathbf{x}_S; \mathbf{y}|\mathbf{x}_{S^c})]$  corresponding to  $|\mathcal{S}| = 1$  to 4, because of the symmetric choice of parameters and user SNR's. In the figure, solid curves correspond to the asymptotic optimum covariance matrices, dashed curves to iid power allocation, and dash-dot curves to Jensen approximation. Each type of curve corresponds to  $|\mathcal{S}| = 1$  to 4, from bottom to top. Markers represent corresponding Monte-Carlo simulation results. Again, it can be noticed that there is always an excellent agreement between analytic asymptotic approximations and Monte-Carlo simulations, which confirms the accuracy of the optimum covariance matrices derived with the asymptotic approximation.

It can be noticed that there is a close correspondence between asymptotic optimum and Jensen approximation results over the SNR range considered. On the contrary, the diagrams display marked differences between optimum/Jensen and iid results, that confirm the consistent suboptimality of iid power allocation.

### 4.4.3 Third scenario ( $K = 4$ , asymmetric)

The capacity of the third scenario is illustrated by Figs. 4.4 and 4.5. The assumption of having three strong users and one weak user (with a 10 dB-lower SNR) implies that the achievable rate region is determined by 7 sum-rates for every SNR, four of which are included in Fig. 4.4 and three in Fig. 4.5. More precisely, Fig. 4.4 shows the sum-rates  $\mathbb{E}[\mathbf{I}(\mathbf{x}_{\mathcal{S}}; \mathbf{y} | \mathbf{x}_{\mathcal{S}^c})]$  corresponding to  $\mathcal{S} = \{1\}, \{k\}, \{1, k\}, \{k, k'\}$  with  $k, k' \neq 1$  (from bottom to top). Fig 4.5 shows the sum-rates corresponding to  $\mathcal{S} = \{1, k, k'\}, \{k, k', k''\}, \{1, 2, 3, 4\}$  with  $k, k', k'' \neq 1$  (from bottom to top).

In all cases, iid power allocation is again shown to be considerably suboptimal while the Jensen approximation covariance matrices turn out to give a very good agreement with the asymptotic optimum results in the range of SNR considered. Markers report corresponding Monte-Carlo simulations and confirm, again, the accuracy of the asymptotic approximation.

## 4.5 Conclusions

We presented an asymptotic analytic method to calculate the *ergodic capacity region* of a multiple-access MIMO channel with correlated Rician fading. The method applies when the number of antennas goes to infinity but yields very accurate approximations even with a small number of antennas.

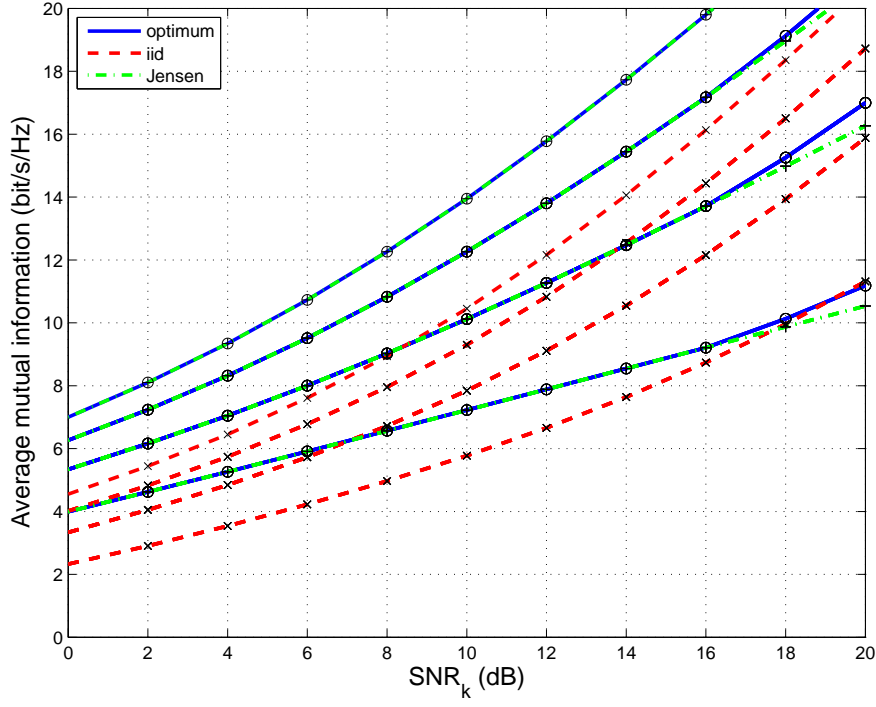


Figure 4.3: Average mutual information  $\mathbb{E}[\mathbf{I}(\mathbf{x}_{\mathcal{S}}; \mathbf{y} | \mathbf{x}_{\mathcal{S}^c})]$  versus  $\text{SNR}_k$  for scenario 2 ( $K = 4$  users) and  $|\mathcal{S}| = 1, \dots, 4$  (from bottom to top). The results refer to the optimum (4.19), Jensen (4.26), and iid power allocation.  $K_k^R = 10$  dB,  $\mathbf{T}_k$  and  $\mathbf{R}$  are exponential with base 0.7,  $r = 4$ ,  $t_k = 4$  for all  $k \in \mathcal{K}$ . Markers correspond to Monte-Carlo simulations.

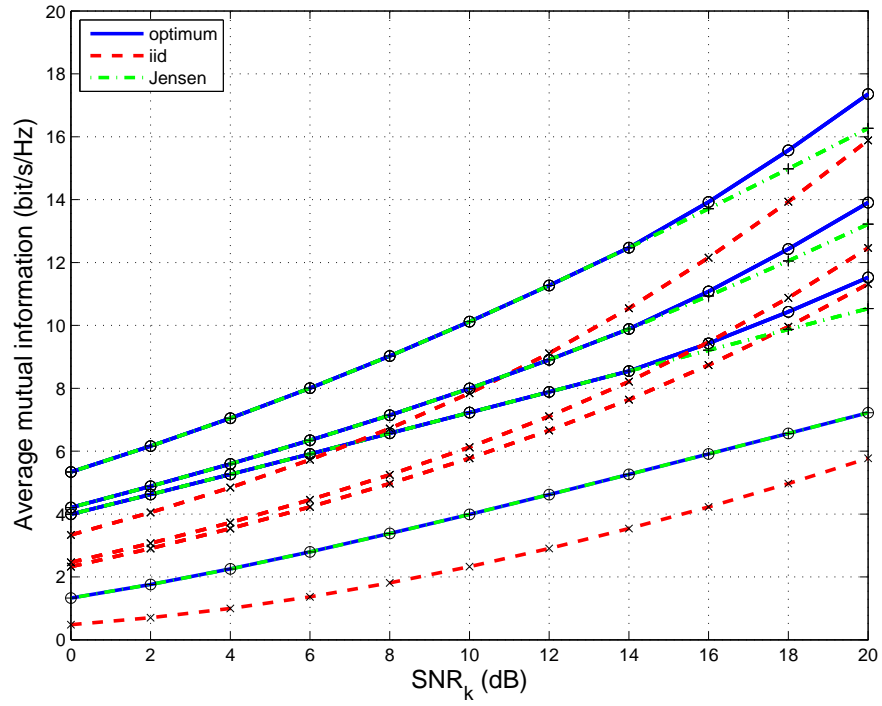


Figure 4.4: Same as Fig. 4.3 but for scenario 3 [asymmetric, weak user 1 with  $(\text{SNR}_1)_{\text{dB}} = (\text{SNR}_k)_{\text{dB}} - 10$ ]. The curves in this figure refer to  $\mathcal{S} = \{1\}, \{k\}, \{1, k\}, \{k, k'\}$  with  $k, k' \neq 1$  (from bottom to top).

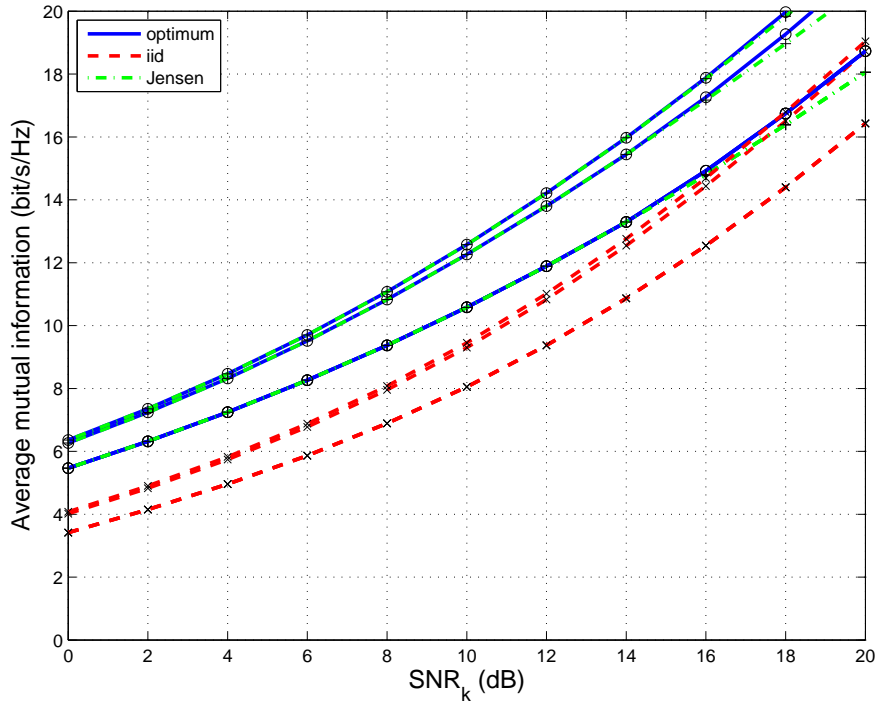


Figure 4.5: Same as Fig. 4.4 but the curves in this figure refer to  $\mathcal{S} = \{1, k, k'\}, \{k, k', k''\}, \mathcal{K}$  with  $k, k', k'' \neq 1$  (from bottom to top).

## Chapter 5

# Characterization of Capacity Regions by Means of Interference Functions

*Holger Boche, Martin Schubert*

### 5.1 Introduction

The analysis of capacity regions is complicated by interference between the communication links (“users”). The achievable capacity of one link can depend on the transmission strategies of other links. This typically results in a coupled system with many degrees of freedom. So well-established communication strategies for point-to-point links are not always applicable to multiuser systems.

A well-known example is the capacity region of the Gaussian MIMO multiple access channel (MAC), and its dual broadcast channel (BC) [68–70, 82]. A thorough understanding of the underlying performance trade-offs is often the basis for the development of efficient multiuser transmission strategies. For example, the characterization of the aforementioned MIMO broadcast region was accompanied by a search for optimum communication strategies.

However, these results only hold under certain conditions. For example, the transmit strategies for MIMO broadcast channels proposed in [68–70] rely on the BC/MAC duality. If the transmit strategy is constrained to be linear, then the existence of such a duality is still unknown. If we further constrain the system by forbidding time sharing or rate splitting, then the capacity region can even be non-convex [81]. For more complicated channels, e.g. relay channels, the characterization of the capacity region is an open problem.

This discussion shows that the characterization of wireless capacity regions

can be quite complicated, especially when additional system constraints are considered. This motivates an abstract approach, which focuses on some core properties. A typical property, for example, is that the performance of any link can be reduced without leaving the feasible region. This “monotonicity behavior” is known as *comprehensiveness* in a game-theoretical context [72].

### 5.1.1 Comprehensive Performance Sets

Before providing detailed definitions for comprehensiveness, we need some notational conventions.

- $\mathcal{K} = \{1, 2, \dots, K\}$  is the set of users (communication links).
- The set of non-negative reals is denoted by  $\mathbb{R}_+$ . The set of positive reals is denoted by  $\mathbb{R}_{++}$ .
- Matrices and vectors are denoted by bold capital letters and bold lowercase letters, respectively. Let  $\mathbf{y}$  be a vector, then  $y_l = [\mathbf{y}]_l$  is the  $l$ th component.
- A vector inequality  $\mathbf{x} > \mathbf{y}$  means  $x_k > y_k$ , for all  $k$ . The same definition holds for  $\mathbf{x} \geq \mathbf{y}$ .

**Definition 1** A set  $\mathcal{V} \subset \mathbb{R}^K$  is said to be upwards-comprehensive if for an arbitrary  $\mathbf{v} \in \mathcal{V}$ ,

$$\mathbf{v}' \geq \mathbf{v} \implies \mathbf{v}' \in \mathcal{V}. \quad (5.1)$$

The set  $\mathcal{V}$  is said to be downwards-comprehensive if for an arbitrary  $\mathbf{v} \in \mathcal{V}$ ,

$$\mathbf{v}' \leq \mathbf{v} \implies \mathbf{v}' \in \mathcal{V}. \quad (5.2)$$

A useful concept for analyzing interference-coupled comprehensive systems is the framework of interference functions [71]. A general overview on this framework was already shown in the previous deliverable D.3.1.1 and the related reference [78].

The main contribution of this deliverable (the corresponding paper was published in [77]) is to show that interference functions can be used for the analysis of capacity regions (and also other performance regions). In particular, every comprehensive capacity region can be expressed as a sub-level set of an interference function. This facilitates a general framework for analyzing performance trade-offs in multiuser networks.

### 5.1.2 Level Sets and Interference Functions

We start with an introduction of general interference functions.

**Definition 2** We say that  $\mathcal{I} : \mathbb{R}_{++}^K \mapsto \mathbb{R}_{++}$  is an interference function if it fulfills the axioms:

- A1** (positivity)  $\mathcal{I}(\mathbf{p}) > 0$
- A2** (scale invariance)  $\mathcal{I}(\alpha \mathbf{p}) = \alpha \mathcal{I}(\mathbf{p}) \quad \forall \alpha \in \mathbb{R}_{++}$
- A3** (monotonicity)  $\mathcal{I}(\mathbf{p}) \geq \mathcal{I}(\mathbf{p}') \quad \text{if } \mathbf{p} \geq \mathbf{p}'$

These properties are quite intuitive when we think of  $\mathbf{p} = [p_1, \dots, p_K]^T$  as a vector of transmission powers, and  $\mathcal{I}(\mathbf{p})$  as the resulting interference. However, other interpretations are possible. Some examples are provided, e.g. in [71, 73, 78].

Let's analyze the interference function  $\mathcal{I}(\mathbf{p})$ . It was shown in [77] that every interference function satisfying A1–A3 has a min-max representation. In order to explain this result, we need the definition of relative closedness.

**Definition 3** A set  $\mathcal{V} \subset \mathbb{R}_{++}^K$  is said to be relatively closed in  $\mathbb{R}_{++}^K$  if there exists a closed set  $\mathcal{A} \subset \mathbb{R}^K$  such that  $\mathcal{V} = \mathcal{A} \cap \mathbb{R}_{++}^K$ .

Consider level sets

$$\underline{L}(\mathcal{I}) = \{\hat{\mathbf{p}} > 0 : \mathcal{I}(\hat{\mathbf{p}}) \leq 1\} \quad (5.3)$$

$$\overline{L}(\mathcal{I}) = \{\hat{\mathbf{p}} > 0 : \mathcal{I}(\hat{\mathbf{p}}) \geq 1\} . \quad (5.4)$$

Since  $\mathcal{I}(\mathbf{p})$  is continuous [71], the sets  $\underline{L}(\mathcal{I})$  and  $\overline{L}(\mathcal{I})$  are relatively closed in  $\mathbb{R}_{++}^K$ . From property A3, it follows that the sets are comprehensive.

We have the following result [77]

**Property 5.1.1** Let  $\mathcal{I}$  be an arbitrary interference function fulfilling A1–A3, then

$$\mathcal{I}(\mathbf{p}) = \min_{\hat{\mathbf{p}} \in \underline{L}(\mathcal{I})} \max_{k \in \mathcal{K}} \frac{p_k}{\hat{p}_k} \quad (5.5)$$

$$= \max_{\hat{\mathbf{p}} \in \overline{L}(\mathcal{I})} \min_{k \in \mathcal{K}} \frac{p_k}{\hat{p}_k} . \quad (5.6)$$

That is, every interference function can be interpreted as an optimization over elementary interference function, where the variable  $\hat{\mathbf{p}}$  is chosen from a closed comprehensive level set.

Conversely, we can start with a closed comprehensive positive set (e.g. a capacity region), and show that this set can be expressed as a level set of an interference function. This leads to the following observation [77]



**Property 5.1.2** *Every closed comprehensive positive set  $\mathcal{U}$  can be expressed as a sub-level set*

$$\mathcal{U} = \{\mathbf{u} \in \mathbb{R}_{++}^K : C(\mathbf{u}) \leq 1\} . \quad (5.7)$$

*The set  $\mathcal{U}$  is closed comprehensive positive if and only if  $C(\mathbf{u})$  is an interference function.*

Hence, there is a direct correspondence between interference functions and comprehensive performance regions.

An extension to other types of performance regions can be found in [80], where the supportable region of a multiuser system with log-convex interference functions is analyzed.

**Definition 4** *We say that  $\mathcal{I} : \mathbb{R}_+^K \mapsto \mathbb{R}_+$  is a log-convex interference function if A1–A3 are fulfilled and in addition  $\mathcal{I}(\exp\{\mathbf{s}\})$  is log-convex on  $\mathbb{R}^K$ .*

It was shown in [71] that every convex interference function is a log-convex interference function, but the converse is not true. Note, that this statement only holds true with the change of variable  $\mathbf{p} = \exp\{\mathbf{s}\}$ . Thus, the family of log-convex interference functions is more general than the family of convex interference functions. The log-convexity property is useful for the analysis of certain types of performance regions (e.g. SIR feasible sets). This is further investigated, e.g. in [76, 79].

In the context of capacity regions, we are mostly interested in convex regions. Convexity simplifies the tasks of resource allocation. A common design goal is to find a Pareto optimal point on the boundary of the region.

The assumption of convexity is typically justified by time sharing or rate splitting arguments. An analysis of convex performance regions was published in [75]. In this publication, the following result was shown:

**Property 5.1.3** *Every closed comprehensive positive convex utility set can be expressed as a sub-level set*

$$\mathcal{U} = \{\mathbf{u} \in \mathbb{R}_{++}^K : C(\mathbf{u}) \leq 1\} . \quad (5.8)$$

*The set  $\mathcal{U}$  is closed comprehensive positive convex if and only if  $C(\mathbf{u})$  is a convex interference function.*

This shows how basic properties of interference functions are transferred to properties of utility sets, and vice versa.

### 5.1.3 Discussion

The results of this work package, published in [75, 77, 80], show that there is a direct correspondence between comprehensive performance regions and interference functions fulfilling the properties A1–A3. Further properties, like convexity or log-convexity, can be added.

This theoretical framework facilitates a general and unifying approach for the analysis of different kinds of capacity regions. By focusing on core properties, we are able to develop a rigorous framework which allows for an analytical treatment. The results provide intuition and a roadmap for the development of algorithms in WP1.

An application example is the iterative algorithm for max-min balancing published in [74]. Other resource allocation strategies are currently being investigated in WP1. For example, interference functions were successfully applied to the analysis of resource allocation strategies based on cooperative game theory in [76, 79].

## Chapter 6

# Asymptotic Ergodic Capacity of Wideband MIMO Channels with Separately-Correlated Rician Fading

*Giorgio Taricco, Erwin Riegler*

### 6.1 Introduction

Many research studies in multiple-input multiple-output (MIMO) communications address the problem of deriving the channel capacity. Probably one of the most successful results in this area is due to Telatar, who found closed-form expressions for the capacity of the independent Rayleigh fading MIMO channel [46]. Subsequent studies showed that this channel model is not always appropriate because it does not take into account spatial correlation and the presence of a line-of-sight (LOS) component. Many experimental and theoretical works addressed the issue of channel modeling [16]. A good trade-off between model complexity and accuracy is achieved by the so-called separately-correlated channel [18].

Another active research area in MIMO communications targeted the extension to the wideband channel, in most cases based on an OFDM approach [13, 19, 22]. Recently, Moustakas and Simon studied the asymptotic mutual information of the wideband Rayleigh fading correlated MIMO channel in [21], extending one of their own earlier studies, relevant to the narrowband channel case [20]. The analytic approach used in these works is based on the so-called *replica method*, widely adopted in theoretical physics because of its effectiveness in addressing complex system scenarios where other methods, such as the one based on Stieltjes

transform proposed in [31], provide equivalent results with a greater mathematical effort.

In this work we address the case of a wideband Rician fading MIMO channel with separate correlation (see the System Model section for the precise assumptions). We derive an expression of the average mutual information in the case of asymptotically large number of antennas. This result represents a nontrivial extension of earlier results [25] addressing the narrowband channel case. By using this result, we derive the ergodic capacity of the channel and the corresponding (optimum) input signal covariance matrix. Then, we study the numerical accuracy of the asymptotic approximation proposed and analyze the effectiveness of covariance optimization with respect to the spatial correlation and the Rice factor.

## 6.2 System model

We consider a wideband frequency-selective separately-correlated Rician fading MIMO channel with  $t$  transmit and  $r$  receive antennas. We assume that the channel bandwidth is  $B$  and the delay spread is  $\tau$ . Then, the maximum number of resolvable paths of the channel is  $L = \lceil B\tau \rceil$ , and the channel equation can be written as follows:

$$\mathbf{y}[n] = \sum_{\ell \in \mathcal{L}} \mathbf{H}[\ell] \mathbf{x}[n - \ell] + \mathbf{z}[n], \quad n \in \mathcal{N}. \quad (6.1)$$

Here, we defined: the frame length  $N$ ; the sets  $\mathcal{L} \triangleq \{0, \dots, L - 1\}$  and  $\mathcal{N} \triangleq \{0, \dots, N - 1\}$ ; the transmitted signal vector  $\mathbf{x}[n] \in \mathbb{C}^{t \times 1}$  at time  $n$ ; the impulse response channel matrix  $\mathbf{H}[\ell] \in \mathbb{C}^{r \times t}$ , sampled at delay time  $\ell$ ; the additive noise vector  $\mathbf{z}[n] \in \mathbb{C}^{r \times 1}$  at time  $n$ ; and the received signal vector  $\mathbf{y}[n] \in \mathbb{C}^{r \times 1}$  at time  $n$ .

We assume that the additive noise is independent over time and uncorrelated:  $\mathbf{z}[n] \sim \mathcal{N}_c(\mathbf{0}, \mathbf{I}_r)$ <sup>1</sup>.

We assume uncorrelated scattering, so that all matrices  $\mathbf{H}[\ell]$  are independent. We consider a separately correlated Rician fading model with common spatial correlation among the users at the receiver. Thus, we can write

$$\mathbf{H}[\ell] = \mathbb{E}[\mathbf{H}[\ell]] + \mathbf{R}[\ell]^{1/2} \mathbf{H}_w[\ell] \mathbf{T}[\ell]^{1/2}. \quad (6.2)$$

Here, the matrices  $\mathbf{H}_w[\ell]$  have all iid entries distributed as  $\mathcal{N}_c(0, 1)$  (in compliance with the uncorrelated scattering assumption),  $\mathbf{R}[\ell]$  is the receive spatial correlation

---

<sup>1</sup> In the presence of spatial correlation,  $\mathbf{z}[n] \sim \mathcal{N}_c(\mathbf{0}, \mathbf{\Sigma}_Z)$ , we can reduce to the uncorrelated case pre-multiplying  $\mathbf{y}[n]$  by  $\mathbf{\Sigma}_Z^{-1/2}$ . The resulting noise vector will be uncorrelated with identity covariance matrix and the channel matrices will be all pre-multiplied by the same constant matrix  $\mathbf{\Sigma}_Z^{-1/2}$  and the resulting problem will be still tractable under our approach.

matrix, and  $\mathbf{T}[\ell]$  is the transmit spatial correlation matrix. Since a nonzero average channel matrix derives from a line-of-sight component, which typically affects only the signal path with the lowest delay, it will be assumed that

$$\mathbb{E}[\mathbf{H}[\ell]] = \begin{cases} \bar{\mathbf{H}} & \ell = 0 \\ \mathbf{0} & \ell > 0 \end{cases}. \quad (6.3)$$

The Rician factor (defined as the ratio between the LOS and the scattered received power when the input signals are iid with power  $P/t$ ) and the signal-to-noise ratio (SNR) are given by:

$$K \triangleq \frac{\|\bar{\mathbf{H}}\|^2}{\sum_{\ell \in \mathcal{L}} \text{Tr}(\mathbf{R}[\ell]) \text{Tr}(\mathbf{T}[\ell])} \quad (6.4)$$

and

$$\rho \triangleq \frac{(K+1)P \sum_{\ell \in \mathcal{L}} \text{Tr}(\mathbf{R}[\ell]) \text{Tr}(\mathbf{T}[\ell])}{rt}, \quad (6.5)$$

respectively.

### 6.2.1 OFDM signaling

We consider a communication system based on orthogonal frequency-division multiplexing (OFDM). It is well known that OFDM allows to convert the frequency-selective fading channel into a set of parallel frequency-flat channels. The key assumption of OFDM is the insertion of a cyclic prefix (CP) in all frames, whose length is equal to the delay spread. The CP assumption implies the following set of conditions:

$$\mathbf{x}[n] = \mathbf{x}[n+N] \quad n = -L+1, \dots, -1. \quad (6.6)$$

It must be noted that the presence of a CP reduces the available throughput by a factor  $\kappa \triangleq \frac{N}{N+L-1}$ .

Next, we calculate the discrete Fourier transform (DFT) of the channel equation. Let  $\omega \triangleq \exp(-j2\pi/N)$  so that, in consideration of the CP condition (6.6), we can transform the channel equation (6.1) into

$$\begin{aligned} \hat{\mathbf{y}}[m] &\triangleq \sum_{n \in \mathcal{N}} \mathbf{y}[n] \omega^{mn} \\ &= \sum_{n \in \mathcal{N}} \sum_{\ell \in \mathcal{L}} \mathbf{H}[\ell] \mathbf{x}[n-\ell] \omega^{mn} + \sum_{n \in \mathcal{N}} \mathbf{z}[n] \omega^{mn} \\ &= \sum_{\ell \in \mathcal{L}} \mathbf{H}[\ell] \omega^{m\ell} \sum_{n \in \mathcal{N}} \mathbf{x}[n] \omega^{mn} + \sum_{n \in \mathcal{N}} \mathbf{z}[n] \omega^{mn} \\ &= \hat{\mathbf{H}}[m] \hat{\mathbf{x}}[m] + \hat{\mathbf{z}}[m], \quad (m \in \mathcal{N}). \end{aligned} \quad (6.7)$$

Here, we defined the following vector/matrix DFTs:  $\hat{\mathbf{H}}[m] \triangleq \sum_{\ell \in \mathcal{L}} \mathbf{H}[\ell] \omega^{m\ell}$ ,  $\hat{\mathbf{x}}[m] \triangleq \sum_{n \in \mathcal{N}} \mathbf{x}[n] \omega^{mn}$ , and  $\hat{\mathbf{z}}[m] \triangleq \sum_{n \in \mathcal{N}} \mathbf{z}[n] \omega^{mn}$ .

### 6.2.2 Frequency-domain equivalent channel

First, we notice that the frequency-domain noise vectors are iid since:

$$\begin{aligned}\mathbb{E}[\hat{\mathbf{z}}[m]\hat{\mathbf{z}}[m']^H] &= \sum_{n \in \mathcal{N}} \mathbb{E}[\mathbf{z}[n]\mathbf{z}[n]^H] \omega^{(m-m')n} \\ &= \begin{cases} \mathbf{0} & m \neq m' \\ N\mathbf{I}_r & m = m' \end{cases}.\end{aligned}\quad (6.8)$$

However, the frequency-domain channel matrices  $\hat{\mathbf{H}}[m]$  may be statistically *dependent*, even though the  $\mathbf{H}[\ell]$  matrices are independent. This complicates the analysis of the mutual information because the resulting channel would be frequency-varying. However, under assumption (6.3), the frequency-domain channel matrices are marginally identically distributed. In fact, it is easy to see that  $\mathbb{E}[\hat{\mathbf{H}}[m]] = \bar{\mathbf{H}}$  for all  $m \in \mathcal{N}$ . Moreover, setting  $\mathbf{h}[\ell] \triangleq (\mathbf{H}[\ell])$ ,  $\hat{\mathbf{h}}[m] \triangleq (\hat{\mathbf{H}}[m])$ ,  $\tilde{\mathbf{h}}[\ell] \triangleq \mathbf{h}[\ell] - \mathbb{E}[\mathbf{h}[\ell]]$ , for  $\ell \in \mathcal{L}$  and  $m \in \mathcal{N}$ , and  $\bar{\mathbf{h}} \triangleq (\bar{\mathbf{H}})$ , we have:

$$\begin{aligned}\mathbb{E}[\hat{\mathbf{h}}[m]\hat{\mathbf{h}}[m']^H] &= \mathbb{E}[\hat{\mathbf{h}}[m]]\mathbb{E}[\hat{\mathbf{h}}[m']^H] + \sum_{\ell \in \mathcal{L}} \mathbb{E}[\tilde{\mathbf{h}}[\ell]\tilde{\mathbf{h}}[\ell]^H] \omega^{(m-m')\ell} \\ &= \bar{\mathbf{h}}\bar{\mathbf{h}}^H + \sum_{\ell \in \mathcal{L}} \mathbf{T}^T[\ell] \xi \mathbf{R}[\ell] \omega^{(m-m')\ell},\end{aligned}\quad (6.9)$$

which is independent of  $m$  when  $m = m'$ , as required.

Then, we follow the arguments reported in [13]. Assuming every OFDM symbol corresponds to independent realizations of the channel impulse response matrices, the mutual information will be given (in nat) by

$$\begin{aligned}I(\hat{\mathbf{Q}}[0], \dots, \hat{\mathbf{Q}}[N-1]) &= \frac{1}{N+L-1} \\ &\cdot \sum_{m \in \mathcal{N}} \ln \det(\mathbf{I}_r + \hat{\mathbf{H}}[m]\hat{\mathbf{Q}}[m]\hat{\mathbf{H}}[m]^H),\end{aligned}\quad (6.10)$$

where  $\hat{\mathbf{Q}}[m] = \mathbb{E}[\hat{\mathbf{x}}[m]\hat{\mathbf{x}}[m]^H]$ .

Typically, the ergodic channel capacity is calculated under an average time-domain power constraint:

$$\sum_{n \in \mathcal{N}} \mathbb{E}[\text{Tr}(\mathbf{x}[n]\mathbf{x}[n]^H)] \leq P. \quad (6.11)$$

However, applying the inverse DFT (IDFT)

$$\mathbf{x}[n] = \frac{1}{N} \sum_{m \in \mathcal{N}} \hat{\mathbf{x}}[m] \omega^{-mn},$$

we can obtain the following equivalent average frequency-domain power constraint:

$$\begin{aligned}
P &\geq \sum_{n \in \mathcal{N}} \mathbb{E}[\text{Tr}(\mathbf{x}[n]\mathbf{x}[n]^H)] \\
&= \sum_{n \in \mathcal{N}} \sum_{m \in \mathcal{N}} \sum_{m' \in \mathcal{N}} \frac{\mathbb{E}[\text{Tr}(\hat{\mathbf{x}}[m]\hat{\mathbf{x}}[m']^H)]}{N^2} \omega^{-(m-m')n} \\
&= \frac{1}{N} \sum_{m \in \mathcal{N}} \mathbb{E}[\text{Tr}(\hat{\mathbf{x}}[m]\hat{\mathbf{x}}[m]^H)].
\end{aligned} \tag{6.12}$$

Therefore, the MIMO-OFDM ergodic capacity is defined as the maximum of  $\mathbb{E}[\mathbf{I}(\hat{\mathbf{Q}})]$  under the constraint

$$\sum_{m \in \mathcal{N}} \text{Tr}(\hat{\mathbf{Q}}[m]) \leq NP. \tag{6.13}$$

Furthermore, in accordance with [21, Sec. III-A], we can show that the ergodic capacity achieving matrices are independent of  $m$ . In fact, let us define:

$$\phi(\hat{\mathbf{Q}}) \triangleq \mathbb{E}[\ln \det(\mathbf{I}_r + \hat{\mathbf{H}}[m]\hat{\mathbf{Q}}\hat{\mathbf{H}}[m]^H)].$$

This is a concave function of  $\hat{\mathbf{Q}}$  from [15, Th. 17.9.1]. Hence, we can apply Jensen's inequality and obtain:

$$\begin{aligned}
\mathbb{E}[\mathbf{I}(\hat{\mathbf{Q}}[0], \dots, \hat{\mathbf{Q}}[N-1])] &= \frac{1}{N+L-1} \sum_{m \in \mathcal{N}} \phi(\hat{\mathbf{Q}}[m]) \\
&\leq \kappa \phi\left(\frac{1}{N} \sum_{m \in \mathcal{N}} \hat{\mathbf{Q}}[m]\right).
\end{aligned}$$

Since the last matrix argument is independent of  $m$ , we conclude that the ergodic capacity is achieved by a covariance matrix that is constant for all values of  $m$ . Therefore, we restrict to consider the case of equal covariance matrices  $\hat{\mathbf{Q}}[m] = \hat{\mathbf{Q}}$  with the power constraint

$$\text{Tr}(\hat{\mathbf{Q}}) \leq P.$$

## 6.3 Asymptotic Analysis

In this section we provide an analytic method to calculate the asymptotic ergodic mutual information of the OFDM-MIMO fading channel described above. The method proposed is a generalization of the one presented in [8, 9] relevant to the

narrowband channel. Recalling that each frequency-domain channel matrix  $\hat{\mathbf{H}}[m]$  has the same probability distribution, our goal is to calculate the mean

$$\mathbb{E}[\ln \det(\mathbf{I}_r + \hat{\mathbf{H}}\hat{\mathbf{Q}}\hat{\mathbf{H}}^H)],$$

where

$$(\hat{\mathbf{H}}) \sim \mathcal{N}_c\left(\bar{\mathbf{h}}, \sum_{\ell \in \mathcal{L}} \mathbf{T}^T[\ell] \xi \mathbf{R}[\ell]\right),$$

according to (6.9). The ergodic mutual information of our OFDM-MIMO fading channel model is obtained as

$$\mathbb{E}[I(\hat{\mathbf{Q}})] = \kappa \mathbb{E}[\ln \det(\mathbf{I}_r + \hat{\mathbf{H}}\hat{\mathbf{Q}}\hat{\mathbf{H}}^H)].$$

### 6.3.1 Asymptotic CGF

Repeating the main steps of the asymptotic analysis developed in [8, 9], we write the cumulant generating function (CGF) of the mutual information as follows:

$$\begin{aligned} G(\nu) &\triangleq \mathbb{E}\left[\det(\mathbf{I}_r + \hat{\mathbf{H}}\hat{\mathbf{Q}}\hat{\mathbf{H}}^H)^{-\nu}\right] \\ &= \mathbb{E}\left[\int_{\mathbb{C}^{r \times \nu}} d\mathbf{U} \int_{\mathbb{C}^{t \times \nu}} \text{etr}\left\{-\pi(\mathbf{U}^H \mathbf{U} + \mathbf{V}^H \mathbf{V})\right\} \right. \\ &\quad \left. \text{etr}\left\{-\pi(\mathbf{U}^H \hat{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \mathbf{V} - \mathbf{V}^H \hat{\mathbf{Q}}^{1/2} \hat{\mathbf{H}}^H \mathbf{U})\right\} d\mathbf{V}\right]. \end{aligned}$$

Using the following result (see, e.g., [8, 9]):

$$\mathbb{E}[\text{etr}(\mathbf{M}\mathbf{W} - \mathbf{W}^H \mathbf{M}^H)] = \exp(-\|\mathbf{M}\|^2),$$

which holds for any random matrix  $\mathbf{W}$  with iid  $\mathcal{N}_c(0, 1)$  entries, we obtain:

$$\begin{aligned} G(\nu) &= \int_{\mathbb{C}^{r \times \nu}} d\mathbf{U} \int_{\mathbb{C}^{t \times \nu}} d\mathbf{V} \text{etr}[-\pi(\mathbf{U}^H \mathbf{U} + \mathbf{V}^H \mathbf{V})] \\ &\quad \text{etr}\left\{-\pi(\mathbf{U}^H \bar{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \mathbf{V} - \mathbf{V}^H \hat{\mathbf{Q}}^{1/2} \bar{\mathbf{H}}^H \mathbf{U})\right. \\ &\quad \left.-\pi^2 \sum_{\ell \in \mathcal{L}} \mathbf{U}^H \mathbf{R}[\ell] \mathbf{U} \mathbf{V}^H \tilde{\mathbf{T}}[\ell] \mathbf{V}\right\} \\ &= \int_{\mathbb{C}^{r \times \nu}} d\mathbf{U} \int_{\mathbb{C}^{t \times \nu}} d\mathbf{V} \text{etr}[-\pi(\mathbf{U}^H \mathbf{U} + \mathbf{V}^H \mathbf{V})] \\ &\quad \text{etr}[-\pi(\mathbf{U}^H \bar{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \mathbf{V} - \mathbf{V}^H \hat{\mathbf{Q}}^{1/2} \bar{\mathbf{H}}^H \mathbf{U})] \\ &\quad \prod_{\ell \in \mathcal{L}} \int d\mu(\mathbf{W}_\ell, \mathbf{Z}_\ell) \text{etr}\left\{\mathbf{W}_\ell \mathbf{Z}_\ell - \pi \mathbf{W}_\ell \mathbf{U}^H \mathbf{R}[\ell] \mathbf{U}\right\} \\ &\quad \text{etr}\left\{-\pi \mathbf{Z}_\ell \mathbf{V}^H \tilde{\mathbf{T}}[\ell] \mathbf{V}\right\}. \end{aligned}$$



Here, integration over  $\mathbf{W}_\ell$  is carried out on  $\mathbb{R}^{\nu \times \nu} + \mathbf{W}_{\ell,0}$  for some  $\mathbf{W}_{\ell,0} \in \mathbb{C}^{\nu \times \nu}$ , and the integration over  $\mathbf{Z}_\ell$  is carried out on  $(j\mathbb{R})^{\nu \times \nu} + \mathbf{Z}_{\ell,0}$  for some  $\mathbf{Z}_{\ell,0} \in \mathbb{C}^{\nu \times \nu}$ . Moreover,  $d\mu(\mathbf{W}_\ell, \mathbf{Z}_\ell) = d\mathbf{W}_\ell d(\mathbf{Z}_\ell / (j2\pi))$ . Applying the identity  $\text{Tr}(\mathbf{M}\mathbf{X}^H\mathbf{N}\mathbf{Y}) = (\mathbf{X})^H(\mathbf{M}^T\xi\mathbf{N})(\mathbf{Y})$  [17] in the last integral, we obtain the following result:

$$G(\nu) = \int \left\{ \prod_{\ell \in \mathcal{L}} \text{etr}(\mathbf{W}_\ell \mathbf{Z}_\ell) d\mu(\mathbf{W}_\ell, \mathbf{Z}_\ell) \right\} \times \det \begin{pmatrix} \mathbf{I}_{\nu r} + \sum_{\ell \in \mathcal{L}} \mathbf{W}_\ell \xi \mathbf{R}[\ell] & \mathbf{I}_\nu \xi \bar{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \\ -\mathbf{I}_\nu \xi \hat{\mathbf{Q}}^{1/2} \bar{\mathbf{H}}^H & \mathbf{I}_{\nu t} + \sum_{\ell \in \mathcal{L}} \mathbf{Z}_\ell \xi \tilde{\mathbf{T}}[\ell] \end{pmatrix}^{-1} \quad (6.14)$$

Here and in the following we define  $\tilde{\mathbf{T}}[\ell] \triangleq \hat{\mathbf{Q}}^{1/2} \mathbf{T}[\ell] \hat{\mathbf{Q}}^{1/2}$ .

### 6.3.2 Ergodic mutual information

We assume that the saddlepoint of the integrand in (6.14) corresponds to the replica symmetric point  $\mathbf{W}_\ell = w_\ell \mathbf{I}_\nu$ ,  $\mathbf{Z}_\ell = z_\ell \mathbf{I}_\nu$ , for  $\ell \in \mathcal{L}$ . Under this assumption, we can get the values of  $w_\ell, z_\ell$  by solving the following set of fixed-point equations:

$$\begin{cases} w_\ell = \text{Tr} \left\{ \left[ \mathbf{I}_t + \sum_{\ell \in \mathcal{L}} z_\ell \tilde{\mathbf{T}}[\ell] + \hat{\mathbf{Q}}^{1/2} \bar{\mathbf{H}}^H \right. \right. \\ \quad \times \left. \left( \mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] \right)^{-1} \bar{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \right]^{-1} \tilde{\mathbf{T}}[\ell] \right\} \\ z_\ell = \text{Tr} \left\{ \left[ \mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] + \bar{\mathbf{H}} \hat{\mathbf{Q}}^{1/2} \right. \right. \\ \quad \times \left. \left( \mathbf{I}_t + \sum_{\ell \in \mathcal{L}} z_\ell \tilde{\mathbf{T}}[\ell] \right)^{-1} \hat{\mathbf{Q}}^{1/2} \bar{\mathbf{H}}^H \right]^{-1} \mathbf{R}[\ell] \right\} \end{cases} \quad (6.15)$$

for all  $\ell \in \mathcal{L}$ . The resulting asymptotic approximation of the ergodic capacity is given by:

$$\begin{aligned} \mu_I(\hat{\mathbf{Q}}) &= \ln \det \left\{ \mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] \right\} \\ &+ \ln \det \left\{ \mathbf{I}_t + \left[ \sum_{\ell \in \mathcal{L}} z_\ell \mathbf{T}[\ell] + \bar{\mathbf{H}}^H \left( \mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] \right)^{-1} \bar{\mathbf{H}} \right] \hat{\mathbf{Q}} \right\} \\ &\quad - \sum_{\ell \in \mathcal{L}} w_\ell z_\ell \end{aligned} \quad (6.16)$$

### 6.3.3 Ergodic capacity

To maximize the ergodic mutual information (6.16) with respect to the input covariance matrix  $\hat{\mathbf{Q}}$  we have to solve the following optimization problem:

$$\begin{aligned} & \text{maximize} && \mu_I(\hat{\mathbf{Q}}) \\ & \text{subject to} && \hat{\mathbf{Q}} \geq \mathbf{0}, \text{Tr}(\hat{\mathbf{Q}}) \leq P. \end{aligned}$$

Paralleling the approach of [26], we can see that the objective function is concave with respect to  $\hat{\mathbf{Q}}$  and the constraints satisfy Slaters qualification condition [14]. Thus, the Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient for optimality [14]. Then, to derive the KKT conditions, we form the Lagrangian

$$\mathcal{L}(\hat{\mathbf{Q}}, \lambda, \Psi) = \mu_I(\hat{\mathbf{Q}}) - \lambda[\text{Tr}(\hat{\mathbf{Q}}) - P] + \text{Tr}(\Psi \hat{\mathbf{Q}}) \quad (6.17)$$

for  $\lambda \geq 0, \Psi \geq \mathbf{0}$ . After nulling the variation of  $\mathcal{L}(\hat{\mathbf{Q}}, \lambda, \Psi)$  with respect to  $\hat{\mathbf{Q}}$ , we obtain the KKT conditions:

$$\begin{cases} (\mathbf{I}_t + \Xi \hat{\mathbf{Q}})^{-1} \Xi &= \lambda \mathbf{I}_t - \Psi \\ \text{Tr}(\hat{\mathbf{Q}}) &= P \\ \text{Tr}(\Psi \hat{\mathbf{Q}}) &= 0 \end{cases} \quad (6.18)$$

where we defined the matrix

$$\Xi \triangleq \sum_{\ell \in \mathcal{L}} z_\ell \mathbf{T}[\ell] + \bar{\mathbf{H}}^H \left( \mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] \right)^{-1} \bar{\mathbf{H}}.$$

The solution of eqs. (6.18) depends on the rank of the matrix  $\Xi$ . Let us assume that  $\Xi = \mathbf{U}_x^H \Lambda_x \mathbf{U}_x$  is the orthogonal decomposition of  $\Xi$  with  $\Lambda_x = \text{diag}(\tilde{\Lambda}_x, \mathbf{0})$  and unitary  $\mathbf{U}_x^H$ , where  $\tilde{\Lambda}_x$  is a positive definite diagonal matrix. Then, if  $P_x = \text{rank}(\Xi) \leq t$ , we can see that

$$\ln \det(\mathbf{I}_t + \Xi \hat{\mathbf{Q}}) = \ln \det(\mathbf{I}_{\rho_x} + \tilde{\Lambda}_x \hat{\mathbf{Q}}_1),$$

where  $\hat{\mathbf{Q}}_1$  is the  $\rho_x \times \rho_x$  upper left submatrix of  $\mathbf{U}_x \hat{\mathbf{Q}} \mathbf{U}_x^H$ . Finally, the ergodic capacity achieving covariance matrix is given by

$$\hat{\mathbf{Q}} = \mathbf{U}_x \text{diag}\left((\lambda^{-1} \mathbf{I}_{\rho_x} - \tilde{\Lambda}_x^{-1})_+, \mathbf{0}\right) \mathbf{U}_x^H \quad (6.19)$$

where  $(\cdot)_+$  denotes the elementwise positive part of the elements of the matrix argument, i.e.,  $(x)_+ = x$  for  $x \geq 0$  and 0 otherwise. The Lagrange multiplier  $\lambda$  is determined by solving the constraint equation  $\text{Tr}(\hat{\mathbf{Q}}) = P$ . It can be noticed that eq. (6.19) can be written as

$$\hat{\mathbf{Q}} = \mathbf{U}_x (\lambda^{-1} \mathbf{I}_t - \Lambda_x^{-1})_+ \mathbf{U}_x^H$$

by assuming that matrix inversion maps the null entries on the main diagonal of  $\Lambda_x$  to positive infinity.

**Remark 6.3.1** It is worth noting that the ergodic capacity achieving covariance matrix found has the typical *waterfilling* form arising in many information theoretic optimization problems (see [15, 26]). The ergodic channel capacity is found by solving the joint set of equations given by (6.15) and (6.19). This can be achieved by an iterative algorithm that starts from an initial value of the variables  $(w_\ell, z_\ell, \hat{\mathbf{Q}})$  and proceeds by applying (6.15) and (6.19) alternately until it converges.

**Remark 6.3.2** In the special case when  $\mathbf{R}[\ell] = \mathbf{R}$ , constant for all  $\ell \in \mathcal{L}$ , we have a substantial simplification of the asymptotic mutual information equations. In fact, eqs. (6.15) and (6.16) become

$$\begin{cases} w = \text{Tr}\{[\mathbf{I}_t + z\tilde{\mathbf{T}} + \hat{\mathbf{Q}}^{1/2}\bar{\mathbf{H}}^H(\mathbf{I}_r + w\mathbf{R})^{-1}\bar{\mathbf{H}}\hat{\mathbf{Q}}^{1/2}]^{-1}\tilde{\mathbf{T}}\} \\ z = \text{Tr}\{[\mathbf{I}_r + w\mathbf{R} + \bar{\mathbf{H}}\hat{\mathbf{Q}}^{1/2}(\mathbf{I}_t + z\tilde{\mathbf{T}})^{-1}\hat{\mathbf{Q}}^{1/2}\bar{\mathbf{H}}^H]^{-1}\mathbf{R}\} \end{cases} \quad (6.20)$$

and

$$\begin{aligned} \mu_I(\hat{\mathbf{Q}}) &= \ln \det(\mathbf{I}_r + w\mathbf{R}) \\ &\quad + \ln \det\{\mathbf{I}_t + [z\mathbf{T} + \bar{\mathbf{H}}^H(\mathbf{I}_r + w\mathbf{R})^{-1}\bar{\mathbf{H}}]\hat{\mathbf{Q}}\} - wz, \end{aligned} \quad (6.21)$$

respectively. In those equations we set  $\mathbf{T} \triangleq \sum_{\ell \in \mathcal{L}} \mathbf{T}[\ell]$  and  $\tilde{\mathbf{T}} \triangleq \hat{\mathbf{Q}}^{1/2}\mathbf{T}\hat{\mathbf{Q}}^{1/2}$ .

## 6.4 Numerical results

We present some numerical results to illustrate the accuracy of the asymptotic analytic method proposed to determine the capacity of the wideband MIMO communication channel considered. Here we disregard the rate reduction due to CP insertion (equivalently, we assume  $N \rightarrow \infty$ ). Specifically, we study the effects of spatial correlation with a strong LOS component corresponding to a Rice factor  $K = 0$  dB. We consider the following parameters of the communication scenarios:

- N. of antennas:  $t = r = 4$ , n. of paths:  $L = 3$ .
- Mean channel matrix:  $(\bar{\mathbf{H}})_{ab} = h_0$  (constant).
- Correlation:  $(\mathbf{R}[\ell])_{ab} = \alpha_\ell^{|a-b|}$  and  $(\mathbf{T}[\ell])_{ab} = \rho_\ell \beta_\ell^{|a-b|}$ , with  $\rho_\ell = 0.9^\ell$ .

The effects of spatial correlation are assessed by comparing the ergodic capacity and the iid mutual information (corresponding to  $\hat{\mathbf{Q}} = q\mathbf{I}_t$ ) in three different scenarios: *i*) uncorrelated (Fig. 6.1); *ii*) partly correlated (i.e., correlated only over

the paths without a LOS component, Fig. 6.2); and *iii*) completely correlated (i.e., correlated over all paths, Fig. 6.3). In all figures, solid and dashed lines with markers correspond to the ergodic capacity and to the iid mutual information, respectively, obtained in both cases via the asymptotic analytic approach developed in this work. Square markers represent the corresponding Monte-Carlo simulation results. We can see that, in all cases considered, simulation results are in close agreement with the analytic ones. The figures also plot the optimum covariance matrix eigenvalue distribution (solid lines without markers, with sum normalized to 1). These curves confirm the well known fact that *beamforming* (corresponding to a single dominant eigenvalue, only one curve stuck to 1) is optimal below a certain SNR. These results show that the threshold increases as the correlation increases, too. The optimization of the covariance matrix is more effective as the level of spatial correlation increases. The results show that covariance optimization can compensate almost completely the iid mutual information degradation due to increased correlation. Similar results, not reported for space limitation, show that the effectiveness of covariance optimization increases as the Rice factor increases. On the contrary, with Rayleigh fading, the advantage is more limited.

## 6.5 Conclusions

We provided an analytic asymptotic (in the number of antennas) method to approximate the mutual information of a wideband Rician fading separately correlated MIMO channel. By using this method, we obtain the asymptotically optimum input signal covariance matrix, yielding the ergodic capacity of the channel considered. As noticed in [31], the asymptotic method is substantially more efficient than numerical optimization methods such as the one proposed in [50] (interior point with barrier). We show by numerical results that the asymptotic approximation is accurate even when the number of antennas is small. Finally, using our proposed method, we analyze the effectiveness of covariance optimization related to the level of spatial correlation and the Rice factor.

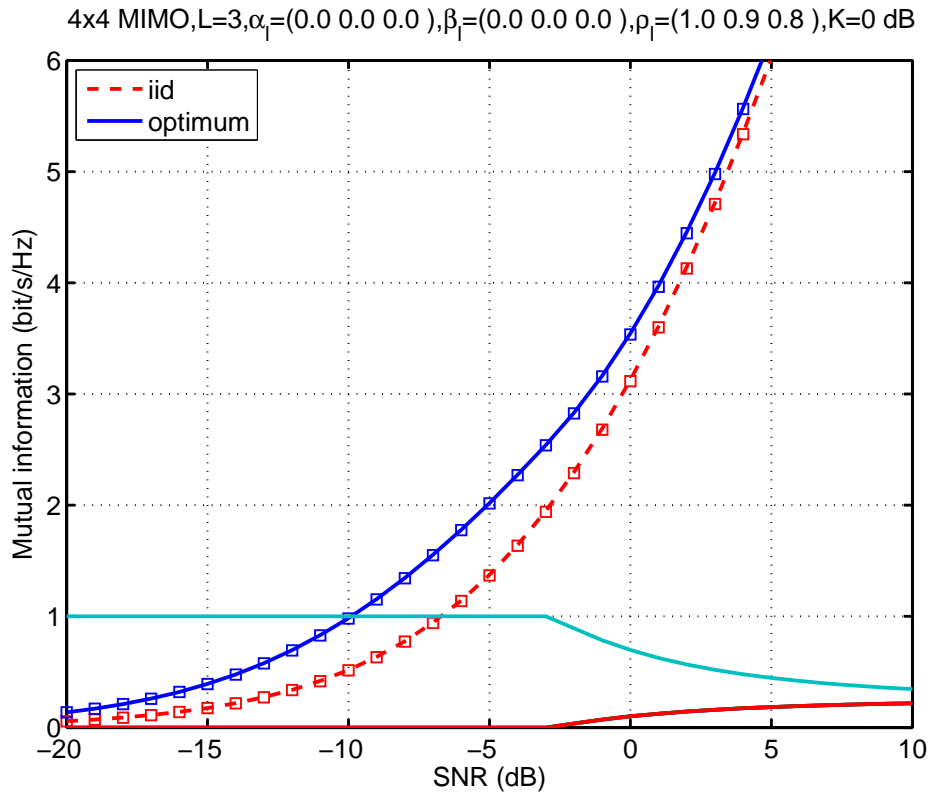


Figure 6.1: System capacity and iid mutual information versus the SNR for the uncorrelated scenario. Optimum covariance normalized eigenvalues are also reported.

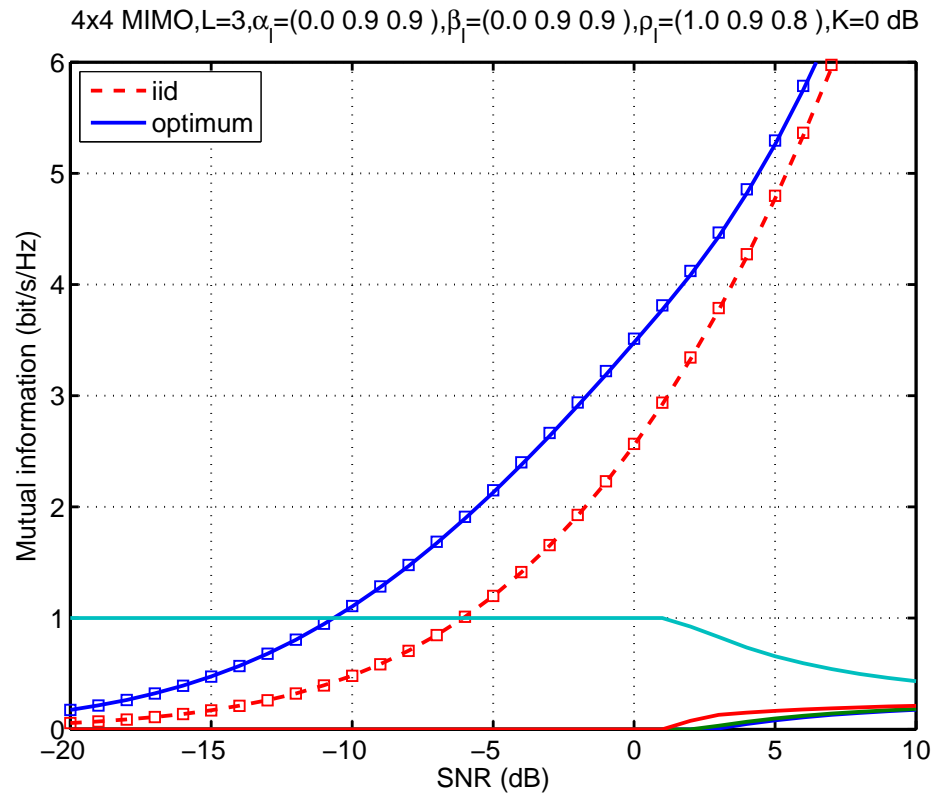


Figure 6.2: Same as Fig. 6.1 for the partly correlated scenario.

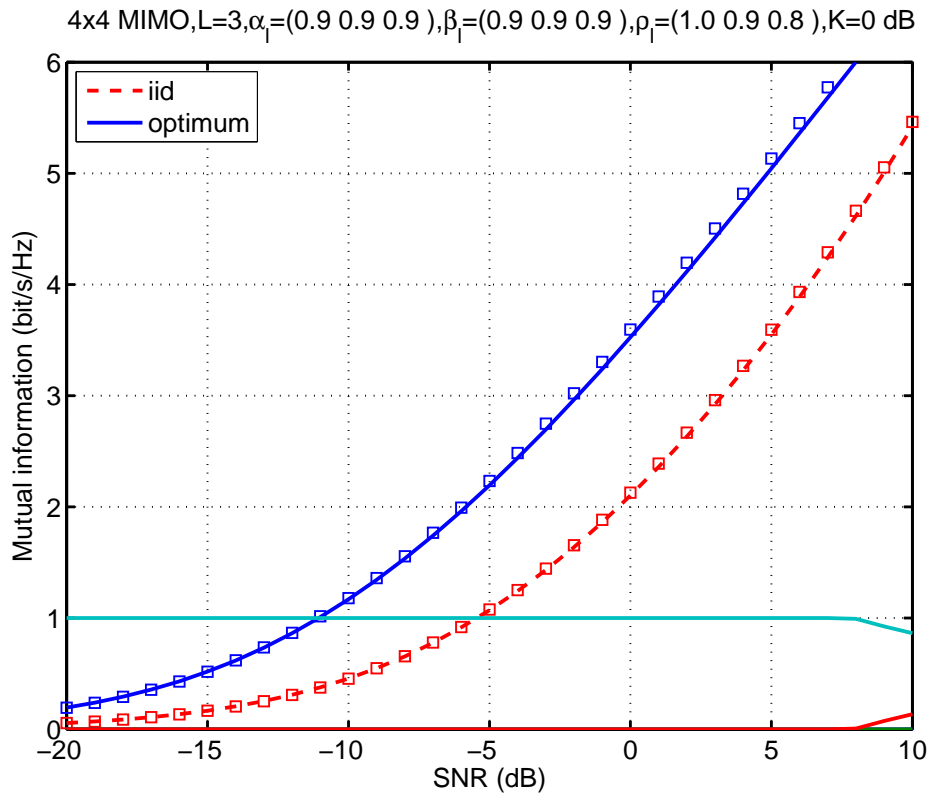


Figure 6.3: Same as Fig. 6.1 for the completely correlated scenario.

## Chapter 7

# Asymptotic Ergodic Capacity Region and Rate Optimization of a Multiple Access OFDM MIMO Channel with Separately-Correlated Rician Fading

*Giorgio Taricco, Erwin Riegler*

### 7.1 Introduction

An important problem of network information theory is the derivation of the capacity region of a multiple access channel. In spite of intense research efforts carried out through the recent decades, there are several open problems in this area that have wide implications to the theory of communications and computation [15].

The multiple access channel capacity region is the largest set of rate vectors that are achievable by the different channel users and has been studied exhaustively in the literature [15]. It admits a simple expression for the Gaussian multiple access channel that has been extended to the Gaussian multiple access MIMO channel by Yu *et al.* [26], who provided an iterative water-filling algorithm aimed at finding the optimum user signal covariance matrices that maximize the sum rate of the channel. Their result applies when the multiple access MIMO channel is perfectly known at the transmitter and at the receiver.

However, in many situations it is more reasonable to assume that only the channel state information at the receiver (CSIR) and the channel distribution in-



formation at the transmitter (CDIT) are available. In this case, the problem of finding the optimum covariance matrices is still open in the general setting [16]. A notable result in this area has been provided by Hösli *et al.* [2], who proved that the ergodic and outage achievable rate region increase (as sets) monotonically with the singular values of the line-of-sight (LOS) component of the channel matrix. In this contribution we assume that full CSIR but only CDIT is available. The channel matrix for each user includes a LOS component and separate spatial correlation at the transmitter and the receiver for each delay tap. In particular, we also allow different receive correlation matrices at the receiver for each delay tap *and* for each user, which is done by introducing "virtual delays". Note that contrary to the semi-correlated case [13], where either the spatial correlation matrix at the transmitter or at the receiver is the identity matrix, the channel matrices in the frequency domain are no longer of the Kronecker type.

Based on these assumptions we provide two algorithms to compute the optimum covariance matrices which maximize the (weighted) rate sums in order to obtain the asymptotic ergodic capacity region. Our analysis is based on a recent result [11] which allows to calculate the wideband asymptotic capacity in the single-user case when the number of transmit and receive antennas grows asymptotically large though the number of users and their SNR's are finite.

Our results generalize previous works where we derived the optimum covariance matrices for a MIMO channel with interference [10] and for the dominant face of a multiple access channel [7] in the narrowband case assuming a common receive correlation matrix. Related works on the asymptotic approximation of the mutual information in the single-user case are [3] and [5] for the separately-correlated Rayleigh fading narrowband and wideband channel, respectively, [4] for the Rician fading narrowband channel without spatial correlation, and [6, 8, 9] for the separately-correlated Rician fading channel. Similar single-user capacity results were also obtained in [31] using an asymptotic method based on Stieltjes transforms and in [50] using the Newton barrier method [14] combined with Monte-Carlo simulations.

An expression for the asymptotic sum mutual information for arbitrary input distributions of a multiple access narrowband separately-correlated Rayleigh fading MIMO channel has been derived in [12].

## 7.2 System model

We consider a wideband frequency selective multiple access separately-correlated Rician fading MIMO channel with  $K$  users transmitting to a single receiver. User  $k$ , for  $k \in \mathcal{K} \triangleq \{1, \dots, K\}$ , is endowed with  $t(k)$  transmit antennas and the receiver with  $r$  receive antennas. We assume that the maximum number of re-

solvable paths of the channel is  $L = \lceil B\tau \rceil = |\mathcal{L} \triangleq \{0, \dots, L-1\}|$  (where  $B$  is the signal bandwidth and  $\tau$  the is the delay spread), so that the channel can be modeled by the following equation:

$$\mathbf{y}[n] = \sum_{k \in \mathcal{K}} \sum_{\ell \in \mathcal{L}} \mathbf{H}_k[\ell] \mathbf{x}_k[n - \ell] + \mathbf{z}[n] \quad (7.1)$$

for  $n \in \mathcal{N} \triangleq \{0, \dots, N-1\}$ . Here,  $n$  is a discrete time index;  $\mathbf{x}_k[n] \in \mathbb{C}^{t(k) \times 1}$  is the transmitted signal vector of user  $k$  at time  $n$ ;  $\mathbf{H}_k[\ell] \in \mathbb{C}^{r \times t(k)}$  is the channel matrix impulse response of user  $k$  sampled at delay time  $\ell$ ;  $\mathbf{z}[n] \in \mathbb{C}^{r \times 1}$  is the additive noise vector at time  $n$ ; and  $\mathbf{y}[n] \in \mathbb{C}^{r \times 1}$  is the received signal vector at time  $n$ . We assume stationary and white (in time) noise samples with

$$\mathbf{Q}_Z \triangleq \mathbb{E}[\mathbf{z}[n]\mathbf{z}[n]^H].$$

We also assume uncorrelated scattering and no correlation between the channel matrices for different users. Spatial correlation is taken into account by setting

$$\begin{aligned} \mathbf{H}_k[0] &= \bar{\mathbf{H}}_k + \mathbf{R}_k[0]^{1/2} \mathbf{W}_k[0] \mathbf{T}_k[0]^{1/2} \\ \mathbf{H}_k[\ell] &= \mathbf{R}_k[\ell]^{1/2} \mathbf{W}_k[\ell] \mathbf{T}_k[\ell]^{1/2} \quad \forall \ell \in \mathcal{L} \setminus \{0\} \end{aligned} \quad (7.2)$$

where  $\bar{\mathbf{H}}_k \triangleq \mathbb{E}[\mathbf{H}_k[0]]$ ,  $\mathbf{R}_k[\ell]$  is the receive spatial correlation matrix of user  $k$  at delay time  $\ell$ , and  $\mathbf{T}_k[\ell]$  is the transmit spatial correlation matrix of user  $k$  at delay time  $\ell$ . The matrices  $\mathbf{W}_k[\ell]$  are iid with entries distributed as  $\mathcal{N}_c(0, 1)$ .

Furthermore, we assume that the communication system is based on an orthogonal-frequency-division multiplexing (OFDM) scheme allowing to convert the frequency selective fading channel into a set of independent (conditionally on the channel matrices) frequency flat channels. A key assumption is the insertion of a cyclic prefix (CP) with length equal to the delay spread, which amounts to the following set of conditions:

$$\mathbf{x}_k[n] = \mathbf{x}_k[n + N] \quad n = -L + 1, \dots, -1. \quad (7.3)$$

The CP condition limits the information throughput by a factor of  $\kappa \triangleq \frac{N}{N+L-1}$ . Next, we calculate the discrete Fourier transform (DFT) of the channel equation. Let  $\omega \triangleq \exp(-j2\pi/N)$  so that, in consideration of the CP condition (7.3), we can transform the channel equation (7.1) into

$$\hat{\mathbf{y}}[m] \triangleq \sum_{n \in \mathcal{N}} \mathbf{y}[n] \omega^{mn} \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}} \sum_{\ell \in \mathcal{L}} \mathbf{H}_k[\ell] \mathbf{x}_k[n - \ell] \omega^{mn} = \sum_{k \in \mathcal{K}} \hat{\mathbf{H}}_k[m] \hat{\mathbf{x}}_k[m] + \hat{\mathbf{z}}[m] \quad (7.4)$$

for  $m \in \mathcal{N}$ , where we defined the DFTs  $\hat{\mathbf{H}}_k[m] \triangleq \sum_{\ell \in \mathcal{L}} \mathbf{H}_k[\ell] \omega^{m\ell}$ ,  $\hat{\mathbf{x}}_k[m] \triangleq \sum_{n \in \mathcal{N}} \mathbf{x}_k[n] \omega^{mn}$ , and  $\hat{\mathbf{z}}[m] \triangleq \sum_{n \in \mathcal{N}} \mathbf{z}[n] \omega^{mn}$  with  $\hat{\mathbf{Q}}_Z \triangleq N \mathbf{Q}_Z = \mathbb{E}[\hat{\mathbf{z}}[m] \hat{\mathbf{z}}^H[m]]$ . We assume an average power constraint for each user  $k \in \mathcal{K}$ :

$$P_k \geq \sum_{n \in \mathcal{N}} \text{Tr}(\mathbb{E}[\mathbf{x}_k[n] \mathbf{x}_k[n]^H]) = \frac{1}{N} \sum_{m \in \mathcal{N}} \text{Tr}(\hat{\mathbf{Q}}_k[m])$$

with  $\hat{\mathbf{Q}}_k[m] \triangleq \mathbb{E}[\hat{\mathbf{x}}[m] \hat{\mathbf{x}}[m]^H]$ .

Note that the entries of  $\hat{\mathbf{H}}_k[m]$  are again Gaussian and

$$\hat{\mathbf{H}}_k[m] \sim \hat{\mathbf{H}}_k \triangleq \bar{\mathbf{H}}_k + \sum_{\ell \in \mathcal{L}} \mathbf{R}_k[\ell]^{1/2} \mathbf{W}_k[\ell] \mathbf{T}_k^{1/2}[\ell] \quad (7.5)$$

for all  $m \in \mathcal{N}$ . Following standard assumptions [1, 9, 10], we define the Rice Factor of user  $k$  as

$$K_k^R \triangleq \frac{\|\bar{\mathbf{H}}_k\|^2}{\sum_{\ell \in \mathcal{L}} \text{Tr}(\mathbf{T}_k[\ell]) \text{Tr}(\mathbf{R}_k[\ell])} \quad (7.6)$$

and the signal-to-noise power ratio  $\text{SNR}_k$  of user  $k$  as

$$\text{SNR}_k \triangleq \frac{P_k (K_k^R + 1) \sum_{\ell \in \mathcal{L}} \text{Tr}(\mathbf{T}_k[\ell]) \text{Tr}(\mathbf{R}_k[\ell])}{t(k) \text{Tr}(\mathbf{Q}_Z)}, \quad (7.7)$$

which corresponds to the received SNR when the transmitted input signal is iid, see the discussion in [7].

### 7.3 Achievable rate region

From the transformation property of the entropy under linear transformations  $h(\mathbf{A}\mathbf{x}) = h(\mathbf{x}) + \log |\det(\mathbf{A})|$  ([15, p. 254]) and the fact that we have independent channels (conditionally on the channel matrices) in the frequency domain (7.4) it is plain to see that the mutual information can be calculated either in the time domain or in the frequency domain:

$$\begin{aligned} I &\triangleq \frac{\kappa}{N} I(\hat{\mathbf{x}}_1[0], \dots, \hat{\mathbf{x}}_K[N-1]; \hat{\mathbf{y}}[0], \dots, \hat{\mathbf{y}}[N-1] \mid \mathfrak{S}) \\ &= \frac{\kappa}{N} I(\mathbf{x}_1[0], \dots, \mathbf{x}_K[N-1]; \mathbf{y}[0], \dots, \mathbf{y}[N-1] \mid \mathfrak{S}) \\ &= \frac{\kappa}{N} \sum_{m \in \mathcal{N}} I(\mathbf{x}_1[m], \dots, \mathbf{x}_K[m]; \mathbf{y}[m], \dots, \mathbf{y}[m] \mid \mathfrak{S}_m) \end{aligned}$$

with  $\mathfrak{S}_m \triangleq \{\mathbf{H}_k[m] \mid k \in \mathcal{K}\}$  and  $\mathfrak{S} \triangleq \bigcup_{m \in \mathcal{N}} \mathfrak{S}_m$ . The ergodic capacity region can therefore be derived in the frequency domain:

**Proposition 1** Denoting by  $\mathcal{S} \subset \mathcal{K}$  an ordered subset of  $\mathcal{K}$  and by  $\mathcal{S}^c \triangleq \mathcal{K} \setminus \mathcal{S}$  its ordered complement in  $\mathcal{K}$ , the ergodic capacity region is given by the following union of polytopes:

$$\begin{aligned}
C &= \bigcup_{\text{Tr}(\hat{\mathbf{Q}}_k \leq P_k)} \{R_{\mathcal{K}} \mid \sum_{k \in \mathcal{S}} R_k \leq \\
&\quad \frac{\kappa}{N} \sum_{m \in \mathcal{N}} \mathbb{E}[I(\hat{\mathbf{x}}_{\mathcal{S}}[m]; \hat{\mathbf{y}}[m] \mid \hat{\mathbf{x}}_{\mathcal{S}^c}[m], \mathfrak{S}_m)] \forall \mathcal{S} \subset \mathcal{K}\} \\
&= \bigcup_{\text{Tr}(\hat{\mathbf{Q}}_k \leq P_k)} \{R_{\mathcal{K}} \mid \sum_{k \in \mathcal{S}} R_k \leq \\
&\quad \kappa \mathbb{E}[\ln \det(\hat{\mathbf{H}}_{\mathcal{S}} \hat{\mathbf{Q}}_{\mathcal{S}} \hat{\mathbf{H}}_{\mathcal{S}}^H + \hat{\mathbf{Q}}_Z)] - \kappa \ln \det(\hat{\mathbf{Q}}_Z) \forall \mathcal{S} \subset \mathcal{K}\}
\end{aligned}$$

where  $R_{\mathcal{K}} \triangleq (R_1, \dots, R_K)$  denotes the rate vector of all users  $k \in \mathcal{K}$ ,  $\hat{\mathbf{x}}_{\mathcal{S}}[m] \triangleq (\hat{\mathbf{x}}_k[m]^T \mid k \in \mathcal{S})^T$ ,  $\hat{\mathbf{H}}_{\mathcal{S}} \triangleq (\hat{\mathbf{H}}_k \mid k \in \mathcal{S})$ ,  $\hat{\mathbf{Q}}_{\mathcal{S}} \triangleq \text{diag}(\hat{\mathbf{Q}}_k \mid k \in \mathcal{S})$ , and  $\hat{\mathbf{Q}}_k[m] = \hat{\mathbf{Q}}_k$  is the same for all  $m \in \mathcal{N}$ .

*Proof* The proof follows from:

$$\begin{aligned}
&\frac{1}{N} \sum_{m \in \mathcal{N}} \mathbb{E}[I(\hat{\mathbf{x}}_{\mathcal{S}}[m]; \hat{\mathbf{y}}[m] \mid \hat{\mathbf{x}}_{\mathcal{S}^c}[m], \mathfrak{S}_m)] \\
&= \frac{1}{N} \sum_{m \in \mathcal{N}} (\mathbb{E}[\ln \det(\hat{\mathbf{H}}_{\mathcal{S}} \hat{\mathbf{Q}}_{\mathcal{S}}[m] \hat{\mathbf{H}}_{\mathcal{S}}^H + \hat{\mathbf{Q}}_Z)] - \ln \det(\hat{\mathbf{Q}}_Z)) \\
&\leq \mathbb{E}[\ln \det(\hat{\mathbf{H}}_{\mathcal{S}} \hat{\mathbf{Q}}_{\mathcal{S}} \hat{\mathbf{H}}_{\mathcal{S}}^H + \hat{\mathbf{Q}}_Z)] - \ln \det(\hat{\mathbf{Q}}_Z),
\end{aligned}$$

with  $\hat{\mathbf{Q}}_{\mathcal{S}}[m] \triangleq \text{diag}(\hat{\mathbf{Q}}_k[m] \mid k \in \mathcal{S})$ , where we used the fact that:

1. The average mutual information depends only on the statistics of the channel matrices defined in (7.5).
2. The average mutual information is concave in the covariance matrices.
3. Jensen's inequality.
4.  $\frac{1}{N} \sum_{m \in \mathcal{N}} \text{diag}(\hat{\mathbf{Q}}_k[m])$  is independent of  $m$  so that we can assume that  $\hat{\mathbf{Q}}_k[m] = \hat{\mathbf{Q}}_k \forall m \in \mathcal{N}$ .

The convex hull operation may be omitted because any convex linear combination of rate vectors is in the polytope defined by the same convex linear combination of covariance matrices (see the proof of Theorem 2 in [26] which holds for any concave function of positive semidefinite matrices).  $\blacksquare$

To maximize the extreme points in each polytope we introduce weighted rate sums [26]. For each permutation  $\pi$  of  $\mathcal{K}$  let  $\mathbf{w} \in \mathcal{D}^\pi$  where

$$\mathcal{D}^\pi \triangleq \{\mathbf{w} = (w_k \mid k \in \mathcal{K}) \in \mathbb{R}_+^K \mid \sum w_k = 1, w_{\pi(K)} \geq \dots \geq w_{\pi(1)} \geq 0\}.$$

Then the extreme point corresponding to  $\mathbf{w}$  is achieved by the set of covariance matrices  $\{\hat{\mathbf{Q}}_k^{\mathbf{w}} \mid k \in \mathcal{K}\}$  that maximize the weighted rate sum

$$\begin{aligned} & \sum_{k \in \mathcal{K}} w_k R_k \\ & \leq \sum_{k \in \mathcal{K}} \Delta_{\pi(k)} \mathbb{E}[\ln \det(\hat{\mathbf{H}}_{\mathcal{S}_{\pi(k)}} \hat{\mathbf{Q}}_{\mathcal{S}_{\pi(k)}} \hat{\mathbf{H}}_{\mathcal{S}_{\pi(k)}}^H + \hat{\mathbf{Q}}_Z)] \\ & \quad - w_{\pi(K)} \ln \det(\hat{\mathbf{Q}}_Z) \\ & \triangleq I_{\mathbf{w}} \end{aligned} \tag{7.8}$$

subject to

$$\text{Tr}(\hat{\mathbf{Q}}_k \leq P_k) \quad \text{and} \quad \hat{\mathbf{Q}}_k \geq 0 \quad \forall k \in \mathcal{K}$$

where  $\mathcal{S}_{\pi(k)}$  is the ordered set defined as  $\mathcal{S}_{\pi(k)} \triangleq \{\pi(j) \mid j \in \mathcal{K}, j \geq k\}$ ,  $w_{\pi(0)} \triangleq 0$ , and  $\Delta_{\pi(k)} \triangleq w_{\pi(k)} - w_{\pi(k-1)}$ .

## 7.4 Optimization

In order to solve this optimization problem we use a recent result allowing to calculate the average mutual information of a single-user *wideband* separately-correlated Rician MIMO fading channel when the number of transmit/receive antennas grows asymptotically large:

**Proposition 2** The asymptotic single-user ergodic mutual information when the channel matrix is of the form

$$\mathbf{H} = \bar{\mathbf{H}} + \sum_{\ell \in \mathcal{L}} \mathbf{R}[\ell]^{1/2} \mathbf{W}[\ell] \mathbf{T}[\ell]^{1/2} \in \mathbb{C}^{r \times t} \tag{7.9}$$

the noise covariance is  $\mathbf{Q}_Z \in \mathbb{C}^{r \times r}$ , and the signal covariance is  $\mathbf{Q} \in \mathbb{C}^{t \times t}$ , is given by

$$\begin{aligned} & \mathbb{E}[I(\mathbf{x}; \mathbf{y} \mid \mathbf{H})] \sim \\ & \mu_I(\mathbf{Q}_Z^{-1/2} \bar{\mathbf{H}}, \{\mathbf{Q}_Z^{-1/2} \mathbf{R}[\ell] \mathbf{Q}_Z^{-1/2}\}, \{\mathbf{T}[\ell]\}, \mathbf{Q}) \end{aligned}$$

nat/complex dimension, where we defined

$$\begin{aligned} \mu_I(\bar{\mathbf{H}}, \{\mathbf{R}[\ell]\}, \{\mathbf{T}[\ell]\}, \mathbf{Q}) &\triangleq \ln \det(\mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell]) + \\ &\ln \det(\mathbf{I}_t + (\sum_{\ell \in \mathcal{L}} z_\ell \mathbf{T}[\ell] + \bar{\mathbf{H}}^H(\mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell])^{-1} \bar{\mathbf{H}}) \mathbf{Q}) \\ &- \sum_{\ell \in \mathcal{L}} w_\ell z_\ell. \end{aligned}$$

Here,  $\{w_\ell\}$  and  $\{z_\ell\}$  are the positive solutions of the fixed point equations:

$$\begin{aligned} w_\ell &= \text{Tr} \left\{ \check{\mathbf{T}}[\ell] [\mathbf{I}_t + \sum_{\ell \in \mathcal{L}} z_\ell \check{\mathbf{T}}[\ell] + \check{\mathbf{H}}^H (\mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell])^{-1} \check{\mathbf{H}}]^{-1} \right\} \\ z_\ell &= \text{Tr} \left\{ \mathbf{R}[\ell] [\mathbf{I}_r + \sum_{\ell \in \mathcal{L}} w_\ell \mathbf{R}[\ell] + \check{\mathbf{H}} (\mathbf{I}_t + \sum_{\ell \in \mathcal{L}} z_\ell \check{\mathbf{T}}[\ell])^{-1} \check{\mathbf{H}}^H]^{-1} \right\} \end{aligned}$$

with  $\check{\mathbf{H}} \triangleq \bar{\mathbf{H}} \mathbf{Q}^{1/2}$  and  $\check{\mathbf{T}}[\ell] \triangleq \mathbf{Q}^{1/2} \mathbf{T}[\ell] \mathbf{Q}^{1/2}$ .

*Proof* see [11]

Applying this result to the upper bound on the weighted rate sum in equation (7.8) yields:

$$I_w \sim \sum_{k \in \mathcal{K}} \Delta_{\pi(k)} \mu_I(\tilde{\mathbf{H}}_{\mathcal{S}_{\pi(k)}}, \{\tilde{\mathbf{R}}[\underline{\ell}]\}, \{\mathbf{T}_{\mathcal{S}_{\pi(k)}}[\underline{\ell}]\}, \hat{\mathbf{Q}}_{\mathcal{S}_{\pi(k)}})$$

where we defined

$$\begin{aligned} \tilde{\mathbf{H}}_{\mathcal{S}} &\triangleq (\hat{\mathbf{Q}}_Z^{-1/2} \bar{\mathbf{H}}_k \mid k \in \mathcal{S}) \\ \mathbf{T}_{\mathcal{S}}[\underline{\ell}] &\triangleq \text{diag}(\delta_{k,j} \mathbf{T}_j[\ell] \mid j \in \mathcal{S}) \forall \underline{\ell} \in \mathcal{L} \\ \tilde{\mathbf{R}}[\underline{\ell}] &\triangleq \hat{\mathbf{Q}}_Z^{-1/2} \mathbf{R}_k[\ell] \hat{\mathbf{Q}}_Z^{-1/2} \forall \underline{\ell} \in \mathcal{L} \end{aligned}$$

for subsets  $\mathcal{S} \subset \mathcal{K}$ . Here, we introduced “virtual delays”  $\underline{\ell} \triangleq (k, \ell) \in \mathcal{L}_{\mathcal{S}} \triangleq \mathcal{S} \times \mathcal{L}$  to cope with the fact that we allow for different received correlation matrices for the transmitting users. Note that there are several systems of fixed point equations, each of which corresponding to a certain set  $\mathcal{S}$ . To distinguish them we write  $\{w_{\underline{\ell}}^{\mathcal{S}}\}$  and  $\{z_{\underline{\ell}}^{\mathcal{S}}\}$  with  $\underline{\ell} \in \mathcal{L}_{\mathcal{S}}$  for the corresponding solution of the system:

$$\begin{aligned} w_{\underline{\ell}}^{\mathcal{S}} &= \text{Tr} \left\{ \check{\mathbf{T}}_{\mathcal{S}}[\underline{\ell}] [\mathbf{I}_{t_{\mathcal{S}}} + \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} z_{\underline{\ell}}^{\mathcal{S}} \check{\mathbf{T}}_{\mathcal{S}}[\underline{\ell}] + \check{\mathbf{H}}_{\mathcal{S}}^H (\mathbf{I}_r + \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} w_{\underline{\ell}}^{\mathcal{S}} \tilde{\mathbf{R}}[\underline{\ell}])^{-1} \check{\mathbf{H}}_{\mathcal{S}}]^{-1} \right\} \\ z_{\underline{\ell}}^{\mathcal{S}} &= \text{Tr} \left\{ \tilde{\mathbf{R}}[\underline{\ell}] [\mathbf{I}_r + \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} w_{\underline{\ell}}^{\mathcal{S}} \tilde{\mathbf{R}}[\underline{\ell}] + \check{\mathbf{H}}_{\mathcal{S}} (\mathbf{I}_{t_{\mathcal{S}}} + \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} z_{\underline{\ell}}^{\mathcal{S}} \check{\mathbf{T}}_{\mathcal{S}}[\underline{\ell}])^{-1} \check{\mathbf{H}}_{\mathcal{S}}^H]^{-1} \right\} \end{aligned}$$

with  $\check{\mathbf{H}}_{\mathcal{S}} \triangleq \tilde{\mathbf{H}}_{\mathcal{S}} \mathbf{Q}_{\mathcal{S}}^{1/2}$ ,  $\check{\mathbf{T}}_{\mathcal{S}}[\underline{\ell}] \triangleq \mathbf{Q}_{\mathcal{S}}^{1/2} \mathbf{T}_{\mathcal{S}}[\underline{\ell}] \mathbf{Q}_{\mathcal{S}}^{1/2}$ , and  $t_{\mathcal{S}} \triangleq \sum_{k \in \mathcal{S}} t(k)$  for  $\mathcal{S} \subset \mathcal{K}$ .

To proceed we define the quantities

$$\mathbf{X}_{\mathcal{S}} \triangleq \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} z_{\underline{\ell}} \mathbf{T}_{\mathcal{S}}[\underline{\ell}] + \tilde{\mathbf{H}}_{\mathcal{S}}^{\mathbf{H}} (\mathbf{I}_r + \sum_{\underline{\ell} \in \mathcal{L}_{\mathcal{S}}} w_{\underline{\ell}} \tilde{\mathbf{R}}[\underline{\ell}])^{-1} \tilde{\mathbf{H}}_{\mathcal{S}}.$$

Furthermore, for each  $\mathcal{S} = \{i_1, \dots, i_{|\mathcal{S}|}\} \subset \mathcal{K}$  and  $k \in \mathcal{S}$  we define the matrices

$$\begin{aligned} \tilde{\mathbf{Q}}_{k,\mathcal{S}} &\triangleq \text{diag}(\mathbf{0}_{t(i_1)}, \dots, \hat{\mathbf{Q}}_k, \dots, \mathbf{0}_{t(i_{|\mathcal{S}|})}) \\ \tilde{\mathbf{Q}}_{k,\mathcal{S}}^c &\triangleq \hat{\mathbf{Q}}_{\mathcal{S}} - \tilde{\mathbf{Q}}_{k,\mathcal{S}}. \end{aligned}$$

Noting that  $(\mathbf{I} + \mathbf{A}\mathbf{B})^{-1}\mathbf{A} \geq \mathbf{0}$  for arbitrary positive semidefinite matrices  $\mathbf{A}, \mathbf{B}$  of the same size [17, p. 19] we can rewrite the expressions:

$$\begin{aligned} &\ln \det(\mathbf{I}_{t_{\mathcal{S}}} + \mathbf{X}_{\mathcal{S}} \hat{\mathbf{Q}}_{\mathcal{S}}) - \ln \det(\mathbf{I}_{t_{\mathcal{S}}} + \mathbf{X}_{\mathcal{S}} \tilde{\mathbf{Q}}_{k,\mathcal{S}}^c) \\ &= \ln \det(\mathbf{I}_{t_{\mathcal{S}}} + (\mathbf{I}_{t_{\mathcal{S}}} + \mathbf{X}_{\mathcal{S}} \tilde{\mathbf{Q}}_{k,\mathcal{S}}^c)^{-1} \mathbf{X}_{\mathcal{S}} \tilde{\mathbf{Q}}_{k,\mathcal{S}}) \\ &= \ln \det(\mathbf{I}_{t(k)} + \mathbf{A}_{k,\mathcal{S}} \hat{\mathbf{Q}}_k) \end{aligned} \quad (7.10)$$

with  $\mathbf{A}_{k,\mathcal{S}} \triangleq [(\mathbf{I}_{t_{\mathcal{S}}} + \mathbf{X}_{\mathcal{S}} \tilde{\mathbf{Q}}_{k,\mathcal{S}}^c)^{-1} \mathbf{X}_{\mathcal{S}}]_k$  for arbitrary  $\mathcal{S} \subset \mathcal{K}$  with  $k \in \mathcal{S}$ , where  $[\mathbf{A}]_k$  denotes the submatrix of  $\mathbf{A}$  obtained by extracting the elements of the rows and columns with indexes from  $\sum_{j \in \mathcal{S}}^{j < k} t(j) + 1$  to  $\sum_{j \in \mathcal{S}}^{j \leq k} t(j)$ . Using these manipulations and the fact that

$$\frac{\partial \mu_I}{\partial w_{\underline{\ell}}^{\mathcal{S}}} = 0 \quad \text{and} \quad \frac{\partial \mu_I}{\partial z_{\underline{\ell}}^{\mathcal{S}}} = 0 \quad \forall \underline{\ell} \in \mathcal{L}_{\mathcal{S}}$$

we further get:

$$\begin{aligned} &\nabla_{\hat{\mathbf{Q}}_{\pi(k)}} I_w \\ &\sim \nabla_{\hat{\mathbf{Q}}_{\pi(k)}} \sum_{j=1}^k \Delta_{\pi(j)} \mu_I(\tilde{\mathbf{H}}_{\mathcal{S}_{\pi(j)}}, \{\tilde{\mathbf{R}}[\underline{\ell}]\}, \{\mathbf{T}_{\mathcal{S}_{\pi(j)}}[\underline{\ell}]\}, \hat{\mathbf{Q}}_{\mathcal{S}_{\pi(j)}}) \\ &= \nabla_{\hat{\mathbf{Q}}_{\pi(k)}} \phi_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)}) \end{aligned} \quad (7.11)$$

with

$$\begin{aligned} \phi_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)}) &\triangleq \\ &\sum_{j=1}^k \Delta_{\pi(j)} \ln \det(\mathbf{I}_{t(\pi(k))} + \mathbf{A}_{\pi(k), \mathcal{S}_{\pi(j)}} \hat{\mathbf{Q}}_{\pi(k)}) \end{aligned} \quad (7.12)$$

for fixed  $\{w_{\underline{\ell}}^{\mathcal{S}_{\pi(j)}}\}$  and  $\{z_{\underline{\ell}}^{\mathcal{S}_{\pi(j)}}\}$  ( $j = 1, \dots, k$ ). This implies that we can alternately solve the systems of fixed point equations and maximize the expressions  $\phi_{\pi(k)}$  with respect to  $\hat{\mathbf{Q}}_{\pi(k)}$ , resulting in the following optimization algorithm:

**Algorithm 3** (*Optimization for arbitrary  $w$* )

initialize  $\hat{\mathbf{Q}}_{\mathcal{K}} = \text{diag}(P_k/t(k)\mathbf{I}_{t(k)} \mid k \in \mathcal{K})$

repeat

    solve fixed point equations  $\{w_{\underline{\ell}}^{\mathcal{S}_{\pi(k)}}\}$  and  $\{z_{\underline{\ell}}^{\mathcal{S}_{\pi(k)}}\} \forall k \in \mathcal{K}$

    for  $k \in \mathcal{K}$

        for fixed  $\{w_{\underline{\ell}}^{\mathcal{S}_{\pi(j)}}\}$  and  $\{z_{\underline{\ell}}^{\mathcal{S}_{\pi(j)}}\}$  where  $j = 1, \dots, k$

        solve  $\hat{\mathbf{Q}}_{\pi(k)}^w = \arg \max(\phi_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)}))$

        subject to  $\text{Tr}(\hat{\mathbf{Q}}_{\pi(k)}) \leq P_{\pi(k)}$  and  $\hat{\mathbf{Q}}_{\pi(k)} \geq \mathbf{0}$

    end

    set  $\hat{\mathbf{Q}}_k = \hat{\mathbf{Q}}_k^w \forall k \in \mathcal{K}$

until the desired accuracy is reached.

For arbitrary weights no *explicit* water-filling solution exists unless all the matrices  $\mathbf{A}_{k, \mathcal{S}_{\pi(j)}}$  which appear in the objective function in equation (7.12) commute, which would imply that we could diagonalize them simultaneously. Nevertheless, this objective function is concave and we can use the Newton barrier method [14, Chapter 11.3] to maximize it. This method has the advantage of avoiding the (highly non-linear) inequality constraints  $\{\hat{\mathbf{Q}}_k \geq \mathbf{0}\}$ . When the positive semidefinite constraint is tight we replace the objective function  $\phi_{\pi(k)}$  by

$$\phi'_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)}) \triangleq \phi_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)}) + (1/t) \ln \det(\hat{\mathbf{Q}}_{\pi(k)}) \quad (7.13)$$

with barrier function  $(1/t) \ln \det(\hat{\mathbf{Q}}_{\pi(k)})$  ( $t > 0$ ) and iterate Algorithm 3 over  $t$ , as proposed in [14, Algorithm 11.1]. Thanks to the manipulations in equations (7.10), the gradient and the Hessian of the objective function (7.13) can be computed with little effort, see [50] equations (9) and (10), respectively.

### 7.4.1 Simultaneous water-filling

Assume that we want to maximize the rate sum of a subset of users  $\mathcal{S} = \{i_1, \dots, i_{|\mathcal{S}|}\} \subset \mathcal{K}$  with *equal* weights:

$$w_k = \begin{cases} 1/|\mathcal{S}| & \text{if } k \in \mathcal{S} \\ 0 & \text{otherwise,} \end{cases}$$

i.e. we want to maximize the *sum rate*. The sum in the objective function (7.12) then collapses to a single term and we get:

$$\begin{aligned} \hat{\mathbf{Q}}_{\pi(k)}^{\text{sum}} &= \arg \max(\phi_{\pi(k)}(\hat{\mathbf{Q}}_{\pi(k)})) \\ &= \arg \max(\ln \det(\mathbf{I}_{t(\pi(k))} + \mathbf{A}_{\pi(k), \mathcal{S}} \hat{\mathbf{Q}}_{\pi(k)})) \end{aligned}$$



for all  $k \in \mathcal{S}$ . Furthermore, we can apply Hadamard's inequality [15, p. 680] and get the following standard water-filling solution for the optimum covariance matrices:

$$\hat{\mathbf{Q}}_k^{\text{sum}} = \mathbf{U}_k \left[ \frac{1}{\lambda_k} \mathbf{I}_{t(k)} - \Lambda_k^{-1} \right]_+ \mathbf{U}_k^H, \quad (7.14)$$

where we used the eigenvalue decomposition

$$\mathbf{A}_{k,\mathcal{S}} = \mathbf{U}_k \Lambda_k \mathbf{U}_k^H$$

and defined  $(\Lambda_k^{-1})_i \triangleq \infty$  in the case of  $(\Lambda_k)_i = 0$ . This gives us the following algorithm:

**Algorithm 4** (*Simultaneous water-filling*)

initialize  $\hat{\mathbf{Q}}_{\mathcal{K}} = \text{diag}(P_k/t(k)\mathbf{I}_{t(k)} \mid k \in \mathcal{K})$

repeat

    solve fixed point equations for  $\{w_{\underline{\ell}}^S\}$  and  $\{z_{\underline{\ell}}^S\}$

    for  $k \in \mathcal{K}$

        for fixed  $\{w_{\underline{\ell}}^S\}$  and  $\{z_{\underline{\ell}}^S\}$

            solve  $\hat{\mathbf{Q}}_k^{\text{sum}} = \mathbf{U}_k \left[ \frac{1}{\lambda_k} \mathbf{I}_{t(k)} - \Lambda_k^{-1} \right]_+ \mathbf{U}_k^H$

    end

    set  $\hat{\mathbf{Q}}_k = \hat{\mathbf{Q}}_k^{\text{sum}} \forall k \in \mathcal{S}$

until the desired accuracy is reached.

## 7.5 Numerical simulations

In this section we present numerical results illustrating the accuracy of the asymptotic optimization of the (weighted) rate sum. Figure 7.1 illustrates the two-user asymptotic ergodic capacity regions for several SNR values, bounded by the thick curves. Red chain dotted lines bound the achievable rate regions using simultaneous water-filling (Algorithm 4), where the optimum covariance matrices maximize the sum rates. Magenta solid curves are obtained by weighted rate sum optimization (Algorithm 3), where we iterated  $w_2$  within  $w_2 \in \{0, 0.05, \dots, 1\}$  and plotted the curves connecting the corresponding extreme points. Blue dashed lines report single-user bounds on  $R_2$  and  $R_1$  corresponding to  $w_2 = 1$  and  $w_2 = 0$ , respectively. Black thin dashed lines show the bounds on the achievable rate regions for iid power allocation. "x"- and "o"- markers refer to Monte-Carlo simulations and label the extreme points and faces in the polytopes, respectively to confirm the accuracy of the asymptotic approximation.

We have chosen the following parameters for our simulations:

1. The number of transmit/receive antennas is equal to  $t(1) = t(2) = r = 4$ .
2. The number of paths is  $L = 3$  with spatial correlation

$$(\mathbf{T}_k[\ell])_{a,b} = \rho(\ell)\alpha_k(\ell)^{|a-b|} \quad (\mathbf{R}_k[\ell])_{a,b} = \beta_k(\ell)^{|a-b|}$$

for  $k \in \mathcal{K}$  and  $\ell \in \mathcal{L}$ . Here,  $\rho(\ell) \triangleq 0.7^\ell$  corresponds to a power delay profile with exponential decay. Furthermore, we set  $\alpha_1(\ell) = \alpha_2(\ell) = \beta_1(\ell) = \beta_2(\ell)$  with  $\alpha_1(0) = 0.7$ ,  $\alpha_1(1) = 0.8$ , and  $\alpha_1(2) = 0.9$ .

3. For the line-of-sight components we took the all one matrix normalized such that  $K_1^R = K_1^R = 0$  dB.
4. The signal to noise ratio is iterated within  $\mathbf{SNR}_1 = \mathbf{SNR}_2 \in \{5, 10, 15, 20, 25\}$  dB,  $\mathbf{Q}_Z = \mathbf{I}_r$ , and  $\kappa = 1$ .

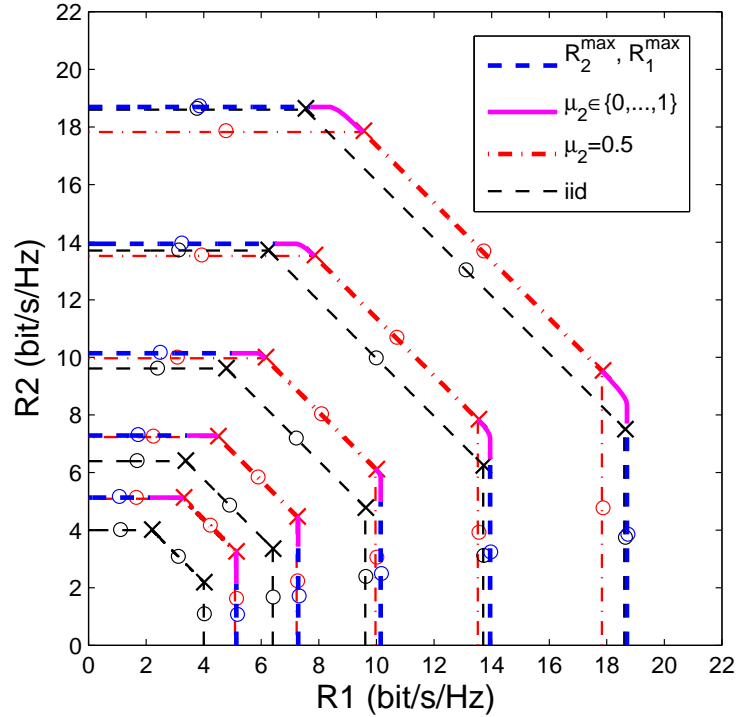


Figure 7.1: Asymptotic ergodic capacity regions of a two user scenario with  $\mathbf{SNR}_1 = \mathbf{SNR}_2 \in \{5, 10, 15, 20, 25\}$  dB.

It can be noticed that there is a considerable distance between the dominant faces of the iid rate regions and the  $w_2 = 0.5$  rate regions, which are obtained by si-

multaneous water-filling, in the low SNR regime. The effect of weighted rate sum optimization is mainly visible in the high SNR regime.

## 7.6 Conclusion

We investigated the ergodic capacity region of a multiple access separately-correlated Rician fading MIMO *wideband* channel using an asymptotic approach. The channel model we used is sufficiently general to allow for considering different spatial correlation matrices for each user and delay at the transmitter and the receiver. Two algorithms have been proposed to optimize the covariance matrices for the (weighted) rate sums in order to obtain the asymptotic ergodic capacity region. We showed by numerical simulations that our approach is very accurate even for a small number of antennas.

# Bibliography

- [1] F. Farrokhi, G. Foschini, A. Lozano, and R. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, March 2001.
- [2] D. Hösl, Y.-H. Kim, and A. Lapidoth, "Monotonicity results for coherent MIMO Rician Channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4334–4339, Dec. 2005.
- [3] A.L. Moustakas, S.H. Simon, and A.M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large n analysis," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.
- [4] A.L. Moustakas and S.H. Simon, "Random matrix theory of multi-antenna communications: the Ricean channel," *J. Phys. A: Math. Gen.*, vol. 38, no. 49, pp. 10859–10872, Dec. 2005.
- [5] A.L. Moustakas and S.H. Simon, "On the Outage Capacity of Correlated Multiple-Path MIMO channels", *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.
- [6] E. Riegler and G. Taricco, "Second-Order Statistics of the Mutual Information of the Asymptotic Separately-Correlated Rician Fading MIMO Channel with Interference," *IEEE GLOBECOM 2007*, Washington, USA, Nov. 26-30, 2007.
- [7] E. Riegler and G. Taricco, "On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel," *IEEE GLOBECOM 2007*, Washington, USA, Nov. 26-30, 2007.
- [8] G. Taricco, "On the capacity of separately-correlated MIMO Rician fading channels," *IEEE Globecom 2006*, San Francisco, USA, Nov. 27-Dec. 1, 2006.

- [9] G. Taricco, "Asymptotic mutual information statistics of separately-correlated Rician fading MIMO channels," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3490–3504, August 2008.
- [10] G. Taricco and E. Riegler, "On the ergodic capacity of the asymptotic separately-correlated Rician fading MIMO channel with interference," *IEEE ISIT 2007*, Nice, France, June 24–30, 2007.
- [11] G. Taricco and E. Riegler, "Asymptotic Ergodic Capacity of Wideband MIMO Channels with Separately-Correlated Rician Fading," *IEEE Globecom 2008*, New Orleans, USA, Nov. 30–Dec. 4, 2008.
- [12] C.K. Wen and K.K. Wong, "Asymptotic Analysis of Spatially Correlated MIMO Multiple-Access Channels With Arbitrary Signaling Inputs for joint and Separate Decoding," vol. 53, no. 1, pp. 252–268, Jan. 2007.
- [13] H. Bölcskei, D. Gesbert, and A.J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 225–234, Feb. 2002.
- [14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., 2006.
- [16] A. Goldsmith, S.A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.
- [17] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [18] 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. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
- [19] M.R. McKay, P.J. Smith, H.A. Suraweera, and I.B. Collings, "Accurate Approximations for the Capacity Distribution of OFDM-Based Spatial Multiplexing", in *Proc. IEEE ICC*, Glasgow, UK, pp. 5377–5382, Jun. 2007.
- [20] A.L. Moustakas, S.H. Simon, A.M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not

- so) large  $N$  analysis,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.
- [21] A.L. Moustakas, S.H. Simon, “On the Outage Capacity of Correlated Multiple-Path MIMO Channels,” *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.
  - [22] G. Raleigh and J. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE Trans. on Commun.*, Vol. 46, No. 3, pp. 357–366, March 1998.
  - [23] E. Riegler and G. Taricco, “Second-order statistics of the mutual information of the asymptotic separately-correlated Rician fading MIMO channel with interference,” *Proc. Globecom 2007*, Washington, DC, USA, November 26–30, 2007.
  - [24] E. Riegler and G. Taricco, “On the ergodic capacity region of the separately correlated Rician fading multiple access MIMO channel,” *Proc. Globecom 2007*, Washington, DC, USA, November 26–30, 2007.
  - [25] G. Taricco and E. Riegler, “On the ergodic capacity of the asymptotic separately-correlated Rician fading MIMO channel with interference,” in *Proc. IEEE ISIT 2007*, Nice, France, June 24–30, 2007.
  - [26] W. Yu, W. Rhee, S. Boyd, and J.M. Cioffi, “Iterative Water-Filling for Gaussian Vector Multiple-Access Channels,” *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
  - [27] D.W. Bliss, K.W. Forsythe, A.O. Hero, III, and A.F. Yegulalp, “Environmental Issues for MIMO Capacity,” *IEEE Trans. Signal Processing*, vol. 50, no. 9, pp. 2128–2142, Sept. 2002.
  - [28] M. Chiani, M. Win, and H. Shin, “Capacity of MIMO systems in the presence of interference,” in *IEEE Global Communications Conference (GLOBECOM)*, Nov. 2006.
  - [29] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
  - [30] S.N. Diggavi and T.M. Cover, “The worst additive noise under a covariance constraint,” *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.
  - [31] J. Dumont, P. Loubaton, and S. Lasaulce, “On the capacity achieving transmit covariance matrices of MIMO correlated Rician channels: A large system approach,” *Globecom 2006*, San Francisco, USA, Nov. 27-Dec. 1, 2006.

- [32] G.J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.
- [33] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [34] D. Gesbert, D. Bolskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.
- [35] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.
- [36] R. Horn and C. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.
- [37] S.A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, July 2004.
- [38] S.K. Jayaweera and H.V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. on Wireless Communications*, vol. 4, no. 3, pp. 1102–1111, March 2005.
- [39] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543–1553, Sept. 2004.
- [40] Y.-H. Kim and S.-J. Kim, "On the convexity of  $\log \det(I + KX^{-1})$ ," url: <http://arxiv.org/pdf/cs.IT/0611043>, 2006.
- [41] S.L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sept. 2001.
- [42] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350–1359, July 2004.

- [43] E. Riegler and G. Taricco, "Second-order statistics of the mutual information of the asymptotic separately-correlated Rician fading MIMO channel with interference," *IEEE Globecom 2007*.
- [44] A. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
- [45] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.
- [46] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T Bell Laboratories Technical Report*, BL0112170-950615-07TM, June 1995.  
Subsequently published as "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November–December 1999.
- [47] A. M. Tulino, A. Lozano and S. Verdú, "Capacity-achieving Input Covariance for Single-User Multi-Antenna Channels," *IEEE Trans. on Wireless Communications*, vol. 5, no. 3, pp. 662–671, March 2006.
- [48] S. Venkatesan, S.H. Simon, and R.A. Valenzuela, "Capacity of a Gaussian MIMO Channel with nonzero mean," *IEEE Vehicular Technology Conference*, Oct. 2003.
- [49] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. on Inform. Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.
- [50] M. Vu and A. Paulraj, "Capacity optimization for Rician correlated MIMO wireless channels," *Asilomar Conference 2005*, ASilomar CA, USA, Nov. 2005.
- [51] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, no. 6, pp. 871–878, June 1987.
- [52] F. A. Berezin and A. A. Kirillov, *Introduction to Superanalysis*. Dordrecht, Netherlands: Reidel Publishing Company, 1987.
- [53] R.S. Blum, J.H. Winters, and N.R. Sollenberger, "On the capacity of cellular systems with MIMO," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 242–244, June 2002.



- [54] R.S. Blum, "MIMO capacity with interference," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 793–801, June 2003.
- [55] S. Catreux, P.F. Driessen, and L.J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 334–336, Nov. 2000.
- [56] J. Dumont, P. Loubaton, S. Lasaulce, and M. Debbah, "On the asymptotic performance of MIMO correlated Ricean channels," *ICASSP05*, vol. 5, pp. 813–816, March 2005.
- [57] G.J. Foschini, "Layered space-time architecture wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech Journal*, pp. 41–59, Autumn 1996.
- [58] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983–2010, June 2005.
- [59] F. Haake, *Quantum signature of Chaos*. Berlin: Springer-Verlag, 1991.
- [60] R.R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2821–2828, Nov. 2003.
- [61] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.
- [62] I.E. Telatar, "Capacity of multi-antenna gaussian channels," AT&T Bell Labs, Tech. Rep., June 1995.
- [63] J.H. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. SAC-5, no. 5, pp. 871–878, June 1987.
- [64] C.-K. Wen and K.-K. Wong, "Asymptotic analysis of spatially correlated MIMO multiple-access channels with arbitrary signaling inputs for joint and separate decoding," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 252–268, Jan. 2007.
- [65] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Ann. Appl. Probab.*, vol. 17, no. 3, 2007, pp. 875–930.

- [66] B. Rimoldi and R. Urbanke, "A Rate-Splitting Approach to the Gaussian Multiple-Access Channel," *IEEE Trans. Inform. Theory*, vol. 42, no. 2, pp. 364–375, March 1996.
- [67] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge Univ. Press, 1998.
- [68] G. Caire and S. Shamai (Shitz), "On the Achievable Throughput of a Multi-Antenna Gaussian Broadcast Channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [69] S. Vishwanath, N. Jindal and A. Goldsmith, "Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [70] W. Yu and J. M. Cioffi, "Sum Capacity of Gaussian Vector Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.
- [71] Martin Schubert and Holger Boche, "QoS-Based Resource Allocation and Transceiver Optimization," *Foundations and Trends in Communications and Information Theory*, vol. 49, no. 10, pp. 383–529, 2005.
- [72] William Thomson, *Handbook of Game Theory*, Vol. 2, Elsevier Science, 1994.
- [73] Roy D. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1348, Sep 1995.
- [74] Holger Boche and Martin Schubert, "Multiuser Interference Balancing for General Interference Functions – A Convergence Analysis," *ICC Conference 2007*, Glasgow, Scotland, Jun, 2007.
- [75] Holger Boche, Martin Schubert, Eduard A. Jorswieck and Aydin Sezgin, "A General Framework for Concave/Convex Interference Coordination Problems and Network Utility Optimization," *ITG-SA-WORKSHOP*, Vienna, Austria, Feb, 2007.
- [76] Holger Boche, Martin Schubert, Nikola Vucic and Siddharth Naik, "Non-Symmetric Nash Bargaining Solution for Resource Allocation in Wireless Networks and Connection to Interference Function Calculus," *EUSIPCO conference 2007*, Poznań, Poland, Sep, 2007.
- [77] Holger Boche and Martin Schubert, "Characterization of the Structure of General Interference Functions," *ISIT 2007*, Nice, France, Jun, 2007.

- [78] Holger Boche and Martin Schubert, “Advanced Interference Calculus – A General Framework for Interference Coordination,” *SPAWC*, Helsinki, Finland, Jun, 2007.
- [79] Martin Schubert and Holger Boche, “Properties and Operational Characterization of Proportionally Fair Resource Allocation,” *SPAWC*, Helsinki, Finland, Jun, 2007.
- [80] Holger Boche and Martin Schubert, “The supportable QoS region of a multiuser system with log-convex interference functions,” *ICASSP*, Hawaii, USA, Apr, 2007.
- [81] Martin Schubert and Holger Boche, “Throughput Maximization for Uplink and Downlink Beamforming with independent Coding,” *CISS Conference 2003*, USA, Mar, 2003.
- [82] H. Weingarten, Y. Steinberg and S. Shamai (Shitz), “The Capacity Region of the Gaussian MIMO Broadcast Channel,” *IEEE Trans. Inform. Theory*, 2006.