Impact of channel state information
on the analysis and design of
multi-antenna communication systems

Ph.D. dissertation by

Miquel Payaró i Llisterri
E-mail: `miquel.payaro@cttc.es`

Thesis advisors:

Prof. Miguel Ángel Lagunas
Director of Centre Tecnològic de Telecomunicacions de Catalunya
Professor at Universitat Politècnica de Catalunya
E-mail: `m.a.lagunas@cttc.es`

Dr. Antonio Pascual Iserte
Assistant Professor at Universitat Politècnica de Catalunya
Associate Researcher at Centre Tecnològic de Telecomunicacions de Catalunya
E-mail: `tonip@gps.tsc.upc.es`

December 2006

A la meva família i a la Núria,

iv

Abstract

During the last decade, there has been a steady increase in the demand of high data rates that are to be supported by wireless communication applications. Among the different solutions that have been proposed by the research community to cope with this new demand, the utilization of multiple antennas arises as one of the best candidates due to the fact that it provides both an increase in reliability and also in information transmission rate. Although the use of multiple antennas at the receiver side dates back from the sixties, the full potential of multiple antennas at both communication ends has been both theoretically and practically recognized in the last few years.

The design of proper multi-antenna communication systems to satisfy the high data rates demand depends not only on the chosen figure of merit or performance metric, but also on the quantity and the quality of the channel state information that is available at the communication ends. In this dissertation we deal with the analysis and design of different architectures for multiple-antenna communication systems for various degrees of quality and quantity of channel state information. The analysis section is devoted to the study of capacity and achievable rates and the part that deals with design is aimed at the synthesis of practical communication systems that maximize a certain performance measure.

Firstly, we focus our attention on multiple antenna single-user communication systems with perfect channel state information, which is an idealization of actual practical systems. In this context, we review well known capacity results and deal with the practical characterization of a linear transmitter that is designed to maximize the reliability of the wireless multi-antenna link. Some analogies between the optimal linear transmitter design and the theory of constellation construction are also pointed out.

Secondly, we stay in a single-user scenario and we move onto the case where the channel state information is incomplete. In this case, a detailed capacity analysis is presented dealing with the ergodic and compound capacity formulations, which arise depending on the model utilized to characterize the channel. While in rapidly varying channels the ergodic capacity is a key measure of the rates that can be achieved by any communication system, in slow varying or fixed channels the compound capacity measures the minimum transmission rate that can be sustained during the transmission of the message.

Next, we shift to the case where the available channel state information is imperfect. Precisely, we deal with a practical communication system called spatial Tomlinson-Harashima precoder and study its achievable rate capabilities. Due to the versatile architecture of the spatial Tomlinson-Harashima precoder we are able to perform the study for the single and multi-user scenarios. For both cases, a design is presented which is robust to the uncertainties of the channel state information and which is aimed at maximizing the transmission rate.

Finally, staying in the multi-user scenario with imperfect channel state information, we present a transmission architecture that is robust to the uncertainties of the side information that is available at both the transmitter and the receiver. The robustness criterion is to minimize the transmitted power while guaranteeing a certain quality of service per user for every possible realization of the channel that is compatible with the available channel state information.

Resum

Al llarg d'aquesta última dècada, s'ha produit un creixement constant en la demanda d'elevades taxes de transmissió de dades que han de suportar les aplicacions sobre comunicacions sense fils. Entre les diferents solucions ideades per la comunitat recercaire per tal de fer front a aquesta nova demanda, la utilització de múltiples antenes s'erigeix com una de les millors candidates degut al fet que proporciona simultàniament una millora en les taxes de transmissió i en la fiabilitat en la recepció de les dades. L'ús d'antenes múltiples en un dels extrems de la comunicació data de la dècada dels seixanta, nogensmenys ha estat en aquests últims anys quan s'ha pogut provar, tant en els camps teòric com pràctic, tot el potenical que possibilita la presència de múltiples antenes en ambdós extrems de la comunicació.

El disseny adequat de sistemes de comunicació amb múltiples antenes per satisfer aquesta demanda no només depèn de la funció de mèrit (o de la mètrica de rendiment) escollida, sinó que també es veu afectat per la quantitat i la qualitat de la informació de l'estat del canal que es troba disponible als extrems de la comunicació. Aquesta tesi tracta sobre l'anàlisi i el disseny d'arquitectures per sistemes de comunicació amb múltiples antenes i amb diferents nivells de quantitat i qualitat de la informació de l'estat del canal. La secció d'anàlisi es centra en l'estudi de la capacitat i les taxes de transmissió assolibles per aquests tipus de sistemes de comunicació i la part de disseny queda més encarada a la síntesi de sistemes de comunicació pràctics amb l'objectiu de maximitzar el rendiment d'acord amb la mètrica de rendiment escollida.

Primerament, l'atenció es centra en sistemes de comunicació amb múltiples antenes per a un únic usuari amb informació perfecte de l'estat del canal, que suposa una idealització dels sistemes pràctics que s'empren en la realitat. En aquest context, es revisen resultats de capacitat que són ben coneguts, i es caracteritza, a més, un transmissor lineal dissenyat per tal de maximitzar la fiabilitat de l'enllaç sense fils amb múltiples antenes. Addicionalment, s'apunten una sèrie d'analogies entre el disseny del transmissor lineal òptim i la teoria de construcció de constel.lacions de símbols.

En segon lloc, es roman en un escenari de comunicacions amb un únic usuari i es considera el cas on la informació sobre l'estat del canal és incompleta. En aquest cas, es presenta un anàlisi detallat sobre la capacitat a través de les formulacions ergòdica i composta (*compound*), les quals prenen significat depenent del model utilitzat per caracteritzar el canal. Mentres que en canals ràpidament variants la capacitat ergòdica és la mesura clau de les taxes de transmissió assolibles per qualsevol sistema de comunicació, en canals fixos o de variació lenta, és la capacitat composta, la que mesura la mínima taxa de transmissió assolible de forma sostinguda durant la transmissió del missatge.

Seguidament, es considera el cas on la informació disponible sobre l'estat del canal és imperfecta. Precisament, es discorre sobre un sistema de comunicació pràctic anomenat Precodifi-

cador Espacial de Tomlinson i Harashima i s'estudien les seves potencialitat en termes de taxes de transmissió assolibles. Gràcies a l'arquitectura versàtil del Precodificador Espacial de Tomlinson i Harashima l'esmentat estudi es duu a terme tant per escenaris amb un únic usuari com per escenaris amb múltiples usuaris. Per aquests dos casos, es presenta així doncs un disseny que és robust a les incerteses de la informació de l'estat del canal i que té per objectiu minimitzar les pèrdues de taxa de transmissió d'informació.

Finalment, restant en un escenari amb múltiples usuaris amb coneixement imperfecte de l'estat del canal, es presenta una arquitectura de transmissió que és robusta a les incerteses de la informació sobre l'estat del canal disponible tant en el transmissor com en el receptor. La variable per al disseny robust és la distribució de potència entre els símbols d'informació destinats a cada usuari, i el criteri d'optimització és minimitzar la potència total transmesa, tot garantint una determinada qualitat de servei per cada usuari i per qualsevol possible realització del canal que sigui compatible amb la informació disponible sobre l'estat del canal.

Acknowledgments

Although a Ph.D. dissertation is a very personal piece of work, it is not less true that the following pages wouldn't have been the same without all the people who, directly or indirectly, have influenced me during these last years. I can do no more than express my gratitude towards all of them.

First and foremost, I want to thank my advisors, Miguel Ángel Lagunas and Antonio Pascual. It was mainly through discussing and working with them that I shaped my understanding of my subject, of Electrical Engineering at large. I feel very lucky to have had Miguel Ángel and Toni as advisors, and look forward to having more opportunities to learn from and collaborate with them in the future. It has been a pleasure to make this long journey under their supervision not only in the good moments, but also in the moments where my enthusiasm was not at its highest.

Another of the many reasons I am grateful to Miguel Ángel and Toni was their encouragement and support of my meeting and working with other researchers. As a result, in a relatively short time, I have collaborated with many wonderful people and I feel deeply indebted to all of them: I thank Xavier Mestre who guided my first steps in the amazing world of information theory and Ana Pérez who introduced me to the field of precoding. I would also like to thank Ami Wiesel, as it has been a pleasure working with him, and Jinhong Yuan for he warmly welcomed me into his research group during the (Australian) summer of 2005 in the University of New South Wales. Finally, I want to thank Daniel Palomar, I am also very happy to have had the opportunity to work with him. Needless to say, I'm looking forward to continuing the collaboration with Dani in the (very near) future.

I wish to thank Mari Kobayashi and Jose Vicario for they kindly agreed to review this thesis, for the time they spent reading it, and for their valuable comments. I would also like to thank Margarita Cabrera and Josep Vidal for they awakened my curiosity for research during my last year as an undergraduate student. Finally, I would like to thank Stephan Pfletschinger for he kindly presented the IST'06 paper for me.

My working time during these last years would have been much duller without my friends and fellows. I thank very deep the *ghetto* people Jose, Ricardo, Toni M., Diego, Toni P., and Fran with whom we had memorable moments; and also Aitor, Aurora, Carme, Christian, Gemma, Javi, Jesús G., Joan Anton, Joana, Jordi C., Josep, Lluis B., Lluis V., Marc P., Marc R., Mari, Pavel, Patrícia, Raül, Xavi N., and all the people at CTTC. I wish also to thank the people that I met at UNSW, specially my gratitude goes to Ido, Tuyen, Noor, Tom, and Julia. Finally, I thank my office mates at the new CTTC building: Jordi S., Nizar, Jesús A., Bego, David, and Panos. Forgive me if your name should be here and is missing.

I am very grateful for the scholarships I received during my studies, which helped me both

immensely, both in allowing me to focus on my research, and in travel funds that helped me present my works and cooperate with other researchers. I thank the people at AGAUR[1], and at the CTTC PhD fellowship program, specially Simó.

Finally, I would like to thank my grandparents Miquel and Conxita, my parents Miquel and Araceli, my siblings Anna and Albert and also my sweet girlfriend Núria. They made completing this project possible and worthwhile. I love you all.

Miquel
December 2006.

Contents

Notation

Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and lower-case italics denote scalars.

$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	The set of all natural, integer, real and complex numbers, respectively.
$\mathbb{R}^+$	The set of all strictly positive real numbers.
$\mathbb{Z}^{n \times m}, \mathbb{R}^{n \times m}, \mathbb{C}^{n \times m}$	The set of $n \times m$ matrices with integer-, real-, and complex-valued entries, respectively. If $m = 1$, the index can be dropped.
$\mathbf{X}^T$	Transpose of the matrix $\mathbf{X}$.
$\mathbf{X}^H$	Complex conjugate and transpose (Hermitian) of the matrix $\mathbf{X}$.
$[\mathbf{X}]_{ij}$	(i, j)th component of the matrix $\mathbf{X}$.
$\mathrm{Tr}\,\mathbf{X}$	Trace of the matrix $\mathbf{X}$.
$\|\mathbf{X}\|$ or $\det(\mathbf{X})$	Determinant of the matrix $\mathbf{X}$.
$\mathbf{X}^{-1}$	Inverse of the matrix $\mathbf{X}$.
$\mathbf{X}^{\#}$	Moore-Penrose pseudo-inverse of the matrix $\mathbf{X}$.
$\mathbf{X}^{1/2}$	Hermitian square root of the positive semidefinite matrix $\mathbf{X}$, *i.e.*, $\mathbf{X}^{1/2}\mathbf{X}^{1/2} = \mathbf{X}$.
$\mathrm{diag}(\mathbf{A})$	Vector constructed with the elements in the diagonal of matrix $\mathbf{A}$.
$\mathrm{vec}(\mathbf{A})$	Vector constructed stacking the columns of matrix $\mathbf{A}$.
$\mathbf{I}$ or $\mathbf{I}_n$	Identity matrix and identity matrix of dimension $n \times n$, respectively.
$\arg$	Argument.
$\max, \min$	Maximum and minimum.
$\sup, \inf$	Supremum (lowest upper bound) and infimum (highest lower bound).
$(x)^{\star}$	Optimal value of variable x according to some criterion deduced from the context.
$(x)^+$	Positive part of the real scalar x, *i.e.*, $(x)^+ = \max\{0, x\}$.

$\lvert x \rvert$	Magnitude of the complex scalar x.
$\lfloor x \rfloor$	Closest integer lesser than x.
$\lVert \mathbf{x} \rVert^2$	Squared Euclidean norm of the vector $\mathbf{x}$: $\lVert \mathbf{x} \rVert^2 = \mathbf{x}^H \mathbf{x}$.
$\lVert \mathbf{x} \rVert_{\mathbf{A}}^2$	Squared norm of the vector $\mathbf{x}$ utilizing the metric defined by $\mathbf{A}$: $\lVert \mathbf{x} \rVert_{\mathbf{A}}^2 = \mathbf{x}^H \mathbf{A} \mathbf{x}$.
$\sim$	Distributed according to.
$\Pr\{\cdot\}$	Probability.
$p_{\mathbf{x}}(\mathbf{x})$, $p(\mathbf{x})$	Probability density function of the random variable $\mathbf{x}$.
$\mathbb{E}[\cdot]$	Mathematical expectation.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\lvert \mathcal{A} \rvert$	Cardinality of the set $\mathcal{A}$, *i.e.*, number of elements in $\mathcal{A}$.
$\mathfrak{Re}\{\cdot\}$	Real part.
$\mathfrak{Im}\{\cdot\}$	Imaginary part.
$\mathcal{L}$	Likelihood or Lagrangian function (deduced from the context).
$\propto$	Equal up to a scaling factor (proportional).
$\gtrsim, \lesssim$	Approximately greater (lesser) than
$\triangleq$	Defined as.
$\simeq$	Approximately equal.
$\lim$	Limit.
$\mathsf{o}(x)$	Landau symbol to denote that, if $f = \mathsf{o}(x)$, then $f/x \to 0$ in the neighborhood of some x_0.
$\log(\cdot)$	Natural logarithm.
$\log_a(\cdot)$	Base-a logarithm.
$\mathcal{Q}(x)$	$\frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp(-t^2/2)\mathrm{d}t$.
i, i', i''	Index variables to refer to a receiver antenna or user. They range from 1 to n_R or to n_U.
j, j', j''	Index variables to refer to a transmitter antenna. They range from 1 to n_T.
k, k', k''	Index variables to refer to a substream. They range from 1 to n_S.

Acronyms

a.s.	almost surely.
AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
BLAST	Bell-Labs Layered Space-Time.
BS	Base Station.
cdf	cumulative density function.
CSI	Channel State Information.
DPC	Dirty Paper Coding.
DSL	Digital Subscriber Line.
FDD	Frequency Division Duplexing.
IEEE	Institute of Electrical and Electronics Engineers.
i.i.d.	independent and identically distributed.
KKT	Karush Kuhn Tucker.
MIMO	Multiple-Input Multiple-Output.
MISO	Multiple-Input Single-Output.
ML	Maximum Likelihood.
MMSE	Minimum Mean Square Error.
MSE	Mean Square Error.
OSTBC	Orthogonal Space-Time Block Code.
pdf	probability density function.
PEP	Pairwise Error Probability.
QAM	Quadrature Amplitude Modulation.
QoS	Quality of Service.
QPSK	Quadrature Phase Shift Keying.
RF	Radio Frequency.
RMT	Random Matrix Theory.
RX	Receiver.
SIMO	Single-Input Multiple-Output.

SINR	Signal to Interference plus Noise Ratio.
SIR	Signal to Interference Ratio.
SISO	Single-Input Single-Output.
SNR	Signal to Noise Ratio.
STBC	Space-Time Block Code.
STC	Space-Time Code.
STHP	Spatial Tomlinson-Harashima Precoder/Precoding.
STTC	Space-Time Trellis Code.
SVD	Singular Value Decomposition.
TDD	Time Division Duplexing.
TX	Transmitter.
w.l.o.g.	without loss of generality.
w.r.t.	with respect to.
ZF	Zero Forcing.

Chapter 1

Introduction

1.1 Motivation

Information theory was formally initiated in 1948 by Shannon in his pioneering work [Sha48] in which he showed that reliable communication between a transmitter and a receiver was possible even in the presence of a noisy channel. For a given channel, the maximum information transmission rate at which the transmitter and receiver can communicate with an arbitrarily low probability of error and no delay constraints is, since then, called the channel *capacity*.

For the particular case of a unit-gain band-limited continuous channel corrupted with additive white gaussian noise (AWGN), Shannon obtained his celebrated formula for the capacity

$$C = W \cdot \log \frac{P_T + P_N}{P_N}, \qquad (1.1)$$

where W, P_T, and P_N represent the bandwidth, the average transmitted power, and the noise power, respectively. The channel capacity is usually measured in *bits* (base-2 logarithm) or in *nats* (base-e logarithm) per transmission. Sometimes it is useful to omit the dependence of the capacity on the bandwidth by normalizing the capacity by the used bandwidth, obtaining the *spectral efficiency*, usually measured in bits (or nats) per second per hertz.

Due to the logarithmic dependence of the spectral efficiency on the transmitted power, it is extremely expensive to increase the capacity by radiating more power. For example, for a sufficiently high signal to noise power ratio (SNR), doubling the transmitted power just yields one extra bit per second per hertz of spectral efficiency. In addition, increasing the transmitted power, specially in the mobile terminal, is not encouraged since it may violate regulation power masks and the actual effects of electromagnetic radiation in people's health is still a topic of research [Lin02].

Another approach to increase the capacity consists in utilizing a wider electromagnetic band. However, as many other natural resources, the electromagnetic spectrum is a scarce good and

thus, its utilization is, in general, regulated by the government through the awarding of licenses, which are extremely expensive. Consequently, it is essential to utilize the bandwidth in the most efficient way.

Recently, theoretical studies, [Tel99, Fos98], and also practical implementations, [Gol99], have shown that there exists an alternative technique which can increase the capacity without neither transmitting more power nor occupying a wider band when the channel exhibits rich scattering and its variations are accurately tracked by, at least, the receiver. This technique consists in utilizing multiple antennas at both the transmitter and the receiver ends.

Utilizing an array of antennas just at the receiver side dates back from the sixties, and then it was used with the only purpose to provide robustness again deep fades. Each receiving antenna was used to collect incoming energy, and then, by combining properly the different received signals, the SNR (and thus capacity) was increased. Although no extra power is needed, the increase in capacity still has a logarithmical dependence on the SNR, which makes this solution rather inefficient.

Combining multiple antennas both at the transmitter and receiver ends gives rise to the so called multi-input multi-output (MIMO) channels. For MIMO channels, when the SNR is high enough and the receiver tracks the variations of the channel, it can be shown that capacity grows *linearly* with the minimum between the number of transmit, n_T, and receive, n_R, antennas,

$$C \propto \min(n_T, n_R). \tag{1.2}$$

This linear growth of the capacity associated with MIMO channels is based on the premise that a rich scattering environment provides independent transmission paths from each transmit to each receive antenna. The constant multiplier associated with the linear scaling depends on the degree of the channel state information (CSI) at the transmitter side.

Much subsequent work has been aimed at characterizing MIMO channel capacity under different assumptions on the degree of CSI at the transmitter and receiver sides, and also under the imposition of different practical (and, in general, suboptimal) transmission schemes, such as, *e.g.*, the non-linear spatial Tomlinson-Harashima Precoder (STHP) [Fis02a].

Despite the fact that the ultimate performance of any communication system is dictated by the theoretical capacity limits, in the vast majority of cases the system designer has practical constraints (specially time delay and complexity) which separate the rates that can be achieved by the system from the capacity limits. In these cases, the system designer will try to maximize some practical performance measure, instead of aiming at reaching the capacity limits.

In this sense, practical MIMO communication system designers have two primary concerns. Firstly, the designer has to define a practical objective function to measure the global performance of the system. In general, this performance metric will be selected depending on which is

the purpose of the communication system. Secondly, the system designer will have to calculate the values of the system parameters that optimize the value of the selected performance metric of the system. Some of the most widely utilized performance measures include (but are not limited to) bit error rate (BER), mean square error (MSE), and distance among the received codewords.

From all that has been said above, the aim of this dissertation will be devoted to further analyze the influence of the quantity and quality of the CSI at the transmitter and receiver sides on the system capacity, and also the effect of imposing a particular transmitter structure on the system performance, which will be evaluated according to some of the above mentioned practical performance measures.

1.2   Outline of the dissertation

In general terms, this dissertation deals with the design and analysis of multi-antenna communication systems, mainly focusing the attention on transmission architectures. Precisely, the impact of the availability and quality of the CSI on the capacity and on the transmitter design is studied and robust transmission architectures are designed. The main advantage of these robust architecures is that their sensitivity to the presence of errors in the CSI is reduced with respect to their classical non-robust counterparts. The outline of each of the chapter is as follows. Note that the chapters with original contributions are chapters 3, 4, 5, and 6.

Chapter 1 has presented the motivation of this dissertation, is presenting the outline, and, in the following section, will list the contributions of this dissertation.

Chapter 2 overviews some basic concepts and reviews the state of the art concerning multi-antenna communication systems. Precisely the chapter is divided in two main parts: the first one is devoted to the state of the art and mathematical formulation of capacity, and the second one focuses on practical communication architectures.

Chapter 3 deals with communication in a single-user MIMO system where the CSI is considered to be perfect at both communication ends. After reviewing the capacity results for this particular scenario, the practical design of a linear transmitter is considered. The design criterion for the transmitter is such that the minimum distance among the received symbols is maximized, which is meaningful figure of merit when the receiver employs a maximum-likelihood (ML) detector. Additionally, it shows that this linear transmitter design is directly related to optimal symbol constellation construction.

Chapter 4 introduces incompleteness of the CSI at the transmitter in single-user MIMO communication systems. Specifically, it is considered that the transmitter is informed only with (or is only able to estimate) the magnitude of the channel complex coefficients. For this particular

choice of incomplete CSI, capacity results are presented encompassing the ergodic and compound formulations. For some configurations with a small number of antennas, closed form expressions for the capacity are found. For configurations with an arbitrary number of antennas, capacity results based on random matrix theory are also presented.

Chapter 5 deals with multi-antenna communication systems with imperfect CSI at the transmitter side. Precisely, it studies the effects of a noisy channel estimate at the transmitter on the achievable rates of a communication architecture equipped with STHP. In addition, it presents a robust design of this precoder so that the rate loss in the presence of errors in the CSI is minimized. Finally, due to the versatility of STHP to encompass multi-user systems, this chapter links the single-user and the multi-user scenarios for multi-antenna systems.

Continuing with the multi-user scenario described in chapter 5, chapter 6 is fully devoted to the study of a multi-antenna and multi-user communication system where it is considered that both the transmitter and the receivers have only access to an imperfect version of the CSI. Imposing a particular architecture at the transmitter side, the power allocation among the signals to be transmitted to the users is robustly designed with the criterion of minimizing the total transmitted power, while guaranteeing a certain predefined quality of service per user. The problem is formulated within the powerful framework of convex optimization, which enables the designer to efficiently find the solution to the robust transmitter design problem.

Chapter 7 concludes the dissertation and gives some topics for future research.

1.3   Research contributions

The main contribution of this dissertation is on the analysis and design of multi-antenna communication systems depending on the degree and the quality of the CSI. In the following, a detailed list of the research contributions in each chapter is presented.

Chapter 3

The main results of this chapter deal with the design of the optimal linear transmitter that maximizes the minimum distance among the received symbols. These results have been published in one conference paper and another paper has been submitted.

- M. Payaró, A. Pascual-Iserte, D. P. Palomar, M. A. Lagunas, "On linear transmitter designs for MIMO systems with ML detection," submitted to IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'07), Honolulu, HW, (USA), May 2006.

- M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Optimum linear transmitter design for MIMO systems with two QPSK data streams," in Proc. IEEE International Conference on Communications (ICC'06), Istanbul (Turkey), June 2006.

Chapter 4

The main contributions of this chapter are on the characterization of the transmit covariance matrix that maximizes the mutual information for a particular case of channel state uncertainty at the transmitter. Reports on these results have been published in six conference papers, and a journal paper has been submitted.

- M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Capacity evaluation in multiantenna systems under phase uncertainty," submitted to the special issue "Signal Processing for Uncertain Systems", in Circuits, Systems, and Signal Processing (CSSP), August 2006.

- M. Payaró, A. Wiesel, J. Yuan, and M. A. Lagunas, "On the capacity of linear vector Gaussian channels with magnitude knowledge and phase uncertainty," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'06), Toulouse (France), May 2006.

- M. Payaró, J. Yuan, and M. A. Lagunas, "On the capacity of MIMO systems with magnitude knowledge and phase uncertainty," in Proc. IEEE Australian Communications Theory Workshop, Perth (Australia), February 2006.

- M. Payaró, X. Mestre, A. I. Pérez-Neira, and M. A. Lagunas, "Robust power allocation techniques for MIMO systems under modulus channel knowledge at the transmitter," in Proc. IEEE Workshop on Signal Proc. Advances in Wireless Commun. (SPAWC'05), New York, NJ (USA), June 2005.

- M. Payaró, X. Mestre, A. I. Pérez-Neira, and M. A. Lagunas, "On power allocation under phase uncertainty in MIMO systems," in Proc. Winterschool on Coding and Information Theory, Bratislava (Slovakia), February 2005.

- M. Payaró, X. Mestre, and M. A. Lagunas, "Ergodic capacity of a 2×2 MIMO system under phase uncertainty at the transmitter," in Proc. IEEE International Conference on Electronics, Circuits and Systems (ICECS'04), Tel Aviv (Israel), December 2004.

- M. Payaró, X. Mestre, and M. A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter," in Proc. IEEE Information Theory Workshop (ITW'04), San Antonio, TX (USA), October 2004.

Chapter 5

The main results in this chapter involve the robust design of a communications system with STHP and the characterization of its achievable rates in the presence of errors in the design parameters. The results of this chapter have been published in one journal paper and two conference papers.

- M. Payaró, A. Pascual-Iserte, A. I. Pérez-Neira, and M. A. Lagunas, "On the achievable rates with spatial Tomlinson-Harashima precoding in the presence of errors in the CSI," to appear in IEEE Trans. on Wireless Communications (submitted December 2005, revised October 2006).

- M. Payaró, A. I. Pérez-Neira, and M. A. Lagunas, "Robustness evaluation of uniform power allocation with antenna selection for spatial Tomlinson-Harashima precoding," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'05), Philadelphia, PA (USA), March 2005.

- M. Payaró, A. I. Pérez-Neira, and M. A. Lagunas, "Achievable rates for generalized spatial Tomlinson-Harashima precoding," in Proc. IEEE Vehicular Technology Conference (VTC-Fall'04), Los Angeles, CA (USA), September 2004.

Chapter 6

The last chapter with original contributions deals with the robust design of the power allocation in a multi-user system where there is uncertainty in the parameters that characterize the channel, in such a way that a certain quality of service per user is guaranteed. Initial results have been published in one conference paper. In addition, a full report of the obtained results has been submitted to a journal publication and a patent application has been prepared (patent pending).

- M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Method and System for Robustly Transmitting the Minimum Power in Multi-User and Multi-Antenna Communications Systems with Imperfect Channel Knowledge," submitted on 2006/06/28 with submission number PCT/EP2006/006244 (patent pending).

- M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Robust Power Allocation Designs for Multiuser and Multiantenna Downlink Communication Systems through Convex Optimization," submitted to IEEE Journal on Selected Areas in Communications, June 2006.

- M. Payaró, A. Pascual-Iserte, J. Yuan, and M. A. Lagunas, "A Convex Optimization Approach for the Robust Design of Multiuser and Multiantenna Downlink Communica-

tion Systems," in Proc. IEEE Workshop on Signal Proc. Advances in Wireless Commun. (SPAWC'06), Cannes (France), July 2006.

Other contributions not presented in this dissertation

During the last four years, some results not directly related with this dissertation or such that the main research was not conducted by the author of this dissertation have also been published in five conference papers.

Some results concerning the analysis and design of transmit architectures with different degrees of knowledge of the channel state have been published in two conference papers:

- A. Pascual-Iserte, M. Payaró, A. I. Pérez-Neira, and M. A. Lagunas, "Impact of a Line of Sight Component on the Performance of a MIMO System Designed under Statistical Channel Knowledge," in Proc. IEEE Workshop on Signal Proc. Advances in Wireless Commun. (SPAWC'06), Cannes (France), July 2006.

- M. Payaró, A. Pascual-Iserte, A. I. Pérez-Neira, and M. A. Lagunas, "Flexible MIMO Architectures: Guidelines in the design of MIMO parameters," in Proc. IEEE Workshop on Signal Proc. Advances in Wireless Commun. (SPAWC'05), Invited Paper, New York, NJ (USA), June 2005.

A joint paper presenting part of the results from the MARQUIS project was published in the proceedings of a conference:

- M. Payaró, M. Realp, D. Bartolomé, A. Pascual-Iserte, X. Mestre, A. I. Pérez-Neira, "Comparative Use of Metrics for Link/System Level Simulations in WLAN MIMO Systems," in Proc. IST Summit 2006 (IST Mobile Communications Summit 2006). Mikonos (Greece), June 2006.

Some designs involving adaptive modulation and multiple antenna channels were presented in two conference papers.

- A. Pascual-Iserte, M. Payaró, A. I. Pérez-Neira, M. A. Lagunas, "Robust Adaptive Modulation for Throughput Maximization in MIMO Systems Combining OSTBC and Beamforming," in Proc. IST Mobile Communications Summit (IST'04), Lyon (France), June 2004.

- M. Payaró, M. A. Lagunas, "Adaptive Modulation in MISO Wireless Systems with Discrete Low-Rate Power Feedback", in Proc. European Signal Processing Conference (EUSIPCO'04), Viena (Austria), September 2004.

Finally, an early conference paper on MIMO channel modeling was presented.

- M. Cabrera, M. Payaró, J. Vidal, M. Hunukumbure and M. Beach, "2GHz MIMO Channel Model from Experimental Outdoor Data Analysis in UMTS", in Proc. IEEE Vehicular Technology Conference (VTC-Fall'03), Orlando, FL (USA), September 2003.

Chapter 2

Overview of multi-antenna communication systems

2.1 The multi-antenna communication system

2.1.1 General overview and benefits

In addition to the capacity advantage described in the previous chapter, communication over multi-antenna channels presents two main practical advantages with respect to traditional communication over single antenna channels. These gains are usually referred to as *diversity* and *multiplexing* gains.[1] An overview on the potential gains of MIMO channels can be found in [Böl02].

On the one hand, a MIMO communication system is said to have diversity gain d if, in the high SNR regime, the average error probability decays as snr^{-d}. Loosely speaking, the diversity gain can be viewed as an enhancement of reliability due to the reception (or transmission) of replicas of the same information that have experienced different fading paths. The diversity gain is based on the assumption that at least one of these paths will not be in a bad fade state. Traditionally, diversity has been exploited in the time or frequency domains, however, the presence of the additional space dimension (due to the use of multiple antennas) yields another source of diversity. Notice that spatial diversity gain is not exclusive of MIMO systems, as it can also be extracted from MISO or SIMO architectures.

On the other hand, a MIMO system is said to achieve multiplexing gain r if, also in the high SNR regime, its achievable rates scale as $r \log \mathsf{snr}$. In other words, multiplexing gain is the increase of rate that can be attained through the use of multiple antennas at both sides of the communication link, with respect to the rate achievable with a single antenna system, without utilizing additional power. Notice that, as opposed to the diversity gain, the multiplexing gain

---

[1]In addition, one can have a third gain, *beamforming* gain, but it will not be discussed here.
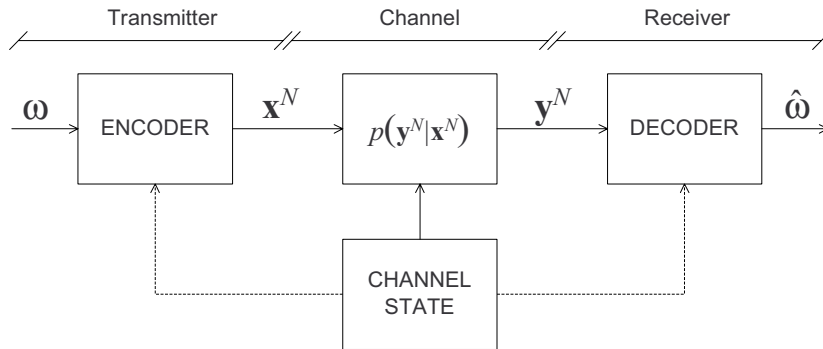
Figure 2.1: Generic MIMO single-user communication system.

can only be obtained with the simultaneous presence of multiple antennas at the transmitter and the receiver ends.

The designers of MIMO communication systems have focused their efforts in trying to obtain transceivers architectures which achieve either the maximum multiplexing gain [Fos96], or the the maximum diversity gain [Tar99a]. In [Zhe03], Zheng and Tse proved that both gains can be achieved but that, actually, there is a fundamental tradeoff between how much of each gain can be extracted. Precisely, for i.i.d. Rayleigh flat-fading MIMO channels with coherence time l, they proved that the optimal diversity gain achievable by any coding scheme of block length l and multiplexing gain r is $(n_T - r)(n_R - r)$ as long as $l \geq n_T + n_R - 1$. Their appealing interpretation for this result was that out of the total resource of n_T transmit and n_R receive antennas, it is as though r transmit and r receive antennas were used for multiplexing and the remaining $n_T - r$ transmit and $n_R - r$ receive antennas provided the diversity.

2.1.2   Mathematical model

As depicted in Figure 2.1, the multi-antenna point to point communication system is composed of three main logical blocks, the single transmitter of the message, the single receiver of the message, and the physical channel between them. Note that both the transmitter and the receiver can have access to some knowledge about the channel state. This knowledge may be partial or imperfect, and it may also include statistical properties in case that the channel can be considered random, or any other kind of information that can be used to describe the properties of the channel. In its most general form, this side information is usually referred to as CSI. More details on the CSI are given in §2.2.

Given the message $\omega \in \Omega$ that is destined to the receiver, the encoder transforms it into the codeword $\mathbf{x}^N(\omega)$ of length N, which is subsequently transmitted as $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ in N uses of the channel. The transmitted signal at time instant n, $\mathbf{x}_n$, is a vector of complex elements of
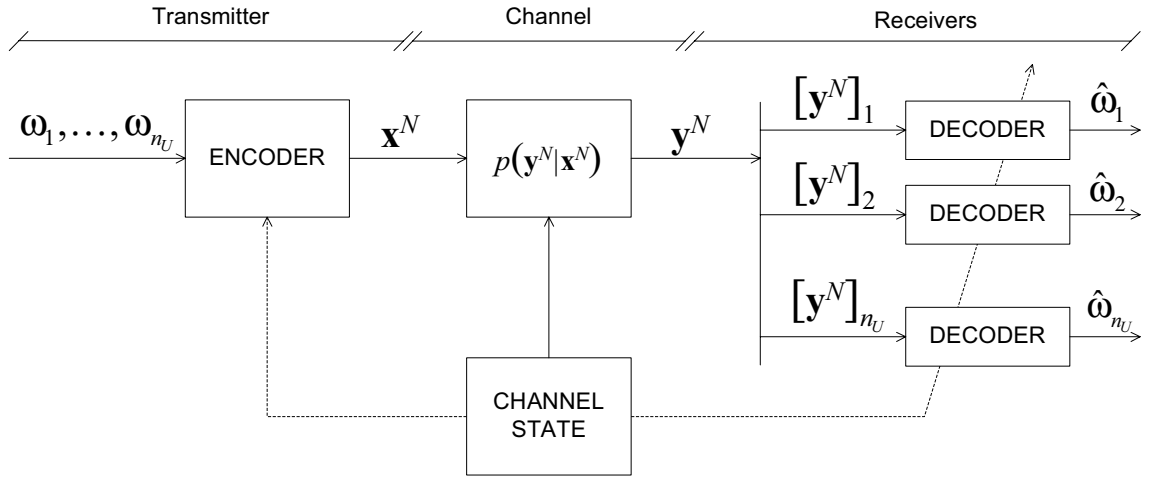
Figure 2.2: Generic MIMO single-user communication system.

dimension n_T such that its j-th element, $[\mathbf{x}_n]_j$, represents the signal transmitted through the j-th antenna.

The transmitted signal $\mathbf{x}^N(\omega)$ propagates through the physical channel characterized by the transition probability $p\left(\mathbf{y}^N|\mathbf{x}^N\right)$ and is received as the random sequence $\mathbf{y}^N$. Each element of this sequence $\mathbf{y}_n$, with $n = 1, \ldots, N$ is a vector of complex elements of dimension n_R such that its i-th element, $[\mathbf{y}_n]_i$, represents the signal received through the i-th antenna. In this dissertation we only deal with memoryless channels, which imply that

$$p\left(\mathbf{y}^N|\mathbf{x}^N\right) = \prod_{n=1}^{N} p(\mathbf{y}_i|\mathbf{x}_i) \tag{2.1}$$

Once the sequence $\mathbf{y}^N$ is received, the decoder takes a decision on which is the sequence that was actually sent, according to a (hopefully properly defined) decision rule given by $\hat{\omega}(\mathbf{y}^N)$.

In this dissertation, in addition to the single-user case we also deal with the more complex scenario where the transmitter, usually a base station (BS), transmits information simultaneously to n_U users, which are equipped with a single antenna (which implies that the total number of receiving antennas is equal to the number of users $n_R = n_U$).

As it will be seen in chapters 5 and 6 the are two differences between the single-user case and the multi-user case where each user has only one antenna, see Figure 2.2. The first difference is that the transmitter has to be able to simultaneously encode up to n_U messages $\omega_1, \omega_2, \ldots, \omega_{n_U}$ into the transmitted codeword $\mathbf{x}^N(\omega_1, \omega_2, \ldots, \omega_{n_U})$, which now depends on all the messages. The second difference is that, since each user is associated with one receiving antenna, the decoded message of i-th user can only depend on the i-th component of the received sequence, which
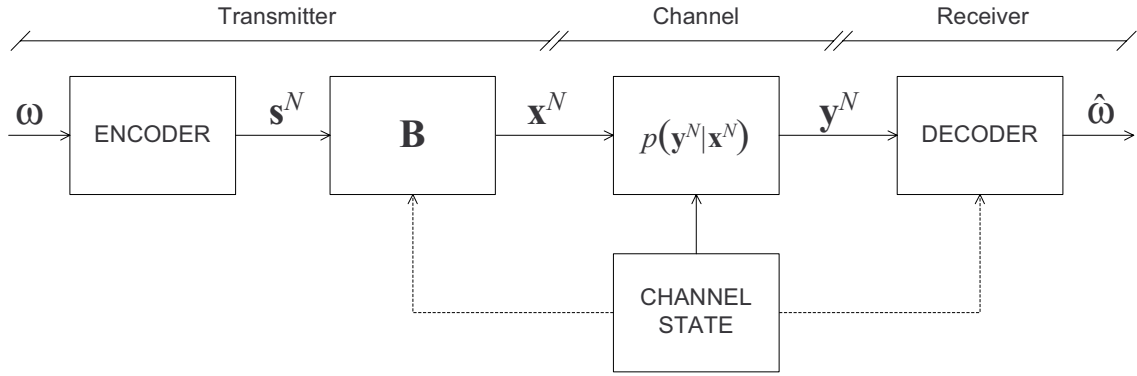
Figure 2.3: General capacity achieving structure for MIMO communication systems with side information at the transmitter. Notice that, as it is said in the main text, the encoder is independent of the CSI, and the dependence on the CSI at the transmitter is only through the linear matrix **B**.

implies that

$$\hat{\omega}_i = \hat{\omega}_i\big(\big[\mathbf{y}^N\big]_i\big), \quad 1 \le i \le n_U, \tag{2.2}$$

which is the reason why, in some works, the multi-user case, is also termed non-cooperative receivers.

Having specified these two differences, in the remainder of this chapter we will only deal with the single-user case. Nonetheless, the reader is aware that the multi-user case can be straightforwardly characterized.

2.1.3   The linear transmitter

The particular case where the transmitter is divided in two main blocks, as it is depicted in Figure 2.3 is of remarkable importance due to its simplicity and performance capabilities.

The first block encodes the message ω into the sequence of data symbols vector $\mathbf{s}^N$, independently of the CSI that is available at the transmitter side, only following the steps dictated by code construction theory. At each time instant n, the symbols vector $\mathbf{s}_n$ is a complex vector of dimension n_S, *i.e.*, $\mathbf{s}_n \in \mathbb{C}^{n_S}$.

The second block is allowed to depend on the CSI available at the transmitter, and it consists of a linear transformation of the sequence of data symbols vector into the transmitted vector $\mathbf{x}_n$. At each time instant n the transmitted vector is given by

$$\mathbf{x}_n = \mathbf{B}_n \mathbf{s}_n, \quad n = 1, \dots, N, \tag{2.3}$$

where $\mathbf{B}_n \in \mathbb{C}^{n_T \times n_S}$ is the linear transformation that is applied to the data symbols vector.

2.1.4   The linear MIMO channel

Although MIMO channels arise in many different communication scenarios such as wire line systems or frequency selective single antenna systems [Pal03a], our focus will spot on multi-antenna digital wireless systems.

As commented above, in its most general form, the MIMO channel is characterized by its transition probability density function, which is given by $p(\mathbf{y}|\mathbf{x})$, and which describes the probability of receiving the vector $\mathbf{y}$ conditioned on the fact that the vector $\mathbf{x}$ was actually transmitted. However, dealing with such a generic type of channel is usually very difficult, and often even unnecessary.

One of the most important and simplest MIMO channels is the *linear* MIMO channel (also termed linear vector channel). As it name suggests, in this case the output of the channel $\mathbf{y}$ is a linear function of the input $\mathbf{x}$. In addition, to model the thermal noise and other undesired effects that are present in the receiving radio-frequency front-ends, a noise term $\mathbf{n}$ is also included.

For the case of linear MIMO channels, the resulting input-output relation is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.4}$$

where $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ represents the linear response of the channel, such that its element $[\mathbf{H}]_{ij}$ denotes the channel path gain between j-th transmitter and i-th receiver.

In the vast majority of cases (and in this dissertation in particular), the noise term is modeled as a circularly symmetric complex Gaussian random variable with zero mean and covariance matrix given by $\mathbf{R}_n$. Or, formally:

$$\mathbb{E}\mathbf{n} = \mathbf{0}, \tag{2.5}$$

$$\mathbb{E}\mathbf{n}\mathbf{n}^H = \mathbf{R}_n. \tag{2.6}$$

In this particular case the channel transition probability $p(\mathbf{y}|\mathbf{x})$ is given by [Mil74]

$$p(\mathbf{y}|\mathbf{x}) = p_{\mathbf{n}}(\mathbf{y} - \mathbf{H}\mathbf{x}) = \frac{1}{\pi^{n_R}|\mathbf{R}_n|}\exp\left\{-(\mathbf{y} - \mathbf{H}\mathbf{x})^H\mathbf{R}_n^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})\right\}. \tag{2.7}$$

2.2   Uncertainty models for the CSI

As mentioned in the previous section, when designing the transmitter and the receiver, both of them may have access to a certain degree of CSI. Consequently, many possible design situations arise depending on the quantity and the quality of the CSI available during the design stage. The CSI is a generic term utilized to encompass all kind of information related to the channel, which may include, for example, the MIMO channel impulse response, the statistics of the received interferences, the statistics of the thermal noise, among many other aspects.

To slightly formalize the presence of CSI in the system we introduce the following notation:

$$\mathcal{K}^X \triangleq \{\text{Set of variables, functions, or descriptions that describe}$$
$$\text{what is known about the channel state at the site } X\} \tag{2.8}$$

$$\mathcal{U}^X \triangleq \{\text{Set of variables, functions, or descriptions that describe}$$
$$\text{what is unknown about the channel state at the site } X\} \tag{2.9}$$

where X can be utilized to design the transmitter or the receiver sides, *e.g.*, $\mathcal{K}^{\text{RX}}$ would be the generic name to represent what is known about the channel state at the receiver side.

In the following, we describe possible situations regarding the presence (or absence) of CSI at the two communication ends. The objective is not to provide an exhaustive description, but only to show a general view of the most common situations of CSI in multi-antenna systems. A more detailed description can be found in [Gol03].

In the most typical communication setup, the channel estimation and the estimation of the statistics of the interferences processes are performed at the receiver during the training period in which the transmitter sends a pilot or training sequence. Consequently, one of the most common situations corresponds to the case in which accurate (and ideally perfect) CSI is available at the receiver side of the communication system. Taking this into account, the only remaining degree of freedom is the availability of CSI at the transmitter side, as given by $\mathcal{K}^{\text{TX}}$ and $\mathcal{U}^{\text{TX}}$, and thus three different situations concerning the quantity and quality of the CSI available at the transmitter side can be identified:

- **No CSI:** the transmitter does not have any knowledge of any parameter concerning the channel or the interferences at the receiver. In this case, a reasonable transmitter strategy consists in utilizing space-time codes [Tar99b, Tar99a, Gan01].

- **Perfect CSI:** the transmitter has full knowledge of the instantaneous channel realization and, possibly, of the interferences statistics at the receiver. In this case, since full information is available, there are many possible strategies and optimization criteria to carry out the design depending on the detection method at the receiver or on the performance metric [Pal03a].

- **Imperfect CSI:** the transmitter has inaccurate knowledge about the parameters describing the channel. For example, the transmitter may be informed of an erroneous channel matrix $\widetilde{\mathbf{H}} \neq \mathbf{H}$. For the case of imperfect CSI, two main strategies can be considered, either the transmitter is designed to attain the maximum performance level for the worst possible situation of the channel among the ones that are compatible with the CSI (maximin or worst-case approach) or the transmitter is designed to have the best mean performance averaged over the unknown parameters of the CSI (statistical or Bayesian approach).

- **Partial CSI:** the transmitter has only partial knowledge of the channel, which can be obtained, for example, through a low rate feedback link or through the channel reciprocity principle in TDD systems that do not have appropriate calibration. An example of this kind of uncertainty is the case where the transmitter has only access to the magnitude of the channel matrix coefficients and has complete uncertainty about the complex phase of the entries of the channel matrix. This model is more carefully analyzed and discussed in chapter 4.

Note that all the above descriptions assume that the receiver has full and perfect access to the CSI. Obviously, as far as the practical implementation of real systems is concerned, this assumption is somewhat too optimistic in the sense that the CSI that is actually available at the receiver is imperfect too. In chapter 6 a multi-user robust design is presented, where both the transmitter and the receiver are considered to have imperfect CSI.

2.3 Capacity of MIMO channels

The capacity of a channel is a fundamental limit in the maximum rates at which information can be transmitted with arbitrarily low probability of error. The channel capacity theorem by Shannon dates back from 1948 [Sha48], and it is the central and most famous success of information theory.

In the following sections, we review the fundamental communication limits of MIMO channels for different degrees of CSI.

2.3.1 Capacity results for different degrees of CSI

Channel capacity, in the Shannon theoretic sense, was originally defined for time-invariant channels as the maximum mutual information between the transmitted and the received signals [Sha48, Sha49]. For wireless communications, however, the channel is time-varying and, in that case, channel capacity admits multiple definitions, depending on the degree of knowledge about the channel state (or its distribution) that is available at the transmit and receive ends. In addition, capacity can be measured by averaging over all possible channel states or by maintaining a fixed minimum rate. See the excellent tutorials [Gol97, Big98] and [Gol03] on capacity of SISO and MIMO fading channels, respectively.

If the channel varies continuously (or for each use of the channel in discrete time models) the capacity is usually referred to as *ergodic capacity*, which is the natural extension of the Shannon capacity for time-varying channels. The ergodic formulation also holds for block-constant fading channel models, where it is assumed that a new realization of the channel is drawn independently

every T symbols and then remains constant for the duration of T symbols. The ergodic formulation is valid in this case because channel coding can be performed over multiple independent fading intervals.

As opposed to this approach, if the unknown channel variations are slow compared with the codeword length, such that the channel can be considered fixed but unknown during the transmission, then either the *compound capacity* or the *outage capacity* formulations are more relevant [Csi81, Oza94]. The compound formulation assumes that the channel is a deterministic quantity that belongs to an uncertainty region, and copes with the problem of finding the maximum rates that can be achieved for the worst channel inside the region. The outage formulation assumes that the channel is an unknown and fixed random variable and it accounts for the fact that, for any given transmission rate, there is a non-zero probability that the channel cannot support it.

The following two subsections give some results on capacity concerning different degrees of knowledge of the CSI at both the transmitter and the receiver.

Capacity results in the non-coherent case

In a fixed wireless environment, it is valid to assume that the fading coefficients vary slowly, in such a way that the transmitter can periodically send pilot signals to allow the receiver to estimate the coefficients accurately. In mobile environments, however, the fading coefficients can change quite rapidly and the estimation of channel parameters becomes much more difficult. The case where neither the transmitter nor the receiver can track the variations of the wireless channel yields the commonly called non-coherent channel models.

Non-coherent channel models arise whenever it is assumed that the actual realization of the fading is not known by the transmitter nor the receiver, however, both of them know the law of the fading channel process. In [Mar99], Marzetta and Hochwald investigated the capacity of the block model for the non-coherent MIMO case. They proved that there is no capacity gain if the number of transmit antennas n_T is bigger than the coherence time T with respect to the case where the number of antennas is equal to the coherence time. In addition, they showed that capacity can be achieved by successively transmitting $T \times n_T$ signal matrices with a certain structure. Finally, for a fixed number of transmit antennas, they found that if the length of the coherence time is allowed to grow without bound, $T \gg n_T$, the capacity approaches the capacity obtained in the coherent case, where the receiver has knowledge of the channel state.

Later, Zheng and Tse [Zhe02] further investigated the capacity for non-coherent MIMO communication. They found the asymptotic capacity of this channel at high SNR in terms of the coherence time T, the number of transmit antennas n_T, and the number of receive antennas

n_R. Precisely, they stated that capacity scales as

$$C = M^* \left(1 - \frac{M^*}{T}\right) \log \mathsf{snr} + \mathsf{o}(1), \quad M^* = \min\left\{n_T, n_R, \left\lfloor \frac{T}{2} \right\rfloor\right\}, \tag{2.10}$$

as SNR tends to infinity. Notice that, as a particular case for SISO sytems, capacity scales as $(T-1)/T$. The model of block-fading for the wireless channel was generalized in [Lia04] allowing for variations inside each fading block. The authors of [Lia04] considered that the correlation matrix that parameterizes the variations of the fading coefficients within a block has rank Q. Consequently, for $Q = 1$ the results of the block model are recovered. For SISO channels it was found that, in the limit of high SNR, capacity scales as

$$C = \left(1 - \frac{Q}{T}\right) \log \mathsf{snr} + \mathsf{o}(1). \tag{2.11}$$

For the MIMO case no variations with respect to the result in (2.10) were found, even in the presence of correlation across blocks in a stationary and ergodic manner.

The only studies that address the non-coherent case with no simplifications on the structure of the fading process are by Moser and Lapidoth [Mos04, Lap03]. There, the authors proved that if the fading process is regular in the sense that have a finite differential entropy rate, then capacity grows double-logarithmically in the SNR as

$$C \propto \log \log \mathsf{snr} + \mathsf{o}(1). \tag{2.12}$$

In addition, in [Lap04], it was demonstrated that, even in the presence of a feedback channel, the multiplicative constant in (2.12) could not be increased. Comparing (2.10) or (2.11) with (2.12) one can readily see that the high-SNR behavior of channel capacity can depend critically on the model, and that simplifications of the model may lead to completely different asymptotic behaviors. Some generalizations of the above results for multiple antenna channels can be found in [Koc04], and for multiple antenna channels with memory in [Lap02].

Capacity results in the coherent case

In a communications environment, where the receiver is (by any means[2]) able to estimate the channel coefficients accurately, the kind of communication that can be established is usually referred to as *coherent* communication. In this particular case of communication, the channel capacity depends on the CSI that is made available at the transmitter side. The ergodic capacity of MIMO communications system over a frequency nonselective fading channel with perfect CSI at the receiver has been a topic that has received a lot of attention in the recent literature as it is detailed in the following.

---

[2]These may include tracking of the pilots sent by the receiver, or blind channel estimation

In his popular work [Tel99], Telatar stated that, when the transmitter has no CSI, uniform power allocation is the most reasonable approach to achieve capacity. The optimality of the uniform power allocation scheme was later proven in [Pal03c], where the authors obtained a robust solution under channel uncertainty by formulating the problem within a game-theoretic framework. The payoff function of the game, which was mathematically formulated as a maximin problem, was the mutual information and the players were the transmitter and a malicious nature. The uniform power allocation is obtained as a robust solution (under a mild isotropy condition).

For the case where the transmitter has perfect CSI, it is also well known [Ral98, Tel99] that the capacity achieving transmit strategy consists in splitting the signal among a set of beamformers giving each one a fraction of transmission power according to a "waterfilling" algorithm. The capacity of wireless communication architectures impaired by both noise and cochannel interference was studied in [Loz02b]. There, the authors found an expression for the capacity in the limit of a large number of antennas using RMT [Wis28, Wig55, Wig58, Wig59, Mar67]. This asymptotic solution depends only on the ratio of the number of transmit antennas over the number of receive antennas, the SNR, and the SIR. One of the main conclusions in [Loz02b] is that antenna diversity can substitute for time and/or frequency diversity at providing ergodicity, even when the number of antennas is very small.

The ergodic capacity of multiple antenna systems with partial CSI at the transmitter side has also been the focus of investigation of numerous researchers. For example, in [Nar98], some initial results were given considering different degrees of quality and quantity of the feedback information. Precisely, the authors determined when the transmission strategy should use some form of beamforming and when it should not (or, equivalently, when the transmit covariance matrix is rank one or not). In addition they showed that, when properly chosen, even a small amount of side information can be quite valuable. In [Vis01], an information-theoretic perspective on optimum transmitter strategies for systems with transmit antenna arrays and imperfect channel feedback was provided. In their paper, two extreme cases were considered: mean feedback (the channel side information resides in the mean of the distribution), and covariance feedback (the channel is assumed to be varying too rapidly to track its mean, so that the mean is set to zero and the side information resides in the covariance of the distribution). For both feedback models, the authors proved that capacity can be achieved by using a Gaussian codebook with a particular covariance matrix, and that when there is a moderate disparity between the strengths of different paths from the transmitter to the receiver, it is nearly optimal to employ the simple beamforming strategy of transmitting all available power in the strongest direction indicated by the feedback. Similar results were given in [Jaf01a], where the authors proved that, in the case of channel covariance feedback, capacity can be achieved by transmitting a Gaussian signal whose covariance matrix has the same eigenvectors as the true channel covariance matrix. The

structure of the corresponding eigenvalues is further investigated in [Jor03]. The results above were further generalized in [Mou02] and [Sim02], taking into account different mutual information formulations such as the ergodic and outage. A unification of all these previous results as well as new insights and results were provided in [Tul05].

In [Sko03], the effect of partial side information at the transmitter is analyzed for a generic MIMO case. One of the most important conclusions of their paper is that capacity can be achieved by a structure that first maps source symbols into space-time codewords independently from the CSI, and then weights these codewords as a function of the CSI at the transmitter, similarly as depicted in Figure 2.3.

It is also important to mention that, when communicating through a MIMO channel, effects of possible correlation in the channel fading coefficients must be taken into account. In [Mes03], Mestre evaluated the asymptotic uniform power allocation capacity of frequency nonselective MIMO channels with fading correlation at either the transmitter or the receiver, concluding that the effects of correlated fading is more harmful at the side with less number of antennas. For a detailed analysis on effects of correlation see also [Kie05].

2.3.2 Capacity formulation for the linear vector Gaussian channel

When a single-user is communicating through a linear MIMO channel corrupted with additive Gaussian noise, the optimum signaling to achieve rates arbitrarily close to the channel capacity is well known as reported in [Bra74, Ral98, Cov91, Tel99].

In the most general setup, the capacity of a channel is the maximum mutual information between the transmitted and the received signals $I(\mathbf{x}; \mathbf{y})$ over all possible input distributions satisfying the power constraint [Bla87, Cov91]:

$$
\begin{aligned}
C = \max_{p(\mathbf{x})} \quad & I(\mathbf{x}; \mathbf{y}), \\
\text{s.\,t.} \quad & \int \mathbf{x}^H \mathbf{x} p(\mathbf{x}) \, \mathrm{d}\mathbf{x} \leq P_T.
\end{aligned}
\tag{2.13}
$$

For the vector Gaussian channel under consideration in this section, it is well known that the maximum mutual information is achieved with a Gaussian input $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$, whose pdf is given by

$$
p(\mathbf{x}) = \frac{1}{\pi^{n_T} |\mathbf{Q}|} \exp\{-\mathbf{x}^H \mathbf{Q}^{-1} \mathbf{x}\}.
\tag{2.14}
$$

This implies that a Gaussian code is used for transmission, where $\mathbf{Q}$ is the covariance matrix of the transmitted vector $\mathbf{x}$, which is yet to be determined. For this particular choice of transmission code the explicit expression of the mutual information $I(\mathbf{x}; \mathbf{y})$ becomes [Cov91, Tel99]

$$
I(\mathbf{x}; \mathbf{y}) = \log \det \left(\mathbf{I} + \mathbf{R}_n^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right).
\tag{2.15}
$$

For notation convenience, this mutual information expression shall be denoted by $\Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U}))$ to explicit its dependence on the transmit covariance matrix $\mathbf{Q}$ and the channel matrix $\mathbf{H}$ through its known $\mathcal{K}$ and unknown $\mathcal{U}$ parameters. Thus, we define

$$\Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U})) = \log \det \left(\mathbf{I} + \mathbf{R}_n^{-1} \mathbf{H}(\mathcal{K}, \mathcal{U}) \mathbf{Q} \mathbf{H}(\mathcal{K}, \mathcal{U})^H \right). \tag{2.16}$$

In the following subsection we particularize the expression of the capacity depending on the characteristics and the model that we utilize to describe the channel.

Ergodic mutual information and capacity

The ergodic mutual information is defined as the expectation over the channel state uncertainty of the instantaneous mutual information expression in (2.16) [Gol03]. The ergodic mutual information is thus given by

$$I_E(\mathcal{K}, \mathbf{Q}) = \mathbb{E}_{\mathcal{U}} \Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U})). \tag{2.17}$$

The ergodic mutual information is a meaningful measure of the achievable rates in situations where the set of known parameters $\mathcal{K}$ remains constant during the transmission of the message, while the channel unknown parameters $\mathcal{U}$ vary sufficiently fast so that its long-term ergodic properties are revealed.

The ergodic capacity, is then defined as the supremum of the ergodic mutual information with respect to the set of possible covariance matrices, subject to a mean transmitted power constraint as

$$\begin{aligned} C_E(\mathcal{K}) = \sup_{\mathbf{Q}} \quad & I_E(\mathcal{K}, \mathbf{Q}), \\ \text{s.t.} \quad & \operatorname{Tr} \mathbf{Q} \leq P_T, \\ & \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{2.18}$$

Note that we have made explicit the dependence of the capacity on the actual state of the known parameters of the channel as given by $\mathcal{K}$.

Compound mutual information and capacity

The compound mutual information is defined [Csi81] as the infimum of the mutual information expression in (2.16) with respect to the channel state uncertainty. Formally, it particularizes to

$$I_C(\mathcal{K}, \mathbf{Q}) = \inf_{\mathcal{U}} \Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U})). \tag{2.19}$$

Noticeworthy, the compound mutual information does not depend on the statistical properties of the uncertainty in the channel, because it only considers the worst-case scenario.

The compound mutual information is a measure of the worst-case achievable rates in situations where no significant channel variability may occur during the transmission of the message and the transmitter is only informed of (or is only able to estimate accurately) the parameters of the channel as given by $\mathcal{K}$. This may be the case of a static communication between the transmitter and the receiver, or when communicating in a slow fading environment.

Similarly as in the previous case, the compound capacity, is then defined as the supremum of the compound mutual information with respect to the set of possible covariance matrices, subject to a mean transmitted power constraint as

$$\begin{aligned} C_C(\mathcal{K}) = \sup_{\mathbf{Q}} \quad & I_C(\mathcal{K}, \mathbf{Q}), \\ \text{s.t.} \quad & \operatorname{Tr} \mathbf{Q} \leq P_T, \\ & \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{2.20}$$

Outage mutual information and capacity

The outage mutual information [Oza94] is the maximum rate R such that the probability that the mutual information in (2.16) is below R is less or equal than ϵ as expressed in (2.21).

$$I_O^\epsilon(\mathcal{K}, \mathbf{Q}) = \sup_R \{ R | \Pr\{\Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U})) \leq R\} \leq \epsilon \}, \tag{2.21}$$

where the probability is with respect to the distribution of the unknown parameters $\mathcal{U}$. The outage mutual information can be technologically relevant in either static scenarios, or when the channel fluctuations are slow enough so that the channel can be considered fixed[3] during the transmission of the message. From its expression in (2.21), we can see that the outage mutual information can be directly related to the achievable rate that can be guaranteed with a certain probability, which is given by $1 - \epsilon$.

The outage capacity problem can be stated from two different formulations, which are essentially equivalent. The first approach is to find the optimal covariance matrix that maximizes the rate given an outage probability. The second approach is to obtain the minimum outage probability given a fixed target rate. Considering the former point of view, the formulation for the problem of finding the outage capacity becomes

$$\begin{aligned} C_O^\epsilon(\mathcal{K}) = \sup_{\mathbf{Q}} \quad & I_O^\epsilon(\mathcal{K}, \mathbf{Q}), \\ \text{s.t.} \quad & \operatorname{Tr} \mathbf{Q} \leq P_T, \\ & \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{2.22}$$

---

[3]The channel is considered fixed, but the transmitter has only partial (or no) knowledge of the channel state matrix.

Considering the latter point of view, for a target rate C_{out} we obtain the following problem formulation for the minimum outage probability $\epsilon_{\min}$,

$$
\begin{aligned}
\epsilon_{\min}(\mathcal{K}, C_{\text{out}}) = \inf_{\mathbf{Q}} \quad & \Pr\left\{\Psi(\mathbf{Q}, \mathbf{H}(\mathcal{K}, \mathcal{U})) \leq C_{\text{out}}\right\} \\
\text{s.\,t.} \quad & \text{Tr}\,\mathbf{Q} \leq P_T, \\
& \mathbf{Q} \succeq \mathbf{0}.
\end{aligned} \tag{2.23}
$$

2.4   Practical communication schemes through multi-antenna channels

Although the theoretical study of the capacity limits in multi-antenna systems is of paramount importance, in the vast majority of cases the system designer has practical constraints (mainly time delay and complexity) which turn the capacity limits into an unreachable objective. Rather, the system designer will try to exploit some other benefits that are provided when communicating through multi-antenna channels.

In this sense, one of the primary concerns of MIMO communication systems designers is, as a first step, the definition of a practical objective function to measure the global performance of the system, and secondly the design of the system accordingly to that specific function definition.

In the following subsections, some of the most important practical communication schemes for multi-antenna channels are reviewed depending on the degree of CSI that is available at both communication ends. This revision is not aimed at being exhaustive but rather at giving a general overview. For the sake of completeness, once this overview is presented a brief section will characterize some of the most utilized practical performance metrics.

2.4.1   Practical schemes with no CSI

When no CSI is available at the transmitter side, there exist two main philosophies to design communication architectures. On one hand there exist techniques whose objective is the increase in the transmission rate by exploiting the multiplexing gain described in §2.1.1. On the other hand some transmission architectures are designed to take advantage of the diversity gain provided by multi-antenna channels.

As far as the techniques that are focused on maximizing the multiplexing gain, the BLAST family [Fos96] lies among the most cited and celebrated ones. The most popular schemes among the BLAST family are the Vertical-BLAST, or V-BLAST [Loz02a], and Diagonal-BLAST (D-BLAST). The V-BLAST technique consists in transmitting different data streams through each antenna. Each data stream is independently coded, which implies that the transmitter can send multiple data streams so that the final rate is increased. Since each antenna is transmitting its

own data stream the V-BLAST system can be seen as a multiple-access channel (MAC), and consequently the receiver is very similar to a multi-user receiver based on successive decoding and cancelation. The only conceptual difference between the V-BLAST and the D-BLAST schemes is that in the latter case the independently coded data streams are transmitted through a different antenna each time the channel is accessed creating the illusion that each data stream is transmitted through a time varying channel.

Concerning the techniques that enable a more reliable communication through the multi-antenna channel without any knowledge of the CSI, one of the first works was [Ses93], where a delay diversity technique was presented in a multiple transmit antenna system. This simple technique essentially consists in transmitting the same information stream through all the antennas, but applying a different delay at each antenna in such a way that different realizations of the same data are acquired at the receiver side. This technique was subsequently generalized leading to the so-called space-time codes (STC).

In [Tar98], the rank and determinant criteria for the optimum design of STC for a MIMO channel where the receiver performs ML detection were derived and also the space-time trellis codes (STTC) were proposed according to these criteria, and whose performance was excellent both in terms of rate and reliability. However, one of the main drawbacks of STTC is that the transceiver complexity increases dramatically as the number of dimensions of the MIMO channel grows.

To overcome this complexity problem, the research community proposed the utilization of space-time block codes (STBC) which were decoded with ML performance assuming that perfect CSI was available at the receiver. The first contribution in this field was given by Alamouti in [Ala98], where a rate one orthogonal STBC for two transmit antennas was proposed. The main interest of orthogonal STBC is that the ML detector can be decoupled into a set of parallel ML detectors with extremely reduced complexity. The pioneering work by Alamouti was further generalized to any number of transmit antennas within the framework of the Hurwitz-Radon family of matrices in [Tar99a] or under the framework of amicable designs in [Gan01, Gan02, Sto02]. In these works it was shown that rate one orthogonal STBC existed only for the case of two transmitting antennas. For a generic number of antennas either the rate one or the orthogonality properties should be dropped to be able to obtain a suitable STBC design. If the orthogonality principle is dropped it is possible to obtain rate one STBC designs as shown in [Jaf01b]. However, as a consequence the simplicity of the decoupled ML detector cannot be achieved. The study of STC assuming the presence of errors in the CSI acquired a the receiver was conducted in [Tar99b].

In the case where even the receiver has no access to the CSI, it was shown in [Mar99] that, in the high SNR regime, the optimum transmitted signal is such that the data streams transmitted through different antennas are mutually orthogonal, giving rise to unitary codes. In [Hoc00a],

some families of unitary code designs are presented, and in [Hoc00b], a systematic design of unitary codes based on Fourier and algebraic models is provided. Moreover, in [Hoc00c], the extension of differential modulation schemes for the multi-antenna case was presented, obtaining the so called differential unitary space-time designs, which do not require CSI neither at the transmitter nor at the receiver.

2.4.2   Practical schemes with perfect CSI

The case where both the transmitter and the receiver have full access to the CSI is a topic that has been extensively studied by the research community. There exists a vast literature on the subject depending on which was the metric utilized to compute the performance of the system and which were the designs constraints such as the kind of transmitter and receiver, which may be linear or nonlinear.

Initially, in works such as [Sal85, Yan94] the transmitter and receiver of a MIMO system were designed based on the presence of linear time-invariant filters both at the transmitter and the receiver whose objective was the minimization of the MSE. In these works, it was shown that the optimal filter design diagonalizes the channel, and, consequently, the data streams are transmitted through the eigenmodes of the channel.

In [Ral98], instead of the minimization of the MSE, it was considered the maximization of the mutual information defined as in [Cov91] for a frequency selective MIMO channel, and two techniques were described, namely spatio-temporal vector-coding and discrete matrix multitone. In both cases the optimal design also diagonalizes the MIMO channel matrix.

Another important performance measure is the BER of the system. This case was considered, for example, in [Din02], where it was shown that the optimal transmitter, in general, does not diagonalize the MIMO channel matrix.

Later, the case of linear transmitters and receivers was solved in [Sca02] for most of the already known performance measures, and has been further generalized, under the framework of convex optimization, in [Pal03b]. Moreover, in [Pal03a, Pal03b], Palomar showed that most of the typical objective functions, such as the minimization of the trace of the MSE matrix or the maximization of the mutual information, fall into two categories extracted from majorization theory: Schur-convex and Schur-concave functions, [Mar79]. In this common case, the design of any MIMO communications system can be framed, in a unified way, in the powerful of theory of convex optimization and consequently, the design problem can be solved very efficiently. Particularly, he demonstrated that for Schur-concave functions (mutual information and weighted arithmetic mean of the MSE among others) the optimal transmit and receive linear filters are such that the system is fully diagonalized, resulting in parallel substreams. In addition, for

Schur-convex objective functions (such as the maximization of the mean BER[4], or the maximization of the harmonic mean of the SNR), he showed that the system is diagonalized up to a rotation transformation at the transmitter side, which uniformly distributes the transmitted symbols among the different eigenmodes.

It has also been shown that the obtained performances can be improved when nonlinear stages are permitted or when the complexity of the receiver is increased, taking as the objective the minimization of the BER (see *e.g.* the application of the nonlinear receiver MMSE-VBLAST in [Jia05]).

Although the case of having a perfect CSI at both sides of the system has been studied deeply, the only work (prior to this one) in which the transmitter is designed assuming that a perfect ML receiver is employed is [Col04]. The utilization of a ML receiver is motivated by recent efficient implementations such as the optimal sphere decoder [Has03, Dam03], or the schemes based on semidefinite relaxation (SDR) [Ma02, Jal05, Wie05b], which reduce the computational complexity by sacrificing optimality.

2.4.3 Practical schemes with imperfect and/or incomplete CSI

In the two previous subsections, the transmission strategies in two extreme cases have been described. In a practical scenario it may be too pessimistic that the transmitter is totally unaware of the CSI and assuming that both the transmitter and the receiver have full access to the CSI may not be too realistic in a practical deployment. Consequently, usually only imperfect and/or incomplete CSI is available (specially at the transmitter side, since, in some scenarios, the receiver can accurately track the channel variations).

As far as the imperfect CSI case is concerned, one of the first works to appear was [Nar98], where a multi-antenna transmitter was considered, which was informed either with a version of the channel corrupted with Gaussian noise or with a quantized version of the channel. For multiple performance criteria (such as the mean SNR or the mutual information) the optimal transmitter consisted in transmitting different data streams through the eigenmodes of the channel with a suitable power allocation among them.

A parallel line of work was developed in [Jön02] where a particular structure for the transmitter was imposed consisting of a concatenation of an orthogonal STBC and a linear transformation, which was designed to maximize the mean SNR according to the available CSI following the Bayesian philosophy. The authors proved that when there is full access to the CSI the solution tends to the structure of a simple beamformer, whereas for the case of no CSI the linear transformation becomes the identity and thus the transmitter is equivalent to performing only the STBC.

---

[4]Assuming that the same constellation is utilized for each of the substreams.

Another design possibility consists in taking the worst-case or maximin approach [Kas85]. In this case, the transmitter is aware of an uncertainty region where the actual channel belongs to, and is aimed at maximizing some performance criterion for the worst possible channel inside the uncertainty region. Precisely, in [Vor03] a robust receive beamformer is designed taking into account that the parameters that describe the actual spatial signature of the incoming signal belong to an uncertainty region. Moreover in [Pal04], the optimal linear transmitter and receiver were designed under the maximin criterion so that the total transmitted power was minimized and some QoS constraints were satisfied for each data substream. Finally, in [PI06] the authors considered a similar structure to the one described in [Jön02], *i.e.*, a concatenation of a STBC plus a fixed linear transformation. In this case, the authors designed the optimal power allocation among the different outputs of the STBC according to the maximin philosophy and under the framework of convex optimization.

Much less literature exists concerning the transmission schemes with incomplete CSI. In [Muk01], for example, it was considered the case where, for a single receiving antenna, the multi-antenna transmitter was informed with either the phases of the channel response or the magnitude. In case of phase information it was shown that the optimal transmitter consisted of a phase-former with the phases matched to those of the actual channel. For the case of magnitude knowledge the best strategy consisted in transmitting only through the antenna with highest gain. Another case of incomplete CSI was studied in [Cho02d], where the author considered the case where the transmitter performs single beamforming of the outgoing data. In the scheme developed by Choi, the transmitter only has partial knowledge of the eigenvector associated with the highest eigenvalue. In this case, transmission beamformer was acquired differentially, as step-by-step updates, from the receiver through a low rate feedback link, at a rate that was related to the Doppler frequency of the channel.

2.4.4   Practical transmission schemes for the multi-user case

For the case of a multi-user scenario, one of the best known and most widely studied transmit architecture is the extension of the non-linear precoding scheme proposed by Tomlinson [Tom71] and Harashima [Miy72, Har72], for temporal intersymbol interference mitigation, into spatial interference equalization for MIMO systems [Fis02a], see Figure 2.4. In their work, they showed that STHP offered good power efficiency and low decoding delays.

The main idea behind STHP is to move the decision feedback equalizer (DFE) structure which is typically placed at the receiver end (*e.g.*, in BLAST-like architectures [Fos96]) to the transmitter side. The most important advantage of this movement of the equalizer is that, as the information symbols are obviously known to the transmitter, it overcomes the frequent problem of error propagation of DFE systems, which occur when wrong symbol decisions are used to
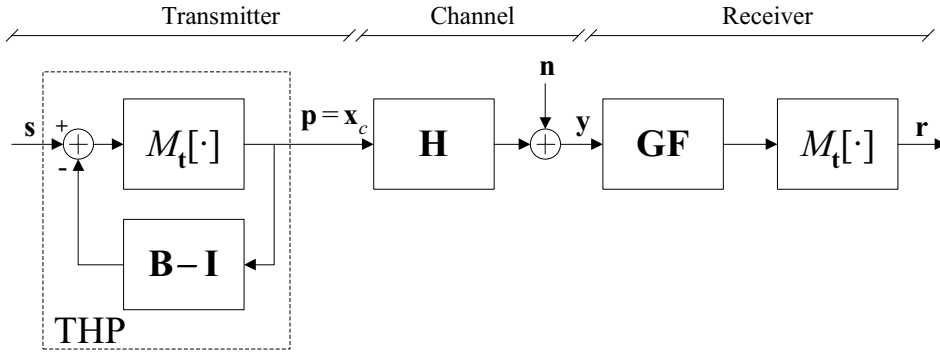
Figure 2.4: Transceiver scheme for STHP.

detect other received symbols by canceling its interference.

However, as many precoding schemes, the main limitation of the STHP structure is that perfect CSI has to be made available at the transmitter side so that the precoder can be matched to the channel. When the CSI at the transmitter is imperfect the system suffers from performance degradation. The effects of imperfect CSI for STHP, in terms of SNR loss, were studied in [Fis02b]. The extension of the SHTP scheme to the multi-user case was done in [Win04a].

In addition to STHP another practical transmission scheme was developed by Peel, Hochwald, and Swindlehurst [Pee05a, Pee05b] motivated in this case by recent theoretical results [Cai03, Yu04, Vis03b, Vis03a, Wei04, Wei06] describing the sum-capacity when using multiple antennas to communicate with multiple users in a known rich scattering environment. The authors introduced an encoding algorithm which is variation on channel inversion that regularizes the inverse and uses a "sphere encoder" to perturb the data to reduce the power of the transmitted signal. Firstly, they showed that while the sum-capacity grows linearly with the minimum of the number of antennas and users, the sum-rate of channel inversion does not. They precised that this poor performance is due to the large spread in the singular values of the channel matrix. As a modification of channel inversion the authors introduced regularization to improve the condition of the inverse and maximize the signal to interference-plus-noise ratio at the receivers as

$$\mathbf{x} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha I \right)^{-1} \mathbf{s}, \tag{2.24}$$

where α is the regularization constant, and $\mathbf{s}$ represents the data vector before precoding. However, regularization was found to be not enough to achieve near-capacity performance at all SNRs. After the regularization of the channel inverse, a certain perturbation of the data using a "sphere encoder" was performed to further reduce the energy of the transmitted signal. The authors claimed that the performance difference with and without this perturbation is shown to be dramatic. In [Win04b], Windpassinger *et al.* presented a reduced complexity scheme similar

to the one presented in [Pee05a, Pee05b] based on lattice-reduction.

2.5 Mathematical characterization of alternative performance measures

As commented in the previous section, despite the interest of the theoretical study of the capacity limits in multi-antenna systems, the system designer has different constraints which turn the capacity limits into an unreachable objective in a practical approach. Rather, one of the primary concerns of MIMO communication systems designers is the definition of a practical objective function to measure the global performance of the system.

In the following some of the most widely utilized performance measures are reviewed.

2.5.1 Minimum distance of the received codewords

Consider the linear MIMO communication system described in §2.1.4 with Gaussian noise. Assume that the set of possible transmitted codewords is given by $\mathbf{x}_m$ with $m = 1, \ldots, |\Omega|$. Then the m-th received codeword is defined by $\mathbf{y}_m = \mathbf{H}\mathbf{x}_m$. The squared distance among the received codewords $\mathbf{y}_m$ and $\mathbf{y}_{m'}$ is then defined as

$$d^2(\mathbf{y}_m, \mathbf{y}_{m'}) \triangleq \|\mathbf{y}_m - \mathbf{y}_{m'}\|_{\mathbf{R}_n^{-1}}^2 = (\mathbf{y}_m - \mathbf{y}_{m'})^H \mathbf{R}_n^{-1} (\mathbf{y}_m - \mathbf{y}_{m'}). \tag{2.25}$$

Once the squared distance has been properly defined, it is straightforward to define the minimum distance as the minimum distance among all the possible pairs of received codewords as

$$d_{\min} = \min_{m, m'} d(\mathbf{y}_m, \mathbf{y}_{m'}). \tag{2.26}$$

As it suggested by its name, the higher the minimum distance among the received codewords the better the performance of the system, as the immunity against Gaussian noise is increased.

The importance of the minimum distance as a performance metric stems from the utilization at the receiver side of a ML detector. The ML detector selects as the best choice for the estimation of the transmitted codeword $\hat{\mathbf{x}}_m$, the one that maximizes the likelihood function $\mathcal{L}(\mathbf{y}, \mathbf{x}_m)$, which, assuming that all the codewords are equally probable, is given by the channel transition probability $p(\mathbf{y}|\mathbf{x}_m)$,

$$\hat{\mathbf{x}}_m = \arg \max_{\mathbf{x}_m} \mathcal{L}(\mathbf{y}, \mathbf{x}_m) = \arg \max_{\mathbf{x}_m} p(\mathbf{y}|\mathbf{x}_m). \tag{2.27}$$

In communication systems corrupted with AWGN the channel transition probability is the one described in (2.7) and consequently maximizing the likelihood function particularizes to mini-

mizing the quadratic form

$$\hat{\mathbf{x}}_m = \arg \max_{\mathbf{x}_m} p(\mathbf{y}|\mathbf{x}_m) = \arg \min_{\mathbf{x}_m} (\mathbf{y} - \mathbf{H}\mathbf{x}_m)^H \mathbf{R}_n^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_m). \tag{2.28}$$

Within this framework, the pairwise error probability (PEP) of the codewords $\mathbf{x}_{m'}$ and $\mathbf{x}_m$ is defined as the probability that the codeword $\mathbf{x}_{m'}$ is detected when $\mathbf{x}_m$ was actually transmitted. This erroneous detection occurs whenever the likelihood of $\mathbf{x}_{m'}$ is higher than that of $\mathbf{x}_m$, conditioned on the fact that $\mathbf{x}_m$ was actually transmitted, *i.e.*, $\mathbf{y} = \mathbf{H}\mathbf{x}_m + \mathbf{n}$.

Mathematically, the probability of such an event can be described as

$$\mathsf{pep}_{m,m'} = \Pr \left\{ \mathcal{L}(\mathbf{H}\mathbf{x}_m + \mathbf{n}, \mathbf{x}_{m'}) > \mathcal{L}(\mathbf{H}\mathbf{x}_m + \mathbf{n}, \mathbf{x}_m) \right\} \tag{2.29}$$

$$= \Pr \left\{ p(\mathbf{H}\mathbf{x}_m + \mathbf{n}|\mathbf{x}_{m'}) > p(\mathbf{H}\mathbf{x}_m + \mathbf{n}|\mathbf{x}_m) \right\} \tag{2.30}$$

$$= \Pr \left\{ (\mathbf{H}(\mathbf{x}_m - \mathbf{x}_{m'}) + \mathbf{n})^H \mathbf{R}_n^{-1} (\mathbf{H}(\mathbf{x}_m - \mathbf{x}_{m'}) + \mathbf{n}) < \mathbf{n}^H \mathbf{R}_n^{-1} \mathbf{n} \right\}, \tag{2.31}$$

where the definition of $p(\mathbf{y}|\mathbf{x})$ in (2.7) has been utilized in last equation.

With some straightforward manipulations from the last expression, the PEP simplifies to

$$\mathsf{pep}_{m,m'} = \Pr \left\{ (\mathbf{x}_m - \mathbf{x}_{m'})^H \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{x}_m - \mathbf{x}_{m'}) < -2\Re e(\mathbf{x}_m - \mathbf{x}_{m'})^H \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{n} \right\}. \tag{2.32}$$

Noticing that the right hand side of the equation is a real Gaussian random variable with variance equal to $(\mathbf{x}_m - \mathbf{x}_{m'})^H \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{x}_m - \mathbf{x}_{m'})/2$ the PEP can be finally expressed as

$$\mathsf{pep}_{m,m'} = \mathcal{Q} \left(\sqrt{\frac{(\mathbf{x}_m - \mathbf{x}_{m'})^H \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H}(\mathbf{x}_m - \mathbf{x}_{m'})}{2}} \right) \tag{2.33}$$

$$= \mathcal{Q} \left(\sqrt{\frac{(\mathbf{y}_m - \mathbf{y}_{m'})^H \mathbf{R}_n^{-1} (\mathbf{y}_m - \mathbf{y}_{m'})}{2}} \right) \tag{2.34}$$

$$= \mathcal{Q} \left(\sqrt{\frac{d_{m,m'}^2}{2}} \right). \tag{2.35}$$

Note that the well known expression in (2.35) coincides with that given in [Woz65, p.265].

From all the possible combinations of m and m', the most important one is the one that gives the minimum distance as defined in (2.26). As it is shown in [Loz06] the minimum distance characterizes the behavior of any practical communication system, specially at high SNRs.

2.5.2 MSE

The MSE refers to the expected value of the squared error committed by an estimator of a random scalar quantity. We denote by s the random scalar quantity and its estimator, by $\hat{s}$. The MSE is then given by

$$\mathsf{mse} \triangleq \mathbb{E}|s - \hat{s}|^2. \tag{2.36}$$

In the common case where the random quantity $\mathbf{s}$ and the estimator $\hat{\mathbf{s}}$ are vectors, the MSE can properly described by the matrix

$$\mathbf{E} = \mathbb{E}(\mathbf{s} - \hat{\mathbf{s}})(\mathbf{s} - \hat{\mathbf{s}})^H, \tag{2.37}$$

whose k-th diagonal entry represents the MSE committed in the estimation of the k-th element of vector $\mathbf{s}$,

$$[\mathbf{E}]_{kk} = \mathbb{E}|s_k - \hat{s}_k|^2 = \mathsf{mse}_k. \tag{2.38}$$

Obviously, when the MSE refers to the error committed at the receiver when estimating the information that is sent from the transmitter, it is desired that the MSE be as small as possible, since it means that the estimation matches closely the desired information. Hence, any reasonable system has to be designed to have a low MSE.

Recently, a deep entanglement has been unveiled between the MMSE, the key measure in the field of estimation theory, and the mutual information (and thus capacity), which is the representative measure in information theory. Some preliminary intuitions on this relation were developed in [For04], and initial formal results for the scalar case were presented in [Guo05] and were extended to the vector formulation in [Pal06]. The relation between the MMSE and the mutual information is through differentiation with respect to the SNR which is described next.

2.5.3   Signal to noise ratio

The signal to noise ratio (SNR) of a given quantity $y = \alpha d + \beta u$ refers to the expected power quotient of the desired part αd over the undesired part βu, which arises from the noise present in the system. The SNR given by

$$\mathsf{snr} \triangleq \frac{|\alpha|^2 \mathbb{E}|d|^2}{|\beta|^2 \mathbb{E}|u|^2} = \frac{P_d}{P_u}. \tag{2.39}$$

As it is intuitive from its definition, the higher the SNR the better the system. Hence, any reasonable system has to be designed to have a SNR as high as possible. The definition of SNR can be further generalized to include the effects of interferences in the undesired signal. In this case, the ratio in (2.39) is usually termed signal to interference-plus-noise ratio (SINR).

An important issue concerning the SINR is how it is determined which part of the given quantity y is desired and which part is undesired. For example, assume that the model for a received signal is

$$y = \alpha d + \beta u, \tag{2.40}$$

where the quantity α is not accurately known by the receiver and only an estimate $\hat{\alpha}$ is available. Accordingly, the receiver will detect the desired signal as a function of $\hat{\alpha}$. For example, in the

case where the receiver utilizes detection thresholds to estimate the desired signal, *i.e.*,

$$\hat{d} = \frac{y}{\hat{\alpha}} = d + \frac{(\alpha - \hat{\alpha})d + \beta u}{\hat{\alpha}}, \tag{2.41}$$

these thresholds will be set-up as a function of the value of $\hat{\alpha}$. Alternatively, if the receiver performs ML detection, the likelihood function will also be evaluated according to the value of $\hat{\alpha}$, leading to a certain mismatch between the actual likelihood function and the utilized one.

From all that has been said above, in such cases it is more convenient to consider the following decomposition of the quantity y:

$$y = \hat{\alpha}d + (\alpha - \hat{\alpha})d + \beta u, \tag{2.42}$$

where $\hat{\alpha}d$ is the *effective* desired signal and $(\alpha - \hat{\alpha})d + \beta u$ is the *effective* undesired contribution. Note that in the *effective* undesired signal the term $(\alpha - \hat{\alpha})d$ acts as a self-interfering quantity.

Consequently, even though the actual powers of the desired and undesired signals are equal to $P_d = |\alpha|^2 \mathbb{E}|d|^2$ and $|\beta|^2 \mathbb{E}|u|^2$ respectively, and the SINR is given by (2.39), it may be more accurate to compute the *effective* SINR, which is given by

$$\mathsf{esinr} = \frac{|\hat{\alpha}|^2 \mathbb{E}|d|^2}{|\alpha - \hat{\alpha}|^2 \mathbb{E}|d|^2 + |\beta|^2 \mathbb{E}|u|^2}, \tag{2.43}$$

where d and u are assumed independent.

2.5.4 Symbol and bit error rates

The symbol error rate (SER) of a communication system is empirically defined as the quotient between the number of symbols received in error and the total number of received symbols. Similarly, the BER is defined as the fraction of bits in error.

The SER is intimately related with the PEP described in §2.5.1, but it is usually difficult to deal with a precise explicit expression. Consequently, assuming that the interference-plus-noise component is Gaussian distributed, the SER, denoted by P_s^{err}, can be approximated by an analytical expression as a function of the SINR [Pro95]:

$$P_s^{\mathrm{err}} = a_1 \mathcal{Q}(a_2 \mathsf{sinr}), \tag{2.44}$$

where a_1 and a_2 are constants that depend on the signal constellation.

Once the SER has been characterized, the BER can be approximately obtained from the symbol error probability P_s^{err} assuming that the bits have been mapped into the constellation points using a Gray encoding. The expression for the BER, ber, is

$$\mathsf{ber} \simeq \frac{P_s^{\mathrm{err}}}{\log_2 M}, \tag{2.45}$$

where M is the number of points in the constellation utilized for transmission. In [Cho02a] a better approximate expression for the BER than the one in (2.45) is found.

Chapter 3

Single-user communication through MIMO channels with perfect CSI

3.1 Introduction

As it has been described in the previous chapter, communicating through MIMO channels provides a great enhancement in capacity, [Fos96, Fos98], and also in system performance in terms of diversity and multiplexing gains under the tradeoff theoretically described in [Zhe03]. In practice, these potential improvements can be achieved by proper design of transmitter and receiver architectures.

It has been also commented that one main aspect that conditions this design process is the quantity and quality of the CSI available at both communication ends. In this chapter the case where both the transmitter and the receiver have full knowledge of the state of the channel is analyzed. First we recall the expression for the capacity in such case and then proceed to describe practical transmission schemes that try to achieve this limit.

3.2 Capacity results

3.2.1 Theoretical limits

When a single-user is communicating through a MIMO channel where perfect CSI is available at both sides of the link, the notion of instantaneous mutual information as defined in (2.16) becomes a meaningful measure of the maximum rates at which information can be reliably transferred from the transmitter to the receiver.

The assumption of perfect CSI holds true either when the channel remains essentially fixed, such as in DSL systems, or when its variations are sufficiently slow with respect to the duration of the transmission. For this particular case of perfect CSI, capacity is achieved by adapting the

transmitted signal to the specific channel realization $\mathbf{H}$ and to the covariance of the noise $\mathbf{R}_n$.

Formally, with perfect CSI the ergodic, compound, and outage capacity formulations presented in chapter 2 collapse, because the set of unknown parameters is void, $\mathcal{U}^{\mathrm{TX}} = \mathcal{U}^{\mathrm{RX}} = \emptyset$. Consequently, the capacity formulation can be expressed as [Bra74, Cov91, Ral98, Tel99, Sca99]

$$
\begin{aligned}
C = \max_{\mathbf{Q}} \quad & \log\det\left(\mathbf{I} + \mathbf{R}_n^{-1}\mathbf{H}\mathbf{Q}\mathbf{H}^H\right), \\
\mathrm{s.\,t.} \quad & \mathrm{Tr}\,\mathbf{Q} \le P_T, \\
& \mathbf{Q} \succeq 0,
\end{aligned}
\tag{3.1}
$$

where, as indicated in the problem formulation, the maximization with respect to the transmit covariance matrix $\mathbf{Q}$ of the Gaussian distributed transmitted signal is restricted to positive semi-definite matrices such that a mean transmitted power constraint is fulfilled.

The maximization problem in (3.1) has a well-known solution, $\mathbf{Q}^\star$, which is aligned[1] with the whitened channel correlation matrix, $\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}$:

$$
\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H} = \mathbf{U_H}\mathbf{\Lambda_H}\mathbf{U_H}^H \Rightarrow \mathbf{Q}^\star = \mathbf{U_H}\mathbf{\Lambda_Q}\mathbf{U_H}^H.
\tag{3.2}
$$

This implies that the capacity achieving architecture transmits independent data streams through each one of the eigenmodes of the whitened channel correlation matrix. With this optimal structure the capacity problem in (3.1) simplifies to

$$
\begin{aligned}
C = \max_{\{\lambda_{\mathbf{Q},j}\}} \quad & \sum_{j=1}^{n_T} \log(1 + \lambda_{\mathbf{Q},j}\lambda_{\mathbf{H},j}) \\
\mathrm{s.t.} \quad & \sum_{j=1}^{n_T} \lambda_{\mathbf{Q},j} \le P_T \\
& \lambda_{\mathbf{Q},j} \ge 0, \quad 1 \le j \le n_T,
\end{aligned}
\tag{3.3}
$$

where $[\mathbf{\Lambda_H}]_{jj} = \lambda_{\mathbf{H},j}$ and $[\mathbf{\Lambda_Q}]_{jj} = \lambda_{\mathbf{Q},j}$ have been used. The value of $\lambda_{\mathbf{Q},j}$ corresponds to the power that is allocated to the j-th eigenmode. It is now straightforward to prove that the expression for the optimal eigenvalues of the covariance matrix $\mathbf{Q}^\star$ is given by

$$
\lambda_{\mathbf{Q},j}^\star = \left(\mu - \frac{1}{\lambda_{\mathbf{H},j}}\right)^+, \quad 1 \le j \le n_T
\tag{3.4}
$$

where μ is chosen so that the power constraint in (3.3) is fulfilled with equality.

In [Cov91] a beautiful interpretation of the optimal power allocation in (3.4) is given which is also depicted in Figure 3.1. Basically, it consists in considering each one of the eigenmodes as water-porous unit-base vessels which are arranged together. Then each vessel is filled with a solid substance such that the height of this substance is equal to the inverse of the corresponding

---

[1]In this context, we consider that two positive semi-definite matrices are aligned when they commute, or, equivalently, when they share the same eigenvectors.
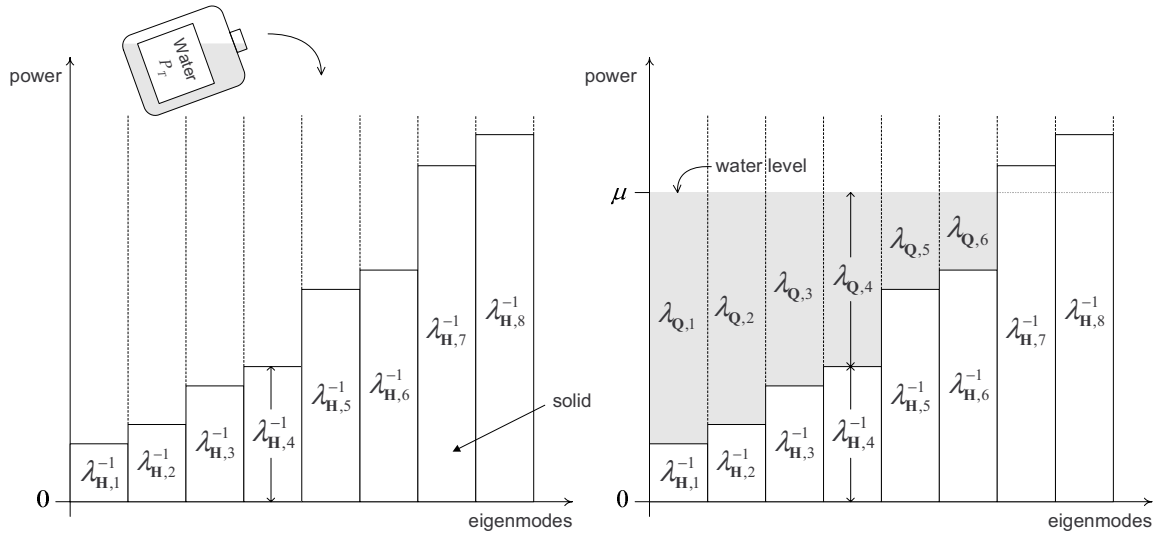
Figure 3.1: Graphical representation of the power allocated to each eigenmode. In this case, no power is allocated to eigenmodes 7 and 8, as their inverse eigenvalues are above the water level.

channel eigenvalues $\lambda_{\mathbf{H},j}$. Next, an amount of water which represents the available power P_T is poured through the vessels. In this interpretation, the value of μ plays the role of the height of the water level, and the value of the optimal power allocation $\{\lambda_{\mathbf{Q},j}^{\star}\}$ is the water height in each of the vessels. If the amount of solid substance of any of the vessels is above the water level no power is assigned to the corresponding eigenmode. This interpretation makes clear the facts that higher power is allocated to eigenmodes with higher eigenvalues, and that no power is given to the eigenmodes whose eigenvalues are below a certain threshold which is determined from the available power P_T and the other channel eigenvalues $\lambda_{\mathbf{H},j}$.

3.2.2 Practical limits: mercury/waterfilling

Although Gaussian inputs are optimum from a mutual information viewpoint, state-of-the-art technology cannot implement them in practice. Rather, technology constrains the inputs to be drawn from discrete constellations, which may differ significantly from the Gaussian idealization.

In [Loz06], the authors found the optimal power allocation that maximizes the input-output mutual information for the case where the MIMO channel consists in a bank of independent parallel channels with AWGN (which implies that $n_T = n_R$ and that $\mathbf{U_H} = \mathbf{R}_n = \mathbf{I}$) and the input distribution to each of these parallel channels is given and fixed. For this case the input-output relation in (2.4) can be scalarized and reduces to

$$y_j = \sqrt{\lambda_{\mathbf{H},j}}x_j + n_j, \quad 1 \leq j \leq n_T, \tag{3.5}$$

where we have assumed that $\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}$ is full rank. Since the authors of [Loz06] only deal with the design of the power allocation, this implies that the transmit covariance matrix is diagonal

and the power allocation coincides with the eigenvalues of the transmit covariance matrix. It is important to mention that for a generic case with a non-diagonal MIMO channel, it is not known whether the optimal covariance matrix diagonalizes the channel or not. Following the model presented in §2.1, the transmitted signal can be expressed as

$$x_j = \sqrt{\lambda_{\mathbf{Q},j}} s_j, \quad 1 \leq j \leq n_T, \tag{3.6}$$

where $\{s_j\}$ are independent symbols drawn from an arbitrary set of constellations. Once the symbol y_j has been received, the MMSE estimate of s_j becomes

$$\hat{s}_j(\lambda) = \mathbb{E}\left\{s_j | \sqrt{\lambda} s_j + n_i\right\}, \tag{3.7}$$

and similarly we can define the MMSE function of j-th parallel channel as

$$\mathsf{mmse}_j(\lambda) = \mathbb{E}\left|s_j - \hat{s}_j(\lambda)\right|^2. \tag{3.8}$$

For this simplified system model, the authors of [Loz06] found that the eigenvalues of the transmit covariance matrix that maximize the mutual information between the input and the output have an analogous expression to that in (3.4) found for the case of utilizing ideal Gaussian codes. In this case the optimal power allocation is given by

$$\lambda_{\mathbf{Q},j}^{\star} = \left(\mu - \frac{G_j\left((\lambda_{\mathbf{H},j}\mu)^{-1}\right)}{\lambda_{\mathbf{H},j}}\right)^+, \tag{3.9}$$

where, just as in the previous case, the value of μ is such that the power constraint is fulfilled with equality and where the only difference with respect to the expression in (3.4) is the factor $G_j\left((\lambda_{\mathbf{H},j}\mu)^{-1}\right)$, which is never below one and takes into account the non-Gaussianity of the codes that are being utilized. The explicit expression for the function G_j is

$$G_j(\zeta) = \begin{cases} 1/\zeta - \mathsf{mmse}_j^{-1}(\zeta), & \zeta \in [0,1] \\ 1, & \zeta > 1 \end{cases}, \tag{3.10}$$

where mmse_j^{-1} is the inverse, with respect to the composition of functions, of the MMSE function described in (3.8).

The expression in (3.9) admits an interpretation which generalizes that of the water-filling presented in the previous section.

1. We represent each parallel channel in (3.5), with a unit-base water-porous mercury-nonporous vessel.

2. Next, we fill each vessel with some solid substance up to a height equal to $\lambda_{\mathbf{H},j}^{-1}$. See Figure 3.2.

3. We then fix the value of μ, and pour mercury onto each of the vessels until its total height (including the solid) reaches $G((\lambda_{\mathbf{H},j}\mu)^{-1})/\lambda_{\mathbf{H},j}$. See Figure 3.3.

4. Finally, we pour a volume of water equal to P_T (the waterlevel reaches a height equal to μ). See Figure 3.4.

Once all these steps have been performed, the water height from the top of the mercury to the water level gives $\lambda_{\mathbf{Q},j}^\star$.
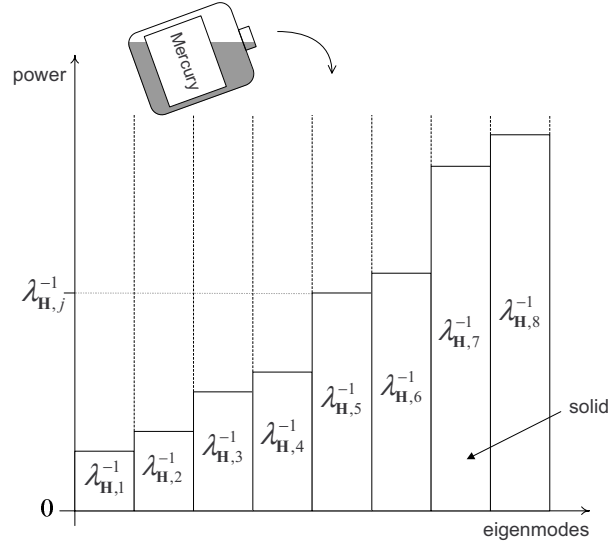
Figure 3.2: Graphical representation of each parallel channel by a unit-base water-porous mercury-nonporous vessel. Each vessel is filled with some solid substance up to a height equal to $\lambda_{\mathbf{H},j}^{-1}$.
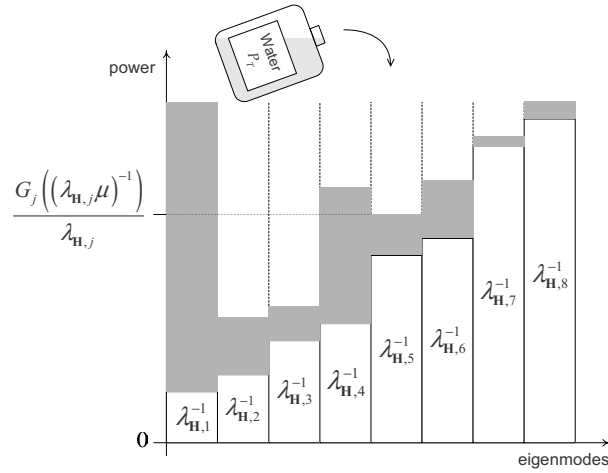
Figure 3.3: Mercury is poured onto each of the vessels until its total height (including the solid) reaches $G((\lambda_{\mathbf{H},j}\mu)^{-1})/\lambda_{\mathbf{H},j}$.

What must be highlighted from the mercury/waterfilling interpretation is that the additional mercury pouring stage regulates the water admitted by each vessel thus tailoring the process
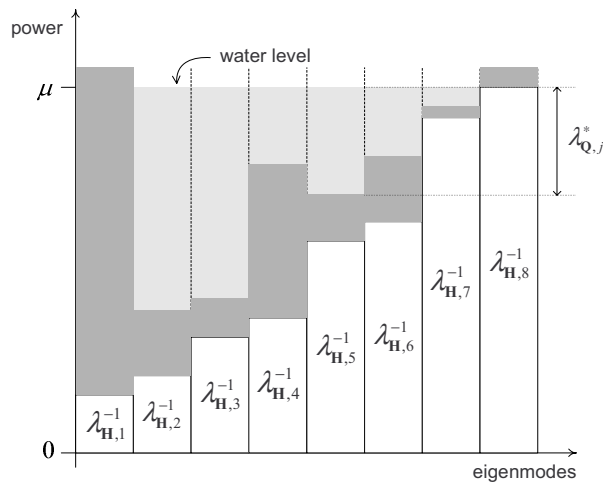
Figure 3.4: A volume of water equal to P_T is poured (the waterlevel reaches a height equal to μ. The water height from the top of the mercury to the water level gives $\lambda^{\star}_{\mathbf{Q},j}$.

to arbitrary input distributions. Pouring mercury onto a vessel amounts to reducing the power that is allocated to that channel by an amount that depends on the deviation of the input distribution from the Gaussian ideal. Note that for Gaussian signaling the function G_j is always equal to one, recovering the original water-filling interpretation.

For example, let us assume that a particular channel has a very high equivalent gain $\lambda_{\mathbf{H},j}$. Consequently, if we perform waterfilling assuming that an ideal Gaussian code is being utilized a lot of power is going to be given to this particular channel to take advantage of its high gain to achieve higher rates. However, if the symbol constellation that is transmitted through this channel deviates significantly from the Gaussian idealization[2], a lot of mercury is going to be poured to its vessel and consequently the power that is actually given to that channel is greatly reduced because otherwise it would be wasted. (See, for example, channels 1 and 4 in Figures from 3.2 to 3.4).

Note also that, similarly to the water-filling solution, some of the channels are not activated (no power is allocated to them) by the mercury/waterfilling solution. In the following section, it will be shown that this same behavior can be present in practical transmission schemes.

3.3   Practical transmission schemes

The notions of capacity and mutual information enable us to obtain a measure of the maximum rates at which information can be reliably transferred from the transmitter to the receiver when utilizing Gaussian codes (waterfilling interpretation) or arbitrary signalling (mer-

---

[2]For example, in the case of supposing that the total available power P_T is high and the chosen constellation has a very low rate.

cury/waterfilling). However, to achieve these limits infinite-length codes should be utilized. Despite their theoretical usefulness, these kind of codes can not be implemented in a practical system, since the codewords have to be necessarily bounded in the time domain (due to, *e.g.*, delay constraints or complexity issues).

Consequently, it is convenient to design the transceiver taking into account that finite-length codewords which belong to discrete constellations are the ones that are going to be actually transmitted. In fact, as described in §2.4.2, practical transceiver designs for the case where the transmitter has full access to the CSI (through the reciprocity principle or a feedback link) have been extensively studied by many researchers. The differences among the works rely on the kind of transmitter and receiver and the chosen performance measure.

In the following, we focus on a communication scheme with the ML detector at the receiver side. This is motivated by recent efficient implementations such as the optimal sphere decoder [Dam03], or the schemes based on semidefinite relaxation [Ma02], which reduce the computational complexity by sacrificing optimality. Some initial results on linear transmitter design for this particular case of receiver structure have been given in [Col04, Pay06b, Pay06c].

Precisely, we give further insight on the problem of designing a linear transmitter for a MIMO system with perfect CSI at both sides, when a ML detector is utilized. According to this choice of receiver, a meaningful performance objective is the minimization of the worst PEP for any possible pair of transmitted symbol vectors. In this sense, the problem is mathematically formulated as a maximin optimization problem, which is, in general, quite complicated to solve due to its non-convex and discrete nature. Despite this fact, we are able to partially characterize its solution. The relation between the utilized performance metric and the mutual information and capacity as described above will be clarified in §3.3.3.

3.3.1 System description

Let us recall the system model described in chapter 2. A narrowband multiplexing system with n_T transmit and n_R receive antennas corrupted with additive Gaussian noise is considered. Let us define $\mathbf{x} \in \mathbb{C}^{n_T}$ as the transmitted signal, where $[\mathbf{x}]_j$ represents the transmitted signal through j-th antenna. We also define $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ as the channel matrix, where $[\mathbf{H}]_{ij}$ represents the baseband equivalent path gain from the j-th transmitter to the i-th receiver. Finally, $\mathbf{n} \in \mathbb{C}^{n_R}$ is defined as the noise vector, where $[\mathbf{n}]_i$ represents the noise component received at the i-th antenna. The noise is modeled as AWGN, with $\mathbb{E}|[\mathbf{n}]_i|^2 = 1$, $\forall i \in [1, n_R]$. The received vector, $\mathbf{y} \in \mathbb{C}^{n_R}$, for can thus be expressed as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$.

As it is stated in the previous section, the transmitted signal $\mathbf{x} \in \mathbb{C}^{n_T}$ is obtained from a linear combination of L independent data symbols, which are stacked to form the data symbols vector $\mathbf{s} \in \mathcal{C} \subset \mathbb{C}^L$, where $\mathcal{C}$ is a discrete set (codebook) containing the possible codewords for
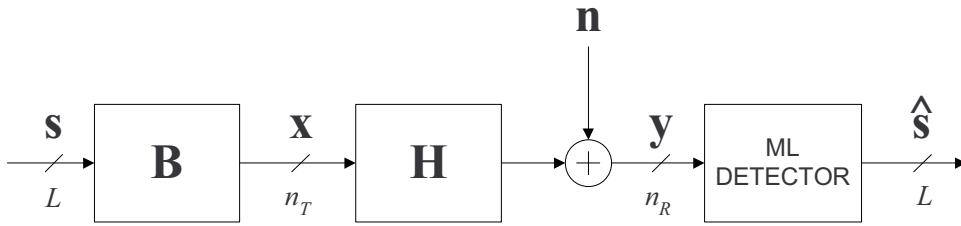
Figure 3.5: MIMO communications scheme considered in this chapter.

the vector $\mathbf{s}$ and where we assume $\mathbb{E}\mathbf{s}\mathbf{s}^H = \mathbf{I}_L$. Consequently, the transmitted signal is given by $\mathbf{x} = \mathbf{B}\mathbf{s}$, where $\mathbf{B} \in \mathbb{C}^{n_T \times L}$ represents a generic linear transformation. In addition, we recall that throughout the remainder of this chapter we assume that the receiver employs a ML detector. See Figure 3.5 for a graphical representation of the described communication scheme.

3.3.2   ML receiver

Since the receiver performs ML detection, the estimate of the data symbols vector is given by the vector $\hat{\mathbf{s}}$ maximizing the log-likelihood function particularized for the case where $\mathbf{R}_n = \mathbf{I}$,

$$\log \mathcal{L}(\mathbf{y}, \mathbf{B}\mathbf{s}) \propto (\mathbf{y} - \mathbf{H}\mathbf{B}\mathbf{s})^H \mathbf{R}_n^{-1} (\mathbf{y} - \mathbf{H}\mathbf{B}\mathbf{s}) = -\|\mathbf{y} - \mathbf{H}\mathbf{B}\mathbf{s}\|^2, \tag{3.11}$$

where the function $\mathcal{L}$ is the same as described in §2.5.1. The detected symbols vector is thus given by

$$\hat{\mathbf{s}} = \arg \max_{\mathbf{s} \in \mathcal{C}} \log \mathcal{L}(\mathbf{y}, \mathbf{B}\mathbf{s}). \tag{3.12}$$

In our case, there are $|\mathcal{C}|$ different data symbols vectors (codewords), which can then be indexed as $\mathbf{s}_m$, where $m \in \{1, 2, \ldots, |\mathcal{C}|\}$. The probability of deciding in favor of $\mathbf{s}_{m'}$ when $\mathbf{s}_m$ is actually transmitted is denoted by $\mathsf{pep}_{m,m'}$, and it is given by the probability that $\mathcal{L}(\mathbf{s}_{m'}) > \mathcal{L}(\mathbf{s}_m)$ conditioned on the fact that $\mathbf{s}_m$ is transmitted. We recall the expression for this probability from (2.35):

$$\mathsf{pep}_{m,m'} = \Pr\left\{\mathcal{L}(\mathbf{y}, \mathbf{s}_{m'}) > \mathcal{L}(\mathbf{y}, \mathbf{s}_m) | \mathbf{y} = \mathbf{H}\mathbf{B}\mathbf{s}_m + \mathbf{n}\right\} = \mathcal{Q}\left(\sqrt{\frac{d_{m,m'}^2}{2}}\right), \tag{3.13}$$

where $d_{m,m'}^2$ represents the squared Euclidean distance between the received constellation points $\mathbf{H}\mathbf{B}\mathbf{s}_m$ and $\mathbf{H}\mathbf{B}\mathbf{s}_{m'}$, *i.e.*,

$$d_{m,m'}^2 = \|\mathbf{H}\mathbf{B}\mathbf{s}_m - \mathbf{H}\mathbf{B}\mathbf{s}_{m'}\|^2 = \mathbf{e}_{m,m'}^H \mathbf{B}^H \mathbf{R}_\mathbf{H} \mathbf{B}\mathbf{e}_{m,m'}, \tag{3.14}$$

where $\mathbf{R}_\mathbf{H} = \mathbf{H}^H\mathbf{H}$ and $\mathbf{e}_{m,m'} = \mathbf{s}_m - \mathbf{s}_{m'}$. It is important to highlight that in case the noise is colored, *i.e.*, the case where $\mathbb{E}\mathbf{n}\mathbf{n}^H = \mathbf{R}_n \neq \mathbf{I}$, can be incorporated in this model just by letting $\mathbf{R}_\mathbf{H} = \mathbf{H}^H \mathbf{R}_n^{-1} \mathbf{H}$.

Now, we denote by $\mathcal{E}_T$ the set of all possible error vectors. Furthermore we define $\mathcal{E}$ as the set of error vectors that yield different values of the squared distance expression in (3.14). Note that there may exist groups of error vectors which yield the same distance, and consequently only one representative of the group need to be considered, lowering the total number of elements in $\mathcal{E}$ with respect to $\mathcal{E}_T$, *i.e.*, $|\mathcal{E}| \leq |\mathcal{E}_T|$.

In any case, since we are only interested in errors that yield different values of the minimum distance in (3.14), we can always index the possible error vectors from 1 to $|\mathcal{E}|$, as $\mathbf{e}_q$, with $q = \{1, 2, \ldots, |\mathcal{E}|\}$. Since the expression for $d_{m,m'}^2$ in (3.14) depends on the particular error vector, we can also index the different values of the set $\{d_{m,m'}^2\}$, with the same label that we use for the error vectors, q. In the following, d_q^2 is utilized instead of $d_{m,m'}^2$, and, equivalently, pep_q replaces $\mathsf{pep}_{m,m'}$.

3.3.3 Transmitter design to minimize the worst-case PEP

We now present the design of the optimum linear transmitter, $\mathbf{B}^\star$. The criterion for optimality is to minimize the maximum (or worst-case) PEP with respect to the set of error vectors, or, formally

$$\mathbf{B}^\star = \arg\min_{\mathbf{B}} \max_q \mathsf{pep}_q, \tag{3.15}$$

where $q \in \{1, 2, \ldots, |\mathcal{E}|\}$. Moreover, the search space of the minimization with respect to $\mathbf{B}$ is restricted to matrices that fulfill that the mean transmitted power does not exceed P_T, or, equivalently, to matrices such that $\operatorname{Tr} \mathbf{B}\mathbf{B}^H \leq P_T$. Notice that, due to the dependence given in (3.13), minimizing the worst-case PEP is equivalent to maximizing the squared minimum Euclidean distance between the received constellation points:

$$\arg\min_{\mathbf{B}} \max_q \mathsf{pep}_q = \arg\max_{\mathbf{B}} \min_q d_q^2 = \arg\max_{\mathbf{B}} d_{\min}^2. \tag{3.16}$$

Since the behavior of the mutual information in the high SNR regime is dictated by the minimum distance between the received constellation symbols, maximizing the minimum distance is a maximum mutual information achieving strategy in the high SNR regime. Consequently, our results agree with those in [Loz06] in the high SNR regime.

Writing explicitly the squared distance as a function of the transmission matrix $\mathbf{B}$ and the error vector $\mathbf{e}_q$ as in (3.14), the optimization problem in (3.15) is equivalently reformulated as

$$\mathbf{B}^\star = \arg\max_{\mathbf{B}} \quad \min_{\mathbf{e} \in \mathcal{E}} \mathbf{e}^H \mathbf{B}^H \mathbf{R}_\mathbf{H} \mathbf{B} \mathbf{e}, \tag{3.17}$$
$$\text{s.t.} \quad \operatorname{Tr} \mathbf{B}\mathbf{B}^H \leq P_T.$$

Utilizing the SVD, the optimum transmitter matrix $\mathbf{B}^\star$ can always be factored as

$$\mathbf{B}^\star = \mathbf{U}^\star \mathbf{\Sigma}^\star \mathbf{V}^{\star H}, \tag{3.18}$$

where $\mathbf{U}^\star \in \mathbb{C}^{n_T \times n_T}$ and $\mathbf{V}^\star \in \mathbb{C}^{L \times L}$ are unitary matrices and $\mathbf{\Sigma}^\star \in \mathbb{C}^{n_T \times L}$ is a matrix with non-zero elements only in its main diagonal. From a signal processing perspective, this matrix decomposition can be viewed as a three-step transformation of the data symbols vector. The matrix $\mathbf{U}^\star$ is a spatial processing matrix with n_T orthonormal columns coupling the transmission through the spatial modes of the channel, and $\mathbf{\Sigma}^\star$ performs the power allocation among these modes. Note also that, with this notation, the eigenvalues of the transmit covariance matrix $\mathbf{Q} = \mathbb{E}\mathbf{x}\mathbf{x}^H = \mathbf{B}\mathbf{B}^H$ are given by $\lambda_{\mathbf{Q},j} = [\mathbf{\Sigma}]_{jj}^2$. Finally, the purpose of $\mathbf{V}^\star$ is to spread the constellation symbols among these spatial modes.

In the following subsections we characterize each one of these matrices.

Optimal spatial processing matrix

The optimal spatial processing matrix $\mathbf{U}^\star$ is found next:

Proposition 3.3.1 *Let the SVD of $\mathbf{R_H}$ be given by $\mathbf{R_H} = \mathbf{U_H}\mathbf{\Lambda_H}\mathbf{U_H}^H$, where the elements of $\mathbf{\Lambda_H}$ are sorted in decreasing order, then it follows that there exists a solution to the optimization problem in (3.17) with the following structure:*

$$\mathbf{B}^\star = \mathbf{U_H}\mathbf{\Sigma}^\star\mathbf{V}^{\star H}, \tag{3.19}$$

where $\mathbf{\Sigma}^\star$ and $\mathbf{V}^\star$ are the same as defined in (3.18).

Proof Consider any given matrix $\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. For this generic choice the quadratic form in the objective function in (3.17) becomes

$$\mathbf{e}^H\mathbf{V}\mathbf{\Sigma}\mathbf{U}^H\mathbf{R_H}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{e}. \tag{3.20}$$

We now consider the SVD decomposition $\mathbf{\Sigma}\mathbf{U}^H\mathbf{R_H}\mathbf{U}\mathbf{\Sigma} = \mathbf{Q}\mathbf{D}\mathbf{Q}^H$, with $\mathbf{D}$ diagonal and $\mathbf{Q}$ unitary, which clearly implies that

$$\mathbf{D} = \mathbf{Q}^H\mathbf{\Sigma}\mathbf{U}^H\mathbf{R_H}\mathbf{U}\mathbf{\Sigma}\mathbf{Q} \tag{3.21}$$

is a diagonal matrix. From [Pal03b, Lemma 12], we can state that there exists a matrix $\mathbf{X} = \mathbf{U_H}\mathbf{\Sigma_X}$, with $\mathbf{\Sigma_X}$ having non-zero elements only in the main diagonal, such that $\mathbf{X}^H\mathbf{R_H}\mathbf{X} = \mathbf{D}$ and that $\mathrm{Tr}\,\mathbf{X}\mathbf{X}^H \leq \mathrm{Tr}\,\mathbf{\Sigma}\mathbf{\Sigma}^H = \mathrm{Tr}\,\mathbf{B}\mathbf{B}^H$. We now only need to check that

$$\mathbf{B}^H\mathbf{R_H}\mathbf{B} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^H\mathbf{R_H}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \mathbf{V}\mathbf{Q}\mathbf{D}\mathbf{Q}^H\mathbf{V}^H = \mathbf{V}\mathbf{Q}\mathbf{X}^H\mathbf{R_H}\mathbf{X}\mathbf{Q}^H\mathbf{V}^H. \tag{3.22}$$

Defining $\widetilde{\mathbf{B}} = \mathbf{X}\mathbf{Q}^H\mathbf{V}^H = \mathbf{U_H}\mathbf{\Sigma_X}\widetilde{\mathbf{V}}^H$, with $\widetilde{\mathbf{V}} = \mathbf{V}\mathbf{Q}$, we have shown by construction that for any given matrix $\mathbf{B}$ we can find another matrix $\widetilde{\mathbf{B}}$ such that the objective function in (3.17) is the same $\mathbf{B}^H\mathbf{R_H}\mathbf{B} = \widetilde{\mathbf{B}}^H\mathbf{R_H}\widetilde{\mathbf{B}}$, whereas the required transmitted power is now lower, $\mathrm{Tr}\,\widetilde{\mathbf{B}}\widetilde{\mathbf{B}}^H = \mathrm{Tr}\,\mathbf{X}\mathbf{X}^H \leq \mathrm{Tr}\,\mathbf{B}\mathbf{B}^H$. Consequently, the optimal transmitter must be of the form $\mathbf{B}^\star = \mathbf{U_H}\mathbf{\Sigma}^\star\mathbf{V}^{\star H}$. ∎

This solution implies that the optimal transmitter $\mathbf{B}^\star$ transmits the processed symbols through the eigenmodes of the channel. The corresponding gains of these eigenmodes are given by $\lambda_{\mathbf{H},j}$.

Once we have derived the left singular vectors of the optimal transmitter structure $\mathbf{B}^\star$ we now proceed to further characterize the optimal structure of the solution.

Optimal power allocation, $\mathbf{\Sigma}^\star$

In this section we derive the structure of the optimal power allocation matrix $\mathbf{\Sigma}^\star$. It is shown that the optimal power allocation inverts the gains of the eigenmodes, $\lambda_{\mathbf{H},j}$, up to a fixed scaling factor which is different for each eigenmode, recovering the results of [Loz06, Theorem 7] in the high SNR regime. For the sake of notation, we restrict ourselves to the case where the number of components of the data symbols vector is equal to the rank of $\mathbf{R_H}$ and also to the number of antennas at the transmitter and the receiver, $L = n_T = n_R = \operatorname{rank} \mathbf{R_H}$. The following is also valid for the general case of arbitrary values of L, n_T, n_R, and rank $\mathbf{R_H}$, but care has to be taken with the dimensions of the matrices.

With $\mathbf{B}^\star = \mathbf{U_H} \mathbf{\Sigma}^\star \mathbf{V}^{\star H}$, the problem in (3.17) simplifies to

$$\{\mathbf{\Lambda_Q^\star}, \mathbf{V}^\star\} = \arg \max_{\mathbf{\Lambda_Q}, \mathbf{V}} \quad \min_{\mathbf{e} \in \mathcal{E}} \quad \mathbf{e}^H \mathbf{V} \mathbf{\Lambda_H} \mathbf{\Lambda_Q} \mathbf{V}^H \mathbf{e},$$
$$\text{s.t.} \quad \operatorname{Tr} \mathbf{\Lambda_Q} \leq P_T, \tag{3.23}$$
$$\mathbf{V}\mathbf{V}^H = \mathbf{V}^H \mathbf{V} = \mathbf{I},$$

where, as commented above, $\mathbf{\Lambda_Q} = \mathbf{\Sigma}^2$ is a diagonal matrix containing the eigenvalues of the transmit covariance matrix. The matrix $\mathbf{V}$ is now assumed fixed and we focus on the optimization of the matrix $\mathbf{\Lambda_Q}$. To this purpose, we index the set of possible error vectors with q and rewrite the objective function in (3.23) as

$$\mathbf{e}_q^H \mathbf{V} \mathbf{\Lambda_H} \mathbf{\Lambda_Q} \mathbf{V}^H \mathbf{e}_q = \operatorname{Tr} \mathbf{\Lambda_H} \mathbf{\Lambda_Q} \mathbf{V}^H \mathbf{e}_q \mathbf{e}_q^H \mathbf{V} = \sum_{j=1}^{n_T} \lambda_{\mathbf{H},j} \lambda_{\mathbf{Q},j} \alpha_{qj}, \tag{3.24}$$

where we have defined the variable $\alpha_{qj} = [\mathbf{V}^H \mathbf{e}_q \mathbf{e}_q^H \mathbf{V}]_{jj}$. Introducing the new optimization variable $t = \min_q \sum_{j=1}^{n_T} \lambda_{\mathbf{H},j} \lambda_{\mathbf{Q},j} \alpha_{qj}$ we can rewrite the maximization with respect to $\mathbf{\Lambda_Q}$ in the problem in (3.23) as

$$\max_{t, \{\lambda_{\mathbf{Q},j}\}} \quad t$$
$$\text{s.t.} \quad \sum_{j=1}^{n_T} \lambda_{\mathbf{H},j} \lambda_{\mathbf{Q},j} \alpha_{qj} \geq t, \quad \forall q \in \{1, 2, \ldots, |\mathcal{E}|\},$$
$$\sum_{j=1}^{n_T} \lambda_{\mathbf{Q},j} \leq P_T, \tag{3.25}$$
$$\lambda_{\mathbf{Q},j} \geq 0, \quad 1 \leq j \leq n_T.$$

Note that the previous optimization problem is a linear (and thus convex) optimization problem in the optimization variables $\{\lambda_{\mathbf{Q},j}\}$ and t.

The set of inequalities $\sum_{j=1}^{n_T} \lambda_{\mathbf{H},j}\lambda_{\mathbf{Q},j}\alpha_{qj} \geq t$, $\forall q \in \{1,2,\ldots,|\mathcal{E}|\}$ will be tight for a subset of indices $q \in \mathcal{T}$. *A priori*, the set $\mathcal{T}$ is unknown, but we assume it given by an oracle. Then, clearly, the optimal solution must fulfill

$$\sum_{j=1}^{n_T} \lambda_{\mathbf{H},j}\lambda_{\mathbf{Q},j}^\star\alpha_{qj} = t^\star, \quad \forall q \in \mathcal{T}, \tag{3.26}$$

which can be more compactly rewritten as

$$\mathbf{A}\mathbf{\Lambda_H}\boldsymbol{\lambda}_\mathbf{Q}^\star = t^\star\mathbf{1}, \tag{3.27}$$

where $[\mathbf{A}]_{n(q)j} = \alpha_{qj}$, $[\boldsymbol{\lambda}_\mathbf{Q}^\star]_j = \lambda_{\mathbf{Q},j}^\star$, $\mathbf{1}$ is the all-one vector, and the function $n(q)$ can be any arbitrary function fulfilling that it assigns to each element $q \in \mathcal{T}$ a different number from 1 to $|\mathcal{T}|$. From the equation $\mathbf{A}\mathbf{\Lambda_H}\boldsymbol{\lambda}_\mathbf{Q}^\star = t^\star\mathbf{1}$ a necessary and sufficient structure of the optimal power allocation is

$$\lambda_{\mathbf{Q},j}^\star = \frac{t^\star\beta_j}{\lambda_{\mathbf{H},j}}, \tag{3.28}$$

where $[\boldsymbol{\beta}]_j = \beta_j$ is such that $\mathbf{A}\boldsymbol{\beta} = \mathbf{1}$ which always has at least one solution. For the case where the rank of $\mathbf{A}$ is equal to L (which can always be assumed in practice) the vector $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta} = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{1}. \tag{3.29}$$

Note that $\boldsymbol{\beta}$ is a function of matrix $\mathbf{V}$, the error vectors $\mathbf{e}_q$, and the set $\mathcal{T}$.

Finally, the value of $t^\star$ can be found from the restriction on the total transmitted power as

$$\sum_{j'=1}^{n_T} \frac{t^\star\beta_{j'}}{\lambda_{\mathbf{H},j'}} = P_T \Rightarrow t^\star = \frac{P_T}{\sum_{j'=1}^{n_T} \frac{\beta_{j'}}{\lambda_{\mathbf{H},j'}}}, \tag{3.30}$$

and we obtain,

$$\lambda_{\mathbf{Q},j}^\star = \frac{\beta_j}{\lambda_{\mathbf{H},j}} \frac{P_T}{\sum_{j'=1}^{n_T} \frac{\beta_{j'}}{\lambda_{\mathbf{H},j'}}}. \tag{3.31}$$

The solution found in (3.31) is such that the equivalent gain of each eigenmode is now $\lambda_{\mathbf{Q},j}\lambda_{\mathbf{H},j} = t^\star\beta_j$, where $t^\star$ is given by (3.30) which is a common factor to all these eigenmodes. Note that this common factor only depends on the set of active constraints $\mathcal{T}$, *i.e.*, the relative ratios between the equivalent gains at the receiver are fixed for any value of the gains $\lambda_{\mathbf{H},j}$, whenever the variation of these gains does not change the set of active constraints $\mathcal{T}$.

It is important to mention that, although the expression in (3.31) sheds some light on the optimal design of a linear transmitter, it can not be utilized in practice as it depends on the

set $\mathcal{T}$ which is *a priori* unknown. The optimal solution must be found by numerically solving the convex optimization problem in (3.25). Once the solution is known, it is very easy to find the set $\mathcal{T}$ and check that (3.31) is actually fulfilled. Note also that (3.31) agrees with the results reported in [Pay06b] for the particular case of transmitting 2 QPSK data streams and which are reproduced in §3.3.4.

Comments on the optimal V

It now remains to calculate $\mathbf{V}^\star$ to obtain the full design of the transmitter $\mathbf{B}^\star$. Unfortunately, the problem of finding the optimal $\mathbf{V}$ matrix appears to be intractable. Consequently, we propose an *ad-hoc* algorithm to obtain a good approximation of $\mathbf{V}^\star$.

1. Generate a set of possible candidates $\mathcal{V} \triangleq \{\mathbf{V}_c\}$. A good choice for this set consists in selecting these matrices from the Grassmannian manifold of the appropriate dimension [Lov03], because it yields a set of matrices that uniformly cover all the unitary matrices space.

2. For a given channel matrix $\mathbf{H}$, solve the optimization problem in (3.25) for each matrix in $\mathcal{V}$ and calculate the minimum distance.

3. Choose the matrix from $\mathcal{V}$ that yields the maximum minimum distance as the transmitter $\mathbf{V}$ matrix

The main drawback of this algorithm is that it can become computationally hard because, in order to obtain a good transmitter design, the cardinality of $\mathcal{V}$ has to be high, and for each candidate in $\mathcal{V}$ an optimization problem has to be solved to obtain the corresponding optimal power allocation.

3.3.4   Closed form solution for the particular case of two QPSK streams

The framework for the design of linear transmitters presented in the preceding sections is general in the sense that the symbols codebook $\mathcal{C}$, the number of data streams L, and the MIMO configuration given by n_T and n_R can be arbitrarily chosen. However, no closed form solution is, in general, available for the design of the optimal transmitter $\mathbf{B}^\star$.

To gain further insight in this problem, we now find a closed form transmitter that maximizes the minimum distance for a particular case. Precisely, we focus on a case with arbitrary values of n_T and n_R and with $L = 2$ data streams where the elements of the data symbols vector belong to a QPSK constellation. Formally, we can state that $\mathbf{s} \in \mathcal{C} \triangleq \mathcal{M} \times \mathcal{M} \subset \mathbb{C}^2$, where $\mathcal{M}$

represents the QPSK constellation, *i.e.*,

$$\mathcal{M} \triangleq \left\{ \frac{1+i}{\sqrt{2}}, \frac{1-i}{\sqrt{2}}, \frac{-1+i}{\sqrt{2}}, \frac{-1-i}{\sqrt{2}} \right\}. \tag{3.32}$$

Consequently, the transmitted signal is given by $\mathbf{x} = \mathbf{Bs}$, with $\mathbf{B} \in \mathbb{C}^{n_T \times 2}$.

From this particular choice of codebook $\mathcal{C}$ and the definition of the error vector, $\mathbf{e}_{n,m} = \mathbf{s}_n - \mathbf{s}_m$, one easily sees that each component of the error vector $\mathbf{e}_{n,m}$ must belong to the set

$$\mathcal{S} \triangleq \left\{ 0, \pm\sqrt{2}, \pm i\sqrt{2}, \pm\sqrt{2} \pm i\sqrt{2} \right\}. \tag{3.33}$$

We recall that we defined the set $\mathcal{E}$ as the set of error vectors that yield different values for the objective function and, $\mathcal{E}_T$ as the set of all possible error vectors. Considering that the zero vector can not belong to $\mathcal{E}_T$, we obtain

$$\mathcal{E}_T = \mathcal{S}^2 \backslash \{\mathbf{0}\}. \tag{3.34}$$

The cardinal of $\mathcal{E}_T$ is $|\mathcal{E}_T| = |\mathcal{S}|^2 - 1 = 80$. This implies that the different error vectors can be indexed from 1 to 80, as $\mathbf{e}_q$, with $q = \{1, 2, \ldots, 80\}$. Similarly, we recall that we defined $\mathcal{E}$ as the set of error vectors that yield different values of the objective function in (3.17). In Appendix 3.A it is shown that the number of errors that yield different values of the objective function in (3.17) is given by $|\mathcal{E}| = 14$.

We are now ready to particularize the general problem in (3.17) for the specific choice of the parameters described above.

Proposition 3.3.2 *Consider the following constrained maximin optimization problem:*

$$\begin{aligned} \max_{\mathbf{B}} \min_{\mathbf{e}} \quad & \mathbf{e}^H \mathbf{B}^H \mathbf{R_H} \mathbf{Be}, \\ s.t. \quad & \mathrm{Tr}\, \mathbf{BB}^H \leq P_T, \\ & \mathbf{e} \in \mathcal{E}, \\ & \mathbf{B} \in \mathbb{C}^{n_T \times 2}, \end{aligned} \tag{3.35}$$

where $\mathbf{R_H} \in \mathbb{C}^{n_T \times n_T}$ is a positive semidefinite hermitian matrix with $\lambda_{\mathbf{H},1}$ and $\lambda_{\mathbf{H},2}$ being its two largest eigenvalues, with $\lambda_{\mathbf{H},1} \geq \lambda_{\mathbf{H},2}$, and $\mathcal{E}$ is the set of error vectors that yield different values of the objective function in (3.35). It then follows that there is an optimal solution, $\mathbf{B}^\star$, which is given by $\mathbf{B}^\star = \mathbf{U_H} \mathbf{\Sigma} \mathbf{V}^H$, where $\mathbf{U_H} \in \mathbb{C}^{n_T \times 2}$ has as columns the eigenvectors of $\mathbf{R_H}$ corresponding to $\lambda_{\mathbf{H},1}$ and $\lambda_{\mathbf{H},2}$. In addition, defining

$$\lambda_c \triangleq \frac{(\sqrt{3}-1)(3-2\sqrt{2})}{1+3\sqrt{3}-2\sqrt{6}} \simeq 9.683 \cdot 10^{-2}, \tag{3.36}$$

the optimal solution is completed with
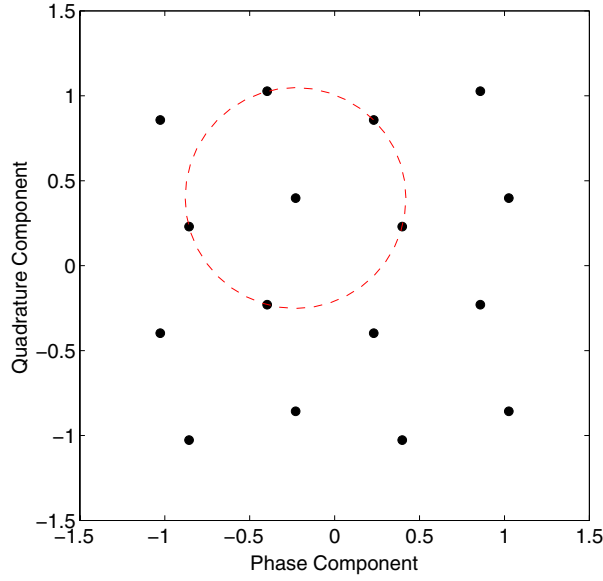
Figure 3.6: Optimal received constellation when $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$. Notice that the four innermost points have five neighbors at minimum distance, as opposed to the 16-QAM constellation where they have only four neighbors.

- If $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$, then

$$\boldsymbol{\Sigma} = \left(\begin{array}{cc} \sqrt{P_T} & 0 \\ 0 & 0 \end{array} \right), \tag{3.37}$$

$$\mathbf{V}^H = \left(\begin{array}{cc} \cos\theta_T & -e^{\mathrm{i}\,\phi_T}\sin\theta_T \\ e^{-\mathrm{i}\,\phi_T}\sin\theta_T & \cos\theta_T \end{array} \right), \tag{3.38}$$

where $\phi_T = \arccos\left(\frac{3+\sqrt{3}}{2\sqrt{6}}\right)$ *and* $\theta_T = \arctan\left(\frac{\sqrt{6}}{3+\sqrt{3}}\right)$.

- If $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \geq \lambda_c$, then

$$\boldsymbol{\Sigma} = \sqrt{\frac{P_T}{\lambda_{\mathbf{H},2} + \alpha_0\lambda_{\mathbf{H},1}}} \left(\begin{array}{cc} \sqrt{\lambda_{\mathbf{H},2}} & 0 \\ 0 & \sqrt{\alpha_0\lambda_{\mathbf{H},1}} \end{array} \right), \tag{3.39}$$

$$\mathbf{V}^H = \frac{1}{\sqrt{2}} \left(\begin{array}{cc} 1 & -e^{\mathrm{i}\,\pi/4} \\ e^{-\mathrm{i}\,\pi/4} & 1 \end{array} \right), \tag{3.40}$$

where $\alpha_0 = 3 - 2\sqrt{2} \simeq 1.716 \cdot 10^{-1}$.

Proof See Appendix 3.A. ∎

Noteworthy, the inner structure of the optimal transmission matrix $\mathbf{B}^\star$ depends on the relation of the two largest eigenvalues of $\mathbf{R_H}$.

On one hand, when this relation is low, $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$, only the strongest eigenmode is found useful for transmission (the rank of $\boldsymbol{\Sigma}$ is one), and then a new constellation is created
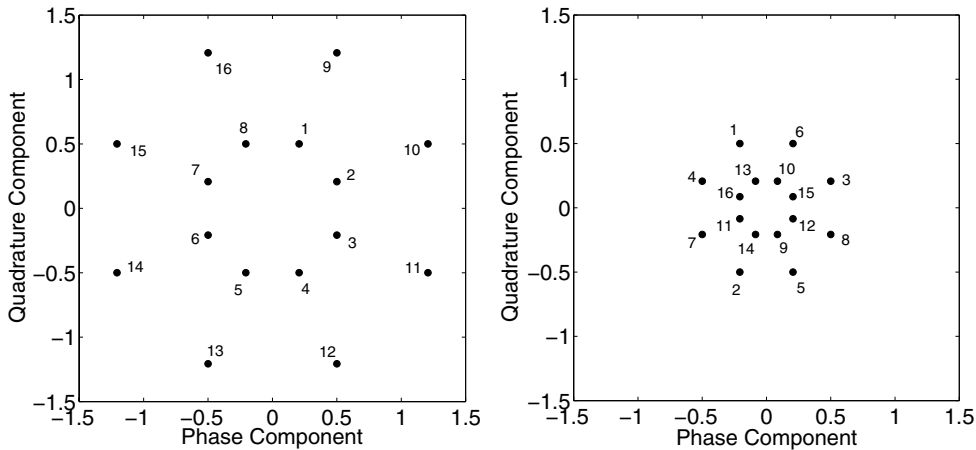
Figure 3.7: Received constellation through the two largest eigenmodes. Left, eigenmode associated with $\lambda_{\mathbf{H},1}$. Right, eigenmode associated with $\lambda_{\mathbf{H},2}$. Notice that the constellation points at the two different eigenmodes are paired by numbers from 1 to 16. Each pair is transmitted and received together.

using the two QPSK streams. This new constellation is very similar to a 16-QAM modulation, with little perturbations over the positions of the constellation points in order to maximize the minimum distance among them and, consequently, maximize the number of neighbors at minimum distance. In Figure 3.6 this new constellation is depicted.

On the other hand, if the relation $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1}$ is bigger than the threshold value λ_c, then the two eigenmodes associated with the two largest eigenvalues are used for transmission. In this case, two similar signal constellations with 16 points are transmitted through the two eigenmodes. If we consider the received constellation $\mathbf{HBs}$, we can express it, up to a unitary transformation $\mathbf{Q} \in \mathbb{C}^{n_R \times 2}$ which preserves the distances, as

$$\mathbf{HBs} = \mathbf{Q}\sqrt{\frac{P_T\lambda_{\mathbf{H},1}\lambda_{\mathbf{H},2}}{\lambda_{\mathbf{H},2} + \alpha_0\lambda_{\mathbf{H},1}}} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\alpha_0} \end{pmatrix} \mathbf{V}^H\mathbf{s}, \tag{3.41}$$

where we have represented the two largest eigenmodes of the channel matrix $\mathbf{H}$ in its SVD, $\mathbf{H} = \mathbf{Q}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{U}_{\mathbf{H}}^H$. See Figure 3.7 for a graphical representation of the received constellation. Note that the constellation points at the two different eigenmodes are paired, in the sense that they are transmitted and received together, giving a total of 16 different symbols. This implies that the symbol rate is the same as in the case of using only one channel eigenmode. In addition, the points in the outer circle in one of the eigenmodes are paired with the points in the inner circle in the other eigenmode, and vice versa. Moreover, equation (3.41) implies that the relation between the sizes of the constellations received through the two eigenmodes is fixed and equal to $\sqrt{\alpha_0} = \sqrt{2} - 1$. This fixed ratio is optimal in the sense that it maximizes the minimum distance between the received constellation points.

To give a visual idea of the two optimal transmission schemes we present a picture of both in Figure 3.8.

It is also important to highlight that the optimal values found for the diagonal entries of matrix $\mathbf{\Sigma}^\star$ for the 2 QPSK stream case agree with the general results reported in (3.31).

- For the case $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$ we have

$$\lambda_{\mathbf{Q},1}^\star = [\mathbf{\Sigma}^\star]_{11}^2 = P_T, \tag{3.42}$$

$$\lambda_{\mathbf{Q},2}^\star = [\mathbf{\Sigma}^\star]_{11}^2 = 0, \tag{3.43}$$

which agrees with the results in (3.31) with $\beta_1 = 1$ and $\beta_2 = 0$.

- For the case $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \geq \lambda_c$ we have

$$\lambda_{\mathbf{Q},1}^\star = [\mathbf{\Sigma}^\star]_{11}^2 = \frac{P_T \lambda_{\mathbf{H},2}}{\lambda_{\mathbf{H},2} + \alpha_0 \lambda_{\mathbf{H},1}} = \frac{1}{\lambda_{\mathbf{H},1}} \frac{P_T}{\frac{1}{\lambda_{\mathbf{H},1}} + \frac{\alpha_0}{\lambda_{\mathbf{H},2}}}, \tag{3.44}$$

$$\lambda_{\mathbf{Q},2}^\star = [\mathbf{\Sigma}^\star]_{11}^2 = \frac{P_T \alpha_0 \lambda_{\mathbf{H},1}}{\lambda_{\mathbf{H},2} + \alpha_0 \lambda_{\mathbf{H},1}} = \frac{\alpha_0}{\lambda_{\mathbf{H},2}} \frac{P_T}{\frac{1}{\lambda_{\mathbf{H},1}} + \frac{\alpha_0}{\lambda_{\mathbf{H},2}}}, \tag{3.45}$$

which agrees with the results in (3.31) with $\beta_1 = 1$ and $\beta_2 = \alpha_0$.
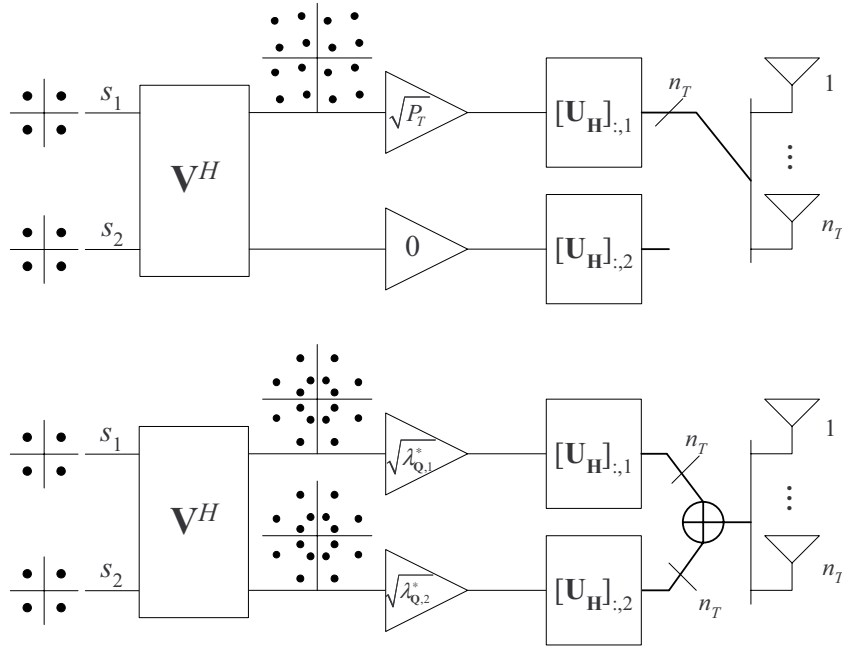
Figure 3.8: Graphical representation of the two optimal transmission architectures. The upper scheme is optimal when $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$, notice that only one eigenmode is used. The lower scheme is optimal when $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \geq \lambda_c$.

3.4   Simulation results

The transmitter design presented in the previous section is optimal in the sense that it maximizes the minimum distance (MMD) between the received constellation points. To validate the goodness of this criterion, we have compared the performance of our transmission scheme – MMD – with the performance of four well known transmission architectures: V-BLAST [Fos96, Wol98], OSTBC schemes [Ala98, Tar99a], the minimum BER optimal linear design (MBOL) [Pal03b], and the maximum minimum SNR eigenvalue design (MMS) [Sca02]. For the sake of fairness, in all these cases the receiver performs ML detection, and the rate is fixed.

3.4.1   Transmission of four bit per channel use

In this case we have fixed the transmission rate at 4 bit per channel use. This implies that the BLAST scheme transmits two QPSK symbols per channel use, the OSTBC scheme (Alamouti) transmits two 16-QAM symbols each two channel uses, and the MMD, MBOL, and MMS designs transmit a linear combination of two QPSK symbols.

We have considered a random 2×2 MIMO channel, with i.i.d. Rayleigh entries. Firstly, we have obtained the pdfs of the squared minimum distance, $d_{\min}^2$, between the received constellation points, when the transmission power is fixed to unity (see Figure 3.9). The mean of the pdfs of the squared minimum distance is summarized in Table 3.1.

Table 3.1: Mean $d_{\min}^2$ for Different Transmission Architectures

Scheme	Mean $d_{\min}^2$
Maximum Minimum SNR Eigenvalue (MMS)	0.8001
16-QAM Alamouti	0.8003
QPSK BLAST	0.9924
Optimum Linear (MBOL)	1.0848
Maximum Minimum Distance (MMD)	1.8696

As expected, the MMD scheme presents the highest mean squared minimum distance, because it yields the maximum minimum distance for each channel realization.

3.4.2   Transmission of eight bit per channel use

Case with $n_T = n_R = 2$

Firstly, we have considered a transmission of eight bit per channel use through a random 2×2 MIMO channel with i.i.d. Rayleigh entries. In the case of the MMD, MBOL, and MMS designs we are transmitting a linear combination of two 16-QAM streams. In the case of V-BLAST we
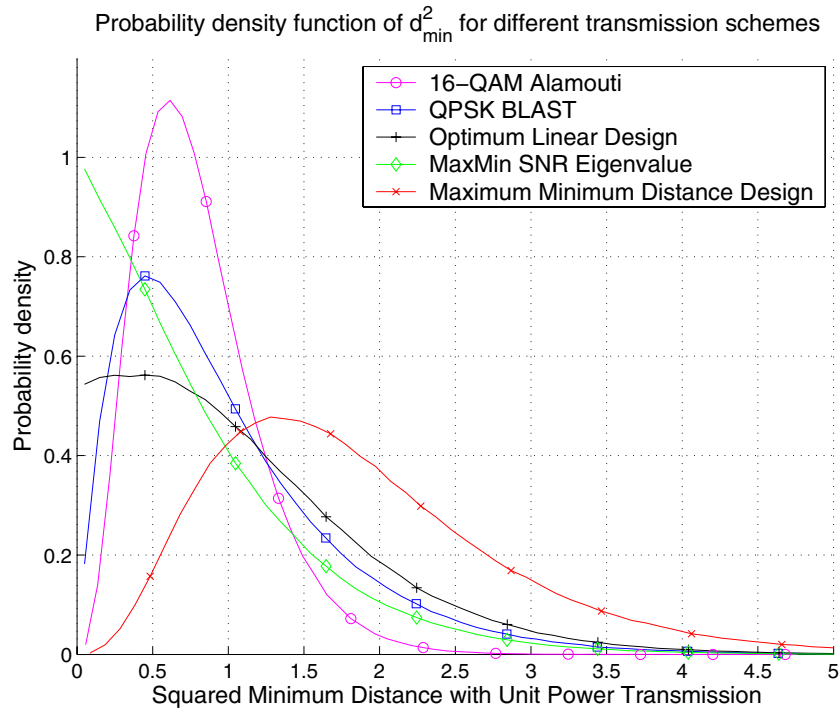
Figure 3.9: Pdf of the squared minimum distance $d^2_{\min}$ for different types of communication schemes. We have fixed $P_T = 1$.

are directly transmitting an independent 16-QAM data stream through each antenna. Finally in the case of OSTBC (Alamouti) we are transmitting two 256-QAM symbols each two channel uses.

We have obtained the cdf of the squared minimum distance, $d^2_{\min}$, between the received constellation points (see Figure 3.10). Note how our proposed scheme, MMD, yields always the highest minimum distance as its plot is always the rightmost for a given probability value. The mean of the squared minimum distance is summarized in Table 3.2.

Table 3.2: Mean $d^2_{\min}$ for Different Transmission Architectures

Scheme	Mean $d^2_{\min}$
256-QAM OSTBC (Alamouti)	0.0465
Optimum Linear (MBOL)	0.1468
Maximum Minimum SNR Eigenvalue (MMS)	0.1555
16-QAM BLAST	0.1790
Maximum Minimum Distance (MMD)	0.2974

Cumulative distribution function of d^2_{min} for different transmission schemes
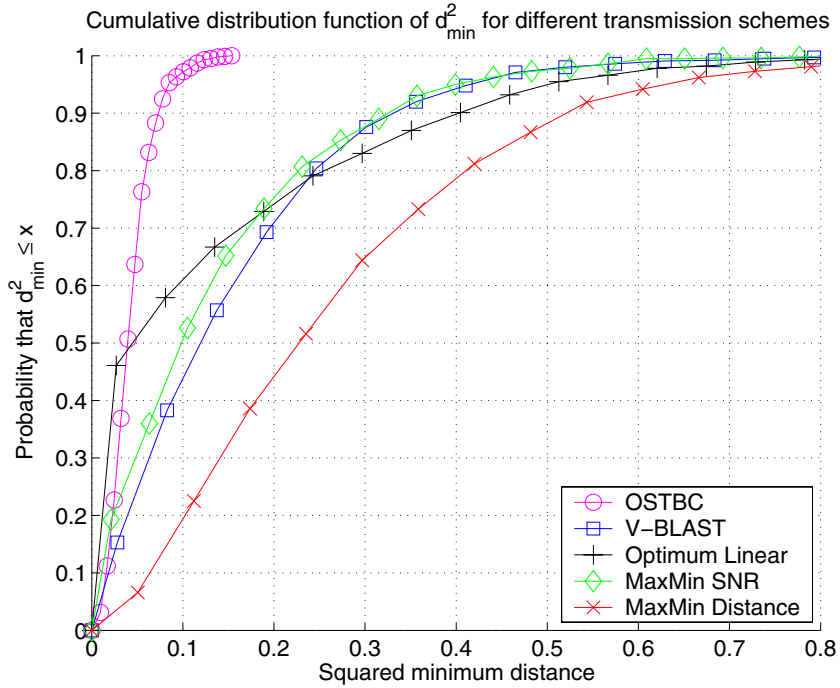
Figure 3.10: Cdf of the squared minimum distance among the received constellation points.

Case with $n_T = n_R = 4$

Secondly, we have considered a transmission of eight bit per channel use through a fixed 4×4 MIMO channel. In the case of the MMD, MBOL, and MMS designs we are transmitting a linear combination of four QPSK data streams. In the case of V-BLAST we are directly transmitting an independent QPSK data stream through each one of the four antennas. Finally in the case of OSTBC we assume that a rate one code exists for four antennas and consequently we are transmitting four 256-QAM symbols each four channel uses[3].

For this case of eight bit per channel use transmission with $n_T = n_R = 4$ the performance comparison will be made in the following subsection in terms of SER.

3.4.3   Performance in terms of SER

Up to this point we have only compared the performance of our scheme with the ones in the existing literature in terms of minimum distance. Since our proposed architecture is aimed at maximizing the minimum distance the results that we have obtained so far are not surprising.

To complete the comparison, in Figure 3.11 (4 bit per channel use) and Figure 3.12 (8 bit

---

[3]Note that this assumption is advantageous to the OSTBC scheme as no rate one codes exist for four antennas, and consequently modulations richer than 256-QAM should be used and the minimum distance would be decreased.
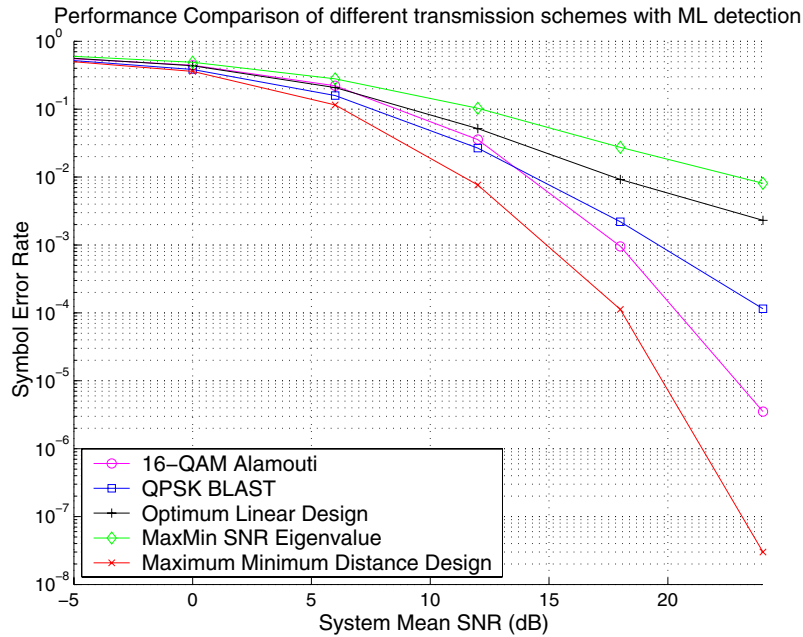
Figure 3.11: SER for different types of communication schemes for a 2×2 MIMO channel when the rate is fixed to 4 bit per channel.

per channel use) we present the mean SER vs. the system mean SNR. Also in this case, the MMD scheme shows the best performance, because it makes an efficient use of the CSI at the transmitter side. As a counterpart, we recall that the BLAST and Alamouti schemes do not need any CSI at the transmitting end.

Noteworthy, although the MMD scheme is not originally designed to minimize the mean SER (but the worst-case PEP), it also presents an excellent performance in its terms. This implies that the minimization of the worst-case PEP is a key point in the design of good transmission schemes where the performance is dictated by the SER. If we consider the BER instead of the SER as a performance metric, similar results to the ones presented here are obtained.

In addition, the poor performance of MBOL and MMS, specially at high SNR, is due to the fact that, for these two cases, $d_{\min}^2$ can take values close to zero with non-zero probability (see Figure 3.9), as opposed to the other cases (BLAST, Alamouti, and MMD) where the probability that $d_{\min}^2$ takes small values tends to zero.

3.5 Chapter summary and conclusions

In this chapter, we have studied the design of the optimal linear transmitter for a MIMO system with full CSI when the objective is the maximization of the minimum distance among the received constellation points, which is a practical design criterion as opposed to mutual information
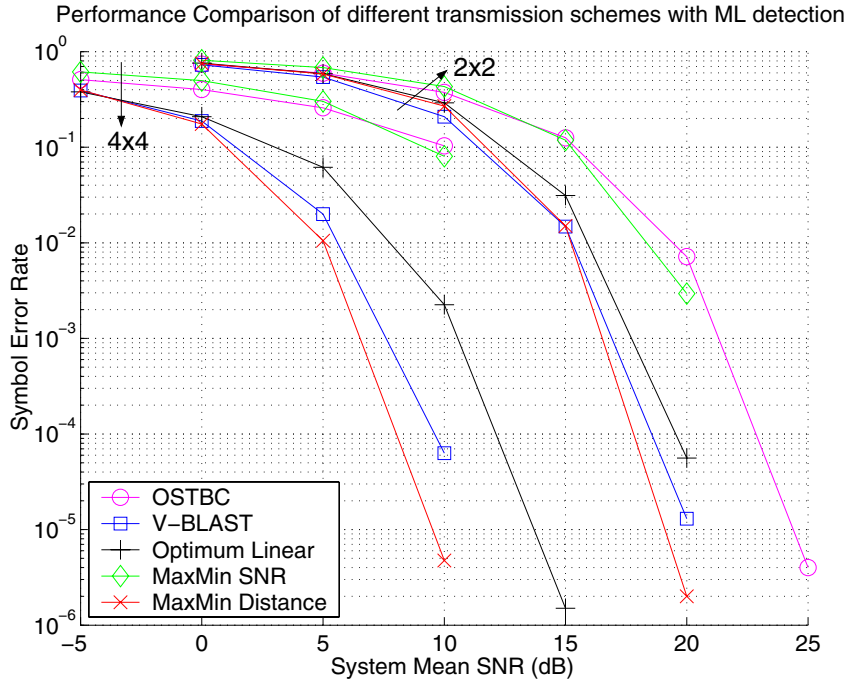
Figure 3.12: SER for different types of communication schemes when the rate is fixed to 8 bit per channel use. Two MIMO channel configuration have been considered: 2×2 and 4×4.

criteria which assume that infinite-length codes are being utilized.

We have optimally characterized the left singular vectors, $\mathbf{U}^\star$, and the singular values, $\mathbf{\Sigma}^\star$, of the transmitter matrix, $\mathbf{B}^\star$, and have given a simple algorithm to obtain a good approximation to $\mathbf{V}^\star$.

We have also found that our proposed scheme yields an excellent performance in terms of mean minimum distance and SER, when compared to well-known schemes such as V-BLAST, OSTBC schemes, the minimum BER optimum linear design, or the maximum minimum SNR eigenvalue design. Unfortunately, for the general case, the computational complexity of our proposed scheme is rather high as it involves a search over a big set of matrices $\{\mathbf{V}\}$, which makes it difficult to implement in fast-fading environments. However, for fixed fading environments such that the transmitter calculation can can be performed off-line we have shown that it presents an excellent performance.

In addition, for the particular case of transmitting two QPSK streams, we have presented the design in closed form, which reduces enormously the computational complexity as the search over the set of matrices $\{\mathbf{V}\}$ is not necessary since an explicit expression for it is available.

3.A Proof of Proposition 3.3.2

From Proposition 3.3.1, the objective function of the optimization problem in (3.35) becomes

$$\mathbf{e}^H \mathbf{V} \boldsymbol{\Sigma} \boldsymbol{\Lambda_H} \boldsymbol{\Sigma} \mathbf{V}^H \mathbf{e}, \tag{3.46}$$

where we have used that $\mathbf{B} = \mathbf{U_H} \boldsymbol{\Sigma} \mathbf{V}^H$ and $\mathbf{R_H} = \mathbf{U_H} \boldsymbol{\Lambda_H} \mathbf{U_H}^H$. Notice that the restriction $\text{Tr} \, \mathbf{BB}^H \leq P_T$ becomes $\text{Tr} \, \boldsymbol{\Sigma} \boldsymbol{\Sigma}^H \leq P_T$.

Now, suppose that we fix $\boldsymbol{\Sigma}$ and we define $\mathbf{L} = \boldsymbol{\Sigma} \boldsymbol{\Lambda_H} \boldsymbol{\Sigma}$. The original problem in (3.35) becomes

$$\max_{\boldsymbol{\Sigma}} \max_{\mathbf{V}} \min_{\mathbf{e}} \quad \mathbf{e}^H \mathbf{V} \mathbf{L} \mathbf{V}^H \mathbf{e}, \tag{3.47}$$

where the dependence of the objective function on $\boldsymbol{\Sigma}$ is implicit in $\mathbf{L}$. Since $\mathbf{L}$ is a diagonal matrix and $\mathbf{V}$ is unitary, they can be parameterized as

$$\mathbf{L} = \begin{pmatrix} \ell_1 & 0 \\ 0 & \ell_2 \end{pmatrix}, \text{ and } \mathbf{V} = \begin{pmatrix} e^{i\delta} \cos\theta & e^{i\epsilon} \sin\theta \\ -e^{-i\epsilon} \sin\theta & e^{-i\delta} \cos\theta \end{pmatrix}. \tag{3.48}$$

The optimization problem becomes

$$\max_{\boldsymbol{\Sigma}} \max_{\theta,\delta,\epsilon} \min_{q \leq |\mathcal{E}_T|} \quad \mathbf{e}_q^H \mathbf{V} \mathbf{L} \mathbf{V}^H \mathbf{e}_q, \tag{3.49}$$

where the error vector has been indexed as explained in §3.3.4. For the sake of notation, let us define the objective function for the possible error vectors as $f_q \triangleq \mathbf{e}_q^H \mathbf{V} \mathbf{L} \mathbf{V}^H \mathbf{e}_q$, with $q \leq |\mathcal{E}_T|$. Since there are elements in the set $\mathcal{E}_T$ that are proportional, *i.e.*, $\mathbf{e}_q = z\mathbf{e}_{q'}$, $z \in \mathbb{C}$, it is possible to discard 66 elements from the total of 80 in $\mathcal{E}_T$ because either they yield the same objective function ($|z| = 1$) or because the objective function is always greater in one case ($|z| > 1$). For example, if we let

$$\mathbf{e}_2 = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix}, \mathbf{e}_{15} = \begin{pmatrix} i\sqrt{2} \\ i\sqrt{2} \end{pmatrix}, \text{ and } \mathbf{e}_{16} = \begin{pmatrix} \sqrt{2} - i\sqrt{2} \\ \sqrt{2} - i\sqrt{2} \end{pmatrix}, \tag{3.50}$$

then $\mathbf{e}_{15} = \exp(i\pi/2)\mathbf{e}_2$ and $\mathbf{e}_{16} = (1 + i)\mathbf{e}_2$. From which, clearly, $f_2 = f_{15}$ and $f_2 < f_{16}$, which implies that f_{15} and f_{16} can be discarded as they will have no effect as long as f_2 is present in the inner minimization in (3.49). Repeating this kind of elimination with the remaining elements in $\mathcal{E}_T$ we end up with the set $\mathcal{E}$ of error vectors that yield different values of the objective function in (3.35). In this case, it can be shown that $|\mathcal{E}| = 14$ and, consequently, there are 14 different objective functions, $f_1, f_2, \ldots, f_{14}$.

$$\min_{q \leq |\mathcal{E}|=14} \mathbf{e}_q^H \mathbf{V} \mathbf{L} \mathbf{V}^H \mathbf{e}_q \leq \min_{q \leq |\mathcal{E}_T|} \mathbf{e}_q^H \mathbf{V} \mathbf{L} \mathbf{V}^H \mathbf{e}_q. \tag{3.51}$$

The explicit expressions for these 14 functions are

$$f_1 = 2\left(\ell_1 \sin^2\theta + \ell_2 \cos^2\theta\right), \tag{3.52}$$

$$f_4 = 2\left(\ell_1 \cos^2\theta + \ell_2 \sin^2\theta\right), \tag{3.53}$$

$$f_2 = 2\left(\ell_1 + \ell_2 + 2(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\phi\right), \tag{3.54}$$

$$f_5 = 2\left(\ell_1 + \ell_2 - 2(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\phi\right), \tag{3.55}$$

$$f_6 = 2\left(\ell_1 + \ell_2 + 2(\ell_2 - \ell_1)\sin\theta\cos\theta\sin\phi\right), \tag{3.56}$$

$$f_7 = 2\left(\ell_1 + \ell_2 - 2(\ell_2 - \ell_1)\sin\theta\cos\theta\sin\phi\right), \tag{3.57}$$

$$f_8 = 2\left(\ell_1 + \ell_2 + \ell_1\cos^2\theta + \ell_2\sin^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi + 3\pi/4\right)\right), \tag{3.58}$$

$$f_9 = 2\left(\ell_1 + \ell_2 + \ell_1\cos^2\theta + \ell_2\sin^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi + \pi/4\right)\right), \tag{3.59}$$

$$f_{10} = 2\left(\ell_1 + \ell_2 + \ell_1\cos^2\theta + \ell_2\sin^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - \pi/4\right)\right), \tag{3.60}$$

$$f_{11} = 2\left(\ell_1 + \ell_2 + \ell_1\cos^2\theta + \ell_2\sin^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - 3\pi/4\right)\right), \tag{3.61}$$

$$f_{12} = 2\left(\ell_1 + \ell_2 + \ell_1\sin^2\theta + \ell_2\cos^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi + 3\pi/4\right)\right), \tag{3.62}$$

$$f_{13} = 2\left(\ell_1 + \ell_2 + \ell_1\sin^2\theta + \ell_2\cos^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi + \pi/4\right)\right), \tag{3.63}$$

$$f_3 = 2\left(\ell_1 + \ell_2 + \ell_1\sin^2\theta + \ell_2\cos^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - \pi/4\right)\right), \tag{3.64}$$

$$f_{14} = 2\left(\ell_1 + \ell_2 + \ell_1\sin^2\theta + \ell_2\cos^2\theta + 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - 3\pi/4\right)\right), \tag{3.65}$$

where $\phi = \delta + \epsilon$. Since the only dependence of $\{f_q\}$ on δ and ϵ is through ϕ, the problem now becomes

$$\max_{\boldsymbol{\Sigma}} \max_{\theta,\phi} \min_{q \leq 14} \quad f_q(\theta, \phi), \tag{3.66}$$

where, in this case, the dependence of $\{f_q\}$ on $\boldsymbol{\Sigma}$ is implicit in ℓ_1 and ℓ_2.

In principle, the search space for the inner maximization part in (3.66) is $(\theta, \phi) \in [-\pi, \pi] \times [-\pi, \pi]$, but, from the specific dependence of $\{f_q\}$ on (θ, ϕ) the search space can be reduced to $(\theta, \phi) \in \mathcal{D} \triangleq [0, \pi/4] \times [0, \pi/4]$, and the index function q can be constrained to be not greater than 3, as we show in Appendix 3.B. As a consequence, we obtain an equivalent formulation of (3.35) as

$$\max_{\boldsymbol{\Sigma}} \max_{\theta,\phi} \min_{q \leq 14} \quad f_q(\theta, \phi) = \max_{\boldsymbol{\Sigma}} \max_{(\theta,\phi)\in\mathcal{D}} \min_{q \leq 3} \quad f_q(\theta, \phi). \tag{3.67}$$

Let us define $\alpha \triangleq \ell_2/\ell_1 \in [0, 1]$, and $\alpha_0 \triangleq 3 - 2\sqrt{2}$. We can distinguish two different cases.

- If $\alpha \in [\alpha_0, 1] \triangleq \mathcal{A}_1$, then the two innermost optimization problems in (3.67) can be reduced to

$$\max_{(\theta,\phi)\in\mathcal{D}} \min_{q\leq 3} f_q(\theta, \phi) = \max_{(\theta,\phi)\in\mathcal{D}} f_1(\theta, \phi) = \ell_1 + \ell_2, \tag{3.68}$$

  where the maximum is attained for $\theta^\star = \phi^\star = \pi/4$. See Appendix 3.B for further details.

- If $\alpha \in [0, \alpha_0) \triangleq \mathcal{A}_0$, then, we also show in Appendix 3.B that the two innermost problems in (3.67) can be bounded by

$$\max_{(\theta,\phi)\in\mathcal{D}} \min_{q\leq 3} f_q \leq K(\lambda_c \ell_1 + \ell_2), \tag{3.69}$$

where $\lambda_c \triangleq \dfrac{(\sqrt{3}-1)(3-2\sqrt{2})}{1+3\sqrt{3}-2\sqrt{6}}$, and $K \triangleq \dfrac{1+\alpha_0}{\lambda_c + \alpha_0}$. In this case, instead of maximizing the objective function, we maximize the bound and then prove that the bound is attained by the objective function at the optimal point. The expressions for the possible optimal values of $(\theta^\star, \phi^\star)$ can be found in Appendix 3.B.

This last two cases allow us to split the original problem in (3.67) into two simpler problems. Namely,

$$\max_{\boldsymbol{\Sigma}} \max_{(\theta,\phi)\in\mathcal{D}} \min_{q\leq 3} f_q \begin{cases} = \max_{\boldsymbol{\Sigma}} \ell_1 + \ell_2 & \text{if } \alpha \in \mathcal{A}_1 \\ \leq \max_{\boldsymbol{\Sigma}} K(\lambda_c \ell_1 + \ell_2) & \text{if } \alpha \in \mathcal{A}_0 \end{cases}, \tag{3.70}$$

where the dependence of ℓ_1 and ℓ_2 on $\boldsymbol{\Sigma}$ is

$$\mathbf{L} = \boldsymbol{\Sigma}\boldsymbol{\Lambda_H}\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \lambda_{\mathbf{H},1} & 0 \\ 0 & \lambda_{\mathbf{H},2} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_{\mathbf{H},1}\sigma_1^2 & 0 \\ 0 & \lambda_{\mathbf{H},2}\sigma_2^2 \end{pmatrix} = \begin{pmatrix} \ell_1 & 0 \\ 0 & \ell_2 \end{pmatrix}, \tag{3.71}$$

with $\operatorname{Tr} \boldsymbol{\Sigma}\boldsymbol{\Sigma}^H = \sigma_1^2 + \sigma_2^2 \leq P_T$. For the sake of notation, we define $\gamma = \sigma_1^2 + \sigma_2^2$, $p = \sigma_1^2/\gamma$, and $\lambda = \lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \leq 1$. Since $\lambda_{\mathbf{H},1} \geq \lambda_{\mathbf{H},2}$, it is clear that at the optimal power allocation the inequality $\sigma_1^2 \geq \sigma_2^2$ must be fulfilled as otherwise by switching the values of σ_1^2 and σ_2^2 the value of the objective function in (3.70) could be increased without increasing the transmitted power. With these definitions and considerations, we obtain an expression for ℓ_1 and ℓ_2

$$\begin{aligned} \ell_1 &= \gamma\lambda_{\mathbf{H},1}p, \\ \ell_2 &= \gamma\lambda_{\mathbf{H},1}\lambda(1-p). \end{aligned} \tag{3.72}$$

Consequently the two optimization problems in (3.70) can be further simplified. When $\alpha \in \mathcal{A}_1$ (or, equivalently, when $p \leq \frac{\lambda}{\lambda+\alpha_0}$) we obtain

$$\max_{p,\gamma} \gamma\lambda_{\mathbf{H},1}(p + \lambda(1-p)) \tag{3.73}$$

and when $\alpha \in \mathcal{A}_0$ ($p \geq \frac{\lambda}{\lambda+\alpha_0}$) the bound becomes

$$\max_{p,\gamma} K\gamma\lambda_{\mathbf{H},1}(\lambda_c p + \lambda(1-p)). \tag{3.74}$$

Notice that, both problems in (3.73) and (3.74), are increasing functions of γ, then it is clear that $\gamma = \sigma_1^2 + \sigma_2^2 \leq P_T$, will take its maximum allowed value, *i.e.*, $\gamma = P_T$. It now only remains to find the optimal value for the parameter p, which we recall is constrained to belong to $p \in [1/2, 1]$.

Thus, the problem in (3.73) becomes

$$\max_{p} \quad p + \lambda - \lambda p = p(1 - \lambda) + \lambda$$
$$\text{s.t.} \quad p \in \left[\frac{1}{2}, \frac{\lambda}{\lambda + \alpha_0}\right], \tag{3.75}$$

where the restriction on p is due to the fact that $\alpha \in \mathcal{A}_1$ must be guaranteed. Since the factor $(1 - \lambda)$ in the objective function in (3.75) is bigger than zero, the optimum will be attained when p equals its maximum value, *i.e.*, $p^\star = \frac{\lambda}{\lambda + \alpha_0}$.

Similarly, the problem in (3.74) can be written equivalently as

$$\max_{p} \quad \lambda_c p + \lambda(1 - p) = (\lambda_c - \lambda)p + \lambda,$$
$$\text{s.t.} \quad p \in \left(\frac{\lambda}{\lambda + \alpha_0}, 1\right], \tag{3.76}$$

where, in this case the restriction on p is to guarantee that $\alpha \in \mathcal{A}_0$. The solution to the maximization depends on a condition on λ. On one hand, if $\lambda < \lambda_c$, then the optimum is $p^\star = 1$. On the other hand, if $\lambda \geq \lambda_c$, the optimum is attained by $p^\star \to \frac{\lambda}{\lambda + \alpha_0}$, which coincides with the optimum solution for the problem in (3.73) and which implies that, when $\lambda \geq \lambda_c$, the solution lies in the boundary of the two regions, *i.e.*, $\alpha = \alpha_0$. Finally, we recall that what we obtained is just the maximization of the bound in (3.69), we now check that the bound is actually attained by the objective function.

- $\lambda < \lambda_c$: In this case we know that $p^\star = 1$, from (3.72) this implies that $\alpha^\star = 0$. Particularizing in (3.111) and (3.112) for $\alpha^\star = 0$ we obtain

$$\theta^\star = \arctan\left(\sqrt{6}/(3 + \sqrt{3})\right), \tag{3.77}$$
$$\phi^\star = \arccos\left((3 + \sqrt{3})/(2\sqrt{6})\right), \tag{3.78}$$

  and evaluating f_q, for $q = \{1, 2, 3\}$, at the point $(\theta^\star, \phi^\star, p^\star)$ we obtain

$$\min_{q \leq 3} f_q(\theta^\star, \phi^\star, p^\star) = 2\left(1 - \frac{1}{\sqrt{3}}\right) P_T \lambda_{\mathbf{H},1} = K \lambda_c P_T \lambda_{\mathbf{H},1} =$$
$$= K(\lambda_c \ell_1 + \ell_2)|_{p^\star = 1} \geq K(\lambda_c \ell_1 + \ell_2) \geq \min_{q \leq 3} f_q, \tag{3.79}$$

  where last inequality follows from the fact that $K(\lambda_c \ell_1 + \ell_2)$ is a bound for $\min_{q \leq 3} f_q$.

- $\lambda \geq \lambda_c$: In this case we know that $p^\star \to \frac{\lambda}{\lambda + \alpha_0}$, from 3.72 this implies that $\alpha^\star \to \alpha_0$. Particularizing in (3.111) and (3.112) for $\alpha^\star = \alpha_0$ we obtain

$$\theta^\star = \pi/4, \tag{3.80}$$
$$\phi^\star = \pi/4, \tag{3.81}$$

and evaluating f_q, for $q = \{1, 2, 3\}$, at the point $(\theta^\star, \phi^\star, p^\star)$ we obtain

$$
\min_{q \leq 3} f_q(\theta^\star, \phi^\star, p^\star) = \frac{\lambda}{\lambda + \alpha_0}(1 + \alpha_0)P_T \lambda_{\mathbf{H},1} = K\ell_1(\lambda_c + \alpha_0) =
$$
$$
= K(\lambda_c \ell_1 + \ell_2)|_{p^* = \frac{\lambda}{\lambda + \alpha_0}} \geq K(\lambda_c \ell_1 + \ell_2) \geq \min_{q \leq 3} f_q, \quad (3.82)
$$

where, as before, $K(\lambda_c \ell_1 + \ell_2)|_{p^\star}$ is the bound in (3.70) evaluated at the argument $p^\star$ in (3.80) and where last inequality follows from the fact that $K(\lambda_c \ell_1 + \ell_2)$ is a bound for $\min_{q \leq 3} f_q$.

Now that we have obtained the optimal values for all the parameters, $\theta^\star$, $\phi^\star$, and $p^\star$, for both problems in (3.73) and (3.74) and that have proved that the objective function in (3.70) attains its bound when $\alpha \in \mathcal{A}_0$, it remains to evaluate the objective function in (3.70) and calculate in each case which function yields the greatest value.

From all the previous obtained results we can write

$$
\max_p \max_{(\theta,\phi) \in \mathcal{D}} \min_{q \leq 3} f_q =
\begin{cases}
\dfrac{\lambda}{\lambda + \alpha_0}(1 + \alpha_0)P_T \lambda_{\mathbf{H},1} & \alpha = \alpha_0 \in \mathcal{A}_1 \\
K\lambda_c P_T \lambda_{\mathbf{H},1} & \alpha = 0 \in \mathcal{A}_0
\end{cases}. \quad (3.83)
$$

Clearly, the case that will yield the greatest value depends on

$$
\frac{\lambda}{\lambda + \alpha_0}(1 + \alpha_0) \gtrless K\lambda_c \Leftrightarrow \lambda \gtrless \lambda_c. \quad (3.84)
$$

Finally, recalling that $\lambda = \lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1}$, we obtain

$$
\max_p \max_{(\theta,\phi) \in \mathcal{D}} \min_{q \leq 3} f_q =
\begin{cases}
\dfrac{\lambda_{\mathbf{H},1}\lambda_{\mathbf{H},2}}{\lambda_{\mathbf{H},2} + \alpha_0 \lambda_{\mathbf{H},1}}(1 + \alpha_0)P_T & \lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \geq \lambda_c, \\
K\lambda_c P_T \lambda_{\mathbf{H},1} & \lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c.
\end{cases} \quad (3.85)
$$

To sum up, we have obtained that

- For $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} < \lambda_c$ the solution to the optimization problem in (3.35) is parameterized by $p^\star = 1$, $\theta^\star = \theta_T = \arctan\left(\sqrt{6}/(3 + \sqrt{3})\right)$, and $\phi^\star = \phi_T = \arccos\left((3 + \sqrt{3})/(2\sqrt{6})\right)$. The optimal transmitter is then

$$
\mathbf{\Sigma}^\star = \sqrt{P_T}\begin{pmatrix} \sqrt{p^\star} & 0 \\ 0 & \sqrt{1 - p^\star} \end{pmatrix} = \begin{pmatrix} \sqrt{P_T} & 0 \\ 0 & 0 \end{pmatrix} \quad (3.86)
$$

$$
\mathbf{V}^\star = \begin{pmatrix} \cos\theta^\star & \mathrm{e}^{\mathrm{i}\phi^\star}\sin\theta^\star \\ -\mathrm{e}^{-\mathrm{i}\phi^\star}\sin\theta^\star & \cos\theta^\star \end{pmatrix} = \begin{pmatrix} \cos\theta_T & \mathrm{e}^{\mathrm{i}\phi_T}\sin\theta_T \\ -\mathrm{e}^{-\mathrm{i}\phi_T}\sin\theta_T & \cos\theta_T \end{pmatrix} \quad (3.87)
$$

$$
\mathbf{B}^\star = \mathbf{U_H}\mathbf{\Sigma}^\star\mathbf{V}^{\star H} = \mathbf{U_H}\begin{pmatrix} \sqrt{P_T} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \cos\theta_T & -\mathrm{e}^{\mathrm{i}\phi_T}\sin\theta_T \\ \mathrm{e}^{-\mathrm{i}\phi_T}\sin\theta_T & \cos\theta_T \end{pmatrix} \quad (3.88)
$$

- For $\lambda_{\mathbf{H},2}/\lambda_{\mathbf{H},1} \geq \lambda_c$ the solution to the optimization problem in (3.35) is parameterized by $p^\star = \frac{\lambda}{\lambda+\alpha_0} = \frac{\lambda_{\mathbf{H},2}}{\lambda_{\mathbf{H},2}+\alpha_0\lambda_{\mathbf{H},1}}$, $\theta^\star = \pi/4$, and $\phi^\star = \pi/4$. The optimal transmitter is then

$$\mathbf{\Sigma}^\star = \sqrt{P_T} \begin{pmatrix} \sqrt{p^\star} & 0 \\ 0 & \sqrt{1-p^\star} \end{pmatrix} = \sqrt{P_T} \begin{pmatrix} \sqrt{\frac{\lambda_{\mathbf{H},2}}{\lambda_{\mathbf{H},2}+\alpha_0\lambda_{\mathbf{H},1}}} & 0 \\ 0 & \sqrt{\frac{\alpha_0\lambda_{\mathbf{H},1}}{\lambda_{\mathbf{H},2}+\alpha_0\lambda_{\mathbf{H},1}}} \end{pmatrix} \tag{3.89}$$

$$\mathbf{V}^\star = \begin{pmatrix} \cos\theta^\star & \mathrm{e}^{\mathrm{i}\phi^\star}\sin\theta^\star \\ -\mathrm{e}^{-\mathrm{i}\phi^\star}\sin\theta^\star & \cos\theta^\star \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \mathrm{e}^{\mathrm{i}\pi/4} \\ -\mathrm{e}^{-\mathrm{i}\pi/4} & 1 \end{pmatrix} \tag{3.90}$$

$$\mathbf{B}^\star = \mathbf{U_H}\mathbf{\Sigma}^\star\mathbf{V}^{\star H} = \mathbf{U_H}\sqrt{\frac{P_T}{2(\lambda_{\mathbf{H},2}+\alpha_0\lambda_{\mathbf{H},1})}} \begin{pmatrix} \sqrt{\lambda_{\mathbf{H},2}} & 0 \\ 0 & \sqrt{\alpha_0\lambda_{\mathbf{H},1}} \end{pmatrix} \begin{pmatrix} 1 & -\mathrm{e}^{\mathrm{i}\pi/4} \\ \mathrm{e}^{-\mathrm{i}\pi/4} & 1 \end{pmatrix} \tag{3.91}$$

3.B   Analysis of the minimum distance function candidates

Throughout this appendix, the matrix $\mathbf{\Sigma}$ will be considered fixed and, consequently, ℓ_1 and ℓ_2, which depend directly on $\mathbf{\Sigma}$ will also be fixed. In addition, by the definition in (3.72) we recall that $\ell_1 \geq \ell_2$ and we define $\mathcal{J} = [0, \pi/4]$, $\mathcal{P} = (-\pi, \pi]$, and $\mathcal{D} = \mathcal{J}^2$.

We first assume that $\phi \in \mathcal{J}$. Then it follows

$$\min_{q=\{1,4\}} f_q = \ell_1 + \ell_2 - (\ell_1 - \ell_2)\left|\cos 2\theta\right|, \tag{3.92}$$

$$\min_{q=\{2,5,6,7\}} f_q = 2\left(\ell_1 + \ell_2 - (\ell_1 - \ell_2)\cos\phi\left|\sin 2\theta\right|\right), \tag{3.93}$$

$$\min_{q=\{3,8,\ldots,14\}} f_q = 3(\ell_1 + \ell_2) - (\ell_1 - \ell_2)\left|\cos 2\theta\right| - 2\sqrt{2}(\ell_1 - \ell_2)\cos\phi\left|\sin 2\theta\right|. \tag{3.94}$$

These three functions are periodic functions with period $T_\theta = \pi/2$ and symmetric, inside k-th period with respect to $\theta_k = \pi/4 + k\pi/2$.

Secondly, we assume that $\theta \in \mathcal{J}$. In this case it follows

$$\min_{q=\{2,5,6,7\}} f_q = 2\left(\ell_1 + \ell_2 + 2(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - k_1\pi/2\right)\right), \tag{3.95}$$

$$\min_{q=\{3,8,\ldots,14\}} f_q = 2(\ell_1 + \ell_2 + \ell_1\sin^2\theta + \ell_2\cos^2\theta+ \tag{3.96}$$

$$+ 2\sqrt{2}(\ell_2 - \ell_1)\sin\theta\cos\theta\cos\left(\phi - \pi/4 - k_2\pi/2\right), \tag{3.97}$$

where $k_1 = \lfloor 2\phi/\pi + 1/2 \rfloor$ and $k_2 = \lfloor 2\phi/\pi \rfloor$. Notice that these two functions are periodic functions with period $T_\phi = \pi/2$ and symmetric, inside k-th period with respect to $\phi_k = \pi/4 + k\pi/2$.

From the double periodicity in θ and ϕ, and the double symmetry with respect to $\pi/4 + k\pi/2$, it can then be directly obtained that the search space can be reduced from $\mathcal{P}^2$ to $\mathcal{D}$.

In addition, it can be seen by simple inspection from (3.92), (3.93), (3.94), (3.95), and (3.97)

that $\forall(\theta, \phi) \in \mathcal{D}$ the following equalities hold

$$f_1 = \min_{q=\{1,4\}} f_q, \tag{3.98}$$

$$f_2 = \min_{q=\{2,5,6,7\}} f_q, \tag{3.99}$$

$$f_3 = \min_{q=\{3,8,...,14\}} f_q. \tag{3.100}$$

Consequently, the original problem in (3.66) can be simplified to

$$\max_{\tilde{\mathbf{U}}, \mathbf{\Sigma}} \max_{\theta, \phi} \min_{q \leq 14} \quad f_q(\theta, \phi) = \max_{\tilde{\mathbf{U}}, \mathbf{\Sigma}} \max_{(\theta, \phi) \in \mathcal{D}} \min_{q \leq 3} \quad f_q(\theta, \phi). \tag{3.101}$$

Now, we have to determine under which conditions each of the three functions yields the minimum value for a given point (θ, ϕ). To do so, we find the domain regions where each function will take the minimum value. Each of these regions will be denoted by $\mathcal{R}_i \subseteq \mathcal{D}$.

$$(\theta, \phi) \in \mathcal{R}_i \Rightarrow f_i(\theta, \phi) \leq f_j(\theta, \phi), \quad j \neq i \tag{3.102}$$

In order to find the frontiers of these regions, let us define (θ_{pq}, ϕ_{pq}) as the set of points such that $f_p(\theta_{pq}, \phi_{pq}) = f_q(\theta_{pq}, \phi_{pq})$. From the expressions in (3.52), (3.54), (3.64), they must fulfill

$$\cos \phi_{12} = \frac{\ell_1 \cos^2 \theta_{12} + \ell_2 \sin^2 \theta_{12}}{2(\ell_1 - \ell_2) \sin \theta_{12} \cos \theta_{12}}, \tag{3.103}$$

$$\cos (\phi_{13} - \pi/4) = \frac{\ell_1 + \ell_2}{2\sqrt{2}(\ell_1 - \ell_2) \sin \theta_{13} \cos \theta_{13}}, \tag{3.104}$$

$$\sin \phi_{23} = \frac{\ell_1 \sin^2 \theta_{23} + \ell_2 \cos^2 \theta_{23}}{2(\ell_1 - \ell_2) \sin \theta_{23} \cos \theta_{23}}. \tag{3.105}$$

Recalling that $\alpha \triangleq \ell_2/\ell_1 \in [0, 1)$, and $\alpha_0 \triangleq 3 - 2\sqrt{2}$ in Figure 3.13 we have plotted the frontiers of each region for diferent values of α. Concerning the frontiers of the three regions $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$, we can distinguish three different cases.

- $\alpha \in (1/3, 1]$: There exist no solutions in $\mathcal{D}$ for equations in (3.103)-(3.105), thus, there are no intersections between any of the three functions. It is straightforward to check that, since the maximum value of function f_1 is lower than the minimum of the values of functions f_2 and f_3, in this case $\mathcal{R}_1 = \mathcal{D}$. The optimization problem is solved by

$$\max_{(\theta, \phi) \in \mathcal{D}} \min_q f_q(\theta, \phi) = \max_{(\theta, \phi) \in \mathcal{D}} f_1(\theta, \phi) = \ell_1 + \ell_2, \tag{3.106}$$

where the maximum is attained for $\theta^\star = \pi/4$, $\forall \phi$.

- $\alpha \in [\alpha_0, 1/3]$: There exist solutions in $\mathcal{D}$ only for equations (3.103) and (3.105), thus there are intersections between f_1 and f_2 and between f_2 and f_3. Since there are no intersections
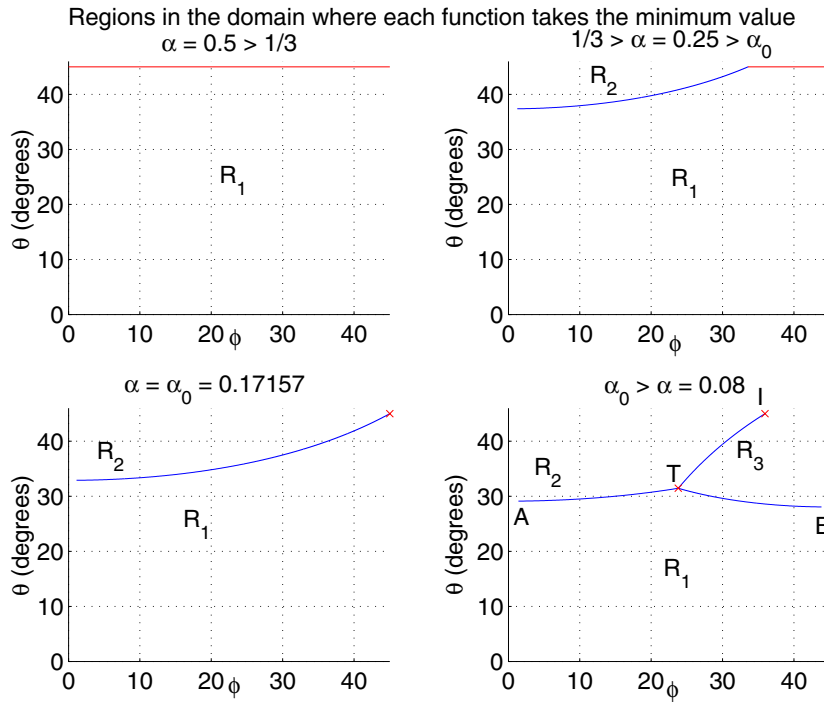
Figure 3.13: Partitions of $\mathcal{D}$ in regions $\mathcal{R}_i$ such that $f_i \leq f_j$. The solid lines represents the boundary between each region. The red lines (or marks) represent possible candidates to the maximum value of the minimum of the three functions.

between f_1 and f_3, in this case $\mathcal{R}_3 = \emptyset$, thus $\mathcal{R}_1 \cup \mathcal{R}_2 = \mathcal{D}$. In addition, the maximum value of f_1 is attained by $\theta = \pi/4$, for all ϕ, but in this case the intersection with f_2 clips part of this zone where the maximum is attained and reduces it to $\theta = \pi/4$ and $\phi \in \left[\arccos\left(\frac{1+\alpha}{2(1-\alpha)}\right), \pi/4 \right]$, see Figure 3.13. Finally, we obtain

$$\max_{(\theta,\phi)\in\mathcal{D}} \ \min_q \ f_q(\theta,\phi) = \max_{(\theta,\phi)\in\mathcal{D}} \ f_1(\theta,\phi) = \ell_1 + \ell_2, \tag{3.107}$$

where the maximum is attained for $\theta^\star = \pi/4$, and $\forall \phi \in \left[\arccos\left(\frac{1+\alpha}{2(1-\alpha)}\right), \pi/4 \right]$.

Since the expression for $\max_{\theta,\phi} \ \min_q \ f_q(\theta,\phi)$ is the same for this case than for the previous one, we can merge these two cases, obtaining

$$\alpha \in [\alpha_0, 1/3] \cup (1/3, 1] = [\alpha_0, 1] \Rightarrow \max_{(\theta,\phi)\in\mathcal{D}} \ \min_q \ f_q(\theta,\phi) = \ell_1 + \ell_2, \tag{3.108}$$

where $\theta^\star = \pi/4$ and $\phi^\star = \pi/4$ is a solution $\forall \alpha \in [\alpha_0, 1]$. Notice that $\arccos\left(\frac{1+\alpha_0}{2(1-\alpha_0)}\right) = \pi/4$.

- $\alpha \in [0, \alpha_0)$: There are intersections between all possible pairs of functions. See Figure 3.13 for details. The region $\mathcal{R}_1$ is delimited by the points A, T, and B; the region $\mathcal{R}_2$, by A, T, and I; and the region $\mathcal{R}_3$, by I, T, and B. Noteworthy, the point-T, which we call *triple point*, (θ_T, ϕ_T), is the intersection between all the three functions. Since f_1 is an increasing function of θ, f_2 is an increasing function of ϕ, and f_3 is a decreasing function

of ϕ, the point $(\theta^\star, \phi^\star)$ that maximizes the minimum of the three functions must lie in the intersection between f_2 and f_3 (The curve that goes from I to T in Figure 3.13). If we plug (3.105) in (3.54), it can be seen that the function is continuous and that there is only a local minimum. Consequently, the maximum in the curve has to be in one of the two extremes, namely the point-T or the point-I. So, the two candidate points for being the location of the maximum $(\theta^\star, \phi^\star)$ are the triple intersection point (θ_T, ϕ_T) and the point in the intersection between f_2 and f_3, (θ_I, ϕ_I). We give here the expressions for (θ_I, ϕ_I) and (θ_T, ϕ_T):

$$\theta_I = \frac{\pi}{4}, \tag{3.109}$$

$$\phi_I = \arcsin\left(\frac{1+\alpha}{2(1-\alpha)}\right), \tag{3.110}$$

$$\theta_T = \arctan\left(\sqrt{\frac{3 - 3\alpha - \sqrt{3}\sqrt{\alpha^2 - 6\alpha + 1}}{3 - 3\alpha + \sqrt{3}\sqrt{\alpha^2 - 6\alpha + 1}}}\right), \tag{3.111}$$

$$\phi_T = \arccos\left(\frac{\alpha\cos^2\theta_T + \sin^2\theta_T - 1 - \alpha}{2\sin\theta_T\cos\theta_T(\alpha - 1)}\right). \tag{3.112}$$

Finally, evaluating the functions at the two candidate points we obtain

$$f_{2,3}(\theta_I, \phi_I) = 2\ell_1\left(1 + \alpha - \frac{1}{2}\sqrt{3\alpha^2 - 10\alpha + 3}\right), \tag{3.113}$$

$$f_{1,2,3}(\theta_T, \phi_T) = \ell_1\left(1 + \alpha - \frac{1}{\sqrt{3}}\sqrt{\alpha^2 - 6\alpha + 1}\right). \tag{3.114}$$

We present now a linear upper bound, on both functions in (3.113) and (3.114). The bound is found by fixing ℓ_1 and by constructing the line that goes from $f_{2,3}(\theta_I, \phi_I)$ evaluated at $\alpha = \alpha_0$ to $f_{1,2,3}(\theta_T, \phi_T)$ evaluated at $\alpha = 0$. See Figure 3.14.

The bound is

$$K\ell_1(\lambda_c + \alpha) = K(\lambda_c\ell_1 + \ell_2), \tag{3.115}$$

where

$$\lambda_c \triangleq \frac{(\sqrt{3} - 1)(3 - 2\sqrt{2})}{1 + 3\sqrt{3} - 2\sqrt{6}} \simeq 9.683 \cdot 10^{-2}, \quad K \triangleq \frac{1 + \alpha_0}{\lambda_c + \alpha_0} \simeq 4.3650. \tag{3.116}$$
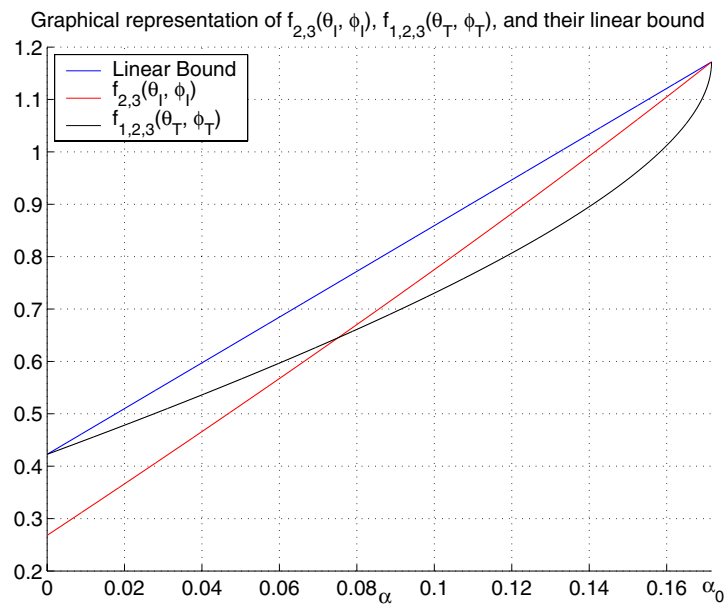
Figure 3.14: Graphical representation of $f_{2,3}(\theta_I, \phi_I)$, $f_{1,2,3}(\theta_T, \phi_T)$, and their linear bound, for ℓ_1 fixed as a function of α according to (3.113), (3.114), and (3.115).

Chapter 4

Single-user communication through MIMO channels with incomplete CSI

4.1 Introduction

In this chapter we study and characterize different mutual information and capacity formulations for a very specific case of incomplete CSI at the transmitter side and perfect CSI at the receiver side.

As commented in chapter 2, in the existing literature, a paradigmatic example of study for a scenario with incomplete CSI is the case of knowing only the channel statistics, but not the actual realization. According to this, the transmit covariance matrix has to be designed to maximize the mutual information averaged over these statistics, *i.e.*, the so-called *ergodic capacity*.

In the following pages a different approach is taken by considering a more engineering perspective. Here it is considered that the transmitter has perfect knowledge of the magnitude of the complex channel coefficients, but a complete lack of knowledge of the channel phases. This uncertainty model encompasses many practical situations of interest. As an illustrative example, consider the case of TDD schemes, where the transmitter can estimate the channel during the uplink and use it as it was the same in the downlink thanks to the electromagnetic reciprocity principle. Note, however, that this is not true since the uplink and downlink channels are seen through different RF chains. Although calibration methods are able to compensate the gains, the phases are much more difficult to be estimated and compensated.

As it is commented in §2.3.2, it is widely known that, in case of having additive Gaussian noise at the receiver, the optimum transmitted signal has to be zero-mean Gaussian distributed if the mutual information is to be maximized. Thus, the only remaining degree of freedom when optimizing the system is the transmit covariance matrix. In this chapter, a design of the transmit covariance matrix to achieve the ergodic and compound capacities is proposed under

the already mentioned uncertainty model. Unfortunately, closed-form solutions do not exist and approximations have to be applied. For both cases, a numerical optimization method is proposed based on the use of a finite set of random realizations of the channel phases, although it may require a high computational load. Additionally, for the case of the ergodic capacity, an approximate method with lower complexity is also given based on RMT [Tul04], whereas for the case of the compound capacity, an approximate finite term Taylor expansion of the mutual information is exploited. In this sense, one of the major contributions of the present chapter with respect to the existing literature, is the application of RMT not only for the evaluation but also for the optimization of the ergodic mutual information for the case of magnitude knowledge and phase uncertainty.

4.2   System model

We stay with the flat-fading MIMO channel model presented in the preceding chapter and which is reviewed in §2.1.4. We recall it here, for the sake of completeness. The received signal vector $\mathbf{y} \in \mathbb{C}^{n_R}$ is given by

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

where $\mathbf{x} \in \mathbb{C}^{n_T}$ is the transmitted signal vector, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ represents the channel matrix, and $\mathbf{n} \in \mathbb{C}^{n_R}$ is the noise vector. In this case, the entries of the noise vector are considered to be proper complex Gaussian random variables with $\mathbb{E}\mathbf{n} = \mathbf{0}$ and $\mathbb{E}\mathbf{n}\mathbf{n}^H = \sigma^2\mathbf{I}$.

As explained in the introduction, we consider that, while the receiver is fully cognizant of the channel state, the transmitter has only knowledge about the magnitude of the entries of the channel matrix, and a complete lack of knowledge about the actual value of their phases. To separate the known from the unknown part of the channel entries, the following variables are defined:

$$[\mathbf{H}]_{ij} = m_{ij}\, \mathrm{e}^{\mathrm{i}\theta_{ij}}, \quad m_{ij} \in \mathbb{R}^+ \cup \{0\}, \quad \theta_{ij} \in [0, 2\pi), \tag{4.2}$$

where θ_{ij} are i.i.d. random variables uniformly distributed in $[0, 2\pi)$ as suggested by, *e.g.*, [Jak74].

By defining the matrices $\mathbf{M}$ and $\mathbf{P}$ such that $[\mathbf{M}]_{ij} = m_{ij}$ and $[\mathbf{P}]_{ij} = \mathrm{e}^{\mathrm{i}\theta_{ij}}$, the uncertainty model described in (4.2) can be compactly rewritten as

$$\mathbf{H} = \mathbf{M} \odot \mathbf{P}, \tag{4.3}$$

where $\odot$ represents the Hadamard element-wise matrix product.

4.3   Instantaneous mutual information

The capacity achieving strategy in MIMO channels with AWGN is signaling using random Gaussian vectors as shown in [Cov91] and commented in chapter 2. In this case, the instantaneous mutual information for a fixed transmit covariance matrix, $\mathbf{Q} = \mathbb{E}\mathbf{x}\mathbf{x}^H$, and channel state as defined by $\{\mathbf{M}, \mathbf{P}\}$ is the particularization of the expression (2.16) in §2.3.2 for the uncertainty model considered here

$$\Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}) = \log \det \left(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{Q}(\mathbf{M} \odot \mathbf{P})^H \right), \tag{4.4}$$

where the transmitted power is given by $\mathbb{E}\mathbf{x}^H\mathbf{x} = \operatorname{Tr}\mathbf{Q} = P_T$. Recall from the previous chapter that utilizing the SVD the matrix $\mathbf{Q}$ can be factored as $\mathbf{Q} = \mathbf{U_Q}\mathbf{\Lambda_Q}\mathbf{U_Q}^H$, with $\mathbf{U_Q}$ being a unitary matrix with the eigenvectors of $\mathbf{Q}$ and $\mathbf{\Lambda_Q}$ being a diagonal matrix containing the eigenvalues of $\mathbf{Q}$ which are given by $\lambda_{\mathbf{Q},j} = [\mathbf{\Lambda_Q}]_{jj}$.

In the following sections, we particularize the definitions of ergodic and compound mutual information for the uncertainty model considered in this chapter and explain in which situations these measures become meaningful and how can they be maximized.

4.4   Ergodic mutual information and capacity

The information theoretic community defines the ergodic mutual information as the expectation with respect to the channel state uncertainty of the instantaneous mutual information in (4.4) as described in (2.17). In our case, since the channel uncertainty is associated with the channel phases $\mathbf{P}$, the expression for the ergodic mutual information particularizes to

$$I_E(\mathbf{Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}}\Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}). \tag{4.5}$$

Note that, by taking the expectation with respect to $\mathbf{P}$, the ergodic mutual information $I_E(\mathbf{Q}, \mathbf{M})$ does not depend on $\mathbf{P}$.

The ergodic capacity is then defined as the supremum of the ergodic mutual information with respect to the set of possible covariance matrices $\mathbf{Q}$, subject to a mean transmitted power constraint. Formally, this can be written as

$$\begin{aligned} C_E = \sup_{\mathbf{Q}} \quad & I_E(\mathbf{Q}, \mathbf{M}) \\ \text{s.t.} \quad & \operatorname{Tr}\mathbf{Q} \le P_T, \\ & \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{4.6}$$

The ergodic capacity is utilized as a measure of the maximum rates that can be achieved in situations where, during the transmission of the message, the magnitude of the channel
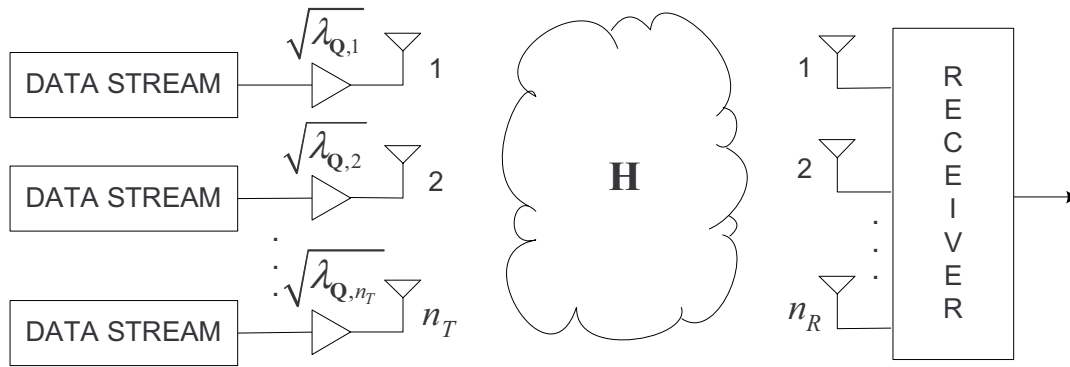
Figure 4.1: Optimum transmitter architecture for the case of complete phase unknowledge of the channel gains. Both for the cases of ergodic and compound capacities, the optimum signalling consists in transmitting independent signals through different transmit antennas with an appropriate power allocation.

matrix, $\mathbf{M}$, remains constant while the channel phases, $\mathbf{P}$, vary sufficiently fast so that its long-term properties are revealed. This model would correspond, for example, to a communication situation with direct line of sight, where the mobile user is moving slowly. In this scenario, while the magnitude of the entries of $\mathbf{M}$ would not change in an appreciable way, the phases in $\mathbf{P}$ would vary very rapidly because of the relative movement.

The optimal structure of the covariance matrix that achieves the ergodic capacity is given next and is depicted in Figure 4.1.

Proposition 4.4.1 *Assume instantaneous perfect CSI at the receiver and magnitude CSI at the transmitter, $\mathbf{M}$. If the channel state, defined as $\mathbf{H} = \mathbf{M} \odot \mathbf{P}$, with $[\mathbf{P}]_{ij} = \mathrm{e}^{\mathrm{i}\,\theta_{ij}}$, is a random matrix whose elements' phases, θ_{ij}, are i.i.d. random variables drawn from a uniform distribution in the interval $[-\pi, \pi)$, then the ergodic capacity of the Gaussian memoryless channel (4.1), subject to a transmit power constraint P_T, can only be achieved by a diagonal covariance matrix $\mathbf{Q}^\star = \mathbf{\Lambda_Q}$, i.e., $\mathbf{U_Q^\star} = \mathbf{I}$. The ergodic capacity is then given by the solution to*

$$
\begin{aligned}
C_E = \sup_{\mathbf{\Lambda_Q}} \quad & I_E(\mathbf{\Lambda_Q}, \mathbf{M}) \\
\mathrm{s.\,t.} \quad & \mathrm{Tr}\, \mathbf{\Lambda_Q} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \ [\mathbf{\Lambda_Q}]_{jj'} = 0, \quad 1 \leq j, j' \leq n_T.
\end{aligned} \tag{4.7}
$$

Proof See Appendix 4.A.   ∎

Because of the difficulty of dealing with the expectation with respect to the channel phases in the expression for $I_E(\mathbf{\Lambda_Q}, \mathbf{M})$, an expression for the optimal values of the diagonal elements of $\mathbf{\Lambda_Q}$ has only been found for the $n_T = n_R = 2$ case [Pay04b, Pay04a], as given in the following section.

Once the exact solution for the $n_T = n_R = 2$ case is presented, in the following sections we derive two methods to numerically calculate an approximation to the optimal values of the diagonal of $\mathbf{\Lambda_Q}$ for any possible values of n_T and n_R.

4.4.1 Solution to the particular case $n_T = n_R = 2$.

As a particular case from what has been proven in the last section, we can state that the capacity of a 2×2 MIMO system with phase uncertainty at the transmitter can be characterized with the following maximization:

$$
\begin{aligned}
C_E = \sup_{\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}} \quad & I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M}) \\
\text{s.\,t.} \quad & \lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \quad 1 \leq j \leq 2.
\end{aligned}
\tag{4.8}
$$

where $I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M})$ is the same as defined in (4.7) with the positive weights $\lambda_{\mathbf{Q},1}$ and $\lambda_{\mathbf{Q},2}$ denoting the diagonal elements of $\mathbf{\Lambda_Q}$. Equivalently, $I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M})$ can be expressed as

$$
I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \log \det \left(\mathbf{I} + \sigma^{-2} \lambda_{\mathbf{Q},1} \mathbf{h}_1 \mathbf{h}_1^H + \sigma^{-2} \lambda_{\mathbf{Q},2} \mathbf{h}_2 \mathbf{h}_2^H \right),
\tag{4.9}
$$

where the dependence of $\mathbf{h}_j$, $j = 1, 2$, on $\mathbf{P}$ has not been written explicitly. For the sake of simplicity we define,

$$
T = 2\lambda_{\mathbf{Q},1} \lambda_{\mathbf{Q},2} \sigma^{-4} m_{11} m_{12} m_{21} m_{22},
\tag{4.10}
$$
$$
S = 1 + \lambda_{\mathbf{Q},1} \sigma^{-2} \left(m_{11}^2 + m_{21}^2 \right) + \lambda_{\mathbf{Q},2} \sigma^{-2} \left(m_{12}^2 + m_{22}^2 \right) + \lambda_{\mathbf{Q},1} \lambda_{\mathbf{Q},2} \sigma^{-4} \left(m_{11}^2 m_{22}^2 + m_{12}^2 m_{21}^2 \right),
\tag{4.11}
$$
$$
\phi = \theta_{21} - \theta_{11} + \theta_{12} - \theta_{22}.
\tag{4.12}
$$

Notice that ϕ is also uniformly distributed in $[-\pi, \pi)$. With the last definitions and expanding the determinant in (4.9), $I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M})$ can be expressed as

$$
I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M}) = \log S + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left(1 - \frac{T \cos \phi}{S} \right) \mathrm{d}\phi.
\tag{4.13}
$$

This last integral can be solved using [Gra00, p.526]

$$
\frac{1}{\pi} \int_0^{\pi} \log \left(1 + \beta \cos x \right) \mathrm{d}x = \log \left(\tfrac{1}{2} \left(1 + \sqrt{1 - \beta^2} \right) \right).
\tag{4.14}
$$

Thus, the mutual information in (4.13) can be expressed as

$$
I_E(\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}, \mathbf{M}) = \log \left(\tfrac{1}{2} \left(S + \sqrt{S^2 - T^2} \right) \right).
\tag{4.15}
$$

Once we have obtained an explicit expression for the ergodic mutual information, I_E, we can characterize the capacity of the system and the optimum power allocation scheme by carrying out

the maximization of (4.15) with respect to the positive weights $\lambda_{\mathbf{Q},1}$, $\lambda_{\mathbf{Q},2}$ under the constraint $\lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} = P_T$. Consequently, the optimization problem can be reformulated as follows,

$$
\begin{aligned}
\{\lambda^{\star}_{\mathbf{Q},1}, \lambda^{\star}_{\mathbf{Q},2}\} = \arg \sup_{\lambda_{\mathbf{Q},1},\lambda_{\mathbf{Q},2}} \quad & \log\left(\tfrac{1}{2}\left(S + \sqrt{S^2 - T^2}\right)\right) \\
\text{s.t.} \quad & \lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} = P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \quad 1 \leq j \leq 2,
\end{aligned}
\tag{4.16}
$$

where the dependence of the objective function on $\lambda_{\mathbf{Q},1}$ and $\lambda_{\mathbf{Q},2}$ is through S and T. We make the slight abuse of notation and denote the optimization variables and their optimum values with the same symbol. It is easy to see that this optimization problem is convex. In order to investigate the properties of the optimum solution, let us first formulate the Lagrangian of the corresponding minimization problem,

$$
\mathcal{L}(\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}) = -S - \sqrt{S^2 - T^2} - \mu_1 \lambda_{\mathbf{Q},1} - \mu_2 \lambda_{\mathbf{Q},2} + \nu\left(\lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} - P_T\right).
\tag{4.17}
$$

Imposing the KKT conditions, we see that, the following are necessary and sufficient conditions for $\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}$ to be the global maximum of (4.16)

$$
\lambda_{\mathbf{Q},j} \geq 0,
\tag{4.18}
$$

$$
\mu_j \geq 0,
\tag{4.19}
$$

$$
\lambda_{\mathbf{Q},j}\mu_j \geq 0,
\tag{4.20}
$$

$$
\lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} - P_T = 0,
\tag{4.21}
$$

$$
\frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},j}}\left(S + \sqrt{S^2 - T^2}\right) + \mu_j - \nu = 0,
\tag{4.22}
$$

for $j = 1, 2$. Now, observe that μ_1 and μ_2 are just slack variables that can be eliminated, leaving

$$
\frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},j}}\left(S + \sqrt{S^2 - T^2}\right) \leq \nu,
\tag{4.23}
$$

$$
\lambda_{\mathbf{Q},j}\left(\nu - \frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},j}}\left(S + \sqrt{S^2 - T^2}\right)\right) = 0,
\tag{4.24}
$$

again for $j = 1, 2$. Let us now analyze the different situations regarding power allocation.

Situation 1: one transmit antenna is switched off.

Assume that $\lambda^{\star}_{\mathbf{Q},1} = P_T$ and $\lambda^{\star}_{\mathbf{Q},2} = 0$ so that the second antenna is switched off. Our objective here is to find necessary and sufficient conditions for this situation to be optimal. Let us first concentrate on the necessary conditions. Obviously, from (4.24) one must have,

$$
\nu = \frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},1}}\left(S + \sqrt{S^2 - T^2}\right)
\tag{4.25}
$$

and consequently

$$\frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},1}}\left(S + \sqrt{S^2 - T^2}\right) \geq \frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},2}}\left(S + \sqrt{S^2 - T^2}\right). \tag{4.26}$$

This is clearly our candidate condition for sufficiency, and can alternatively be written as

$$\frac{P_T}{\sigma^2} \leq \frac{\left(m_{11}^2 + m_{21}^2\right) - \left(m_{12}^2 + m_{22}^2\right)}{m_{11}^2 m_{22}^2 + m_{21}^2 m_{12}^2} \triangleq \xi. \tag{4.27}$$

From this point, we see that the necessary and sufficient conditions for transmitting with the first antenna are (4.27) together with $\left(m_{11}^2 + m_{21}^2\right) > \left(m_{12}^2 + m_{22}^2\right)$. By the symmetry of the problem, one can also see that, whenever $\left(m_{11}^2 + m_{21}^2\right) < \left(m_{12}^2 + m_{22}^2\right)$ and $P_T\sigma^{-2} \leq -\xi = |\xi|$, the first antenna will be switched off and the whole power will be allocated to the second one, i.e., $\lambda_{\mathbf{Q},1}^\star = 0$, $\lambda_{\mathbf{Q},2}^\star = P_T$.

Situation 2: both antennas are active.

Without loss of generality, we assume here that $\xi > 0$ (the case $\xi < 0$ can readily be studied exploiting the inherent symmetry of the problem, and the case $\xi = 0$ leads to the trivial solution $\lambda_{\mathbf{Q},1}^\star = \lambda_{\mathbf{Q},2}^\star = P_T/2$).

We define $\lambda_{\mathbf{Q},1} = \lambda$ and then $\lambda_{\mathbf{Q},2} = P_T - \lambda$. The last KKT conditions in (4.18)–(4.22) tell us that

$$\nu = \frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},1}}\left[S + \sqrt{S^2 - T^2}\right] = \frac{\mathrm{d}}{\mathrm{d}\lambda_{\mathbf{Q},2}}\left[S + \sqrt{S^2 - T^2}\right]. \tag{4.28}$$

This condition can alternatively be written as

$$\frac{S}{T} + \sqrt{\frac{S^2}{T^2} - 1} = \frac{2m_{11}m_{12}m_{21}m_{22}}{m_{11}^2 m_{22}^2 + m_{21}^2 m_{12}^2}\frac{P_T - 2\lambda}{\xi\sigma^2 + P_T - 2\lambda}. \tag{4.29}$$

From this equation, the expression of the ergodic mutual information in (4.15), and the fact that $T > 0$ by its definition in (4.10), one readily sees that $(P_T - 2\lambda)$ and $\left(\xi\sigma^2 + P_T - 2\lambda\right)$ must always have the same sign (otherwise, the logarithm could not be properly defined). In addition, it can be stated that $P_T - 2\lambda < 0$, because if the contrary was assumed it would lead to $\sqrt{S^2 - T^2} < 0$. Thus, since $P_T - 2\lambda$ and $\xi\sigma^2 + P_T - 2\lambda$ have the same sign, it implies $\lambda > (P_T + \xi\sigma^2)/2$, i.e., the optimum power allocation algorithm gives more power to the best antenna (namely the first one whenever $\xi > 0$).

Note that, up to now, we have only proven that the optimum λ is such that

$$\frac{P_T + \xi\sigma^2}{2} < \lambda < P_T. \tag{4.30}$$

Next, we give a more specific characterization of the optimum transmit power allocation.
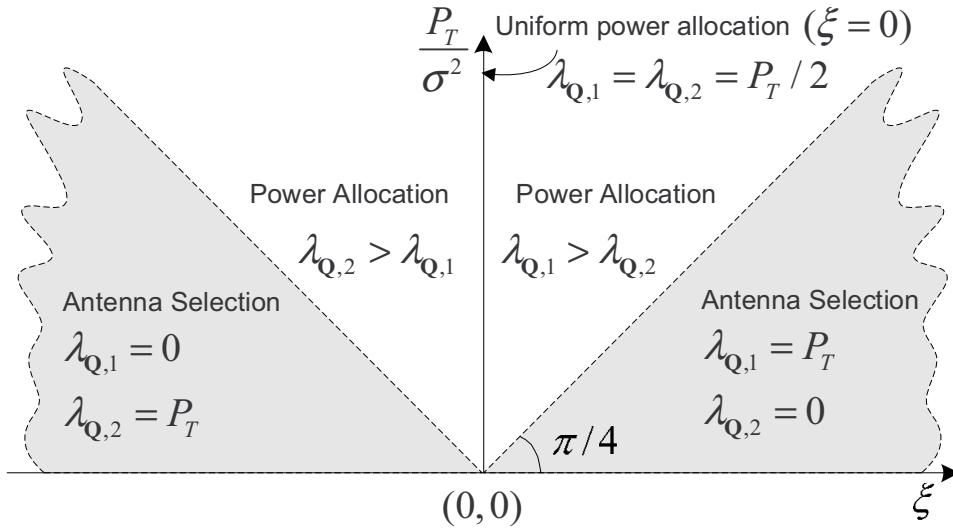
Figure 4.2: Graphical scheme of the optimal power allocation as a function of γ and ξ.

Our objective now is to obtain a closed form expression for the optimum power allocation strategy in the case where the two antennas are active. So, we define

$$x = \frac{2\lambda - P_T}{\sigma^2}, \tag{4.31}$$

$$\Delta = \frac{m_{11}^2 m_{22}^2 - m_{21}^2 m_{12}^2}{m_{11}^2 m_{22}^2 + m_{21}^2 m_{12}^2}, \tag{4.32}$$

$$\Upsilon = \frac{2 + P_T \sigma^{-2} \left(m_{11}^2 + m_{21}^2\right) + P_T \sigma^{-2} \left(m_{12}^2 + m_{22}^2\right)}{\left(m_{11}^2 + m_{21}^2\right) - \left(m_{12}^2 + m_{22}^2\right)}. \tag{4.33}$$

Using basic algebraic manipulations, one can show that (4.29) can be reformulated (adding, of course, new solutions) as:

$$\Delta^2 x^4 - 4\xi x^3 + \left[(3\xi - 4\Upsilon)\xi - \Delta^2 P_T^2 \sigma^{-4}\right] x^2 + 4\Upsilon\xi^2 x + P_T^2 \sigma^{-4}\xi^2 = 0. \tag{4.34}$$

The power assignment can be determined finding the root of the fourth order polynomial in (4.34) in the range $x \in (\xi, P_T\sigma^{-2})$. Note that there exist closed form expressions for the roots of a fourth order polynomial, and therefore the optimum power allocation strategy can be derived in exact form (no root-searching procedures are needed). In Appendix 4.D we prove that there is a single root of the polynomial (4.34) located on $x \in (\xi, P_T\sigma^{-2})$. Thereby, using (4.31), the optimal λ can be found.

Summary: optimum power allocation scheme.

Let us now briefly summarize the optimum power allocation policy in a 2×2 MIMO system under phase uncertainty at the transmitter.

- If $P_T/\sigma^2 < |\xi|$, then the best transmit antenna should be selected (namely $\lambda^\star_{\mathbf{Q},1} = P_T$, $\lambda^\star_{\mathbf{Q},2} = 0$ if $\xi > 0$ and vice versa if $\xi < 0$).

- If $P_T/\sigma^2 \geq |\xi|$, and $\xi > 0$, then $\lambda^\star_{\mathbf{Q},1} = (P_T + \sigma^2 x)/2$, where x is the unique root of the fourth order polynomial (4.34) in the region $\xi < x < P_T/\sigma^2$. If $\xi < 0$, an equivalent characterization can be given simply by swapping the role of antenna 1 and antenna 2.

In Figure 4.2, the optimum power allocation scheme above described is depicted as a function of ξ and P_T/σ^2.

4.4.2 Approximation by finite sample size

In the previous section we presented a closed-form expression for the optimal transmit covariance matrix for the particular case of a 2×2 MIMO system. In this section we characterize the diagonal elements of the transmit covariance matrix utilizing numerical methods for the general case.

From its expression in (4.7), it becomes clear that the mutual information maximization problem is a convex optimization problem because it is defined as the supremum of a concave function and the constraints are linear in the design variables, [Boy04]. However, the stochastic nature of the objective function (defined as the expectation of a random quantity) complicates the optimization by making difficult to compute the exact value of the objective function, its gradient, and its Hessian, which are needed at each optimization step.

To overcome this problem, in a practical set-up the expectation operator is approximated by taking the mean of a sufficiently large sample of realizations of the channel phases matrix $\mathbf{P}$ for a fixed magnitude matrix $\mathbf{M}$.

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) \simeq I_E^{\mathrm{num}}(\mathbf{\Lambda_Q}, \mathbf{M}) = \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda_Q}, \mathbf{M}, \mathbf{P}), \qquad (4.35)$$

where $\mathcal{M}$ is the set of realizations of the phases matrix $\mathbf{P}$. With the utilization of this approximation, we obtain the following problem

$$
\begin{aligned}
C_E \simeq \sup_{\mathbf{\Lambda_Q}} \quad & \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda_Q}, \mathbf{M}, \mathbf{P}) \\
\text{s.t.} \quad & \mathrm{Tr}\, \mathbf{\Lambda_Q} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \; [\mathbf{\Lambda_Q}]_{jj'} = 0, \quad 1 \leq j, j' \leq n_T,
\end{aligned}
\qquad (4.36)
$$

where the concavity of the objective function is preserved and, consequently, it is a convex optimization problem. The speed of convergence of these kind of problems can be greatly increased by providing to the algorithm the analytical expressions of the gradient vector and the Hessian matrix of the cost function, which are given next.

Since the dependence of $I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})$ on $\boldsymbol{\Lambda}_{\mathbf{Q}}$ is only through its diagonal elements, we can calculate the gradient of $I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})$ with respect to the vector $\boldsymbol{\lambda}_{\mathbf{Q}}$ which is given by $[\boldsymbol{\lambda}_{\mathbf{Q}}]_j = [\boldsymbol{\Lambda}_{\mathbf{Q}}]_{jj}$. Defining the vectors $\mathbf{m}_j$ and $\mathbf{p}_j$, with $1 \leq j \leq n_T$, as the vectors such that $[\mathbf{m}_j]_i = [\mathbf{M}]_{ij}$ and that $[\mathbf{p}_j]_i = [\mathbf{P}]_{ij}$, the gradient can be expressed as [Mag99]

$$[\nabla_{\boldsymbol{\lambda}_{\mathbf{Q}}} I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})]_j = \frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} (\mathbf{m}_j \odot \mathbf{p}_j)^H (\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\boldsymbol{\Lambda}_{\mathbf{Q}}(\mathbf{M} \odot \mathbf{P})^H)^{-1}(\mathbf{m}_j \odot \mathbf{p}_j).$$
$$(4.37)$$

Similarly, the Hessian matrix $\nabla^2_{\boldsymbol{\lambda}_{\mathbf{Q}}} I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$[\nabla^2_{\boldsymbol{\lambda}_{\mathbf{Q}}} I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})]_{jj'} = \frac{\partial^2 I_E^{\text{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})}{\partial \lambda_{\mathbf{Q},j} \partial \lambda_{\mathbf{Q},j'}} =$$
$$= -\frac{1}{|\mathcal{M}|} \sum_{\mathbf{P} \in \mathcal{M}} |(\mathbf{m}_j \odot \mathbf{p}_j)^H (\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\boldsymbol{\Lambda}_{\mathbf{Q}}(\mathbf{M} \odot \mathbf{P})^H)^{-1}(\mathbf{m}_{j'} \odot \mathbf{p}_{j'})|^2. \quad (4.38)$$

In the simulations section we show some ergodic capacity results, which have been obtained by solving the convex optimization problem in (4.36). We highlight that, although being a convex optimization problem, for large values of n_T and n_R the solution is computationally hard to obtain. In the following section, we present an alternative method to overcome this problem.

4.4.3   Optimization utilizing results from the theory of random matrices

Although the method presented in the last section to compute the ergodic mutual information is generic, in the sense that the values of n_T and n_R are arbitrary, and it can achieve any desired accuracy, its main drawback is that it is computationally hard to obtain a solution, which makes difficult its implementation in real-time systems.

Recall that, in the previous section, the expectation with respect to the channel phases of the instantaneous mutual information was approximated by a mean over a sample of phases realizations, as in (4.35). An alternative approximation for this expectation can be obtained within the framework of RMT. See further [Tul04] for an excellent and exhaustive introduction to RMT with applications to the analysis of wireless communications systems.

One of the objects of study of the theory of random matrices is the eigenvalue probability density of random matrices such as $\mathbf{X}\mathbf{X}^H$, with $\mathbf{X} \in \mathbb{C}^{n_R \times n_T}$ when the number of columns, n_T, and rows, n_R, of the matrix grows without bound but keeping the ratio n_T/n_R held constant, as

$$\lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \frac{n_T}{n_R} = \beta. \quad (4.39)$$

An interesting result in RMT is that the asymptotic eigenvalue density of $\mathbf{X}\mathbf{X}^H$ is independent of the particular distribution of the entries of $\mathbf{X}$ as long as they are independently distributed.

In particular, the entries of $\mathbf{X}$ have to fulfill

$$[\mathbf{X}]_{ij} = \sqrt{\frac{d_{ij}}{n_T}}\, z_{ij}, \tag{4.40}$$

with z_{ij} being i.i.d. random variables with $\mathbb{E}z_{11} = 0$ and $\mathbb{E}|z_{11}|^2 = 1$, and some other technical requirements on higher order moments which are out of the scope of the present dissertation, but which are fulfilled by the presented model. The parameter d_{ij} controls the variance of the random variable $[\mathbf{X}]_{ij}$ as it is described below.

The matrix definition in (4.40) fits excellently in the uncertainty model considered in this chapter as we explain in the following. Firstly, recall that the interest is in the evaluation of the quantity

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \log \det \left(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H \right). \tag{4.41}$$

Utilizing the spectral decomposition of the matrix $(\mathbf{M}\odot\mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M}\odot\mathbf{P})^H$ we obtain an equivalent expression for the ergodic mutual information.

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \sum_{i=1}^{n_R} \log \left(1 + \sigma^{-2}\zeta_i \right), \tag{4.42}$$

where ζ_i, with $1 \leq i \leq n_R$, represents the eigenvalues of the matrix $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H$. Note that the eigenvalues ζ_i are random quantities that depend on the current realization of the channel phases $\mathbf{P}$. By introducing the spectral measure of $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H$ as the following random probability measure

$$\mu(\zeta) = \frac{1}{n_R} \sum_{i=1}^{n_R} \delta(\zeta - \zeta_i), \tag{4.43}$$

the mutual information expression in (4.42) can be reformulated as

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \int_0^\infty \log \left(1 + \sigma^{-2}\zeta \right)\mu(\zeta)\mathrm{d}\zeta, \tag{4.44}$$

where the expectation with respect to the channel phases now acts upon the probability measure $\mu(\zeta)$. As commented above, the probability measure $\mu(\zeta)$ is the eigenvalue density of the matrix $(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H$. Defining $\mathbf{X} = (\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}^{1/2}$ it can be readily seen that

$$\mathbf{X}\mathbf{X}^H = (\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H. \tag{4.45}$$

Now, the entries of $\mathbf{X}$ fulfill that

$$[\mathbf{X}]_{ij} = [\mathbf{M}]_{ij}\lambda_{\mathbf{Q},j}^{1/2}\, \mathrm{e}^{\mathrm{i}\,\theta_{ij}}, \tag{4.46}$$

which is the same model as in (4.40) by identifying

$$\sqrt{\frac{d_{ij}}{n_T}} = [\mathbf{M}]_{ij} \lambda_{\mathbf{Q},j}^{1/2}, \tag{4.47}$$

$$z_{ij} = \mathrm{e}^{\mathrm{i}\,\theta_{ij}}, \tag{4.48}$$

and noting that $\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\theta_{ij}} = 0$ and $\mathbb{E}|\,\mathrm{e}^{\mathrm{i}\,\theta_{ij}}\,|^2 = 1$. Consequently the asymptotical distribution $\mu(\zeta)$ can be computed with the results of RMT. However, the interest here is not in the asymptotical distribution of $\mu(\zeta)$ but rather, on a deterministic approximation of it, for finite values of n_T and n_R. Fortunately, such approximation of $\mu(\zeta)$ has been developed in, *e.g.*, [Tul04] and [Hac05], and it is reproduced in the following, utilizing the guidelines described in [Hac06] or [Pay06a].

First of all, by differentiating and integrating the expression (4.41) with respect to σ^2, the following expression is obtained

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \int_{\sigma^2}^\infty \left(\frac{1}{\xi} - \mathbb{E}_{\mathbf{P}} \frac{1}{n_R} \operatorname{Tr}\left((\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H + \xi\mathbf{I}\right)^{-1}\right) \mathrm{d}\xi \tag{4.49}$$

It can be readily seen that the second term inside the integral can be equivalently expressed as

$$\frac{1}{n_R} \operatorname{Tr}\left((\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H + \xi\mathbf{I}\right)^{-1} = \frac{1}{n_R}\sum_{i=1}^{n_R}\frac{1}{\zeta_i + \xi} = \int \frac{1}{\zeta + \xi}\mu(\zeta)\mathrm{d}\zeta = m_\mu(-\xi), \tag{4.50}$$

where $m_\mu(z)$ is the Stieltjes transform of the random measure $\mu(\zeta)$. Having introduced this transformation, the mutual information expression in (4.49) becomes

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \int_{\sigma^2}^\infty \left(\frac{1}{\xi} - \mathbb{E}_{\mathbf{P}} m_\mu(-\xi)\right) \mathrm{d}\xi. \tag{4.51}$$

As commented above, the key idea is to replace the expectation of the Stieltjes transform of the random measure $\mu(\zeta)$ by the Stieltjes transform of a deterministic measure $\nu(\zeta)$ in such a way that

$$\lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \mu(\zeta) - \nu(\zeta) = 0, \text{ a.s.,} \quad \text{with } \lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \frac{n_T}{n_R} = \beta. \tag{4.52}$$

Consequently, the approximation

$$I_E(\mathbf{\Lambda_Q}, \mathbf{M}) \simeq \bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M}) = \int_{\sigma^2}^\infty \left(\frac{1}{\xi} - m_\nu(-\xi)\right) \mathrm{d}\xi, \tag{4.53}$$

is obtained, which fulfills that

$$\lim_{\substack{n_T \to \infty \\ n_R \to \infty}} I_E(\mathbf{\Lambda_Q}, \mathbf{M}) - \bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M}) = 0, \text{ a.s.,} \quad \text{with } \lim_{\substack{n_T \to \infty \\ n_R \to \infty}} \frac{n_T}{n_R} = \beta. \tag{4.54}$$

and for any finite values of n_T and n_R, $\bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M})$ approximates $I_E(\mathbf{\Lambda_Q}, \mathbf{M})$. Unfortunately, the mean and variance of this estimator are not fully characterized yet and its study is a subject of ongoing research.

Now it is only needed to obtain a procedure to calculate $m_\nu(z)$ as a function of $\mathbf{\Lambda_Q}$ and $\mathbf{M}$. In [Tul04] and [Hac05], the authors found that the Stieltjes transform of the deterministic probability measure $\nu(\zeta)$, fulfilling the property in (4.52)) is given by

$$m_\nu(z) = \frac{1}{n_R} \mathbf{1}^H \mathbf{t}(z), \tag{4.55}$$

where $\mathbf{1}^H$ is the all-one row vector of the appropriate dimension, and $\mathbf{t}(z)$ is the solution to the following system of equations

$$[\mathbf{t}(z)]_i = \frac{1}{-z\left(1 + n_T^{-1}\tilde{\mathbf{d}}_i^H \tilde{\mathbf{t}}(z)\right)}, \quad \text{for } 1 \le i \le n_R, \tag{4.56}$$

$$[\tilde{\mathbf{t}}(z)]_j = \frac{1}{-z\left(1 + n_T^{-1}\mathbf{d}_j^H \mathbf{t}(z)\right)}, \quad \text{for } 1 \le j \le n_T, \tag{4.57}$$

where $\mathbf{d}_j \in \mathbb{R}^{n_R}$, with $1 \le j \le n_T$, and $\tilde{\mathbf{d}}_i \in \mathbb{R}^{n_T}$, with $1 \le i \le n_R$ are column vectors with real non-negative entries such that

$$[\mathbf{d}_j]_i = d_{ij} = n_T [\mathbf{M}]_{ij}^2 \lambda_{\mathbf{Q},j} \tag{4.58}$$

$$[\tilde{\mathbf{d}}_i]_j = d_{ij} = n_T [\mathbf{M}]_{ij}^2 \lambda_{\mathbf{Q},j}. \tag{4.59}$$

With these definitions and integrating (4.53) it is now possible to obtain an expression for the deterministic approximation of the mutual information as

$$\bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M}) = -\sum_{i=1}^{n_R} \log(\sigma^2 [\mathbf{t}(-\sigma^2)]_i) - \sum_{j=1}^{n_T} \log(\sigma^2 [\tilde{\mathbf{t}}(-\sigma^2)]_j) -$$
$$-\sigma^2 \sum_{i=1}^{n_R} \sum_{j=1}^{n_T} [\mathbf{t}(-\sigma^2)]_i [\tilde{\mathbf{t}}(-\sigma^2)]_j [\mathbf{M}]_{ij}^2 \lambda_{\mathbf{Q},j}. \tag{4.60}$$

Although this expression for $\bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M})$ is rather complicated, it can be numerically evaluated in a very efficient manner. The computationally heavier part is the calculation of $\mathbf{t}$ and $\tilde{\mathbf{t}}$ which has to be done with numerical methods (for example, utilizing the fixed-point technique, which, for this particular case, is a very efficient and reliable method).

Once a method to compute $\bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M})$ has been obtained, the ergodic capacity can now be approximated as $C_E \simeq \bar{C}_E$, where $\bar{C}_E$ is the solution to the following optimization problem

$$\begin{aligned} \bar{C}_E = \sup_{\mathbf{\Lambda_Q}} \quad & \bar{I}_E(\mathbf{\Lambda_Q}, \mathbf{M}) \\ \text{s.t.} \quad & \text{Tr}\,\mathbf{\Lambda_Q} \le P_T, \\ & \lambda_{\mathbf{Q},j} \ge 0, \ [\mathbf{\Lambda_Q}]_{jj'} = 0, \quad 1 \le j, j' \le n_T, \end{aligned} \tag{4.61}$$

which is solved numerically utilizing standard methods. Recall again that the proposed method to approximate C_E is the result from the application of RMT.

In the simulations section, the goodness of this approximation is validated experimentally, and some insight into the implications of RMT in this problem is given.

4.5 Compound mutual information and capacity

The expression for the compound mutual information can be obtained from the particularization of the expression given in (2.19), which recall that is defined in [Csi81] as the infimum, with respect to the channel state uncertainty, of the mutual information expression in (4.4). For the considered uncertainty model, it particularizes to:

$$I_C(\mathbf{Q}, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{P}} \Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}), \tag{4.62}$$

where $\mathcal{P} \triangleq \{\mathbf{X} \in \mathbb{C}^{n_R \times n_T} \mid |[\mathbf{X}]_{ij}| = 1\}$ defines the set of all possible channel phases compatible with the incomplete knowledge about $\mathbf{H}$. It must be emphasized that the compound mutual information does not depend on the statistical properties of the uncertainty in the channel, because it only considers the worst-case scenario.

The compound capacity is then naturally defined as the maximum with respect to the transmit covariance matrix of the compound mutual information as in (2.20), subject to a total mean transmitted power constraint.

$$
\begin{aligned}
C_C = \sup_{\mathbf{Q}} \quad & I_C(\mathbf{Q}, \mathbf{M}) \\
\text{s.t.} \quad & \operatorname{Tr} \mathbf{Q} \leq P_T, \\
& \mathbf{Q} \succcurlyeq \mathbf{0}.
\end{aligned}
\tag{4.63}
$$

The compound mutual information is a measure of the worst-case achievable rates in situations where no significant channel variability may occur during the transmission of the message and the transmitter is only informed of (or is only able to estimate accurately) the magnitude matrix $\mathbf{M}$. This may be the case of a static communication between the transmitter and the receiver, or when communicating in a slow fading environment.

The optimal structure of the covariance matrix that achieves the compound capacity is given next and it is depicted in Figure 4.1.

Proposition 4.5.1 *Assume instantaneous perfect CSI at the receiver and magnitude CSI at the transmitter, $\mathbf{M}$. If the channel state, which is an unknown matrix, is defined as $\mathbf{H} = \mathbf{M} \odot \mathbf{P}$, with $[\mathbf{P}]_{ij} = \mathrm{e}^{\mathrm{i}\theta_{ij}}$ then the compound capacity of the Gaussian memoryless channel (4.1), subject to a transmit power constraint P_T, can only be achieved by a diagonal covariance matrix $\mathbf{Q}^\star = \mathbf{\Lambda_Q}$, i.e., $\mathbf{U_Q^\star} = \mathbf{I}$. The compound capacity is then given by the solution to*

$$
\begin{aligned}
C_C = \sup_{\mathbf{\Lambda_Q}} \quad & I_C(\mathbf{\Lambda_Q}, \mathbf{M}) \\
\text{s.t.} \quad & \operatorname{Tr} \mathbf{\Lambda_Q} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \ [\mathbf{\Lambda_Q}]_{jj'} = 0, \quad 1 \leq j, j' \leq n_T.
\end{aligned}
\tag{4.64}
$$

Proof See Appendix 4.A.   ∎

As stated in Appendix 4.A, similarly as in the ergodic case, the compound capacity can also be achieved by a diagonal transmit covariance matrix, *i.e.*, transmitting independent data streams through different antennas with an appropriate power allocation (see Figure 4.1).

In the following section, a closed-form solution for the above optimization problem is given for the particular case with $n_T = 2$ [Pay05b, Pay05a]. For a generic value of n_T numerical methods need to be utilized.

4.5.1 Solution for the particular case with $n_T = 2$

Before proceeding to present the solution to the problem of finding the power allocation that maximizes the compound mutual information, some preliminary results have to be described.

It can be proved that the second order Taylor expansion of the determinant $\det(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H)$, formulated as $\det\left(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H\right) = \mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) + \mathsf{o}\left(\sum_{j,j'} \lambda_{\mathbf{Q},j}\lambda_{\mathbf{Q},j'}\right)$, is given by the following function:

$$\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) = 1 + \sigma^{-2}\sum_j \lambda_{\mathbf{Q},j}\alpha_{jj} + \sigma^{-4}\sum_{j,j'>1}\lambda_{\mathbf{Q},j}\lambda_{\mathbf{Q},j'}(\alpha_{jj}\alpha_{j'j'} - |\alpha_{jj'}|^2), \tag{4.65}$$

where $\alpha_{jj'} = \mathbf{h}_j^H\mathbf{h}_{j'}$ and $\mathbf{h}_j$ represents the j-th column of the matrix $\mathbf{H}$. Note that for $n_T = 2$ the second order approximation becomes exact,

$$n_T = 2 \Rightarrow \det\left(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H\right) = \mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}). \tag{4.66}$$

The objective now is to find a lower bound of the function $\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H})$ over all the possible phases realizations which model the uncertainty in the channel knowledge. In order to do it, note that $|\alpha_{jj'}|^2$ can be upper bounded as

$$|\alpha_{jj'}|^2 = |\mathbf{h}_j^H\mathbf{h}_{j'}|^2 = \left(\sum_k h_{kj}^*h_{kj'}\right)\left(\sum_l h_{lj}h_{lj'}^*\right) \tag{4.67}$$

$$= \sum_k |h_{kj}|^2|h_{kj'}|^2 + \sum_{k,l\neq k} h_{kj}^*h_{kj'}h_{lj}h_{lj'}^* \tag{4.68}$$

$$= \sum_k m_{kj}^2 m_{kj'}^2 + \sum_{k,l>k} 2m_{kj}m_{kj'}m_{lj}m_{lj'}\cos(\theta_{kj} - \theta_{kj'} - \theta_{lj} + \theta_{lj'}) \tag{4.69}$$

$$\leq \sum_k m_{kj}^2 m_{kj'}^2 + \sum_{k,l>k} 2m_{kj}m_{kj'}m_{lj}m_{lj'} \tag{4.70}$$

$$= |\mathbf{m}_j^H\mathbf{m}_{j'}|^2, \tag{4.71}$$

where $\mathbf{m}_j$ is the j-th column of the matrix $\mathbf{M}$. Noticing that $\alpha_{jj} = \|\mathbf{h}_j\|^2 = \|\mathbf{m}_j\|^2$, and using the upper bound above, it can be easily seen that

$$\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) \geq \mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M}), \tag{4.72}$$

i.e., the minimum of the second order Taylor expansion of the determinant over all the possible phases corresponds to the evaluation of this Taylor expansion for the concrete case of having zero-phases.

From what has been said above, for the particular case where $n_T = 2$, the expression of the compound mutual information in (4.62) is simplified to

$$I_C(\mathbf{Q}, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{P}} \Psi(\mathbf{Q}, \mathbf{M}, \mathbf{P}) = \inf_{\mathbf{P} \in \mathcal{P}} \log \mathcal{D}(\mathbf{\Lambda_Q}, \mathbf{H}) = \log \mathcal{D}(\mathbf{\Lambda_Q}, \mathbf{M}). \tag{4.73}$$

Thus, the problem in (4.63) can be reformulated as

$$\begin{aligned} C_C = \sup_{\{\lambda_{\mathbf{Q},1}, \lambda_{\mathbf{Q},2}\}} \quad & \log \mathcal{D}(\mathbf{\Lambda_Q}, \mathbf{M}) \\ \text{s.t.} \quad & \lambda_{\mathbf{Q},1} + \lambda_{\mathbf{Q},2} = P_T, \\ & \lambda_{\mathbf{Q},j} \geq 0, \quad 1 \leq j \leq 2, \end{aligned} \tag{4.74}$$

and, using the KKT conditions and supposing $\|\mathbf{m}_1\|^2 < \|\mathbf{m}_2\|^2$ and $\mathbf{m}_1 \neq k\mathbf{m}_2$, the optimal solution can be readily obtained as

$$\lambda_{\mathbf{Q},1}^\star = \left(\frac{P_T}{2} - \frac{(\|\mathbf{m}_2\|^2 - \|\mathbf{m}_1\|^2)\sigma^2}{2\left(\|\mathbf{m}_1\|^2\|\mathbf{m}_2\|^2 - |\mathbf{m}_2^T\mathbf{m}_1|^2\right)} \right)^+, \tag{4.75}$$

$$\lambda_{\mathbf{Q},2}^\star = P_T - \lambda_{\mathbf{Q},1}^\star. \tag{4.76}$$

The solution to the case where $\|\mathbf{m}_1\|^2 > \|\mathbf{m}_2\|^2$ is symmetric to the solution presented above, and if $\mathbf{m}_1 = k\mathbf{m}_2$ all the power must be given to the best channel. Notice that, although the problem set-up is different, a similar result was obtained in [Chu01].

In the following subsection approximate solutions to compute the optimal power allocation without imposing the restriction that $n_T = 2$ are presented.

4.5.2   Approximation by finite sample size

Similarly as in the ergodic channel capacity presented in the previous section, in this case, the compound mutual information is proposed to be calculated approximately by using a large sample of realizations of the channel phases matrix $\mathbf{P}$ for a fixed magnitude matrix $\mathbf{M}$. Using the same notation as before, let $\mathcal{M}$ be the set of samples of random phases realizations, *i.e.*, $\mathcal{M} = \{\mathbf{P}_1, \ldots, \mathbf{P}_{|\mathcal{M}|}\}$, where k will be used as the index corresponding to the k-th element, *i.e.*, $\mathbf{P}_k$. Using this, the compound mutual information can be approximated as

$$I_C(\mathbf{\Lambda_Q}, \mathbf{M}) \simeq I_C^{\text{num}}(\mathbf{\Lambda_Q}, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{M}} \Psi(\mathbf{\Lambda_Q}, \mathbf{M}, \mathbf{P}) = \Psi(\mathbf{\Lambda_Q}, \mathbf{M}, \mathbf{P}_{k_{\min}}), \tag{4.77}$$

where $k_{\min}$ is the index corresponding to the phase matrix at which the minimum in the expression above is achieved, which depends on $\mathbf{\Lambda_Q}$ for a fixed $\mathbf{M}$, *i.e.*, $k_{\min} = k_{\min}(\mathbf{\Lambda_Q})$. The

optimization problem, corresponding to the maximization of $I_C^{\text{num}}(\mathbf{\Lambda_Q}, \mathbf{M})$, can now be written as the following convex optimization problem:

$$
\begin{aligned}
\bar{C}_C \simeq \sup_{\mathbf{\Lambda_Q}} \quad & \Psi(\mathbf{\Lambda_Q}, \mathbf{M}, \mathbf{P}_{k_{\min}(\mathbf{\Lambda_Q})}) \\
\text{s.t.} \quad & \operatorname{Tr} \mathbf{\Lambda_Q} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \ [\mathbf{\Lambda_Q}]_{jj'} = 0, \quad 1 \leq j, j' \leq n_T.
\end{aligned} \tag{4.78}
$$

The numerical algorithms for convex problems based, for example, on interior point methods, can be applied to obtain the solution to the above problem. The speed of convergence of these kind of procedures can be greatly increased by providing to the algorithm the analytical expressions of the gradient vector and the Hessian matrix of the cost function. In this case, the gradient and Hessian can be directly found by adapting the expressions given in §4.4.2 to the compound case. This results in the following expressions:

$$
[\nabla_{\boldsymbol{\lambda_Q}} I_C^{\text{num}}(\mathbf{\Lambda_Q}, \mathbf{M})]_j = (\mathbf{m}_j \odot \mathbf{p}_j)^H \big(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H\big)^{-1}(\mathbf{m}_j \odot \mathbf{p}_j)\Big|_{\mathbf{P}=\mathbf{P}_{k_{\min}(\mathbf{\Lambda_Q})}}. \tag{4.79}
$$

Similarly, the Hessian matrix $\nabla^2_{\boldsymbol{\lambda_Q}} I_C^{\text{num}}(\mathbf{\Lambda_Q}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$
\begin{aligned}
[\nabla^2_{\boldsymbol{\lambda_Q}} I_C^{\text{num}}(\mathbf{\Lambda_Q}, \mathbf{M})]_{jj'} = \\
= - \,|(\mathbf{m}_j \odot \mathbf{p}_j)^H \big(\mathbf{I} + \sigma^{-2}(\mathbf{M} \odot \mathbf{P})\mathbf{\Lambda_Q}(\mathbf{M} \odot \mathbf{P})^H\big)^{-1}(\mathbf{m}_{j'} \odot \mathbf{p}_{j'})|^2\Big|_{\mathbf{P}=\mathbf{P}_{k_{\min}(\mathbf{\Lambda_Q})}}. \tag{4.80}
\end{aligned}
$$

4.5.3 Approximated approach inspired by Taylor expansion

The previous approach based on the approximation of the minimum of the mutual information over a finite set of realizations of the phase matrix $\mathbf{P}$ may have an unaffordable computational load in order to obtain accurate approximations since the sample size may be required to be too high, specially for moderate and high values of n_T and n_R.

In this subsection a different approach is taken. First, the second order Taylor expansion of the determinant in the expression of the mutual information is considered and, afterwards, the minimum value of this Taylor expansion over all the possible phases realizations is presented recalling the results from §4.5.1. Once this expression has been made explicit, a convex problem is formulated inspired by the previous result in order to calculate an approximation of the maximum value of the worst-case original mutual information, *i.e.*, of the compound capacity. In the cases where the Taylor expansion coincides with the original exact determinant, this convex optimization problem is able to find the compound capacity exactly. This happens when either the number of transmit or receive antennas is equal to 2 or less, *i.e.*, $\min\{n_T, n_R\} \leq 2$. In a general case, the approximation will improve as the terms of order equal to or greater than

3 in the Taylor expansion have a low contribution, which happens, for example, in low SNR conditions. The goodness of this approximation will be evaluated in the simulations section.

Recalling the results from §4.5.1, the second order Taylor expansion of the determinant $\det(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H)$, formulated as $\det\left(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H\right) = \mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) + \mathsf{o}\left(\sum_{j,j'} \lambda_{\mathbf{Q},j}\lambda_{\mathbf{Q},j'}\right)$, is given by the following function:

$$\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) = 1 + \sigma^{-2}\sum_j \lambda_{\mathbf{Q},j}\alpha_{jj} + \sigma^{-4}\sum_{j,j'>1} \lambda_{\mathbf{Q},j}\lambda_{\mathbf{Q},j'}(\alpha_{jj}\alpha_{j'j'} - |\alpha_{jj'}|^2), \qquad (4.81)$$

where $\alpha_{jj'} = \mathbf{h}_j^H\mathbf{h}_{j'}$ and $\mathbf{h}_j$ represents the j-th column of the matrix $\mathbf{H}$. As it was also seen in §4.5.1 a lower bound of the function $\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H})$ is given by

$$\mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{H}) \geq \mathcal{D}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M}), \qquad (4.82)$$

i.e., the minimum of the second order Taylor expansion of the determinant over all the possible phases corresponds to the evaluation of this Taylor expansion for the concrete case of having zero-phases.

Inspired by this result, in the following it will be assumed that

$$\det(\mathbf{I} + \sigma^{-2}\mathbf{H}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{H}^H) \gtrsim \det(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H), \qquad (4.83)$$

which will be more accurate as the terms of order greater than 2 in the Taylor expansion are more negligible. Using this approximation, the compound capacity will now be found numerically as the solution to the following convex optimization problem:

$$\begin{aligned}
\bar{C}_C \simeq \sup_{\boldsymbol{\Lambda}_{\mathbf{Q}}} \quad & \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H) \\
\text{s.t.} \quad & \operatorname{Tr}\boldsymbol{\Lambda}_{\mathbf{Q}} \leq P_T, \\
& \lambda_{\mathbf{Q},j} \geq 0, \ [\boldsymbol{\Lambda}_{\mathbf{Q}}]_{jj'} = 0, \quad 1 \leq j,j' \leq n_T.
\end{aligned} \qquad (4.84)$$

Similarly as presented above, this convex problem can be solved easily using efficient numerical algorithms based, for example, on interior point methods. Also as noted before, the speed of convergence of these kind of procedures can be greatly increased by providing to the algorithm the analytical expressions of the gradient vector and the Hessian matrix of the cost function. In this case, the gradient and Hessian can be directly found by adapting the expressions given in §4.5.2 to the objective function of the problem in (4.84). This results in the following expression:

$$[\nabla_{\boldsymbol{\lambda}_{\mathbf{Q}}} \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H)]_j = \mathbf{m}_j^H\left(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H\right)^{-1}\mathbf{m}_j, \quad 1 \leq j \leq n_T. \qquad (4.85)$$

Similarly, the Hessian matrix $\nabla_{\boldsymbol{\lambda}_{\mathbf{Q}}}^2 I_C^{\mathrm{num}}(\boldsymbol{\Lambda}_{\mathbf{Q}}, \mathbf{M})$ is defined as the matrix whose (j, j')-th element fulfills

$$[\nabla_{\boldsymbol{\lambda}_{\mathbf{Q}}}^2 \log\det(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H)]_{jj'} = -|\mathbf{m}_j^H\left(\mathbf{I} + \sigma^{-2}\mathbf{M}\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{M}^H\right)^{-1}\mathbf{m}_{j'}|^2, \quad 1 \leq j,j' \leq n_T. \qquad (4.86)$$
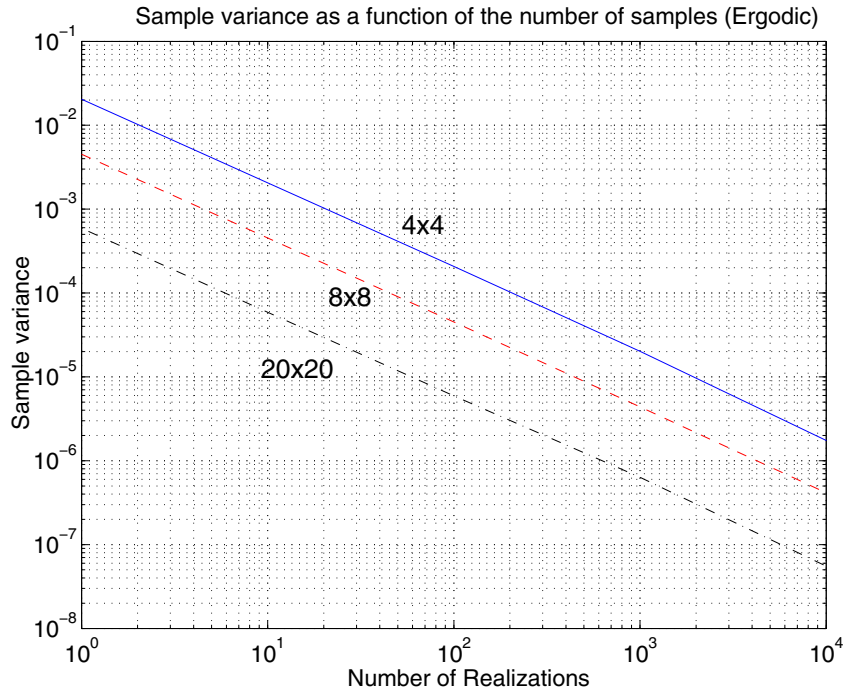
Figure 4.3: Sample variance evaluation of the ergodic mutual information when using a set of random realizations of the phases matrix. Different numbers of random realizations of the phases matrix and numbers of transmit and receive antennas have been evaluated.

4.6 Simulations

In this section, some simulation results are presented to illustrate the concepts that have been presented previously and evaluate the proposed solutions in terms of achieved performance.

First, in Figure 4.3, the accuracy of the approximation of the ergodic mutual information using a sample set of random realizations of the phases matrix is evaluated. This evaluation is performed in terms of the sample variance of the estimation vs. the size of the sample set and for different numbers of transmit and receive antennas. As a general conclusion, it is observed that when the number of random samples and antennas increases, the sample variance decreases, *i.e.*, the quality of the approximates improves, as expected. On the other hand, a qualitative evaluation of the accuracy of the technique based on random matrix theory is shown in Figure 4.4. There, the estimated histogram of the eigenvalues of a random MIMO channel (*i.e.*, the eigenvalues of the expression $\mathbf{HH}^H$) for different numbers of transmit and receive antennas is presented jointly with the asymptotic non-random probability density function $\nu(\zeta)$ derived from RMT. As expected, as the size of the MIMO channel increases, the accuracy gets better. Note, however, that even in the case of a low number of antennas, the asymptotic distribution is able to reproduce the mean behavior of the random distribution.
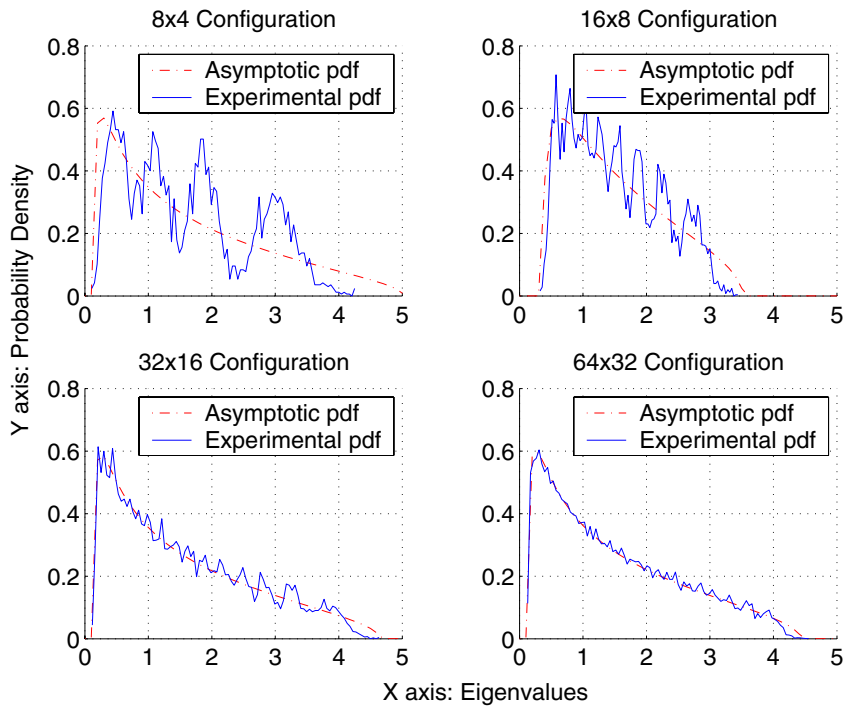
Figure 4.4: Comparison between the experimental and, therefore, random histogram of the eigenvalues of a random channel, *i.e.*, of the expression $\mathbf{H}\mathbf{H}^H$, and the asymptotic non-random probability density function $\nu(\zeta)$ given by RMT for different numbers of transmit and receive antennas.

Figure 4.5 shows again the random histogram of the eigenvalues and the asymptotic non-random probability density function $\nu(\zeta)$ derived from RMT for two different situations. First, the situation corresponding to a transmitter in which a uniform power allocation is used, *i.e.*, a situation in which the knowledge of the modulus of the channel gains is not exploited (eigenvalues of $\mathbf{H}\mathbf{I}\mathbf{H}^H$). Secondly, the case where the power allocation is derived so that the ergodic mutual information is maximized, *i.e.*, the ergodic capacity is achieved (eigenvalues of $\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^H$, with $\boldsymbol{\Lambda}$ containing the optimized power allocation). As a conclusion, it is shown that thanks to the optimization of the power allocation, the distribution of the eigenvalues is shifted to higher argument values and, therefore, the mutual information increases.

A more concrete numerical evaluation of the ergodic capacity is shown in Figures 4.6 and 4.7. From the first one it is concluded that both the numerical algorithm based on the sample set of random realizations of the phases matrix and the algorithm derived from the application of random matrix theory achieve almost the same ergodic capacity. This suggests to choose the second technique, since it has a much lower computational load than the first one. In the second figure, *i.e.*, Figure 4.7, the optimized power allocation is compared with the uniform power allocation, that is, that corresponding to the complete uncertainty of both the magnitude and the phases of the channel gains. As seen in the figure, the gain provided by the optimized power
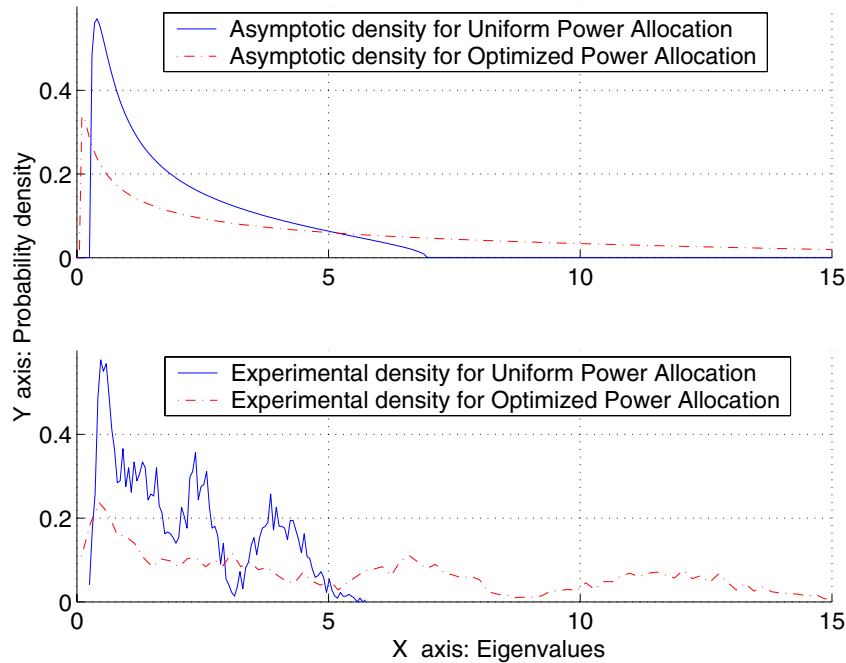
Figure 4.5: Comparison between the experimental and the asymptotic probability density functions of the eigenvalues resulting from uniform power allocation (*i.e.*, of the expression $\mathbf{H}\mathbf{I}\mathbf{H}^{H}$) and the optimized power allocation to achieve ergodic capacity (*i.e.*, of the expression $\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^{H}$).

allocation is specially important at low SNR, since at high SNR, the optimized power allocation tends to be the uniform power allocation, as expected.

Concerning the simulations corresponding to the compound capacity, some results are given in the remaining figures. Figure 4.8 also evaluates the sample variance of the compound capacity corresponding to the application of a sample set of random realizations of the phases matrix. The same conclusion as in Figure 4.3 can be obtained. Note, however, that the sample variances obtained in the case of the compound capacity are approximately an order of magnitude higher than in the case of the ergodic capacity. Finally, the same conclusions as before are derived from the observation of Figs. 4.9 and 4.10. Note that in Figure 4.9 the techniques that are compared are the one based on the application of the sample set of random realizations of the phases matrix and the algorithm inspired by the second order Taylor expansion of the determinant in the expression of the mutual information. As previously, both algorithms provide almost the same result and, therefore, it is suggested to use the second one, since it has a much lower computational complexity.
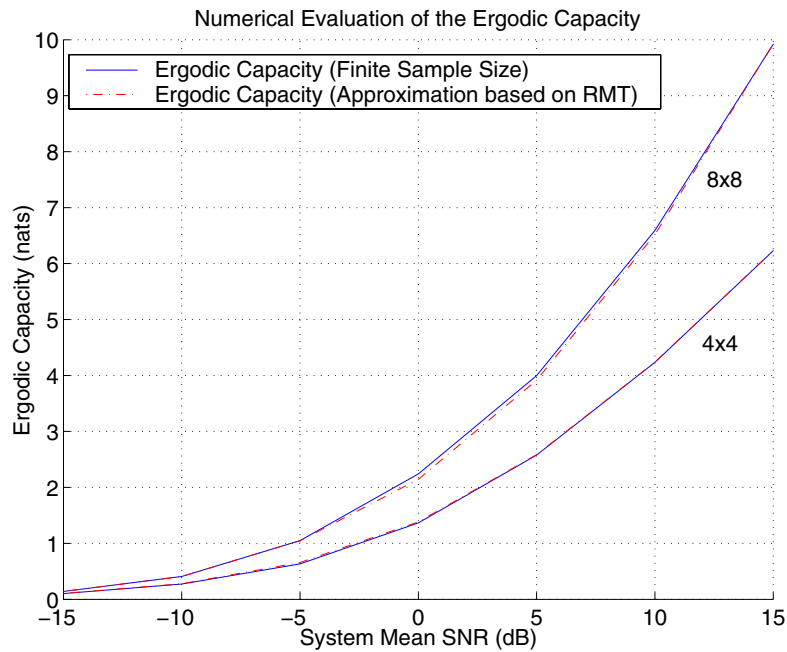
Figure 4.6: Evaluation of the optimized power allocation in terms of ergodic mutual information using the approximated algorithm with a finite sample set of random realization of the phases matrix and the algorithm based on RMT.
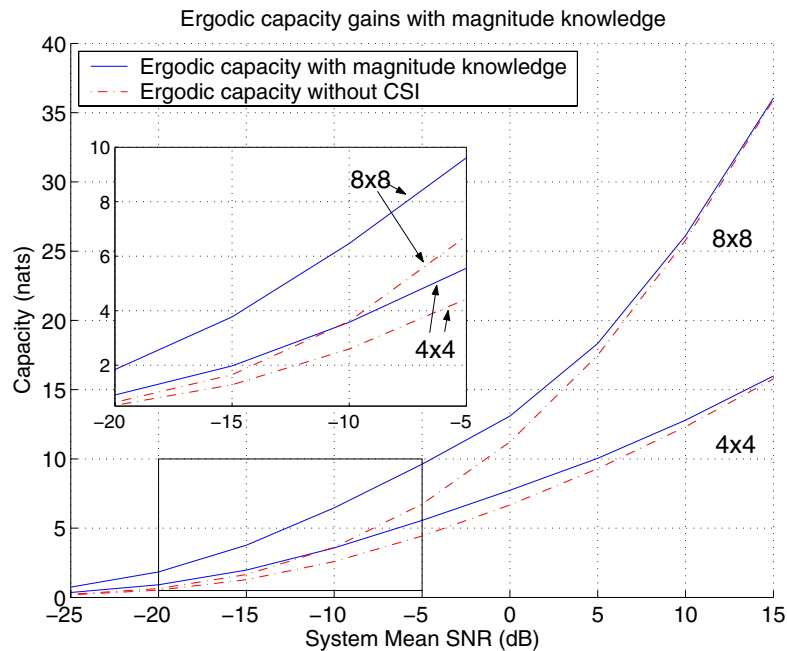
Figure 4.7: Comparison of the ergodic capacity that can be achieved when the transmitter has magnitude knowledge and phase uncertainty using the optimized power allocation with the capacity that can be achieved when the transmitter has no CSI, *i.e.*, with uniform power allocation.

Figure 4.8: Sample variance evaluation of the compound mutual information when using a set of random realizations of the phases matrix. Different numbers of random realizations of the phases matrix and numbers of transmit and receive antennas have been evaluated.
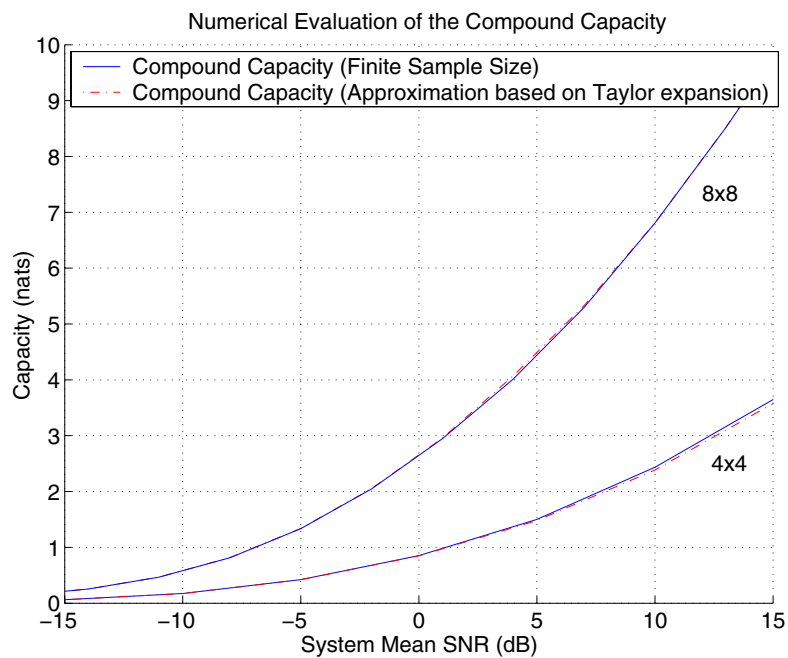
Figure 4.9: Evaluation of the optimized power allocation in terms of compound mutual information using the approximated algorithm with a finite sample set of random realization of the phases matrix and the algorithm inspired by the second order Taylor expansion of the determinant in the expression of the mutual information.
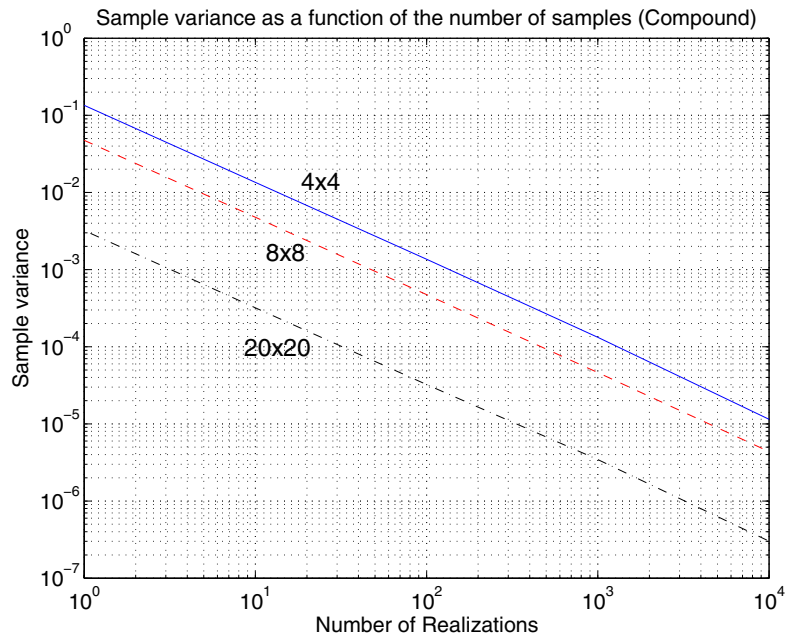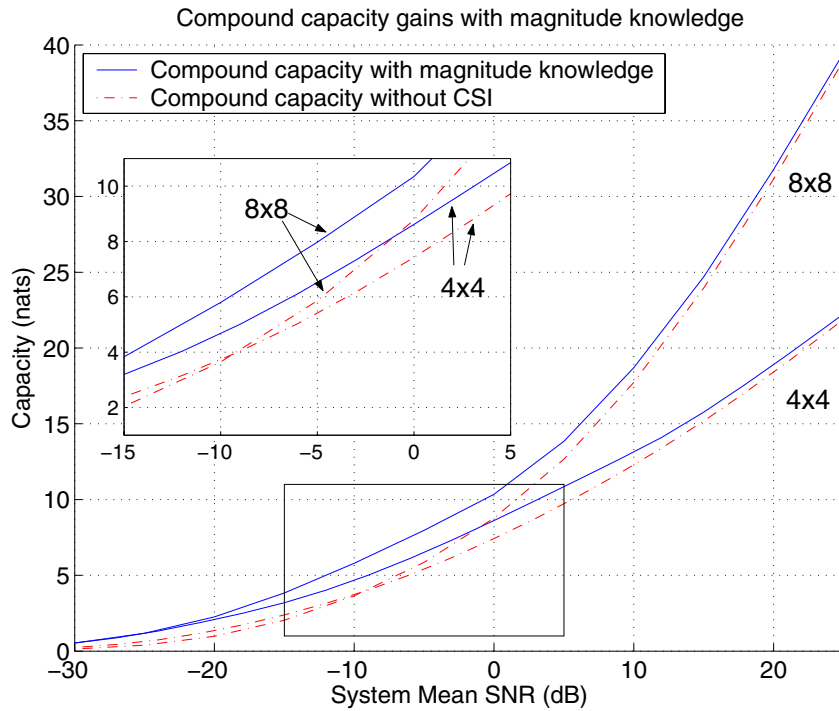
Figure 4.10: Comparison of the ergodic capacity that can be achieved when the transmitter has magnitude knowledge and phase uncertainty using the optimized power allocation with the capacity that can be achieved when the transmitter has no CSI, *i.e.*, with uniform power allocation.

4.7   Chapter summary and conclusions

In this chapter, the problem of the design of a transmitter in a MIMO channel to achieve capacity has been addressed. More concisely, the MIMO channel has been considered as a system with uncertainty in the channel knowledge in terms of known magnitude of the channel gains and completely unknown phases. The design problem has been derived from two different perspectives that deal with the uncertainty in the system through two different approaches. On the one hand, in the ergodic capacity case, the uncertainty has been addressed from a statistical point of view, *i.e.*, the capacity has been calculated as the average of the mutual information over the statistics of the unknown phases. On the other hand, in the compound capacity case, the capacity has been found as the infimum of the mutual information over all the possible phases realizations.

The main problem when dealing with the two designs of the transmit covariance matrix to achieve ergodic and compound capacities is that, except for particular cases of values of n_T and n_R, closed-form expressions do not exist. This chapter has then focused the attention on the application of numerical methods to find the optimal transmission signaling schemes that, in both cases, can be implemented by transmitting independent data streams through different

antennas with an appropriate power allocation. In both cases, two different numerical approaches have been proposed. One of them is based on the utilization of a set of random samples of the phases matrix, which are used to approximate the ergodic and compound mutual information. For the case of the ergodic capacity, an additional technique based on the tools provided by RMT is also presented, which reduces drastically the computational complexity. This reduction in the computational load can also be achieved in the compound capacity problem by using an approximation of the worst-case mutual information inspired by the minimization of a second order Taylor expansion of the determinant in the mutual information.

Finally, the simulations have been useful to show two different aspects. First, the accuracy of the proposed approximated solutions has been tested, showing that the low complexity techniques provide nearly the same results as the high load algorithms based on the set of random realizations of the phases matrix. Finally, the ergodic and compound capacities themselves have been evaluated for different scenarios, leading to the conclusion that the optimization of the power allocation with only the knowledge of the magnitude of the channel gains allows to improve the system performance very importantly.

4.A   Proof that a diagonal covariance matrix is ergodic and compound optimal

The proof for the ergodic case is first presented, and then the necessary steps to obtain the proof for the compound case are simply added.

4.A.1   Ergodic case

The following lemmas are first presented.

Lemma 1 *Let $\mathbf{J}_n \in \mathbb{Z}^{n_T \times n_T}$ be a diagonal matrix such that its non-zero entries satisfy $[\mathbf{J}_n]_{jj} \in \{1, -1\}$, $\forall j$. There are $L = 2^{n_T}$ different such matrices which are indexed from $n = 1$ to $n = L$. It then follows that*

$$\mathbf{\Lambda} = \frac{1}{L} \sum_{n=1}^{L} \mathbf{J}_n \mathbf{Q} \mathbf{J}_n \tag{4.87}$$

is a diagonal matrix such that $[\mathbf{\Lambda}]_{jj} = [\mathbf{Q}]_{jj}$, $\forall \mathbf{Q} \in \mathbb{C}^{n_T \times n_T}$.

Proof From the definition of $\mathbf{J}_n$ we readily see that $[\mathbf{J}_n \mathbf{Q} \mathbf{J}_n]_{jj} = [\mathbf{Q}]_{jj}$, $\forall n$, and that

$$[\mathbf{J}_n \mathbf{Q} \mathbf{J}_n]_{jj'} = \begin{cases} [\mathbf{Q}]_{jj'} & \text{if } [\mathbf{J}_n]_{jj} = [\mathbf{J}_n]_{j'j'} \\ -[\mathbf{Q}]_{jj'} & \text{if } [\mathbf{J}_n]_{jj} \neq [\mathbf{J}_n]_{j'j'} \end{cases} . \tag{4.88}$$

Since the sum in (4.87) is over all the possible $\{\mathbf{J}_n\}$ then, while the terms in the diagonal, $[\mathbf{J}_n \mathbf{Q} \mathbf{J}_n]_{jj}$, always contribute constructively, the terms off the diagonal, $[\mathbf{J}_n \mathbf{Q} \mathbf{J}_n]_{jj'}$, will have the same number of positive contributions than of negative ones, and, thus, the total sum vanishes. ■

Lemma 2 *The function $I_E(\mathbf{Q}, \mathbf{M})$ is invariant under the transformation $\mathbf{Q} \mapsto \mathbf{J}_n \mathbf{Q} \mathbf{J}_n$, $\forall n$, where the definition of $\mathbf{J}_n$ is the same as in Lemma 1.*

Proof We begin by noticing that

$$(\mathbf{M} \odot \mathbf{P}) \mathbf{J}_n \mathbf{Q} \mathbf{J}_n (\mathbf{M} \odot \mathbf{P})^H = (\mathbf{M} \odot \widetilde{\mathbf{P}}) \mathbf{Q} (\mathbf{M} \odot \widetilde{\mathbf{P}})^H, \tag{4.89}$$

where we have introduced the following change of variables

$$[\widetilde{\mathbf{P}}]_{ij} = \begin{cases} [\mathbf{P}]_{ij} & \text{if } [\mathbf{J}_n]_{jj} = 1 \\ [\mathbf{P}]_{ij} \, \mathrm{e}^{\mathrm{i}\pi} & \text{if } [\mathbf{J}_n]_{jj} = -1 \end{cases} . \tag{4.90}$$

Notice that the change of variable $[\widetilde{\mathbf{P}}]_{ij} = [\mathbf{P}]_{ij} \, \mathrm{e}^{\mathrm{i}\pi}$ does not affect its distribution, since we have assumed that the phase variables θ_{ij} are uniformly distributed on $[0, 2\pi)$ and consequently the distribution of $[\mathbf{P}]_{ij} = \mathrm{e}^{\mathrm{i}\theta_{ij}}$ is invariant under phase shifts.

In (4.93), the transformation $\mathbf{Q} \mapsto \mathbf{J}_n \mathbf{Q} \mathbf{J}_n$ is performed on $I_E(\mathbf{Q}, \mathbf{M})$ and it is shown that it remains invariant.

$$I_E(\mathbf{Q}, \mathbf{M}) \mapsto I_E(\mathbf{J}_n \mathbf{Q} \mathbf{J}_n, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \log \det \left(\mathbf{I} + (\mathbf{M} \odot \mathbf{P}) \mathbf{J}_n \mathbf{Q} \mathbf{J}_n (\mathbf{M} \odot \mathbf{P})^H \right) \tag{4.91}$$

$$= \mathbb{E}_{\widetilde{\mathbf{P}}} \log \det \left(\mathbf{I} + (\mathbf{M} \odot \widetilde{\mathbf{P}}) \mathbf{Q} (\mathbf{M} \odot \widetilde{\mathbf{P}})^H \right) \tag{4.92}$$

$$= I_E(\mathbf{Q}, \mathbf{M}), \tag{4.93}$$

where last equality follows from the fact that $\widetilde{\mathbf{P}}$ is a dummy integration variable. ∎

Lemma 3 *The function $I_E(\mathbf{Q}, \mathbf{M})$ is strictly concave in $\mathbf{Q}$, where $\mathbf{Q}$ belongs to the set of positive semidefinite matrices.*

Proof The function $I_E(\mathbf{Q}, \mathbf{M})$ is defined in (4.5) as the expectation with respect to $\mathbf{P}$, of a strictly concave function in $\mathbf{Q}$. Consequently, from [Boy04], $I_E(\mathbf{Q}, \mathbf{M})$ is also strictly concave. ∎

Once these lemmas have been established the optimality of a diagonal covariance matrix is pretty straightforward as it is shown next.

Proof To prove achievability we only need to show that

$$I_E(\mathbf{Q}, \mathbf{M}) = \sum_{n=1}^{L} \frac{1}{L} I_E(\mathbf{Q}, \mathbf{M}) \tag{4.94}$$

$$= \sum_{n=1}^{L} \frac{1}{L} I_E(\mathbf{J}_n \mathbf{Q} \mathbf{J}_n, \mathbf{M}) \tag{4.95}$$

$$\leq I_E \left(\frac{1}{L} \sum_{n=1}^{L} \mathbf{J}_n \mathbf{Q} \mathbf{J}_n, \mathbf{M} \right) \tag{4.96}$$

$$= I_E(\mathbf{\Lambda}, \mathbf{M}), \tag{4.97}$$

where $\mathbf{\Lambda}$ is a diagonal matrix such that $\operatorname{Tr} \mathbf{\Lambda} = \operatorname{Tr} \mathbf{Q}$. Note that (4.95) follows from Lemma 2; inequality (4.96), from Lemma 3; and, finally, we invoked Lemma 1 to write (4.97). It is now straightforward to see that, if $\mathbf{Q}$ is a solution to the optimization problem in (4.6), then there exists a diagonal matrix $\mathbf{\Lambda}$ that is also optimal or, formally,

$$\sup_{\mathbf{Q}} I_E(\mathbf{Q}, \mathbf{M}) \leq \sup_{\mathbf{\Lambda}} I_E(\mathbf{\Lambda}, \mathbf{M}), \tag{4.98}$$

where $\mathbf{\Lambda}$ is a diagonal matrix such that $\operatorname{Tr} \mathbf{\Lambda} = \operatorname{Tr} \mathbf{Q}$. ∎

Proof To prove the converse we use a similar argument as in [Eld04]. Since $I_E(\mathbf{Q}, \mathbf{M})$ is strictly concave in $\mathbf{Q}$ it has a unique global maximum, $\mathbf{Q}^\star$. From Lemma 2, we deduce that, for $\mathbf{Q}^\star$ to be unique, it has to satisfy $\mathbf{Q}^\star = \mathbf{J}_n \mathbf{Q}^\star \mathbf{J}_n$, $\forall n$, which implies that $\mathbf{Q}^\star$ has to be diagonal. ∎

4.A.2   Compound case

To prove the optimality of a diagonal covariance matrix for the compound case we follow a very similar development as in the previous section.

Lemma 4 *The function $I_C(\mathbf{Q}, \mathbf{M})$ is invariant under the transformation $\mathbf{Q} \mapsto \mathbf{J}_n\mathbf{Q}\mathbf{J}_n$, $\forall n$, where the definition of $\mathbf{J}_n$ is the same as in Lemma 1.*

Proof We recall the change of variable introduced in the proof of Lemma 2. Notice that $\mathbf{P} \in \mathcal{P} \Rightarrow \widetilde{\mathbf{P}} \in \mathcal{P}$. In (4.101), we perform the transformation $\mathbf{Q} \mapsto \mathbf{J}_n\mathbf{Q}\mathbf{J}_n$ on $I_C(\mathbf{Q}, \mathbf{M})$ and show that it remains invariant.

$$I_C(\mathbf{Q}, \mathbf{M}) \mapsto I_C(\mathbf{J}_n\mathbf{Q}\mathbf{J}_n, \mathbf{M}) = \inf_{\mathbf{P} \in \mathcal{P}} \log \det \left(\mathbf{I} + (\mathbf{M} \odot \mathbf{P})\mathbf{J}_n\mathbf{Q}\mathbf{J}_n(\mathbf{M} \odot \mathbf{P})^H\right) \tag{4.99}$$

$$= \inf_{\widetilde{\mathbf{P}} \in \mathcal{P}} \log \det \left(\mathbf{I} + (\mathbf{M} \odot \widetilde{\mathbf{P}})\mathbf{Q}(\mathbf{M} \odot \widetilde{\mathbf{P}})^H\right) \tag{4.100}$$

$$= I_C(\mathbf{Q}, \mathbf{M}), \tag{4.101}$$

where last equality follows from the fact that $\widetilde{\mathbf{P}}$ is a dummy minimization variable. ∎

Lemma 5 *The function $I_C(\mathbf{Q}, \mathbf{M})$ is strictly concave in $\mathbf{Q}$, where $\mathbf{Q}$ belongs to the set of positive semidefinite matrices.*

Proof The function $I_C(\mathbf{Q}, \mathbf{M})$ is defined in (4.62) as the pointwise infimum, with respect to $\mathbf{P} \in \mathcal{P}$, of a set of strictly concave functions in $\mathbf{Q}$. Consequently, from [Boy04, p. 81], $I_C(\mathbf{Q}, \mathbf{M})$ is also strictly concave. ∎

Again, we are now ready to prove the optimality of a diagonal covariance matrix for the maximization of the compound mutual information.

Proof To prove achievability we only need to show that

$$I_C(\mathbf{Q}, \mathbf{M}) = \sum_{n=1}^{L} \frac{1}{L} I_C(\mathbf{Q}, \mathbf{M}) \tag{4.102}$$

$$= \sum_{n=1}^{L} \frac{1}{L} I_C(\mathbf{J}_n\mathbf{Q}\mathbf{J}_n, \mathbf{M}) \tag{4.103}$$

$$\leq I_C\left(\frac{1}{L}\sum_{n=1}^{L} \mathbf{J}_n\mathbf{Q}\mathbf{J}_n, \mathbf{M}\right) \tag{4.104}$$

$$= I_C(\mathbf{\Lambda}, \mathbf{M}), \tag{4.105}$$

where $\mathbf{\Lambda}$ is a diagonal matrix such that $\operatorname{Tr}\mathbf{\Lambda} = \operatorname{Tr}\mathbf{Q}$. Note that (4.103) follows from Lemma 4; inequality (4.104), from Lemma 5; and, finally, we invoked Lemma 1 to write (4.105). It is now

straightforward to see that, if $\mathbf{Q}$ is a solution to the optimization problem in (4.63), then there exists a diagonal matrix $\boldsymbol{\Lambda}$ that is also optimal or, formally,

$$\sup_{\mathbf{Q}} I_C(\mathbf{Q}, \mathbf{M}) \leq \sup_{\boldsymbol{\Lambda}} I_C(\boldsymbol{\Lambda}, \mathbf{M}), \tag{4.106}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix such that $\text{Tr}\,\boldsymbol{\Lambda} = \text{Tr}\,\mathbf{Q}$. ∎

Proof To prove the converse a similar argument as in [Eld04] is used. Since $I_C(\mathbf{Q}, \mathbf{M})$ is strictly concave in $\mathbf{Q}$ it has a unique global maximum, $\mathbf{Q}^\star$. From Lemma 4, we deduce that, for $\mathbf{Q}^\star$ to be unique, it has to satisfy $\mathbf{Q}^\star = \mathbf{J}_n \mathbf{Q}^\star \mathbf{J}_n$, $\forall n$, which implies that $\mathbf{Q}^\star$ has to be diagonal. ∎

4.B   Alternative proof of Proposition 4.4.1

An alternative proof of Proposition 4.4.1 to that given in [Pay06f, Pay06e] is presented next. In §4.4 it was stated that the ergodic capacity was given by

$$\begin{aligned} C_E = \sup_{\mathbf{Q}} \quad & I_E(\mathbf{Q}, \mathbf{M}) \\ \text{s.\,t.} \quad & \text{Tr}\,\mathbf{Q} \leq P_T, \\ & \mathbf{Q} \succcurlyeq \mathbf{0}. \end{aligned} \tag{4.107}$$

The objective here is to show that the maximization with respect to $\mathbf{Q}$ can be restricted to the set of diagonal positive semidefinite matrices. The same procedure as in [Vis01] will be followed.

Let us first consider the following optimization problem above, restricting the maximization to the set of positive semidefinite diagonal matrices such that $\text{Tr}\mathbf{Q} = P_T$. The optimization function is strictly concave in $\mathbf{Q}$, while the constraint set is convex[1] and compact. This implies that there exists a unique $\boldsymbol{\Lambda}$, diagonal and positive semidefinite, that solves the restricted optimization problem. We now show that this $\boldsymbol{\Lambda}_{\mathbf{Q}}$ is actually the solution to the first optimization problem. Indeed, a necessary and sufficient condition for the overall optimality of $\boldsymbol{\Lambda}$ is

$$dI_E(\boldsymbol{\Lambda}, \mathbf{Q} - \boldsymbol{\Lambda}; \mathbf{M}) \leq 0, \tag{4.108}$$

where $dI_E(\mathbf{A}, \mathbf{B}; \mathbf{M})$ is the Fréchet differential of $I_E(\mathbf{Q}, \mathbf{M})$ in the direction of $\mathbf{B}$ evaluated at $\mathbf{A}$ (see further [Lue98]). Since the explicit expression for $I_E(\mathbf{Q}, \mathbf{M})$ is given by

$$I_E(\mathbf{Q}, \mathbf{M}) = \mathbb{E}_{\mathbf{P}} \log \det \left(\mathbf{I} + \sigma^{-2} (\mathbf{M} \odot \mathbf{P}) \mathbf{Q} (\mathbf{M} \odot \mathbf{P})^H \right), \tag{4.109}$$

the necessary and sufficient condition for optimality in (4.108) is equivalent to:

$$\mathbb{E}_{\mathbf{P}} \text{Tr} \left[\mathbf{H}^H \left(\mathbf{I} + \mathbf{H} \boldsymbol{\Lambda} \mathbf{H}^H \right)^{-1} \mathbf{H} \left(\mathbf{Q} - \boldsymbol{\Lambda} \right) \right] \leq 0, \tag{4.110}$$

---

[1] It is obviously bounded due to the constraint, and it is closed because of the way it is defined (preimage of a closed set by a continuous application).

where $\mathbf{H} = \mathbf{M} \odot \mathbf{P}$ has been utilized for the sake of notation.

Let $\mathbf{D_Q}$ be a diagonal matrix containing the diagonal entries of $\mathbf{Q}$, and $\tilde{\mathbf{Q}} = \mathbf{Q} - \mathbf{D_Q}$. The above condition can be expressed as

$$\mathbb{E}_{\mathbf{P}} \text{Tr} \left[\mathbf{H}^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{H} \left(\mathbf{D_Q} - \mathbf{\Lambda} \right) \right] + \mathbb{E}_{\mathbf{P}} \text{Tr} \left[\mathbf{H}^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{H}\tilde{\mathbf{Q}} \right] \le 0. \quad (4.111)$$

Notice that the first term is always negative or zero, thanks to the optimality of $\mathbf{\Lambda}$ among the set of positive semidefinite diagonal matrices and that $\text{Tr}\,\mathbf{D_Q} = \text{tr}\,\mathbf{Q} = P_T$, so that the constraint is fulfilled. As for the second term, it can be written as

$$\mathbb{E}_{\mathbf{P}} \text{Tr} \left[\mathbf{H}^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{H}\tilde{\mathbf{Q}} \right] = \sum_{i=1}^{n_T} \sum_{\substack{k=1 \\ k \ne i}}^{n_T} q_{ik} \mathbb{E}_{\mathbf{P}} \left[\mathbf{h}_i^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{h}_k \right]. \quad (4.112)$$

From symmetry considerations, notice that it is sufficient to show that, for $1 < i \le n_T$,

$$\mathbb{E}_{\mathbf{P}} \left[\mathbf{h}_i^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{h}_1 \right] = 0 \quad (4.113)$$

in order to prove that the second term in (4.111) is identically zero. With this, we will have proven that a diagonal covariance matrix $\mathbf{Q}$ is the optimum structure to attain capacity. Equation (4.113) can be alternatively written as:

$$\frac{1}{J} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \mathbf{h}_i^H \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{h}_1 \mathrm{d}\theta_{11} \ldots \mathrm{d}\theta_{n_R n_T} = 0, \quad (4.114)$$

where $J = (2\pi)^{n_R n_T}$. This is equivalent to

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \mathbf{h}_i^H \mathbf{g} \mathrm{d}\theta_{12} \ldots \mathrm{d}\theta_{n_R n_T} = 0, \quad (4.115)$$

with

$$\mathbf{g} = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left(\mathbf{I} + \mathbf{H\Lambda H}^H \right)^{-1} \mathbf{h}_1 \mathrm{d}\theta_{11} \ldots \mathrm{d}\theta_{n_R 1}. \quad (4.116)$$

Defining $\mathbf{f} = (\mathbf{I} + \mathbf{H\Lambda H}^H)^{-1}\mathbf{h}_1$, it can be noticed that each element of vector function $\mathbf{f} = [f_1 \ldots f_{n_R}]^T$ can be considered as a function of arguments $(\theta_{11}, \ldots, \theta_{n_R 1})$ with the remaining $(\theta_{12}, \ldots, \theta_{n_R n_T})$ considered constant parameters, *i.e.*, $f_k : \mathbb{R}^{n_R} \to \mathbb{C}$. Each f_k, $1 \le k \le n_R$, is a measurable function and has the following properties:

$$f_k(\theta_{11} + \pi, \ldots, \theta_{n_R 1} + \pi) = -f_k(\theta_{11}, \ldots, \theta_{n_R 1}), \quad (4.117)$$

$$f_k(\theta_{11}, \ldots, \theta_{i1} \pm 2\pi, \ldots, \theta_{n_R 1}) = f_k(\theta_{11}, \ldots, \theta_{i1}, \ldots, \theta_{n_R 1}), \quad (4.118)$$

for all $1 \le i \le n_R$. Next lemma gives a result concerning functions with the above properties.

Lemma 6 *Let $f : \mathbb{R}^n \to \mathbb{C}$ be a measurable function such that $f(\theta_1, \ldots, \theta_n) = -f(\theta_1 + \pi, \ldots, \theta_n + \pi)$ and that $f(\theta_1, \ldots, \theta_i \pm 2\pi, \ldots, \theta_n) = f(\theta_1, \ldots, \theta_i, \ldots, \theta_n)$, with $1 \le i \le n$. Then*

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_n) \, \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n = 0. \quad (4.119)$$

Proof See appendix 4.C. ∎

Last lemma implies that $\mathbf{g} = \mathbf{0}$, which is a sufficient condition for (4.115) to be true. With this, we have proven that a diagonal $\mathbf{Q}$ is enough to attain the MIMO channel capacity.

4.C   Proof of Lemma 6

Although Lemma 6 is quite intuitive, its proof is given here for completeness. The proof will be done by induction.

- $n = 1$

Let $f : \mathbb{R} \to \mathbb{C}$ be a measurable function such that $f(\theta) = -f(\theta + \pi)$ and that $f(\theta) = f(\theta \pm 2\pi)$. Then,

$$\int_{-\pi}^{\pi} f(\theta) \mathrm{d}\theta = \int_{-\pi}^{0} f(\theta) \mathrm{d}\theta + \int_{0}^{\pi} f(\theta) \mathrm{d}\theta. \tag{4.120}$$

Defining the change $\omega = \theta - \pi$ in the second term of the right hand side of last equation and using the property $f(\theta) = -f(\theta + \pi)$, we can write

$$\int_{-\pi}^{\pi} f(\theta) \mathrm{d}\theta = \int_{-\pi}^{0} f(\theta) \mathrm{d}\theta - \int_{-\pi}^{0} f(\omega) \mathrm{d}\omega = 0. \tag{4.121}$$

- n true $\to n+1$ true

Let $f : \mathbb{R}^n \to \mathbb{C}$ be a measurable function such that $f(\theta_1, \ldots, \theta_n) = -f(\theta_1 + \pi, \ldots, \theta_n + \pi)$ and that $f(\theta_1, \ldots, \theta_i \pm 2\pi, \ldots, \theta_n) = f(\theta_1, \ldots, \theta_i, \ldots, \theta_n)$, with $1 \le i \le n$, then

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_n) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n = 0. \tag{4.122}$$

Let us now consider a measurable function $f : \mathbb{R}^{n+1} \to \mathbb{C}$ such that $f(\theta_1, \ldots, \theta_{n+1}) = -f(\theta_1 + \pi, \ldots, \theta_{n+1} + \pi)$ and that $f(\theta_1, \ldots, \theta_i \pm 2\pi, \ldots, \theta_{n+1}) = f(\theta_1, \ldots, \theta_i, \ldots, \theta_{n+1})$, with $1 \le i \le n+1$. Then

$$\int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_{n+1}) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_{n+1} = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \varphi(\theta_1, \ldots, \theta_n) \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_n, \tag{4.123}$$

where $\varphi(\theta_1, \ldots, \theta_n) = \int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_{n+1}) \mathrm{d}\theta_{n+1}$. Evidently, $\varphi(\theta_1, \ldots, \theta_i \pm 2\pi, \ldots, \theta_n) = \varphi(\theta_1, \ldots, \theta_i, \ldots, \theta_n)$ holds for all i such that $1 \le i \le n$, and from the Fubini-Torelli theorem we get that $\varphi(\theta_1, \ldots, \theta_n)$ is a measurable function. Thus, from n case assumption, it is now sufficient to prove that $\varphi(\theta_1, \ldots, \theta_n) = -\varphi(\theta_1 + \pi, \ldots, \theta_n + \pi)$. By definition

$$\varphi(\theta_1 + \pi, \ldots, \theta_n + \pi) = \int_{-\pi}^{\pi} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \theta_{n+1}) \mathrm{d}\theta_{n+1}. \tag{4.124}$$

Considering the change $\omega = \theta_{n+1} - \pi$ last equation reads

$$\varphi(\theta_1 + \pi, \ldots, \theta_n + \pi) =$$
$$= \int_{-2\pi}^{-\pi} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \omega + \pi) \mathrm{d}\omega + \int_{-\pi}^{0} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \omega + \pi) \mathrm{d}\omega. \quad (4.125)$$

Using the change $\varpi = \omega + 2\pi$ in the first term of last equation and the property $f(\theta_1, \ldots, \theta_i \pm 2\pi, \ldots, \theta_{n+1}) = f(\theta_1, \ldots, \theta_i, \ldots, \theta_{n+1})$, with $1 \leq i \leq n$, we get

$$\varphi(\theta_1 + \pi, \ldots, \theta_n + \pi) = \int_{0}^{\pi} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \varpi + \pi) \mathrm{d}\varpi +$$
$$+ \int_{-\pi}^{0} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \omega + \pi) \mathrm{d}\omega = \int_{-\pi}^{\pi} f(\theta_1 + \pi, \ldots, \theta_n + \pi, \omega + \pi) \mathrm{d}\omega. \quad (4.126)$$

Using that $f(\theta_1, \ldots, \theta_{n+1}) = -f(\theta_1 + \pi, \ldots, \theta_{n+1} + \pi)$ we can finally write

$$\varphi(\theta_1 + \pi, \ldots, \theta_n + \pi) = -\int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_n, \omega) \mathrm{d}\omega = -\varphi(\theta_1, \ldots, \theta_n). \quad (4.127)$$

4.D   Roots of polynomial (4.34) in (ξ, γ)

Let us express (4.34) as
$$(x^2 \Delta^2 - \xi^2) = \frac{4x\xi(x - \xi)(x + \Upsilon)}{x^2 - \gamma^2}. \quad (4.128)$$

When $\xi < x < \gamma$ the derivative of the left hand side of last equation is $2x\Delta^2 > 0$, while the derivative of the right hand side can be expressed as

$$4\xi \frac{(x - \xi)(x + \Upsilon) + x(x + \Upsilon) + x(x - \xi)}{x^2 - \gamma^2} - \frac{8x^2\xi(x - \xi)(x + \Upsilon)}{(x^2 - \gamma^2)^2} < 0. \quad (4.129)$$

Consequently, it can be seen that for the range of values of x under consideration, the left hand side of (4.128) is monotonically increasing, while the right hand side is monotonically decreasing. Since there is always a solution within $\xi < x < \gamma$ for which these two side coincide, the solution must be unique.

Chapter 5

From single to multi-user communication with imperfect CSI

5.1   Introduction

Following the notions and ideas presented in last chapter, in this chapter we study the performance, in terms of mutual information (or achievable rates), of a particular single-user transceiver architecture, the so-called STHP in the presence of errors in the CSI. One of the main advantages of this scheme is that its design can be very easily extended to the multi-user case, *i.e.*, a situation where more than one user is receiving information from the transmitter as described in §2.1.2. This study will make us wonder what is the impact of an imperfect CSI in a more general multi-user scenario, which will be discussed in more detail in chapter 6.

5.2   Background

In 1983, Costa surprised the electrical engineering research community with a result, [Cos83], which loosely speaking stated that the capacity of an interfered communications system, such that the interference is known non-causally at the transmitter, is the same as if the interference were not present. He also showed that the capacity can be achieved by a transmission scheme which is, since then, known as "dirty paper" coding (DPC). Costa used an analogy to give an intuitive idea of this coding technique: in his particular vision, the signal space is a blank paper, the interference is dirt in this paper, and the transmitted signal is ink. The optimal strategy is not to clean the paper and then write, but to take advantage of the dirt by writing as much aligned with it as possible. From this point of view, the optimal "dirty paper" strategy can be seen as the transmitter counterpart of the ML receiver, whose philosophy is to estimate the data symbols without trying to cancel or equalize the incoming interferences.

This pioneering work by Costa and recent theoretical results describing the sum-capacity

when using multiple antennas to communicate with multiple users in a known rich scattering channel [Cai03, Vis03b, Yu04, Vis03a, Wei06], have motivated a significant research on MIMO downlink transmit strategies. One of such downlink strategies [Fis02a, Win04a] was the extension of the non-linear precoding scheme for temporal intersymbol interference mitigation proposed by Tomlinson [Tom71] and Harashima [Miy72, Har72], into spatial interference equalization for MIMO systems, leading to the STHP scheme. The main idea behind STHP is to pre-subtract the interference that is produced by the presence of the MIMO channel before the signal is actually transmitted, which turns out to be computationally very simple to implement.

In addition to STHP, other downlink transmission strategies have been recently proposed, such as the transmission architecture for the MIMO broadcast channel developed by Peel, Hochwald, and Swindlehurst [Pee05a, Pee05b]. In their work, they proposed a transmission scheme based on a two-step process. In the first step, the channel is partially equalized utilizing a regularized inversion that improves the performance, specially at low SNRs. In the second step, a vector perturbation, which is calculated using a combinatorial search, is applied to the transmitted signal to obtain good performance at all SNRs. This scheme is also known as sphere encoder, because it is the transmitter version of the sphere decoder [Fin85]. A similar, but computationally less demanding scheme, was proposed in [Win04b] where optimality was sacrificed to reduce the computational load. There, the key idea was to replace the vector perturbation technique with Babai's closest-point approximate solution which offers an excellent tradeoff between complexity and optimality loss. An analogous approach was taken in [Sha05a], where the authors proposed another alternative method to calculate the vector perturbation, based on the MMSE feedback precoder. Other downlink schemes include the STHP optimized with respect to the sum of mean square errors as given in [Joh04], or the non-linear precoding scheme with vector perturbation described in [Sch05] which shows the best performance among all the cited schemes.

A commonality among all these transmission architectures is that, since the precoder has to be matched to the channel, some degree of CSI has to be made available at the transmitter. In this chapter it is assumed that this CSI is imperfect, due to errors in the estimation of the channel or due to the presence of a noisy feedback channel, which implies that there is a performance degradation of the system. The effects of imperfect CSI for STHP in a single-user scenario, in terms of SNR loss, were studied in [Fis02b]. The authors in [Die05] proposed a robust design of the transceiver matrices in STHP in a multi-user broadcast scenario taking a Bayesian modeling of the errors in the CSI and looking for the minimization of the sum of the MSE for all the users. An additional zero-forcing constraint but keeping the same design objective and error modeling was given in [Hun04].

In this chapter, the problem of evaluating the *rate* (or mutual information) loss of STHP due to small errors in the available CSI both for the single and multi-user cases is addressed from an

information theoretic point of view. A maximin *robust* design of the set of moduli used in STHP is presented and the corresponding power allocation among the transmitted data streams, that maximizes the worst-case achievable rates, is derived. The choice of analyzing the STHP scheme is due to the fact that it is a widely studied transmission architecture, which can be implemented in real systems because it requires a small amount of computational effort and also for reasons of simplicity of analysis.

5.3   System model

The downlink of a flat fading wireless communications system with n_T antennas at the BS and n_R antennas at the receiver side is considered. Recalling the model reviewed in §2.1.4, the received signal vector $\mathbf{y} \in \mathbb{C}^{n_R}$ is given by

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

where $\mathbf{x} \in \mathbb{C}^{n_T}$ is the transmitted signal vector, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ represents the channel matrix, and $\mathbf{n} \in \mathbb{C}^{n_R}$ is the noise vector. In this case, the entries of the noise vector are considered to be proper complex Gaussian random variables with zero mean and variance σ^2, *i.e.*, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{n_R})$.

As it was stated in §5.2, the single and multi-user scenarios are analyzed in the following sections. In the single-user case, see Figure 5.1, the i-th element of the received vector, y_i, represents the received signal at the i-th antenna of that single-user, *i.e.*, there is only one user with a receiver equipped with n_R antennas. In the multi-user case, see Figure 5.2, y_i stands for the received signal by i-th user, *i.e.*, there are n_R users with single antenna terminals. In both cases, the n_T transmitting antennas belong to the BS. Moreover, the BS is equipped with STHP, which is a non-linear transformation applied to the data symbols. The output of this transformation is matched to the channel in such a way that the interference caused by the channel matrix once the signal is transmitted is precisely the inverse of the precoding transformation. Consequently, the original data symbols can be easily recovered at the receiving end. Notice that, an additional linear processing matrix placed at the receiver (single-user) or at the transmitter (multi-user) is necessary to fully invert the STHP transformation. A full description of this transformation is detailed in the following section.

5.3.1   The spatial Tomlinson-Harashima precoder

The input of the STHP is a vector, $\mathbf{s} \in \mathbb{C}^{n_S}$, whose entries are the n_S data symbols that are to be transmitted, s_k, $k = 1, \ldots, n_S$. In (5.2), the output of the precoder, $\mathbf{p} \in \mathbb{C}^{n_S}$, is obtained from these data symbols in a similar way to the recursive causal spatial relation described in
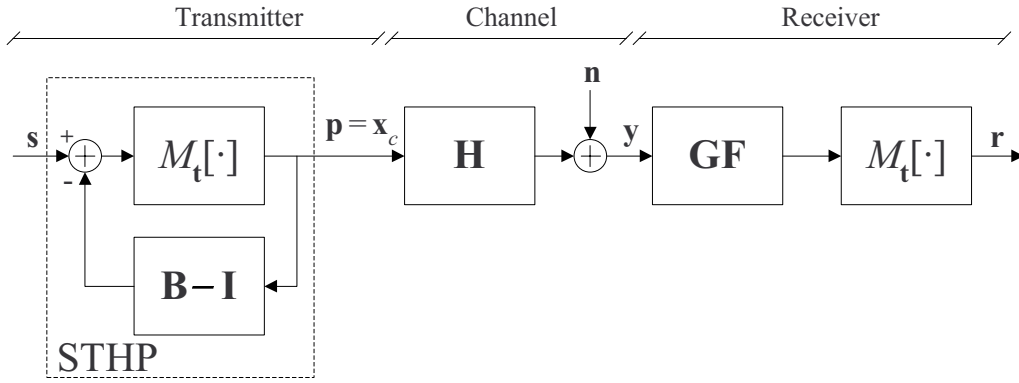
Figure 5.1: Single-user communication scheme with STHP.
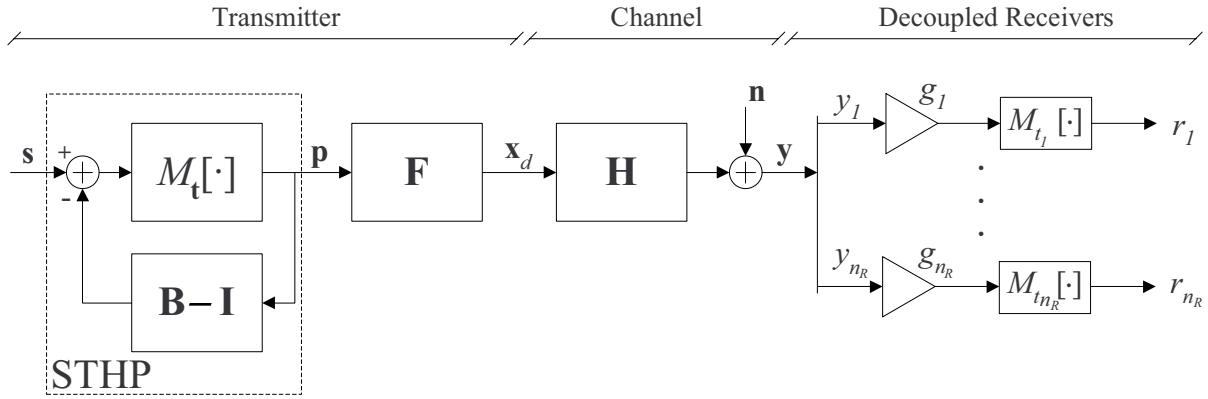
Figure 5.2: Multi-user communication scheme with STHP.

[Fis02a].

$$p_k = [\mathbf{p}]_k = M_{t_k}\left[s_k - \sum_{k'=1}^{k-1} b_{kk'}p_{k'}\right], \quad \text{for } k = 1,\ldots,n_S. \tag{5.2}$$

In (5.2), $b_{kk'}$ are a set of coefficients to be determined, and $M_{t_k}[p]$ represents a complex modulo reduction of p into the complex square region $\mathcal{D}_k \triangleq [-t_k, t_k) \times [-\mathrm{i}\,t_k, \mathrm{i}\,t_k)$, which is used to limit the dynamic range of the output signal to overcome the problem of the increase of the transmitted power in precoding schemes where the channel response is inverted. Notice that the main difference between (5.2) and the scheme in [Fis02a] is that our scheme allows different moduli, t_k, for each component of the precoder output $\mathbf{p}$.

An alternative and more compact notation can be used for the output of the precoder by arranging the coefficients $b_{kk'}$ into the elements below the diagonal of a lower triangular matrix $\mathbf{B} \in \mathbb{C}^{n_S \times n_S}$ with ones in the main diagonal, such that $[\mathbf{B}]_{kk'} = b_{kk'}$. The output of the precoder is then

$$\mathbf{p} = \mathbf{B}^{-1}(\mathbf{s} + \mathbf{a}), \tag{5.3}$$

where the elements of vector $\mathbf{a}$ must be of the form $a_k = 2t_k u_k + \mathrm{i}\, 2t_k v_k$ for some $u_k, v_k \in \mathbb{Z}$, which must be chosen such that the resulting elements of $\mathbf{p}$ lie inside the modulo region defined by the elements of vector $\mathbf{t}$, *i.e.*, $p_k \in \mathcal{D}_k$.

In the following two subsections we will consider the cases mentioned earlier, namely: (a) single-user (coupled receivers), *i.e.*, the different receivers are allowed to cooperate by jointly processing the received data; (b) multi-user (decoupled receivers), *i.e.*, the receiving antennas belong to different users and therefore only individual processing of the received symbols is possible. It is also important to mention (as it was similarly noted in [Win04a]) that, while the coupled receivers structure can never model a multi-user situation, the decoupled receivers structure can also be applied in the single-user scenario. Consequently, in the single-user case, the optimum design would be to have a versatile transceiver structure that could switch between the coupled and decoupled architectures depending on which of them yields a better performance for a given channel realization. Bearing this fact in mind, we consider in the following, the analysis of the coupled and decoupled receivers structures.

5.3.2 Single-user scenario (coupled receivers)

For the case where the receivers are coupled (see Figure 5.1), the transmitted signal vector, $\mathbf{x}_c$, is

$$\mathbf{x}_c = \mathbf{p} = \mathbf{B}^{-1}(\mathbf{s} + \mathbf{a}). \tag{5.4}$$

From dimensional analysis, last equation implies that the number of transmit antennas has to be equal to or higher than the number of transmitted signals, *i.e.*, $n_T \geq n_S$. If the number of antennas is higher than n_S then, a subset of n_S antennas is selected, where the selection is made according to an achievable rates maximization criterion as it is commented later. Moreover, since the single-user has to recover all the n_S transmitted symbols, its mobile terminal has to be equipped with $n_R \geq n_S$ antennas. Let us assume that the selected antennas are numbered from 1 to n_S, then, the transmitted power for the coupled case, P_T^c, is calculated as

$$P_T^c = \mathbb{E}\, \mathbf{x}_c^H \mathbf{x}_c = \sum_{k=1}^{n_S} P_k, \tag{5.5}$$

where P_k is the power transmitted through k-th antenna.

The design of the transmitter and the receiver consists in specifying the precoding matrix $\mathbf{B}$ at the transmitter side and the product of matrices, $\mathbf{GF}$ at the receiver end. These matrices were found in [Win04a] for a zero-forcing criterion[1], and here we recall their results. From the

---

[1]Notice that, since we are interested in high data transmission rates, the SNR can be considered to be high enough so that the zero-forcing and the MMSE criteria yield the same result. A similar argument is considered in [Win04a].

ql-factorization of the channel,

$$\mathbf{H} = \mathbf{F}^H\mathbf{S}, \tag{5.6}$$

where $\mathbf{F}$ is unitary and $\mathbf{S}$ is lower triangular, we obtain

$$s_{kk} = [\mathbf{S}]_{kk}, \quad \mathbf{G} = \mathbf{diag}(\{s_{kk}^{-1}\}), \quad \mathbf{B} = \mathbf{GS}. \tag{5.7}$$

With these definitions, we reproduce in (5.8) the estimate, $\mathbf{r} \in \mathbb{C}^{n_S}$, of the data symbols, $\mathbf{s}$.

$$\mathbf{r} = M_\mathbf{t}\left[\mathbf{GFHx}_c + \mathbf{GFn}\right] = M_\mathbf{t}\left[\mathbf{Bx}_c + \mathbf{GFn}\right] = M_\mathbf{t}\left[\mathbf{s} + \mathbf{GFn}\right]. \tag{5.8}$$

Notice that $M_\mathbf{t}[\mathbf{z}]$ performs a modulo-t_k operation for each element z_k. Since $\mathbf{F}$ is a unitary transformation matrix, in (5.8) it is possible to define a new noise vector $\tilde{\mathbf{n}} = \mathbf{Fn}$ with the same statistical behavior as the original one, $i.e.$, $\tilde{\mathbf{n}} \sim \mathbf{n}$. In addition, since $\mathbf{G}$ is a diagonal matrix with $g_{kk} = [\mathbf{G}]_{kk}$, the estimate of the data symbols vectors $\mathbf{r}$ can be expressed in the form of n_T parallel data streams as

$$r_k = M_{t_k}\left[s_k + g_{kk}\tilde{n}_k\right], \quad k = 1, \ldots, n_S. \tag{5.9}$$

This model of parallel data streams clarifies the purpose of utilizing a precoder to presubtract the intersymbol interference that is caused by the channel matrix: we obtain n_S parallel (or independent) data streams between the transmitter and the receiver. Note that the quality of the data streams is dictated by the components of the diagonal of the lower triangular matrix $\mathbf{S}$, which fulfill that $s_{kk} \neq 0$, $\forall k$, as

$$\prod_{k=1}^{n_S} s_{kk}^2 = \det(\mathbf{S}^H\mathbf{S}) = \det(\mathbf{H}^H\mathbf{H}), \tag{5.10}$$

and $\mathbf{H}$ is assumed full rank.

5.3.3   Multi-user scenario (decoupled receivers)

In the multi-user scenario, it is necessary that the receivers be decoupled (see Figure 5.2), or alternatively, since the components of the received vector may be located at geographically separated places, only individual processing of each element of the received signal vector is permitted. Last statement implies that the receiver processing matrix must be diagonal. Bearing this constraint in mind, in [Win04a], it was found that the unitary matrix $\mathbf{F}$ must be placed at the transmitter side to linearly process the output of the STHP, obtaining the expression in (5.11) for the transmitted signal. Notice that the diagonal processing matrix $\mathbf{G}$ remains at the receiving end to adjust the dynamic range of the received signal.

$$\mathbf{x}_d = \mathbf{Fp}. \tag{5.11}$$

Since in the decoupled case n_R represents the number of users, we assume $n_R = n_S$, so that each user receives its own data symbol. In addition, if the BS has to transmit these n_S symbols simultaneously, then $n_T \geq n_S$ must be guaranteed. Notice that, since $\mathbf{F}$ is a unitary matrix, the transformation in (5.11) does not increase the transmitted power, P_T^d, with respect to the coupled case, P_T^c,

$$P_T^d = \mathbb{E}\,\mathbf{x}_d^H\mathbf{x}_d = \mathbb{E}\,\mathbf{p}^H\mathbf{F}^H\mathbf{F}\mathbf{p} = \mathbb{E}\,\mathbf{p}^H\mathbf{p} = \mathbb{E}\,\mathbf{x}_c^H\mathbf{x}_c = P_T^c, \tag{5.12}$$

and consequently, in the sequel $P_T = P_T^c = P_T^d$ will be used.

Again, following [Win04a], the design, for a zero-forcing criterion, of the matrices at the transmitter side ($\mathbf{B}$ and $\mathbf{F}$) and at the receiver side, $\mathbf{G}$, is based, in this case on the *lq*-factorization in (5.13) of the channel matrix:

$$\mathbf{H} = \mathbf{SF}^H, \tag{5.13}$$

where $\mathbf{S}$ is a lower triangular matrix and $\mathbf{F}$ is unitary. The definitions

$$s_{kk} = [\mathbf{S}]_{kk}, \quad \mathbf{G} = \mathbf{diag}(\{s_{kk}^{-1}\}), \quad \mathbf{B} = \mathbf{GS}. \tag{5.14}$$

are identical as in the previous section. With these definitions, and some algebra it can be found that the estimate of the data symbols vector is given by

$$\mathbf{r} = M_\mathbf{t}\left[\mathbf{s} + \mathbf{Gn}\right]. \tag{5.15}$$

Analogously as in (5.9), the estimate, $\mathbf{r}$, of the data symbols vector can be expressed in terms of n_S parallel streams, as

$$r_k = M_{t_k}\left[s_k + g_{kk}n_k\right], \quad k = 1, \ldots, n_S. \tag{5.16}$$

Noteworthy, the models presented in (5.9) and (5.16) enable a scalable structure in the number of active streams. In this sense, the k-th stream can be disabled by letting $t_k = 0$, and (5.9) and (5.16) still hold.[2] As it will be shown in §5.4.2, while in the coupled case, letting $t_k = 0$ means that the k-th transmit antenna is switched off; in the decoupled scenario, it implies that k-th user does not receive its data symbol. Consequently, there is an analogy between the number of active transmit *antennas* and the actual number of data streams in the single-user scenario; and between the number of active *users* and the number of data streams in the multi-user case.

From this point, since the information-theoretic analysis in the case of perfect CSI is based on the expressions (5.9) and (5.16) and they represent the same input-output relation, in the next section, a unified approach will be taken.

---

[2]If a stream is to be disabled, the optimal procedure would be to consider a previous spatial ordering block in the communications diagram. This spatial ordering block would indicate which would be the best stream to be disabled. The best stream to be disabled would be the stream that when disabled produced an equivalent channel with lowest values of $\{g_{kk}\}$, *i.e.*, with the lowest noise powers at the receiver.

5.4 Information-theoretic analysis with perfect CSI

In this section the maximum rates that can be achieved when communicating utilizing the STHP scheme are analyzed. Apart from the analysis in itself, one of the novelties in this section is that the design parameters are the set of variables $\{t_k\}$, which control the modulo operations in the STHP system. In addition, the objective function to be optimized is the achievable rates, as opposed to the minimization of the MSE, which is widely considered in the literature.

5.4.1 Rate per stream

The unified model deduced in (5.9) and (5.16) describes the input-output relation of a set of parallel communication streams. Assuming independence between the elements in the data symbols vector, $\mathbf{s}$, the mutual information between the transmitted symbols and the output of the receiver end, $\mathbf{r}$, can be expressed as the sum of the mutual information between the elements of each vector, $I(\mathbf{s}; \mathbf{r}) = \sum_{k=1}^{n_S} I(s_k; r_k)$. Each term in the sum is independently upper-bounded by [Wes98]

$$I(s_k; r_k) \leq 2\log(2t_k) - h(M_{t_k}[g_{kk}n_k]), \tag{5.17}$$

where $h(\cdot)$ denotes differential entropy (see further [Cov91]), so that $h(M_{t_k}[g_{kk}n_k])$ represents the differential entropy of a modulo-$2t_k$ complex Gaussian random variable. The reader is referred to Appendix 5.A for a description on the properties of $h(M_{t_k}[g_{kk}n_k])$.

Now the aim is to find the maximum achievable rates for each independent data stream, and this rate is denoted by $R(t_k, g_{kk}\sigma)$, whose dependence on t_k and the noise standard deviation $g_{kk}\sigma$ has been explicitly indicated. One would want $R(t_k, g_{kk}\sigma)$ to be equal to the bound in (5.17) and this can be achieved by choosing s_k to be a random variable drawn from a uniform distribution in the complex square region $\mathcal{D}_k$ [Wes98]. For such a case, the rate corresponding to the k-th data stream, $R(t_k, g_{kk}\sigma)$ equals the maximum achievable mutual information, *i.e.*,

$$R(t_k, g_{kk}\sigma) = 2\log(2t_k) - h(M_{t_k}[g_{kk}n_k]), \tag{5.18}$$

where $g_{kk}\sigma$ is the square root of the variance of $g_{kk}n_k$. Noteworthy, the only degree of freedom that we have to control the rate of k-th data stream is through t_k, because $g_{kk}\sigma$ is externally determined by the zero-forcing design in (5.7), and the Gaussian noise determines σ.

5.4.2 Power per stream and total transmitted power

From what was stated in §5.4.1, we now assume that the data symbols, s_k, are chosen to be uniformly distributed in the complex region $\mathcal{D}_k$ so that the mutual information per stream is maximized. This implies that the power assigned to k-th stream is $\mathbb{E}|s_k|^2 = 2t_k^2/3$.

For this particular distribution of s_k, it can be shown that the entries, p_k, of the output vector of the precoder, $\mathbf{p}$, are also independent random variables following a uniform distribution in $\mathcal{D}_k$, i.e., $p_k \sim s_k$ (see Appendix 5.B for details).

Since in the coupled case the transmitted signal is directly the output of the precoder, $\mathbf{x}_c = \mathbf{p}$, the entries of the transmitted signal will also be uniformly distributed in $\mathcal{D}_k$ and therefore the peak power and dynamic range of the transmitted signal are limited. We define

$$P_k^c = \mathbb{E}\left[\mathbf{x}_c\mathbf{x}_c^H\right]_{kk} \tag{5.19}$$

as the power transmitted through k-th antenna in the coupled case. It is straightforward to prove that, in the coupled case, the power transmitted through k-th antenna coincides with the power assigned to k-th stream,

$$\mathbb{E}|s_k|^2 = \mathbb{E}|p_k|^2 = \mathbb{E}\left[\mathbf{x}_c\mathbf{x}_c^H\right]_{kk} = P_k^c = 2t_k^2/3. \tag{5.20}$$

In addition, the total transmitted power is the sum of the individual contributions of each antenna,

$$P_T = \sum_{k=1}^{n_S} \frac{2t_k^2}{3}. \tag{5.21}$$

From what was seen in §5.3.3, the transmitted power for the decoupled case is the same as in the coupled case, P_T, consequently (5.21) is still valid for the decoupled case. However, the terms in the summation in (5.21) are not individually associated to the power transmitted through any particular antenna, which we shall denote as P_k^d. In the decoupled case, the power transmitted through k-th antenna is

$$P_k^d = \mathbb{E}\left[\mathbf{x}_d\mathbf{x}_d^H\right]_{kk} = \mathbb{E}\sum_{k'=1}^{n_S} f_{kk'}p_{k'} \sum_{k''=1}^{n_S} f_{kk''}^* p_{k''}^* = \sum_{k'=1}^{n_S} \frac{2t_{k'}^2}{3}|f_{kk'}|^2, \tag{5.22}$$

where $f_{kk'} = [\mathbf{F}]_{kk'}$. Notice that we have used that $\mathbb{E}p_k p_{k'}^* = \delta_{kk'}2t_k^2/3$, as we show in Appendix 5.B.

5.4.3 Achievable rates with STHP

From what was stated in the last two sections, we can now state the problem of obtaining the maximum achievable rates for a system with STHP. The maximum achievable rates are the maximum of the sum of the rates of each stream subject to a total transmitted power constraint,

$$\begin{aligned} C_{\text{THP}} = \max_{\{t_k\}} \quad & \sum_{k=1}^{n_S} R(t_k, g_{kk}\sigma), \\ \text{s.t.} \quad & \sum_{k=1}^{n_S} \frac{2t_k^2}{3} = P_T. \end{aligned} \tag{5.23}$$

This problem is, in general, non-convex in the design variables, $\{t_k\}$, due to the fact that $R(t_k, g_{kk}\sigma)$ is non-convex in t_k. This implies that the problem in (5.23) has to be solved by exhaustive search or by some non-convex optimization algorithm. However, it turns out that, in the high SNR regime, the problem becomes convex and can be solved very easily in a closed form. In addition, the high SNR solution inspires a quasi-optimal solution valid for all SNRs.

Solution in the high SNR regime

In the high SNR regime, P_T can be chosen as large as needed, which implies that the values of t_k in the constraint in (5.23) can also be considered to be unbounded. Clearly, for high values of t_k, the effect of the modulo reduction, M_{t_k}, in the random variable $M_{t_k}[g_{kk}\tilde{n}_k]$ is lost. This means that the differential entropy in (5.17) will tend to the entropy of a complex Gaussian random variable, *i.e.*,

$$h(M_{t_k}[g_{kk}n_k]) \rightarrow h(g_{kk}n_k) = \log\left(g_{kk}^2\sigma^2\pi e\right), \tag{5.24}$$

(see Appendix 5.A). Consequently, the rate of k-th stream tends to

$$R(t_k, g_{kk}\sigma) \rightarrow 2\log\left(2t_k\right) - \log\left(g_{kk}^2\sigma^2\pi e\right), \tag{5.25}$$

which is the mutual information when the input is uniformly distributed in $\mathcal{D}_k$ and the noise is Gaussian with power given by $g_{kk}^2\sigma^2$. In the high SNR regime, the problem in (5.23) can be thus reformulated as

$$
\begin{aligned}
C_{\text{THP}}^{\infty} = \max_{\{t_k\}} \quad & \sum_{k=1}^{n_S} 2\log\left(2t_k\right) - \log\left(g_{kk}^2\sigma^2\pi e\right), \\
\text{s.t.} \quad & \sum_{k=1}^{n_S} \frac{2t_k^2}{3} = P_T.
\end{aligned}
\tag{5.26}
$$

The solution to this problem is easily obtained by letting all the design variables take the same value, $t_k = \sqrt{3P_T/2n_S}$, which is determined so that the power constraint is fulfilled. Substituting this solution in (5.26) we may rewrite C_{THP}^{∞} as

$$C_{\text{THP}}^{\infty} = n_S \log\left(\frac{6P_T}{n_S\sigma^2\pi e}\right) + \sum_{k=1}^{n_S} \log\left(g_{kk}^{-2}\right) = n_S \log\left(\frac{6P_T}{n_S\sigma^2\pi e}\right) + \log\left(\prod_{k=1}^{n_S} g_{kk}^{-2}\right), \tag{5.27}$$

where g_{kk}^{-1} are the elements of the triangular matrix $\mathbf{S}$, such that $\mathbf{S}^H\mathbf{S} = \mathbf{H}^H\mathbf{H}$, therefore it follows that

$$\prod_{k=1}^{n_S} g_{kk}^{-2} = \det\left(\mathbf{H}^H\mathbf{H}\right), \tag{5.28}$$

and we can now write

$$C_{\text{THP}}^{\infty} = n_S \log\left(\frac{P_T}{n_S\sigma^2}\right) - n_S \log\left(\frac{\pi e}{6}\right) + \log\det\left(\mathbf{H}\mathbf{H}^H\right), \tag{5.29}$$

where we have explicitly indicated the minus sign in the second term. Recall that the maximum achievable rates in the high SNR regime when the STHP is *not* present coincides with the MIMO channel capacity for the single-user case, *i.e.*,

$$C = n_S \log \left(\frac{P_T}{n_S \sigma^2} \right) + \log \det \left(\mathbf{H} \mathbf{H}^H \right). \tag{5.30}$$

Finally, comparing the expressions for the system capacity, C, and the maximum achievable rates with STHP, C_{THP}^∞, we see that there is a constant gap between them given by

$$\Delta C = C - C_{\text{THP}} = n_S \log \left(\frac{\pi e}{6} \right) \simeq 0.353 \, n_S \quad \text{(nats)}. \tag{5.31}$$

With this last expression we have obtained the spatial version of the shaping loss for temporal Tomlinson-Harashima precoding in [Wes98]. Notice that, although ΔC does not decrease as the SNR tends to infinity, its relative importance, defined as $\Delta C/C$, does, as both C and C_{THP} tend to grow without bound.

Suboptimal solution

In the previous section we found that, in the high SNR regime, the solution to the problem in (5.23) consists in letting all t_k take the same value. This solution can shed some light to find a quasi-optimal solution for the problem in (5.23) that can be valid at all SNRs. We consider the two following approximations for the function $R(t_k, g_{kk}\sigma)$, (see Appendix 5.A):

- For low values of $t_k/(g_{kk}\sigma)$,

$$R(t_k, g_{kk}\sigma) \simeq 2 \log 2t_k - 2 \log 2t_k = 0. \tag{5.32}$$

  Note that the entropy of $M_{t_k}[g_{kk}\tilde{n}_k]$ in (5.18) tends to that of a complex uniform random variable.

- For high values of $t_k/(g_{kk}\sigma)$,

$$R(t_k, g_{kk}\sigma) \simeq 2 \log(2t_k/(g_{kk}\sigma)) - \log(\pi e), \tag{5.33}$$

  because $h(M_{t_k}[g_{kk}\tilde{n}_k])$ in (5.18) tends to the entropy of a complex Gaussian random variable with variance $g_{kk}^2\sigma^2$.

See Figure 5.3 for a graphical representation of the function $R(t_k, g_{kk}\sigma)$ and these two approximations. Notice that, for values of $t_k/(g_{kk}\sigma) > \sqrt{\pi e/4} \simeq 1.46$, the function $R(t_k, g_{kk}\sigma)$ is better characterized by its high $t_k/(g_{kk}\sigma)$ approximation than by its low argument approximation. The opposite happens for $t_k/(g_{kk}\sigma) < \sqrt{\pi e/4}$.

Let us now suppose that all the available streams are classified into two sets ($\mathcal{G}$-Gaussian and $\mathcal{U}$-uniform) depending on which approximation of the two presented above is more accurate for
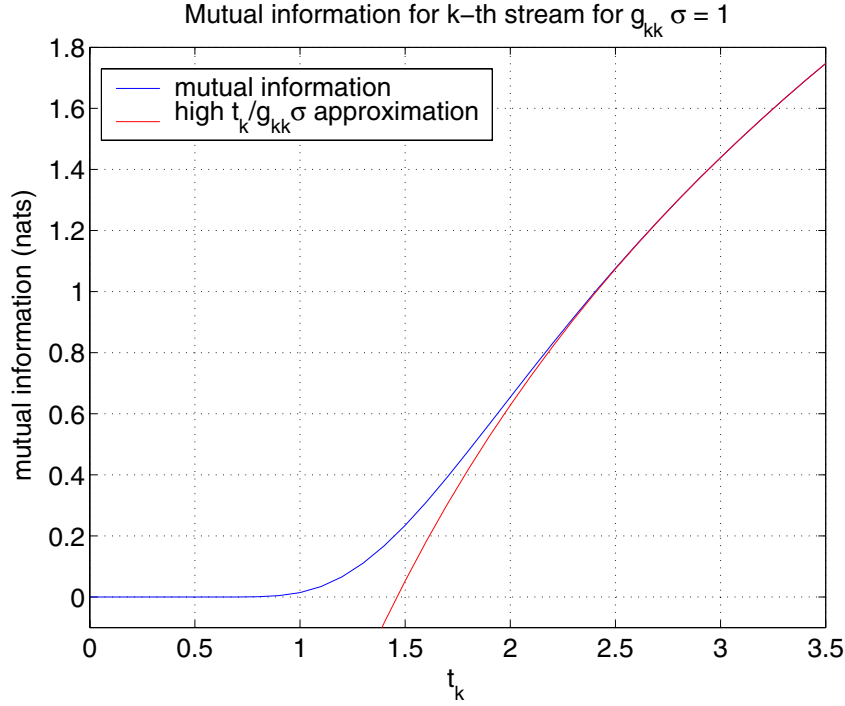
Figure 5.3: Graphical representation of the function $R(t_k, g_{kk}\sigma)$ for $g_{kk}\sigma = 1$ as a function of t_k and its approximation for high values of $t_k/(g_{kk}\sigma)$. Notice also that for low values of $t_k/(g_{kk}\sigma)$ the function $R(t_k, g_{kk}\sigma)$ is very well approximated by $R(t_k, g_{kk}\sigma) \simeq 0$.

its rate. As the contribution to the system rate (5.23) of streams belonging to $\mathcal{U}$ tends to zero no power is assigned to these streams, so that no power is wasted. Thus, the problem becomes

$$C_{\text{THP}} \simeq \max_{\{t_{k \in \mathcal{G}}\}} \quad \sum_{k \in \mathcal{G}} 2\log(2t_k/(g_{kk}\sigma)) - \log(\pi e),$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{G}} P_k = \frac{2}{3}\sum_{k \in \mathcal{G}} t_k^2 = P_T, \tag{5.34}$$

$$t_k = 0, \quad \forall k \in \mathcal{U}$$

The solution to the maximization problem above can be found very easily and it corresponds to assigning the same power to all the streams in $\mathcal{G}$ (and no power to those in $\mathcal{U}$), *i.e.*,

$$k \in \mathcal{G} \rightarrow t_k = \sqrt{3P_T/(2|\mathcal{G}|)},$$

$$k \in \mathcal{U} \rightarrow t_k = 0, \tag{5.35}$$

where $|\mathcal{G}|$ denotes the number of active streams. The set of active streams $\mathcal{G}$ is found by sorting the set $\{g_{kk}\}_{k=1}^{n_S}$ in increasing order and constructing the new sets $\{g_{kk}\}_{k=1}^{N}$, for $N = 1, \ldots, n_S$. For each set, the total power, P_T, is uniformly distributed among the streams and the Gaussian assumption is checked for each stream. The set of active streams, $\mathcal{G}$, is defined as the set that yields the highest achievable rates and all its streams fulfill the Gaussian assumption.

As it is clear from (5.35), the solution consists in performing uniform power allocation among the set of active streams (stream selection). We have found this simple transmit strategy as a result of a rate optimization problem, without imposing it as an *a priori* structure. In the case of coupled receivers, since each stream is transmitted through a different antenna, when a stream is not active it means that its corresponding antenna is switched off, in an analogous way to the well known antenna-selection algorithms, *e.g.*, [Gor00]. In the case of decoupled receivers (multi-user scenario), each stream corresponds to a particular user, consequently, we have to select a group of active *users*, which simultaneously receive their transmitted symbols.

Noteworthy, since some of the streams may be switched off as indicated by (5.35), a new channel matrix without the elements corresponding to the disabled streams *can* be considered. In this case, its corresponding Cholesky decomposition gives different values of $\{g'_{kk}\}$, $k \in \mathcal{G}$, which are lower than the original ones, $\{g_{kk}\}$, [Gol96], and then the Gaussian approximations for these streams continue to be valid, giving the same solution as obtained in (5.35).

5.5   Rate loss with imperfect CSI

In the previous section we have analyzed the achievable rates of the STHP structure for the case where the transmitter can be designed with perfect CSI. In practice, however, this situation is fictional and the transmitter usually acquires knowledge of the current channel realization through an imperfect feedback link (or by estimation of the reciprocal channel in TDD systems).

In this section we will consider that there is a mismatch between the ideal transmitter design (as given by $\mathbf{B}$ in the coupled case, and by $\mathbf{B}$ and $\mathbf{F}$ in the decoupled one) and the actual transmitter which we shall denote by $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$, which are erroneous versions of their ideal counterparts. Precisely, the main contribution of this section is that in both cases (coupled and decoupled) we obtain a robust design, whose objective is the maximization of the sum rate in the presence of errors in the CSI. In the coupled case (see Figure 5.1), we consider that the transmitter is informed directly of $\widehat{\mathbf{B}}$ through a feedback link with errors. Noteworthy, since $\mathbf{B}$ is a lower triangular matrix with ones in the main diagonal, only the elements below the diagonal are to be fed back. Consequently, feeding back these elements requires a lower amount of feedback than transmitting the whole channel matrix $\mathbf{H}$ from the receiver to the transmitter, and that the reason why we consider that the feedback information is the matrix $\mathbf{B}$ rather than the matrix $\mathbf{H}$.

In the decoupled case (see Figure 5.2), two precoding matrices are necessary at the transmitter side: $\mathbf{B}$ and $\mathbf{F}$, which cannot be calculated at the receivers side because each receiver only knows one of the rows of the full matrix $\mathbf{H}$. Consequently, we assume that the transmitter is informed, *e.g.*, through a feedback link or by direct estimation, with an erroneous channel matrix $\widehat{\mathbf{H}}$, and that, based upon that estimate, the transmitter calculates $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$ following

(5.13) and (5.7).

Since we are interested in high data rate transmission, the following analysis is done assuming that the SNR is high (for practical purposes this implies that it has to be greater than 10 dB) and that the error in the precoding matrices $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$ is kept small.

5.5.1   Single-user scenario (coupled case)

In the coupled case, we consider that the receiver feeds back the elements below the main diagonal of the matrix $\mathbf{B}$ to the transmitter. If some kind of error is considered in the feedback link, the transmitter will be informed with an erroneous precoding matrix, $\widehat{\mathbf{B}} = \mathbf{B} + \boldsymbol{\Delta}$ with $\boldsymbol{\Delta}$ being a strictly lower triangular matrix. In this case, using the compact notation as in (5.4), the transmitted signal, $\mathbf{x}_{c,\widehat{\mathbf{B}}}$, becomes

$$\mathbf{x}_{c,\widehat{\mathbf{B}}} = \widehat{\mathbf{B}}^{-1}(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{B}}}) = (\mathbf{B} + \boldsymbol{\Delta})^{-1}(\mathbf{s} + \mathbf{a}_{\widehat{\mathbf{B}}}). \tag{5.36}$$

Notice that $\mathbf{a}_{\widehat{\mathbf{B}}}$ in (5.36) must now be chosen so that $\mathbf{x}_{c,\widehat{\mathbf{B}}}$ lies inside the modulo region defined by vector $\mathbf{t}$. In addition, it is important to state that, in general, $\mathbf{a}_{\widehat{\mathbf{B}}} \neq \mathbf{a}$, where $\mathbf{a}$ is defined as in (5.4). However, if the variance of the elements of the error matrix $\boldsymbol{\Delta}$ is kept small it can be assumed that

$$(\mathbf{As1}): \quad \mathbf{a} = \mathbf{a}_{\widehat{\mathbf{B}}}. \tag{5.37}$$

See §5.6 for a validity evaluation. Under $(\mathbf{As1})$, substituting (5.36) in (5.8) and using the matrix inversion lemma [Hor85], $(\mathbf{B}+\boldsymbol{\Delta})^{-1} = \mathbf{B}^{-1} - \mathbf{B}^{-1}(\boldsymbol{\Delta}\mathbf{B}^{-1}+\mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}\mathbf{B}^{-1}$, the received signal can be expressed as $\mathbf{r}_{\widehat{\mathbf{B}}} = M_{\mathbf{t}}[\mathbf{s} + (\boldsymbol{\Delta}\mathbf{B}^{-1}+\mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}\mathbf{B}^{-1}(\mathbf{s}+\mathbf{a}) + \mathbf{G}\widetilde{\mathbf{n}}]$. Notice that $\mathbf{B}^{-1}(\mathbf{s}+\mathbf{a}) = \mathbf{x}_c$, *i.e.*, we can reduce $\mathbf{B}^{-1}(\mathbf{s}+\mathbf{a})$ to $\mathbf{x}_c$, which would be the transmitted signal in the coupled case if no errors were present in $\widehat{\mathbf{B}}$ matrix, and whose components are bounded in $[-t_k, t_k) \times [-j\,t_k, j\,t_k)$. Defining $\mathbf{L} = (\boldsymbol{\Delta}\mathbf{B}^{-1} + \mathbf{I}_{n_T})^{-1}\boldsymbol{\Delta}$ the received signal reads

$$\mathbf{r}_{\widehat{\mathbf{B}}} = M_{\mathbf{t}}[\mathbf{s} + \mathbf{L}\mathbf{x}_c + \mathbf{G}\widetilde{\mathbf{n}}], \tag{5.38}$$

where $\mathbf{L}$ is a strictly lower triangular matrix.[3] Attention must be paid to the fact that if the elements of $\boldsymbol{\Delta}$ are sufficiently low, the first order approximation $\mathbf{L} = \boldsymbol{\Delta} + o(\boldsymbol{\Delta})$ becomes valid.

Once we have obtained an expression for the input-output relation in the presence of feedback errors we can proceed to maximize the mutual information between $\mathbf{s}$ and $\mathbf{r}_{\widehat{\mathbf{B}}}$. As it was described in §5.3.2, the power transmitted through k-th antenna, P_k, is controlled by t_k by the relation $P_k = 2t_k^2/3$. In §5.4, it was found that, if no feedback errors are present, the STHP gets very close to achieve its capacity when each element of data vector $\mathbf{s}$ is uniformly distributed in the

---

[3]Notice that the inverse of a triangular matrix is also triangular. Additionally, the product of a lower triangular matrix and a strictly lower triangular matrix is a strictly lower triangular matrix.

interval defined by vector $\mathbf{t}$ and the transmission power, P_T, is equally distributed among the set of active antennas, which is defined as the set of active antennas that would correspond to the perfect CSI case (see §5.4.3). Noteworthy, the optimal set in the imperfect CSI case need not be the same as in the perfect CSI case, but, in practice, it can be checked that the two sets coincide (at least with the small errors assumption). From now on, the set of active antennas are numbered from 1 to N, i.e., $P_1, \ldots, P_N > 0$, $P_{N+1}, \ldots, P_{n_S} = 0$. In the following, the power distribution and, consequently, the moduli $\{t_k\}$ will be adapted to maximize the worst-case achievable rates when feedback errors are present, giving thus rise to a robust design.

The first order approximation in $\boldsymbol{\Delta}$ of the received signal vector is

$$(\mathbf{As2}): \quad \mathbf{r}_{\widehat{\mathbf{B}}} \simeq M_{\mathbf{t}}[\mathbf{s} + \boldsymbol{\Delta}\mathbf{x} + \mathbf{G}\tilde{\mathbf{n}}], \tag{5.39}$$

where $\mathbf{x}_c$ has been substituted by $\mathbf{x}$ for the sake of notation. If $\boldsymbol{\Delta}\mathbf{x}$ is treated as an unknown interference the mutual information between k-th element of data vector, s_k, and the corresponding element of received signal, $r_k = [\mathbf{r}_{\widehat{\mathbf{B}}}]_k$, is

$$I(s_k; r_k) = \log(6P_k) - h\left(M_{t_k}\left[g_{kk}\tilde{n}_k + \sum_{j<k}\delta_{kj}x_j\right]\right), \tag{5.40}$$

where $\delta_{kj} = [\boldsymbol{\Delta}]_{kj}$ and $h(\cdot)$ denotes differential entropy in an analogous way as in (5.17). Notice that, as it was stated in §5.4.2, since x_k is the output of the precoder, it is uniformly distributed in $\mathcal{D}_k$ and thus its variance equals P_k.

Let us define the random variable $z_k = g_{kk}\tilde{n}_k + \sum_{j<k}\delta_{kj}x_j$ with power $\mathbb{E}\{|z_k|^2\} = g_{kk}^2\sigma^2 + \sum_{j<k}|\delta_{kj}|^2P_j$. It can be easily verified [Pap91] that z_k can be approximately modeled as a complex Gaussian random variable as long as

$$(\mathbf{As3}): \quad \max_j |\delta_{kj}|^2 P_j \lesssim g_{kk}^2\sigma^2/3. \tag{5.41}$$

Under $(\mathbf{As3})$, we found in §5.4.3 that the mutual information expression (5.40) is very well approximated by

$$(\mathbf{As4}): \quad I(s_k; r_k) \simeq \log^+\left(\frac{6P_k}{\pi e(g_{kk}^2\sigma^2 + \sum_{j<k}|\delta_{kj}|^2P_j)}\right), \tag{5.42}$$

where $\log^+(x) = \max(\log(x), 0)$. The achievable rates for the STHP structure will then be the sum of the mutual information of each active stream, $C = \sum_{k=1}^N I(s_k; r_k)$.

In order to describe the noise worst-case scenario we consider that the squared moduli of the components of the error matrix $\boldsymbol{\Delta}$ are bounded, i.e., $|\delta_{kj}|^2 < \alpha_{kj}$. In addition, for the sake of simplicity we assume that $\alpha_{kj} = \alpha, \forall k, j$. Notice that, since we are interested in the worst-case, in case that the values of α_{kj} were different for some k, j, the more pessimistic case could be considered by letting $\alpha = \max_{k,j} \alpha_{kj}$.

From all the considerations above, the power distribution that maximizes the worst-case achievable rates when the CSI is imperfect is the solution to the following maximin problem:

$$
\begin{aligned}
C_{\text{THP}}^{rob,c} = \max_{\{P_k\}} \min_{\{\delta_{ij}\}} \quad & \sum_{k=1}^{N} I(s_k; r_k), \\
\text{s.t.} \quad & \sum_{k=1}^{N} P_k = P_T, \\
& |\delta_{ij}|^2 \leq \alpha, \ \forall i, j.
\end{aligned}
\tag{5.43}
$$

The solution to the minimization part is trivial, since each term $I(s_k; r_k)$ is a decreasing function of $|\delta_{ij}|^2$ and each $|\delta_{ij}|^2$ is upper bounded independently of the others. Thus, the minimum will be attained when $|\delta_{ij}|^2 = \alpha$, $\forall i, j$. The resulting maximization problem is a standard constrained optimization problem, and can be solved with the use of the Lagrange method. The Lagrange equation is, up to a constant,

$$
\mathcal{L}(\{P_k\}; \lambda) = \sum_{k=1}^{N} \log \left(\frac{P_k}{(g_{kk}^2 \sigma^2 + \alpha \sum_{k' < k} P_{k'})} \right) + \lambda \left(\sum_{k=1}^{N} P_k - P_T \right).
\tag{5.44}
$$

The optimal power allocation should satisfy

$$
\frac{\partial \mathcal{L}(\{P_k\}; \lambda)}{\partial P_k} = 0 \text{ for } k = 1, \ldots, N \quad \text{and} \quad \frac{\partial \mathcal{L}(\{P_k\}; \lambda)}{\partial \lambda} = 0,
\tag{5.45}
$$

with the additional constraint that $\{P_k\}$ is non-negative $\forall k$. With some basic manipulations from (5.44) and (5.45) a recursive relation of the type $P_{N-k} = f(P_{N-k+1}, \ldots, P_N)$ for $k = 1, \ldots, N-1$ between the assigned power to each antenna can be found as

$$
P_{N-k} = P_{N-k+1} \frac{w_{N-k+1} + \alpha \left(P_T - \sum_{k' > N-k} P_{k'} \right)}{w_{N-k+1} + \alpha \left(P_T - \sum_{k' > N-k+1} P_{k'} \right)}, \quad \text{with } w_k = g_{kk}^2 \sigma^2
\tag{5.46}
$$

for $k = 1, \ldots, N - 1$. Notice that the second factor in last equation is always lower than unity, which implies necessarily that $P_1 \leq P_2 \leq \ldots \leq P_{N-1} \leq P_N$, which is a reasonable solution since power interference is progressive in the sense that P_k interferes with streams from $k + 1$ to N, but not with streams from 1 to $k - 1$, see (5.42). The set of equations in (5.46) together with $\sum_{k=1}^{N} P_k = P_T$ can be solved numerically obtaining the robust power allocation.

5.5.2   Multi-user scenario (decoupled receivers case)

In the decoupled case two precoding matrices are necessary at the transmitter side: $\mathbf{B}$ and $\mathbf{F}$. As commented above, we assume that the transmitter is informed, *e.g.*, through a feedback link or by direct estimation, with an erroneous channel matrix $\widehat{\mathbf{H}} = \mathbf{H} + \boldsymbol{\Delta}$, and that, based upon that estimate, the transmitter calculates $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$ following (5.13) and (5.7). Furthermore, we consider that the entries of $\boldsymbol{\Delta}$ are i.i.d. circularly symmetric complex Gaussian random variables.

The error present in the estimate $\widehat{\mathbf{H}}$ propagates to the estimates $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$ as $\widehat{\mathbf{B}} = \mathbf{B} + \mathbf{\Delta_B}$, and $\widehat{\mathbf{F}} = \mathbf{F} + \mathbf{\Delta_F}$. This error propagation has been recently studied in [Bah05] to characterize the BER in STHP systems. Fortunately, the explicit expressions for $\mathbf{\Delta_B}$ and $\mathbf{\Delta_F}$ are not needed here, because the authors of [Bah05] also obtained a very simple expression, which is reproduced in (5.47), for the estimate of the data symbols vector in the decoupled case when $\widehat{\mathbf{B}}$ and $\widehat{\mathbf{F}}$ are utilized in the transmitter design.

$$\mathbf{r} = M_{\mathbf{t}} \left[\mathbf{s} + \boldsymbol{\xi}(\mathbf{s}) + \mathbf{Gn} \right]. \tag{5.47}$$

In last equation, $\boldsymbol{\xi}(\mathbf{s})$ is a random vector that represents the effects of having imperfect versions of $\mathbf{B}$ and $\mathbf{F}$ at the transmitter side, and is independent of $\mathbf{n}$ but depends on the transmitted symbols sequence. In [Bah05], the authors also found that $\boldsymbol{\xi}$ is distributed as a zero mean circularly symmetric complex Gaussian random vector with $\mathbb{E}\,\boldsymbol{\xi}\boldsymbol{\xi}^H = \mathbf{\Xi}$.[4] An explicit expression for the calculation of $\mathbf{\Xi}$ is given in [Bah05] and is not reproduced here for the sake of space. We simply need to know that $\mathbf{\Xi}$ can be calculated for the worst possible sequence of transmitted symbols as a function of $\mathbf{H}$ and the power of the entries of $\mathbf{\Delta}$. Notice that, in this case, the structure of the interference is not progressive, as it was found to be in the coupled case (5.38). This is due to the fact that, in the decoupled case, before the output of the precoder $\mathbf{p}_{\widehat{\mathbf{B}}}$ is being transmitted it is multiplied by the matrix $\widehat{\mathbf{F}}$ which is unitary and consequently distributes uniformly the interference that is present in $\mathbf{p}_{\widehat{\mathbf{B}}}$ among all the components of the transmitted signal $\mathbf{x}_d = \widehat{\mathbf{F}}\mathbf{p}_{\widehat{\mathbf{B}}}$.

Moreover, in general, the matrix $\mathbf{\Xi}$ is not diagonal, which implies that the entries of the interference term $\boldsymbol{\xi}$ in (5.47), can be correlated. However, since the receivers are not allowed to cooperate, no advantage can be taken from the correlation of the interference term in the decoding process, and thus we only need to be concerned about the diagonal entries of $\mathbf{\Xi}$, $[\mathbf{\Xi}]_{kk}$, which represent the power of the interference term in the received signal by each user.

From what has been said above, the entries of the interference term $\boldsymbol{\xi}$ can be accurately modeled as an additional source of Gaussian noise independent of $\mathbf{Gn}$, and whose power is given by the diagonal elements of the matrix $\mathbf{\Xi}$. Consequently, since the sum of two independent Gaussian variables is another Gaussian variable whose power is given by the sum of the individual powers, the two noise terms can be easily grouped into a single noise term, *i.e.*, $\xi_k + g_{kk}n_k \sim \mathcal{CN}(0, g_{kk}^2\sigma^2 + [\mathbf{\Xi}]_{kk})$. Now, the received vector in (5.47) can thus be equivalently expressed as

$$\mathbf{r} = M_{\mathbf{t}} \left[\mathbf{s} + \mathbf{D}\boldsymbol{\nu} \right], \ \text{or} \ r_k = M_{t_k} \left[s_k + d_{kk}\nu_k \right], \ k = 1, \ldots, n_S, \tag{5.48}$$

where $\mathbf{D}$ is a diagonal matrix, with $[\mathbf{D}]_{kk}^2 = d_{kk}^2 = g_{kk}^2\sigma^2 + [\mathbf{\Xi}]_{kk}$, and where $\boldsymbol{\nu} \sim \mathcal{CN}(0, \mathbf{I}_{n_S})$.

---

[4]More precisely, in [Bah05] the authors state that the vector $\boldsymbol{\xi}$ in (5.47) is *asymptotically* distributed as a circularly symmetric complex Gaussian random vector. In this case, asymptotically refers to the fact that in the limit where the entries of $\mathbf{\Delta}$ are Gaussian distributed then so are the entries in the vector $\boldsymbol{\xi}$. Since we are assuming that the entries of $\mathbf{\Delta}$ are Gaussian distributed we can drop the adjective asymptotically without losing precision.

Now we can express the robust achievable rates for the decoupled case as the solution to

$$
\begin{aligned}
C_{\text{THP}}^{rob,d} &= \max_{\{t_k\}} \quad \sum_{k=1}^{n_S} R\left(t_k, \sqrt{g_{kk}^2\sigma^2 + [\boldsymbol{\Xi}]_{kk}}\right), \\
&\text{s.t.} \quad \sum_{k=1}^{n_S} \frac{2t_k^2}{3} = P_T.
\end{aligned}
\tag{5.49}
$$

However, one readily sees that the problem in (5.49) is the same that was solved in §5.4.3, but with a new set of noise powers given according to $d_{kk}^2 = g_{kk}^2\sigma^2 + [\boldsymbol{\Xi}]_{kk}$. Consequently, the problem can be solved quasi-optimally by selecting an active set of streams (*users*) and performing uniform power allocation among them as was explained in §5.4. Noteworthy, in the imperfect CSI case, the robust set of active users does not necessarily be the same as in the perfect CSI case. This is due to the fact that, when the CSI is imperfect, the ordering of the new set of noise powers d_{kk}^2 may be different than the set $g_{kk}^2\sigma^2$. Consequently, the algorithm may deactivate a user whose channel was good when no feedback errors were present, but whose channel quality has decreased significantly due to the presence of the term $[\boldsymbol{\Xi}]_{kk}$.

5.6   Simulation results

To validate and give graphical representations of our results, numerical simulations have been conducted. We have considered a flat-fading 3×3 MIMO channel matrix, whose entries have been assumed to be zero mean unit variance i.i.d. circularly symmetric complex Gaussian random variables. For the case where the presence of errors in the feedback is considered we have fixed a mean SNR of $P_T/\sigma^2 = 15$ dB.

For the perfect CSI case, in Figure 5.4 the maximum achievable rates for STHP and the system capacity are plotted versus the mean SNR. At the high SNR regime, the rate loss of STHP with respect to the system capacity is approximately of 1 nat, as expected from (5.31). The achievable rates given by the quasi-optimal solution in §5.4.3 are not plotted because they overlap with the ones obtained with numerical optimization of the problem in (5.23), which is a graphical indication of the quasi-optimality of our uniform power allocation with stream selection solution.

Concerning the case of imperfect CSI, on the one hand, for the single-user case, as the robust capacity analysis has been done using numerous approximations (**As1**), (**As2**), (**As3**), and (**As4**), before presenting the simulations results, the validity of the approximations is shown in Figure 5.5 by plotting the fraction of realizations with respect to the noise in which the approximations are valid. Precisely, (**As4**) is considered to hold true when the relative difference between the approximation (5.42) and the actual value (5.40), calculated numerically, is lower than 10^{-3}. It can be seen that, for the particular values of the simulation parameters taken in
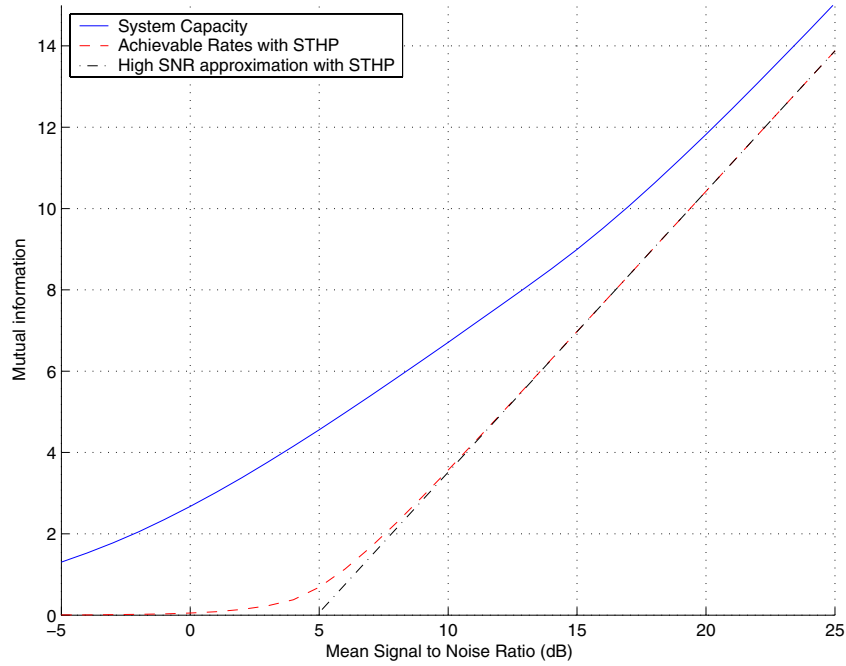
Figure 5.4: System capacity and achievable rates for STHP, as a function of mean SNR. The high SNR approximation has also been plotted.

this section, the capacity analysis is valid for values of α up to -21 dB. For values of α higher than -21 dB our analysis is not valid. However, from the slope of the robust achievable rates curves at $\alpha = -21$ dB in Figure 5.6, the rates that can be achieved for $\alpha > -21$ dB seem to decrease fast, so even with a robust technique it would be difficult to cope with the problems associated with having an imperfect feedback link. Notice that the validity of (**As 2**) has not been plotted as the differences between **L** and $\mathbf{\Delta_B}$ have always been found negligible. In Figure 5.6, we have plotted the maximum achievable rates for the robust power allocation strategy and also for the non-robust uniform power allocation with antenna selection scheme described in §5.4.3 for different realizations of the MIMO channel. In Figure 5.7 we have plotted the fraction of the total power that is transmitted through each one of the antennas, which is related to the set of moduli used in STHP, for a particular realization of the channel. It can be seen that, as α gets close to -21 dB, the robust power allocation differs substantially from the uniform power allocation.

On the other hand, in the multi-user case, similar simulations have been conducted but in this case the parameter α represents the noise power of each component of the estimation error in the channel matrix $\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{\Delta_H}$, *i.e.*, $\mathbb{E}|[\mathbf{\Delta_H}]_{ij}|^2 = \alpha$. For each channel realization the worst-case matrix $\mathbf{\Xi}$ has been computed following [Bah05] and then a set of active users has been selected according to the set of noise powers given by $d_{kk}^2 = g_{kk}^2\sigma^2 + [\mathbf{\Xi}]_{kk}$. The resulting maximum achievable sum rate has been plotted in Figure 5.8 for various channel realizations.

In addition, since in the multi-user case we are not so strongly conditioned on the validity of the approximations done in the single-user case, we can extend the domain of α. As the value of α increases the maximum achievable sum rate decreases very rapidly, which implies that, as the estimation noise increases, even the robust technique is not able to cope with the presence of errors in the feedback link. Note that this conclusion, also valid for the single-user case, is not surprising since the robust design presented in this chapter has been derived assuming small errors and, therefore, the degradation can be quite high when this assumption does not hold.

5.7   Chapter summary and conclusion

In this chapter we have analyzed some issues concerning the achievable rates of the STHP scheme. Initially, we have added new degrees of freedom in the design of the STHP by allowing different modulo operations at the output of the precoder. Next, the loss in mutual information with respect to the system capacity has been calculated for the high SNR regime, finding that there is a gap of approximately $0.353n_S$ nats. Finally, we have found two robust power allocation strategies, for the coupled and decoupled cases, that maximize the mutual information with imperfect CSI. We have observed that the proposed robust techniques are able to cope with the imperfections in the CSI by minimizing the loss in terms of rate when the CSI is not perfect.

Finally, we wish to highlight that the proposed algorithm could work in a realistic deployment since the presence of errors in the feedback link has been explicitly taken into account in the design process.
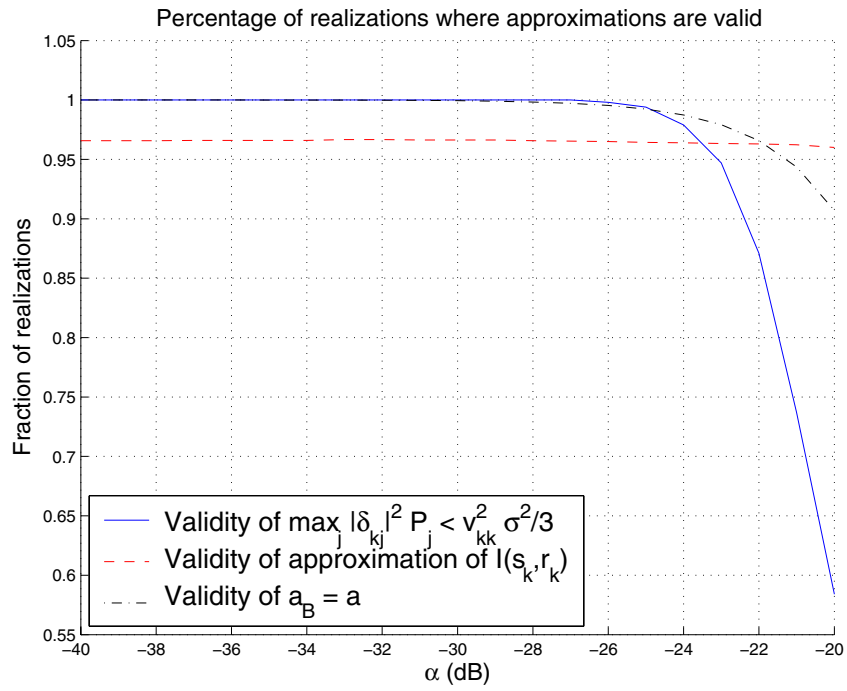
Figure 5.5: Validation of the approximations done in the analysis.
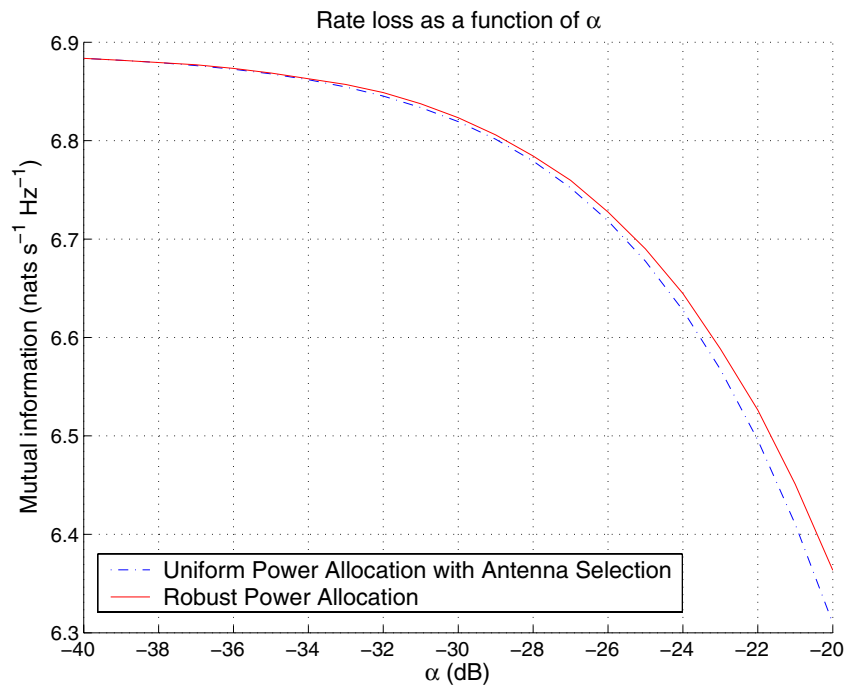
Figure 5.6: Capacity for different power allocation strategies as a function of α for the single-user scenario.
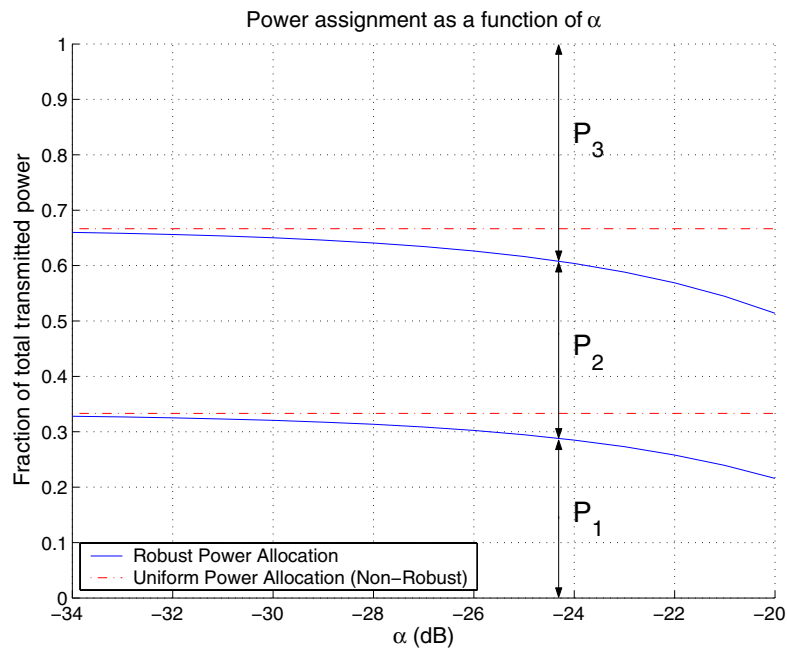
Figure 5.7: Graphical representation of the robust power allocation with respect to the uniform power allocation. In the figure we plotted the fraction of the total transmitted power that is assigned to each of the three antennas as a function of α for the single-user scenario.
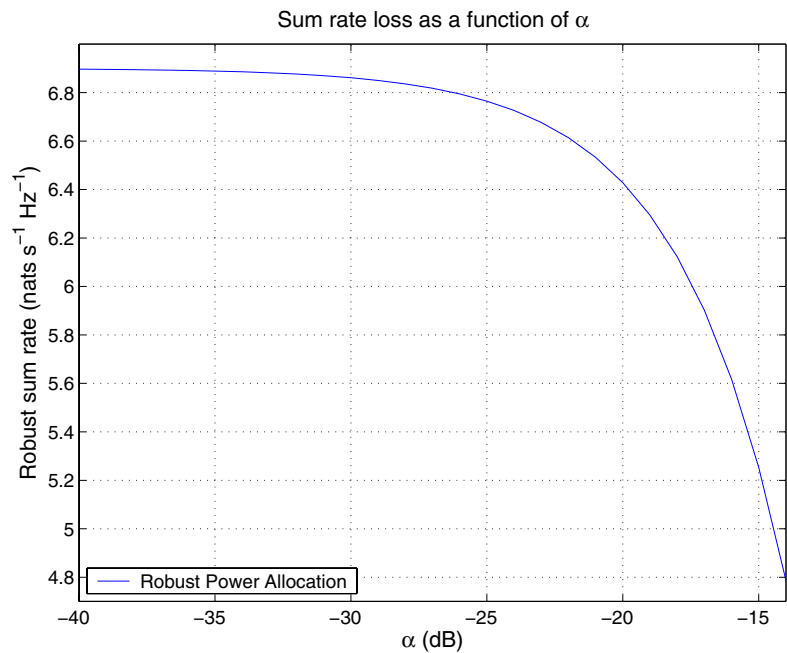
Figure 5.8: Sum rate for the robust power allocation technique as a function of α in the multi-user scenario.

5.A Properties of the differential entropy of a clipped Gaussian random variable

For the sake of simplicity, the properties presented in this section for the differential entropy of a clipped Gaussian random variable, $h(M_t[z])$, with $z \in \mathbb{C}$, will be based on extensions from the differential entropy of a real random variable $h(M_t[x])$, with $x \in \mathbb{R}$. First, the pdf of $M_t[x]$ is characterized, and then some properties of its differential entropy are presented.

5.A.1 Probability density function of a clipped Gaussian random variable

Let x be a real zero-mean Gaussian random variable with variance σ^2. Let us also define $y = M_t[x]$ where now $M_t[x]$ is a modulo-$2t$ reduction of x into the interval $[-t, t)$. The pdf of y, is

$$f_Y(y, t, \sigma) = \begin{cases} \sum_{k=-\infty}^{\infty} f_X(y - 2tk, \sigma) & y \in [-t, t) \\ 0 & y \notin [-t, t) \end{cases}, \tag{5.50}$$

where the dependence of f_Y on variables t, σ has been explicitly written, and where

$$f_X(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{5.51}$$

It is important to notice that, if t tends to infinity with σ fixed, f_Y tends to f_X as the only relevant term in the summation in (5.50) is for $k = 0$. In addition, if σ tends to infinity with t fixed, f_Y tends to the pdf of a uniform random variable in $[-t, t)$, *i.e.*, $f_Y(y) = (2t)^{-1}$ for $y \in [-t, t)$.

5.A.2 Differential entropy

The differential entropy of y can be written as [Cov91]

$$h(y, t, \sigma) = -\int_{-t}^{t} f_Y(y, t, \sigma) \log f_Y(y, t, \sigma) \mathrm{d}y. \tag{5.52}$$

In the previous section, it was stated that, in some limiting values, f_Y tends to the pdf of a uniform ($t \to \infty$, σ fixed) or Gaussian (t fixed, $\sigma \to \infty$) random variables. Consequently, the limiting values of the differential entropy in these cases will correspond to the well known expressions for the differential entropy of uniform and Gaussian random variables:

$$\lim_{t\to\infty} h(y, t, \sigma) = \frac{1}{2} \log 2\pi e\sigma^2, \tag{5.53}$$

$$\lim_{\sigma\to\infty} h(y, t, \sigma) = \log 2t, \tag{5.54}$$

which for the complex case these tend to $\log \pi e\sigma^2$ and $2 \log 2t$, respectively.

5.B Probability density function of $p = M_t[s + z]$

The purpose of this appendix is to show that the elements p_k in equation (5.2), are independent of the others $p_{k'}$, with $k' \neq k$, i.e., $f_{P_k P_{k'}}(p_k, p_{k'}) = f_{P_k}(p_k) f_{P_{k'}}(p_{k'})$. Just as in the previous section, we will present the analysis in the real case, bearing that the extension to the complex case is straightforward.

First of all, we recall the expression in (5.2),

$$p_k = M_{t_k} \left[s_k - \sum_{k'=1}^{k-1} b_{kk'} p_{k'} \right], \quad \text{for } k = 1, \dots, n_S. \tag{5.55}$$

and for the sake of notation, we define $z = -\sum_{k'=1}^{k-1} b_{kk'} p_{k'}$, and we omit the subscript index k. With this new notation, the expression in (5.55) simplifies to $p = M_t[s + z]$. The joint pdf of p and z is

$$f_{PZ}(p, z) = f_{P|Z}(p \mid z) f_Z(z). \tag{5.56}$$

In order to prove independence between p and z, it is sufficient to show that $f_{P|Z}(p \mid z) = f_P(p)$. Defining $q = s + z$, we have $f_{Q|Z}(q \mid z) = f_S(q - z)$, from which it can be stated that

$$f_{Q|Z}(q \mid z) = \begin{cases} 1/(2t) & q \in [z - t, z + t) \\ 0 & q \notin [z - t, z + t) \end{cases}, \tag{5.57}$$

and consequently

$$f_{P|Z}(p \mid z) = \begin{cases} \sum_{k=-\infty}^{\infty} f_{Q|Z}(p - 2tk \mid z) & p \in [-t, t) \\ 0 & p \notin [-t, t) \end{cases}. \tag{5.58}$$

Notice that the function $f_{Q|Z}(p - 2tk \mid z)$ is flat in the interval $[z + t(2k - 1), z + t(2k + 1))$ and zero elsewhere. This implies that $\sum_{k=-\infty}^{\infty} f_{Q|Z}(p - 2tk \mid z)$ is a non-overlapping superposition of flat intervals separated a distance equal to the length of the interval, $2t$. Consequently, $\sum_{k=-\infty}^{\infty} f_{Q|Z}(p - 2tk \mid z) = 1/(2t)$ independently of z, from which it can be stated that

$$f_{P|Z}(p \mid z) = f_P(p) = \begin{cases} 1/(2t) & p \in [-t, t) \\ 0 & p \notin [-t, t) \end{cases}, \tag{5.59}$$

which implies that p is uniformly distributed in $[-t, t)$ and it is independent of z. Since z is defined as a linear combination of $\{p_{k'}\}_{k'=1}^{k-1}$, and the result we have shown is independent of the distribution of z it follows that, for any two pair of values k and k', with $k' < k$, $f_{P_k P_{k'}}(p_k, p_{k'}) = f_{P_k}(p_k) f_{P_{k'}}(p_{k'})$. Notice that if $k' > k$ the independence is trivial since in (5.2) and (5.55) it can be seen that p_k does not depend on $p_{k'}$ with $k' > k$.

Extending the above results to the complex case, it implies that, when the entries, s_k, of the precoder input symbols vector, $\mathbf{s}$, are independent and uniformly distributed in the intervals

$\{[-t_k, t_k) \times [-\mathrm{i}\, t_k, \mathrm{i}\, t_k)\}_k$, the entries, p_k , of the precoder output vector $\mathbf{p}$ are also independent and uniformly distributed in the intervals $\{[-t_k, t_k) \times [-\mathrm{i}\, t_k, \mathrm{i}\, t_k)\}_k$. This implies that

$$\mathbb{E} p_k p_{k'}^* = \delta_{kk'} \frac{2t_k^2}{3},$$

(5.60)

where $\delta_{kk'}$ represents the kronecker delta.

Chapter 6

Multi-user communication through MIMO channels with imperfect CSI

6.1 Introduction

In connection with last chapter, where the multi-user scenario has been introduced, in this chapter we stay in this multi-user setup and present the robust design of a linear transmitter that is aimed at guaranteeing a minimum QoS for every user taking into account the fact that the CSI that both the transmitter and the receivers have available is imperfect.

Before presenting this robust linear transmitter design some capacity results are briefly discussed in the following section.

6.2 Capacity results

As it has been commented in the previous chapters, multiple-antenna communications systems have generated a great deal of interest among the research community since they have the potential of considerably increasing the capacity. In chapter 3 we saw that, if perfect CSI is made available at the transmitter and the receiver, then they can jointly create a number of parallel channels and thus, loosely speaking, increase the capacity of the channel by this same factor. It was later shown that the same capacity scaling is true if the channel is not known at the transmitter [Tel99, Fos98] and even if it is not known at the receiver [Zhe02, Has02, Lap98, Mos04, Lap03] (provided that the coherence interval of the channel is not too short).

Since these results only hold for single-user communication systems, there has been recent interest in the role of MIMO systems in a multi-user setup, and especially in broadcast channels. There have been two main approaches of conducting research in these multi-user scenarios. On one hand, there has been a line of work studying transmission architectures [Kob05b, Kob05a]

and also cross-layer scheduling algorithms in MIMO broadcast channels [Hoc04, Vic06, Rea06]. One of the main results of [Hoc04] being that, due to channel hardening in MIMO systems, many of the multi-user gains disappear. On the other hand, a significant research effort has been devoted to the study the sum-rate capacity and also the capacity region of MIMO BC [Cai03, Vis03b, Vis03a, Yu04]. The main conclusion of these excellent works is that the sum-rate capacity is achieved by DPC [Cos83]. Recently, it has been shown that DPC in fact achieves the capacity region of the Gaussian MIMO broadcast channel [Wei04, Wei06].

While the above results suggest that in the asymptotic regime capacity increases linearly in the number of transmit antennas, they all rely on the assumption that the channel is known perfectly at the transmitter. Moreover, the DPC scheme, especially in the multi-user context, is extremely computationally intensive (although suboptimal schemes such as channel inversion [Yoo05] or Tomlinson-Harashima precoding [Yu01, Zam02, Pee05a, Pee05b] give relatively close performance to the optimal schemes). It is then reasonable to speculate whether, as in the single-user case, it is possible to get the same gains in the multi-user case without having channel knowledge at the transmitter. Unfortunately, if no channel knowledge is available at the transmitter, capacity scales only logarithmically in the number of transmit antennas [Sha05b]. In fact, in this case, increasing the number of transmit antennas yields no gains since the same performance can be obtained with a single transmit antenna operating at higher power.

In many applications, especially if the users are mobile and are moving rapidly, it is not reasonable to assume that all the channel coefficients from the transmitter to every user can be made available at both communication ends. Since perfect CSI may be impractical, yet no CSI is useless, it is very important to study the capacity of broadcasting systems where only partial or imperfect CSI at the transmitter is available.

However, this subject has proved extremely difficult and almost no works have been found in the existing literature that deal with the topic of broadcast capacity with imperfect or incomplete CSI. In [Sha05b] the achievable sum-rate was studied for a particular case or orthogonal beamforming with SNR feedback, and in [Lap06] Lapidoth, Shamai, and Wigger studied the effects of imperfect CSI at the transmitter in the capacity of a simplified MIMO broadcast channel with two antennas at the transmitter and two users equipped with single antenna terminals.

Since the general topic of broadcast capacity with imperfect CSI appears to be untractable, we focused more specifically on the design of practical transceiver architectures, where the imperfections in the CSI are taken into account explicitly in the design process. Unfortunately, due to the lack of general results in broadcast capacity with imperfect CSI it is, for the moment, impossible to design the transmitter according to information-theoretic criteria. Consequently, in the following sections we present the design of a multi-antenna transmitter for the broadcast channel which is aimed at minimizing the total transmitted power while guaranteeing a predefined QoS per user.

6.3   Practical transmission scheme

Although the benefits for single-user communications have been studied extensively in the previous chapters, one of the main potentials of multi-antenna communications is that they can afford multi-user communications, where the different signals can be separated by spatial processing techniques. As opposed to single-user communications, in the design of multi-user systems, several QoS measures have to be considered simultaneously, each one corresponding to each user. This leads to an inherent problem in the design of such systems, which is how to handle with these different quality measures.

A possible solution, inspired by single-user designs, aims at optimizing the mean value of these measures [Ser04]. However, it must be taken into account that the optimization of the mean does not guarantee a minimum acceptable quality for all the users. Consequently, a more suitable approach, is to guarantee a minimum QoS independently for the data stream corresponding to each user, while optimizing a global network parameter such as the total transmitted power.

Similar to the single-user case, the performance in multi-user communications also depends on the available CSI. Initially, most researchers concentrated their efforts on the design of multi-user transmission architectures assuming that both the transmitter and the user receivers have perfect knowledge of the CSI, giving rise to the so-called solution to the downlink beamforming problem, [Ben01, Sch04]. Very recently, a unified framework with a very powerful and general model to deal with the problem of power allocation design in a multi-user and multi-antenna downlink scenario with perfect CSI has been presented in [Boc05].

Note that, as it has been already commented in this dissertation, in a realistic implementation of the system, the assumption of the availability of a complete and perfect CSI is too optimistic, specially for wireless systems where the fluctuations of the channel can be fast, and also due to the presence of estimation noise or quantization effects in the CSI. In the case of imperfect CSI, the simplest approach consists in utilizing the available CSI as if it was perfect, giving rise to naive (non-robust) designs. It has been shown that these designs are extremely sensitive to the errors in the CSI [Cho02b, Cho02c, Zho04], which translates into a decrease of the system performance, or, equivalently, into an inefficient use of the network resources, such as the power consumption.

Thus, a more desired approach in this case is to consider a robust design where the presence of the errors in the CSI is explicitly taken into consideration in the design. Different designs are possible depending on the model assumed for the errors. On one hand, in the Bayesian philosophy the errors are modeled from a statistical point of view, and, on the other hand, the maximin approach does not need a statistical description of the error, because it is assumed that the error belongs to a predefined uncertainty region, whose shape and size are linked to the

physical phenomenon producing the error in the CSI.

In the following, we consider the downlink of a multi-user communications system with several single antenna receivers and a multiantenna BS. The transmitter is composed of two blocks: a power allocation among the symbols for different users, and a linear transformation. The robustness of our system is achieved by a maximin design of the power allocation under two considerations. On one hand, the objective is to minimize the total transmitted power, and, on the other hand, we wish to guarantee a certain minimum QoS per user for any possible error of the CSI inside the uncertainty region. The design of the power allocation fulfilling these two considerations is formulated as a convex optimization problem, which is next solved for several uncertainty regions, modeling the most practical cases of errors in the CSI (estimation Gaussian noise and/or quantization effects). The main advantage of formulating our optimization problem within the convex optimization framework is that numerical solutions can be computed very efficiently, and, even in some cases, a quasi-closed form solution can be found.

Although a significant research effort has been devoted to the design robust receivers (see [Vor03] and references therein), to the best of our knowledge, the existing literature dealing with linear transmitter design in multi-user systems with imperfect CSI is more scarce [Ben00, Ben01, Big04], and some references therein by the same authors. In these works it has been considered that an imperfect estimate of the channel *covariance* matrix is made available at the transmitter, whereas we consider the alternative case where an imperfect estimate of the channel matrix itself is available at both communication ends.

6.4   System model

We consider the same downlink communication system model as in the previous chapter for the multi-user case. Recall that in this model the BS utilizes n_T antennas to simultaneously transmit information symbols to n_U users, whose terminals are equipped with a single antenna. The baseband model for the samples of the received signal vector $\mathbf{y} \in \mathbb{C}^{n_U}$ is then

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{6.1}$$

where $\mathbf{x} \in \mathbb{C}^{n_T}$ represents the transmitted signal by the BS through all the antennas, $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{n_U}]^H \in \mathbb{C}^{n_U \times n_T}$ is the channel matrix, and $\mathbf{n} \in \mathbb{C}^{n_U}$ represents the noise vector.

As opposed to the previous chapter, where we considered both the single- and multi-user cases within the same model, in this chapter only the multi-user case is dealt with and consequently it will be convenient to decouple the model for the received vector in (6.1) as

$$y_i = [\mathbf{y}]_i = \mathbf{h}_i^H \mathbf{x} + n_i, \quad i = 1, \ldots, n_U, \tag{6.2}$$

where $\mathbf{h}_i^H \in \mathbb{C}^{1 \times n_T}$ is the flat fading spatial channel response from the n_T transmission antennas to the i-th user, $\mathbf{x} \in \mathbb{C}^{n_T}$ represents the transmitted signal by the BS through all the antennas, and $n_i = [\mathbf{n}]_i$ is the noise contribution with $\mathbb{E}|n_i|^2 = \sigma^2, \forall i$.

The transmitted signal, $\mathbf{x}$, is designed as a linear function of the information symbols vector, $\mathbf{s} \in \mathbb{C}^{n_U}$, where s_i represents the symbol to be communicated to i-th user and where $\mathbb{E}\mathbf{s}\mathbf{s}^H = \mathbf{I}_{n_U}$ is assumed w.l.o.g. This linear combination is expressed as a product of two linear transformations as

$$\mathbf{x} = \mathbf{B}\mathbf{P}^{1/2}\mathbf{s}, \tag{6.3}$$

where $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_{n_U}]$ is the transmission matrix and the diagonal matrix $\mathbf{P}^{1/2}$, with elements $[\mathbf{P}^{1/2}]_{ii} = \sqrt{p_i}$, takes into account the power allocation among the information symbols.

The objective is now to design the transmitter according to the available information about the actual channel matrix at both communications ends, which is assumed to be imperfect. In order to optimize the global system performance, the presence of these imperfections has to be taken into account explicitly, leading to robust solutions that are less sensitive to these errors. As we have commented previously, there are different ways to incorporate robustness in the system design, such as the Bayesian and the maximin approaches, where the whole transmitter $\mathbf{B}\mathbf{P}^{1/2}$ is designed according to these criteria. In both cases, the formulated mathematical problems are generally much more complicated than the classical non-robust solutions. This too demanding complexity requires to make some assumptions and simplifications in the design, as seen in many works such as [Jön02, Zho02, Zho03, Sch04, Ben01] (indeed, these simplifications may be required not only to solve the mathematical problem itself, but also to obtain a solution that can be implemented in a realistic system with restrictions on the allowed computational load). Concretely, in our case, the design of the transmitter is simplified by dividing it into two parts taking an engineering and practical perspective. The transmission matrix $\mathbf{B}$ is allowed to depend only on the channel estimate and it is designed in a non-robust way according to a predefined performance criterion. On the other hand, the design of the power allocation $\mathbf{P}^{1/2}$ is much more general and is allowed to depend not only on the channel estimate, but also on the model of the imperfections in the CSI. In other words, the robustness is achieved through the addition of a power allocation block before the symbols are processed by the matrix $\mathbf{B}$ as depicted in Figure 6.1 as opposed to the naive design in Figure 6.2 where there is no power allocation and the transmission matrix $\mathbf{B}$ is designed as if no errors were present in the estimate of the channel matrix. Note then that we focus on the power allocation itself. Indeed, the transmitter separation into two blocks has also been taken in excellent works such as [Sch04, Boc05], and references therein by the same authors, where the power allocation is designed assuming perfect CSI, *i.e.*, they do not analyze the robustness problem. As commented before, a similar work as the one presented here on the design of a robust power allocation has been performed by [Big04], where the focus is not given to the design of the linear transmission matrix $\mathbf{B}$.
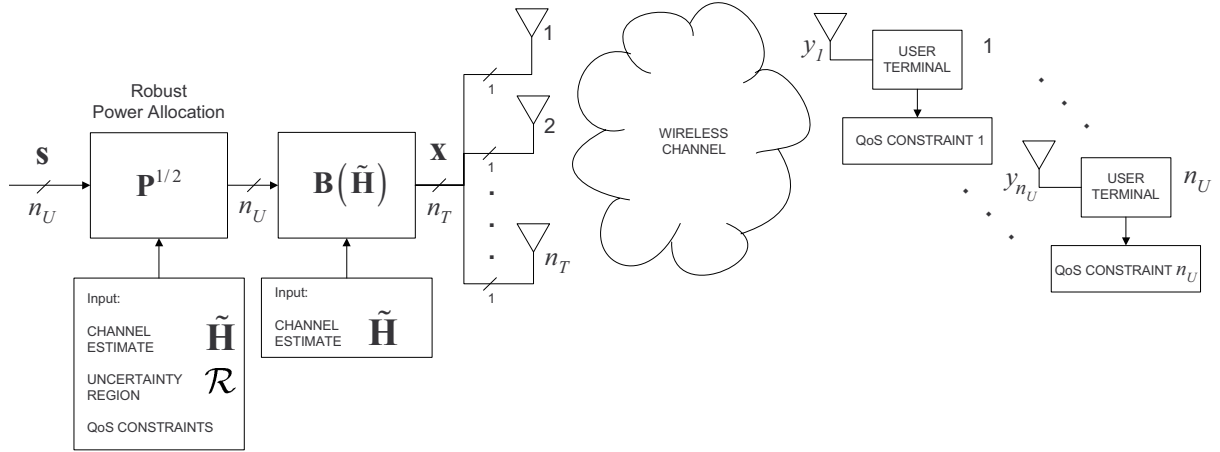
Figure 6.1: Robust downlink multi-user communication scheme. In the robust case we have added a power allocation block, that weights the information symbols prior to transforming them with the linear filter **B**. The power allocation is robustly designed to minimize the transmitted power while guaranteeing a certain QoS for each user. The optimal power allocation is found as a function of the channel estimate $\widetilde{\mathbf{H}}$, the uncertainty region $\mathcal{R}$, and the set of QoS constraints $\{\mathsf{qos}_i\}$.

6.5 Imperfect CSI and problem statement

As it has been widely commented in this dissertation, in a practical communications scenario, the assumption of perfect CSI at the transmitter and receiver sides is rather unrealistic. At the receivers side, the channel is usually estimated through training sequences (pilot symbols), and at the transmitter side, the CSI can be acquired through a feedback channel in FDD systems or from previously received symbols by exploiting the channel reciprocity in TDD systems. In both cases, different sources of errors can be identified depending on how the CSI is obtained, such as estimation Gaussian noise or the effects of a quantized feedback, among others.

In this section, we analyze the case where, due to the aforementioned imperfection in the CSI acquisition, both the transmitter and the receiver have only access to the same noisy estimate, $\widetilde{\mathbf{H}} = [\tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_{n_U}]^H$, of the actual channel $\mathbf{H}$. More precisely, it is assumed that the actual channel is inside an uncertainty region, $\mathcal{R} \subset \mathbb{C}^{n_U \times n_T}$, around its estimate, which formally can be expressed as

$$\mathbf{H} = \widetilde{\mathbf{H}} + \boldsymbol{\Delta}, \tag{6.4}$$

for some $\boldsymbol{\Delta} = [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_{n_U}]^H \in \mathcal{R}$. The shape and size of $\mathcal{R}$ model the kind of uncertainty in the channel estimate. For example, if the uncertainty stems from the fact that the CSI is a uniformly quantized version of the actual channel, then the entries of the error matrix are inside the interval $[\boldsymbol{\Delta}]_{ij} \in [-\rho, \rho] \times [-\mathrm{i}\,\rho, \mathrm{i}\,\rho]$, where 2ρ is the quantization step, and thus the uncertainty region is a hypercube, whose side length is the quantization step (more details are given in §6.6 or see further [PI06]).
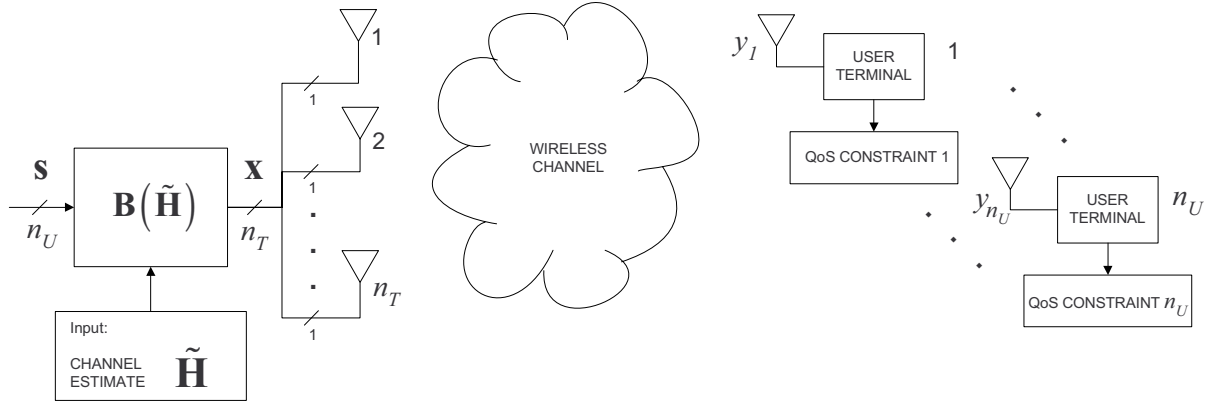
Figure 6.2: Naive downlink multi-user communication scheme. In this non-robust case, the transmitter performs a linear transformation of the information symbols, which is allowed to depend on the channel estimation $\widetilde{\mathbf{H}}$. Note that the design of the transmitter matrix $\mathbf{B}$ is designed like as if the estimate of the channel matrix $\widetilde{\mathbf{H}}$ were perfect.

As mentioned in the introduction, given a design for the transmitter matrix $\mathbf{B}$ as a function of the channel estimate $\widetilde{\mathbf{H}}$, the most general formulation of our problem is to robustly design the power allocation matrix $\mathbf{P}$, as in (6.5), so that the total transmitted power, $P_T(\mathbf{P})$, is minimized, and the QoS indicator for every user, $\mathsf{qos}_i(\mathbf{P}, \boldsymbol{\Delta})$, is always above a certain pre-required minimum quality threshold for each user, qos_i^0, for any possible realization of the error matrix $\boldsymbol{\Delta}$ inside the uncertainty region $\mathcal{R}$. Formally, this optimization problem can be stated as

$$
\begin{aligned}
\min_{\mathbf{P}} \quad & P_T(\mathbf{P}) \\
\text{s.t.} \quad & \mathsf{qos}_i(\mathbf{P}, \boldsymbol{\Delta}) \geq \mathsf{qos}_i^0, \quad 1 \leq i \leq n_U, \quad \forall \boldsymbol{\Delta} \in \mathcal{R}. \\
& p_i \geq 0, \quad 1 \leq i \leq n_U
\end{aligned}
\tag{6.5}
$$

Note that, by explicitly taking into account the imperfections of the CSI in the design process we obtain a communications system which is robust to uncertainties in the channel estimate. The robustness of the proposed system stems from the fact that, by explicitly guaranteeing that the QoS for every user is above a certain different threshold for any possible error realization inside the uncertainty region, an increase in the reliability against estimation errors is provided to the users.

The QoS indicator is chosen to be the perceived SINR for each user, sinr_i, and, consequently, the QoS requirements are expressed as SINR thresholds sinr_i^0, such that the user SINR must be above its corresponding threshold, *i.e.*, $\mathsf{sinr}_i \geq \mathsf{sinr}_i^0$ must be guaranteed. The choice of SINR as a QoS indicator is, however, rather arbitrary. For example, if the QoS requirement is desired to be expressed in terms of rate $\mathsf{r}_i \geq \mathsf{r}_i^0$, it is always possible to transform it into a SINR constraint as $\mathsf{sinr}_i^0 = \exp \mathsf{r}_i^0 - 1$. Similar expressions can be obtained when the QoS is selected to be the BER $\mathsf{ber}_i \leq \mathsf{ber}_i^0$ or the mean square error $\mathsf{mse}_i \leq \mathsf{mse}_i^0$ (see further our preliminary work [Pay06d] where the mean square error criterion has been taken).
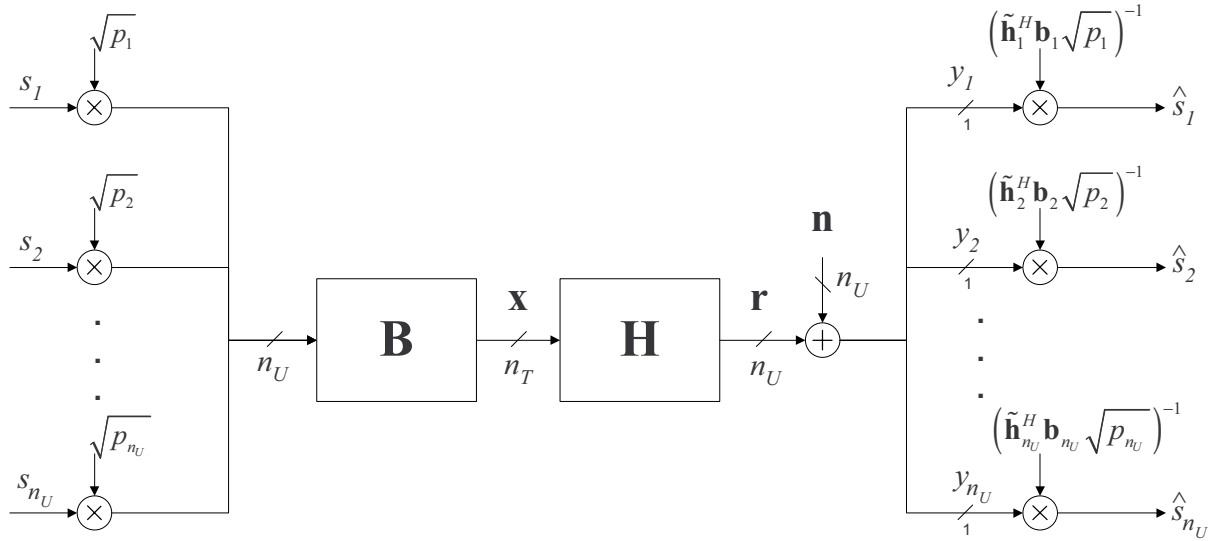
Figure 6.3: Schematic representation of the downlink communication system. Each receiver estimates its own symbol, $\hat{s}_i$, *e.g.*, dividing the incoming signal by the coefficient of s_i in the expression for the received signal in (6.9). Note that the SINR is not modified by this division.

In the following, we particularize the general optimization problem in (6.5) for our communications scheme. From (6.3), the transmitted power for this architecture is readily obtained as the linear function

$$P_T(\mathbf{P}) = \mathbb{E}\operatorname{Tr}\mathbf{x}\mathbf{x}^H = \operatorname{Tr}\mathbf{B}\mathbf{P}\mathbf{B}^H. \tag{6.6}$$

The expression for the SINR for each user can be obtained as follows. We begin from the received signal for i-th user in (6.2) when considering the uncertainty model in (6.4):

$$y_i = \mathbf{h}_i^H\mathbf{B}\mathbf{P}^{1/2}\mathbf{s} + n_i \tag{6.7}$$

$$= \tilde{\mathbf{h}}_i^H\mathbf{B}\mathbf{P}^{1/2}\mathbf{s} + \boldsymbol{\delta}_i^H\mathbf{B}\mathbf{P}^{1/2}\mathbf{s} + n_i \tag{6.8}$$

$$= \tilde{\mathbf{h}}_i^H\mathbf{b}_i\sqrt{p_i}s_i + (\tilde{\mathbf{h}}_i^H\bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H\mathbf{B})\mathbf{P}^{1/2}\mathbf{s} + n_i, \tag{6.9}$$

where $\bar{\mathbf{B}}_i \triangleq [\mathbf{b}_1, \ldots, \mathbf{b}_{i-1}, \mathbf{0}, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_{n_U}]$ (see Figure 6.3 for an schematic representation of the received signal). Expanding the second term in last equation, we obtain

$$(\tilde{\mathbf{h}}_i^H\bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H\mathbf{B})\mathbf{P}^{1/2}\mathbf{s} = \tilde{\mathbf{h}}_i^H\sum_{j\neq i}\mathbf{b}_j\sqrt{p_j}s_j + \boldsymbol{\delta}_i^H\sum_{j\neq i}\mathbf{b}_j\sqrt{p_j}s_j + \boldsymbol{\delta}_i^H\mathbf{b}_i\sqrt{p_i}s_i. \tag{6.10}$$

Now it is important to determine precisely the way how each user is going to estimate its own received symbol according to the reception of y_i. From the expansion in (6.10) it can be seen that the term $(\tilde{\mathbf{h}}_i^H\bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H\mathbf{B})\mathbf{P}^{1/2}\mathbf{s}$ in (6.9) does not have any known contribution to the reception of the symbol s_i because the only term that depends on s_i (the last one in (6.10)) also contains the factor $\boldsymbol{\delta}_i^H$ associated with the error in the channel estimate and which is unknown

for both communication ends. Since it has been assumed that each receiver has available the same channel estimation than the transmitter $\tilde{\mathbf{h}}_i$ and is aware of the power allocated to its signal p_i and the transmitter matrix $\mathbf{B}$, the most reasonable estimate of the received symbol as a function of the received signal is given by

$$\hat{s}_i = \frac{y_i}{\tilde{\mathbf{h}}_i^H \mathbf{b}_i \sqrt{p_i}} = s_i + \frac{(\tilde{\mathbf{h}}_i^H \bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H \mathbf{B})\mathbf{P}^{1/2}\mathbf{s} + n_i}{\tilde{\mathbf{h}}_i^H \mathbf{b}_i \sqrt{p_i}}, \tag{6.11}$$

where, from all that has been said, the second term has to be computed as interference.

Consequently, although in the formal definition of SINR all the terms containing the desired information should be computed as signal power, as it was shown in §2.5.3, it may be more useful to compute the *effective* SINR as in (2.43). With this consideration in mind, the effective SINR[1] experienced by i-th user can be directly obtained as

$$\mathsf{esinr}_i = \frac{\mathbb{E}|\tilde{\mathbf{h}}_i^H \mathbf{b}_i \sqrt{p_i} s_i|^2}{\mathbb{E}|(\tilde{\mathbf{h}}_i^H \bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H \mathbf{B})\mathbf{P}^{1/2}\mathbf{s} + n_i|^2} = \frac{|\tilde{\mathbf{h}}_i^H \mathbf{b}_i|^2 p_i}{(\tilde{\mathbf{h}}_i^H \bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H \mathbf{B})\mathbf{P}(\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i + \mathbf{B}^H \boldsymbol{\delta}_i) + \sigma^2}. \tag{6.12}$$

Particularizing the problem in (6.5) with the expression for the transmitted power in (6.6) and where the the QoS constraint $\mathsf{qos}_i \geq \mathsf{qos}_i^0$ is rewritten with the SINR expression in (6.12), we obtain the equivalent problem for our case:

$$\begin{aligned}
\min_{\mathbf{P}} \quad & \mathrm{Tr}\,\mathbf{BPB}^H, \\
\mathrm{s.\,t.} \quad & \big(\tilde{\mathbf{h}}_i^H \bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H \mathbf{B}\big)\mathbf{P}\big(\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i + \mathbf{B}^H \boldsymbol{\delta}_i\big) - \gamma_i^0\big|\tilde{\mathbf{h}}_i^H \mathbf{b}_i\big|^2 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \quad \forall \boldsymbol{\Delta} \in \mathcal{R}, \\
& p_i \geq 0, \quad 1 \leq i \leq n_U,
\end{aligned} \tag{6.13}$$

where we have defined $\gamma_i^0 = 1/\mathsf{esinr}_i^0$. Note that since the restrictions for each user (indexed by i) in the optimization problem (6.13) have to be guaranteed for all $\boldsymbol{\Delta} \in \mathcal{R}$, it is sufficient to impose that they be satisfied for the worst-case situation, *i.e.*, the case where $\boldsymbol{\Delta}$ is such that it maximizes the value of the restrictions as

$$\begin{aligned}
\min_{\mathbf{P}} \quad & \mathrm{Tr}\,\mathbf{BPB}^H, \\
\mathrm{s.\,t.} \quad & \sup_{\boldsymbol{\Delta} \in \mathcal{R}} \big(\tilde{\mathbf{h}}_i^H \bar{\mathbf{B}}_i + \boldsymbol{\delta}_i^H \mathbf{B}\big)\mathbf{P}\big(\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i + \mathbf{B}^H \boldsymbol{\delta}_i\big) - \gamma_i^0\big|\tilde{\mathbf{h}}_i^H \mathbf{b}_i\big|^2 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \quad (6.14) \\
& p_i \geq 0, \quad 1 \leq i \leq n_U
\end{aligned}$$

whose solution is the same as in the problem (6.13). From [Boy04], we know that $\sup_{\mathbf{z}_2 \in \mathcal{A}} f(\mathbf{z}_1, \mathbf{z}_2)$ is a convex function in $\mathbf{z}_1$ if $f(\mathbf{z}_1, \mathbf{z}_2)$ is also convex in $\mathbf{z}_1$ for all $\mathbf{z}_2$, unaffected by the shape of $\mathcal{A}$. This implies that the restrictions in (6.14) are convex in $\mathbf{P}$ for every possible shape and size of the uncertainty region $\mathcal{R}$ because they are defined as the supremum of a linear (and thus convex) function of $\mathbf{P}$.

---

[1]For the sake of clarity, from this point the effective SINR will be referred to as SINR.

In the following section, we particularize the convex problem in (6.14) for a number of uncertainty regions that model the most interesting practical situations, like, *e.g.*, the fact that the channel estimation is a quantized version of the actual channel, or that the channel estimation has a Gaussian noise contribution with respect to the actual channel or, even the case with a combination of both sources of errors. The cases of spherical or elliptical uncertainty regions have already been considered in [AS05, Wie05a, Wie06] when dealing with robust designs. We try to go a step beyond by doing a generalization effort to include many different uncertainty regions. In addition, for each one of these uncertainty regions, a ready-to-program particularization of the convex optimization general problem in (6.14) is given.

6.6   Uncertainty regions

The definition of the uncertainty region $\mathcal{R}$ should take into account the quality of the channel estimate and the imperfections in the estimation process that generate the error in such a way that the mathematical optimization problem as in (6.14) is directly related to the physical phenomenon producing the error.

In this section, we focus our attention on the particularization of the general problem in (6.14), of finding the power allocation that minimizes the transmitted power while guaranteeing the QoS for the users, for some interesting uncertainty regions derived from the previously described error sources and combinations of them. In addition, we particularize the obtained expressions for the case where $\mathbf{B} = \mathbf{B}_{\mathrm{ZF}} = \widetilde{\mathbf{H}}^H \big(\widetilde{\mathbf{H}} \widetilde{\mathbf{H}}^H \big)^{-1}$. This choice is made for the sake of simplicity, because, in this case, $|\tilde{\mathbf{h}}_i^H \mathbf{b}_i|^2 = 1$ and $\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i = \mathbf{0}$, for all i, and the general robust problem in (6.14) becomes

$$
\begin{aligned}
\min_{\mathbf{P}} \quad & \operatorname{Tr} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H, \\
\mathrm{s.\,t.} \quad & \sup_{\boldsymbol{\Delta} \in \mathcal{R}} \boldsymbol{\delta}_i^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H \boldsymbol{\delta}_i - \gamma_i^0 p_i + \sigma^2 \le 0, \quad 1 \le i \le n_U, \\
& p_i \ge 0, \quad 1 \le i \le n_U.
\end{aligned}
\tag{6.15}
$$

In addition to the simplification of the obtained optimization problem, the choice $\mathbf{B} = \mathbf{B}_{\mathrm{ZF}}$ has been proven optimal in terms of signal reception quality at high SNRs and it is also widely utilized in the wireless downlink literature, *e.g.*, [Yoo05]. However, it is important to recall that the procedures described below are valid no matter what kind of transmission matrix $\mathbf{B}$ is chosen (see further §6.6.5).

6.6.1   Estimation white Gaussian noise

In this section, we deal with the case where the corresponding channel vector of each of the users, $\mathbf{h}_i$, is estimated independently from the others. This model is valid, for example, if each

user estimates its own channel and feeds it back to the transmitter through an ideal feedback link. We thus consider that the estimate $\tilde{\mathbf{h}}_i$ is a noisy version of the actual channel $\mathbf{h}_i$ corrupted with AWGN,

$$\tilde{\mathbf{h}}_i = \mathbf{h}_i + \mathbf{w}_i, \ \forall i, \quad \text{or, compactly,} \quad \widetilde{\mathbf{H}} = [\tilde{\mathbf{h}}_1, \dots, \tilde{\mathbf{h}}_{n_U}]^H = \mathbf{H} + \mathbf{W} \tag{6.16}$$

where $\mathbf{w}_i \in \mathbb{C}^{n_T \times 1}$ represents the estimation noise, and whose entries are proper i.i.d. complex Gaussian random variables, $\mathbf{w}_i \sim \mathcal{CN}(\mathbf{0}, \varsigma_i^2 \mathbf{I})$, and where $\mathbf{W}$ is related with $\mathbf{\Delta}$ as $\mathbf{\Delta} = -\mathbf{W}$. The estimation noise power is characterized by ς_i^2 and can be different for each user, so that different qualities in the channel estimation per user can be modeled (due to, *e.g.*, different distances from the users to the BS).

We must now relate the estimation error model in (6.16) with the uncertainty model that we have considered in this work, $\mathbf{H} = \widetilde{\mathbf{H}} + \mathbf{\Delta}$, where $\mathbf{\Delta}$ belongs to an uncertainty region, $\mathcal{R}$. By inspection from (6.16), one simply obtains $\mathbf{\Delta} = -\mathbf{W}$ (and $\boldsymbol{\delta}_i = -\mathbf{w}_i$). There are a couple of implications of this equality that need to be further commented:

- The first fact is that, since for this error model each one of the estimation noise vectors $\{\boldsymbol{\delta}_i\}$ are white and Gaussian, their probability distribution has spherical symmetry around the origin. The usual approach in this case is to define a spherical uncertainty region, as in [PI06] or [AS05], for the error committed in the estimation of the channel for each user formulated as follows

$$\mathcal{R} = \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}} \triangleq \left\{ \mathbf{\Delta} \in \mathbb{C}^{n_U \times n_T} \ \big| \ \|\boldsymbol{\delta}_i\|^2 = \boldsymbol{\delta}_i^H \boldsymbol{\delta}_i \leq R_i^2, \ \forall i \right\}, \tag{6.17}$$

where R_i is related with the noise power ς_i^2 and also with the probability that the actual channel is inside the uncertainty region, as commented in the following. Note that our formulation allows to consider spherical regions with different radii R_i.

- Another important point comes from the fact that, since the uncertainty region $\mathcal{R}$ is a bounded set and the Gaussian entries of the estimation error $\mathbf{\Delta}$ are unbounded, the actual channel error will belong to the uncertainty region with a certain probability, $P_{\mathrm{in}} = \mathrm{Pr}\{\mathbf{\Delta} \in \mathcal{R}\} < 1$. Consequently, in this case, the QoS required for the users will only be guaranteed with a probability equal to P_{in}. For the case where the error is outside the uncertainty region and an outage event is declared because the QoS can not be guaranteed. See [PI06] and also [Ron06] for a detailed description of the relation between R_i, P_{in}, and the power of the estimation noise, ς_i^2.

From the expression of the uncertainty region in (6.17), it can be seen that the quality of the estimation of the channel of i-th user is determined by the radius of the uncertainty region R_i. Since R_i is related to the power of the estimation noise, ς_i^2, the bigger the uncertainty in the

channel estimation, the bigger the radius of the uncertainty region. This allows us to consider different channel estimation qualities for each user, which models, for example, a situation where the users are placed at different distances of the BS, *i.e.*, the uncertainty radius, R_i, of a user which is close to the BS is lower than that of a user which is placed far from it.

Once the uncertainty region has been properly defined as in (6.17) for the case where the error in the channel estimate is due to estimation Gaussian noise, the optimization problem in (6.15) can now be solved. Its solution is given in the following proposition.

Proposition 6.6.1 *Let us define the diagonal matrices $\mathbf{\Gamma}$ and $\mathbf{\Sigma}$, with $[\mathbf{\Gamma}]_{ii} = R_i^2/\gamma_i^0$ and $[\mathbf{\Sigma}]_{ii} = \sigma^2/\gamma_i^0$. Then it follows that the convex problem in (6.15) for the case $\mathbf{\Delta} \in \mathcal{R} = \mathcal{S}_{\{R_i^2\}}^{ind}$ has a feasible solution if, and only if, $\lambda_{\max}(\mathbf{B}_{\mathrm{ZF}}\mathbf{\Gamma}\mathbf{B}_{\mathrm{ZF}}^H) < 1$. In this case, the entries of its solution, $\mathbf{P}^\star$, are given by*

$$p_i^\star = \frac{R_i^2\mu + \sigma^2}{\gamma_i^0} = \mathsf{esinr}_i^0\big(R_i^2\mu + \sigma^2\big), \quad \forall i \in \{1,\ldots,n_U\} \tag{6.18}$$

where μ is the unique solution to the fixed point equation $\lambda_{\max}(\mathbf{B}_{\mathrm{ZF}}(\mathbf{\Gamma}\mu + \mathbf{\Sigma})\mathbf{B}_{\mathrm{ZF}}^H) = \mu$, which can be very efficiently solved, utilizing, e.g., the Newton method, applying the expression for the differential of $\lambda_{\max}(\mathbf{X})$ found in [Mag99, p. 161].

Proof See Appendix 6.A.1.   ∎

The solution for the optimal power allocation in (6.18) can be particularized for the case where there is no error in the channel estimate and, consequently, $R_i = 0$, $\forall i$. In this case, $p_{i,\mathrm{perf}}^\star = \mathsf{esinr}_i^0\sigma^2$. This allows us to interpret that, so that the QoS are guaranteed in the imperfect CSI case, there is an increase of $\mathsf{esinr}_i^0 R_i^2\mu$ in the power allocated to i-th symbol with respect to the perfect CSI, which is the minimum price to pay to obtain a robust design for the case $\mathbf{B} = \mathbf{B}_{\mathrm{ZF}}$.

6.6.2   Colored noise case

It is possible to generalize the case presented above to include the effects of a colored Gaussian estimation noise. In this case, the model for the estimation error is the same as in (6.16) but we consider $\mathbf{w}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_i)$, with $\mathbf{C}_i$ being a general positive definite matrix. Due to the presence of the general correlation matrix $\mathbf{C}_i$, it can be shown (see, *e.g.*, [Ron06]) that this case corresponds to an elliptical uncertainty region defined as

$$\mathcal{R} = \mathcal{E}_{\{R_i^2\},\{\mathbf{C}_i\}}^{\mathrm{ind}} \triangleq \big\{\mathbf{\Delta} \in \mathbb{C}^{n_U \times n_T} \,\big|\, \boldsymbol{\delta}_i^H \mathbf{C}_i^{-1} \boldsymbol{\delta}_i \leq R_i^2, \,\forall i\big\}. \tag{6.19}$$

The solution to the problem in (6.15) for this case is described in the following proposition.
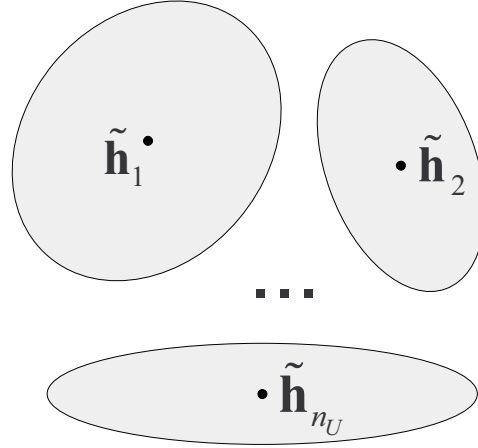
Figure 6.4: Graphical representation of the elliptical/spherical uncertainty region, for the case where there is an individual constraint on the error committed in each channel, $\mathbf{h}_i$.

Proposition 6.6.2 *The solution to the optimal power allocation in (6.15) for the case where* $\boldsymbol{\Delta} \in \mathcal{E}^{ind}_{\{R_i^2\},\{\mathbf{C}_i\}}$, *is given by the solution to the following convex optimization problem:*

$$
\begin{aligned}
\min_{\mathbf{P}} \quad & \operatorname{Tr} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^{H}, \\
\mathrm{s.\,t.} \quad & R_i^2 \lambda_{\max}(\mathbf{C}_i^{1/2} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^{H} \mathbf{C}_i^{1/2}) - \gamma_i^0 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \\
& p_i \geq 0, \quad 1 \leq i \leq n_U,
\end{aligned}
\tag{6.20}
$$

which can be solved numerically in a very efficient manner following the methods described in [Boy04]. Note also that the gradient and Hessian of the first constraint in (6.20) can be found in [Mag99].

Proof See Appendix 6.A.2. ∎

Corollary 1 *For the particular case where all the correlation matrices* $\mathbf{C}_i$ *are equal for all* i, $\mathbf{C}_i = \mathbf{C}$, *the convex problem in (6.20) has a feasible solution if, and only if,* $\lambda_{\max}(\mathbf{C}^{1/2} \mathbf{B}_{\mathrm{ZF}} \boldsymbol{\Gamma} \mathbf{B}_{\mathrm{ZF}}^{H} \mathbf{C}^{1/2}) < 1$. *In this case, the entries of its solution,* $\mathbf{P}^{\star}$, *are given by the expression in (6.18) where in this case* μ *is the unique solution to the fixed point equation* $\lambda_{\max}(\mathbf{C}^{1/2} \mathbf{B}_{\mathrm{ZF}} (\boldsymbol{\Gamma}\mu + \boldsymbol{\Sigma}) \mathbf{B}_{\mathrm{ZF}}^{H} \mathbf{C}^{1/2}) = \mu$.

Proof The proof follows directly from the steps performed in the proof in Appendix 6.A.1 replacing the equation (6.33) by the one obtained in (6.20) with $\mathbf{C}_i = \mathbf{C}$, for all i. ∎

6.6.3   Effects of a quantized CSI

In the previous sections, we have considered the case where the estimation error is modeled as Gaussian noise. In this section we deal with the case where the available CSI, $\widetilde{\mathbf{H}}$, is a quantized
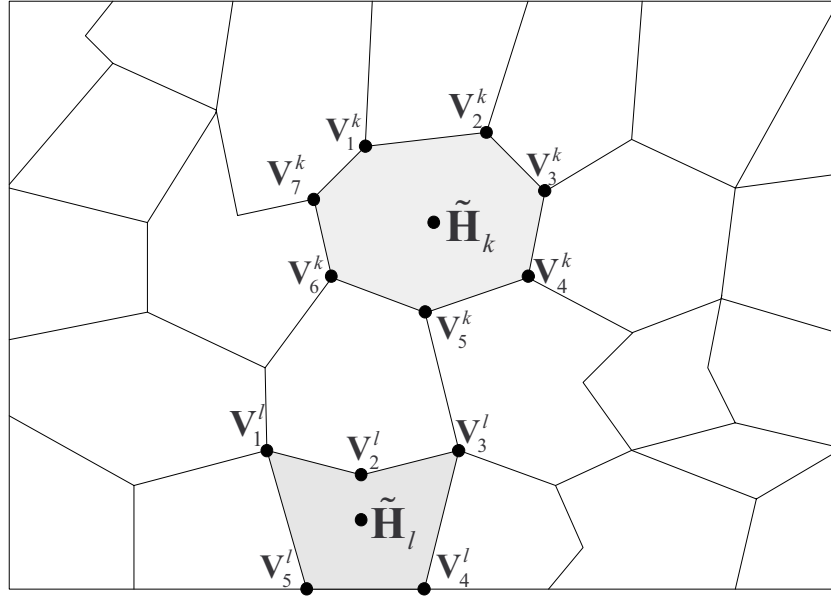
Figure 6.5: Graphical representation of the uncertainty regions that arise when the available CSI is a quantized version of the actual channel. Our formulation allows that each quantization region can have a different shape and that each region can be defined by a different number of vertices, as well as the fact that the quantization regions are not necessarily convex.

version of the actual channel $\mathbf{H}$. This would correspond to the practical case where each receiver quantizes a perfect estimate of its own channel[2] and then feeds back this quantized information to the transmitter through a digital feedback link.

The quantization procedure that we deal with in this section is described in the following. We consider the channel matrices space with K points, $\{\mathbf{H}_k\}$. Each one of these points $\mathbf{H}_k$ is the representative of the region $\mathcal{R}_k$ surrounding it (see Figure 6.5). Each region $\mathcal{R}_k$ is a polyhedron[3] (not necessarily convex) with M_k vertices given by the set $V_k = \{\mathbf{V}_1^k, \mathbf{V}_2^k, \ldots, \mathbf{V}_{M_k}^k\} \subset \mathbb{C}^{n_U \times n_T}$, where the rows of $\mathbf{V}_m^k$ are defined as $\mathbf{V}_m^k = [\mathbf{v}_{m,1}^k, \ldots, \mathbf{v}_{m,n_U}^k]^H$.

The transmitter is informed with the index k of the region where the actual channel belongs to and then the estimate of the channel becomes $\widetilde{\mathbf{H}} = \mathbf{H}_k$ and, consequently, the quantization uncertainty region, $\mathcal{Q}_{V_k}$, becomes the polyhedron around $\mathbf{H}_k$, $\mathcal{Q}_{V_k} \triangleq \mathcal{R}_k$. For the sake of notation, we drop the index k w.l.o.g., and present the characterization of the solution to the problem in (6.15) in the next proposition.

Proposition 6.6.3 *Consider the case where $\mathbf{\Delta} \in \mathcal{Q}_V$, then it follows that the convex problem*

---

[2]The case where this estimate is imperfect is considered in the next subsection.

[3]The term polyhedron is understood as a geometric polytope composed of its boundary and its interior.

in (6.15) can be rewritten as the following linear program:

$$\min_{\mathbf{P}} \quad \operatorname{Tr} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^{H},$$

$$\text{s.t.} \quad \mathbf{v}_{m,i}^{H} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^{H} \mathbf{v}_{m,i} - \gamma_{i}^{0} p_{i} + \sigma^{2} \leq 0, \quad 1 \leq i \leq n_{U}, \quad 1 \leq m \leq M, \qquad (6.21)$$

$$p_{i} \geq 0, \quad 1 \leq i \leq n_{U}.$$

Proof See Appendix 6.A.3. ∎

Note that the set of restrictions in (6.21) is now just a list of linear restrictions in $\mathbf{P}$.

In addition, it must be highlighted that, the versatility of the quantization method described above allows us to consider, as a particular case, the situation where each user utilizes a different quantization rule. As an example of application, we consider the simple case where each user uniformly quantizes the real and imaginary parts of the entries of their corresponding channel utilizing a fixed step of size $2\rho_i$ (different for each user). In this case, all the regions defined by V^k are equal $V^k = V$ (which corresponds to uniform quantization). For this case the vertices can be expressed as

$$\mathbf{V}_m = \mathbf{V}_m^{\mathrm{re}} + \mathrm{i}\,\mathbf{V}_m^{\mathrm{im}}, \quad m = 1, \dots, 2^{n_U \times n_T}, \qquad (6.22)$$

where the elements of the matrices $\mathbf{V}_m^{\mathrm{re}}$ and $\mathbf{V}_m^{\mathrm{im}}$ are of the form $[\mathbf{V}_m^{\mathrm{xx}}]_{ij} = \pm \rho_i$.

6.6.4 Combination of regions

In realistic setups, the error in the channel matrix may come from more than one source. A typical example of this situation is the case where the available channel $\widetilde{\mathbf{H}}$ is a quantized version of a corrupted version of the actual channel (see [PI06]). This would correspond, *e.g.*, to scenarios where each user imperfectly estimates its own channel and then feeds back, to the transmitter, a quantized version of this noisy estimate.

In this case, the estimation error matrix $\boldsymbol{\Delta}$ can be considered to be the sum of two terms, $\boldsymbol{\Delta} = \mathbf{S} + \mathbf{Q}$, the first one takes into account the contribution due to the Gaussian noise, thus $\mathbf{S} \in \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}}$, and the second one models the effects that the channel estimate comes from a quantization process, which implies that $\mathbf{Q} \in \mathcal{Q}_V$. In this case, the shape of the uncertainty region where the global error matrix $\boldsymbol{\Delta}$ belongs to is the hyper-convolution of the two considered regions as we have depicted in Figure 6.6, and the solution characterizing the optimal power allocation is given next.

Proposition 6.6.4 *Consider the case where the uncertainty matrix is $\boldsymbol{\Delta} = \mathbf{S} + \mathbf{Q}$, with $\mathbf{S} \in \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}}$ and $\mathbf{Q} \in \mathcal{Q}_V$. It then follows that the optimization problem in (6.15) is equivalent to*
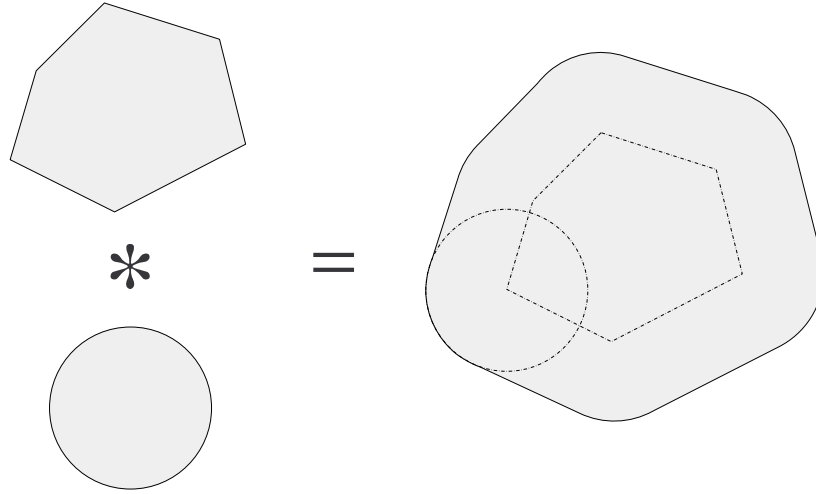
Figure 6.6: Graphical representation of the uncertainty region that arises when two different sources of errors come into consideration. The left side of the picture describes the uncertainty regions associated with two sources of errors, *e.g.*, quantization effects (up) and estimation Gaussian noise (down). The resulting uncertainty region (right) is the "convolution" of the two left regions as indicated by the dashed lines.

solving the convex program

$$
\begin{aligned}
\min_{\mathbf{P}} \quad & \mathrm{Tr}\,\mathbf{B}_{\mathrm{ZF}}\mathbf{P}\mathbf{B}_{\mathrm{ZF}}^H, \\
\mathrm{s.\,t.} \quad & \big(\mathbf{s}_{m,i}^\star(\mathbf{P}) + \mathbf{v}_{m,i}\big)^H \mathbf{B}_{\mathrm{ZF}}\mathbf{P}\mathbf{B}_{\mathrm{ZF}}^H \big(\mathbf{s}_{m,i}^\star(\mathbf{P}) + \mathbf{v}_{m,i}\big) - \\
& \qquad\qquad - \gamma_i^0 p_i + \sigma^2 \le 0, \quad 1 \le i \le n_U, \quad 1 \le m \le M, \\
& p_i \ge 0, \quad 1 \le i \le n_U,
\end{aligned}
\tag{6.23}
$$

where $\mathbf{s}_{m,i}^\star(\mathbf{P})$ in (6.23) depends on $\mathbf{P}$ and it is the solution to the optimization problem described in Appendix 6.B with $\mathbf{A} = \mathbf{B}_{\mathrm{ZF}}\mathbf{P}\mathbf{B}_{\mathrm{ZF}}^H$, $\tilde{\mathbf{y}} = \mathbf{v}_{m,i}$, and $b = R_i^2$.

Proof See Appendix 6.A.4.   ■

Since no closed-form expression for the solution to the convex optimization problem in (6.23) is apparently available, iterative algorithms are needed to obtain a numerical solution. In this case, at each iteration of the algorithm, the restrictions in (6.23) have to be numerically evaluated and, consequently, $\mathbf{s}_{m,i}^\star(\mathbf{P})$ has to be computed as a function of the value of $\mathbf{P}$ in the current iteration, as described in Appendix 6.B.

6.6.5   Example of extension to other types of transmitters

To illustrate the generality of the methods presented in this work, in this section we present the convex optimization problem whose solution gives the robust power allocation for the general

case where we do not impose a particular structure to the transmitter matrix $\mathbf{B}$. As for the definition of the uncertainty region, we consider the very generic case where the estimation error $\boldsymbol{\Delta}$ is modeled as $\boldsymbol{\Delta} = \mathbf{E} + \mathbf{Q}$ where $\mathbf{E} \in \mathcal{E}_{\{R_i^2\},\{\mathbf{C}_i\}}^{\text{ind}}$ takes into account the imperfections in the available channel due to colored Gaussian estimation noise and $\mathbf{Q} \in \mathcal{Q}_V$ models the effects of the quantization of the channel estimate. With these assumptions, the general problem in (6.14) becomes the following convex program:

$$
\begin{aligned}
\min_{\mathbf{P}} \quad & \operatorname{Tr} \mathbf{BPB}^H, \\
\text{s.t.} \quad & \left(\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i + \mathbf{B}^H(\mathbf{e}_i^\star(\mathbf{P}) + \mathbf{v}_{m,i})\right)^H \mathbf{P}\left(\bar{\mathbf{B}}_i^H \tilde{\mathbf{h}}_i + \mathbf{B}^H(\mathbf{e}_i^\star(\mathbf{P}) + \mathbf{v}_{m,i})\right) \\
& \qquad - \gamma_i^0 \big|\tilde{\mathbf{h}}_i^H \mathbf{b}_i\big|^2 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \quad 1 \leq m \leq M, \\
& p_i \geq 0, \quad 1 \leq i \leq n_U,
\end{aligned}
\tag{6.24}
$$

where $\mathbf{e}_i^\star(\mathbf{P})$ is the solution to the optimization problem described in Appendix 6.B with $\mathbf{A} = \mathbf{C}_i^{1/2}\mathbf{BPB}^H\mathbf{C}_i^{1/2}$, $\tilde{\mathbf{y}} = \mathbf{C}_i^{-1/2}\left(\mathbf{v}_{m,i} + \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\bar{\mathbf{B}}_i^H\tilde{\mathbf{h}}_i\right)$, and $b = R_i^2$. This statement is left without proof because it is very similar to that of Proposition 6.6.4. Note that the convex optimization problem in (6.24) admits, as particular cases, most of the convex problems obtained in the previous sections.

6.7 Practical issues

When numerically solving convex optimization problems such as (6.20), (6.21), (6.23), and (6.24) where no closed form solution is available, there are two main considerations that have to be taken into account before proceeding to the numerical optimization itself. First of all, it is important to study the feasibility of the problem because if the feasible set is empty, then, no solution exists. In §6.7.1, we give a generic feasibility test to verify the non-emptiness of the feasible region. Secondly, if the feasible region is non-empty, a feasible initial value of the optimization variable, $\mathbf{P}^{(0)}$, has to be provided to the iterative numerical optimization algorithm. A procedure to obtain this initial value is described in 6.7.2.

6.7.1 Feasibility

First of all, note that the restrictions in the problems (6.20), (6.21), (6.23), and (6.24) can be equivalently expressed as a list of restrictions indexed by i and/or m. For the sake of notation we now define u as the index of the list, which implies that the restrictions in (6.20), (6.21) and (6.23) are of the general form

$$
r_u(\mathbf{P}) + \sigma^2 \leq 0, \quad \forall u,
\tag{6.25}
$$

where $r_u(\mathbf{P})$ is a convex function in $\mathbf{P}$ and homogenous of degree 1[4], *i.e.*, it fulfills that $r_u(\alpha\mathbf{P}) = \alpha r_u(\mathbf{P})$. Clearly, from all this set of restrictions, if the condition (6.25) is met for the maximum w.r.t. u then is met for all u. Thus, we define $r(\mathbf{P}) = \max_u r_u(\mathbf{P})$, which is also homogeneous of degree 1. An equivalent restriction to (6.25) for the feasibility problem is then

$$r(\mathbf{P}) + \sigma^2 \leq 0. \tag{6.26}$$

Now, if $\mathbf{P}$ is feasible, then $\alpha\mathbf{P}$ with $\alpha > 1$ is also feasible. In particular, the limit case for $\alpha \to \infty$ is also feasible, and, in this limit case, the noise variance σ^2 does not have any influence in the inequality constraint in (6.26), and therefore the feasibility problem is equivalent to proving the existence of a $\mathbf{P}$ matrix such that $r(\mathbf{P}) < 0$, which can be further simplified to

$$r(\mathbf{P}) < 0 \Rightarrow r(\bar{\mathbf{P}}) < 0, \ \mathrm{Tr}\,\bar{\mathbf{P}} = 1, \tag{6.27}$$

where we have defined $\bar{\mathbf{P}} = \mathbf{P}/\mathrm{Tr}\,\mathbf{P}$ and the homogeneity property $r(\alpha\mathbf{P}) = \alpha r(\mathbf{P})$ has been utilized. Since we are only interested in proving the existence of a matrix that fulfills (6.27) we can restrict our attention to the matrix that is most *likely* to fulfill it, *i.e.*, the matrix that minimizes the term $r(\bar{\mathbf{P}})$ in (6.27), which is the solution to

$$\begin{aligned}
\min_{\bar{\mathbf{P}}} \quad & r(\bar{\mathbf{P}}), \\
\mathrm{s.\,t.} \quad & \mathrm{Tr}\,\bar{\mathbf{P}} = 1, \\
& \bar{p}_i \geq 0, \quad 1 \leq i \leq n_U.
\end{aligned} \tag{6.28}$$

Since $r(\bar{\mathbf{P}})$ is defined as the maximum of a finite set of convex functions it is also convex, and consequently the optimization problem in (6.28) is also convex and its solution always exists and can be efficiently calculated, and is denoted by $\bar{\mathbf{P}}^\star$.

Once we have obtained this solution, it only remains to check wether $r(\bar{\mathbf{P}}^\star)$ is lower than zero (which implies that (6.27) is fulfilled and the problem is feasible), or $r(\bar{\mathbf{P}}^\star)$ is greater than or equal to zero. Clearly, if the minimum value of $r(\bar{\mathbf{P}})$ does not fulfill (6.27) then no other value of $r(\bar{\mathbf{P}})$ can fulfill it and the problem becomes infeasible.

When the problem becomes infeasible, it means that there exists no power allocation such that all the QoS constraints can be fulfilled. In this case, an outage event can be declared, or, alternatively, some of the QoS constraints could be relaxed.

6.7.2 Starting point

If the problem is feasible, *i.e.*, if $r(\bar{\mathbf{P}}^\star) < 0$, then there exists a solution to the considered original problem (6.20), (6.21), (6.23), or (6.24) and efficient numerical algorithms can be utilized to

---

[4]The function $r_u(\mathbf{P})$ is homogeneous because it is defined as the supremum with respect to $\boldsymbol{\Delta} \in \mathcal{R}$ of a set of linear, and thus homogeneous of degree 1, functions of $\mathbf{P}$.

calculate it. However, the numerical algorithms need a starting feasible point, $\mathbf{P}^{(0)}$, to begin the iterative procedure to compute the solution.

We propose a heuristic starting point of the form $\mathbf{P}^{(0)} = \beta\bar{\mathbf{P}}^\star$, where β is calculated so that $\mathbf{P}^{(0)}$ is feasible and its corresponding transmitted power is minimized so that the speed of convergence is increased.

Since we want the starting point $\mathbf{P}^{(0)}$ to be feasible, it has to fulfill all the restrictions in (6.25). Thus, it is only necessary to impose that it fulfills (6.26), from which we obtain

$$r(\mathbf{P}^{(0)}) + \sigma^2 \leq 0 \Rightarrow r(\beta\bar{\mathbf{P}}^\star) + \sigma^2 \leq 0 \Rightarrow \beta r(\bar{\mathbf{P}}^\star) + \sigma^2 \leq 0 \Rightarrow \beta \geq \frac{-\sigma^2}{r(\bar{\mathbf{P}}^\star)}. \tag{6.29}$$

Note that the feasibility of the problem ($r(\bar{\mathbf{P}}^\star) < 0$) is necessary to guarantee that β exists and is positive, and consequently $\mathbf{P}^{(0)}$ is positive semi-definite as expected.

Since we also want to minimize the transmitted power associated with the starting point $\mathbf{P}^{(0)}$, we have to choose the minimum β that fulfills the feasibility condition, which is clearly given by

$$\beta = \frac{-\sigma^2}{r(\bar{\mathbf{P}}^\star)}. \tag{6.30}$$

6.8 Simulations

In the following, some simulation results are provided in order to give insight into the benefits of the proposed robust design for the power allocation.

In Figure 6.7, we have considered a two user scenario where the uncertainty matrix belongs to a independently constrained spherical region as discussed in §6.6.1. We have plotted the feasibility region (*i.e.*, the set of powers p_1 and p_2 for which the QoS are fulfilled $\forall \mathbf{\Delta} \in \mathcal{R}$) for different values of the uncertainty radius, $R_1 = R_2 = R$. Note that as the uncertainty radius increases the region becomes smaller. If we continued to increase the radius there would be a point where the feasibility region would become void and, thus, the optimization problem is infeasible.

In Figure 6.8, we have also considered a two user scenario such that the uncertainty region is an independently constrained spherical region (see §6.6.1 again) and we have fixed the value of the uncertainty radius for user 2, R_2. In the upper plot, we have drawn the feasibility region (*i.e.*, the set of esinr_i^0 such that the problem is feasible) for different values of the uncertainty radius of user 1, R_1. The feasibility region corresponds to the area below each one of the curves. In the lower plot we have represented the total transmitted power along the red dotted line. Note that, as we approach the limit of the feasibility region, the necessary transmitted power to guarantee the QoS constraints becomes arbitrarily large, as expected.
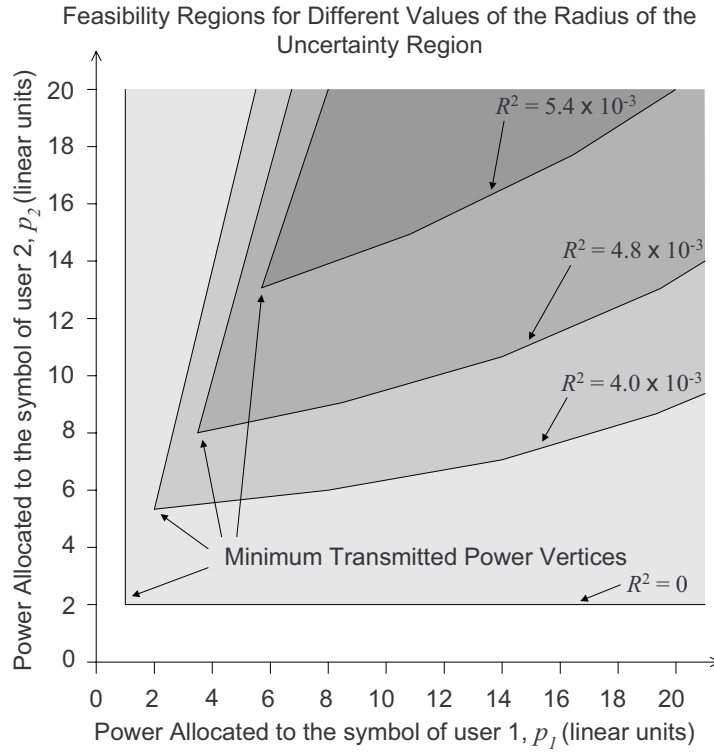
Figure 6.7: Feasibility region as a function of the power allocated to the symbol of each user for different values of the radius of the uncertainty region. The lower-left corner of each region corresponds to the feasible point (the QoS constraints are met) such that the transmitted power is the lowest.

In addition we have conducted some simulations to show the goodness of our proposed robust method. Since no references have been found in the literature that utilized our uncertainty model, we have designed two simple *ad-hoc* power allocation techniques, which are very briefly described. The first technique consists in forcing $\mathbf{P} = \lambda\mathbf{I}$, and obtain the λ such that all QoS constraints for every user are fulfilled $\forall \mathbf{\Delta} \in \mathcal{R}$. The second technique consists in forcing $[\mathbf{P}]_{ii} = \nu\mathsf{sinr}_i^0\sigma^2$, where again ν is selected so that the QoS requirements per user are met $\forall \mathbf{\Delta} \in \mathcal{R}$.

In Figure 6.9 we have considered a two user scenario with independently constrained spherical uncertainty regions (see §6.6.1). The necessary total transmitted power, is plotted as a function of the QoS requirement for user 1, esinr_1^0, for the case of perfect CSI, and also for the case of imperfect CSI, keeping fixed esinr_2^0 in all the cases. In the latter case, we have plotted the necessary power for our robust solution and also for the two *ad-hoc* approaches. As expected, the robust solution yields the minimum necessary transmitted power to guarantee the QoS constraints. Note also that, with the same QoS constraints, there is an increase in the minimum necessary transmitted power for the case of imperfect CSI with respect to the perfect CSI case. In Figure 6.10, a similar simulation has been conducted with 8 users. Note how the goodness of our proposed approach is remarkably higher in this scenario with more users.
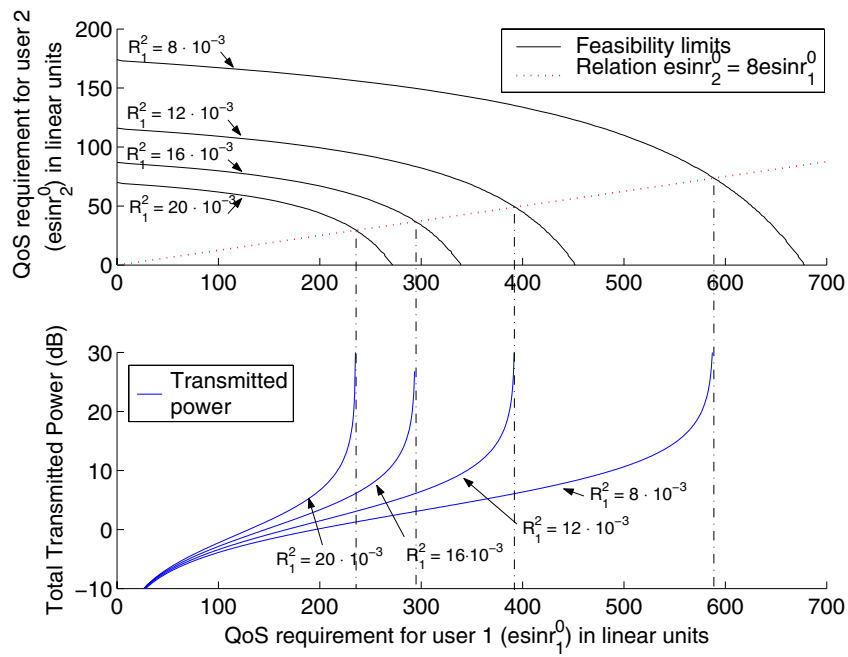
Figure 6.8: The upper plot depicts the feasibility region as a function of the QoS requirement for each user for different values of the radius of the uncertainty region. The lower plot represents the total transmitted power following the dotted red line.
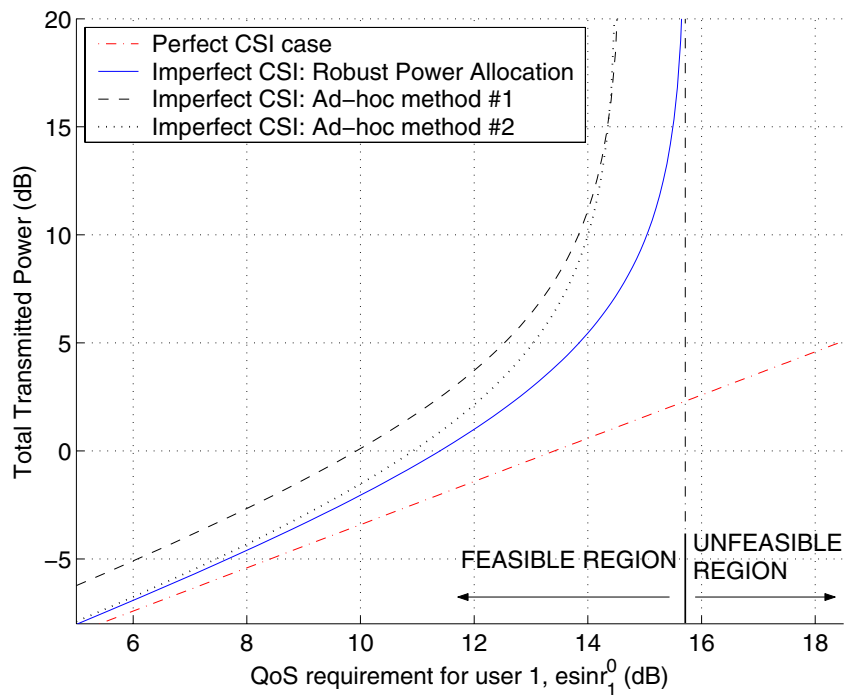
Figure 6.9: Total transmitted power versus QoS requirement for user 1, esinr_1^0. The scenario considered in this simulations has two users, $n_U = 2$.
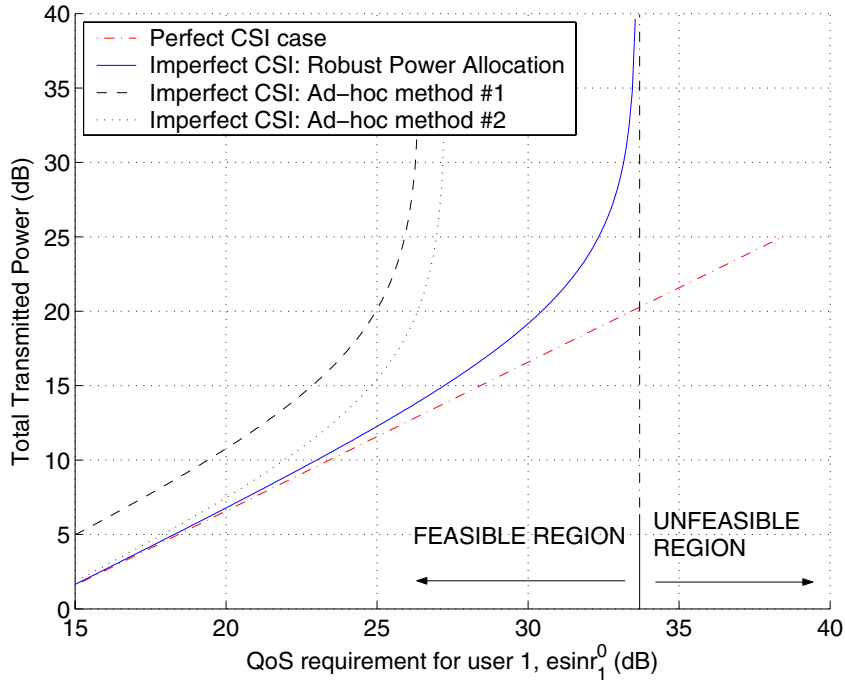
Figure 6.10: Total transmitted power versus QoS requirement for user 1, esinr_1^0. In this scenario we considered eight users, $n_U = 8$. Note how the goodness of our proposed method with respect to the two ad-hoc approaches has increased in relation to the scenario with two users depicted in Figure 6.9.

6.9 Chapter summary and conclusions

In this chapter, a multi-antenna downlink multi-user system has been considered, where the power allocation among the data streams of different users has been designed in a robust way against uncertainties and errors in the available CSI. The robustness has been formulated under a worst-case framework, where the objective has been the minimization of the total transmitted power while still guaranteeing a minimum QoS (in terms of a minimum SINR) per user for any possible channel realization within the so called uncertainty region around the available CSI. This uncertainty region models the imperfections and errors in the CSI, and therefore, its shape and size are directly connected with the physical source of the errors. In the chapter, we have considered two different sources of errors and combinations of them, such as the Gaussian noise and the quantization effects. The robust power allocation design has been solved using the tools provided by convex optimization theory, obtaining closed-form solutions when possible, and, when not possible, simplified optimization problems in convex form that can be solved very efficiently with existing numerical methods. Besides, some practical aspects related to the feasibility of the problem and the starting point in case of using numerical methods have also been studied.

By means of simulations results, it has been proved that the proposed design improves

the performance achieved by other non-robust power allocation policies. In particular, it has been shown that the robust technique needs much less power than other solutions while still guaranteeing the same QoS, *i.e.*, the same quality can be achieved even requiring less energetic resources. The improvement obtained by the robust design increases as the number of users also increases, thus, it is concluded that this design is specially suitable in networks where many users are coexisting in the same region.

6.A   Proofs of equivalence of convex problems

6.A.1   Proof of Proposition 6.6.1

We first need to particularize the general convex problem in (6.15) for the case where

$$\mathcal{R} = \mathcal{S}^{\text{ind}}_{\{R_i^2\}} \triangleq \left\{ \boldsymbol{\Delta} \in \mathbb{C}^{n_U \times n_T} \mid \|\boldsymbol{\delta}_i\|^2 = \boldsymbol{\delta}_i^H \boldsymbol{\delta}_i \leq R_i^2, \, \forall i \right\}. \tag{6.31}$$

Since in the definition of the uncertainty region, there are independent restrictions for each row $\boldsymbol{\delta}_i^H$ of the uncertainty matrix $\boldsymbol{\Delta}$, the supremum in the restriction of the problem in (6.14) particularizes to

$$\sup_{\boldsymbol{\Delta} \in \mathcal{S}^{\text{ind}}_{\{R_i^2\}}} \boldsymbol{\delta}_i^H \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \boldsymbol{\delta}_i \quad \triangleq \quad \begin{array}{ll} \sup\limits_{\boldsymbol{\delta}_i} & \boldsymbol{\delta}_i^H \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \boldsymbol{\delta}_i, \\ \text{s.t.} & \boldsymbol{\delta}_i^H \boldsymbol{\delta}_i \leq R_i^2, \end{array} \tag{6.32}$$

whose solution is well known to be given by $\boldsymbol{\delta}_i$ being proportional to the eigenvector associated with the maximum eigenvalue of the matrix $\mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H$ and such that $\boldsymbol{\delta}_i^H \boldsymbol{\delta}_i = R_i^2$. In this case, the supremum in (6.32) particularizes to $R_i^2 \lambda_{\max}(\mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H)$, and the convex problem in (6.14) becomes

$$\begin{array}{ll} \min\limits_{\mathbf{P}} & \text{Tr}\, \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H, \\ \text{s.t.} & R_i^2 \lambda_{\max}(\mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H) - \gamma_i^0 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \\ & p_i \geq 0, \quad 1 \leq i \leq n_U. \end{array} \tag{6.33}$$

It is straightforward to prove that the solution $\mathbf{P}^\star$ to the problem in (6.33) has to fulfill the restrictions with equality. Then, defining $\mu = \lambda_{\max}(\mathbf{B}_{\text{ZF}} \mathbf{P}^\star \mathbf{B}_{\text{ZF}}^H) \geq 0$, we obtain that the optimal power allocation must fulfill

$$p_i^\star = \frac{R_i^2 \mu + \sigma^2}{\gamma_i^0}, \tag{6.34}$$

where μ has to be determined. First of all, we utilize the definitions of the diagonal matrices $[\boldsymbol{\Gamma}]_{ii} \triangleq \frac{R_i^2}{\gamma_i^0}$ and $[\boldsymbol{\Sigma}]_{ii} \triangleq \frac{\sigma^2}{\gamma_i^0}$ to express $\mathbf{P}^\star = \boldsymbol{\Gamma} \mu + \boldsymbol{\Sigma}$. Then, utilizing these definitions, from the restriction in (6.33) for the optimal power allocation $\mathbf{P}^\star$ we obtain the following equation for the μ parameter

$$\lambda_{\max}(\boldsymbol{\Gamma}_{\mathbf{B}} \mu + \boldsymbol{\Sigma}_{\mathbf{B}}) = \mu, \quad \text{with } \boldsymbol{\Gamma}_{\mathbf{B}} \triangleq \mathbf{B}_{\text{ZF}} \boldsymbol{\Gamma} \mathbf{B}_{\text{ZF}}^H, \ \boldsymbol{\Sigma}_{\mathbf{B}} \triangleq \mathbf{B}_{\text{ZF}} \boldsymbol{\Sigma} \mathbf{B}_{\text{ZF}}^H \tag{6.35}$$

where both sides are convex functions of μ.

We now find under which conditions, equation (6.35) has a solution (and how many). Since $\boldsymbol{\Gamma}_{\mathbf{B}}$ and $\boldsymbol{\Sigma}_{\mathbf{B}}$ are positive definite matrices, then

$$\lambda_{\max}(\boldsymbol{\Gamma}_{\mathbf{B}} \mu + \boldsymbol{\Sigma}_{\mathbf{B}}) \leq \lambda_{\max}(\boldsymbol{\Gamma}_{\mathbf{B}})\mu + \lambda_{\max}(\boldsymbol{\Sigma}_{\mathbf{B}}), \tag{6.36}$$

$$\lambda_{\max}(\boldsymbol{\Gamma}_{\mathbf{B}} \mu + \boldsymbol{\Sigma}_{\mathbf{B}}) > \lambda_{\max}(\boldsymbol{\Gamma}_{\mathbf{B}})\mu. \tag{6.37}$$

Last inequality clearly implies that if $\lambda_{\max}(\mathbf{\Gamma_B}) \geq 1$ then (6.35) has no solution because no intersection in the fixed point equation (6.35) is possible, as for all $\mu \geq 0$

$$\lambda_{\max}(\mathbf{\Gamma_B}) \geq 1 \Rightarrow \lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B}) > \lambda_{\max}(\mathbf{\Gamma_B})\mu \geq \mu. \tag{6.38}$$

On the contrary, if $\lambda_{\max}(\mathbf{\Gamma_B}) < 1$, then

$$\lambda_{\max}(\mathbf{\Gamma_B}) < 1 \Rightarrow \exists \mu \geq 0 \mid \mu > \lambda_{\max}(\mathbf{\Gamma_B})\mu + \lambda_{\max}(\mathbf{\Sigma_B}) \geq \lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B})$$
$$\mu = 0 \Rightarrow \lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B}) > \mu, \tag{6.39}$$

which implies that there exists a value of μ such that $\mu > \lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B})$, and another value such that $\lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B}) > \mu$; from continuity this implies that (6.35) must have a solution. From the fact that $\lambda_{\max}(\mathbf{\Gamma_B}\mu + \mathbf{\Sigma_B})$ is a monotonically increasing and convex function only one solution can exist.

6.A.2   Proof of Proposition 6.6.2

We first particularize the general convex problem in (6.15) for the case where

$$\mathcal{R} = \mathcal{E}^{\text{ind}}_{\{R_i^2\},\{\mathbf{C}_i\}} \triangleq \left\{ \mathbf{\Delta} \in \mathbb{C}^{n_U \times n_T} \mid \boldsymbol{\delta}_i^H \mathbf{C}_i^{-1} \boldsymbol{\delta}_i \leq R_i^2, \, \forall i \right\}. \tag{6.40}$$

Clearly, the supremum in the restriction of the problem in (6.14) particularizes to

$$\sup_{\mathbf{\Delta} \in \mathcal{E}^{\text{ind}}_{\{R_i^2\},\{\mathbf{C}_i\}}} \boldsymbol{\delta}_i^H \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \boldsymbol{\delta}_i, \quad \triangleq \quad \begin{aligned} &\sup_{\boldsymbol{\delta}_i} \quad \boldsymbol{\delta}_i^H \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \boldsymbol{\delta}_i, \\ &\text{s.t.} \quad \boldsymbol{\delta}_i^H \mathbf{C}_i^{-1} \boldsymbol{\delta}_i \leq R_i^2, \end{aligned} \tag{6.41}$$

whose solution can be obtained by performing the change $\tilde{\boldsymbol{\delta}}_i = \mathbf{C}_i^{-1/2} \boldsymbol{\delta}_i$ and it is well known to be given by $\tilde{\boldsymbol{\delta}}_i$ being proportional to the eigenvector associated with the maximum eigenvalue of the matrix $\mathbf{C}_i^{1/2} \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \mathbf{C}_i^{1/2}$ and such that $\tilde{\boldsymbol{\delta}}_i^H \tilde{\boldsymbol{\delta}}_i = R_i^2$. In this case, the supremum in (6.32) particularizes to $R_i^2 \lambda_{\max}(\mathbf{C}_i^{1/2} \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \mathbf{C}_i^{1/2})$, and the convex problem in (6.14) becomes

$$\begin{aligned} &\min_{\mathbf{P}} \quad \text{Tr}\, \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H, \\ &\text{s.t.} \quad R_i^2 \lambda_{\max}(\mathbf{C}_i^{1/2} \mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H \mathbf{C}_i^{1/2}) - \gamma_i^0 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \\ &\qquad p_i \geq 0, \quad 1 \leq i \leq n_U. \end{aligned} \tag{6.42}$$

6.A.3   Proof of Proposition 6.6.3

In this case, the uncertainty region $\mathcal{Q}_V$ is the polyhedron (possibly non-convex) whose vertices are given by the set $V = \{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_M\} \subset \mathbb{C}^{n_U \times n_T}$.

The convex hull of the set of points V, denoted by $\text{conv}\, V$, is the minimal convex set containing V, [Boy04], which clearly implies that the polyhedron $\mathcal{Q}_V$ is inside the convex hull of V,

i.e., $\mathcal{Q}_V \subseteq \operatorname{conv} V$. In [Boy04], it is shown that the convex hull of V is

$$\operatorname{conv} V = \left\{ \sum\nolimits_{m=1}^{M} \theta_m \mathbf{V}_m \mid \sum\nolimits_{m=1}^{M} \theta_m = 1, \ \theta_m \geq 0, \ \forall m \right\}. \tag{6.43}$$

We define $f(\boldsymbol{\Delta}) \triangleq \boldsymbol{\delta}_i^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H \boldsymbol{\delta}_i$, which is a convex function in $\boldsymbol{\Delta}$ (and consequently also on $\boldsymbol{\delta}_i$). The supremum in the restriction in (6.15) particularizes to $\sup_{\boldsymbol{\Delta} \in \mathcal{Q}_V} f(\boldsymbol{\Delta})$ which can then be bounded as,

$$\sup_{\boldsymbol{\Delta} \in \mathcal{Q}_V} f(\boldsymbol{\Delta}) \leq \sup_{\boldsymbol{\Delta} \in \operatorname{conv} V} f(\boldsymbol{\Delta}), \tag{6.44}$$

where $\mathcal{Q}_V \subseteq \operatorname{conv} V$ has been used. From (6.43), for any $\boldsymbol{\Delta} \in \operatorname{conv} V$, it exists a set $\{\theta_1, \ldots, \theta_M : \sum_{m=1}^{M} \theta_m = 1, \ \theta_m \geq 0, \forall m\}$ such that

$$\boldsymbol{\Delta} \in \operatorname{conv} V \Rightarrow f(\boldsymbol{\Delta}) = f\left(\sum_{m=1}^{M} \theta_m \mathbf{V}_m \right) \leq \sum_{m=1}^{M} \theta_m f(\mathbf{V}_m) \leq \max_m f(\mathbf{V}_m), \tag{6.45}$$

from which we can deduce that $\sup_{\boldsymbol{\Delta} \in \operatorname{conv} V} f(\boldsymbol{\Delta}) \leq \max_m f(\mathbf{V}_m)$. Since $\mathbf{V}_m \in \mathcal{Q}_V$, this implies that $\sup_{\boldsymbol{\Delta} \in \mathcal{Q}_V} f(\boldsymbol{\Delta}) = \max_m f(\mathbf{V}_m)$, which means that the supremum of $f(\boldsymbol{\Delta})$, with $\boldsymbol{\Delta} \in \mathcal{Q}_V$, has to be necessarily placed in one of the vertices of the polyhedron (a priori we do not know which one, though). Consequently, the supremum operation $\sup_{\boldsymbol{\Delta} \in \mathcal{Q}_V} f(\boldsymbol{\Delta})$ can be substituted by a simple list of the function $f(\boldsymbol{\Delta})$ evaluated at all the different vertices since necessarily one of them has to be the supremum (and if the restriction is met for the supremum is also met for all other values of $\boldsymbol{\Delta} \in \mathcal{Q}_V$). The following two optimization problems are thus equivalent:

$$\begin{aligned}
\min_{\mathbf{P}} \quad & \operatorname{Tr} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H, \\
\mathrm{s.\,t.} \quad & f(\mathbf{V}_m) - \gamma_i^0 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \quad 1 \leq m \leq M, \\
& p_i \geq 0, \quad 1 \leq i \leq n_U,
\end{aligned} \tag{6.46}$$

and

$$\begin{aligned}
\min_{\mathbf{P}} \quad & \operatorname{Tr} \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H, \\
\mathrm{s.\,t.} \quad & \mathbf{v}_{m,i}^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H \mathbf{v}_{m,i} - \gamma_i^0 p_i + \sigma^2 \leq 0, \quad 1 \leq i \leq n_U, \quad 1 \leq m \leq M, \\
& p_i \geq 0, \quad 1 \leq i \leq n_U,
\end{aligned} \tag{6.47}$$

where $\mathbf{v}_{m,i}^H$ is the i-th row of the vertex $\mathbf{V}_m$.

6.A.4 Proof of Proposition 6.6.4

In this section we have to consider the case where $\boldsymbol{\Delta} = \mathbf{S} + \mathbf{Q}$, with $\mathbf{S} \in \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}}$ and $\mathbf{Q} \in \mathcal{Q}_V$. Consequently, in this case we have $\boldsymbol{\delta}_i = \mathbf{s}_i + \mathbf{q}_i$ and the particularization of the restriction of the problem in (6.15) for this uncertainty region becomes

$$\sup_{\mathbf{S} \in \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}}, \mathbf{Q} \in \mathcal{Q}_V} (\mathbf{s}_i + \mathbf{q}_i)^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{q}_i) \triangleq \sup_{\mathbf{S} \in \mathcal{S}_{\{R_i^2\}}^{\mathrm{ind}}} \sup_{\mathbf{Q} \in \mathcal{Q}_V} (\mathbf{s}_i + \mathbf{q}_i)^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{q}_i)$$

$$\tag{6.48}$$

where we have utilized the fact that any maximization problem $\sup_{\mathbf{x},\mathbf{y}} f(\mathbf{x},\mathbf{y})$ can be decomposed as $\sup_{\mathbf{x}} \sup_{\mathbf{y}} f(\mathbf{x},\mathbf{y})$, [Boy04]. Now, we note that $(\mathbf{s}_i + \mathbf{q}_i)^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{q}_i)$ is a convex function of $\mathbf{q}_i$ and, similarly as we have done in the previous section, the solution to the inner maximization is given by the function $(\mathbf{s}_i + \mathbf{q}_i)^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{q}_i)$ evaluated at one of the vertices that define the convex hull $\{\mathbf{V}_m\}$. From the original problem in (6.48), we obtain a set of maximization problems indexed by the variable m that represent the vertices index, as

$$\sup_{\mathbf{S} \in \mathcal{S}^{\mathrm{ind}}_{\{R_i^2\}}} (\mathbf{s}_i + \mathbf{v}_{m,i})^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{v}_{m,i}) \triangleq \sup_{\mathbf{s}_i^H \mathbf{s}_i \leq R_i^2} (\mathbf{s}_i + \mathbf{v}_{m,i})^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H (\mathbf{s}_i + \mathbf{v}_{m,i}), \quad (6.49)$$

where we have utilized the definition of the hyper-spherical uncertainty region $\mathcal{S}^{\mathrm{ind}}_{\{R_i^2\}}$. The problem expressed in (6.49) is solved in Appendix 6.B, with $\tilde{\mathbf{x}} = \mathbf{s}_i$, $\tilde{\mathbf{y}} = \mathbf{v}_{m,i}$, $\mathbf{A} = \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H$, and $b = R_i^2$ and its solution is denoted by $\mathbf{s}_{i,m}^\star(\mathbf{P})$ which substituted in the maximization problem in (6.49) becomes

$$\left(\mathbf{s}_{m,i}^\star(\mathbf{P}) + \mathbf{v}_{m,i}\right)^H \mathbf{B}_{\mathrm{ZF}} \mathbf{P} \mathbf{B}_{\mathrm{ZF}}^H \left(\mathbf{s}_{m,i}^\star(\mathbf{P}) + \mathbf{v}_{m,i}\right), \quad (6.50)$$

which completes the proof.

6.B Maximization of a general quadratic form with a norm constraint

Let $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ be vectors in the field $\mathbb{C}^{n_T \times 1}$ and let $\mathbf{A} \in \mathbb{C}^{n_T \times n_T}$ be a positive semi-definite matrix with $n_U \leq n_T$ non-zero eigenvalues. We want to solve the following non-convex optimization problem

$$\begin{aligned} \underset{\tilde{\mathbf{x}}}{\text{maximize}} \quad & (\tilde{\mathbf{x}} + \tilde{\mathbf{y}})^H \mathbf{A} (\tilde{\mathbf{x}} + \tilde{\mathbf{y}}) \\ \text{subject to} \quad & \tilde{\mathbf{x}}^H \tilde{\mathbf{x}} \leq b, \end{aligned} \quad (6.51)$$

with b being a positive real value. Performing the SVD decomposition of the positive semi-definite $\mathbf{A}$ matrix we obtain $\mathbf{A} = \mathbf{U} \mathbf{\Omega} \mathbf{U}^H$, with $\mathbf{\Omega} \in \mathbb{C}^{n_T \times n_T}$ being a diagonal matrix with non-negative diagonal elements sorted in decreasing order, and with $\mathbf{U} \in \mathbb{C}^{n_T \times n_T}$ being a unitary matrix. Introducing the changes $\mathbf{x}' = \mathbf{U}^H \tilde{\mathbf{x}}$ and $\mathbf{y}' = \mathbf{U}^H \tilde{\mathbf{y}}$ the problem in (6.51) becomes

$$\begin{aligned} \underset{\{x_i'\}_{i=1}^{n_T}}{\text{maximize}} \quad & \sum_{i=1}^{n_U} |x_i' + y_i'|^2 \omega_i \\ \text{subject to} \quad & \sum_{i=1}^{n_T} |x_i'|^2 \leq b, \end{aligned} \quad (6.52)$$

where $\omega_i = [\mathbf{\Omega}]_{ii} > 0$ for $i \in [1, n_U]$, and where the remaining $\omega_i = 0$ for $i \in [n_U + 1, n_T]$ have been discarded from the summation in the objective function in (6.52). Then clearly the optimal

solution $\mathbf{x}'^{\star}$ fulfills $x_i'^{\star} = 0$ for $i \in [n_U + 1, n_T]$. Consequently, we have reduced the search space from dimension n_T to dimension n_U.

Noting that the inequality $|x_i' + y_i'|^2 \leq ||x_i'| + |y_i'||^2$ becomes an equality if the complex phase of x_i' is chosen equal to the phase of y_i', then we can state that $x_i'^{\star} = \angle y_i'$ and we only need to specify the complex magnitude of the elements of the solution $\mathbf{x}'^{\star}$. Once the complex phase of the solution has been established, the problem can be further simplified to a real optimization problem by defining $x_i = |x_i'|$ and $y_i = |y_i'|$, and then the problem in (6.52) becomes

$$
\begin{aligned}
\underset{\{x_i\}_{i=1}^{n_U}}{\text{maximize}} \quad & \sum_{i=1}^{n_U}(x_i + y_i)^2 \omega_i \\
\text{subject to} \quad & \sum_{i=1}^{n_U} x_i^2 \leq b, \\
& x_i \geq 0, \;\; \forall i
\end{aligned}
\tag{6.53}
$$

where all the parameters $\{y_i\}$ and b are real non-negative numbers. We now separate the terms in the summation in (6.53) in two different groups, depending on whether its corresponding y_i is zero or not. Thus, we define the partitions $\mathcal{I}_1$ and $\mathcal{I}_2$ as the sets of indices such that

$$
i \in \mathcal{I}_1 \subset [1, n_U] \Leftrightarrow y_i > 0, \tag{6.54}
$$

$$
i \in \mathcal{I}_2 \subset [1, n_U] \Leftrightarrow y_i = 0, \tag{6.55}
$$

and that $\mathcal{I}_1 \cup \mathcal{I}_2 = [1, n_U]$. Finally, defining $z_i = (x_i + y_i)^2$, for all i, we obtain a convex optimization problem, which is equivalent to that in (6.53), as

$$
\begin{aligned}
\underset{\{z_i\}_{i=1}^{n_U}}{\text{minimize}} \quad & -\sum_{i \in \mathcal{I}_1} z_i \omega_i - \sum_{i \in \mathcal{I}_2} z_i \omega_i \\
\text{subject to} \quad & \sum_{i \in \mathcal{I}_1}\left(\sqrt{z_i} - y_i\right)^2 + \sum_{i \in \mathcal{I}_2} z_i - b \leq 0, \\
& y_i^2 - z_i \leq 0, \quad \forall i.
\end{aligned}
\tag{6.56}
$$

Note that we can assume w.l.o.g. that for any two indices $k, l \in \mathcal{I}_2$ then

$$
k \neq l \Rightarrow \omega_k \neq \omega_l, \quad \forall k, l \in \mathcal{I}_2 \tag{6.57}
$$

This can be assumed because in case $\omega_k = \omega_l$ for some different $k, l \in \mathcal{I}_2$, then we can always define a new variable $z_{kl} = z_k + z_l$ such that the equivalent problem has the same structure as the original one and is one dimension smaller.

Furthermore, we can also assume w.l.o.g. that

$$
k \in \mathcal{I}_1, \; l \in \mathcal{I}_2 \Rightarrow \omega_k \neq \omega_l. \tag{6.58}
$$

In this case, the proof follows from a primal decomposition of the original problem in (6.56) with $\omega_l = \omega_k$ for some $k \in \mathcal{I}_1$ and $l \in \mathcal{I}_2$. The primal decomposition is performed by separating the

terms with indices k and l in the objective function and by adding the auxiliary variable $c \geq 0$ that allows us to decouple the first restriction in (6.56) for the indices k and l.

$$
\begin{aligned}
\underset{\{z_i\}_{i=1}^{n_U} \backslash \{z_k, z_l\}, c \geq 0}{\text{minimize}} \quad & -\sum_{i \in \mathcal{I}_1 \backslash \{k\}} z_i \omega_i - \sum_{i \in \mathcal{I}_2 \backslash \{l\}} z_i \omega_i + \begin{cases} \underset{z_k, z_l}{\text{minimize}} \quad -\omega_k z_k - \omega_k z_l \\ \text{subject to} \quad (\sqrt{z_k} - y_k)^2 + z_l - c \leq 0, \end{cases} \\
\text{subject to} \quad & \sum_{i \in \mathcal{I}_1 \backslash \{k\}} (\sqrt{z_i} - y_i)^2 + \sum_{i \in \mathcal{I}_2 \backslash \{l\}} z_i + c - b \leq 0, \\
& y_i^2 - z_i \leq 0, \quad \forall i.
\end{aligned}
$$

(6.59)

It can be shown that the solution to the inner minimization problem yields $z_l^\star = 0$ for all possible values of c, y_k, and ω_k which implies that in case $\omega_k = \omega_l$ for some $k \in \mathcal{I}_1$ and $l \in \mathcal{I}_2$ the term corresponding to the index l can be eliminated w.l.o.g.

From all that has been said, it follows that the multiplicity of ω_i is only possible among indices inside the set $\mathcal{I}_1$ but not inside the set $\mathcal{I}_2$ or between one element of $\mathcal{I}_1$ and one of $\mathcal{I}_2$. Because in the latter cases an equivalent problem can be obtained where this multiplicity does not exist.

From the KKT conditions of the optimization problem in (6.56), we readily obtain that the solution $\{z_i^\star\}$ must fulfill:

$$
-\omega_i + \lambda^\star \left(1 - \frac{y_i}{\sqrt{z_i^\star}}\right) - \mu_i^\star = 0, \quad \forall i \in \mathcal{I}_1, \tag{6.60}
$$

$$
-\omega_i + \lambda^\star - \mu_i^\star = 0, \quad \forall i \in \mathcal{I}_2, \tag{6.61}
$$

$$
\lambda^\star \left(\sum_{i \in \mathcal{I}_1} \left(\sqrt{z_i^\star} - y_i\right)^2 + \sum_{i \in \mathcal{I}_2} z_i^\star - b\right) = 0, \tag{6.62}
$$

$$
\mu_i^\star \left(y_i^2 - z_i^\star\right) = 0, \quad \forall i \in \mathcal{I}_1, \tag{6.63}
$$

$$
-\mu_i^\star z_i^\star = 0, \quad \forall i \in \mathcal{I}_2, \tag{6.64}
$$

$$
\mu_i^\star \geq 0, \quad \forall i, \tag{6.65}
$$

$$
\lambda^\star \geq 0. \tag{6.66}
$$

In the following, we deduce that $\mu_i^\star = 0$, $\forall i \in \mathcal{I}_1$. Note that from the last restriction in (6.56), either $z_i^\star = y_i^2$ or $z_i^\star > y_i^2$ must hold. If $z_i^\star > y_i^2$, then from (6.63) we obtain $\mu_i^\star = 0$. On the contrary, if $z_i^\star = y_i^2$ then (6.60) becomes $-\omega_i - \mu_i^\star = 0$. The inequalities $\omega_i > 0$ (by definition) and $\mu_i^\star \geq 0$ (from (6.65)) imply that $-\omega_i - \mu_i^\star < 0$ which is a contradiction thus $z_i^\star = y_i^2$ is not a possible solution. Consequently, for $i \in \mathcal{I}_1$ we obtain $z_i^\star > y_i^2$ and $\mu_i^\star = 0$. From (6.60) with $\mu_i^\star = 0$, we obtain the solution to (6.56) as

$$
z_i^\star = \left(\frac{y_i \lambda^\star}{\lambda^\star - \omega_i}\right)^2, \quad \forall i \in \mathcal{I}_1, \tag{6.67}
$$

where $\lambda^\star$ has yet to be determined.

We now focus on the set of indices $\mathcal{I}_2$. Note that it is impossible that more than one optimal $z_{i'}^\star$ be greater than zero with $i' \in \mathcal{I}_2$. If we assume that $z_{i'}^\star > 0$ and $z_{j'}^\star > 0$, from (6.63) it would imply that $\mu_{i'}^\star = 0$ and $\mu_{j'}^\star = 0$ which in its turn would mean that $\lambda^\star = \omega_{i'}$ and $\lambda^\star = \omega_{j'}$ which is impossible, because, as expressed in (6.57), $\omega_{i'} \neq \omega_{j'}, \forall i', j' \in \mathcal{I}_2$, and $\lambda^\star$ can only take one value. Consequently, at most there exists one index $i' \in \mathcal{I}_2$ such that $z_{i'}^\star > 0$. Obviously, in case this index exists it would correspond to the biggest $\omega_{i'}$, *i.e.*, $i'_{\max} = \arg\max_{i \in \mathcal{I}_2} \omega_i$.

Then only two cases need to be considered.

1. We first consider that $z_i^\star = 0, \forall i \in \mathcal{I}_2$. Then, the optimal solution is completed with (6.67), where $\lambda^\star$ is determined from the restriction in (6.56) similarly as in [Lor05], as the greatest solution to the equation

$$\sum_i \left(\frac{y_i \omega_i}{\lambda^\star - \omega_i} \right)^2 - b = 0. \tag{6.68}$$

2. We now consider that there exists an index $i'_{\max} \in \mathcal{I}_2$ such that $z_{i'_{\max}}^\star > 0$. Then, from (6.64), $\mu_{i'_{\max}}^\star = 0$ and thus $\lambda^\star = \omega_{i'_{\max}}$ as indicated by (6.61). Plugging this value for $\lambda^\star$ in (6.67) we obtain the solution

$$z_i^\star = \left(\frac{y_i \omega_{i'_{\max}}}{\omega_{i'_{\max}} - \omega_i} \right)^2, \quad \forall i \in \mathcal{I}_1, \tag{6.69}$$

where the denominator is always different of zero as $i \in \mathcal{I}_1$ and $i'_{\max} \in \mathcal{I}_2$ which implies that $\omega_i \neq \omega_{i'_{\max}}$. The solution is completed with $z_{i'_{\max}}^\star$, which is determined such that the power constraint is fulfilled with equality as

$$z_{i'_{\max}}^\star = b - \sum_{i \in \mathcal{I}_1} \left(\sqrt{z_i^\star} - y_i \right)^2. \tag{6.70}$$

To determine which of the two cases is the optimal one, we only need to calculate (6.69) and then check the sign of $b - \sum_{i \in \mathcal{I}_1} \left(\sqrt{z_i^\star} - y_i \right)^2$. In case it is negative then (6.70) becomes meaningless because $z_{i'}^\star$ has to be positive or zero and thus the solution is given by the first case. Otherwise the solution is given by the second case.

Once the solution to the convex problem in (6.56), $\{z_i^\star\}_{i=1}^{n_U}$, is determined, we need simply to construct the solution $\tilde{\mathbf{x}}^\star$ of the original problem in (6.51) as

$$\tilde{\mathbf{x}}^\star = \mathbf{U}\mathbf{x}'^\star, \quad \text{with} \begin{cases} x_i'^\star = (\sqrt{z_i^\star} - y_i) \cdot \exp\left(\mathrm{j}\angle\left(y_i'\right)\right) & i \in [1, n_U] \\ x_i'^\star = 0 & i \in [n_U + 1, n_T] \end{cases}. \tag{6.71}$$

Chapter 7

Conclusions and future work

7.1 Conclusions

In general terms, this dissertation has dealt with the design and analysis of multi-antenna communication systems, mainly focusing the attention on transmission architectures. Precisely, the impact of the availability and quality of the CSI on the capacity and on the transmitter design has been studied and robust transmission architectures have been designed in such a way that their sensitivity to the presence of errors in the CSI is reduced with respect to their classical non-robust counterparts.

In chapter 1 we have presented the motivation, the outline, and the contributions of this dissertation and in chapter 2 we have reviewed some basic concepts and the state-of-the-art concerning multi-antenna communication systems.

Chapter 3 has dealt with the design of a linear transmitter in a single-user MIMO system where the CSI is considered to be perfect at both communication ends. The design criterion for the transmitter has been such that the minimum distance among the received symbols has been maximized. We have optimally characterized the left singular vectors, and the singular values of the optimal transmitter matrix and have given a simple algorithm to obtain a good approximation to the right singular vectors. We have also found that our proposed scheme yields an excellent performance when compared to well-known schemes such as V-BLAST, OSTBC schemes, the minimum BER optimum linear design, or the maximum minimum SNR eigenvalue design. Unfortunately, we have also shown that, for the general case, the computational complexity of our proposed scheme is rather high as it involves a search over a big set of matrices, which makes it difficult to implement in fast-fading environments. However, for fixed fading environments such that the transmitter calculation can can be performed off-line we have shown that it presents an excellent performance. In addition, for the particular case of transmitting two QPSK streams, we have presented the design in closed form, which reduces enormously the

computational complexity as the search over the set of matrices is not necessary since an explicit expression for it is available. Additionally, it has been shown that the linear transmitter design is directly related to optimal symbol constellation construction.

Chapter 4 has introduced incompleteness of the CSI at the transmitter in single-user MIMO communication systems. Specifically, it has been considered that the transmitter is informed (or is able to estimate) only the magnitude of the channel complex coefficients. For this particular choice of incomplete CSI, we have shown that a diagonal covariance matrix is sufficient and necessary to achieve capacity in both the ergodic and compound formulations. Moreover, for some configurations with small number of antennas we have been able to find closed form expressions for the elements of the covariance matrix and for the capacity. For configurations with an arbitrary number of antennas, we have presented numerical methods to calculate the optimal values of the diagonal elements of the covariance matrix. We have also shown that these methods may be computationally too demanding. To overcome this problem, for the ergodic case we have presented a reduced complexity method based on RMT and for the compound case, based on a Taylor approximation. Next, the accuracy of the proposed approximated solutions has been tested, showing that the low complexity techniques provide nearly the same results as the high load algorithms. Finally, the ergodic and compound capacities themselves have been evaluated for different scenarios, leading to the conclusion that the optimization of the power allocation with only the knowledge of the magnitude of the channel gains allows to improve the system performance very importantly.

Chapter 5 has dealt with multi-antenna communication systems with imperfect CSI at the transmitter side. Precisely, it has studied the effects of having a noisy channel estimate in the achievable rates of a communication architecture equipped with STHP. It has been shown that, unless some kind of robust architecture is considered the rate loss can be quite high. Consequently, we have presented a robust design of this precoder both for the single- and multi-user scenarios and we have shown that, with the robust design, the rate loss in the presence of errors in the CSI is minimized. Finally, due to the versatility of STHP to encompass multi-user systems, this chapter has linked the single-user and the multi-user scenarios in multi-antenna systems.

Chapter 6 has been fully devoted to the study of a multi-antenna and multi-user communication system where it has been considered that both the transmitter and the receivers have access only to an imperfect version of the CSI. Imposing a particular architecture at the transmitter side, the power allocation among the signals to be transmitted to the users has been robustly designed with the criterion of minimizing the total transmitted power, while guaranteeing a certain predefined quality of service per user. The problem has been formulated within the powerful framework of convex optimization, which enabled the designer to efficiently find the solution to the robust transmitter. It has also been proved that the proposed design improves the perfor-

mance achieved by other non-robust power allocation policies. In particular, it has been shown that the robust technique needs much less power than other solutions while still guaranteeing the same QoS, *i.e.*, the same quality can be achieved even requiring less energetic resources. The improvement obtained by the robust design increases as the number of users also increases, thus, it is concluded that this design is specially suitable in networks where many users are coexisting in the same region.

As a conclusion to this dissertation we would like to highlight that the imperfections in or incompleteness of the CSI can decrease significatively the achievable performance of multi-antenna communication systems, unless robust architectures are properly designed to overcome this situation.

7.2   Future work

There exist many possibilities for future work that may extend the results obtained in this dissertation.

Concerning the setup of a MIMO channel with a single-user with perfect CSI some issues are still open:

- Find the linear transmitter that maximizes the capacity for an arbitrary distribution of the input symbols and an arbitrary structure of the MIMO channel.

- Further characterize the optimal structure of the constellation composition matrix $\mathbf{V}$, for the case of designing a linear transmitter when the receiver performs ML detection of the received symbols.

For the cases where we considered imperfect and incomplete CSI it be interesting to further research on

- Relating the uncertainty models utilized in the chapters that deal with imperfect or incomplete CSI with the algorithms that are utilized for channel classification as in [Bou06].

- Including additional types of partial and incomplete CSI. For example in the case of magnitude knowledge and phase uncertainty it would be worth trying to characterize the effects where the magnitude knowledge is not perfect either.

Referring to the robust multi-user design in chapter 6 it would be interesting to

- Solve the problem that arises when different estimates of the channel matrix are available at the two sides of the communication link.

- Generalize the results by allowing the design of the full transmission matrix, not limiting the optimization procedure to the power allocation.

- Compare the performance of the proposed transmitter with the broadcast capacity region with imperfect CSI.

Bibliography

[Ala98] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[AS05] A. Abdel-Samad, T. N. Davidson, and A. Gershman, "Robust transmit eigen-beamforming based on imperfect channel state information", *IEEE Trans. on Signal Processing (to appear)*, May 2005.

[Bah05] S. Bahng, J. Liu, A. Høst-Madsen, and X. Wang, "The effects of channel estimation on Tomlinson-Harashima precoding in TDD MIMO systems", *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC'05)*, New York, NJ (USA), June 2005.

[Ben00] M. Bengtsson, "Robust and constrained downlink beamforming", *Proc. European Signal Processing Conference*, pp. 1433–1436, Sep. 2000.

[Ben01] M. Bengtsson, and B. Ottersten, "Optimal and Suboptimal Transmit Beamforming", L. C. Godara (ed.), *Handbook of Antennas in Wireless Communications*, CRC Press, Aug. 2001.

[Big98] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects", *IEEE Trans. on Information Theory*, vol. 47, pp. 2619–2691, Oct. 1998.

[Big04] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust downlink power control in wireless cellular systems", *EURASIP Journal on Wireless Communications and Networking*, vol. 2004, no. 2, pp. 261–272, 2004.

[Böl02] H. Bölcskei, and A. J. Paulraj, *The Communications Handbook*, Chap. Multiple-input multiple-output (MIMO) wireless systems, pp. 90.1–90.14, J. Gibson, Ed. CRC Press, $2^{\underline{nd}}$ ed., 2002.

[Bla87] R. E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, 1987.

[Boc05] H. Boche, M. Schubert, and S. Stańczak, "A unifying approach to multiuser receiver design under QoS constraints", *Proc. IEEE International Symposium on Information Theory*, Sep. 2005.

[Bou06] J. J. Boutros, F. Kharrat-Kammoun, and H. Randriambololona, "On the capacity of MIMO systems with magnitude knowledge and phase uncertainty", *Proc. IEEE International Zurich Seminar on Communications (IZS'06)*, Feb. 2006.

[Boy04] S. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[Bra74] L. H. Brandenburg, and A. D. Wyner, "Capacity of the Gaussian channel with memory: the multivariate case", *The Bell System Technical Journal*, vol. 53, no. 5, pp. 745–778, May-June 1974.

[Cai03] G. Caire, and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel", *IEEE Trans. on Information Theory*, vol. 43, pp. 1691–1706, July 2003.

[Cho02a] K. Cho, and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations", *IEEE Trans. on Communications*, vol. 50, no. 7, pp. 1074–1080, July 2002.

[Cho02b] J. Choi, "Performance analysis for transmit antenna diversity with/without channel information", *IEEE Trans. on Vehicular Technology*, vol. 51, no. 1, pp. 101–113, Jan. 2002.

[Cho02c] J. Choi, "Performance limitation of closed-loop transmit antenna diversity over fast Rayleigh fading channels", *IEEE Trans. on Vehicular Technology*, vol. 51, no. 4, pp. 771–775, July 2002.

[Cho02d] J. Choi, H. K. Choi, and H. W. Lee, "An adaptive technique for transmit antenna diversity with feedback", *IEEE Trans. on Vehicular Technology*, vol. 51, no. 4, pp. 617–623, July 2002.

[Chu01] S. T. Chung, A. Lozano, and H. C. Huang, "Approaching eigenmode BLAST channel capacity using BLAST with rate and power feedback", *Proc. Vehicular Technology Conference 2001-Fall (VTC'01)*, vol. 2, pp. 915–919, 2001.

[Col04] L. Collin, O. Berder, P. Rostaing, and G. Burel, "Optimal minimum distance-based precoder for MIMO spatial multiplexing systems", *IEEE Trans. on Signal Processing*, vol. 52, no. 3, Mar. 2004.

[Cos83] M. H. M. Costa, "Writing on dirty paper", *IEEE Transactions on Information Theory*, vol. 29, no. 5, pp. 439–441, May 1983.

[Cov91] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[Csi81] I. Csiszar, and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[Dam03] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point", *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[Die05] F. A. Dietrich, P. Breun, and W. Utschick, "Tomlinson-Harashima precoding: a continuous transition from complete to statistical channel knowledge", *Proc. IEEE Global Telecommunications Conference (GLOBECOM'05)*, vol. 4, pp. 2379–2384, Nov. 2005.

[Din02] Y. Ding, T. N. Davidson, J-K. Zhang, Z-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol. 3, pp. 2261–2264, May 2002.

[Eld04] Y. C. Eldar, and N. Merhav, "A competitive minimax approach to robust estimation of random parameters", *IEEE Trans. on Signal Processing*, vol. 52, no. 7, pp. 1931–1946, July 2004.

[Fin85] U. Fincke, and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis", *Mathematics of Computation*, vol. 44, pp. 463–471, Apr. 1985.

[Fis02a] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Space-time transmission using Tomlinson-Harashima precoding", *Proc. Proceedings of 4th ITG Conference on Source and Channel Coding*, Jan. 2002.

[Fis02b] R. Fischer, C. Windpassinger, A. Lampe, and J. Huber, "Tomlinson-Harashima precoding in space-time transmission for low-rate backward channel", *Proc. IEEE International Zurich Seminar on Broadband Communications (IZS'02)*, Feb. 2002.

[For04] G. D. Forney, Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener", *Available on-line at http://arxiv.org/abs/cs.IT/0409053*, 2004.

[Fos96] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", *Bell Labs Technical Journal*, vol. 1, no. 2, Autumn 1996.

[Fos98] G. J. Foschini, and M. J. Gans, "On limits of wireless communications in a fading en-
         virontment when using multiple antennas", *Wireless Personal Communications*, vol. 6,
         no. 3, pp. 311–335, Mar. 1998.

[Gan01] G. Ganesan, and P. Stoica, "Space-time block codes: A maximum SNR approach",
         IEEE Trans. on Information Theory, vol. 47, no. 4, May 2001.

[Gan02] G. Ganesan, *Designing space-time codes using orthogonal designs*, PhD Thesis, Uppsala
         University, 2002.

[Gol96] G. Golub, and C. Van Loan, *Matrix Computations*, Johns Hopkins, Baltimore, 1996.

[Gol97] A. J. Goldsmith, and P. P. Varaiya, "Capacity of fading channels with channel side
         information", *IEEE Trans. on Information Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[Gol99] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection
         algorithm and initial laboratory results using V-BLAST space-time communication ar-
         chitecture", *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[Gol03] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO
         channels", *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–
         702, June 2003.

[Gor00] D. Gore, R. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas
         in wireless systems", *Proc. IEEE International Conference on Acoustics, Speech, and
         Signal Processing (ICASSP'00)*, vol. 5, pp. 2785–2788, June 2000.

[Gra00] I. Gradshteyn, and I. Ryzhik, *Table of Integrals, Series and Products*, London: Academic
         Press, 2000.

[Guo05] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error
         in Gaussian channels", *IEEE Trans. on Information Theory*, vol. 51, no. 4, pp. 1261–
         1282, 2005.

[Hac05] W.   Hachem,   P.   Loubaton,   and   J.   Najim,   "Deterministic   equivalents
         for   certain   functionals   of   large   random   matrices",   *Available   on-line   at
         http://arxiv.org/abs/math.PR/0507172*, 2005.

[Hac06] W. Hachem, "Random matrices and the Stieltjes transform. MIMO channels capac-
         ity approximation", *Lecture notes of the NEWCOM Summer School on Random Ma-
         trix Theory for Wireless Communications*, Centre Tecnològic de Telecomunicacions de
         Catalunya, July 2006.

[Har72] H. Harashima, and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference", *IEEE Trans. on Communications*, vol. 20, no. 4, pp. 774–780, 1972.

[Has02] B. Hassibi, and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: thereceived signal density in closed form", *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1473–1484, June 2002.

[Has03] B. Hassibi, and H. Vikalo, *Multiantenna Channels: Capacity, Coding and Signal Processing*, Chap. Maximum-likelihood decoding and integer least-squares: The expected complexity, G. J. Foschini and S. Verdú, Eds., American Mathematical Society, 2003.

[Hoc00a] B. M. Hochwald, and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading", *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[Hoc00b] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations", *IEEE Trans. on Information Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000.

[Hoc00c] B. M. Hochwald, and W. Sweldens, "Differential unitary space-time modulation", *IEEE Trans. on Communications*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[Hoc04] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling", *IEEE Trans. on Information Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[Hor85] R. A. Horn, and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[Hun04] R. Hunger, F. A. Dietrich, M. Joham, and W. Utschick, "Robust transmit zero-forcing filters", *Proc. of the ITG Workshop on Smart Antennas*, Munich (Germany), Mar. 2004.

[Jaf01a] S. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback", *Proc. IEEE International Conference on Communications (ICC'01)*, vol. 7, pp. 2266–2270, Jun. 2001.

[Jaf01b] H. Jafharkani, "A quasi-orthogonal space-time block code", *IEEE Trans. on Communications*, vol. 49, no. 1, pp. 1–4, Jan. 2001.

[Jak74] W. C. Jakes, Jr., *Microwave Mobile Communications*, New York: John Wiley & Sons, Inc., 1974.

[Jal05]   J. Jaldén, B. Ottersten, and W.-K. Ma, "Reducing the average complexity of ML detec-
          tion using semidefinite relaxation", *Proc. IEEE International Conference on Acoustics,
          Speech, and Signal Processing (ICASSP'05)*, Mar. 2005.

[Jia05]   Y. Jiang, and J. Li, "Adaptable channel decomposition for MIMO communications",
          *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing
          (ICASSP'05)*, vol. 4, Mar. 2005.

[Jön02]   G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal
          space-time block coding", *IEEE Trans. on Information Theory*, vol. 48, no. 3, pp. 611–
          627, Mar. 2002.

[Joh04]   M. Joham, J. Brehmer, and W. Utschick, "MMSE approaches to multiuser spatio-
          temporal Tomlinson-Harashima precoding", *Proc. of the ITG Conference on Source
          and Channel Coding*, pp. 387–394, Erlangen (Germany), Jan. 2004.

[Jor03]   E. Jorswieck, and H. Boche, "Optimal transmission with imperfect channel state in-
          formation at the transmit antenna array", *Wireless Personal Communications*, vol. 27,
          pp. 33–56, 2003.

[Kas85]   S. A. Kassam, and H. V. Poor, "Robust techniques for signal processing: A survey",
          Proceedings of the IEEE, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[Kie05]   M. Kiessling, "Unifying analysis of ergodic MIMO capacity in correlated rayleigh fading
          environments", *European Transactions on Telecommunications*, pp. 17–35, Jan. 2005,
          published online in Wiley InterScience (www.interscience.wiley.com).

[Kob05a] M. Kobayashi, and G. Caire, "A low-complexity approach to space-time coding for
          multipath fading channels", *EURASIP Journal of Wireless Communications and Net-
          working*, Aug. 2005.

[Kob05b] M. Kobayashi, G. Caire, and D. Gesbert, "Transmit diversity vs. opportunistic beam-
          forming in data packet mobile downlink transmission", *IEEE Trans. on Communica-
          tions*, 2005.

[Koc04]   T. Koch, "On the asymptotic capacity of multiple-input single-output fading channels
          with memory", M.S. Thesis, Signal and Information Processing Laboratory, ETH Zurich,
          Switzerland, Apr. 2004, supervised by Prof. Amos Lapidoth.

[Lap98]   A. Lapidoth, and P. Narayan, "Reliable communication under channel uncertainty",
          IEEE Trans. on Information Theory, vol. 44, no. 6, pp. 2148–2177, June 1998.

[Lap02] A. Lapidoth, and S. M. Moser, "On the fading number of multi-antenna systems over flat fading channels with memory and incomplete side information", *Proc. IEEE International Symposium on Information Theory (ISIT'02)*, Lausanne (Switzerland), June-July 2002.

[Lap03] A. Lapidoth, and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels", *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[Lap04] A. Lapidoth, and S. M. Moser, "Feedback increases neither the fading number nor the pre-log", *Proc. 23rd IEEE Convention of Electrical and Electronics Engineers in Israel (IEEEI'04)*, pp. 2266–2270, Herzlia, Israel, Sep. 2004.

[Lap06] A. Lapidoth, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information", *Available on-line at http://arxiv.org/abs/cs.IT/0605079*, 2006.

[Lia04] Y. Liang, and V. V. Veeravalli, "Capacity of noncoherent time-selective rayleigh-fading channels", *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.

[Lin02] J. C. Lin, "Health effects: Scientific research and mobile phone testing", *IEEE Microwave Magazine*, vol. 3, no. 4, pp. 17–35, Dec. 2002.

[Lor05] R. G. Lorenz, and S. P. Boyd, "Robust minimum variance beamforming", *IEEE Trans. on Signal Processing*, vol. 53, no. 5, pp. 1684–1696, May 2005.

[Lov03] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems", *IEEE Trans. on Information Theory*, vol. 49, Oct. 2003.

[Loz02a] A. Lozano, and C. B. Papadias, "Layered space-time receivers for frequency-selective wireless channels", *IEEE Trans. on Wireless Communications*, vol. 50, no. 1, pp. 65–73, Jan. 2002.

[Loz02b] A. Lozano, and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures", *IEEE Trans. on Information Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[Loz06] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions", *IEEE Trans. on Information Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

[Lue98] D. Luenberger, *Optimization by Vector Space Methods*, New York: John Wiley & Sons, Inc., 1998.

[Ma02] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA", *IEEE Trans. on Signal Processing*, vol. 50, no. 4, pp. 912–922, Apr. 2002.

[Mag99] J. R. Magnus, and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 1999.

[Mar67] V. A. Marčenko, and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices", *Math. USSR-Sbornik*, vol. 1, pp. 457–483, 1967.

[Mar79] A. W. Marshall, and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[Mar99] T. L. Marzetta, and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading", *IEEE Trans. on Information Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[Mes03] X. Mestre, J. R. Fonollosa, and A. Pagès-Zamora, "Capacity of MIMO channels: asymptotic evaluation under correlated fading", *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 829–838, June 2003.

[Mil74] K. S. Miller, *Complex Stochastic Processes*, Addison-Wesley, Reading, MA, 1974.

[Miy72] H. Miyakawa, and H. Harashima, "Information transmission rate in matched transmission systems with peak transmitting power limitation", *National Conference Record, Institute of Electronics, Information, and Communication Engineers of Japan*, vol. 1269, pp. 1268, Aug. 1972.

[Mos04] S. M. Moser, *Duality-based bounds on channel capacity*, PhD Thesis, ETH Zurich, Switzerland, Oct. 2004.

[Mou02] A. L. Moustakas, and S. H. Simon, "Optimizing multi-transmitter-single-receiver (MISO) antenna systems with partial channel knowledge", *Bell Laboratories Technical Memorandum*, May 2002.

[Muk01] K. K. Mukkavilli, A. Sabharwal, and B. Aazhang, "Design of multiple antenna coding schemes with channel feedback", *Proc. Asilomar Conference on Signals, Systems, and Computers (ASILOMAR'01)*, vol. 2, pp. 1009–1013, Nov. 2001.

[Nar98] A. Narula, M. J. Lopez, M. D. Trot, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels", *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[Oza94] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio", *IEEE Trans. on Vehicular Technology*, vol. 43, May 1994.

[Pal03a] D. P. Palomar, *A unified framework for communications through MIMO channels*, PhD Thesis, Universitat Politècnica de Catalunya (UPC), 2003.

[Pal03b] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization", *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[Pal03c] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas Hernández, "Uniform power allocation in MIMO channels: A game-theoretic approach", *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1707–1727, July 2003.

[Pal04] D. P. Palomar, M. A. Lagunas Hernández, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints", *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[Pal06] D. P. Palomar, and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels", *IEEE Trans. on Information Theory*, vol. 52, no. 1, pp. 141–154, 2006.

[Pap91] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 3$^{\text{rd}}$ ed., 1991.

[Pay04a] M. Payaró, X. Mestre, and M. A. Lagunas, "Ergodic capacity of a 2×2 MIMO system under phase uncertainty at the transmitter", *Proc. IEEE International Conference on Electronics, Circuits and Systems (ICECS'04)*, Dec. 2004.

[Pay04b] M. Payaró, X. Mestre, and M. A. Lagunas, "Optimum transmit architecture of a MIMO system under modulus channel knowledge at the transmitter", *Proc. IEEE Information Theory Workshop (ITW'04)*, Oct. 2004.

[Pay05a] M. Payaró, X. Mestre, A. I. Pérez-Neira, and M. A. Lagunas, "On power allocation under phase uncertainty in MIMO systems", *Proc. Winterschool on Coding and Information Theory*, Feb. 2005.

[Pay05b] M. Payaró, X. Mestre, A. I. Pérez-Neira, and M. A. Lagunas, "Robust power allocation techniques for MIMO systems under modulus channel knowledge at the transmitter", *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC'05)*, June 2005.

[Pay06a] M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Capacity evaluation in multi-antenna systems under phase uncertainty", *Circuits, Systems, and Signal Processing (CSSP)*, Aug. 2006, submitted.

[Pay06b] M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Optimum linear transmitter design for MIMO systems with two QPSK data streams", *Proc. IEEE International Conference on Communications (ICC'06)*, June 2006.

[Pay06c] M. Payaró, A. Pascual-Iserte, D. P. Palomar, and M. A. Lagunas, "On linear transmitter designs for MIMO systems with ML detection", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'07)*, May 2006, submitted.

[Pay06d] M. Payaró, A. Pascual-Iserte, J. Yuan, and M. A. Lagunas, "A convex optimization approach for the robust design of multiuser and multiantenna downlink communication systems", *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'06)*, July 2006.

[Pay06e] M. Payaró, A. Wiesel, J. Yuan, and M. A. Lagunas, "On the capacity of linear vector Gaussian channels with magnitude knowledge and phase uncertainty", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'06)*, May 2006.

[Pay06f] M. Payaró, J. Yuan, and M. A. Lagunas, "On the capacity of MIMO systems with magnitude knowledge and phase uncertainty", *Proc. IEEE Australian Communications Theory Workshop (AusCTW'06)*, Feb. 2006.

[Pee05a] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-Part I: Channel inversion and regularization", *IEEE Trans. on Communications*, vol. 53, no. 1, Jan. 2005.

[Pee05b] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-Part II: Perturbation", *IEEE Trans. on Communications*, vol. 53, no. 3, pp. 537–544, Mar. 2005.

[PI06] A. Pascual-Iserte, D. P. Palomar, A. I. Pérez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization", *IEEE Trans. on Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[Pro95] J. G. Proakis, *Digital Communications*, McGraw-Hill, 3$^{\rm rd}$ ed., 1995.

[Ral98] G. G. Raleigh, and J. M. Cioffi, "Spatio-temporal coding for wireless communication", *IEEE Trans. on Communications*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[Rea06] M. Realp, *PHY-MAC scheduling and multiplexing with multi-user diversity*, PhD Thesis, Universitat Politècnica de Catalunya (UPC), 2006.

[Ron06] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block-coded MIMO systems: A probabilistically constrained approach", *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1560–1570, Aug. 2006.

[Sal85] J. Salz, "Digital transmission over cross-coupled linear channels", *AT&T Technical Journal*, vol. 64, no. 6, pp. 1147–1159, July 1985.

[Sca99] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels", *IEEE Trans. on Information Theory*, vol. 45, no. 3, pp. 1019–1032, Apr. 1999.

[Sca02] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders", *IEEE Trans. on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[Sch04] M. Schubert, and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints", *IEEE Trans. on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[Sch05] D. Schmidt, M. Joham, and W. Utschick, "Minimum mean square error vector precoding", *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'05)*, Berlin (Germany), Sep. 2005.

[Ser04] S. Serbetli, and A. Yener, "Transceiver optimization for multiuser MIMO systems", *IEEE Trans. on Signal Processing*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[Ses93] N. Seshadri, and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity", *Proc. IEEE Vehicular Technology Conference-Spring (VTC'93)*, pp. 508–511, May 1993.

[Sha48] C. E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[Sha49] C.E. Shannon, "Communicating in the presence of noise", *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[Sha05a] X. Shao, J. Yuan, and P. Rapajic, "Precoder design for MIMO broadcast channels", *Proc. IEEE International Conference on Communications (ICC'05)*, May 2005.

[Sha05b] M. Sharif, and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information", *IEEE Trans. on Information Theory*, vol. 51, no. 2, pp. 506–522, 2005.

[Sim02] S. H. Simon, and A. L. Moustakas, "Optimality of beamforming in multiple transmitter multiple receiver communication systems with partial channel knowledge", *Proc. DIMACS Workshop on Signal Processing for Wireless Transmission*, Oct. 2002.

[Sko03] M. Skoglund, and G. Jöngren, "On the capacity of a multiple-antenna communication link with channel side information", *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 395–405, Apr. 2003.

[Sto02] P. Stoica, and G. Ganesan, "Maximum-SNR spatial-temporal formatting designs for MIMO channels", *IEEE Trans. on Signal Processing*, vol. 50, no. 12, pp. 3036–3042, Dec. 2002.

[Tar98] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction", *IEEE Trans. on Information Theory*, vol. 44, Mar. 1998.

[Tar99a] V. Tarokh, H. Jafharkani, and A. R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[Tar99b] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths", *IEEE Trans. on Communications*, vol. 47, no. 2, pp. 199–207, Feb. 1999.

[Tel99] I. E. Telatar, "Capacity of multi-antenna Gaussian channels", *European Transactions on Telecommunications (ETT)*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999, see also a previous version of the paper in AT&T Bell Labs Internal Tech. Memo, June 1995.

[Tom71] M. Tomlinson, "New automatic equalizer employing modulo arithmetic", *Electronic Letters*, vol. 7, no. 2, Mar. 1971.

[Tul04] A. M. Tulino, and S. Verdú, *Random Matrices and Wireless Communications. Foundations and Trends in Communications and Information Theory*, NOW Publishers, 1$^{\text{st}}$ ed., 2004.

[Tul05] Tulino, A. M. and Lozano, A. and Verdú, S., "Capacity-achieving input covariance for single-user multi-antenna channels", *IEEE Trans. on Wireless Communications*, vol. 5, no. 3, pp. 662–671, Mar. 2005.

[Vic06] J. L. Vicario, *Antenna selection techniques in single- and multi-user systems: A cross-layer approach*, PhD Thesis, Universitat Politècnica de Catalunya (UPC), 2006.

[Vis01] E. Visotsky, and U. Madhow, "Space-Time Transmit Precoding With Imperfect Feedback", *IEEE Trans. on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[Vis03a] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum capacity of Gaussian MIMO broadcast channels", *IEEE Trans. on Information Theory*, vol. 49, no. 8, pp. 2658–2668, Aug. 2003.

[Vis03b] P. Viswanath, and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality", *IEEE Trans. on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[Vor03] S. A. Vorobyov, A. B. Gershman, and Z-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem", *IEEE Trans. on Signal Processing*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[Wei04] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel", *Proc. Conference on Information Sciences and Systems (CISS'04)*, pp. 7–12, Princeton, NJ (USA), Mar. 2004.

[Wei06] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel", *IEEE Trans. on Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[Wes98] R. D. Wesel, and J. M. Cioffi, "Achievable rates for Tomlinson-Harashima precoding", *IEEE Trans. on Information Theory*, vol. 2, no. 2, pp. 824–831, Mar. 1998.

[Wie05a] A. Wiesel, Y. C. Eldar, and S. Shamai, "Beamforming maximizes the rank one Ricean MIMO compound capacity", *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC'05)*, June 2005.

[Wie05b] A. Wiesel, Y. C. Eldar, and S. Shamai, "Semidefinite relaxation of the 16-QAM maximum likelihood detector", *Proc. Conference on Information Sciences and Systems (CISS'05)*, Mar. 2005.

[Wie06] A. Wiesel, Y. C. Eldar, and S. Shamai, "Optimization of the MIMO compound capacity", *IEEE Trans. on Wireless Communications*, 2006, accepted for publication.

[Wig55] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions", *Annals of Mathematics*, vol. 62, pp. 546–564, 1955.

[Wig58] E. P. Wigner, "On the distribution of roots of certain symmetric matrices", *Annals of Mathematics*, vol. 67, pp. 325–327, 1958.

[Wig59] E. P. Wigner, "Statistical properties of real symmetric matrices with many dimensions",
         Proc. 4th Canadian Mathemathical Congress, pp. 174–176, 1959.

[Win04a] C. Windpassinger, R. Fischer, T. Vencel, and J. Huber, "Precoding in multiantenna
          and multiuser communications", *IEEE Trans. on Wireless Communications*, vol. 3,
          no. 4, pp. 1305–1316, July 2004.

[Win04b] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reduction-aided broad-
          cast precoding", *IEEE Trans. on Communications*, vol. 52, no. 12, pp. 2057–2060, Dec.
          2004.

[Wis28] J. Wishart, "The generalized product moment distribution in samples from a normal
         multivariate population", *Biometrika*, vol. 20 A, pp. 32–52, 1928.

[Wol98] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An
         architecture for realizing very high data rates over the rich-scattering wireless channel",
         Proc. IEEE International Symposium on Signals, Systems, and Electronics (ISSSE'98),
         Oct. 1998.

[Woz65] J. M. Wozencraft, and I. M. Jacobs, *Principles of Communication Engineering*, John
         Wiley & Sons, 1965.

[Yan94] J. Yang, and S. Roy, "On joint transmitter and receiver optimization for multiple-
         input-multiple-output (MIMO) transmission systems", *IEEE Trans. on Communica-
         tions*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[Yoo05] T. Yoo, and A. Goldsmith, "Optimality of zero-forcing beamforming with multiuser
         diversity", *Proc. IEEE International Conference on Communications (ICC'05)*, pp. 542–
         546, May 2005.

[Yu01] W. Yu, and J. M. Cioffi, "Trellis precoding for the broadcast channel", *Proc. IEEE Global
        Telecommunications Conference, (GLOBECOM'01)*, vol. 2, pp. 1344–1348, 2001.

[Yu04] W. Yu, and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels", *IEEE
        Trans. on Information Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[Zam02] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multi-
         terminal binning", *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1250–1276,
         June 2002.

[Zhe02] L. Zheng, and D. N. C. Tse, "Communicating on the grassmann manifold: A geometric
         approach to the noncoherent multiple antenna channel", *IEEE Trans. on Information
         Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[Zhe03] L. Zheng, and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels", *IEEE Trans. on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[Zho02] S. Zhou, and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback", *IEEE Trans. on Signal Processing*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[Zho03] S. Zhou, and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations", *IEEE Trans. on Information Theory*, vol. 49, no. 7, pp. 1673–1690, July 2003.

[Zho04] S. Zhou, and G. B. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?", *IEEE Trans. on Wireless Communications*, vol. 3, no. 4, pp. 1285–1294, July 2004.